**JOHANNES GUTENBERG-UNIVERSITÄT**

Institute of Physics


# Variational Monte Carlo study of the Hydrogen Electronic Structure at Ultra-High Pressure

---

Dissertation submitted for the award of the title
"Doctor of Natural Sciences"
to the Faculty of Physics, Mathematics, and Computer Science
of the Johannes Gutenberg University in Mainz

Author:
Francesco Calcavecchia

# Contents

# Notation and conventions

$a_0$: Bohr units;

ASWF: Antisymmetric Shadow Wave Function (see section 1.3);

atm-bcc: Atomic bcc phase (see subsection 3.1.1);

BOP: Born-Oppenheimer potential;

CG: Conjugate Gradient;

DAC: Diamond Anvil Cell (see subsection 2.2.2);

DFT: Density Functional Theory;

DMC: Diffusion Monte Carlo;

$E$: Energy per particle;

$E_{\mathrm{JF}}$: Modified Jackson-Feenberg kinetic energy per particle (see subsection 3.1.5);

$E_{\mathrm{kin}}$: Pandharipande-Bethe kinetic energy per particle (see subsection 3.1.5);

$E_{\mathrm{pot}}$: Potential energy;

FSWF: Fermionic Shadow Wave Function (see section 1.3);

GSS: Golden Section Search;

H: Hamiltonian;

JF: Jackson-Feenberg (see subsection 3.1.5);

$\mathcal{H}$: Hilbert space;

$M$: Number of Monte Carlo sampled points;

$m_i$: Mass of the $i$-th particle;

MC: Monte Carlo;

MD: Molecular Dynamics;

mol-hcp: Molecular hcp phase (see subsection 3.1.1);

$M(RT)^2$ algorithm: Also known as Metropolis algorithm;

$N$: Number of particles in the simulation box;

NM-M: Non-metal to metal (phase transition);

pdf: Probability density function;

PB: Pandharipande-Bethe (see subsection 3.1.5);

PC: Periodic Coordinates;

PI: Parabolic Interpolation;

QMC: Quantum Monte Carlo;

Q-E: The Quantum-Espresso DFT code [1];

$r_s$: Alternative expression for the density (see subsection 3.1.1);

$\mathbf{r_i}$: 3-dimensional coordinates of the $i$-th particle;

$R$: Global coordinates of the simulated particles, i.e. $R = (\mathbf{r}_1, \mathbf{r}_2, \dots \mathbf{r}_N) \in \mathbb{R}^{3N}$;

$R_p$: Global coordinates of the protons;

SR: Stochastic Reconfiguration;

SWF: Shadow Wave Function;

TABC: Twisted Averaged Boundary Conditions (see subsection 3.1.7);

$\mathcal{V}$: Volume of the system;

var: Variance associated to the Monte Carlo estimator;

VMC: Variational Monte Carlo;

$V(r)$: Pair potential;

$\Gamma$: Integrand of the Shadow Wave Function (see section 1.2);

$\sigma$: Standard deviation associated to the Monte Carlo estimator;

Convention about Figures: When error bars are not visible, it means that they are smaller than symbols.

# Introduction

The *causa finalis* of the present work is the understanding of the hydrogen phase diagram at ultra-high pressures, which ranges from insulating $H_2$ to metallic H. Since conditions of ultra-high pressure are difficult to achieve in laboratory, computer simulations represent an essential alternative tool of investigation. However, such calculations are also very challenging. One of the major problems is represented by the accurate evaluation of the Born-Oppenheimer potential, which must be suitable for both the insulating and metallic phase, and has to account for the strong correlations entailed by the $H_2$ covalent bond and by the eventual phase transition. This problem was the target of our efforts.

In the context of Variational Monte Carlo (VMC), the Shadow Wave Function (SWF) is a very promising option. Due to its flexibility in describing both localized and delocalized systems, and its ability to account for high order correlations, it is an ideal candidate for our purposes. Unfortunately, its formulation for electrons involves a sign problem which limits its applicability. However, it is possible to circumvent this difficulty by imposing its nodal surface *a priori*. By means of this formalism we have been able to achieve significant improvements in the description of the electronic structure of hydrogen, opening for very promising prospects.

Along this line of research, we have also investigated the nature of the sign problem that affects the SWF, gaining a deeper understanding of its origin.

The present thesis is organized in four chapters. The first chapter introduces VMC and the SWF, with special attention to the fermionic systems. Chapter 2 outlines the literature on the phase diagram of hydrogen at ultra-high pressure. The third chapter presents the implementations of our VMC code and the results obtained. Finally, Chapter 4 summarizes our attempts to solve the sign problem associated with the SWF.

# Einleitung

Die *causa finalis* der vorliegenden Arbeit ist das Verständnis des Phasendiagramms von Wasserstoff bei ultrahohen Drücken, welche von nichtleitendem $H_2$ bis hin zu metallischem H reichen. Da die Voraussetzungen für ultrahohen Druck im Labor schwer zu schaffen sind, bilden Computersimulationen ein wichtiges alternatives Untersuchungsinstrument. Allerdings sind solche Berechnungen eine große Herausforderung. Eines der größten Probleme ist die genaue Auswertung des Born-Oppenheimer Potentials, welches sowohl für die nichtleitende als auch für die metallische Phase geeignet sein muss. Außerdem muss es die starken Korrelationen berücksichtigen, die durch die kovalenten $H_2$ Bindungen und die eventuellen Phasenübergänge hervorgerufen werden. Auf dieses Problem haben unsere Anstrengungen abgezielt.

Im Kontext von Variationellem Monte Carlo (VMC) ist die Shadow Wave Function (SWF) eine sehr vielversprechende Option. Aufgrund ihrer Flexibilität sowohl lokalisierte als auch delokalisierte Systeme zu beschreiben sowie ihrer Fähigkeit Korrelationen hoher Ordnung zu berücksichtigen, ist sie ein idealer Kandidat für unsere Zwecke. Unglücklicherweise bringt ihre Formulierung ein Vorzeichenproblem mit sich, was die Anwendbarkeit limitiert. Nichtsdestotrotz ist es möglich diese Schwierigkeit zu umgehen indem man die Knotenstruktur *a priori* festlegt. Durch diesen Formalismus waren wir in der Lage die Beschreibung der Elektronenstruktur von Wasserstoff signifikant zu verbessern, was eine sehr vielversprechende Perspektive bietet.

Während dieser Forschung haben wir also die Natur des Vorzeichenproblems untersucht, das sich auf die SWF auswirkt, und dabei ein tieferes Verständnis seines Ursprungs erlangt.

Die vorliegende Arbeit ist in vier Kapitel unterteilt. Das erste Kapitel führt VMC und die SWF mit besonderer Ausrichtung auf fermionische Systeme ein. Kapitel 2 skizziert die Literatur über das Phasendiagramm von Wasserstoff bei ultrahohem Druck. Das dritte Kapitel präsentiert die Implementierungen unseres VMC Programms und die erhaltenen Ergebnisse. Zum Abschluss fasst Kapitel 4 unsere Bestrebungen zur Lösung des zur SWF zugehörigen Vorzeichenproblems zusammen.

# Chapter 1

# Variational Monte Carlo and the Shadow Wave Function

Variational Monte Carlo (VMC) [2] is a well established method that allows for the approximate ground state of the time-independent Schrödinger equation. It applies the variational principle by taking advantage of the Monte Carlo (MC) ability to efficiently estimate the variational energy.

Compared to other Quantum Monte Carlo (QMC) methods, VMC requires small computational resources and, nevertheless, provides the same qualitative results. Moreover, its approximated solution is necessary for the actual implementation of QMC projection methods, like Diffusion Monte Carlo (DMC). In fact, without the optimized VMC trial wave function, DMC suffers from large population oscillations, which undermine its applicability.

In this chapter we introduce the reader to the VMC method and to the development of the trial wave function necessary to approximate the ground state. Afterwards, we present a special class of trial wave functions called Shadow Wave Function (SWF), which possess several unique features [3–20]. Finally, we present two different versions of the Shadow Wave Function, allowing for its application to fermions.

The chapter outline is:

- Section 1.1 introduces the Variational Monte Carlo method;
- Section 1.2 presents the Shadow Wave Function;
- Section 1.3 shows how to modify the SWF in order to describe fermionic systems.

## 1.1 Introduction to Variational Monte Carlo (VMC)

The subject of this chapter is the solution of the time-independent Schrödinger equation

$$\mathrm{H}\Psi = E\Psi \ , \tag{1.1}$$

formally an eigenvalue equation. Satisfactory[1] exact numerical methods which permit for the solution of this equation for many interacting bosons at $T = 0$ K are known, but not for fermions. In other words, the ground state (which is symmetric, and therefore bosonic) of a many-particle system can be determined numerically, whereas the excited states, including the antisymmetric (fermionic) ground state, in general cannot. However, many interesting particles are fermions, like the electrons. For this reason many attempts have been devoted to develop a general method that allows for a good approximation of the fermionic ground state. In the following we are going to outline the most famous and successful ones.

From now on, we restrict our discussion to systems with a two-body local interaction, and therefore the Hamiltonian is considered to be in the form

$$\mathrm{H} = -\hbar^2 \sum_{i=1}^{N} \frac{\nabla_i^2}{2m_i} + \sum_{i<j} V(r_{ij}) \, . \tag{1.2}$$

We are interested in 3D systems, so the particle coordinates are represented by vectors $\mathbf{r}_i \in \mathbb{R}^3$. For sake of shortness we also introduce the global coordinates $R \equiv (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots \mathbf{r}_N) \in \mathbb{R}^{3N}$.

A first approach to solve the SE is to consider the solution in the form of a product of single-particle orbitals[2]

$$\Psi_{\mathrm{MF}}(R) = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)\phi_3(\mathbf{r}_3)\cdots\phi_N(\mathbf{r}_N) \, , \tag{1.3}$$

introduce the mean-field potential[3] $v_{\mathrm{MF}}$, and decompose the SE accordingly:

$$\left(-\frac{\hbar^2}{2m_i}\nabla^2 + v_{\mathrm{MF}\,i}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}) \, . \tag{1.5}$$

This is known as the Independent-Particle Approximation (IPA), and leads to mean-field methods, like Hartree-Fock. In Equation (1.3) we have formulated the wave function as simple product of orbitals, but it must be recast in a

[1] By satisfactory we mean a method that does not increase exponentially in computational time with the number of particles.

[2] In the many-body physics language, orbitals are functions of single-particle coordinates. In other words, orbitals are one-body terms.

[3] Its expression changes accordingly to the chosen wave function. Anyhow, the mean field potential is obtained by averaging the potential energy over the positions of all the other particles. In this way the $N$-body interaction is replaced by an interaction with a mean field potential, formally assimilable as external. For example, with the wave function reported in Equation (1.3), the corresponding mean field potential for the particle $i$ reads

$$v_{\mathrm{MF}\,i}(\mathbf{r}) = \int dR^{\bar{i}} \prod_{j=1}^{i-1}\prod_{j=i+1}^{N} \left|\phi_j(\mathbf{r_j})\right|^2 \left(\sum_{j=1}^{i-1}\sum_{j=i+1}^{N} V(|\mathbf{r}-\mathbf{r}_j|)\right) \, , \tag{1.4}$$

where $dR^{\bar{i}} \equiv d\mathbf{r}_1 \cdots d\mathbf{r}_{i-1} d\mathbf{r}_{i+1} \cdots d\mathbf{r}_N$.

symmetric or antisymmetric form in order to describe bosons or fermions, respectively. In the case of bosons, since the ground state will be the product of the orbital $\phi_0$ that grants the lowest eigenvalue $\varepsilon_0$, the symmetric form is

$$\Psi_{\text{MF-bosons}}(R) = \phi_0(\mathbf{r}_1)\phi_0(\mathbf{r}_2)\phi_0(\mathbf{r}_3)\cdots\phi_0(\mathbf{r}_N)\,. \tag{1.6}$$

To determine the antisymmetric ground state instead, a Slater determinant (SD) is adopted:

$$\Psi_{\text{MF-fermions}}(R) = \det\begin{pmatrix} \phi_0(\mathbf{r}_1) & \phi_0(\mathbf{r}_2) & \dots & \phi_0(\mathbf{r}_N) \\ \phi_1(\mathbf{r}_1) & \phi_1(\mathbf{r}_2) & \dots & \phi_1(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \\ \phi_N(\mathbf{r}_1) & \phi_N(\mathbf{r}_2) & & \phi_N(\mathbf{r}_N) \end{pmatrix}. \tag{1.7}$$

Here we have presented the simplest way to write the ground state, corresponding to the well-known Hartree and Hartree-Fock methods. The major limitation of this approach is that it neglects the *correlations* between particles. More accurate forms can be obtained if superpositions including excited states are considered. This is, by the way, the leading idea of some sophisticated (and computationally demanding) methods like Coupled Cluster (CC) or Full Configuration Interaction (Full CI), which we are not going to discuss here.

In 1964 Kohn and Hohenberg established the Density Functional Theory (DFT), proving that it is possible to describe a quantum system supplying its density distribution $\rho$ rather than the wave function [21]. One year later, Kohn and Sham introduced the homonymous (KS) equation [22]

$$\left(-\frac{\nabla^2}{2m_i} + v_{\text{eff}}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r})\,, \tag{1.8}$$

which is formally identical to the mean field equation (Equation (1.5)). It should be noted that the DFT orbitals $\phi_i$ are conceptually different from the ones used in the IPA approximation (equation (1.3)). In DFT they are merely a formal expedient for expressing the density

$$\rho_{\text{DFT}}(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2\,. \tag{1.9}$$

Hence, it should be kept in mind that DFT issues the density of the system, and not its wave function. In summary, the similarity between equation (1.5) and (1.8) is only formal. The substantial difference between mean field theory and DFT originates from the $v_{\text{eff}}$ term, which includes the correlation effect and, in the case of a fermionic system, exchange contributions, written as functional of the density. Even though $v_{\text{eff}}$ could in principle be exact, in general its expression is unknown, and thus DFT is in practice an approximated method. DFT has been very successful, and it has been largely used in quantum chemistry. Its major drawback is the impossibility to asses the accurateness of $v_{\text{eff}}$ in describing the system, other than by comparison with the experiment or with more accurate methods.

We finally present the approach that we adopted, which is based on the *variational principle* underlying the time-independent Schrödinger equation, which

states that (we will not prove it here)

$$\left.\frac{\delta E(\Psi_\alpha)}{\delta\alpha}\right|_{\alpha=\alpha_S} = 0 \quad \Longleftrightarrow \quad \mathrm{H}\Psi_{\alpha_S} = E\Psi_{\alpha_S}\,, \tag{1.10}$$

where $\Psi_\alpha$ is a *trial wave function* characterized by some *variational parameters* $\alpha$, and $E(\Psi_\alpha)$ is its corresponding energy functional

$$E(\Psi_\alpha) = \frac{\int dR\,\Psi^*_\alpha(R)\mathrm{H}\Psi_\alpha(R)}{\int dR\,\Psi^*_\alpha(R)\Psi_\alpha(R)}\,. \tag{1.11}$$

$\alpha_S$ is a *stationary point*, and corresponds to a solution of the Schrödinger equation. By changes in $\alpha$, the trial wave function $\Psi_\alpha$ spans a subspace $\mathcal{H}_\Psi$ of the Hilbert space. Even though $\mathcal{H}_\Psi$ does not contain the exact ground state, as this would require the knowledge of its analytical functional form, one can determine its best approximation within this subspace by finding the stationary point with the lowest associated energy.

In summary, the variational principle states that, in order to find a solution of the SE that best approximates the ground state, it is sufficient to introduce an appropriate trial wave function $\Psi_\alpha$ and locate the stationary point $\alpha_0$ that minimize the energy $E(\Psi_\alpha)$. The choice of the trial wave function is crucial, because it determines the quality of the approximation. In this respect, the application of the variational principle has some similarity with DFT, since an ansatz is required. However, whereas the variational principle provides a criterion to determine the best choice among a certain set of ansätze, DFT requires a comparison with data outside its own framework in order to assess the appropriateness of the correlation and exchange energy functionals contained in $v_{\mathrm{eff}}$. In this sense DFT is said to be non-variational.

We remark that the knowledge of the wave function grants access to all the properties of the system that it describes, as long as the operator $O$ connected to the observable of interest is known. In fact, its expected value can be computed by means of

$$\langle O\rangle = \frac{\int dR\,\Psi^*(R)O\Psi(R)}{\int dR\,\Psi^*(R)\Psi(R)}\,. \tag{1.12}$$

Variational Monte Carlo (VMC) combines the variational method and the ability of Monte Carlo (MC) technique to efficiently evaluate the multidimensional integral for the energy (Equation (1.11)). In fact, through a simple manipulation, one can derive the expression

$$E(\Psi) = \frac{\int dR\,\Psi^*(R)\Psi(R)\frac{\mathrm{H}\Psi(R)}{\Psi(R)}}{\int dR\,\Psi^*(R)\Psi(R)}\,, \tag{1.13}$$

which can be easily evaluated by means of MC, sampling $M$ points $R_i$ from the probability density function

$$\rho_{\mathrm{MC}}(R) = \frac{\Psi^*(R)\Psi(R)}{\int dR\,\Psi^*(R)\Psi(R)} \tag{1.14}$$

by means of the $M(RT)^2$ algorithm (also known as Metropolis algorithm) [23, 24], and accumulating the *local energy* terms

$$E_{\mathrm{loc}}(R) = \frac{\mathrm{H}\Psi(R)}{\Psi(R)}\,. \tag{1.15}$$

The estimated energy is then

$$E = \frac{1}{M} \sum_{i=1}^{M} E_{\mathrm{loc}}(R_i)\,. \tag{1.16}$$

One merit of VMC is the possibility to unambiguously determine which is the trial wave function to favor over a certain selection of them: The lower the associated variational energy, the more accurate the approximation of the ground state. This gives VMC a tremendous advantage over DFT, wherewith it is often unclear which energy functional form to adopt, since there is no underlying variational principle guiding the decision. A disadvantage of VMC is its significantly more demanding cpu-time requirement, even though VMC falls in the embarrassingly parallel methods category, and therefore its work load can be conveniently distributed among any number of processors without drops in performance. The last mentioned feature is particularly appealing, since supercomputers are constantly growing in number of cores and not in clock frequency. Finally, we remark that VMC includes correlations and exchange effects only if the trial wave function takes them into account. That is to say that it is not a granted feature, but must be included by elaborating an appropriate ansatz for the trial wave function.

As we have remarked several times now, the quality of the VMC results relies completely on the trial wave function adopted, and is therefore the most important ingredient. A good starting point for the ansatz is:

$$\Psi_{\mathrm{T}}(R) = \Psi_{\mathrm{MF}}(R)\, \Psi_{\mathrm{correlation}}(R)\,. \tag{1.17}$$

$\Psi_{\mathrm{MF}}$ is a mean field solution, which can be quite rough since the second factor normally amends it, and it consists of a symmetrized product sequence (bosons) or a Slater determinant (fermions). It is very common, especially for electronic systems, to employ the DFT orbitals in $\Psi_{\mathrm{MF}}$. The convenience of this choice can be assessed by means of the variational principle.

Correlations can be thought as partitioned in two-body contribution, three-body contribution and so on. Accordingly, one can introduce them in the wave function following the Feenberg expansion [25]:

$$\Psi_{\mathrm{correlation}}(R) = e^{\sum_{i<j} u_2(r_{ij})} \cdot e^{\sum_{i<j<k} u_3(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k)} \cdots\,. \tag{1.18}$$

The functions $u_k$ are called *pseudopotentials*, since they resemble the potential in the Boltzmann distribution. The two-body term $e^{\sum_{i<j} u_2(r_{ij})}$ and the three-body term $e^{\sum_{i<j<k} u_3(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k)}$ are usually referred to as *Jastrow J* (after the physicist Robert Jastrow who first introduced it) and *Triplet T*, respectively. In VMC one often refers to the *order* of the many-body correlations introduced. Such a terminology refers precisely to the classification of the many body-correlations suggested by the Feenberg expansion. For example, the Jastrow introduces many-body correlations of the second order (or two-body correlations), while the Triplet contains many-body correlations of the third order (or three-body correlations). It should also be kept in mind that a Jastrow does not account for the physical two-body correlations, but only part of them, depending on the quality of the ansatz. We remark that the Feenberg expansion includes correlations in a symmetric form, and is therefore not completely satisfactory for fermions.

As already mentioned, the trial wave function describing a fermionic system must contain an antisymmetric term, usually in a SD form. In the following, we want to briefly comment on two important aspects related to it:

- The first one is the *nodal surface*, i.e. the set of coordinates where the trial wave function is equal to zero. The problem of finding the correct nodal surface is very challenging, in fact knowing it exactly would allow for finding an exact solution, for example via fixed-node Diffusion Monte Carlo[4] (DMC). In other words, the knowledge of the nodal surface would give access to the exact fermionic ground state. When a SD is employed, the embedded orbitals determine the resulting nodal surface in a non-trivial way. For this reason, the choice of the orbitals is not a simple task, but nevertheless of great importance and therefore crucial. When Hartree-Fock or DFT orbitals are used in conjunction with fixed-node DMC, the resulting energy significantly improves.
- The second aspect is the so-called *backflow*, first introduced by Feynman and Cohen [26]. It is a modification to the plane waves functions that takes into account the backflow effect represented in Figure 1.1. To reproduce this phenomenon, they used terms of the form

$$e^{i\left(\mathbf{k}\cdot\mathbf{r}_i+\lambda\sum_{i\neq j} b(r_{ij})\mathbf{r}_{ij}\right)}\,, \tag{1.19}$$

  where $\lambda$ is a variational parameter and $b$ a function to be set. In this way, correlations for the fermi momenta $\mathbf{k}$ are introduced. When this modified plane waves are embedded in the Slater determinant, not only the nodal surface is modified, but antisymmetrical correlations are introduced. The backflow contribution can be extremely important, but it is also one of the most difficult terms to include and optimize.

In conclusion, leaving out the sophisticated Triplet and backflow terms, a typical trial wave function for fermions reads

$$\Psi_{\mathrm{T}} = \mathrm{J}(R)\mathrm{SD}(\phi_\alpha(\mathbf{r}_\beta))\,, \tag{1.20}$$

where $\phi_\alpha$ can be simple plane wave functions, Hartree-Fock or DFT orbitals, or any of the previous ones written on a modified basis set that includes the backflow effects.

## 1.2 The Shadow Wave Function (SWF)

The *Shadow Wave Function* (SWF) is a trial wave function first introduced by Kalos [3], and further developed by Reatto [4]. Its most general form is

$$\Psi_{\mathrm{SWF}}(R) = \int dS\,\Gamma(R,S)\,, \tag{1.21}$$

where $\Gamma$ is a function integrated over the auxiliary degrees of freedom $S$ named *shadows*. This definition is extremely general and allows for a great flexibility,

[4]The latter mentioned QMC method is a projection technique that permits for sampling from the lowest possible energy ground state within the given nodal surface.

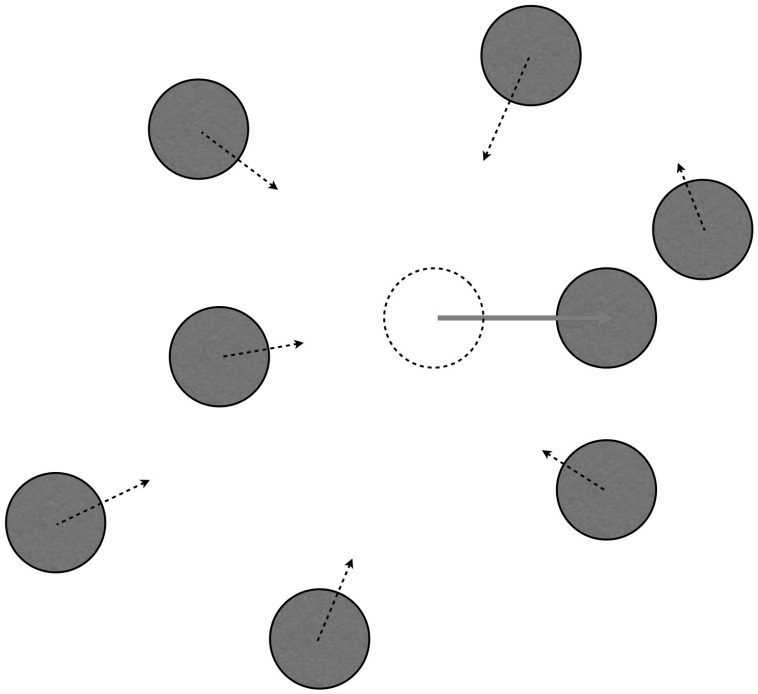

Figure 1.1: Visualization of backflow effect. The surrounding particles are adapting to the new position of a particle, filling the empty space left behind.

but a more specific form is necessary for practical applications. A typically used form is

$$\Psi_{\mathrm{SWF}}(R) = J_{\mathrm{p}}(R) \int dS\, \Xi(R,S) J_{\mathrm{s}}(S)\,, \tag{1.22}$$

where $J_{\mathrm{p}}$ and $J_{\mathrm{s}}$ are assumed to have the Jastrow form, and $\Xi$ is the so-called *kernel* that connects the particle coordinates $R$ to the shadows, and usually consists of a gaussian $e^{-C(R-S)^2}$, with $C$ a variational parameter.

The evaluation of the variational energy with the SWF requires two different coordinates sets for the shadows, which we label $S_1$ and $S_2$. In fact, both $\Psi^*_{\mathrm{SWF}}(R)$ and $\Psi_{\mathrm{SWF}}(R)$ result from an integral over $S$, and therefore the energy has to be calculated by performing the $9N$-dimensional integral

$$E(\Psi_{\mathrm{SWF}}) = \frac{\int dR dS_1 dS_2 \Gamma^*(R,S_1)\mathrm{H}\Gamma(R,S_2)}{\int dR dS_1 dS_2 \Gamma^*(R,S_1)\Gamma(R,S_2)}\,. \tag{1.23}$$

The simulation scheme is illustrated in Figure 1.2.

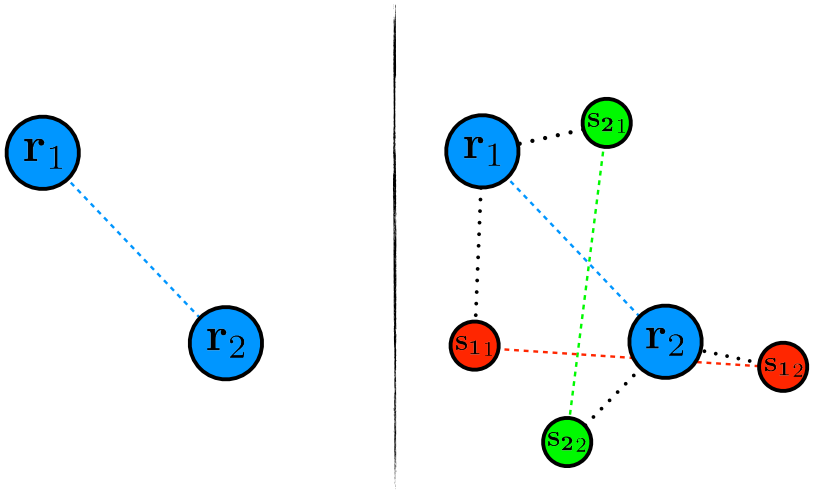


Figure 1.2: Schematic diagram of the interaction between two particles for a standard trial wave function that consists of a Jastrow (left side), and for the SWF presented in equation (1.22) (right side). In the SWF framework, each particle is connected with two shadows by means of a gaussian term, whilst the shadows are interacting in their turn via $J_{\mathrm{s}}$. We remark that $S_1$ and $S_2$ have no direct connections.

Along this section we do not mention the symmetry requirements of the trial wave function, but we rather leave the possibility open for both bosons and fermions. Section 1.3 is devoted to the application of the SWF to fermionic systems.

We now present two different ways of deriving the SWF functional form.

**First way.** It is well known that the imaginary time propagator $e^{-\tau \mathrm{H}}$ projects onto the ground state $\psi_{\mathrm{GS}}$ any trial wave function $\psi \not\perp \psi_{\mathrm{GS}}$. In order

to demonstrate it, we decompose the trial wave function into

$$\psi = \sum_{n=0}^{+\infty} c_n \phi_n \,, \tag{1.24}$$

where $\phi_n$ are the eigenfunctions of the Schrödinger equation, i.e. $\mathrm{H}\phi_n = E_n\phi_n$ for all $n \in \mathbb{N}$, with $E_n$ indicating the associated eigenenergies. If we apply the imaginary time propagator onto $\psi$, we obtain

$$e^{-\tau \mathrm{H}}\psi = \sum_{n=0}^{+\infty} c_n e^{-\tau E_n}\phi_n \,. \tag{1.25}$$

The projector $e^{-\tau \mathrm{H}}$ cause all the components to decay exponentially[5], so that eventually the one with the lowest energy $E_0$, i.e. the ground state, is the dominant term

$$\lim_{\tau\to\infty} e^{-\tau \mathrm{H}}\psi = \lim_{\tau\to\infty} \sum_{n=0}^{+\infty} c_n e^{-\tau E_n}\phi_n \quad \propto \phi_0 \,. \tag{1.26}$$

Any given trial wave function $\Psi_\mathrm{T}$ can then be improved systematically by (the normalization factors are inessential and therefore omitted)

$$\begin{aligned} e^{-\tau \mathrm{H}}\Psi_\mathrm{T}(R) &= \langle R \mid e^{-\tau \mathrm{H}} \mid \Psi_\mathrm{T}\rangle & (1.27)\\ &= \int dS \, \langle R \mid e^{-\tau \mathrm{H}} \mid S\rangle\langle S \mid \Psi_\mathrm{T}\rangle \,, & (1.28) \end{aligned}$$

where we have introduced an integral over a complete set of Dirac deltas $\mid S\rangle$. Assuming that $\tau \ll 1$, we now use the Trotter formula to approximate

$$e^{-\tau(K+V)} \sim e^{-\frac{\tau}{2}V} e^{-\tau K} e^{-\frac{\tau}{2}V} \,, \tag{1.29}$$

where $K$ represents the kinetic term of the Hamiltonian. It is well known that

$$\langle x \mid e^{-\tau K} \mid y\rangle = \frac{e^{-\frac{(x-y)^2}{4\tau}}}{a} \,, \tag{1.30}$$

where $a$ is a normalization factor. Finally, we can derive the final expression for the improved trial wave function

$$e^{-\tau \mathrm{H}}\Psi_\mathrm{T}(R) = e^{-\frac{\tau}{2}V(R)} \int dS \, e^{-\frac{\tau}{2}V(S)} e^{-\frac{(R-S)^2}{4\tau}} \langle S \mid \Psi_\mathrm{T}\rangle \,. \tag{1.31}$$

According to our derivation, it is necessary to set $\tau \ll 1$. Hence, the imaginary time propagation would be very short, meaning a small change in the trial wave function. To elongate the propagation, the described procedure could be repeated, including additional auxiliary degrees of freedom and thus ending up with a formalism similar to the path integral approach, that would eventually exactly solve the Schrödinger equation. Nevertheless, our original intention was

[5] If some $E_n$ are negative, the corresponding terms grow exponentially rather than decay. Nevertheless, all the following conclusions apply anyway. This can be shown by means of a simple addition of an appropriate constant term to H, in order to restore the decaying terms.

to find an improved trial wave function to use within VMC, whereas a path integral approach would considerably increase the computational cost of the calculation. Considering this, we rather truncate the projection after one step and refine the obtained functional form at a variational level.

In other words, we wipe out the limit $\tau \to 0$, and substitute $\tau$ with a variational parameter $C$ into the gaussian term. Then, we interpret the factor $e^{-V(R)}$ as two-body correlation factor, namely a Jastrow. We do the same with $e^{-V(S_1)}$ and $e^{-V(S_2)}$. Last, but not least, the terms in the form $\langle S \mid \Psi_{\mathrm{T}} \rangle = \Psi_{\mathrm{T}}(S)$ imply that the original trial wave function has to be evaluated on the shadow coordinates. This last observation is particularly important for the term that dictates the symmetry of the shadow wave function (a product of orbitals for bosons, a SD for fermions). In conclusion, any trial wave function $\Psi_{\mathrm{T}}$ can be improved by means of the shadow formalism, assuming the form

$$\Psi_{\mathrm{SWF}}(R) = J_{\mathrm{p}}(R) \int dS \, e^{-C \sum_{i=1}^{N} (\mathbf{r}_i - \mathbf{s}_i)^2} \, J_{\mathrm{s}}(S) \, \Psi_{\mathrm{T}}(S) \, . \tag{1.32}$$

**Second way.** The second derivation of the SWF should be contextualized at the time when it was first derived [3]. The inclusion of the DFT orbitals into VMC was computationally very demanding. Hence, the SD was often written in terms of simple functions, like plane waves or gaussians.

Roughly speaking, we can divide all possible phases of materials in two categories: Delocalized (liquids and gas) and localized ones (solids). The mean-field solution for a delocalized phase is given by a product of simple plane waves orbitals (in the case of bosons it results into a constant function, since the first Fermi momentum is $\mathbf{k}_0 = 0$), whereas localized phases can be well described by a product of gaussians. Even if the localization originates from the interacting term in the Hamiltonian, this effect cannot be effectively reproduced only including correlation terms as the Jastrow, but requires a one-body term, such as a gaussian centered on an equilibrium position. In fact, even though a localized trial wave function is in principle possible employing only a Jastrow factor, the resulting kinetic energy is too high to entail this possibility.

In other words, neglecting for the moment the possibilities offered by DFT, one should decide a-priori whether the trial wave function should describe a localized system or not. In the first case, the equilibrium positions for the particles, i.e. the crystal lattice, should be specified. This is a highly limiting fact, as it precludes appropriate descriptions of liquids near the crystallization point, when particles are almost localized and diffuse very slowly. Moreover, it prevents the possibility of obtaining the crystal lattice as output.

A different perspective is given by the following observation: The Hamiltonian describing a solid is translationally invariant, contrary to the resulting phase. This is a very simple example of spontaneous symmetry breaking. Similarly to what happens for real systems, one wishes a broken symmetry to emerge spontaneously from the calculation, and not from an a-priori imposition. In other words, it is desirable to have a trial wave function that allows localization, albeit formally translationally invariant.

The SWF was originally developed to overcome this difficulty. Instead of

fixing some lattice positions, new auxiliary degrees of freedom were introduced (*shadows*), and sampled from an appropriate distribution that allowed them to be localized or delocalized. Finally, the shadows were connected to the particles through a convolution integral by means of a gaussian kernel:

$$\Psi_{\mathrm{SWF}}(R) = J_{\mathrm{p}}(R) \int dS\, e^{-C(R-S)^2} J_{\mathrm{s}}(S)\,. \tag{1.33}$$

$J_{\mathrm{s}}$ contains the variational parameters which determine whether the shadows are localized or not. We remark that $J_{\mathrm{s}}$ does not explicitly contribute to the kinetic energy, and therefore the possibility of a resulting high kinetic energy is circumvented.

We now want to show that, in the asymptotic limits of strong localization and of complete delocalization, the SWF assumes the correct form. In the first case, the shadows will be localized around the equilibrium positions $R_0$ through the Jastrow $J_{\mathrm{s}}$. In other words $J_{\mathrm{s}} \sim e^{-A(S-R_0)^2}$ with $A \gg 1$ and hence

$$\Psi_{\mathrm{SWF}}(R) \sim J_{\mathrm{p}}(R) \int dS\, e^{-C(R-S)^2} e^{-A(S-R_0)^2} \tag{1.34}$$

$$\sim J_{\mathrm{p}}(R) e^{-\alpha(R-R_0)^2}\,, \tag{1.35}$$

which coincidentally is a functional form well-suited for describing solids. In the case of a gas, or free particles, $J_{\mathrm{s}} \sim 1$, and therefore

$$\Psi_{\mathrm{SWF}}(R) \sim J_{\mathrm{p}}(R) \int dS\, e^{-C(R-S)^2} \tag{1.36}$$

$$\sim J_{\mathrm{p}}(R)\,, \tag{1.37}$$

resulting into a trial wave function that can well represent a gas phase.

In conclusion, the SWF provides the possibility to describe localized and delocalized systems within the same functional form, simply optimizing its variational parameters.

The Shadow Wave Function possesses several unique features. In the second derivation, we already mentioned its ability of describing both localized and delocalized phases while retaining its formal translational invariance, which is probably its most notable characteristic. Another remarkable quality of the SWF has been proven through the first derivation: It possesses a formal improved similarity to the exact ground state. In fact, given any trial wave function $\Psi_{\mathrm{T}}$, a SWF version of it can be devised via imaginary time propagation, which is known to eventually lead to the exact ground state. The resulting SWF is reported in equation (1.32).

A further quality of the SWF is the inclusion of many-body correlations of any order, possibly infinite. Even though only Jastrows are included formally, the two-body term $J_{\mathrm{s}}$, once integrated, entails inter-particle correlations of all order beyond the pair [4]. This peculiarity is of great importance in the study of second-order phase transitions, since particles are strongly interrelated by high-order many-body correlations. For this reason, the SWF has been successfully employed in several studies on the liquid-solid phase transition of $^4$He [10, 13].

Once the SWF is conveniently modified, writing the shadow variational parameters as functions of the local density, it allows the description of inhomogeneous systems. It is then possible to investigate the solid-liquid coexistence and interface in $^4$He [5]. Subsequently, the SWF has been devised to study the vacancies formation and related dynamics in solid $^4$He [7].

Moreover, the SWF permitted significant progresses in the study of the excited states of $^4$He [12, 14–16].

Despite the great success in the application of the SWF to bosons, equally convenient results were not achieved for fermions. In the next subsection we will explain the reasons behind it, and show that an extension of the SWF to the fermions is not straightforward.

## 1.3 The Shadow Wave Function for Fermions

The simplest way to formulate an antisymmetric version of the SWF is to append a Slater determinant $\mathrm{SD}(R)$ to a symmetric SWF, obtaining the *Antisymmetric Shadow Wave Function* (ASWF):

$$\Psi_{\mathrm{ASWF}}(R) = \mathrm{SD}(R) J_{\mathrm{p}}(R) \int dS\, \Xi(R,S) J_{\mathrm{s}}(S)\,. \tag{1.38}$$

Since we are discussing at a very abstract level, the choice of the orbitals embedded in the SD is irrelevant at the moment, hence it is not specified. The ASWF has certain merits, since it introduces higher order correlations. Nevertheless, such an improvement is restricted only to the symmetric part.

By inspection of equation (1.32), it appears manifest that the most natural way to formulate an antisymmetric version of the SWF is to evaluate the Slater determinant on the shadow coordinates. This justifies the introduction of the *Fermionic Shadow Wave Function* (FSWF):

$$\Psi_{\mathrm{FSWF}}(R) = J_{\mathrm{p}}(R) \int dS\, \Xi(R,S) J_{\mathrm{s}}(S) \mathrm{SD}(S)\,. \tag{1.39}$$

It can be easily proven that $\Psi_{\mathrm{FSWF}}$ changes sign upon the pair exchange of like-spin particles, as required for a fermionic wave function.

In general, the FSWF provides a lower variational energy than the ASWF, and is thus to be favored. This improvement stems from two main features of the FSWF:

1. The nodal surface results from a non-trivial integral that contains optimizable variational parameters. Consequentially, the nodal surface can be adjusted via the variational parameters embedded in the integral. By contrast, in the ASWF the nodal surface is defined by $\mathrm{SD}(R)$, and is not modified by the shadow formalism;

2. As already mentioned, the integral over the shadows has an antisymmetric structure, and, moreover, includes correlation contributions. On this regard the integral mimics the backflow structure, which is therefore implicitly included. Furthermore, since correlations of higher orders are taken into account, the resulting backflow includes high-order many-body effects and can be interpreted as a *super-backflow*.

Despite its limitations, the ASWF has proven to be useful in describing the liquid-solid transition in $^3$He [6]. Anyhow, the clear superiority of the FSWF has been proven during the analysis of the vacancy formation energy in solid $^3$He [9].

As we already pointed out, the FSWF provides lower variational energies than the ASWF, but unfortunately it is plagued by a sign problem that makes its adoption not always recommendable. Delocalized systems prove to be especially challenging on this regard. Simulations of liquid $^3$He, for example, require substantially more computational resources compared to the solid phase. Moreover, the efficiency of the computation decreases drastically with the number of simulated particles $N$. As a consequence, simulations with a $N$ large enough to infer conclusions on the thermodynamic limit are often impracticable.

The sign problem emerges in the evaluation of the variational energy

$$E = \frac{\int dR\, dS_1\, dS_2\, \Gamma^*(R,S_1)\mathrm{H}\Gamma(R,S_2)}{\int dR\, dS_1\, dS_2\, \Gamma^*(R,S_1)\Gamma(R,S_2)}\,, \tag{1.40}$$

where $\Gamma$ is the function introduced in equation (1.21). In fact, contrary to the case of the ASWF, which leads to a squared Slater determinant $\mathrm{SD}^2(R)$, the energy evaluation for the FSWF contains the non-positive-definite product $\mathrm{SD}(S_1)\mathrm{SD}(S_2)$. Hence, when the FSWF is employed, $\Gamma^*(R,S_1)\Gamma(R,S_2)$ cannot be interpreted as a probability density function to use in the $\mathrm{M(RT)}^2$ algorithm. However, the latter difficulty can be easily overcome by considering its absolute value, and evaluate the integral as

$$E = \frac{\int dR\, dS_1\, dS_2\ |\Gamma^*(R,S_1)\Gamma(R,S_2)|\ \frac{\Gamma^*(R,S_1)\Gamma(R,S_2)}{|\Gamma^*(R,S_1)\Gamma(R,S_2)|}\ \frac{\mathrm{H}\Gamma(R,S_2)}{\Gamma(R,S_2)}}{\int dR\, dS_1\, dS_2\ |\Gamma^*(R,S_1)\Gamma(R,S_2)|\ \frac{\Gamma^*(R,S_1)\Gamma(R,S_2)}{|\Gamma^*(R,S_1)\Gamma(R,S_2)|}}\,, \tag{1.41}$$

sampling from the probability density function

$$\rho_{\mathrm{MC}} = \frac{|\Gamma^*(R,S_1)\Gamma(R,S_2)|}{\int dR\, dS_1\, dS_2\ |\Gamma^*(R,S_1)\Gamma(R,S_2)|} \tag{1.42}$$

and accumulating the local energy

$$E_i = \frac{\mathrm{H}\Gamma(R,S_2)}{\Gamma(R,S_2)} \tag{1.43}$$

and the *weights*

$$w_i = \frac{\Gamma^*(R,S_1)\Gamma(R,S_2)}{|\Gamma^*(R,S_1)\Gamma(R,S_2)|} = \pm 1\,. \tag{1.44}$$

The resulting estimator for the energy is

$$E \simeq \frac{\sum_{i=1}^{M} w_i E_i}{\sum_{i=1}^{M} w_i}\,. \tag{1.45}$$

Due to the oscillating nature of the weights $w_i = \pm 1$, this series can be extremely slow converging and cause what we call the *sign problem of the FSWF*. Chapter 4 is completely dedicated to it.

# Chapter 2

# Hydrogen At Ultrahigh Pressure

Hydrogen at ultra-high pressure ($P > 10$ GPa) is one of the most studied systems, due to its abundance in the universe and theoretical interest [27–29]. In fact, despite its apparent simplicity, it reveals many interesting properties, such as high-temperature superconductivity, high zero point energy, and strong anharmonic effects.

In principle, at extremely high pressure, molecular hydrogen is predicted to dissociate into an atomic metallic phase (although this issue is still debated). Moreover, according to the BCS theory, this hypothetical material is a room-temperature superconductor. This prevision immediately attracted the interest of the scientific community, which began to investigate hydrogen at progressively higher pressures. Since advancements in experimental techniques are very hard to achieve, computer simulations provide an excellent alternative tool of investigation. However, the computational challenges also increase, leading to contradicting results.

In conclusion, the phase diagram of hydrogen at ultra-high pressure is still elusive.

This chapter is organized in the following manner:

- Section 2.1 provides a general introduction to the properties of hydrogen, with a special emphasis on ultrahigh pressure conditions;
- Section 2.2 reviews the most important experimental and theoretical contributions to this topic;
- Section 2.3 introduces the reader to the problem that we intend to address in the present work.

## 2.1 Introduction

Hydrogen is the first element in the periodic table and the simplest one, as it consists of only one proton and one electron. The two hydrogen isotopes

observed in nature are deuterium, with a natural abundance of $\sim 0.015\%$, and tritium, present in trace amounts only.

Due to its simplicity, when quantum mechanics was first developed, the atom of hydrogen was one of the first studied systems. The exact analytical solutions of the associated electronic Schrödinger equation are well-known, and referred to as $1s$, $2s$, $2p$, $3s$. $3p$, etc. atomic orbitals.

Hydrogen atoms are found isolated only very rarely in nature, because they tend to associate in $H_2$ pairs, as result of the H-H covalent bond. This bond is very stable, in fact its energy is equal to 4.52 eV, that corresponds to a temperature $T \sim 5 \times 10^4$ K.

In standard laboratory conditions, hydrogen is in its gas phase. But when it is cooled down to 20.37 K and 13.95 K, a liquid and a solid phase are observed respectively, as illustrated in Figure 2.1. The solid state is known to be structured in freely rotating $H_2$ molecules on a hcp lattice.

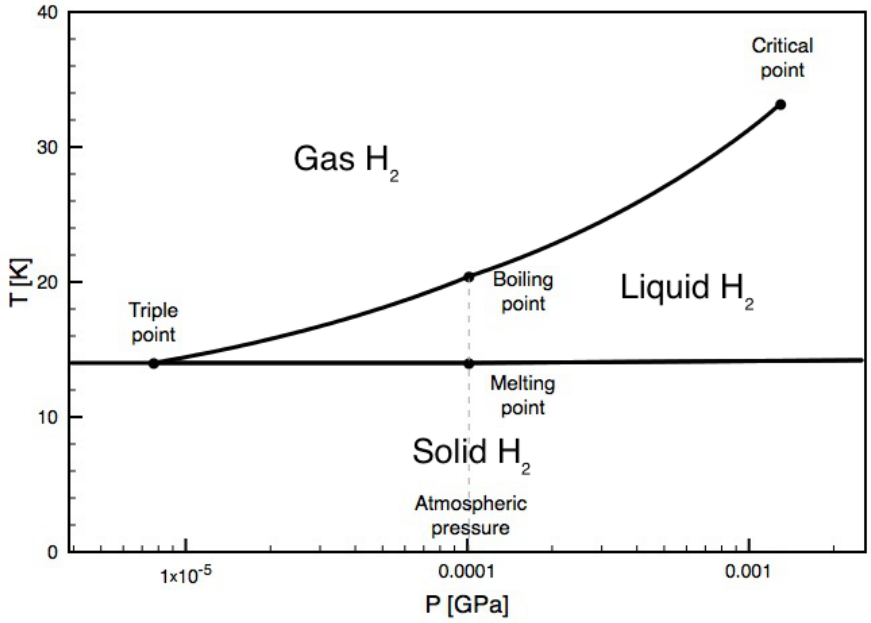


Figure 2.1: Schematic representation of the $P - T$ phase diagram of hydrogen near the atmospheric pressure. Data were taken from [30].

Not only hydrogen is the simplest element, but also the most abundant in the universe, as it constitutes around 75% of the universe [31] and 98% of the solar system mass. 90% of the Jupiter mass is formed by hydrogen subject to high pressure and relatively low temperature. It is not known which phases exhibits hydrogen in these conditions, but they are believed to be significantly different from the ones we commonly observe. The possibility of a metallic phase would explain the strong magnetic field generated by Jupiter. Therefore it is widely accepted that at very high pressure hydrogen undergoes an insulator-metallic phase transition.

The last observation leads us directly to notice that hydrogen belongs to

the group of alkali metals, even if it does not possess any metallic behavior. Nevertheless, hydrogen could manifest its metallic nature in very different thermodynamic conditions from the ones we have on Earth.

What makes hydrogen special compared to the other alkali metals is its extremely light nucleus, around seven times lighter than lithium. This peculiarity is linked to its very high zero-point energy, one of the characteristics that makes it so interesting.

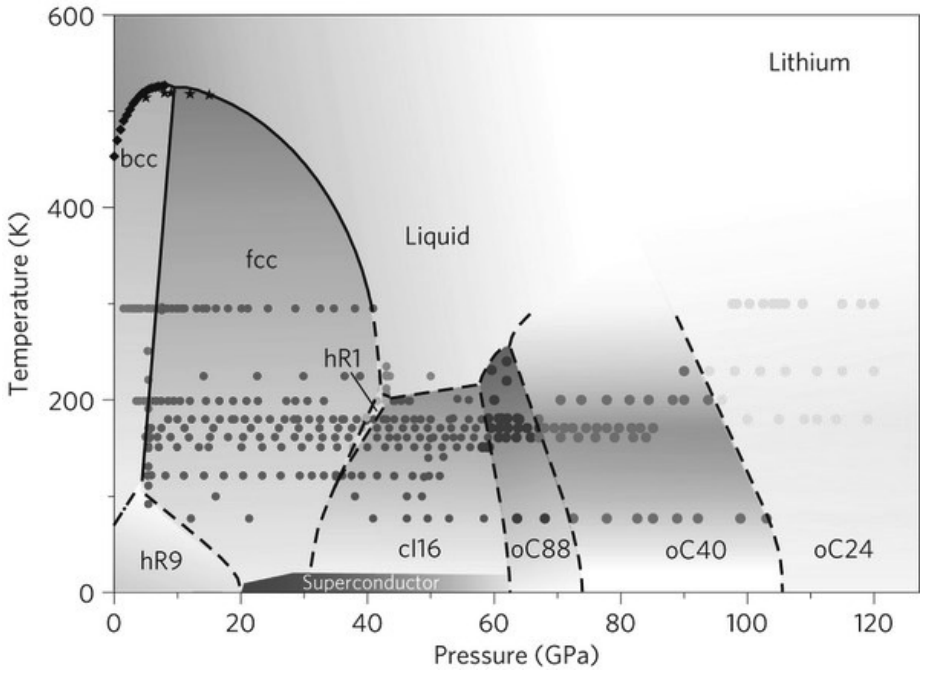


Figure 2.2: P-T phase diagram of lithium at very high pressure [32].

In 1935, Wigner and Huntington suggested the possibility that hydrogen exhibits a metallic phase at very high pressure and $T = 0$ K [33]. They first pointed out that any translationally invariant lattice of hydrogen atoms (not molecules) would be metallic. Then, to support their hypothesis, they evaluated the ground state energy for an atomic bcc lattice for different densities, finding a minimum for $r_s = 1.63$, where $r_s \equiv a/a_0$, $a_0$ equal to the Bohr radius, and $a = \sqrt[3]{\frac{3}{4\pi\rho}}$ ($\rho$ is the considered density). Given this result, they were justified to state that starting from a density of approximative $r_s = 1.63$, a solid atomic metallic phase was possible. They were aware of the fact that other atomic phases would have been possible (like hcp or fcc), but the conclusion would have been qualitatively the same. Nevertheless they suggested that a completely different phase, as a metallic layer-like lattice, could be more stable and appear at lower densities.

Wigner and Huntington gave an approximate analytical estimate of the energy including zero-point and correlation energies. As they acknowledge, it was not obvious how to relate the density $r_s = 1.63$ to the pressure, since the equation of state was not known up to that extreme condition. They gave a rough

estimate of 25 GPa, a pressure completely out of the range of the experimental possibilities at that time. Nowadays, however, we know that $r_s = 1.63$ corresponds to a pressure of $\sim 75$ GPa (for $T = 0$ K) [34]. Nevertheless, even by increasing the pressure to values considerably higher than 75 GPa, the atomic metallic phase predicted by Wigner and Huntington has not been detected.

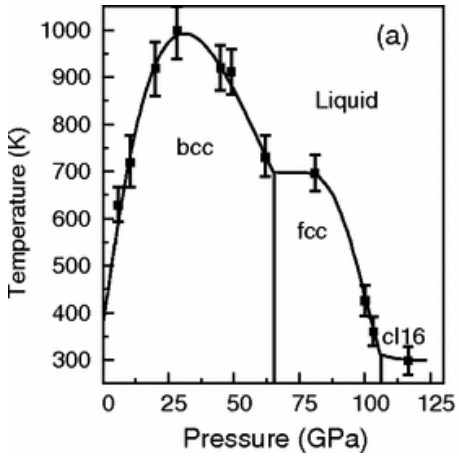


Figure 2.3: P-T phase diagram of sodium at very high pressure [35].

In 1968 Ashcroft applied the BCS theory to the hypothetical metallic hydrogen predicted by Wigner and Huntington [36] and he found that, thanks to the very light nucleus and corresponding high Debye temperature, metallic hydrogen is a good candidate for high-temperature superconductor, unlike other alkali metals. In fact, he estimated 64 K as the minimum transition temperature, but he recognized that it could be sensibly higher, possibly room-temperature. This prediction greatly rose the interest of the scientific community towards the phase diagram of hydrogen at very high pressure.

It is legitimate to wonder how the phase diagrams of other alkali metals look like. In Figures 2.2 and 2.3 we show the most recent results found for lithium and sodium [32, 35, 37]. In particular, we remark two interesting facts:

- Both lithium and sodium present a very unusual negative slope in the melting line;
- They both have bcc crystal structures which change into fcc at a high enough pressure, before undergoing other phase transitions and structuring themselves into more complicated lattices.

## 2.2 Literature Review

The literature on hydrogen at ultrahigh pressure is too vast to be summarized in few pages, hence we do not aim to write a complete overview on the topic (for more complete reviews please refer to [27–29]), but we rather provide the reader with a brief survey on the major findings and the encountered difficulties. The

first two subsections 2.2.1 and 2.2.2 are devoted to the two most advanced experimental techniques to reach ultra high pressure. We then introduce the results of both experiments and simulations, dividing them into investigations done at high temperatures (subsection 2.2.3) and low temperatures (subsection 2.2.4). Finally, in subsection 2.2.5, we report the literature about the melting line of hydrogen, which exhibits an unusual behavior.

### 2.2.1 Dynamic Compression: Shock-Wave Experiments

In shock-wave experiments compression to very high pressure is obtained through the collision of the sample with a bullet. The advantage of this technique is that it is possible to obtain high pressure and high temperature simultaneously, allowing for the study of the liquid region of the hydrogen phase diagram. We are now going to describe the experimental setup used by Weir, Mitchell and Nellis [38, 39].

The equipment consists mainly of two parts: A 19 meters long two-stage light-gas gun and a target containing the hydrogen sample, as illustrated in Figure 2.4. The bullet (or impactor) shot by the gun consists of a 20 g metal plate, that hits the target at an approximate speed of 7 Km/s. The target is illustrated in Figure 2.5, and is constituted by an aluminum container, two disks of sapphire which confine the hydrogen sample, and four electrodes necessary to measure the electrical resistance. The sample holder is deformed by the impact, therefore it was replaced with a new one after each measurement.

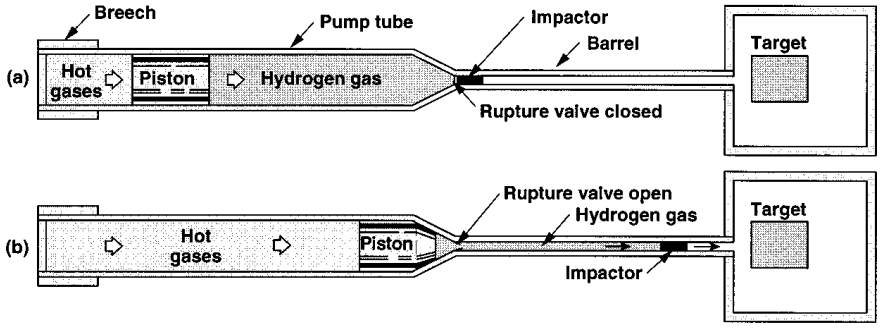


Figure 2.4: Illustration of the gun used in the shock-wave experiments. Initially an explosive is used to move a piston that, compressing the hydrogen, causes the rupture of the valve that separates the pump tube where the piston is, from the barrel. When the pressure is high enough and the valve breaks, the highly pressured hydrogen (around 0.1 GPa) push the impactor along the barrel. [38]

In order to obtain a not-too-high temperature, hydrogen was pre-cooled at 20 K, obtaining a liquid dense phase at atmospheric pressure. The aluminum layer was thus necessary, as part of the cryogenic procedure. Furthermore, aluminum is ductile and soft, allowing the impactor to transfer almost all his momentum to the sapphire disk.

Sapphire was chosen because of its stiffness, which is important to transmit the bullet momenta to the sample, and electrically insulation, so that it does

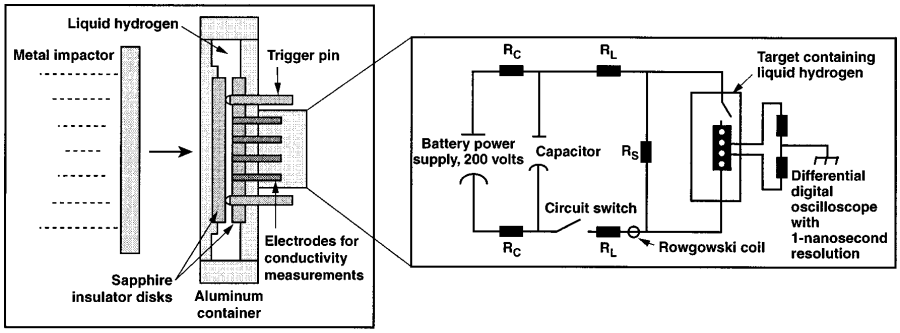


Figure 2.5: Illustration of the target assembled in the shock-wave experiments [38].

not interfere with the conductivity measurement.

The electrodes were made of stainless steel to have a similar shock impedance and thermal contraction of sapphire.

The thickness of the hydrogen layer was chosen so that the duration of the compression was long enough to reach thermal and electrical equilibrium, but short enough so that other not desired effects could not happen, such as chemical reactions or lowering of the temperature through diffusion. The typical duration of the experiment was 100 ns.

The final pressure is not reached immediately after the first impact, which generates a shock-wave with a pressure far lower than the final one, but it is obtained by the overlap of several shock-waves reflected by the sapphire disks, when the reverberation process attains the maximum of the pressure peak. This process allows to have a much lower temperature compared to a single shock compression that achieves the same pressure, and it is shown in Figure 2.6.

Since the experiment is so short, the transformation can be considered adiabatic and isentropic.

The pressure was estimated measuring the speed of the impactor trough flash X-ray radiography, and matching it with the known Hugoniot equations of state of the impactor, which describe the propagation of a shock-wave in a one-dimensional fluid or solid. The estimated uncertainty for pressure is 1%.

Finally, the resistivity $\rho$ was determined by measuring the resistance $R$ in two different and independent ways (this is why four electrodes were used, and not two) and evaluating the cell factor $C$ with computer simulations, so that

$$\rho = R/C\,. \tag{2.1}$$

We remind the reader that the conductivity $\sigma$ is defined as the inverse of the resistivity $\sigma \equiv 1/\rho$.

### 2.2.2 Static Compression: The Diamond Anvil Cell

The Diamond Anvil Cell (DAC) is an experimental instrument used to measure the properties of various materials in static conditions of extremely high pressure.

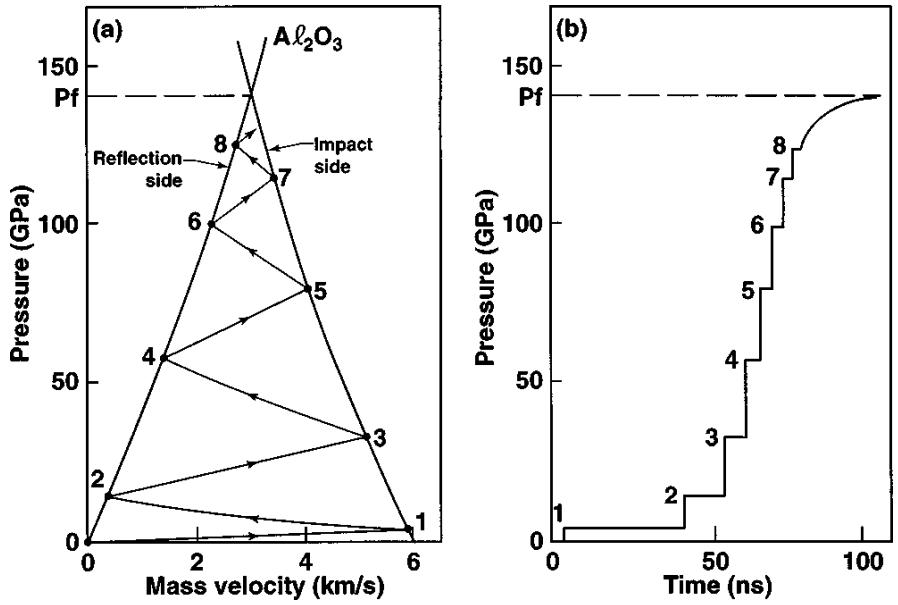


Figure 2.6: Reverberation dynamic inside the sample container [38].

The leading idea is that, since the pressure is the ratio of force to the area over which that force is distributed, a moderate force can generate an extremely high pressure on a sufficiently small surface.

In the first half of the 20th century, experimental research on high pressure physics was dominated by the Bridgman anvil and the piston-cylinder device, both developed by Percy Williams Bridgman (famous also for introducing the concept of operationalism [40]). These instruments were able to reproduce the remarkable pressure of 10 GPa. In 1950 Lawson and Tang introduced the DAC, that was lately further developed by other groups in different practical realization, differing mainly in the source of force acting on the diamonds and the diamonds alignment system [27]. Other differences were connected to measurement of interest: X-ray diffraction, spectroscopy, low or high temperatures conditions.

The essential structure of a DAC is represented in Figure 2.7, and it consists of two diamond anvils and a metal gasket that is used to contain the sample. The gasket is usually deformed by the diamonds during the compression, and has to be replaced after each experiment.

The most common method to quantify the internal pressure is to measure the fluorescence lines of a tiny ruby component that is inserted in the DAC together with the sample, whose trend is known for the range of pressures reached by the DAC.

Temperature can be set and measured in different ways. When the sample temperature is controlled by means of a thermal bath, the knowledge of its temperature is sufficient. This includes experiments at low temperature, employing special containers designed to contain cryogenic liquids as He, $N_2$, $O_2$, or $H_2$. High temperatures are more difficult to handle because external heat-

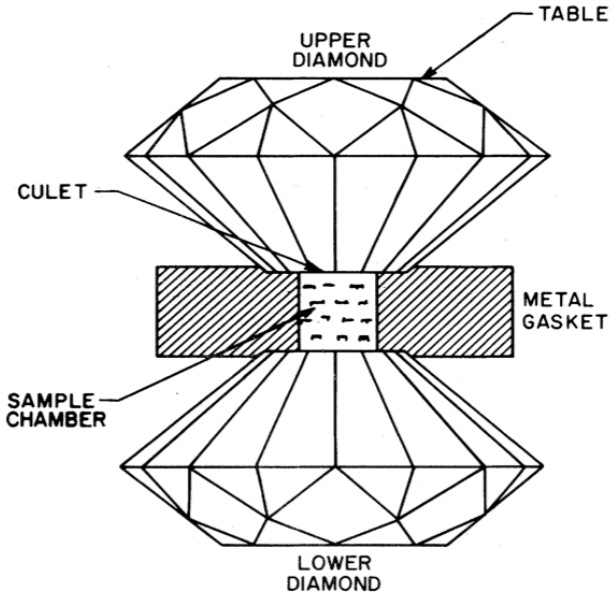


Figure 2.7: Illustration of the basic structure of a DAC [27].

ing of the sample has some limitation, since oxidation and/or graphitization of diamond limit the temperature to $700 - 800$ °C. Moreover, the hardness of diamond decreases with temperature, therefore this method is practically limited to $\sim$ 500 °C. To overcome this limitation, impulsed laser has been employed to heat the sample internally [41]. In this way, the different thermal diffusing coefficient of the sample (in our case, hydrogen), the gasket, and the diamond prevents the latter to become hot during the short time during which the sample is heated. The sample peak temperature can be determined through a blackbody spectra analysis.

X-ray diffractometry and vibrational spectroscopy (i.e. Raman spectroscopy) are also possible in conjunction with the DAC, and has been successfully applied to many materials [27]. In fact electromagnetic radiations can easily access the sample from the top and the bottom of the DAC, through the diamond.

Electrical conductivity measurements are more complicated but nevertheless possible [42]. It is accomplished by sputtering electrodes at the opposite sides of the sample, for example on the two culets of the diamonds, and inserting two or more probes to measure the current flow and the electrical potential at the two extremities.

The major difficulty in applying the DAC to hydrogen arises from the intrusion of H atoms into the diamond structure and its consequent rupture. Pressure and temperature are the two most important variables which increase the penetrative power of hydrogen. Hence, it is more difficult to enhance the pressure at room temperature rather than in a cryogenic thermal bath. This is why shock-wave experiments are preferred over DAC to explore the high-temperature phase diagram regions.

### 2.2.3 High Temperature

In 1995, single and double-shock experiments (i.e. shock wave experiments where the measurements are performed before the reverberation process ends) provided the first $P-T$ points up to 83 GPa [43]. These data permit to check the quality of theoretical models used to obtain other important information, like the dissociated fraction (i.e., fraction of atoms not bound into $H_2$ molecules) and the electrical conductivity.

In 1996 shock-wave experiments reported the first Non Metal to Metal (NM-M) transition of hydrogen ever observed [39], confirmed by more extensive measurements in 1999 [38]. In Figure 2.8 the experimental resistivity measurements at different pressures are illustrated. It should be mentioned that at very high temperatures and densities the zero-point vibrational energies of deuterium and hydrogen can be neglected. Therefore D has the same properties of H and can be used to investigate different thermodynamical conditions. Detailed data are reported in Figure 2.9. The temperature was obtained using theoretical equations of state which are strongly model-dependent, and therefore estimated within an uncertainty of 30%. In fact, at such a high pressure sapphire looses its transparency, preventing optical measurement, the only known technique to measure the temperature in shock-wave experiments [43]. Densities and dissociation fractions were also estimated with a model; the uncertainty for the first quantity is about 6%, while for the latter one none is specified.

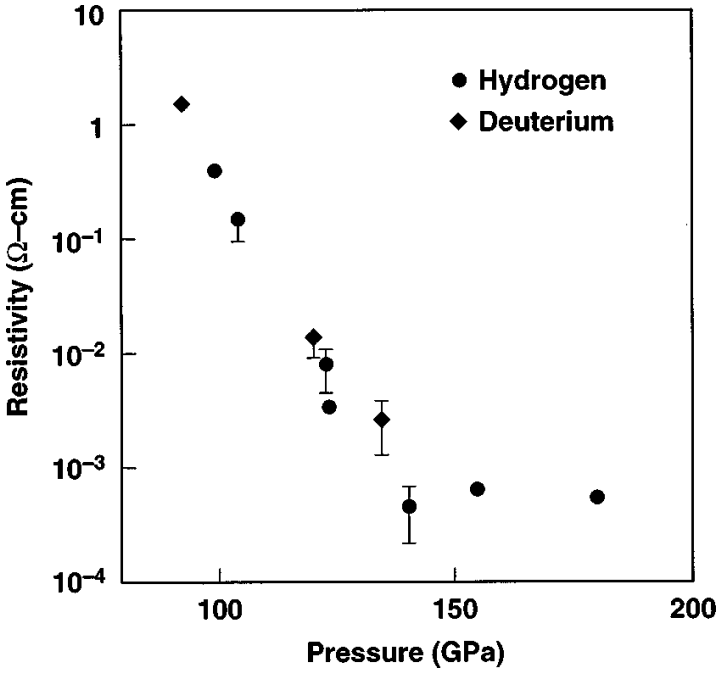


Figure 2.8: Resistivity versus pressure for liquid hydrogen. The temperature was in the range of $2000-3000$ K [38, 39].

The phase transition appears to be of the second order, and it is completed

| Experiment | Impactor | $U_I$ (km/sec) | Resistivity ($\Omega$ cm) | Pressure (GPa) | Density (mol/cm$^3$) | Temperature (K) | Dissociation fraction |
|---|---|---|---|---|---|---|---|
| SLDMS4-$D_2$ | Al | 5.59 | 1.4(0.3) | 93.1 | 0.291 | 2090 | $7.0\times10^{-3}$ |
| SLDMS5-$D_2$ | Al | 6.76 | $1.3(0.5)\times10^{-2}$ | 121 | 0.321 | 2760 | $4.9\times10^{-2}$ |
| SLDMS8-$D_2$ | Al | 7.33 | $2.4(1.2)\times10^{-3}$ | 135 | 0.334 | 3090 | $9.4\times10^{-2}$ |
| SLDMS6-$H_2$ | Al | 5.90 | $3.8(0.8)\times10^{-1}$ | 100 | 0.303 | 1670 | $2.7\times10^{-3}$ |
| SLDMS13-$H_2$ | Al | 6.10 | $1.4(0.5)\times10^{-1}$ | 105 | 0.308 | 1810 | $5.0\times10^{-3}$ |
| SLDMS7-$H_2$ | Al | 6.90 | $7.4(3.0)\times10^{-3}$ | 124 | 0.326 | 2230 | $2.3\times10^{-2}$ |
| SLDMS9-$H_2$ | Al | 6.91 | $3.2(0.7)\times10^{-3}$ | 124 | 0.326 | 2230 | $2.3\times10^{-2}$ |
| SLDMS12-$H_2$ | Cu | 5.58 | $4.2(2.2)\times10^{-4}$ | 141 | 0.342 | 2560 | $5.7\times10^{-2}$ |
| SLDMS10-$H_2$ | Cu | 5.96 | $6.0(1.2)\times10^{-4}$ | 155 | 0.357 | 2730 | $9.2\times10^{-2}$ |
| SLDMS11-$H_2$ | Cu | 6.65 | $5.0(2.0)\times10^{-4}$ | 180 | 0.380 | 2910 | $1.6\times10^{-1}$ |

Figure 2.9: Table summarizing the results obtained with shock-wave experiments [38].

at 140 GPa and an estimated temperature of 1600 K, where the resistivity drops from 1 $\Omega$ cm to $5 \times 10^{-4}$ $\Omega$ cm. These results show that the metallic phase can be more easily accessed at high temperature rather than at low temperature, in the solid phase.

Soon after that the NM-M transition of hydrogen was experimentally observed, many calculations were performed in order to reproduce the same results and provide a better understanding of the new liquid metallic phase.

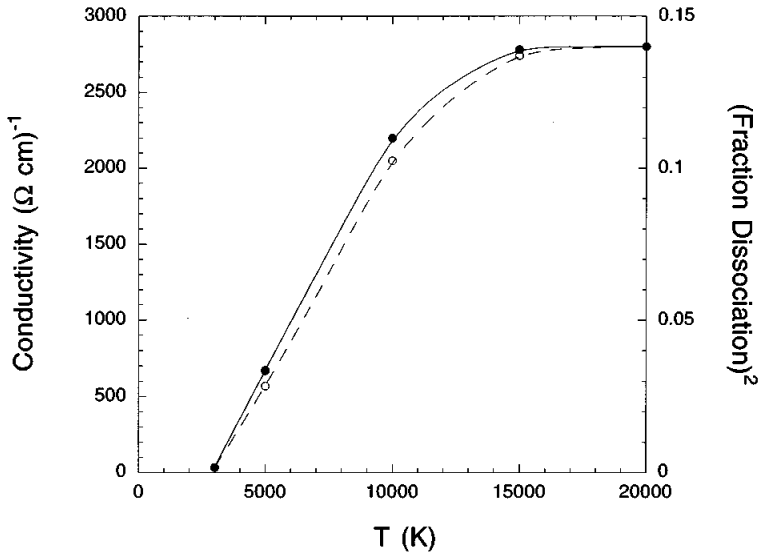


Figure 2.10: Electrical conductivity and squared dissociation fraction versus temperature, $r_s = 2$ [44].

In 1997, Lanosky et al. performed simulations of dense hot ($3000 - 10000$ K) liquid hydrogen using the molecular dynamic scheme in conjunction with an accurate tight-binding model for describing the H-H interaction [44]. This tight-binding model was built by fitting experimental and ab initio results, in order to be as accurate as possible. Protons quantum effects were neglected, as they

are irrelevant at very high temperature. They estimated the conductivity using an approximate expression, so that it can be expected to be correct within a factor of 2-3. Using this model, they observed that at increasing density and/or temperature, the electrical conductivity increased significantly, as reported from experiments. Moreover, they detected a molecular dissociation happening in the simulation, and they were able to relate it to conductivity, as illustrated in Figure 2.10, so that it is possible to state that the two processes are correlated. Finally, they observed that the dissociation process happens in an unusual way: $H_2$ molecules interchange H atoms between themselves very quickly ( 7 fs), and, during these exchanges, H atoms are isolated for very short time. This mechanism, even if it allows only a small portion of atoms to be dissociated (around 14%), is responsible for the electrical conductivity, acting similarly to dopants in semiconductors. In 1997 Bunker et al. performed a MC simulation using an effective potential to describe the atomic interaction, and reported the same dissociation process to happen [45].

In 2003, Scandolo simulated hydrogen at 1500 K in a pressure range of $75 - 175$ GPa using the Car-Parinello method [46]. He observed a sudden phase transition at $\sim$ 125 GPa accompanied by a 6% volume reduction. When the pressure was larger than 125 GPa, dissociation was evident by inspection of the proton-proton pair correlation function $g_{pp}(r)$. Moreover at 125 GPa $g_{pp}(r)$ was oscillating and the thermal equilibrium could not be reached. These results were consistent with the possibility of a first-order phase transition, even though the author could not exclude a rapid continuous transition. Hence there was an apparent conflict concerning the kind of phase transition (first or second order) with prior calculations, even if the latter ones were obtained with less accurate electronic energy evaluation, and therefore this most recent one was to be considered more reliable.

In 2004, Ceperley, Pierleoni et al. developed a new accurate algorithm to simulate hydrogen, called Coupled Electronic-Ionic Monte Carlo (CEIMC), based on the Born-Oppenheimer approximation [47]. Protons are simulated with classical MC (or PIMC, to include quantum effects), while the electronic structure is accurately determined by means of Quantum Monte Carlo techniques (Variation Monte Carlo [2] or Reptation Monte Carlo [48]). Using this method, they investigated the nature of the liquid NM-M transition [49], finding no evidence for a first-order phase transition at $T = 1500$ K and $T = 2000$ K when the more accurate RQMC is used.

In 2010, Bonev et al. simulated liquid hydrogen with a first-principle MD simulation, reproducing the NM-M phase transition [50]. They found that the polarizability smoothly increases while the transition happens. They were also able to characterize the liquid structure, finding that the metallic phase shows a short-range orientational order, which could explain the drop in volume observed by Scandolo in 2003.

In the same year Ceperley et al. completed an extensive study on the NM-M phase transition of liquid hydrogen, with First-Principle MD (FPMD) and an improved version of CEIMC [51]. They found that FPMD tends to underestimate the pressure of the phase transition, but qualitatively reproduced the same results of CEIMC: A plateau in the pressure appears when the transition happens as well as a sudden change in the conductivity. Because of these results the authors stated that the transition is of the first order, in contrast to what

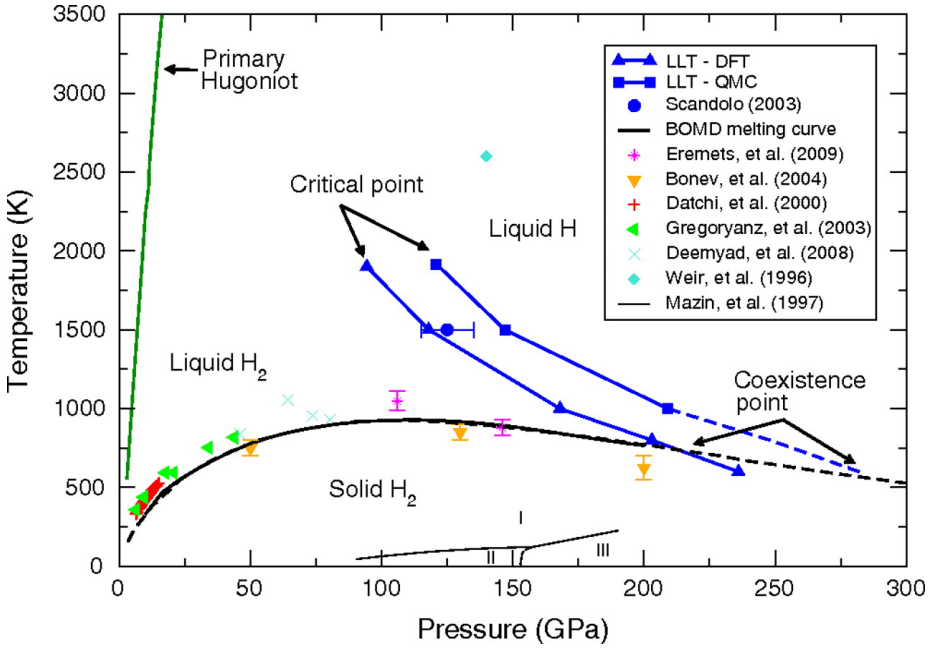


Figure 2.11: Phase diagram of hydrogen summarizing the latest results [51].

they have previously claimed when they first used CEIMC on hydrogen. This discrepancy with their own previous result is probably due to the improvement of the CEIMC algorithm. By contrast, above 2000 K the transition is smooth, therefore a critical point is expected to exist around this temperature, as shown in Figure 2.11. These results are currently the most accurate available, and give an explanation for the different type of phase transition observed by different research groups. In fact, above 2000 K, a second-order phase transition is expected, whereas below 2000 K a first-order phase transition is to be observed, solving the controversy between Lanosky et al. and Scandolo.

Lorenzen et al. confirmed the results obtained by Ceperley et al. providing, as further proofs of the first order nature of the phase transition for $T < 2000$ K, a sudden change in $g_{pp}(r)$ [52].

### 2.2.4 Low Temperature

In 1993 the DAC technique was evolved enough to permit an optical study of hydrogen up to 216 GPa at low-temperature (85 K) [28, 53]. IR and Raman spectra allowed to detect two novel phase transitions, one at 110 GPa and the other at 150 GPa. The first transition is characterized by the presence of a new peak in the IR spectra, which disappears again at 150 GPa, and by a small but detectable discontinuity in the Raman vibron. This new phase is commonly called phase II, whereas the solid state of hydrogen at lower pressure, including atmospheric pressure, is called phase I. At 150 GPa a new phase (first reported in [54] and [55]), called phase III, was revealed by the Raman spectra. In fact, a sharp shift in the IR and Raman vibron frequencies can be displayed.

In 1996 an extensive X-ray diffraction study of hydrogen at high pressure permitted to establish the $P - V$ equation of state up to 120 GPa at room temperature [34]. In Figure 2.12, the temperature-reduced results for $T = 0$ K and its extrapolation up to 350 GPa are reported.

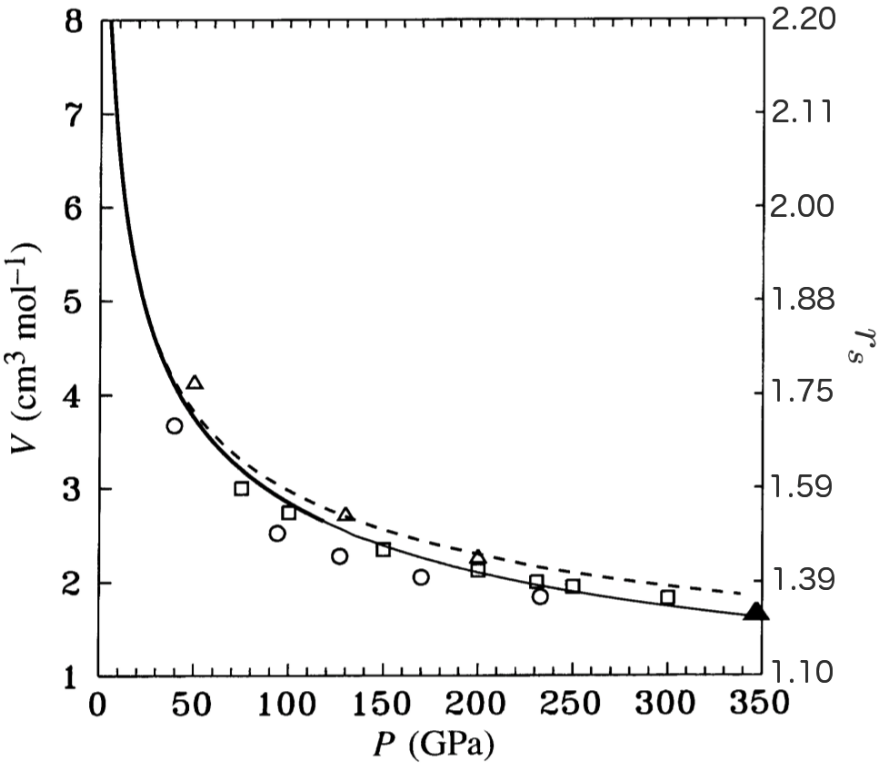


Figure 2.12: Equation of state of hydrogen at $T = 0$ K. The line up to 120 GPa represents the actual data, whereas from 120 GPa to 350 GPA an extrapolation using the Vinet function has been used. The symbols and the dashed line refer to various calculations, and are of no relevance for our discussion. On the right hand side of the graph we have depicted in blue the corresponding $r_s$. [34]

In 2002 the range of pressure that was possible to explore increased significantly, and optical measurement up to 320 GPa [56] and $T = 100$ K became possible. At 320 GPa the hydrogen sample turned completely opaque, a sign that the optical bandgap was $\sim$ 2 eV. This is referred to as black or dark hydrogen.

In 2011, at the Max Planck Institute for Chemistry in Mainz, Eremets and Troyan were able to compress $H_2$ using a DAC up to 315 GPa at room temperature, using some layers of Cu, Au, alumina and other materials to prevent the diffusion of hydrogen into the diamond, which was the reason why previous experiments failed to further increase the pressure [57]. They also reported the darkening of hydrogen at $220 - 260$ GPa, a condition that remained up to 315 GPa. Moreover, at 220 GPa new Raman lines appear and the Raman shift decays quickly without discontinuity, in contrast with experiments at low

temperature (transition to phase III is characterized by a discontinuity in the vibron frequency). Even though optical measurements are no longer possible when hydrogen turns dark, the electrical conductivity is still attainable, and revealed a sharp decay of four oder of magnitude in the resistance at $\sim$ 270 GPa (Figure 2.13). This new metallic phase was suggested by Eremets and Troyan to have a liquid structure, and has been lately named phase IV [58].

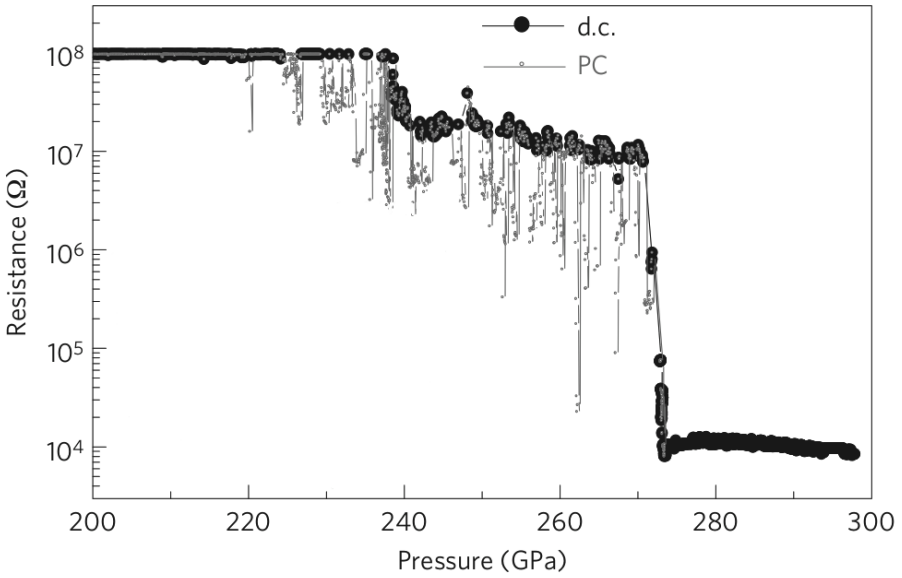


Figure 2.13: Resistance versus pressure at room temperature. The label "d.c." refers to measurements done without illumination, whereas for the photoconductivity "PC" data a He-Ne laser of 5 mW power was employed [57].

In 2012 the results found by Eremets and Troyan were confirmed by an independent experiment of Howie et al. [58]. The electrical conductivity was not measured and therefore, even though the two phase transitions were detected, it was not possible to confirm the high conductivity of phase IV. By means of a detailed analysis of the Raman, phonons, and vibrons spectra, and a comparison with simulations, they suggested phase IV to be a "a mixture of graphenelike layers, consisting of elongated $H_2$ dimers experiencing large pairing fluctuations, and unbound $H_2$ molecules".

We finally report an experimental result that at a first glance conflicts with the previously presented ones [59]. In 1975 three researchers in USSR found a decay of 6 orders of magnitude in the resistance of hydrogen at $P = 100$ GPa and $T = 4.2$ K. This result is in contrast with more recent results, but it might be justified since the authors reported that this transition was observed only for extremely thin layers of hydrogen: "In the case of relatively large thickness of the solid-hydrogen layer, the resistance does not depend on the force applied to the anvils and remains constantly high." Hence, a two-dimensional layer of hydrogen could have fundamentally different properties than a three-dimensional sample. To our knowledge there has not been any further investigation of this result.

In 1980 the first extensive study of the ground state of hydrogen by means of QMC was conducted [60]. Using Variational Monte Carlo (VMC), a wave function for protons (approximated as bosons) interacting with a screened Coulomb potential was employed, instead of an explicit inclusion of the electrons. The trial wave function of choice was a Jastrow multiplied by a product of gaussians centered on the fcc lattice positions, in the case of the solid phase, and simple plane waves for the liquid. They simulated hydrogen and deuterium with both the fcc and the liquid structure for $r_s$ equal to 0.8, 1.2, 1.36, 1.488, and 1.6, and compared the obtained variational energies. For all considered densities the fcc phase was preferred over the liquid one but in the case $r_s = 1.6$, when the two energies overlap within the MC error bar, opening the possibility for a liquid ground state.

In 1987 the VMC trial wave function used to simulate the hydrogen ground state was substantially improved, and new calculations were performed for $1 \leq r_s \leq 3$ [61]. Electrons were included in the trial wave function, and the Hamiltonian included the Coulomb interaction between electrons and protons. Also in this case protons were treated as bosons, considered that the strong localization of the hydrogen atoms makes the bosonic approximation to be reliable. Many different crystal structures were considered, and it has been found that starting from $\sim$ 100 GPa a rotationally ordered structure seems to be preferred. Moreover, the results suggest that a transition between the molecular and the atomic phase has to be expected between $r_s = 1.39$ (molecular) and $r_s = 1.30$ (atomic) with a change in volume of 20%.

In 1993-1995 the VMC calculation made in 1987 were repeated with an improved electronic trial wave function that employed the DFT orbitals for the SD [62, 63]. With the new improved results, they established that for $P > 300$ GPa the diamond crystal structure is favored over the bcc, fcc, and other lattices, and confirmed that for $P > 150$ GPa (phase III) a rotationally ordered phase is energetically favored.

In 2007 Pickard and Needs performed an extensive study in order to determine the crystal structure of hydrogen phase III by using DFT to search the crystal structure that minimizes the enthalpy [64]. The data referred to $T = 0$ K, including the zero point energy (ZPE) of the protons within the harmonic approximation. The resulting favorite structures are, for increasing pressure: $P6_3/m$ ($< 105$ GPa), $C2/c$ ($105 - 240$ GPa), $Cmca - 12$ ($240 - 385$ GPa), $Cmca$ ($385 - 490$ GPa), and $I4_1/amd$ ($> 490$ GPa). Nevertheless, comparing the Raman spectra of these phases with the experimental data, they were able to exclude most of them. Through these detailed analysis, the $C2/c$ crystal lattice seemed to be the best representative for phase III. Moreover, they were able to evaluate the bandgap and therefore establish the electrical conductivity of these phases, excluding metallization up to 320 GPa.

In 2011 a study similar to the one done in 2007 was repeated to predict the crystal structure of atomic hydrogen for higher pressures ($500 - 4500$ GPa) [65]. Unfortunately the authors were not able to state with certainty which was the favorite phase, because there were many extremely competitive ones. Moreover the enthalpies differences were dominated by the strong impact of the ZPE, that would have required the inclusion of complicated anharmonic effects. Nevertheless, their results suggest the following favored phases: $I4_1/amd$ ($500 - 1000$ GPa), $R - 3m$ ($1000 - 3500$ GPa), and fcc ($3500 - 4500$ GPa).

In 2012, because of the experimental results of Troyan and Eremets conflicting with their predictions, Pickard and Needs repeated their study including temperature effects (350 K) [66]. In this way, a new favorite mixed phase ($Pc$) emerged in the range of $250 - 295$ GPa when $T = 350$ K, whereas this was not the case for $T = 0$ K. Raman spectra and other observables of the new phase were in reasonable agreement with experiments, suggesting the mixed-structure $Pc$ as best candidate for phase IV (see Figure 2.14). This new phase consists of "alternate layers of strongly bonded molecules and weakly bonded graphene-like sheets", and is therefore in agreement with the suggestion given by Howie et al. [58].

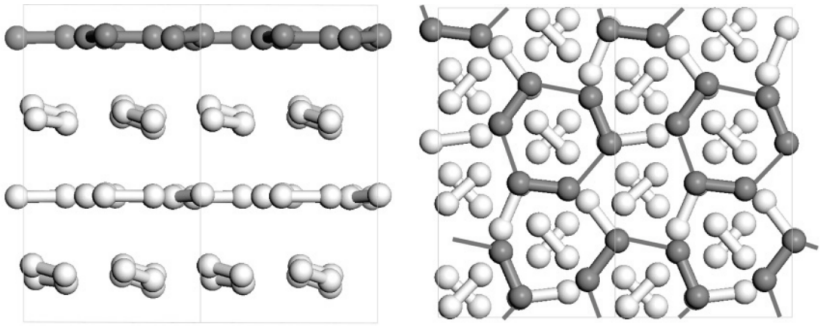

Figure 2.14: Side and top views of the mixed-structure $Pc$, the proposed hydrogen phase IV structure [66].

In the same year, Kühne et al. calculated the enthalpies of the phase suggested by Pickard and Needs in 2007 at a very accurate QMC level [67]. $P6_3/m$ (illustrated in Figure 2.15) resulted to be the most stable structure, and since also the Raman spectra is in agreement with the experimental data, it is nowadays the best candidate for phase III.

### 2.2.5 The Melting Line

Progresses on the DAC technique allowed the study of the melting line of hydrogen up to 15 GPa in 2000 [68]. Data supported the possibility of a maximum of the melting line in the $P$-$T$ phase diagram and a consecutive negative slope, in agreement with some *ab initio* calculations [69, 70]. In fact the Simon–Glatzel equation, which predicts a continuous increase in the melting temperature, could not describe satisfactorily the experimental data, whereas the Kechin one, which implies the existence of a maximum, perfectly fits them.

An extended study employing Car-Parinello *ab initio* calculations was performed in 2004 in order to establish the behavior of the melting line [71]. Hysteresis effects were prevented by running two-phase simulations, i.e. starting the simulation with half of the cell configured in a solid phase and the other half in a liquid phase, and waiting for the favorite phase to spontaneously emerge. In this way they were able to predict behavior of the melting line up to 200 GPa.

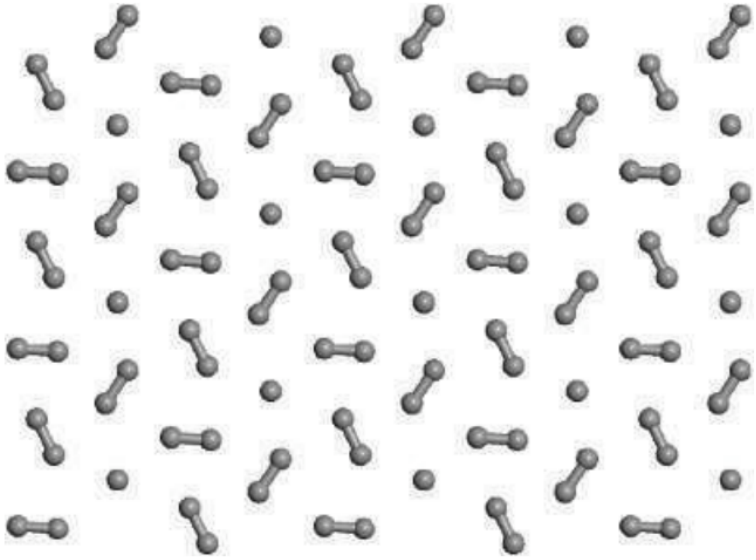

Figure 2.15: Representation of hydrogen at 300 GPa in the $P6_3/m$ structure, the most accredited candidate for phase III [64].

Progresses in the experimental methods allowed in 2008 [72] to confirm these theoretical predictions. In 2009 there was a further corroboration by an independent experiment [42].

In 2013 a computational study on the melting line of hydrogen employing ab-initio path integral molecular dynamic was published [73]. The focus was on the range of pressure $500 - 1200$ GPa, performing two-phase simulations (half of the initial simulation cell contained solid hydrogen with the $I4_1/amd$ structure, and the other half with liquid hydrogen) to find the melting line. The melting lines dropped almost linearly from $150 - 175$ K at 500 GPa to $75 - 100$ K at 800 GPa (see Figure 2.16). Even more interestingly, in the range $900 - 1200$ GPa, the melting line could not be detected down to 50 K, which was the lowest temperature attainable. The same results were not obtained when standard molecular dynamic (MD) was employed, revealing the importance of the quantum effects of the protons. Following the linear drop of the melting line, the authors predicted a liquid atomic metallic phase to exist starting from 1100 GPa.

## 2.3 Open Problems and Suitability of the VMC Method and the SWF

As we have outlined along section 2.2, the knowledge of the hydrogen phase diagram at ultra-high pressure is still obscure.

First of all only the crystal structure of phase I is known with certainty,

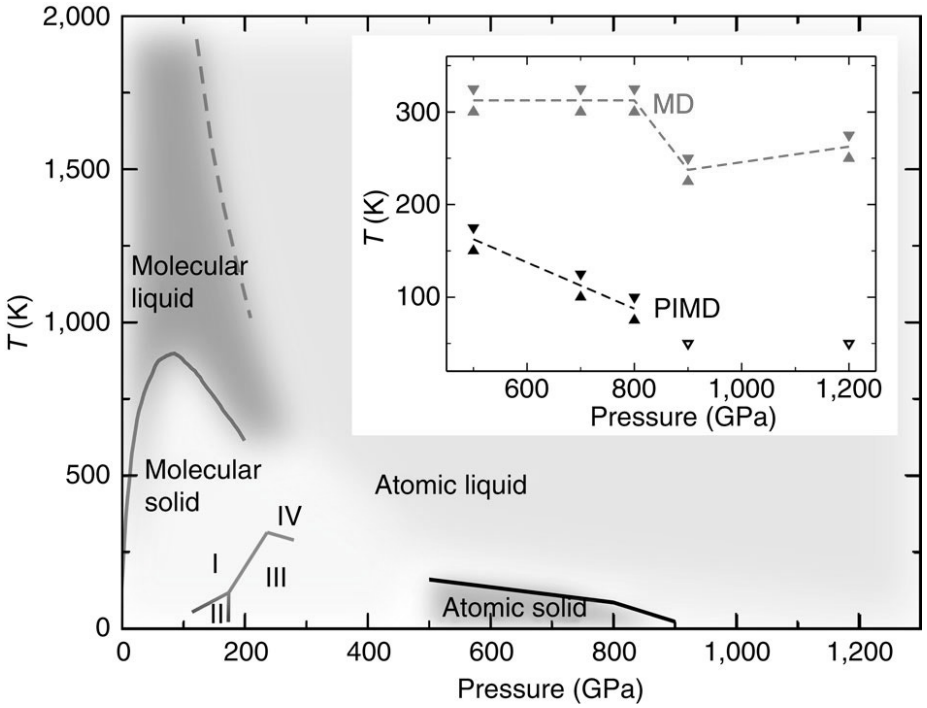


Figure 2.16: Prediction of the melting line of hydrogen for ultra-high pressure [73].

whereas the suggested structures for phase II, III and IV are controversial. Phase II, in particular, results to be especially elusive.

Secondly, computer calculations do not agree whether the liquid NM-M transition observed in 1995 [39] is of the first or second order, and could not resolve its collocation in the phase diagram. Recently the issue seemed to be solved [51], but more recent results re-opened the debate [74].

Of even greater importance, the question raised by Wigner and Huntington has still not found an answer. The pressure that will eventually reveal their predicted atomic metallic solid phase has been increased year by year, precluding resolutive experiments in the immediate future. Furthermore, the possibility that the metallic phase found by Eremets and Troyan in 2011 contains dissociated hydrogen atoms has been discarded.

In order to predict the transition to the Wigner-Huntington phase, it would be desirable to know the most stable atomic crystal lattices, in order to compare their enthalpy with the molecular insulating ones. The most recent and reliable study that attempted to determine the most favorable atomic solid structures for $P > 500$ GPa has not been entirely satisfactory, because the enthalpy differences between various structures were dominated by the anharmonic effects, which are difficult to estimate accurately [65]. The possibility of a liquid ground state, further complicates the problem.

The melting line of hydrogen is also a debated theme. Its negative slope opens the possibility to an intersection with the $T = 0$ K axis and consequentially to the existence of a liquid ground state. A recent study, which employed very sophisticated simulating techniques, supported this idea, and reported the stability of the liquid state for pressures between 900 and 1200 GPa at low temperature ($T = 50$ K) [73]. If confirmed, such a metallic liquid would have several unique and fascinating properties, as superconductivity and superfluidity [75–78].

In conclusion, hydrogen at ultra-high pressure is a topic that requires significant advancements in both theoretical and experimental techniques. Even though the progress in high-pressure laboratory techniques along the last decades has been remarkable, simulations are very important to predict and better interpret experimental results, and should be pursued. It is somehow puzzling that the simplest element cannot be accurately simulated yet. Many modern and accurate codes have been developed for this end, but very often with conflicting results, hence arising legitimate doubts on the validity of their predictions.

Undoubtedly, the most important ingredient for an accurate simulation is the electronic energy, or Born-Oppenheimer potential (BOP) energy. In fact, the dynamics of the protons is dominated by it, and small differences can lead to effects much larger than protonic quantum effects. Concurrently to accuracy, the BOP evaluation should be as modest as possible in terms of computational time, to allow its employment into efficient protonic dynamic. In this respect, the Variational Monte Carlo (VMC) method has many appealing properties:

1. It is very accurate, because it treats the exchange and correlation energies without any approximation, within the trial wave function ansatz. In general VMC provides more accurate results than DFT, which is the most widely used method;

2. It permits to verify the quality of the adopted trial wave function by means of the underlying variational principle;

3. It is not too demanding in terms of cpu-time. It requires more resources than DFT, but considerably less than more accurate methods like Diffusion Monte Carlo (DMC), or Coupled Cluster;

4. VMC falls in the category of embarrassingly parallel algorithm, thence it has in principle the possibility to take full advantage of an unlimited number of processors. This feature will most probably turn to be more and more important in the future, since supercomputers are growing in number of cpu and not in clock speed.

Moreover, VMC was the method of choice to determine the electronic structure in the most recent and accurate calculations on hydrogen, demonstrating its suitability [51, 74]. For all these reasons we believe that VMC is the method with the most promising prospectives.

In the context of VMC, the Shadow Wave Function (SWF) [3] is a trial wave function that has several unique characteristics. In the past, the SWF has been proven to be particularly appropriate to describe phase transitions and inhomogeneous systems, and therefore the study of the hydrogen phase diagram could benefit from its employment. In particular, the SWF is able to describe localized and delocalized systems within the same functional form while retaining its formal translational invariance [6], and this would be an especially welcome feature for describing the electronic structure of hydrogen in the eventuality of a NM-M transition. In conclusion, investigation on this trial wave function could open new unexpected possibilities.

# Chapter 3

# Study of the Electronic Structure of Hydrogen by means of VMC and the SWF

As we have pointed out in the previous chapter, Variational Monte Carlo (VMC) [2] and the Shadow Wave Function (SWF) [3] provide a very promising tool of investigation for the phase diagram of hydrogen. In particular, the simulation of its electronic structure could benefit from the flexibility of the SWF. In order to investigate the actual advantages, we have developed a VMC code that allows for the computation of the electronic ground state energy of hydrogen, assuming arbitrary protonic configurations within an orthorhombic crystal cell. We have included, with opportune modifications, several modern techniques which are the state of the art in the context of QMC, like Twist Averaged Boundary Conditions (TABC) [79], Periodic Coordinates (PC) [80], and Stochastic Reconfiguration (SR) [81–83]. Even though further improvements are possible, the results attainable with our code are accurate and reliable.

Our investigation mainly focuses on the methodological improvement attainable through modifications of the trial wave functions. Therefore the largest part of the discussion is dedicated to this matter. However, such an analysis allows also for physical insights.

The chapter is organized in the following manner:

- Section 3.1 introduces the most important technical aspects of the VMC code that we have developed;
- Section 3.2 presents and discusses our results;
- Finally, section 3.3 contains the conclusion and the future prospects.

# 3.1 Implementations of Our VMC Code

## 3.1.1 Considered Hydrogen Structures

Hydrogen can be characterized by the Hamiltonian

$$\begin{aligned} \mathrm{H} &= -\sum_{i_e=1}^{N} \hbar^2 \frac{\nabla_{i_e}^2}{2m_e} - \sum_{i_p=1}^{N} \hbar^2 \frac{\nabla_{i_p}^2}{2m_p} + \\ & \sum_{i_e<j_e} \frac{K_C}{|\mathbf{r}_{i_e} - \mathbf{r}_{j_e}|} + \sum_{i_p<j_p} \frac{K_C}{|\mathbf{r}_{i_p} - \mathbf{r}_{j_p}|} - \sum_{i_e=1,i_p=1}^{N} \frac{K_C}{|\mathbf{r}_{i_e} - \mathbf{r}_{i_p}|}, \end{aligned} \quad (3.1)$$

where $e$ and $p$ refer to the electrons and the protons respectively, and $K_C$ is the Coulomb constant. Spin contributions are neglected, as well as relativistic effects. Since in the present work we focus only on the electronic structure, the protons are considered fixed on a lattice, and their kinetic contribution is neglected. In other words, the protons are treated as the source of an external potential.

We consider hydrogen to be in the thermodynamical limit of a canonical ensemble. Therefore we set the density $\rho$ as constant. The calculation is restricted to $N$ atoms of hydrogen. However, the periodic boundary conditions are employed in order to mimic an essentially infinite system. The volume of the simulation box can be then determined by the ratio $\mathcal{V} = N/\rho$. It is a common convention to express the density in units of $r_s = \frac{a}{a_0}$, where $a_0$ is the Bohr radius and $a$ is defined by $\frac{\mathcal{V}}{N} = \frac{4}{3}\pi a^3$.

The two most relevant phases for solid hydrogen are the molecular insulating and the atomic metallic one. As we have already discussed, the corresponding crystal structures are not known. We have then considered only two simple protonic lattices, representative of the molecular insulating and the atomic metallic phases: The molecular hcp (mol-hcp) crystal structure illustrated in Figure 3.1, and the atomic bcc (atm-bcc) phase, represented in Figure 3.2. The atm-bcc structure is important because of the Wigner-Huntington prediction, who used this crystal lattice to predict the possibility of an atomic metallic phase for hydrogen at $T = 0$ K. The choice of the mol-hcp lattice is instead justified by noticing that phase I has freely rotating $H_2$ molecules arranged in a hcp lattice. Therefore, the $H_2$ orientation is relatively unimportant in order to describe a molecular insulating phase. We are aware that the atm-bcc and mol-hcp structures are not energetically favorable compared to others reported in literature, but they reproduce the most important qualitative characteristics of interest in order to test the quality of the Shadow Wave Function without dealing with the complications introduced by non-orthorhombic crystal lattices. However, the methods described along this chapter are general, and can be extended to more complicated structures.

The hydrogen energy is always expressed in units of Rydberg for atom, as commonly found in literature.

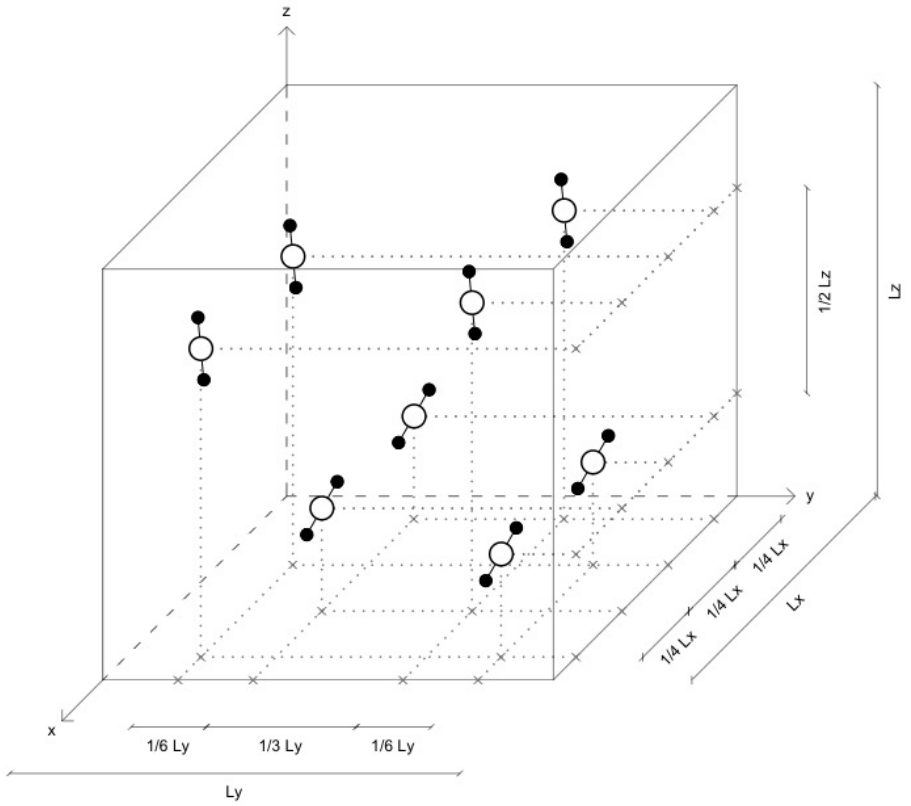


Figure 3.1: 3D representation of the molecular hcp primitive crystal cell (mol-hcp). The empty circles represent the centers of the molecules, whereas the filled ones represent the hydrogen atoms. The molecules centers form a hcp lattice, with sides $L_x = \sqrt{2}L$, $L_y = \sqrt{3/2}L$, and $L_z = \sqrt{1/3}L$, where $L = r_s \left(\frac{4\pi N}{3}\right)^{1/3}$. The $H_2$ protons pairs are set at the equilibrium distance 0.74 Å = 1.396 $a_0$, i.e. the covalent bond length. The molecules are arranged in two horizontal layers, with different molecular orientations. Starting with the molecules alined along the x direction, the lower (according to the $z$ axis) layer molecular orientation is obtained by rotating the molecules anti-clockwise by 60° in the $xz$ plane and then clockwise by 60° in the $xy$ plane. The upper layer results instead from an anti-clockwise rotation by 60° in the $xz$ plane.

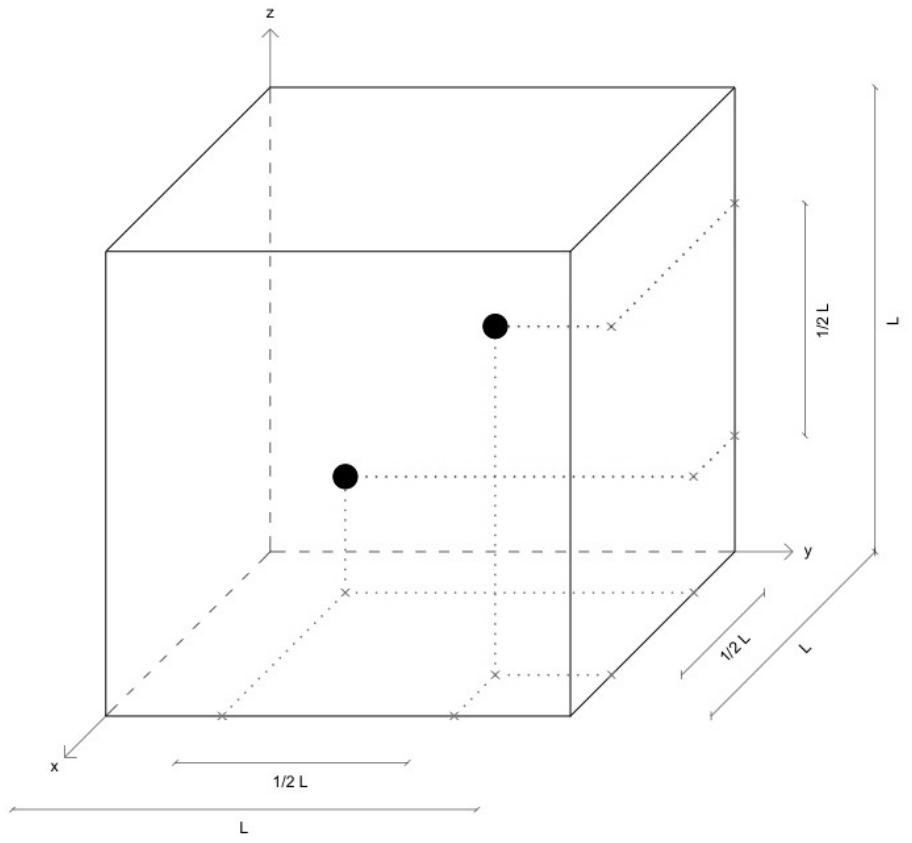


Figure 3.2: 3D representation of the atomic bcc primitive crystal cell (atm-bcc), where $L = r_s \left(\frac{4\pi N}{3}\right)^{1/3}$.

### 3.1.2 The Sampling Technique

We will now outline a technical aspect of our actual implementation of the $\mathrm{M(RT)}^2$ algorithm used within VMC.

It is well known that *global moves* lead to very strong autocorrelations between successive sampled points. A global move is a displacement of all the particles of the system:

$$R_{\mathrm{new}} = R_{\mathrm{old}} + \Delta\left(\eta_1, \dots \eta_{3N}\right) , \tag{3.2}$$

where $\eta_i$ are random numbers sampled with a flat distribution from the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$. As a consequence of the strong autocorrelations, the efficiency of the VMC algorithm decreases substantially. Therefore, a more convenient way to propose walker moves is desirable.

A significant improvement is achievable by means of two different modifications, both introduced in [84].

**Single-particle moves:** Only one particle is displaced before submitting the new position to the acceptance/rejection process. In other words, a random particle $l \in \{1, \dots N\}$ is selected and shifted by

$$\mathbf{r}_{l\,\mathrm{new}} = \mathbf{r}_{l\,\mathrm{old}} + \Delta\left(\eta_1, \eta_2, \eta_3\right) ; \tag{3.3}$$

**Gradient multi-particle moves:** The global move is not sampled from a flat distribution, but instead biased towards the maximum of the probability density function $|\Psi|^2$ by means of the gradient $\frac{\nabla\Psi}{\Psi}$. Such a modification of the transition probability requires an opportune correction of the acceptance probability, in order to avoid a bias in the final estimator.

In this work we have exploited the first possibility.

The implementation of single-particle moves is rather simple, and it only requires the modification of the subroutines that compute the wave function, in order to update them efficiently. While the update of a Jastrow pseudopotential is straightforward, an optimal method for SD is non trivial. However, such a procedure is available [84, 85], and reads

$$\mathrm{SD}_{\mathrm{new}} = \mathrm{SD}_{\mathrm{old}} \sum_j \left(A^{-1}\right)_{jl\,\mathrm{old}} A_{lj\,\mathrm{new}} , \tag{3.4}$$

where $A$ is the matrix that generates the SD, i.e $\det(A) = \mathrm{SD}$, and $l$ the moved particle. Similarly, also the inverse matrix $\left(A^{-1}\right)$ can be conveniently updated by means of

$$\begin{cases} \left(A^{-1}\right)_{il\,\mathrm{new}} = \left(A^{-1}\right)_{il\,\mathrm{old}} \frac{\mathrm{SD}_{\mathrm{old}}}{\mathrm{SD}_{\mathrm{new}}} \\ \left(A^{-1}\right)_{ij\,\mathrm{new}} = \left(A^{-1}\right)_{ij\,\mathrm{old}} - \left(A^{-1}\right)_{il\,\mathrm{old}} \frac{\mathrm{SD}_{\mathrm{old}}}{\mathrm{SD}_{\mathrm{new}}} \sum_s \left(A^{-1}\right)_{sj\,\mathrm{old}} A_{ls\,\mathrm{new}} , \quad j \neq l \end{cases} . \tag{3.5}$$

At the beginning of each simulation, $\Delta$ is set in order to have an acceptance rate of $\sim 50\%$. In fact, it is heuristic knowledge that such a rate provides a good efficiency. Moreover, we have performed $\frac{3}{2}N$ single-particle moves between every successive evaluation of the estimators, in order to reduce the autocorrelations.

to our code. In fact, assuming a JS-pw trial wave function (defined in the next subsection), $M = 10^5$ and a mol-hcp phase at $r_s = 2.61$ with $N = 16$, we obtained $E = -0.891(5)$ Ry (computed in 174 seconds) with single-particle moves, to be compared with $E = -0.887(9)$ (computed in 89 seconds). The corresponding efficiencies[1] are $23 \times 10^{-4}$ Ry$^{-1}$s$^{-2}$ and $14 \times 10^{-4}$ Ry$^{-1}$s$^{-2}$, therefore the efficiency gain is of almost a factor 2. This improvement is reflected also in the values of $\Delta$, which are 2.76 $a_0$ and 0.54 $a_0$ respectively. Moreover, the advantage provided by single-particle moves rise with $N$. For example, repeating the previous test with $N = 128$, the gained factor is equal to $\sim$20.

### 3.1.3 Trial Wave Functions

The trial wave functions considered in this work are assembled starting from the simple terms listed in the following:

**$J_{ee}$-YUK:** The *Yukawa Jastrow* is suitable for the description of the two-body electron-electron correlation [86]. It reads

$$J_{ee\text{-YUK}}(R) = \exp\left(-\frac{1}{2}\sum_{i<j}^{N}\frac{A_{ee}^{\sigma_i\sigma_j}\left(1-\exp\left(-F_{ee}^{\sigma_i\sigma_j}||\mathbf{r}_i-\mathbf{r}_j||\right)\right)}{||\mathbf{r}_i-\mathbf{r}_j||}\right), \tag{3.7}$$

where $\sigma_i$ represents the spin of the i-th particle, and $A_{ee}^{\uparrow\uparrow} \equiv A_{ee}^{\downarrow\downarrow}$, $A_{ee}^{\uparrow\downarrow} \equiv A_{ee}^{\downarrow\uparrow}$, $F_{ee}^{\uparrow\uparrow} \equiv F_{ee}^{\downarrow\downarrow}$, and $F_{ee}^{\uparrow\downarrow} \equiv F_{ee}^{\downarrow\uparrow}$ are variational parameters.

**$J_{ep}$-YUK:** The Yukawa Jastrow can also be used to take into account the proton-electron correlation:

$$J_{ep\text{-YUK}}(R,R_p) = \exp\left(-\frac{1}{2}\sum_{i,j=1}^{N}\frac{A_{ep}^{\sigma_i\sigma_j}\left(1-\exp\left(-F_{ep}^{\sigma_i\sigma_j}||\mathbf{r}_i-\mathbf{r}_{\mathbf{p}j}||\right)\right)}{||\mathbf{r}_i-\mathbf{r}_{\mathbf{p}j}||}\right), \tag{3.8}$$

where $R_p$ are the protonic coordinates and $A_{ep}^{\uparrow\uparrow} \equiv A_{ep}^{\downarrow\downarrow}$, $A_{ep}^{\uparrow\downarrow} \equiv A_{ep}^{\downarrow\uparrow}$, $F_{ep}^{\uparrow\uparrow} \equiv F_{ep}^{\downarrow\downarrow}$, and $F_{ep}^{\uparrow\downarrow} \equiv F_{ep}^{\downarrow\uparrow}$ are variational parameters.

**$\mathbf{SD_{pw}}$:** The *plane wave Slater determinant* is the ground state of a Fermi gas, and even if it is not appropriate for characterizing localized systems, it encloses the Fermi statistics and thus it allows for a simple representation of a fermionic system. Usually, it is convenient to decompose it into the product of a SD of the first $N/2$ spin-up electrons times the SD of the

[1] The efficiency $\eta$ of a MC calculation is defined as

$$\eta = \frac{1}{\text{time} \times \text{variance}}. \tag{3.6}$$

remaining spin-down electrons:

$$\mathrm{SD}_{\mathrm{pw}}(R) = \det\begin{pmatrix} e^{i\mathbf{k}_1\cdot\mathbf{r}_1} & e^{i\mathbf{k}_1\cdot\mathbf{r}_2} & \dots & e^{i\mathbf{k}_1\cdot\mathbf{r}_{\frac{N}{2}}} \\ e^{i\mathbf{k}_2\cdot\mathbf{r}_1} & e^{i\mathbf{k}_2\cdot\mathbf{r}_2} & \dots & e^{i\mathbf{k}_2\cdot\mathbf{r}_{\frac{N}{2}}} \\ \vdots & \vdots & \ddots & \\ e^{i\mathbf{k}_{\frac{N}{2}}\cdot\mathbf{r}_1} & e^{i\mathbf{k}_{\frac{N}{2}}\cdot\mathbf{r}_2} & \dots & e^{i\mathbf{k}_{\frac{N}{2}}\cdot\mathbf{r}_{\frac{N}{2}}} \end{pmatrix} \times \det\begin{pmatrix} e^{i\mathbf{k}_1\cdot\mathbf{r}_{\frac{N}{2}+1}} & e^{i\mathbf{k}_1\cdot\mathbf{r}_{\frac{N}{2}+2}} & \dots & e^{i\mathbf{k}_1\cdot\mathbf{r}_N} \\ e^{i\mathbf{k}_2\cdot\mathbf{r}_{\frac{N}{2}+1}} & e^{i\mathbf{k}_2\cdot\mathbf{r}_{\frac{N}{2}+2}} & \dots & e^{i\mathbf{k}_2\cdot\mathbf{r}_N} \\ \vdots & \vdots & \ddots & \\ e^{i\mathbf{k}_{\frac{N}{2}}\cdot\mathbf{r}_{\frac{N}{2}+1}} & e^{i\mathbf{k}_{\frac{N}{2}}\cdot\mathbf{r}_{\frac{N}{2}+2}} & \dots & e^{i\mathbf{k}_{\frac{N}{2}}\cdot\mathbf{r}_N} \end{pmatrix}, \tag{3.9}$$

where the $\mathbf{k}$-vectors have the form $\mathbf{k} = \frac{2\pi}{L}(i,j,k)$, with $i,j,k \in \mathbb{N}$. No variational parameters are involved.

**$\mathbf{SD_{DFT}}$:** The *DFT Slater determinant* embodies the solutions of the Kohn-Sham equation. Also this SD does not entail any variational parameter[2], even though one has to set the cutoff energy $E_{\mathrm{ctf}}$ used by Quantum-Espresso (Q-E) [1], our code of choice for solving the Kohn-Sham equation. $E_{\mathrm{ctf}}$ is used to determine the maximum momenta of the plane wave basis set. Given the first $N/2$ Kohn-Sham orbitals $\phi_i$, one can build the Slater determinant

$$\mathrm{SD}_{\mathrm{DFT}}(R) = \det\begin{pmatrix} \phi_1(\mathbf{r}_1) & \phi_1(\mathbf{r}_2) & \dots & \phi_1(\mathbf{r}_{\frac{N}{2}}) \\ \phi_2(\mathbf{r}_1) & \phi_2(\mathbf{r}_2) & \dots & \phi_2(\mathbf{r}_{\frac{N}{2}}) \\ \vdots & \vdots & \ddots & \\ \phi_{\frac{N}{2}}(\mathbf{r}_1) & \phi_{\frac{N}{2}}(\mathbf{r}_2) & \dots & \phi_{\frac{N}{2}}(\mathbf{r}_{\frac{N}{2}}) \end{pmatrix} \times \det\begin{pmatrix} \phi_1(\mathbf{r}_{\frac{N}{2}}+1) & \phi_1(\mathbf{r}_{\frac{N}{2}}+2) & \dots & \phi_1(\mathbf{r}_N) \\ \phi_2(\mathbf{r}_{\frac{N}{2}}+1) & \phi_2(\mathbf{r}_{\frac{N}{2}}+2) & \dots & \phi_2(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \\ \phi_{\frac{N}{2}}(\mathbf{r}_{\frac{N}{2}}+1) & \phi_{\frac{N}{2}}(\mathbf{r}_{\frac{N}{2}}+2) & \dots & \phi_{\frac{N}{2}}(\mathbf{r}_N) \end{pmatrix}. \tag{3.10}$$

The DFT solutions that we have used are always referred to the PBE exchange-correlation energy functional and the bare Coulomb pseudopotential.

**$\mathbf{SD_{dnfH}}$:** The *diagonalized nuclear field Hamiltonian Slater determinant* [87] is obtained by diagonalizing a simplified Hamiltonian matrix which contains the kinetic contribution in the diagonal components and the Coulomb interaction with the protons in the off-diagonal ones. In other words, the electrons are treated as non-interacting, and the protons provide an

[2]It is possible, however, to optimize all the orbitals coefficients provided by Quantum-Espresso. Nevertheless, this implies the optimization of thousands of parameters, and thus it requires a substantially larger computational cost. Therefore, we have not exploited this possibility.

external electric field. The Hamiltonian then reads

$$\begin{pmatrix} \hbar^2 \frac{\mathbf{k}_1^2}{2m} & -4\pi \frac{K_{\mathrm{C}} \sum_{j=1}^{N_F} e^{-i(\mathbf{k}_1-\mathbf{k}_2)\cdot \mathbf{r}_{\mathbf{p}j}}}{\mathcal{V}|\mathbf{k}_1-\mathbf{k}_2|^2} & \dots \\ -4\pi \frac{K_{\mathrm{C}} \sum_{j=1}^{N_F} e^{-i(\mathbf{k}_2-\mathbf{k}_1)\cdot \mathbf{r}_{\mathbf{p}j}}}{\mathcal{V}|\mathbf{k}_2-\mathbf{k}_1|^2} & \hbar^2 \frac{\mathbf{k}_2^2}{2m} & \dots \\ \vdots & \vdots & \\ -4\pi \frac{K_{\mathrm{C}} \sum_{j=1}^{N_F} e^{-i(\mathbf{k}_{N_F}-\mathbf{k}_1)\cdot \mathbf{r}_{\mathbf{p}j}}}{\mathcal{V}|\mathbf{k}_{N_F}-\mathbf{k}_1|^2} & -4\pi \frac{K_{\mathrm{C}} \sum_{j=1}^{N_F} e^{-i(\mathbf{k}_{N_F}-\mathbf{k}_2)\cdot \mathbf{r}_{\mathbf{p}j}}}{\mathcal{V}|\mathbf{k}_{N_F}-\mathbf{k}_2|^2} & \dots \end{pmatrix}, \tag{3.11}$$

where $K_{\mathrm{C}}$ is the Coulomb constant, and $\mathbf{k}_i$ are the first $N_F$ Fermi momenta. The eigenstates with the lowest eigenenergies are then used to generate a Slater determinant formally identical to the one in Equation (3.10). The advantage of employing these orbitals rather than the DFT ones, is given by the lower computational cost involved, and by the fact that they do not require an external code.

**$\Xi_{es}$:** The kernel connects the electronic coordinates with the associated shadows by means of the gaussian term

$$\Xi_{es}(R,S) = \exp\left(-C\sum_{i=1}^{N}\left(\mathbf{r}_i-\mathbf{s}_i\right)^2\right), \tag{3.12}$$

where $C$ is a variational parameter.

**$J_{ss}$-YUK:** The Yukawa Jastrow is suitable for the description of the shadow-shadow correlation. It reads

$$J_{ss\text{-YUK}}(S) = \exp\left(-\frac{1}{2}\sum_{i<j}^{N}\frac{A_{ss}^{\sigma_i\sigma_j}\left(1-\exp\left(-F_{ss}^{\sigma_i\sigma_j}||\mathbf{s}_i-\mathbf{s}_j||\right)\right)}{||\mathbf{s}_i-\mathbf{s}_j||}\right), \tag{3.13}$$

where $A_{ss}^{\uparrow\uparrow} \equiv A_{ss}^{\downarrow\downarrow}$, $A_{ss}^{\uparrow\downarrow} \equiv A_{ss}^{\downarrow\uparrow}$, $F_{ss}^{\uparrow\uparrow} \equiv F_{ss}^{\downarrow\downarrow}$, and $F_{ss}^{\uparrow\downarrow} \equiv F_{ss}^{\downarrow\uparrow}$ are variational parameters.

**$J_{sp}$-YUK:** Also the the shadow-proton correlation is taken into account by means of the Yukawa Jastrow

$$J_{sp\text{-YUK}}(R,R_p) = \exp\left(-\frac{1}{2}\sum_{i,j=1}^{N}\frac{A_{sp}^{\sigma_i\sigma_j}\left(1-\exp\left(-F_{sp}^{\sigma_i\sigma_j}||\mathbf{s}_i-\mathbf{r}_{\mathbf{p}j}||\right)\right)}{||\mathbf{s}_i-\mathbf{r}_{\mathbf{p}j}||}\right), \tag{3.14}$$

where $A_{sp}^{\uparrow\uparrow} \equiv A_{sp}^{\downarrow\downarrow}$, $A_{sp}^{\uparrow\downarrow} \equiv A_{sp}^{\downarrow\uparrow}$, $F_{sp}^{\uparrow\uparrow} \equiv F_{sp}^{\downarrow\downarrow}$, and $F_{sp}^{\uparrow\downarrow} \equiv F_{sp}^{\downarrow\uparrow}$ are variational parameters.

We now concisely introduce the trial wave functions considered in this work:

**JS-pw** is the simplest trial wave function, given by

$$\Psi_{\text{JS-pw}}(R,R_p) = \mathrm{SD}_{\mathrm{pw}}(R) J_{ee\text{-YUK}}(R) J_{ep\text{-YUK}}(R,R_p)\,; \tag{3.15}$$

**JS-DFT** is very commonly employed in electronic structure calculations, and reads

$$\Psi_{\text{JS-DFT}}(R,R_p) = \mathrm{SD}_{\mathrm{DFT}}(R) J_{ee\text{-YUK}}(R) J_{ep\text{-YUK}}(R,R_p)\,; \tag{3.16}$$

**JS-dnfH** is a trial wave function that employs a novel SD,

$$\Psi_{\text{JS-dnfH}}(R, R_p) = \text{SD}_{\text{dnfH}}(R) J_{ee\text{-YUK}}(R) J_{ep\text{-YUK}}(R, R_p)\,; \tag{3.17}$$

**ASWF-pw** has been already presented in section 1.3,

$$\begin{aligned}\Psi_{\text{ASWF-pw}}(R, R_p) &= \text{SD}_{\text{pw}}(R) J_{ee\text{-YUK}}(R) J_{ep\text{-YUK}}(R, R_p) \times \\ &\int dS\, \Xi_{es}(R, S) J_{ss\text{-YUK}}(S) J_{sp\text{-YUK}}(S, R_p)\,;\end{aligned} \tag{3.18}$$

**ASWF-DFT** employs the DFT orbitals rather than simple plane waves,

$$\begin{aligned}\Psi_{\text{ASWF-DFT}}(R, R_p) &= \text{SD}_{\text{DFT}}(R) J_{ee\text{-YUK}}(R) J_{ep\text{-YUK}}(R, R_p) \times \\ &\int dS\, \Xi_{es}(R, S) J_{ss\text{-YUK}}(S) J_{sp\text{-YUK}}(S, R_p)\,;\end{aligned} \tag{3.19}$$

**FSWF-pw** has been introduced in section 1.3,

$$\begin{aligned}\Psi_{\text{FSWF-pw}}(R, R_p) &= J_{ee\text{-YUK}}(R) J_{ep\text{-YUK}}(R, R_p) \int dS\, \Xi_{es}(R, S) \times \\ &\text{SD}_{\text{pw}}(S) J_{ss\text{-YUK}}(S) J_{sp\text{-YUK}}(S, R_p)\,.\end{aligned} \tag{3.20}$$

**FSWF-DFT** takes advantage of the DFT orbitals,

$$\begin{aligned}\Psi_{\text{FSWF-DFT}}(R, R_p) &= J_{ee\text{-YUK}}(R) J_{ep\text{-YUK}}(R, R_p) \int dS\, \Xi_{es}(R, S) \times \\ &\text{SD}_{\text{DFT}}(S) J_{ss\text{-YUK}}(S) J_{sp\text{-YUK}}(S, R_p)\,.\end{aligned} \tag{3.21}$$

As we have already mentioned, the DFT orbitals are determined within a given cutoff energy, which must be specified in the input file of Q-E. A high cutoff implies accurate orbitals, but also a long computational time both for Q-E and our VMC code, because of the large basis set. Therefore, a good compromise must be found. In order to determine the optimal $E_{\text{ctf}}$, we have evaluated the variational energy for different values of $E_{\text{ctf}}$, as shown in Figure 3.3. From this analysis it emerges that when a Jastrow in used in conjunction with $\text{SD}_{\text{DFT}}$, the variational energy is almost independent from $E_{\text{ctf}}$. Thus, it is sufficient to select the minimum $E_{\text{ctf}}$ necessary to reach the convergence within Q-E.

The same discussion about the cutoff energy holds for the JS-dnfH trial wave function, with analogous conclusions. In this case, we replaced the cutoff energy with the factor $F_{\text{ctf}}$, which is related to the basis set used to diagonalize the dnfH matrix. More specifically, we have considered plane waves up to the momentum $F_{\text{ctf}} \times n_{\text{shell}}$, where $n_{\text{shell}}$ is the highest momenta shell employed in the $\text{SD}_{\text{pw}}$. In Figure 3.4 we report some extensive calculations obtained with different $F_{\text{ctf}}$. The results demonstrate that there is a fundamental problem, since for the mol-hcp phase higher $F_{\text{ctf}}$ corresponds to significantly lower variational energies. The reason for this unexpected effect may originate from the fact that, in the covalent bond, the interaction of the electrons with the protons in the molecule is screened, and this aspect is completely neglected by the approximated Hamiltonian in equation (3.11). Because of this limitation, we did not further investigate this trial wave function, as it requires a fundamental revision. Nevertheless, it is very encouraging that the atm-bcc phase is better represented by the JS-dnfH rather than by the JS-DFT, considering also the lower computational cost involved.

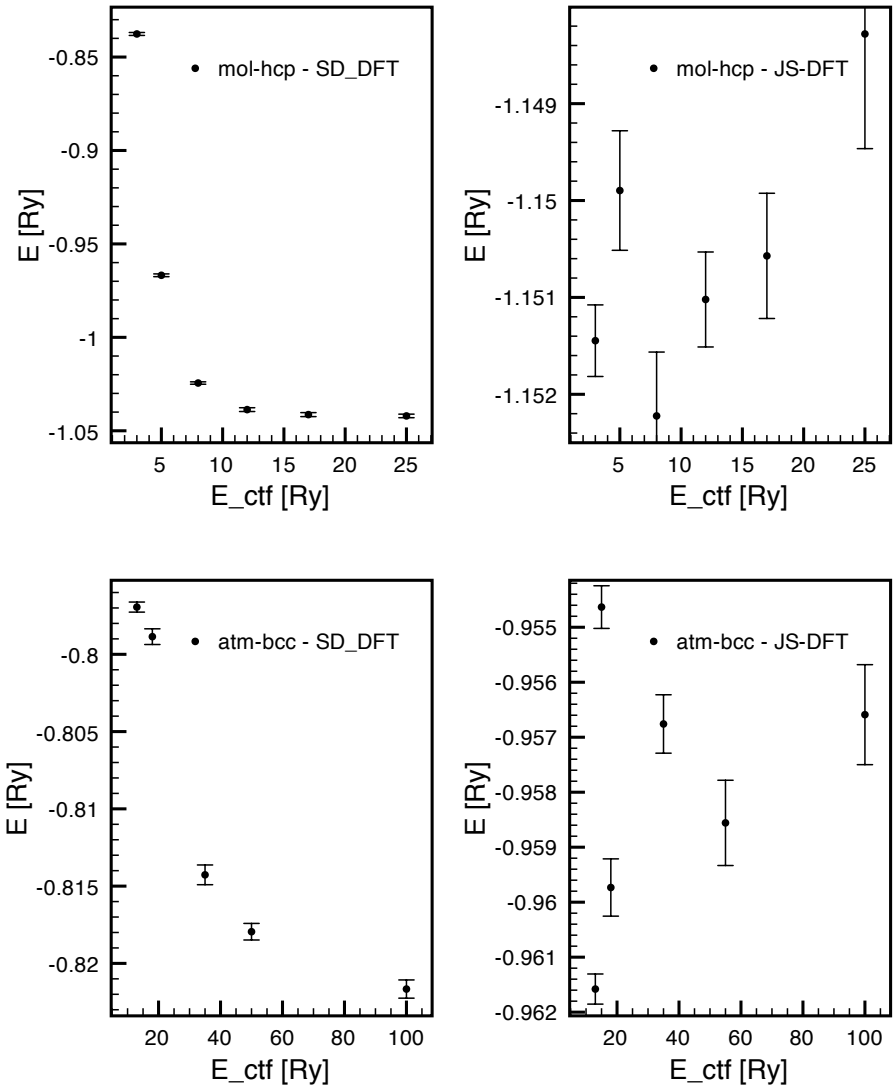


Figure 3.3: Relation between the variational energy $E$ and the cutoff energy used within Quantum-Espresso. The data refers to two different cases: The mol-hcp phase at relatively low density $r_s = 2.61$, and the atm-bcc one at high density $r_s = 1.31$. In both systems we have considered both the simple Slater determinant $\mathrm{SD}_{\mathrm{DFT}}$ (left column) and the complete JS-DFT wave function (right column). All the simulations have been executed with $N = 16$.

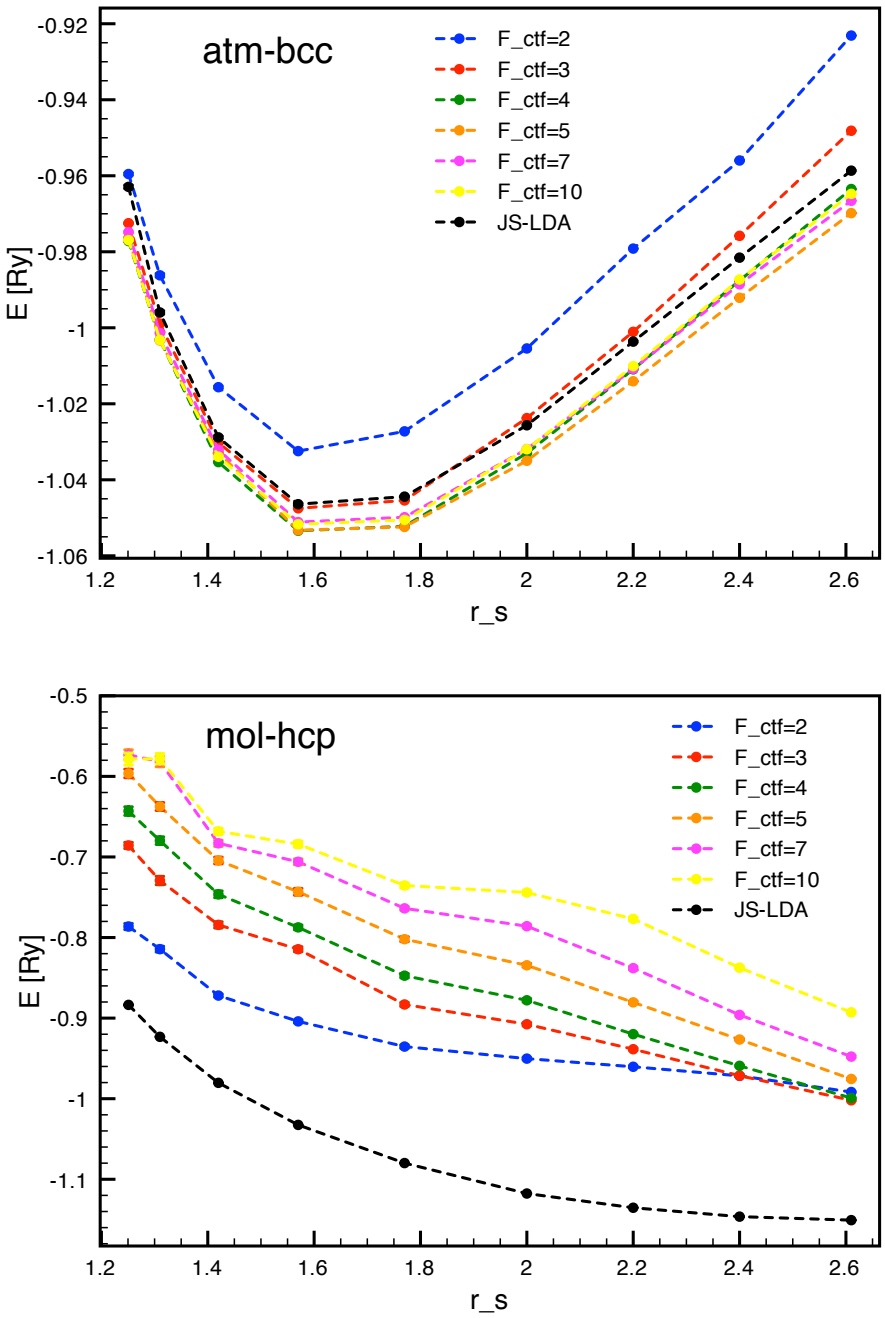


Figure 3.4: JS-dnfH variational energy for 128 hydrogen atoms in the mol-hcp and atm-bcc phase, employing different $F_{\mathrm{ctf}}$ and $r_s$.

### 3.1.4 Ewald Summation

The Coulomb potential energy within periodic boundary conditions cannot be computed straightforwardly, but it requires some extra efforts. In the following we explain the difficulties involved, and the most commonly used technique to overcome them: The *Ewald summation* [88, 89].

The direct evaluation of the Coulomb energy in a periodic cell reads

$$E_{\text{pot}} = K_C \frac{1}{2} \sum_{\mathbf{R}} \sum_{i=1}^{N} \sum_{j \neq i, j=1}^{N} \frac{q_i q_j}{||\mathbf{r}_i - \mathbf{r}_j + \mathbf{R}||} , \tag{3.22}$$

where $q_i$ is the charge of the $i$-th particle, $K_C$ the Coulomb constant, and $\sum_{\mathbf{R}}$ refers to the sum over the periodic lattice vectors ($\mathbf{R}_0 = (0,0,0)$, $\mathbf{R}_1 = (L_x, 0, 0)$, $\mathbf{R}_2 = (0, L_y, 0)$ etc.). The slow decay $\propto 1/r$ implies a large number of vectors $\mathbf{R}$ in order to attain an accurate estimation of $E_{\text{pot}}$, making it very demanding in terms of computational time.

In 1921, Ewald suggested to decompose the Coulomb potential in a short and long range contributions [88]:

$$v(r) = \underbrace{\text{erfc}(\alpha r) v(r)}_{\text{short range}} + \underbrace{\text{erf}(\alpha r) v(r)}_{\text{long range}} . \tag{3.23}$$

In this form, the sum over $\mathbf{R}$ is no more necessary, since the short range potential decays after a short distance, whereas the long range part can be conveniently computed in the $k$-space, i.e. by means of the Fourier series. In fact, the Fourier transform of the long range potential is

$$\tilde{v}_{\text{L}}(k) = K_C \frac{4\pi}{k^2} e^{-\frac{k^2}{4\alpha^2}} , \tag{3.24}$$

and permits a truncation at relatively small $k$. We then recast the Equation (3.22) as

$$\begin{aligned} E_{\text{pot}} \quad = \quad & K_C \frac{1}{2} \sum_{\mathbf{R}} \sum_{i=1}^{N} \sum_{j \neq i, j=1}^{N} \frac{q_i q_j}{||\mathbf{r}_i - \mathbf{r}_j + \mathbf{R}||} \text{erfc}(\alpha ||\mathbf{r}_i - \mathbf{r}_j + \mathbf{R}||) + \\ & K_C \frac{1}{2} \sum_{\mathbf{R}} \sum_{i=1}^{N} \sum_{j \neq i, j=1}^{N} \frac{q_i q_j}{||\mathbf{r}_i - \mathbf{r}_j + \mathbf{R}||} \text{erf}(\alpha ||\mathbf{r}_i - \mathbf{r}_j + \mathbf{R}||) . \end{aligned} \tag{3.25}$$

In the first term, for a small enough $\alpha$ (a typical value is $5/L$), the sum over $\mathbf{R}$ is not required anymore, since the complementary error function vanishes rapidly. Finally, in order to introduce the Fourier series in the second term, it is sufficient to include the case $i = j$ to the summation, and subtract it afterwards. In conclusion, we obtain

$$\begin{aligned} E_{\text{pot}} \quad = \quad & K_C \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i, j=1}^{N} q_i q_j \frac{\text{erfc}(\alpha ||\mathbf{r}_i - \mathbf{r}_j||)}{||\mathbf{r}_i - \mathbf{r}_j||} + \\ & K_C \frac{1}{2\mathcal{V}} \sum_{\mathbf{k} \neq \mathbf{0}} \sum_{i,j} q_i q_j e^{i\mathbf{k}(\mathbf{r}_i - \mathbf{r}_j)} \tilde{v}_{\text{L}}(k) - \sqrt{\frac{\alpha}{\pi}} \sum_{i=1}^{N} q_i^2 , \end{aligned} \tag{3.26}$$

where the summed $\mathbf{k}$ are such that they fulfill the condition $\mathbf{k} \cdot \mathbf{R} = 2\pi n$, with $n \in \mathbb{N}$.

### 3.1.5 Periodic Coordinates

The Yukawa Jastrow, like any other slowly decaying Jastrow, causes a spurious bias in the kinetic energy if it is computed in its straightforward version. In fact one should take into account the contributions given by many mirror cells in order to avoid discontinuities in the derivatives of the wave function when the particle distances switch from one closest image to the other. Needless to say that this approach is time consuming, and a more convenient strategy is desirable.

Before presenting the solution to the aforementioned problem, we introduce a useful check to verify if the correlations are correctly treated within the periodic boundary conditions. We start by integrating by parts the expression for the kinetic energy (for simplicity we consider the kinetic contribution of only one particle $j$)

$$\begin{aligned}\int_\Omega dR\,\psi^*(R)\nabla_j^2\psi(R) &= \sum_{x_\alpha=x_j,y_j,z_j}\int_\Omega dR\,\psi^*(R)\left(\frac{\partial^2}{\partial x_\alpha^2}\right)\psi(R)\\ &= \sum_{x_\alpha=x_j,y_j,z_j}\int_\Omega dR^{\bar{x}_\alpha}\int_{-L_{x_\alpha}/2}^{L_{x_\alpha}/2} dx_\alpha\,\psi^*(R)\frac{\partial^2}{\partial x_\alpha^2}\psi(R)\\ &= \sum_{x_\alpha=x_j,y_j,z_j}\left\{\int_\Omega dR^{\bar{x}_\alpha}\left[\psi^*(R)\frac{\partial}{\partial x_\alpha}\psi(R)\right]_{-L_{x_\alpha}/2}^{L_{x_\alpha}/2} - \right.\\ &\quad \left.\int_\Omega dR\left(\frac{\partial}{\partial x_\alpha}\psi^*(R)\right)\left(\frac{\partial}{\partial x_\alpha}\psi(R)\right)\right\}, \qquad (3.27)\end{aligned}$$

where $dR^{\bar{x}_\alpha}$ is the same as $dR$ but excluding the infinitesimal element $dx_\alpha$, $\Omega$ represents the domain of integration, i.e., the simulation box, and $L_{x_\alpha}$ is the length of the edge of $\Omega$ along the $x_\alpha$ axis. When working with periodic boundary conditions, the wave function $\psi(R)$ and its derivatives are required to be periodic, meaning that they are invariant for particle translations $\mathbf{v} = (n_x L_x, n_y L_y, n_z L_z)$ with $n_x$, $n_y$, and $n_z$ integers. Hence the term

$$\left[\psi^*(R)\frac{\partial}{\partial x_\alpha}\psi(R)\right]_{-L_{x_\alpha}/2}^{L_{x_\alpha}/2} \qquad (3.28)$$

vanishes, and we obtain the modified *Jackson-Feenberg kinetic energy* expression[3]

$$E_{\mathrm{JF}} = \hbar^2\sum_{j=1}^{N}\frac{1}{2m_j}\int_\Omega dR\,\nabla_j\psi^*(R)\cdot\nabla_j\psi(R)\,. \qquad (3.29)$$

As a consequence, the Pandharipande-Bethe (PB)

$$E_{\mathrm{PB}} = -\hbar^2\sum_{i=1}^{N}\frac{1}{2m_j}\int_\Omega dR\,\psi^*(R)\nabla_j^2\psi(R) \qquad (3.30)$$

[3]The original Jackson-Feenberg expression is:

$$\frac{\hbar^2}{2}\sum_{j=1}^{N}\frac{1}{2m_j}\int_\Omega dR\,\left(\nabla_j\psi^*(R)\cdot\nabla_j\psi(R) - \psi^*(R)\nabla_j^2\psi(R)\right).$$

and Jackson-Feenberg (JF) expressions must be equivalent. The equivalence of the JF and PB kinetic energies is a *necessary but not sufficient condition* for the required properties of the wave function. In all our calculations we have computed both expressions and verified that, within the MC error bars, they were equal. In the following we refer to this comparison as the *JF test*.

Equation 3.28 must be correctly interpreted when correlation terms, like a Jastrow, are introduced since the inter-particle distances are computed using the closest image. In fact, even though the particle coordinate $\mathbf{r}_j$ range within the box of volume $\mathcal{V} = L_x L_y L_z$, the distance $\mathbf{r}_{ij}$ between the particle $i$ (assumed as fixed) and $j$ does not range within

$$\sqrt{(\mathbf{r}_j - \mathbf{r}_i)^2} \qquad \text{with } \mathbf{r}_j \text{ spanning } \mathcal{V} \tag{3.31}$$

but always within the box of volume $\mathcal{V}$ centered on the particle $i$. This concept is illustrated in Figure 3.5. It is then possible and convenient to fix the origin on the position of the $i$-th particle that we are analyzing. In the following we will then set $\mathbf{r}_i = 0$, so that $\mathbf{r} = (x, y, z) \equiv \mathbf{r}_{ij} = \mathbf{r}_j$ and $r = \sqrt{x^2 + y^2 + z^2}$.

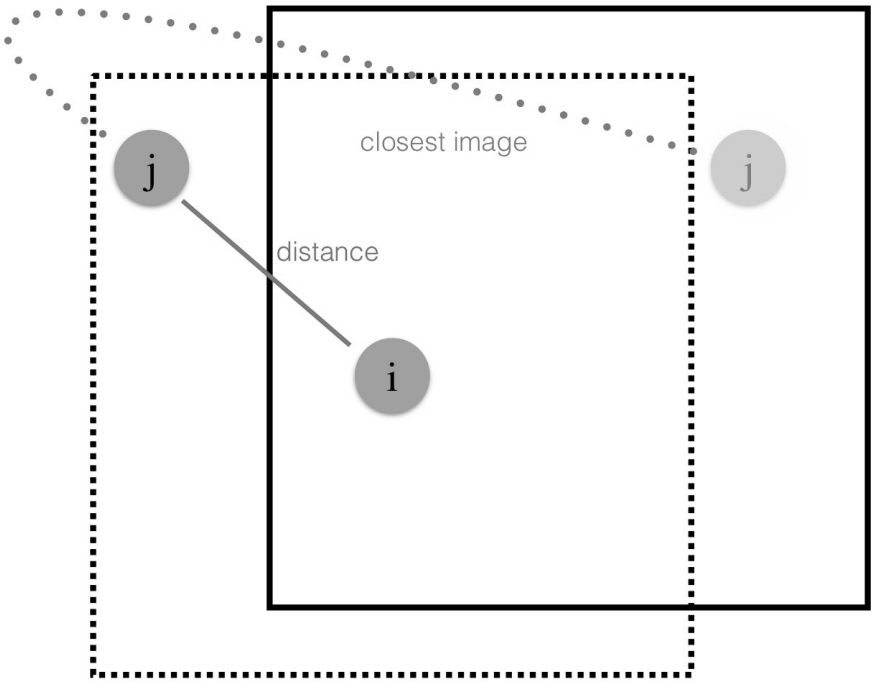


Figure 3.5: Integration of $\mathbf{r}_j$ within a box with periodic boundary conditions, from the point of view of the particle $i$. The continuous black line represents the simulation box, whereas the dotted black line represents the effective volume of integration for the distance between the particles $i$ and $j$.

We now want to demonstrate that the Yukawa Jastrow does not fulfill the JF test. In order to do so, we consider the simple case of only two interacting

particles, assuming

$$J(r) = e^{-\frac{A\left(1-e^{-Fr}\right)}{r}} . \tag{3.32}$$

Its first derivative along the $x$ axis reads

$$\begin{aligned}
\frac{\partial J(r)}{\partial x} &= \frac{\partial e^{-\frac{A\left(1-e^{-Fr}\right)}{r}}}{\partial x} \\
&= \frac{\partial e^{-\frac{A\left(1-e^{-Fr}\right)}{r}}}{\partial r}\frac{\partial r}{\partial x} \\
&= e^{-\frac{A\left(1-e^{-Fr}\right)}{r}}\left(A\frac{1-e^{-Fr}}{r^2} - AF\frac{e^{-Fr}}{r}\right)\frac{x}{r} \\
&= \frac{\partial J(r)}{\partial r}\frac{x}{r} .
\end{aligned} \tag{3.33}$$

It is then manifest that the first derivative is not continuous in $x = \pm L/2$, which is the border between two closest images. As a consequence

$$\lim_{\varepsilon\to 0}\left[J(r)\frac{\partial}{\partial x}J(r)\right]_{-(L_x/2)+\varepsilon}^{(L_x/2)-\varepsilon} = J(r)\frac{\partial J(r)}{\partial r}\frac{L_x}{r} \neq 0 . \tag{3.34}$$

Therefore, the JF and PB kinetic energies cannot be expected to have the same value, since the term in Equation 3.28 does not vanish. Notice that if $\frac{1}{r}\frac{\partial J}{\partial r}$ is small enough in $x = \pm\frac{L}{2}$, then this effect is negligible.

A deeper understanding can be obtained by means of the distribution theory. In fact, one can notice that

$$\frac{\partial^2}{\partial x^2}J(r) = \frac{x}{r}\left(\frac{\partial^2 J(r)}{\partial r^2}\frac{x}{r} + \frac{\partial J(r)}{\partial r}\frac{1}{r} - \frac{\partial J(r)}{\partial r}\frac{1}{r^2}\right) - \left(\frac{\partial J(r)}{\partial x}\frac{x}{r}\right) 2\,\delta(x - \frac{L_x}{2}) , \tag{3.35}$$

since

$$\int_{L_x/2-\varepsilon}^{L_x/2+\varepsilon}\frac{\partial}{\partial x}\left(\frac{\partial J(r)}{\partial x}\right) = -\frac{\partial J(r)}{\partial x}\frac{L_x}{r} . \tag{3.36}$$

In simple words, the discontinuity in the first derivative produces a Dirac delta in the second derivative. It is then clear that such an artifact must be prevented in order to avoid bias in the computation of the kinetic energy.

We remark that a discontinuity in the first derivative does not affect only the validity of the JF expression, but also of the PB one. In other words, the kinetic energy contribution provided by the Yukawa Jastrow is biased. Mathematically, the problem can be eliminated by enforcing a smooth change between closest images. Physically, this problem appears because the simulation box is not large enough to contain the correlation between the particles.

A well-known solution to this problem consists of emulating the Ewald summation decomposition [90]. More specifically, the Jastrow is decomposed into a quickly and a slowly decaying part, which are computed separately in the real coordinates and in the $k$-space respectively. However, this method requires a summation over the momentum space, which is rather expensive in terms of computational time.

A different possibility is provided by exploiting the *Periodic Coordinates* (PC), first introduced by C. Attaccalite and S. Sorella [80]. We repute this solution to be the best option for its elegance, effectiveness and simplicity. The only modification required is the substitution of the coordinates

$$x' = \frac{L}{\pi}\sin(\frac{\pi x}{L})\,, \tag{3.37}$$

$$y' = \frac{L}{\pi}\sin(\frac{\pi y}{L})\,, \tag{3.38}$$

$$z' = \frac{L}{\pi}\sin(\frac{\pi z}{L})\,, \tag{3.39}$$

and the consequent evaluation of the distances as

$$r' = \frac{L}{\pi}\sqrt{\sin^2(\frac{\pi x}{L}) + \sin^2(\frac{\pi y}{L}) + \sin^2(\frac{\pi z}{L})}\,. \tag{3.40}$$

The employment of the Periodic Coordinates enforces the correct periodicity of the wave function. For example, the first derivative

$$\frac{\partial J(r')}{\partial x} = \frac{\partial J(r')}{\partial r'}\frac{\partial r'}{\partial x'}\frac{\partial x'}{\partial x} \tag{3.41}$$

$$= \frac{\partial J(r')}{\partial r'}\frac{x'}{r'}\cos(\frac{\pi x}{L})\,, \tag{3.42}$$

is continuous in $x = \pm L/2$, i.e. on the borders of the simulation box. The same property holds also for the higher order derivatives. Figure 3.6 illustrates the modification introduced by the Periodic Coordinates in the Yukawa Jastrow.

As demonstration for the effectiveness of the PC, we have calculated the kinetic energy of 16 atoms of hydrogen employing the JS-PW trial wave function for describing the atm-bcc structure, with $r_s = 1.31$. We obtained (in 54 minutes)

$$\text{[without PC]} \qquad \begin{array}{rcl} E_{\rm kin} &=& 5.5766(5) \\ E_{\rm JF} &=& 2.9841(32) \end{array} \tag{3.43}$$

without using the PC, whereas, by applying them, we attained (in 56 minutes)

$$\text{[with PC]} \qquad \begin{array}{rcl} E_{\rm kin} &=& 1.5480(6) \\ E_{\rm JF} &=& 1.5454(16) \end{array}\,. \tag{3.44}$$

The success is confirmed for the mol-hcp lattice at $r_s = 2.61$:

$$\text{[without PC]} \qquad \begin{array}{rcl} E_{\rm kin} &=& 2.2428(6) \\ E_{\rm JF} &=& 2.1252(10) \end{array} \tag{3.45}$$

$$\text{[with PC]} \qquad \begin{array}{rcl} E_{\rm kin} &=& 1.0307(11) \\ E_{\rm JF} &=& 1.0290(5) \end{array} \tag{3.46}$$

In conclusion, the PC eliminates completely the problem, with a negligible additional computational cost. We remark that the PC modify the functional form of the Yukawa Jastrow, which it may then become unsuitable for describing the electron-electron and the electron-proton correlations. However, the suitability of the Jastrow embedding Periodic Coordinates can be checked by means of the variational principle. In practice, PC performs very well, as be seen by comparing our variational energy with the results found in the literature (see subsection 3.2.1).

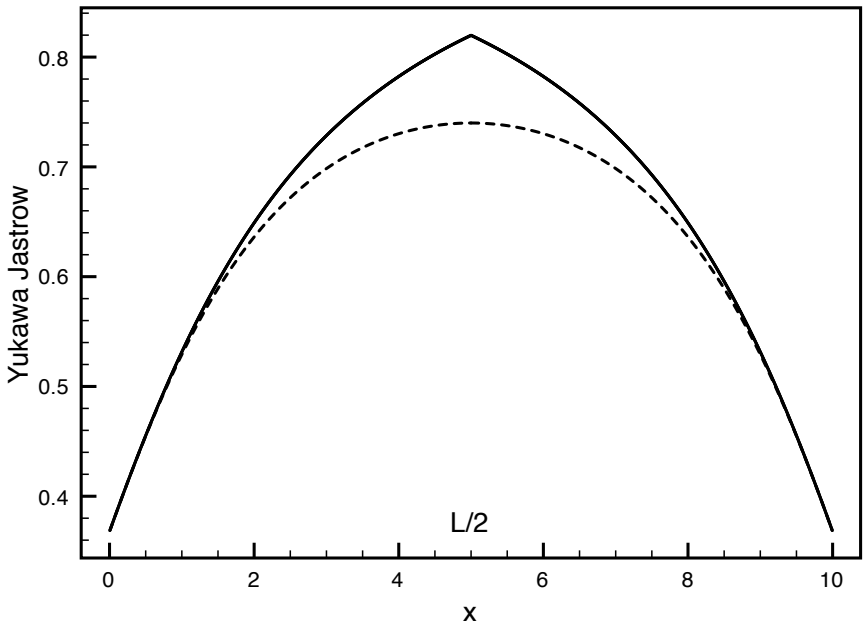


Figure 3.6: Comparison between the Jukawa Jastrow factor for two interacting particles $\exp\left(-\frac{A(1-\exp(-Fr))}{r}\right)$ (black line) and its modified version obtained by implementing the Periodic Coordinates (dashed line). We have set $y = z = 0$, $L = 10$, and $A = F = 1$.

### 3.1.6 SWF Kernel Truncation

If the kernel of the SWF is loose, i.e. the variational parameter $C$ is very small, the simulation box is not big enough to constrain each particle to its corresponding shadows $\mathbf{s}_1$ and $\mathbf{s}_2$ within the limit $L/2$. This eventuality must be prevented, because it introduces spurious terms in the kinetic energy, as it is revealed by comparing the Pandharipande-Bethe kinetic energy with the Jackson-Feenberg expression[4]:

$$E_{\text{kin}} = 1.624(4) \tag{3.47}$$
$$E_{\text{JF}} = 1.478(5)\,. \tag{3.48}$$

In order to eliminate this problem, the kernel must be modified so that it vanishes for $|\mathbf{r}-\mathbf{s}| \to L/2$. In practice, remembering the expression for the Kernel

$$\Xi_{es}(R,S) = e^{-C\sum_{i=1}^{N}(\mathbf{r}_i-\mathbf{s}_i)^2}\,, \tag{3.49}$$

we have connected the exponent $Cx^2$ to the function

$$f(x) = \alpha_0 + \frac{\alpha_1}{x - L/2}\,, \tag{3.50}$$

and required the continuity of the function and its first and second derivatives in the connecting point $x = l$, i.e.

$$\begin{cases} Cl^2 = f(l) \\ 2Cl = \left.\frac{df(x)}{dx}\right|_{x=l} \\ 2C = \left.\frac{d^2f(x)}{dx^2}\right|_{x=l} \end{cases}. \tag{3.51}$$

The solution of this system of three equations and three variables is

$$\begin{cases} l = \frac{L}{6} \\ \alpha_0 = -\frac{CL^2}{12} \\ \alpha_1 = -\frac{CL^3}{27}\,. \end{cases} \tag{3.52}$$

Therefore, the modified expression for the kernel reads

$$\Xi_{es}(R,S) = \prod_{i=1}^{N} e^{-\varsigma(|\mathbf{r}_i-\mathbf{s}_i|)}\,, \tag{3.53}$$

where

$$\varsigma(x) = \begin{cases} Cx^2 & \text{if } x \le \frac{L}{6} \\ \alpha_0 + \frac{\alpha_1}{x-L/2} & \text{if } \frac{L}{6} < x \le \frac{L}{2} \\ +\infty & \text{if } x > \frac{L}{2} \end{cases}. \tag{3.54}$$

This modification is sufficient to solve the aforementioned problem, but is not practical on a computer, because of the infinite term. This difficulty can

[4]The following estimated kinetic energies have been computed employing ASWF-pw on 16 hydrogen atoms in the atm-bcc phase at $r_s = 1.31$. The employed variational parameters are: $A_{ee}^{\uparrow\uparrow} = 0.423$, $A_{ee}^{\uparrow\downarrow} = 0.829$, $F_{ee}^{\uparrow\uparrow} = 2.568$, $F_{ee}^{\uparrow\downarrow} = 1.834$, $A_{ep}^{\uparrow\uparrow} = -74.930$, $A_{ep}^{\uparrow\downarrow} = -68.191$, $F_{ep}^{\uparrow\uparrow} = 0.231$, $F_{ep}^{\uparrow\downarrow} = 0.242$, $C = 0.542$, $A_{ss}^{\uparrow\uparrow} = 2.400$, $A_{ss}^{\uparrow\downarrow} = 2.112$, $F_{ss}^{\uparrow\uparrow} = 5.508$, $F_{ss}^{\uparrow\downarrow} = 19.039$, $A_{sp}^{\uparrow\uparrow} = 2.400$, $A_{sp}^{\uparrow\downarrow} = 2.112$, $F_{sp}^{\uparrow\uparrow} = 5.508$, $F_{sp}^{\uparrow\downarrow} = 19.039$.

be circumvented by refusing any proposed move that leads to $x \geq L/2$, or by setting $\varsigma(x)$ equal to a constant for $(x > L/2 - \varepsilon)$, with $\varepsilon \ll 1$, i.e. by assuming

$$\varsigma(x) = \begin{cases} Cx^2 & \text{if } x \leq \frac{L}{6} \\ \alpha_0 + \frac{\alpha_1}{x - L/2} & \text{if } \frac{L}{6} < x \leq \frac{L}{2} - \varepsilon \\ \alpha_0 + \frac{\alpha_1}{-\varepsilon} & \text{if } x > \frac{L}{2} - \varepsilon \end{cases} . \tag{3.55}$$

However, a more sophisticated version of the truncated kernel is necessary if the Gaussian Determinant (presented in the subsection 4.2.1) is used. In fact, in that case the condition $x > L/2$ is not only possible, but also unpreventable, due to the sum of the permuted terms. Thus, it is necessary to connect the divergent term with a function that grows quickly, but that does not diverge. An appropriate choice is

$$\varsigma(x) = \begin{cases} Cx^2 & \text{if } x \leq \frac{L}{6} \\ \alpha_0 + \frac{\alpha_1}{x - L/2} & \text{if } \frac{L}{6} < x \leq \frac{L}{2} - \varepsilon \\ \beta_0 + \beta_1 x^n & \text{if } x > \frac{L}{2} - \varepsilon \end{cases} , \tag{3.56}$$

with

$$\begin{cases} \varepsilon = \frac{L}{n+1} \\ \beta 1 = -\frac{2\alpha_1}{\varepsilon^3 n(n-1)\left(\frac{L}{2} - \varepsilon\right)^{n-2}} \\ \beta_0 = \alpha_0 - \frac{\alpha_1}{\varepsilon} - \beta_1 \left(l - \varepsilon\right)^n \end{cases} \tag{3.57}$$

and $n \geq 2$. All our simulations have proven that a suitable choice is $n = 12$. Figure 3.7 illustrates the modification introduced by equation (3.56).

We recompute now the kinetic energies in equations (3.47) and (3.48), employing the truncated kernel reported in equation (3.56), obtaining:

$$E_{\text{kin}} = 1.538(6) \tag{3.58}$$

$$E_{\text{JF}} = 1.533(6) . \tag{3.59}$$

This result demonstrates that the suggested truncation completely solves the problem.

### 3.1.7 Finite Size Effects

As we have already discussed in section 3.1.4, the periodic boundary conditions do not automatically allow for a proper description of an infinite system. In fact, simulations that employ different $N$ provide sensibly different results. This occurrence goes under the name of *finite-size effect*. We now present a method that permits its reduction, called *Twist Averaged Boundary Conditions* (TABC) [79].

The TABC have been devised in the context of metallic systems, wherein the Slater determinant of plane waves provides a good description of the electronic structure. However, the embedded **k**-vectors do not well represent an infinite system, since in general a discrete grid of points cannot reproduce the whole Fermi sphere (see Figure 3.8). The TABC method amends this limitation by means of an integration over the Fermi sphere, prescribing a recurrent random shift

$$\mathbf{v}_{\text{twist}} = \frac{2\pi}{L} \left(\eta_1, \eta_2, \eta_3\right) \tag{3.60}$$

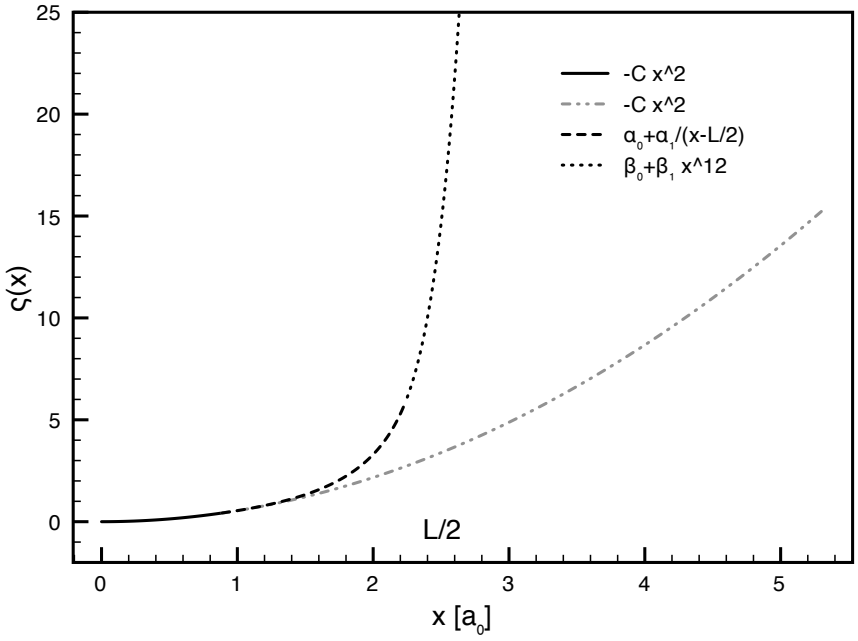


Figure 3.7: Illustration of the kernel truncation prescribed in equation (3.56) with $C = 0.542$, for 16 hydrogen atoms in the atm-bcc phase at $r_s = 1.31$ ($L/2 \simeq 2.66$ a$_0$).

of the **k**-grid, where $\eta_i$ are random numbers sampled in the range $\left[-\frac{1}{2}, \frac{1}{2}\right]$. In the context of the TABC, such a $\mathbf{v}_{\text{twist}}$ translation is called *twist*. By their introduction, an improved convergence to the thermodynamic limit is achieved, as can be seen in Figure 3.9. In summary, the TABC method consists of an integral over the **k**-space, computed by means of Monte Carlo.

The TABC not only have the merit of reducing the finite size effects, but also permit for calculations with a number of particles $N$ different from a magic number[5], without spurious drifts and anisotropy effects.

A VMC simulation of a 3D unpolarized system employing the TABC with $N_{\text{twist}}$ twists is outlined in the following points:

1. Determine the next magic number larger than or equal to $N/2$, and then select the next one, which we label as $n$. Examples: if $N = 54$ then $n = 33$, if $N = 128$ then $n = 93$;
2. Find the first $n$ Fermi $k$-vectors, obtaining $\Gamma_0 = \{\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_n\}$;
3. Generate $\mathbf{v}_{\text{twist}}$ as in Equation (3.60);
4. $K = \Gamma_0 + \{\mathbf{v}_{\text{twist}}\}_n$;
5. Sort the $k$-vectors in $K$ by increasing magnitude, and then use the first $N/2$ $k$-vectors to build up $\text{SD}_{\text{pw}}$;
6. Perform $M_{\text{relax}}$ relaxation steps;
7. Sample $M/N_{\text{twist}}$ points and accumulate the estimators of the observables of interest (normally the kinetic and potential energies);
8. Repeat the points 3-7 $N_{\text{twist}}$ times.

The $M_{\text{relax}}$ relaxation steps are necessary to prevent the possibility of a bias in the calculation. This process could be avoided by submitting the twist to the acceptance/refuse process of the $\text{M(RT)}^2$ algorithm, but we have not exploited this possibility, since the number of relaxation steps is small and its computational cost negligible.

A momentum in the external shell, initially not included in the $\text{SD}_{\text{pw}}$, may occur to have a lower magnitude than the employed ones, as a consequence of the twist. Consequentially, such a momentum replaces the one with the actual highest magnitude. In order to account for this eventuality, more $k$-vectors than strictly necessary to generate $\text{SD}_{\text{pw}}$ are considered in step 1, and then selected accordingly in point 5. This procedure generates the energy lines reported in Figure 3.10.

We now present our extension of the TABC to the $\text{SD}_{\text{DFT}}$. In fact, the DFT method also suffers from the finite-size effects, and requires a sum of contributions from different $K$-points in the Brillouin zone. In particular, Quantum-Espresso allows for considering a grid of $K$ points to average over. The simplest grid simply consists of one point, labelled as $\Gamma_0$, which corresponds to the Fermi gas momenta. However, finite-size effects can be drastically reduced by considering several $K$-points, in order to have a better average, similarly to TABC. This

[5]The magic numbers are the ones that close the Fermi momenta shell in a simple cubic box. For a three dimensional system are $1, 7, 19, 27, 33, 57, 81, 93, 123, 147, 171, 179, 203, 251 \ldots$.

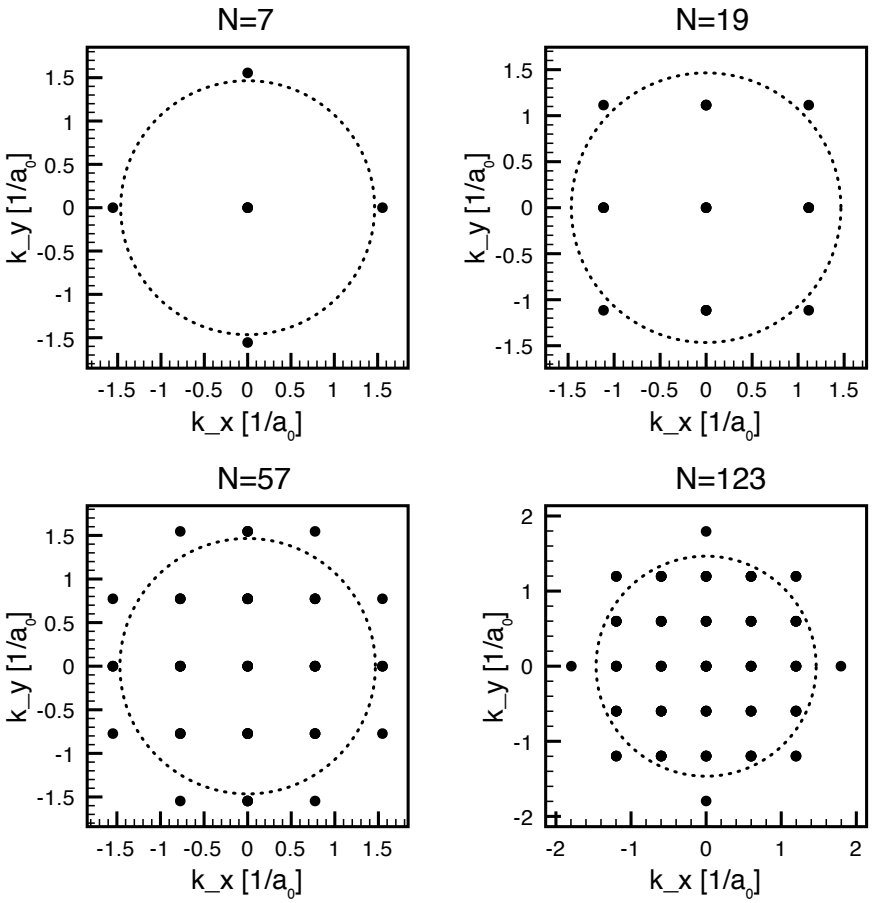


Figure 3.8: Two-dimensional cross section of the momenta $\mathbf{k}$ considered in a $N \times N$ $\mathrm{SD}_{\mathrm{pw}}$ matrix, for several $N$. The dotted circles delineate the Fermi sphere.

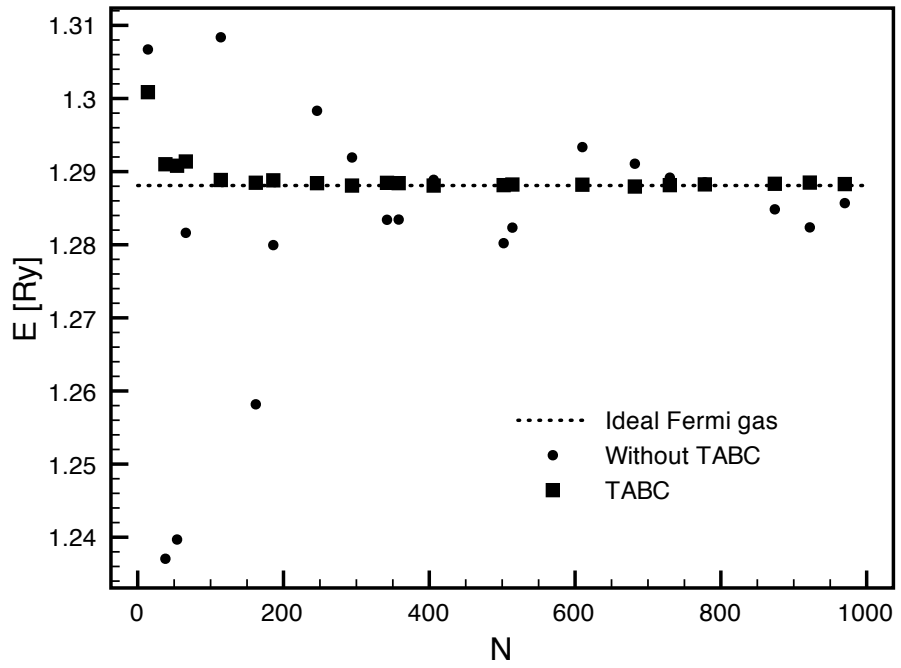


Figure 3.9: Electronic kinetic energy provided by a $\mathrm{SD_{pw}}$, in the case of hydrogen at $r_s = 1.31$. We have considered only magic numbers (see footnote 5).

can be accomplished within Quantum-Espresso by selecting a grid, for example $5 \times 5 \times 5$, and letting the algorithm find the optimal $K$-points to average over, according to the Monkhorst-Pack rule [91]. Each $K$-point has an associated weight reported in the output file of Quantum-Espresso. Instead of summing over all the weighted configurations, we used the scheme of TABC to sample from them, with a probability proportional to their weight. In other words, we averaged over the $K$-points taking advantage of importance sampling.

The implementation of the TABC for $\mathrm{SD_{DFT}}$ can be summarized by the following instructions:

1. Use Quantum-Espresso with a certain energy cutoff $E_{\mathrm{ctf}}$ and $K$-grid, obtaining as output $n_{\mathrm{DFT}}$ solutions which use different $K$-points $K_i$ and associated weights $w_i$;
2. Sample a $K$-point $K_j$ with probability
$$P_j = \frac{w_j}{\sum_{l=1}^{n_{\mathrm{DFT}}} w_l}$$
and employ its associated solutions in the $\mathrm{SD_{DFT}}$;
3. Perform $M_{\mathrm{relax}}$ relaxation steps;
4. Sample $M/N_{\mathrm{twist}}$ points and accumulate the estimators;
5. Repeat the points 2-4 $N_{\mathrm{twist}}$ times.

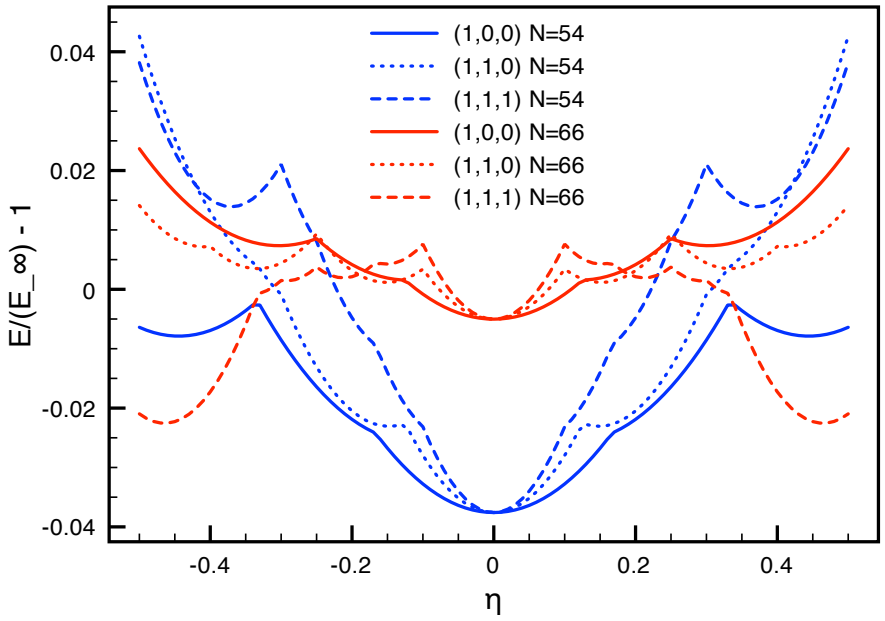


Figure 3.10: Kinetic energy change due to the twists $\mathbf{v}_{\text{twist}} = \frac{2\pi}{L}\eta(1,0,0)$, $\mathbf{v}_{\text{twist}} = \frac{2\pi}{L}\eta(1,1,0)$ and $\mathbf{v}_{\text{twist}} = \frac{2\pi}{L}\eta(1,1,1)$, for 54 and 66 particles.

In Figure 3.11 we report the results obtainable employing the TABC for a JS-DFT trial wave function. In all cases $n_K = 5$ is sufficient for accounting the finite size effects, and we have therefore employed this number also for $N$ larger than 16. We remark that a larger $n_K$ requires more space on the hard disk, but it does not demand longer computational times nor larger RAM within VMC.

Finally, Figure 3.12 demonstrates the effectiveness of the TABC. The TABC provide a quicker convergence to the thermodynamic limit especially in the case of atm-bcc hydrogen, which is subject to strong finite size effects.

### 3.1.8 Conjugate Gradient

The computational possibilities offered by modern computers allow for many methods which can efficiently determine the minimum of a given function $f$ (Chapter 10 [92], [93]). In the following, we present the *Conjugate Gradient* (CG) algorithm, one of the most used ones.

The CG requires the possibility of computing the function $f$ and its gradient $\nabla f$. Given these ingredients, the most obvious procedure is to search a minimum in the direction $-\nabla f$, then re-evaluate the gradient in this new point, search a minimum along the new direction, and so on. This method is called *Steepest Descent* and it is known to be very inefficient, especially near the minimum. In fact, as illustrated in Figure 3.13, even the minimization of a perfect quadratic form leads to a very large number of successive local minima, making this procedure extremely inefficient.

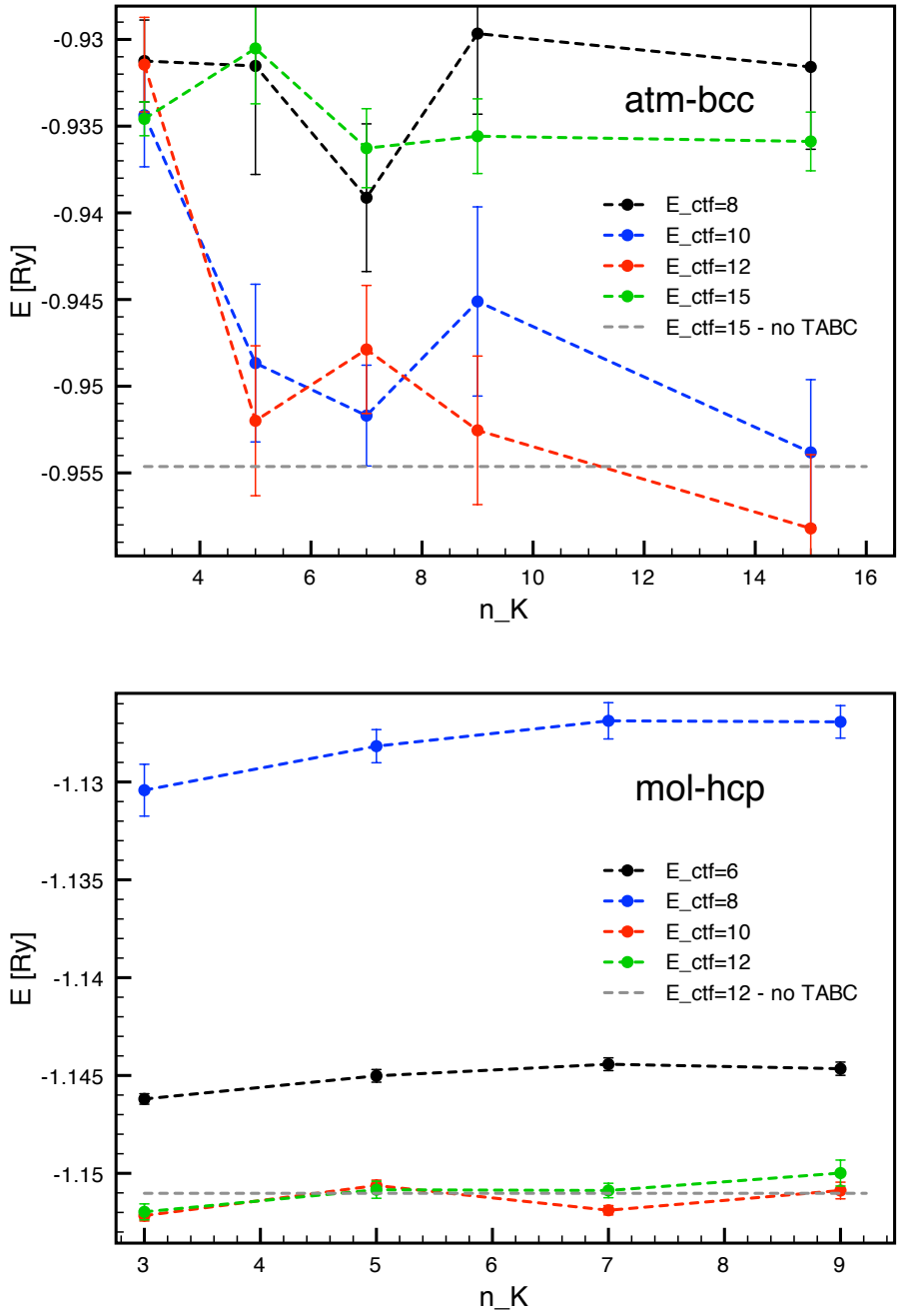


Figure 3.11: Variational energies of 16 atoms of hydrogen in the atm-bcc configuration at $r_s = 1.31$ and mol-hcp at $r_s = 2.61$, for various $E_{\text{ctf}}$ and $n_K \times n_K \times n_K$ grids.

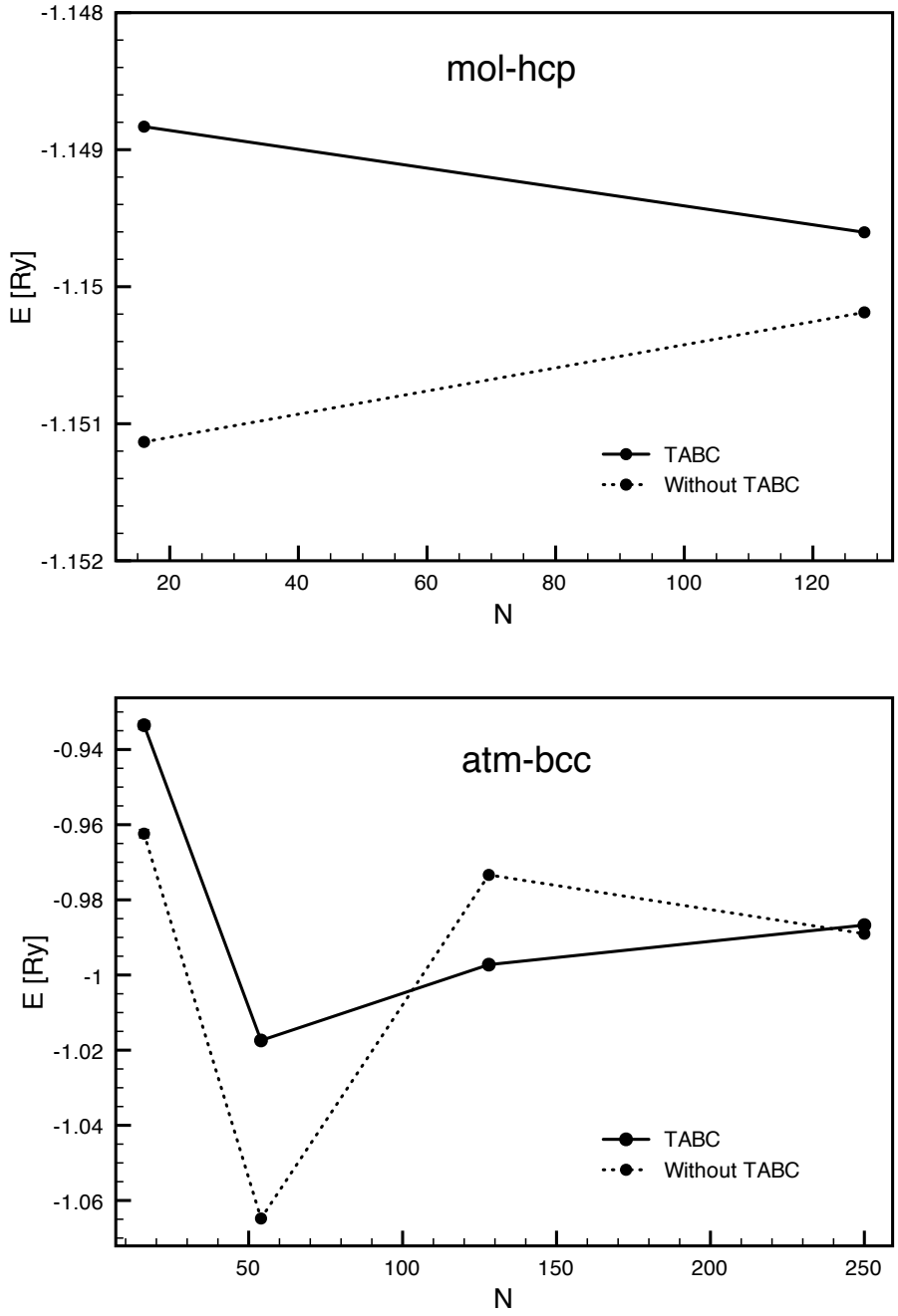


Figure 3.12: Variational energy versus number of simulated atoms $N$, for the atm-bcc phase at $r_s = 1.31$ and the mol-hcp phase at $r_s = 2.61$. All the results refer to the JS-DFT trial wave function.

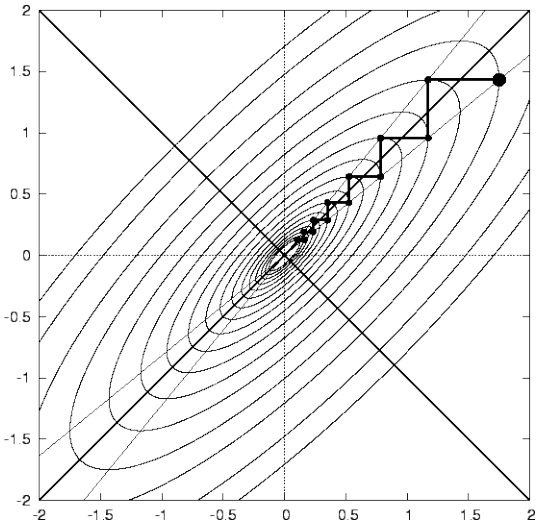


Figure 3.13: Typical minimization path of the Steepest Descent method [94].

To overcome the aforementioned problem, the direction can be adjusted by mixing it with the previous one in order to proceed straightly to the minimum. The resulting Conjugate Gradient method, can be summarized in the following instructions (see [92] and [95] for their derivation):

1. Select a starting point $x_0$;
2. Evaluate the gradient $g_0 = -\nabla f(x_0)$;
3. Find the local minimum $x_1$ in the direction $g_0$;
4. $h_0 = g_0$ and $i = 1$;
5. Evaluate the new gradient $g_i = -\nabla f(x_i)$;
6. Determine the conjugate vector $h_i = g_i + \frac{g_i \cdot g_i}{g_{i-1} \cdot g_{i-1}} h_{i-1}$;
7. Find the local minimum $x_{i+1}$ in the direction $h_i$;
8. $i := i + 1$;
9. Repeat the instructions 4-8 until convergence has been reached.

This method is very general and can be used in conjunction with VMC. However, the evaluation of the energy gradient cannot be computed by a straightforward finite difference, because the associated variance would be very large. For

this reason, a correlated sampling technique must be devised:

$$\begin{aligned}\nabla_\alpha E(\alpha) &= \frac{1}{\delta\alpha}\left(\frac{\int dR\,\Psi^*_{\alpha+\delta\alpha}(R)\mathrm{H}\Psi_{\alpha+\delta\alpha}(R)}{\int dR\,\Psi^*_{\alpha+\delta\alpha}(R)\Psi_{\alpha+\delta\alpha}(R)} - \frac{\int dR\,\Psi^*_{\alpha}(R)\mathrm{H}\Psi_{\alpha}(R)}{\int dR\,\Psi^*_{\alpha}(R)\Psi_{\alpha}(R)}\right) \\ &= \frac{1}{\delta\alpha}\left(\frac{\int dR\,|\Psi_\alpha(R)|^2 w(\alpha+\delta\alpha;\alpha)\mathrm{H}^{\mathrm{loc}}(\alpha+\delta\alpha)}{\int dR\,|\Psi_\alpha(R)|^2 w(\alpha+\delta\alpha;\alpha)} - \right. \\ &\qquad \left. \frac{\int dR\,|\Psi_\alpha(R)|^2 w(\alpha;\alpha)\mathrm{H}^{\mathrm{loc}}(\alpha)}{\int dR\,|\Psi_\alpha(R)|^2 w(\alpha;\alpha)}\right). \quad (3.61)\end{aligned}$$

Here $\alpha$ represents the variational parameters, while $w(\beta;\alpha) = \frac{|\Psi_\beta(R)|^2}{|\Psi_\alpha(R)|^2}$ is a weight function. In other words, the gradient $\nabla_\alpha E$ can be conveniently estimated by selecting a small enough $\delta\alpha$, sample from $|\Psi_\alpha(R)|^2$, and accumulate the estimators $\mathrm{H}^{\mathrm{loc}}(\alpha)$, $\mathrm{H}^{\mathrm{loc}}(\alpha+\delta\alpha)$ and $w(\alpha+\delta\alpha;\alpha)$. Finally,

$$\nabla_\alpha E(\alpha) \simeq \frac{1}{\delta\alpha}\left(\frac{\sum_{i=1}^M w_i(\alpha+\delta\alpha;\alpha)\mathrm{H}_i^{\mathrm{loc}}(\alpha+\delta\alpha)}{\sum_{i=1}^M w_i(\alpha+\delta\alpha;\alpha)} - \sum_{i=1}^M \mathrm{H}_i^{\mathrm{loc}}(\alpha)\right). \quad (3.62)$$

Besides the computation of the gradient, the CG method requires a one-dimensional algorithm for determining the minimum along a given direction. Our implemented method is a combination of the *Parabolic Interpolation* (PI) and the *Golden Section Search* (GSS) techniques (Chapter 10 [92]). Both PI and GSS determine the minimum of a target function $f$ within a certain domain $[a,b]$. The Parabolic Interpolation uses the parabola that passes through $f(a)$, $f(b)$, and a middle point $f(m)$, to select the next point, corresponding to the minimum of the parabola $p$. For the next iteration, it employs the lowest of the four points, that will be named $m$, and its next neighbors, which will be labelled as $a$ and $b$. The procedure continues until the bracket $[a,b]$ is reduced to an appropriately small interval. The GSS is similar to the PI, but instead of using a parabolic function to select the next point, it considers the larger interval $l$ between $[a,m]$ and $[m,b]$, and takes the next point at a distance $\frac{3-\sqrt{5}}{2}l$ (golden ratio) from $m$.

PI is expected to be more efficient than GSS, and in most cases it is. Nevertheless, for certain functions it turns out to be less efficient. For example, the minimization of the function

$$f(x) = \sum_{i=1}^{4}(x+10)^i \quad (3.63)$$

with the initial bracket ($a=-12., m=-9., b=-7.$), requires 20 iterations with GSS and 83 with PI. In the following we introduce our combined method, which is able to determine the minimum within only 15 iterations. PI happens to be inefficient when the parabolic minimum is always on the same side of $m$, so that the bracket does not shrink symmetrically. Therefore, when PI proposes the new point on the same side three times in a row, it is convenient to switch to select the next point according to the GSS criteria. We call this hybrid method *Para-Gold Search* (PGS). We have tested PGS with several different functions and it always performed as well as the best between PI and GSS, and sometimes even better.

By means of these two ingredients, the $\nabla_\alpha E$ and the PGS search algorithm, it is possible to successfully optimize most of the trial wave functions. Unfortunately, the SWF is not included in such a category, because of the large fluctuations that affect the gradient estimator. To demonstrate it, we first compute the gradient for 16 hydrogen atoms in the mol-hcp phase at $r_s = 2.61$ with JS-pw, obtaining[6]:

$$\begin{array}{lcrcl} A_{ee}^{\uparrow\uparrow}, A_{ee}^{\uparrow\downarrow} & - & -0.02073 & \pm & 0.00083 \\ F_{ee}^{\uparrow\uparrow} & - & -0.00763 & \pm & 0.00054 \\ F_{ee}^{\uparrow\downarrow} & - & -0.02088 & \pm & 0.00089 \\ A_{ep}^{\uparrow\uparrow}, A_{ep}^{\uparrow\downarrow} & - & 0.02107 & \pm & 0.00042 \\ F_{ep}^{\uparrow\uparrow}, F_{ep}^{\uparrow\downarrow} & - & -0.03984 & \pm & 0.00077\,. \end{array}$$

When we employ the ASWF-DFT trial wave function instead, the resulting gradient estimator is

$$\begin{array}{lcrcl} A_{ee}^{\uparrow\uparrow}, A_{ee}^{\uparrow\downarrow} & - & 0.001 & \pm & 0.015 \\ F_{ee}^{\uparrow\uparrow} & - & 0.012 & \pm & 0.015 \\ F_{ee}^{\uparrow\downarrow} & - & -0.002 & \pm & 0.016 \\ A_{ep}^{\uparrow\uparrow}, A_{ep}^{\uparrow\downarrow} & - & 0.038 & \pm & 0.015 \\ F_{ep}^{\uparrow\uparrow}, F_{ep}^{\uparrow\downarrow} & - & -0.022 & \pm & 0.015 \\ C & - & -0.069 & \pm & 0.019 \\ A_{ss}^{\uparrow\uparrow}, A_{ss}^{\uparrow\downarrow} & - & 0.014 & \pm & 0.016 \\ F_{ss}^{\uparrow\uparrow}, F_{ss}^{\uparrow\downarrow} & - & 0.014 & \pm & 0.016 \\ A_{sp}^{\uparrow\uparrow}, A_{sp}^{\uparrow\downarrow} & - & 0.016 & \pm & 0.016 \\ F_{sp}^{\uparrow\uparrow}, F_{sp}^{\uparrow\downarrow} & - & 0.014 & \pm & 0.015\,. \end{array}$$

Essentially, $\nabla_\alpha E$ is affected by such a large variance that it is meaningless.

Even by increasing $M$ by a factor of 100, which requires a considerable amount of CPU-time, the gradient is not accurate enough for some shadow variational parameters:

$$\begin{array}{lcrcl} A_{ee}^{\uparrow\uparrow}, A_{ee}^{\uparrow\downarrow} & - & -0.01386 & \pm & 0.00085 \\ F_{ee}^{\uparrow\uparrow} & - & -0.00302 & \pm & 0.00084 \\ F_{ee}^{\uparrow\downarrow} & - & -0.01717 & \pm & 0.00087 \\ A_{ep}^{\uparrow\uparrow}, A_{ep}^{\uparrow\downarrow} & - & 0.02096 & \pm & 0.00085 \\ F_{ep}^{\uparrow\uparrow}, F_{ep}^{\uparrow\downarrow} & - & -0.03870 & \pm & 0.00088 \\ C & - & -0.0822 & \pm & 0.00280 \\ A_{ss}^{\uparrow\uparrow}, A_{ss}^{\uparrow\downarrow} & - & -0.00436 & \pm & 0.00090 \\ F_{ss}^{\uparrow\uparrow}, F_{ss}^{\uparrow\downarrow} & - & -0.00503 & \pm & 0.00091 \\ A_{sp}^{\uparrow\uparrow}, A_{sp}^{\uparrow\downarrow} & - & 0.00092 & \pm & 0.00084 \\ F_{sp}^{\uparrow\uparrow}, F_{sp}^{\uparrow\downarrow} & - & -0.00217 & \pm & 0.00088\,. \end{array}$$

This problem is intrinsic of the SWF, and is a known issue that until now has been circumvented by manual tuning of the variational parameters or automatized linear searches along all the variational parameters directions.

[6]For sake of readability, we have set the constraints $A_{ee}^{\uparrow\uparrow} \equiv A_{ee}^{\uparrow\downarrow}$, $A_{ep}^{\uparrow\uparrow} \equiv A_{ep}^{\uparrow\downarrow}$, $F_{ep}^{\uparrow\uparrow} \equiv F_{ep}^{\uparrow\downarrow}$, $A_{ss}^{\uparrow\uparrow} \equiv A_{ss}^{\uparrow\downarrow}$, $F_{ss}^{\uparrow\uparrow} \equiv F_{ss}^{\uparrow\downarrow}$, $A_{sp}^{\uparrow\uparrow} \equiv A_{sp}^{\uparrow\downarrow}$, and $F_{sp}^{\uparrow\uparrow} \equiv F_{sp}^{\uparrow\downarrow}$.

### 3.1.9 Stochastic Reconfiguration

The *Stochastic Reconfiguration* method is a very efficient algorithm developed specifically for the VMC energy minimization [81–83]. The leading idea is to change the variational parameters according to the ground state projector.

First of all, we demonstrate that given a trial wave function $\Psi_{\mathrm{T}(\alpha_1 \dots \alpha_n)}$, it is possible to project it to the ground state by applying many consecutive times the operator $(\Lambda - \mathrm{H})$, where $\Lambda \gg 1$. In order to prove the latter statement, we decompose $\Psi_{\mathrm{T}}$ into a superposition of eigenstates $\{\phi_i\}_i$:

$$\Psi_{\mathrm{T}} = \sum_i c_i \phi_i \,. \tag{3.64}$$

Then, by applying the aforementioned projector $m$ times, we obtain

$$(\Lambda - \mathrm{H})^m \, \Psi_{\mathrm{T}} = \sum_i c_i \, (\Lambda - E_i)^m \, \phi_i \,. \tag{3.65}$$

If $\Lambda \gg 1$, the term $(\Lambda - E_0)^m$ dominates the summation. Therefore, the ground state spontaneously emerges by successive applications of the projector $(\Lambda - \mathrm{H})$.

Secondly, we note that small changes of the variational parameters

$$(\alpha_1' \dots \alpha_n') = (\alpha_1 \dots \alpha_n) + (\delta\alpha_1 \dots \delta\alpha_n) \;, \tag{3.66}$$

modify the trial wave function according to

$$\Psi_{\mathrm{T}\left(\alpha_1' \dots \alpha_n'\right)} = \Psi_{\mathrm{T}(\alpha_1 \dots \alpha_n)} + \sum_{l=1}^{n} \delta\alpha_l \frac{\partial \Psi_{\mathrm{T}(\alpha_1 \dots \alpha_n)}}{\partial \alpha_l} \tag{3.67}$$

$$= \Psi_{\mathrm{T}(\alpha_1 \dots \alpha_n)} + \sum_{l=1}^{n} \delta\alpha_l O_l \Psi_{\mathrm{T}(\alpha_1 \dots \alpha_n)} \;, \tag{3.68}$$

where

$$O_l = \frac{\partial}{\partial \alpha_l} \ln(\Psi_{\mathrm{T}(\alpha_1 \dots \alpha_n)}) \,. \tag{3.69}$$

In the following we adopt the notation $\Psi_{\mathrm{T}} \equiv \Psi_{\mathrm{T}(\alpha_1 \dots \alpha_n)}$ and ${\Psi_{\mathrm{T}}}' \equiv \Psi_{\mathrm{T}\left(\alpha_1' \dots \alpha_n'\right)}$. We also introduce the wave function $\Psi_{\mathrm{P}} \equiv (\Lambda - \mathrm{H})\, \Psi_{\mathrm{T}}$.

With the intention of determining the $\delta\alpha_i$ that lead to the variational ground state, we require that

$$\frac{\langle \Psi_{\mathrm{T}} \mid O_k \mid {\Psi_{\mathrm{T}}}' \rangle}{\langle \Psi_{\mathrm{T}} \mid {\Psi_{\mathrm{T}}}' \rangle} = \frac{\langle \Psi_{\mathrm{T}} \mid O_k \mid \Psi_{\mathrm{P}} \rangle}{\langle \Psi_{\mathrm{T}} \mid \Psi_{\mathrm{P}} \rangle} \,, \qquad \forall k \in [1, n]. \tag{3.70}$$

From this condition follows

$$\Lambda \sum_{l=1}^{n} \delta\alpha_l \left( \langle O_k O_l \rangle - \langle O_l \rangle \langle O_k \rangle \right) = \langle \mathrm{H} \rangle \langle O_k \rangle - \langle O_k \mathrm{H} \rangle \,, \qquad \forall k \in [1, n], \tag{3.71}$$

where $\langle \cdot \rangle \equiv \langle \Psi_{\mathrm{T}} \mid \cdot \mid \Psi_{\mathrm{T}} \rangle$. Then, we can explicitly write the final expression for the optimal changes of the variational parameters:

$$\delta\alpha_l = \frac{1}{\Lambda} \sum_{k=1}^{n} f_k \left( s^{-1} \right)_{kl} \;, \tag{3.72}$$

where $s_{lk} \equiv \langle O_k O_l \rangle - \langle O_l \rangle \langle O_k \rangle$ is a symmetric matrix, and $f_k \equiv \langle \mathrm{H} \rangle \langle O_k \rangle - \langle O_k \mathrm{H} \rangle$ a vector. We remark that the computation of $\delta\alpha_l$ requires only the accumulation of the estimator $\mathrm{H}^{\mathrm{loc}}$ and $O_l$, where the latter ones can be typically evaluated with a very small computational cost. In other words, the estimation of $\delta\alpha_l$ requires substantially the same resources as the energy alone.

Even though equation (3.72) provides the direction that leads to the optimal variational parameters, its optimal magnitude $\lambda$ is not clear, because of the coefficient $\Lambda$. In fact, even though $\Lambda \gg 1$ would suggest very small $\delta\alpha_l$, in practice this would result computationally disadvantageous. In principle, a linear search algorithm could be used to find the minimum along this direction, but in practice this approach is not convenient because a new direction can be computed together with the energy with a negligible additional computational cost. Given these considerations, we introduce an iterative method that adjusts $\lambda$ according to the successive directions provided by equation (3.72). Specifically, our procedure is:

1. $i = 0$;
2. Sample from the selected trial wave function with starting variational parameters $\alpha^{(0)}$, and estimate the $n$-dimensional vector $\delta\alpha^{(0)}$ by means of equation (3.72). Then, normalize $\delta\alpha^{(0)}$:
$$\delta\alpha^{(0)} := \frac{\delta\alpha^{(0)}}{|\delta\alpha^{(0)}|};$$
3. $\lambda = \sqrt{\frac{|\alpha^{(0)}|^2}{n}}$;
4. $\alpha^{(1)} = \lambda\, \delta\alpha^{(0)}$;
5. $i := i + 1$;
6. Estimate $\delta\alpha^{(i)}$ by sampling with variational parameters $\alpha^{(i)}$, and normalize it:
$$\delta\alpha^{(i)} := \frac{\delta\alpha^{(i)}}{|\delta\alpha^{(i)}|};$$
7. If $i > 2$ then $\lambda := \lambda \left(1 + 0.1 \times \delta\alpha^{(i)} \cdot \delta\alpha^{(i-1)}\right)$
8. If $i > 3$ then $\lambda := \lambda \left(0.85 + 0.3 \times \frac{|\alpha^{(i)} - \alpha^{(i-2)}|}{\sum_{j=1}^{2} |\alpha^{(i+1-j)} - \alpha^{(i-j)}|}\right)$
9. If $i > 5$ then $\lambda := \lambda \left(0.75 + 0.5 \times \frac{|\alpha^{(i)} - \alpha^{(i-4)}|}{\sum_{j=1}^{4} |\alpha^{(i+1-j)} - \alpha^{(i-j)}|}\right)$
10. $\alpha^{(i+1)} = \lambda\, \delta\alpha^{(i)}$;
11. Repeat from point 5 until convergence.

By means of the list of instructions provided above, we have been able to successfully optimize the JS-pw and JS-DFT trial wave functions, as shown in figures 3.14, 3.15, 3.16, and 3.17. However, large oscillations in the energy arise for low densities, in particular when the JS-pw is employed. These fluctuations

are due to the electron-proton variational parameters, as demonstrated in figure 3.18. However, this effect is considerably less pronounced with the JS-DFT trial wave function.

We have compared the effectiveness of SR to CG, by optimizing the variational parameters of the JS-DFT trial wave function for 16 hydrogen atoms in the mol-hcp phase at $r_s = 2.61$. Initially all the parameters were set equal to 1. In 234 minutes CG reached the minimum $E = -1.1200 \pm 0.0003$ Ry, whereas SR provided $E = -1.150 \pm 0.001$ Ry in 77 minutes. This demonstrates that SR is more efficient and reliable than CG.

As we have just pointed out, Stochastic Reconfiguration is very effective in the case of the standard trial wave functions JS-pw and JS-DFT. Even more notably, SR proves to be very reliable also to ptimize the Shadow Wave Function, even though additional care is required. In the following, we analyze in detail the SWF optimization attainable by means of SR. In particular, we will consider three different strategies:

**SR1:** optimize the variational parameters all together;

**SR2:** first optimize the shadow variational parameters ($A_{ss}^{\uparrow\uparrow}$, $A_{ss}^{\uparrow\downarrow}$, $F_{ss}^{\uparrow\uparrow}$, $F_{ss}^{\uparrow\downarrow}$, $A_{sp}^{\uparrow\uparrow}$, $A_{sp}^{\uparrow\downarrow}$, $F_{sp}^{\uparrow\uparrow}$, and $F_{sp}^{\uparrow\downarrow}$) and subsequently all the variational parameters together;

**SR3:** optimize only the shadow variational parameters ($A_{ss}^{\uparrow\uparrow}$, $A_{ss}^{\uparrow\downarrow}$, $F_{ss}^{\uparrow\uparrow}$, $F_{ss}^{\uparrow\downarrow}$, $A_{sp}^{\uparrow\uparrow}$, $A_{sp}^{\uparrow\downarrow}$, $F_{sp}^{\uparrow\uparrow}$, and $F_{sp}^{\uparrow\downarrow}$), then only the particle parameters ($A_{ee}^{\uparrow\uparrow}$, $A_{ee}^{\uparrow\downarrow}$, $F_{ee}^{\uparrow\uparrow}$, $F_{ee}^{\uparrow\downarrow}$, $A_{ep}^{\uparrow\uparrow}$, $A_{ep}^{\uparrow\downarrow}$, $F_{ep}^{\uparrow\uparrow}$, and $F_{ep}^{\uparrow\downarrow}$), and finally all together.

We start our analysis with the ASWF, which is easier to treat than the FSWF, since it does not suffer from the sign problem. Figure 3.19 illustrates the results of our tests. Both SR2 and SR3 procedures prove to be very effective, even though SR2 performs slightly better. However, in the case of the atm-bcc phase at high density, the variance of the estimated energy considerably increases at the end of the minimization. This fact is surprising, since normally a lower variational energy is connected to a lower variance. By direct inspection, we have noticed that this behavior is related to a decrease of the variational parameter $C$, which in turn implies a broad function to integrate, and therefore a large variance. Thus, the growing of the variance is inevitable, and indeed proves the ability of SR to determine the proper $\delta\alpha_l$ even in presence of large fluctuations.

We now consider the variational optimization of the FSWF. The sign problem impeded the analysis for the atm-bcc phase at $r_s = 1.31$, thus we report the SR minimization only for mol-hcp hydrogen (Figure 3.20). Differently from the ASWF, SR3 seems to be the most effective procedure for the FSWF. Interestingly, the variational energy obtainable with SR3 can be achieved simply by employing the variational parameters of the optimized ASWF. Since the sign problem prevents, or makes computationally very demanding, the optimization process, we have adopted this strategy. In section 3.2, all our results for the FSWF were not obtained by means of a minimization, but merely adopting the ASWF variational parameters.

However, it should be mentioned that the optimization of the FSWF is subject to a fundamental problem, not imputable to the minimization algorithm.

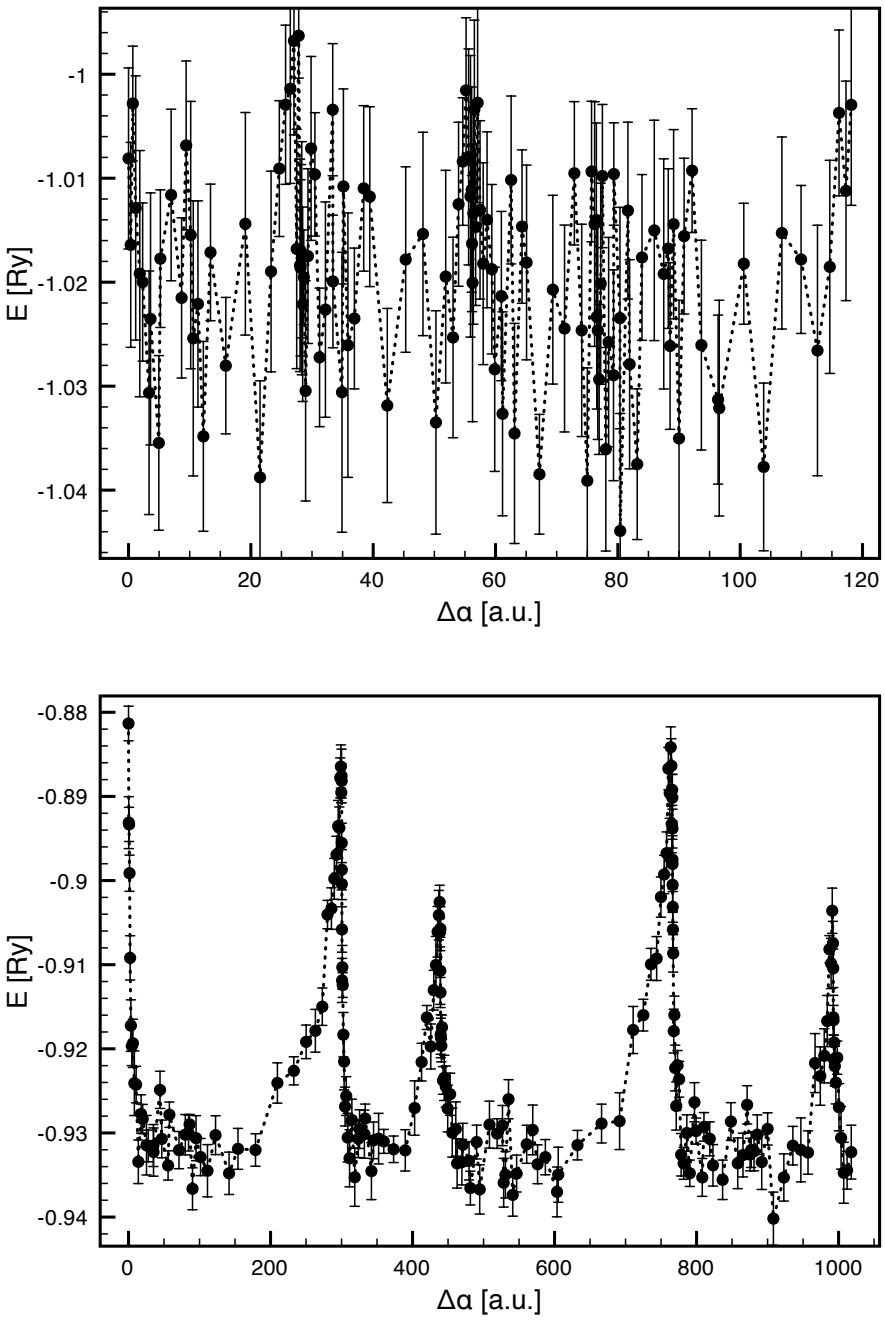


Figure 3.14: SR optimization of the JS-pw trial wave function for 16 hydrogen atoms in the atm-bcc phase at $r_s = 1.31$ (upper graph) and $r_s = 2.61$ (lower graph). Here, the spin-dependent split of the variational parameters has been turned off (i.e. $A_{ee}^{\uparrow\uparrow} \equiv A_{ee}^{\uparrow\downarrow}$, $F_{ee}^{\uparrow\uparrow} \equiv F_{ee}^{\uparrow\downarrow}$, $A_{ep}^{\uparrow\uparrow} \equiv A_{ep}^{\uparrow\downarrow}$, and $F_{ep}^{\uparrow\uparrow} \equiv F_{ep}^{\uparrow\downarrow}$) for simplicity, since they are plotted in figure 3.18.

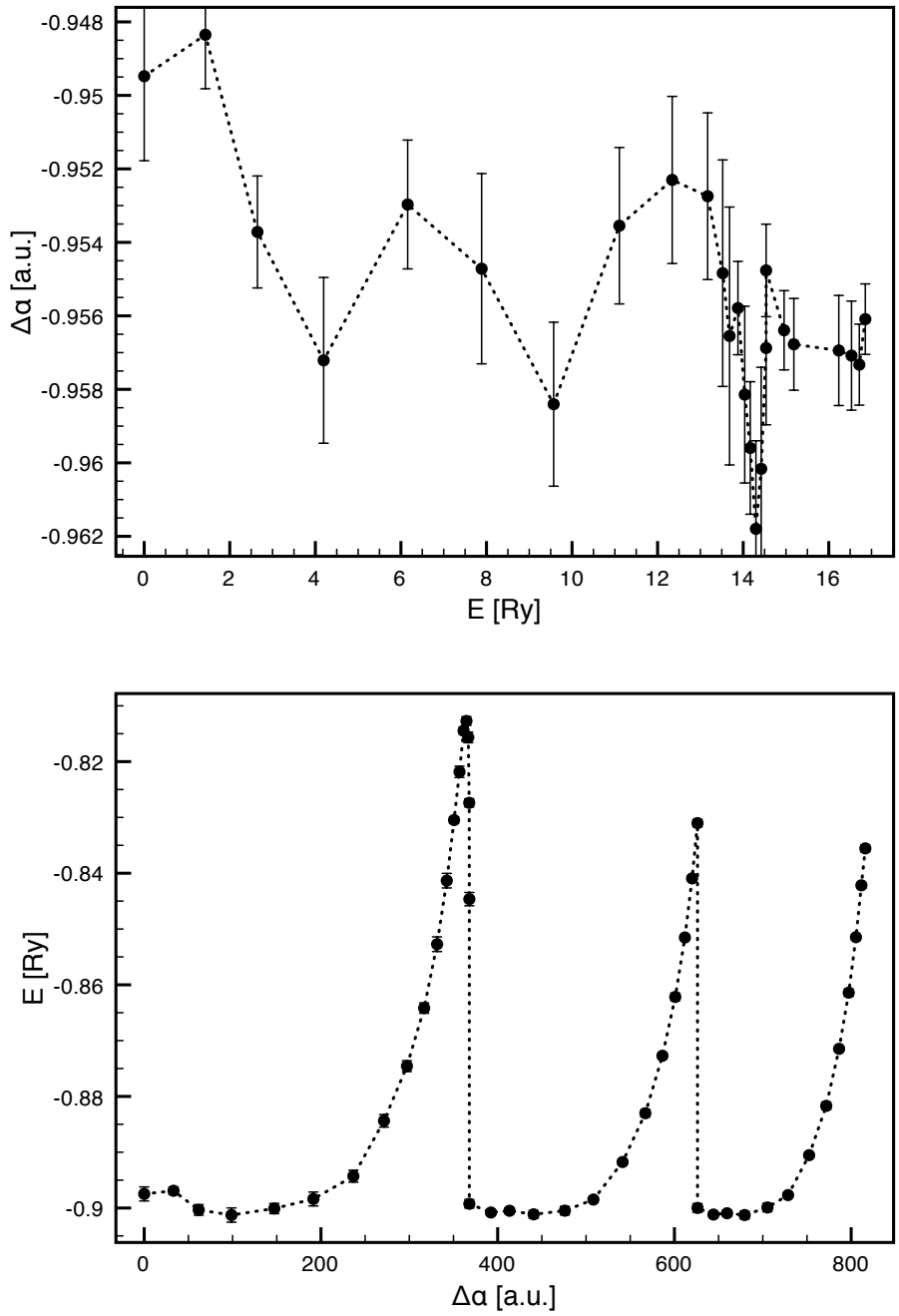


Figure 3.15: SR optimization of the JS-pw trial wave function for 16 hydrogen atoms in the mol-hcp phase at $r_s = 1.31$ (upper graph) and $r_s = 2.61$ (lower graph).

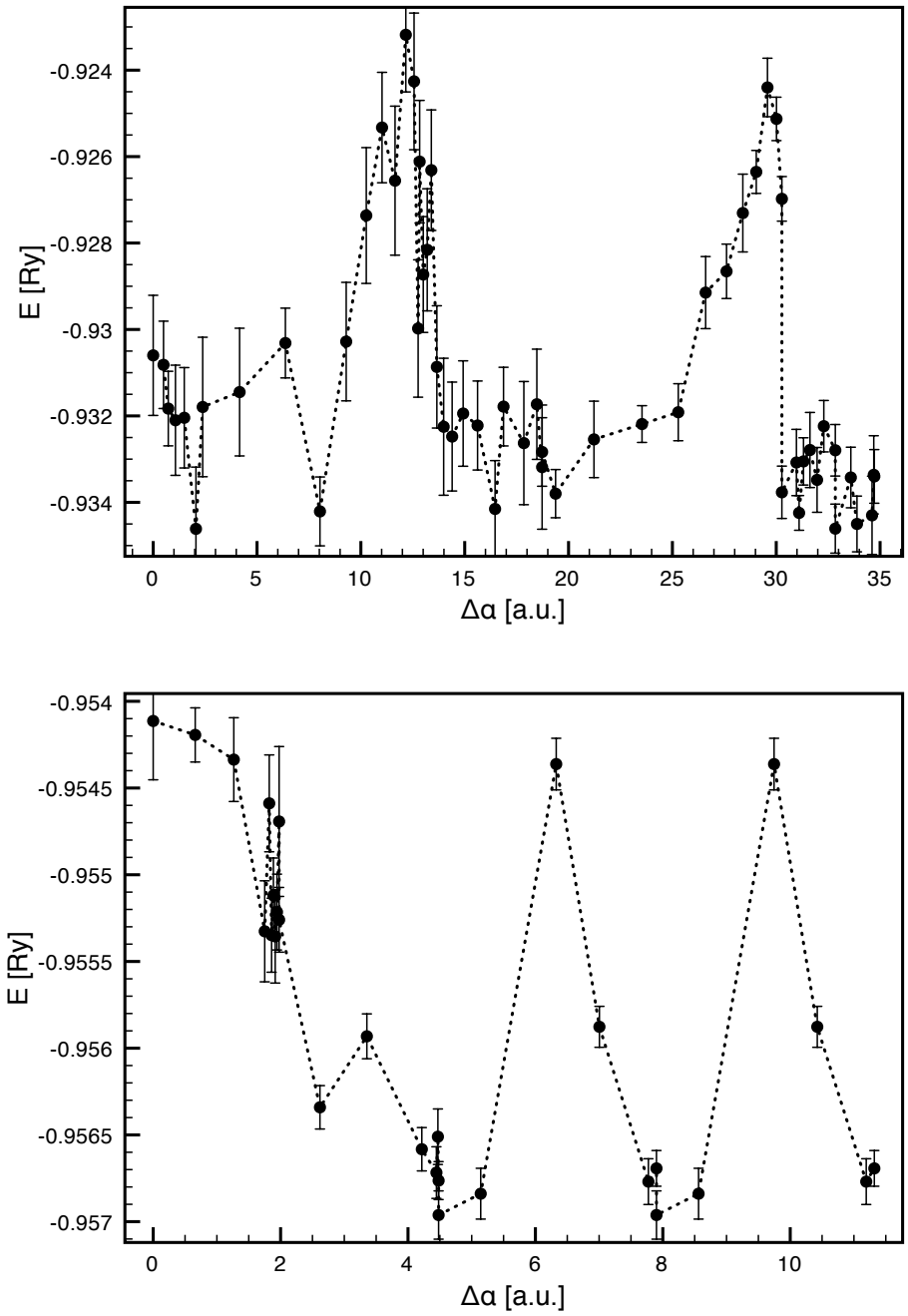


Figure 3.16: SR optimization of the JS-DFT trial wave function for 16 hydrogen atoms in the atm-bcc phase at $r_s = 1.31$ (upper graph) and $r_s = 2.61$ (lower graph).

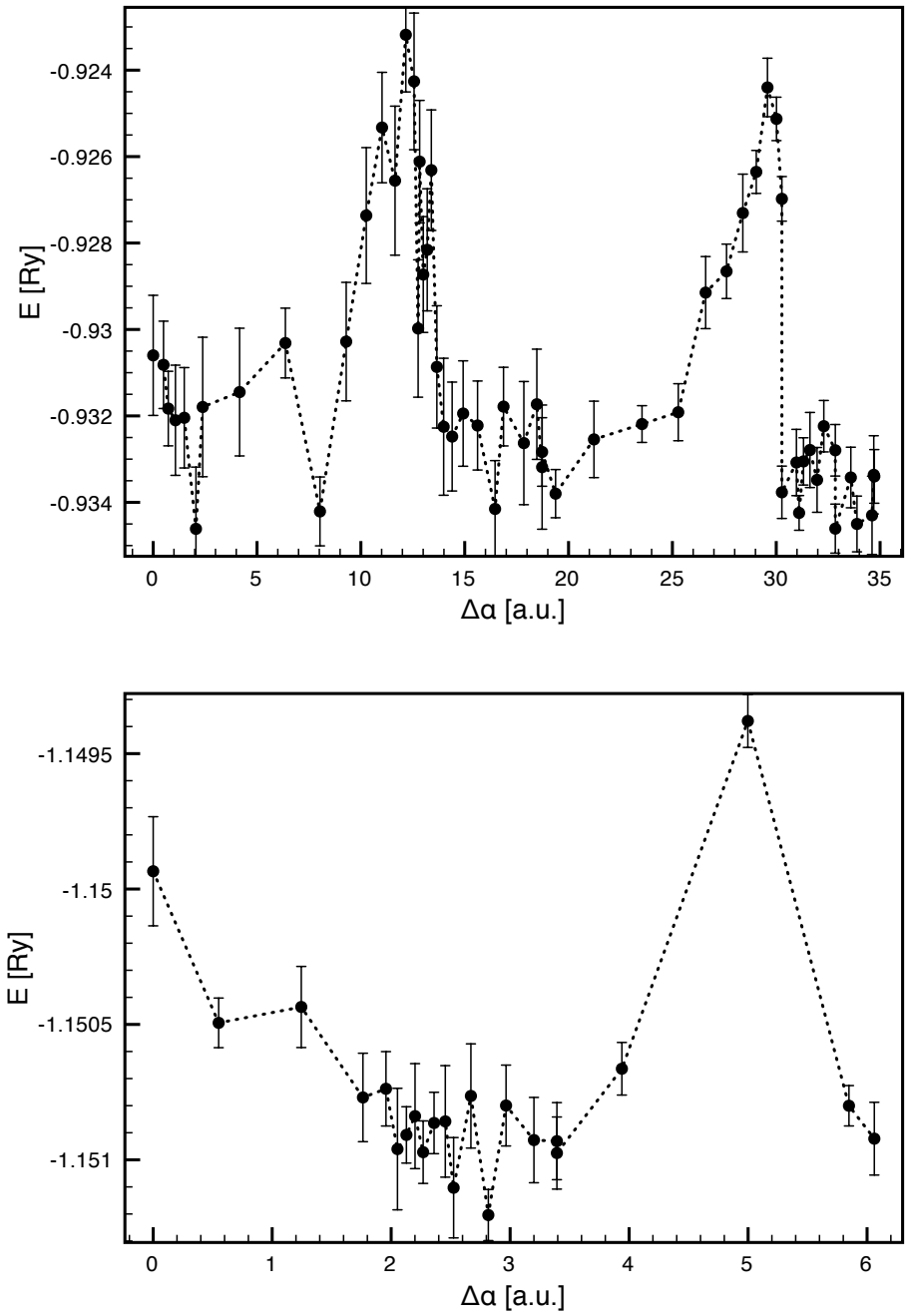


Figure 3.17: SR optimization of the JS-DFT trial wave function for 16 hydrogen atoms in the mol-hcp phase at $r_s = 1.31$ (upper graph) and $r_s = 2.61$ (lower graph).

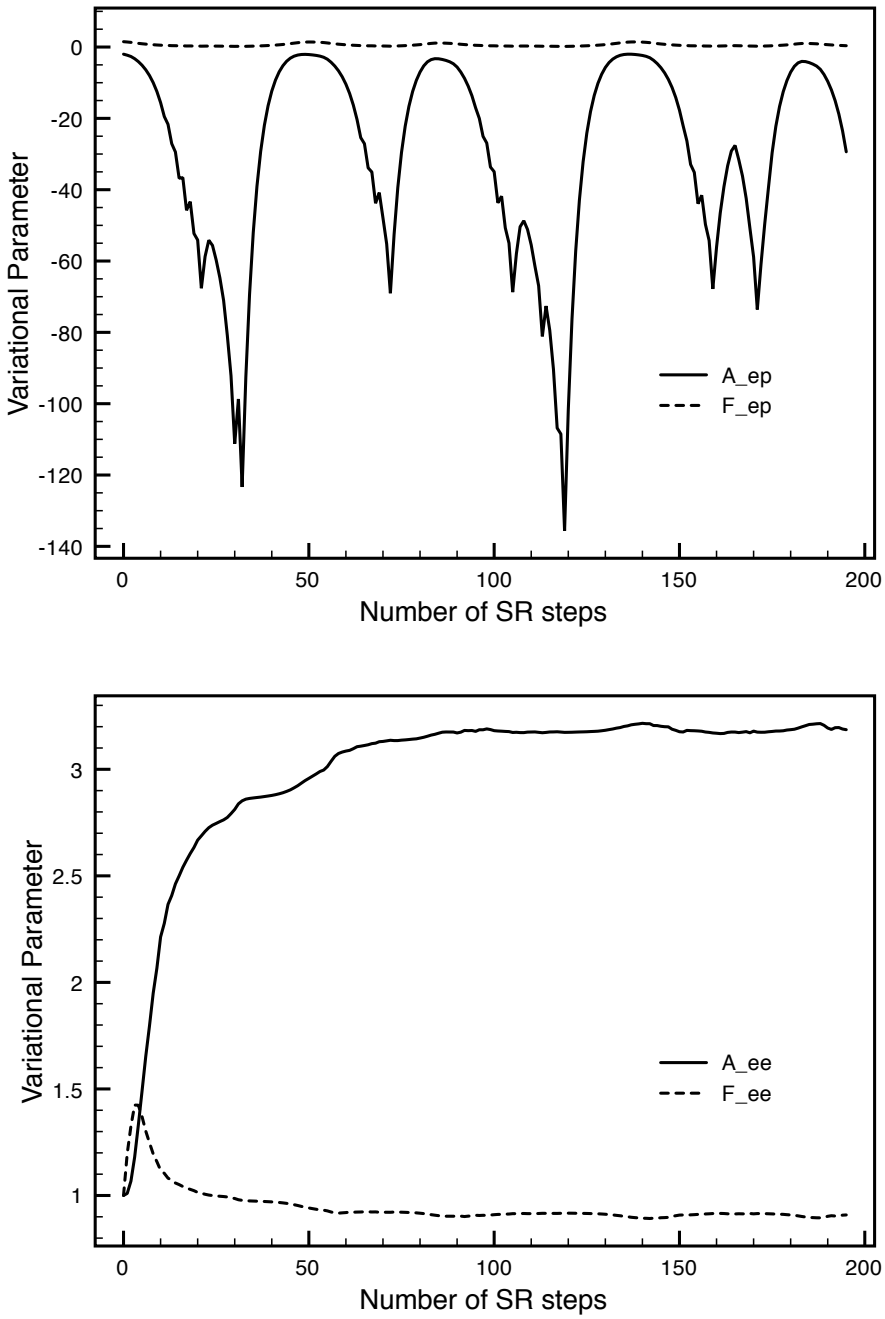


Figure 3.18: Values of the variational parameters during the SR minimization illustrated in the lower graphic in figure 3.14.

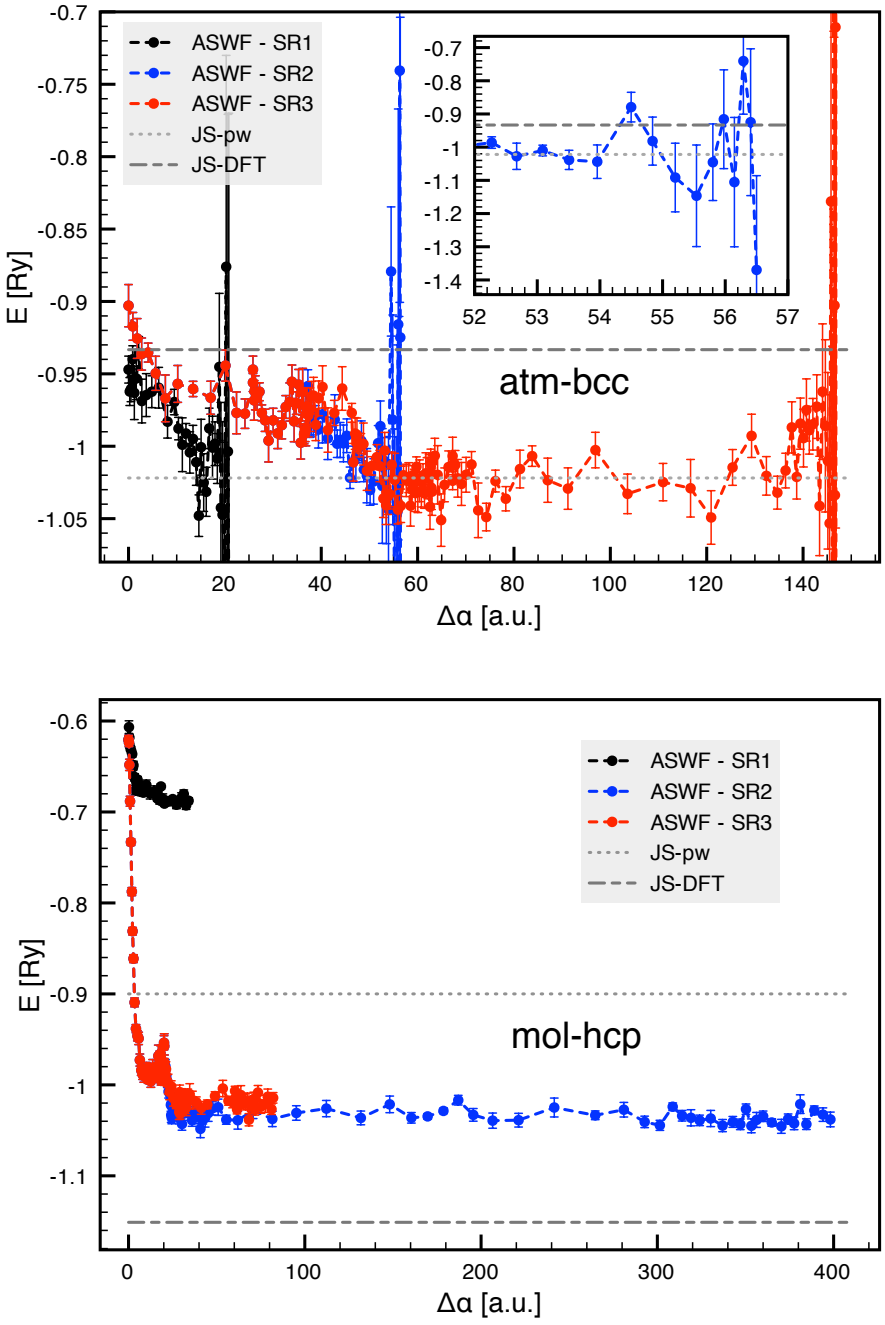


Figure 3.19: ASWF variational energy minimization by means of the three flavors of SR described in the text. The upper graph refers to atm-bcc hydrogen at $r_s = 1.31$ and the lower graph to mol-hcp at $r_s = 2.61$, with $N = 16$. The starting variational parameters were set equal to 1. Two grey lines are used to indicate the variational energy obtainable with JS-pw and JS-DFT. The inset in the upper graph illustrate in detail the end of the optimization process with SR2.

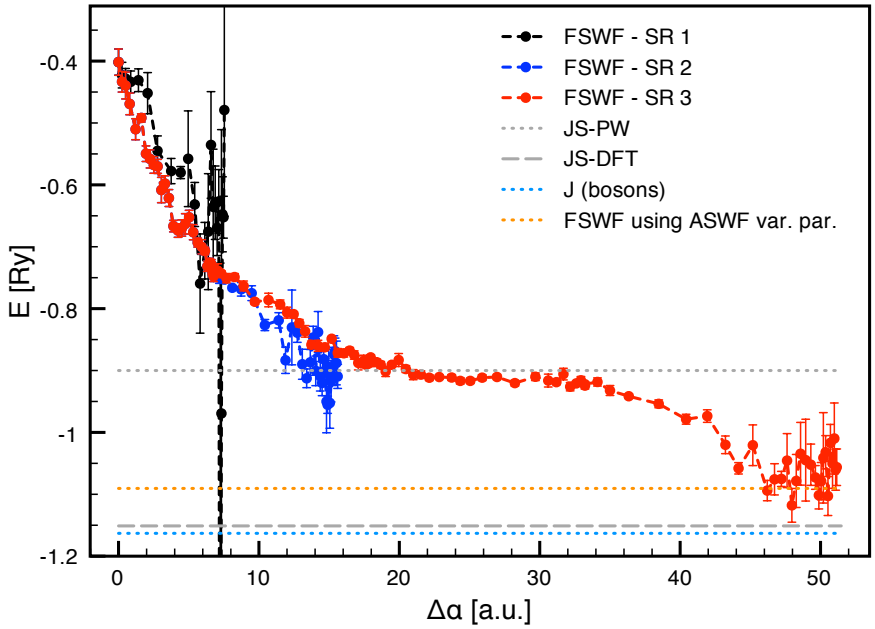


Figure 3.20: FSWF variational energy minimization by means of the three flavors of SR described in the text. The data refers to mol-hcp hydrogen at $r_s = 2.61$, with $N = 16$. The two grey lines represent the variational energy obtainable with JS-pw and JS-DFT. The orange line illustrates the variational energy attainable by employing the ASWF optimized variational parameters in the FSWF. The dotted light-blue line represent the bosonic energy resulting from a simple $J_{ee\text{-YUK}} J_{ep\text{-YUK}}$ trial wave function.

In fact, by setting $C = 0$, the FSWF is reduced to a bosonic wave function consisting of a Jastrow, which has a lower energy than the one obtainable by the FSWF (see Figure 3.20). Therefore, the minimization procedure will favor this possibility, leading to the suppression of the shadow integral. The situation is further complicated by the substantial increase of the sign problem fluctuations associated with the decrease of $C$. This observation elucidates the appearance of large error bars near the end of the minimization process, and the superiority of SR3 over SR2. Moreover, it better justifies the adoption of the ASWF variational parameters for the FSWF, since a complete optimization is precluded.

| Trial wave function | Literature [96] | Our results |
|---|---|---|
| JS-pw | −0.9636(2) Ry | −0.9828 Ry |
| JS-DFT | −0.9924(4) Ry | −0.9961(2) Ry |

Table 3.1: Variational energies of atm-bcc hydrogen at $r_s = 1.31$, with $N = 128$.

## 3.2 Results and Discussion

All the energy values for $N = 128$ hydrogen atoms which are illustrated in this section are reported in Appendix A.

### 3.2.1 Preliminary Check

Before presenting the results, we demonstrate the reliability of our code by means of a comparison with results available in the literature.

The work which better suits for a fair comparison is the one from Holzmann et al. [96]. In fact, they simulated the atm-bcc structure employing both the JS-pw and the JS-DFT trial wave functions. Moreover, they included the TABC, and accounted for the long range decay of the employed Jastrow by means of a modified version of the Ewald summation (see subsection 3.1.5). The most relevant difference is the adopted functional form for the Jastrow, which they deduced from the RPA approximation, and does not include any variational parameter.

In Table 3.1 we report our variational energies and the ones by Holzmann et al. The agreement is excellent. Moreover, $J_{ee\text{-YUK}}$ proves to be very accurate, since it allows for lower variational energies than the ones obtained by means of the RPA Jastrow.

In conclusion, our code proves to be both accurate and reliable.

### 3.2.2 JS-pw and JS-DFT

The JS-DFT and JS-pw variational energies are illustrated in Figure 3.21. Their comparison allows for assessing the improvement provided by DFT orbitals. Moreovere, JS-DFT is considered to be very reliable for electronic structures QMC simulations, and therefore its variational energy can be used as reference value for all the other ansätze.

As expected, the JS-DFT trial wave function is superior to JS-pw in all cases, especially for the mol-hcp phase at low densities. The enhancement is not only quantitative but also qualitative. In fact, the positive slope of the JS-pw energy for mol-hcp in Figure 3.21 can be related to a negative pressure, which is obviously unphysical, and therefore implies the inexistence of such a structure. Since the mol-hcp is known to be the most stable phase in that range of densities, such a prediction is wrong.

Contrary to the mol-hcp, the atm-bcc phase can be described quite accurately by JS-pw. In fact, the two energy curves are rather similar, even though the difference is increasingly larger for low densities. However, both trial wave

functions agree on the possible existence of the atm-bcc phase for $r_s \lesssim 1.6$, in accordance with the Wigner and Huntington prediction. We remark that atm-bcc hydrogen is supposed to be an electrical conductor by the band theory. This explains why JS-pw is reasonably accurate for the atm-bcc structure, since JS-pw is suitable for delocalized phases.

In Figures 3.22, 3.23, 3.24, and 3.25 we report the electron-electron and electron-proton pair correlation functions corresponding to the already presented variational energies. All of them exhibit more structure when JS-DFT is employed, especially in the low density regime, confirming our previous observations.

The pair correlation functions allow for a direct inspection of the $H_2$ covalent bond. In fact, the peaks of $g_{ee}(r)$ around $r = 0.75$ Bohr (Figure 3.24) originate from it. While JS-DFT reproduces it very well, the same does not hold for JS-pw. This confirm the unsuitability of JS-pw in describing the $H_2$ molecule.

In order to further investigate the accurateness in describing the covalent bond, we have simulated an isolated $H_2$ molecule and estimated its bond energy. More specifically, we have set $r_s = 50$, and simulated a single $H_2$ molecule for various proton-proton distances. In order to avoid interactions with the replica originating from the periodic boundary conditions, we have switched off the Ewald summation and the Periodic Coordinates. Also the TABC were not necessary for JS-pw, and therefore not employed. In contrast, DFT profits from their enforcement, therefore we have used a $3\times3\times3$ $K$-grid to generate the orbitals. The results are illustrated in Figure 3.26.

The three most important characteristics of the covalent bond energy are the depth and position $r_{\min}$ of the minimum, and the absence of a positive lobe for $r > r_{\min}$. The last requirement is especially challenging, and is known to be an issue for DFT [74, 97]. The JS-pw and JS-DFT curves are rather similar, and they both fail in preventing the positive lobe. However, in the case of two electrons with opposite spin ($H_2$ molecule), it holds the identity $\mathrm{SD}^{\uparrow}_{\mathrm{pw}}\mathrm{SD}^{\downarrow}_{\mathrm{pw}} = 1$, and hence $\Psi_{\text{JS-pw}} = J_{ee\text{-YUK}}(R)$, i.e. the electrons are assimilable to bosons. On the contrary, when more than two electrons are simulated, the Slater determinant $\mathrm{SD}_{\mathrm{pw}}$ plays a relevant role. It then appears that the main difference between JS-pw and JS-DFT is related to the non-physical momentum imposed to the electrons by the plane waves contained in $\mathrm{SD}_{\mathrm{pw}}$.

To summarize, JS-pw is not adequate for the description of the electronic structure of hydrogen. By contrast, JS-DFT seems at first glance a suitable choice. However, by a deep analysis, it emerges that it is not able to account for the $H_2$ binding curve, which is a crucial ingredient for any dynamic simulation which aim to reproduce the $H_2$ dissociation. Therefore, also JS-DFT cannot be considered a completely reliable choice.

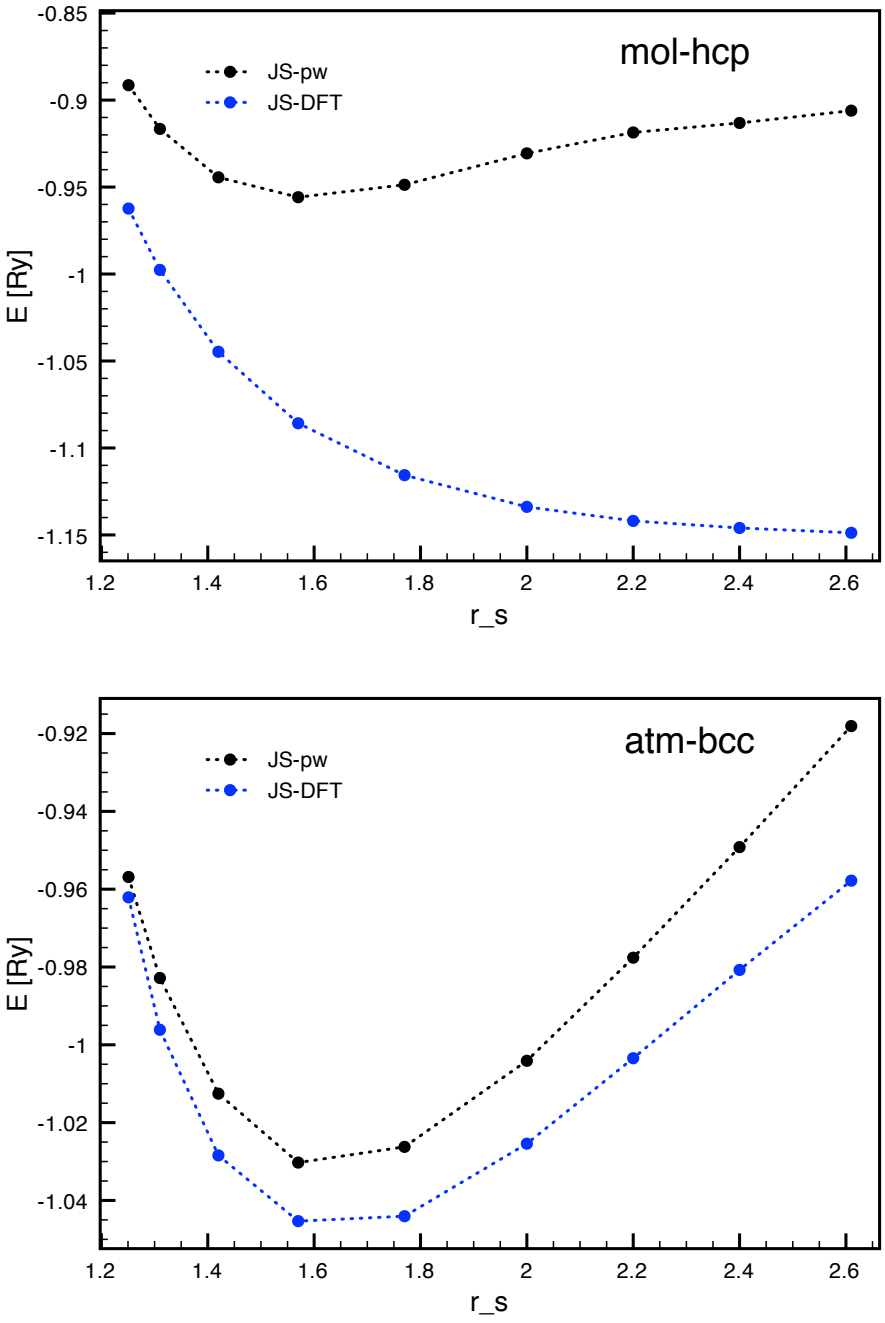


Figure 3.21: JS-pw and JS-DFT variational energies for atm-bcc and mol-hcp hydrogen at various densities. All the results were obtained simulating $N = 128$ atoms.

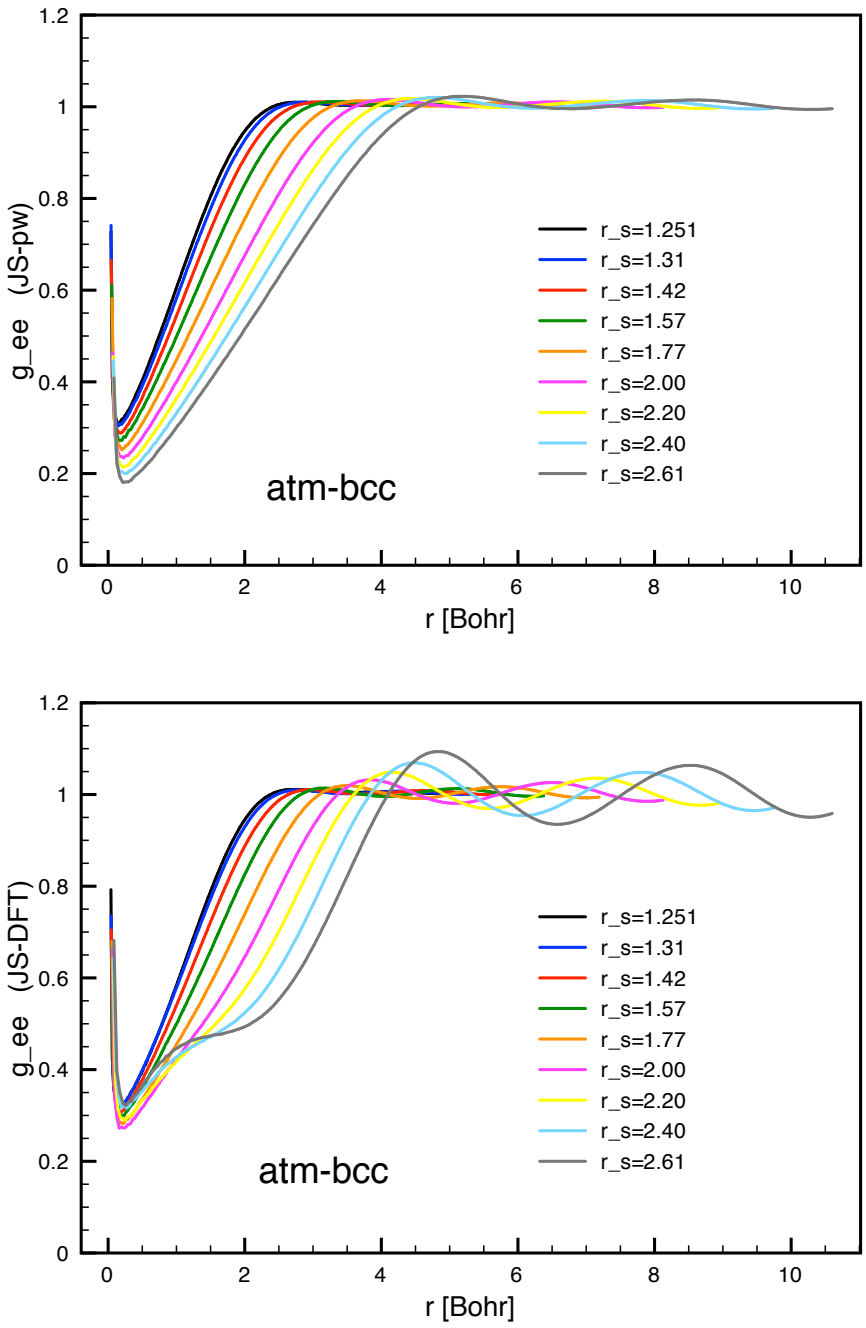


Figure 3.22: JS-pw and JS-DFT electron-electron pair correlation functions in atm-bcc hydrogen for various $r_s$.

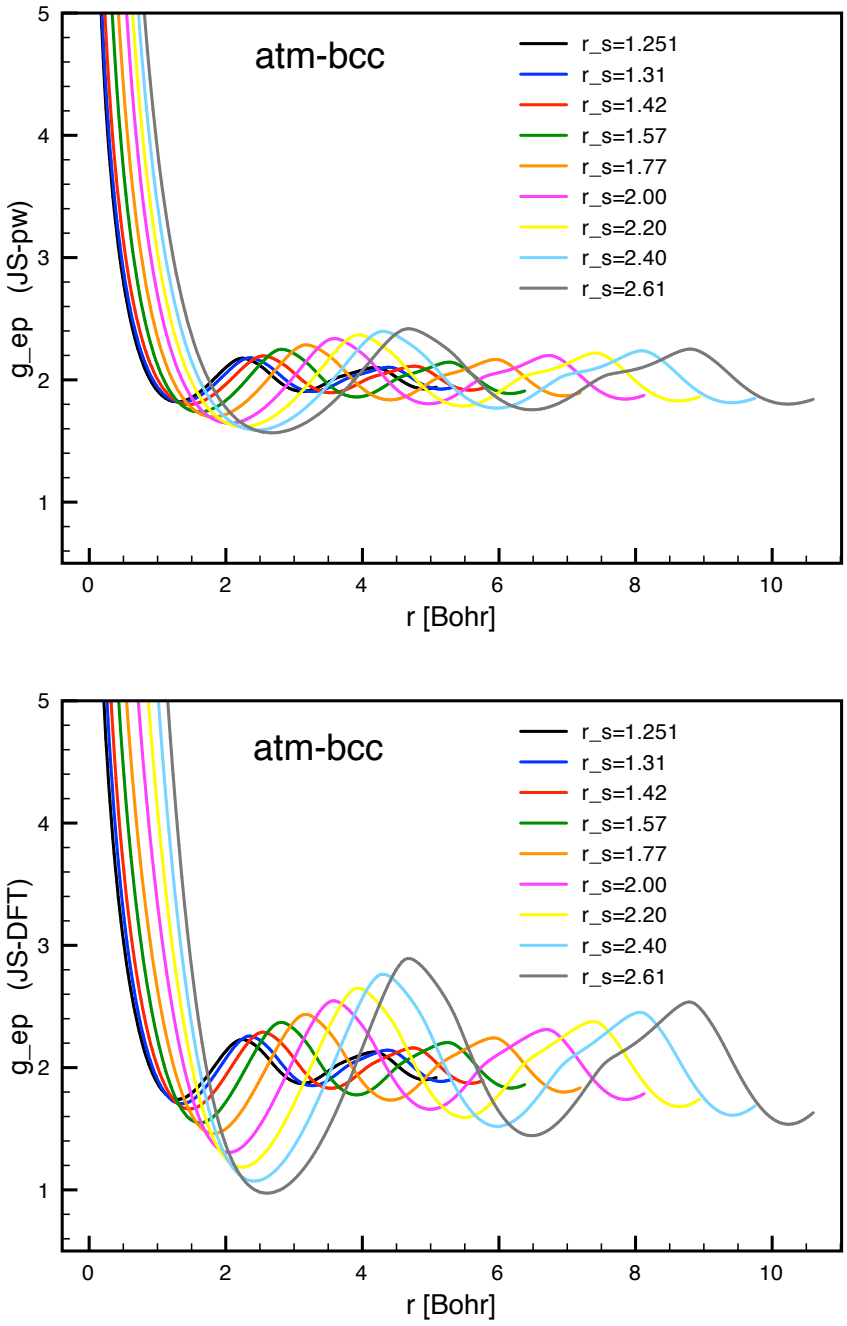


Figure 3.23: JS-pw and JS-DFT electron-proton pair correlation functions in atm-bcc hydrogen for various $r_s$.

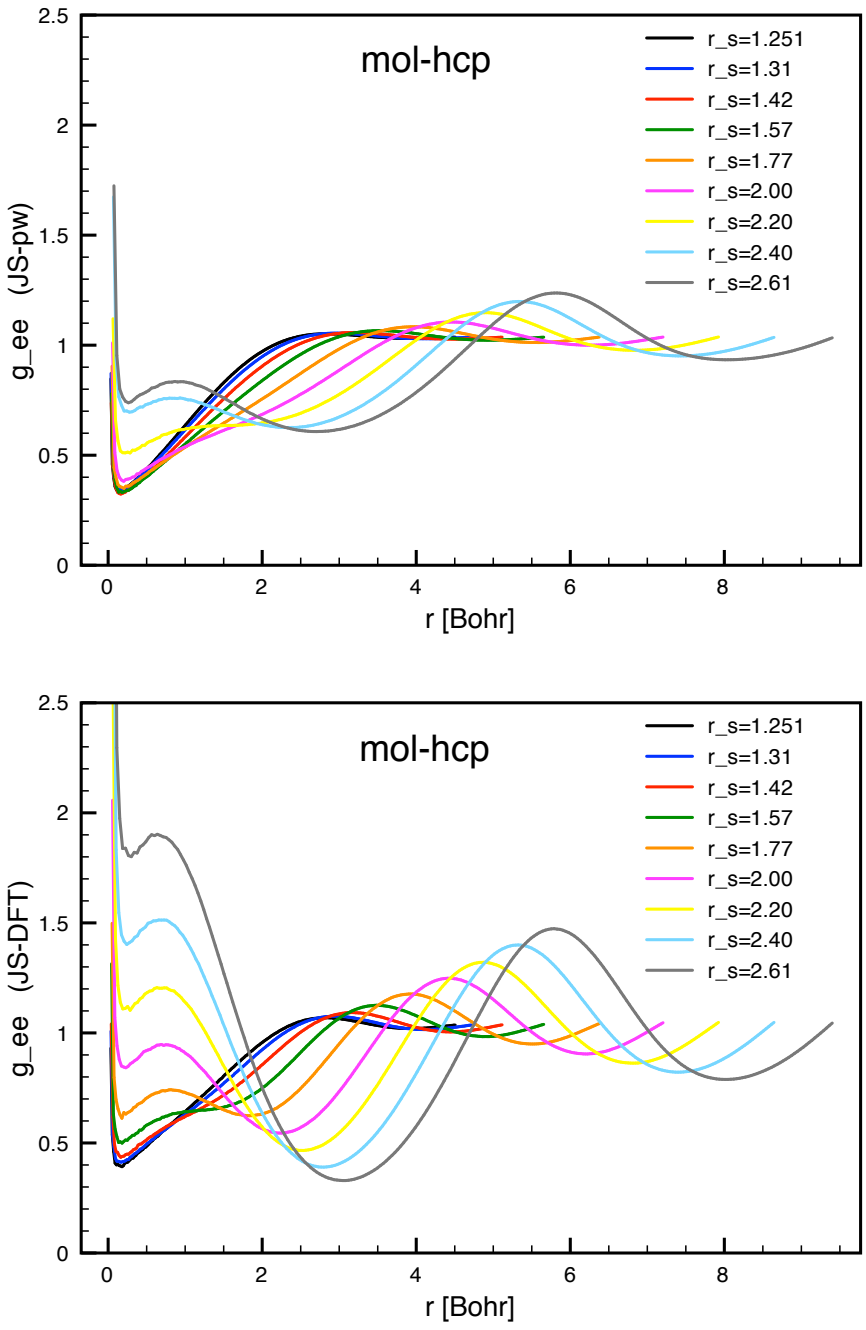


Figure 3.24: JS-pw and JS-DFT electron-electron pair correlation functions in mol-hcp hydrogen for various $r_s$.

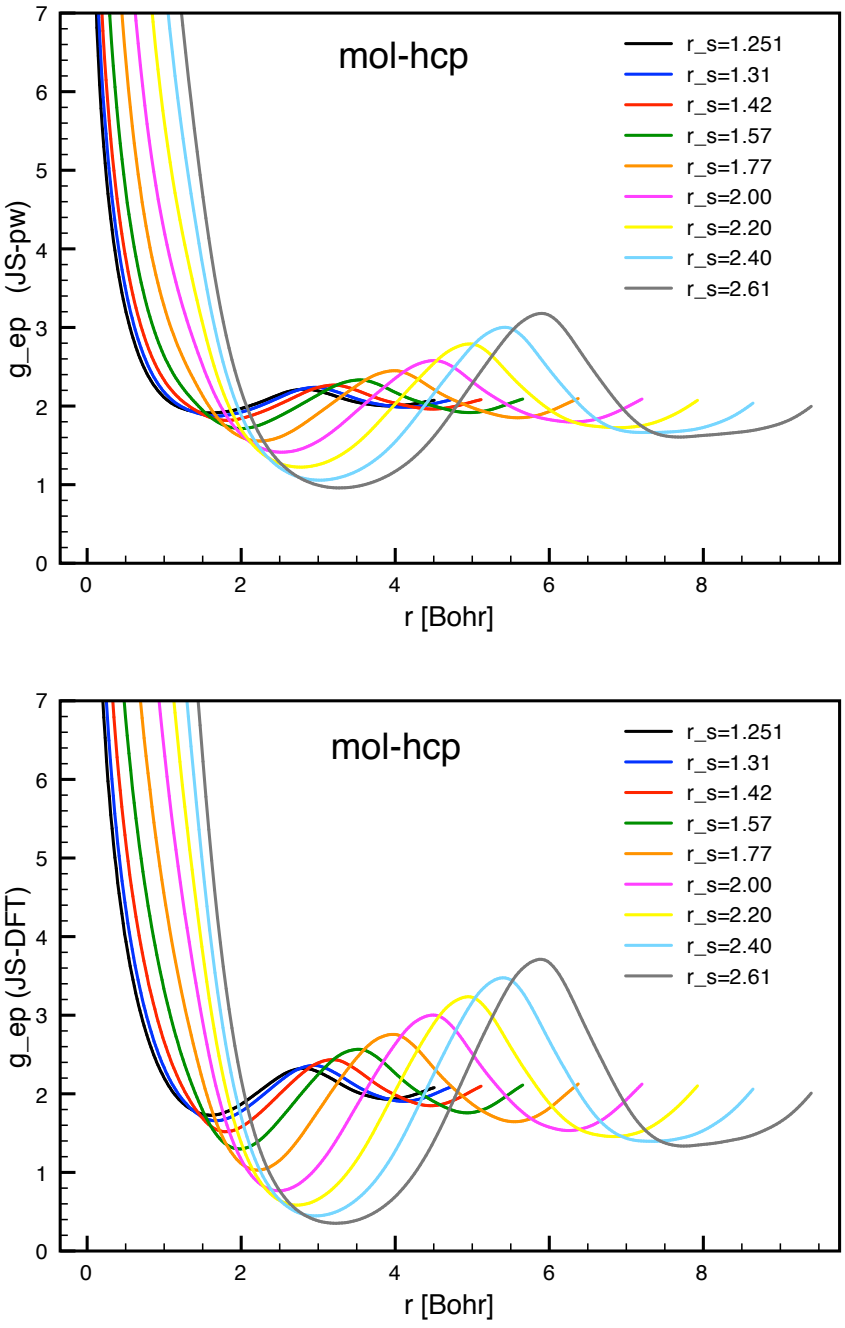


Figure 3.25: JS-pw and JS-DFT electron-proton pair correlation functions in mol-hcp hydrogen for various $r_s$.

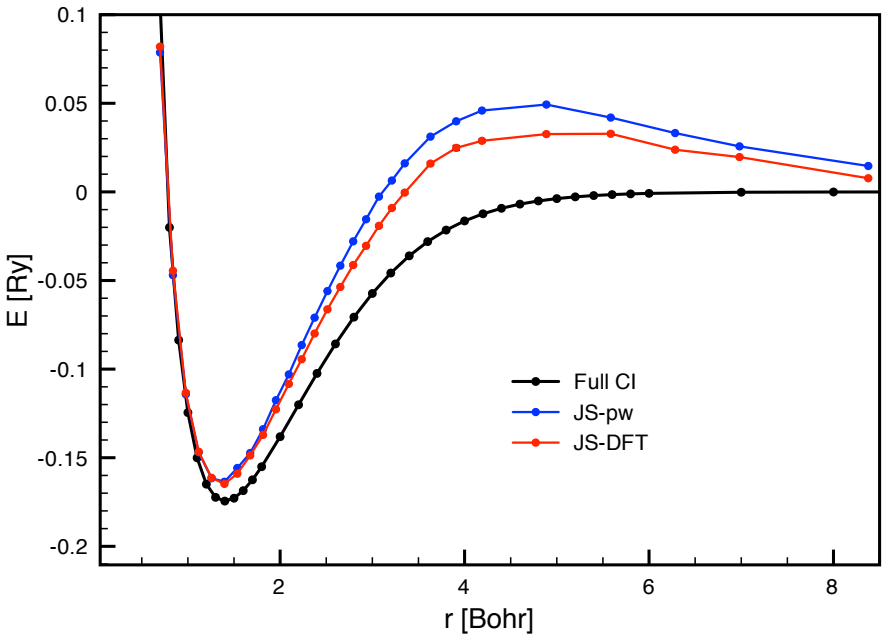


Figure 3.26: $H_2$ binding curve according to JS-pw and JS-DFT. Each curve uses a different off-set, namely the variational energy obtained with the trial wave function of interest at $r = 10$ Bohr. Full Configuration Interaction (Full CI) provides exact solutions up to the sixth digit [98, 99], and for our purpose can be considered exact.

### 3.2.3 ASWF-pw

Figure 3.27 illustrates the ASWF-pw variational energies.

The ASWF-pw trial wave function is able to reproduce the negative slope for the mol-hcp phase, hence allowing its possibility of existence. Moreover, the low density energies are significantly improved also in the atm-bcc phase. Therefore, ASWF-pw entails a major enhancement, compared to JS-pw. Nevertheless, it still cannot be considered competitive with JS-DFT, especially for low densities.

For high densities the ASWF-pw does not provide any improvement over the JS-pw, hence symmetric correlations beyond the second order do not seem to play an essential role.

The pair correlation functions reported in Figures 3.28, 3.29, 3.30, and 3.31 confirm that ASWF-pw approaches the accuracy of JS-DFT, especially for low densities. In particular, $g_{ee}(r)$ appears to better reproduce the covalent bond peak in the mol-hcp phase, compared to JS-pw. The electron-electron pair correlation function in the atm-bcc phase is also closer to the one obtained with JS-DFT. In fact, the two peaks for $r > 4$ Bohr are reproduced with similar amplitude, in contrast to JS-pw.

In Figure 3.32 we report the $H_2$ binding curve obtained with ASWF-pw. We remark that in the case of only two opposite-spin electrons ASWF-pw is identical to FSWF-pw, since $\mathrm{SD}^{\uparrow}_{\mathrm{pw}} = \mathrm{SD}^{\downarrow}_{\mathrm{pw}} = 1$. Considered this, the high accuracy of the results is not too surprising. In particular, the spurious positive lobe that appears with JS-pw and JS-DFT is not present anymore.

Results demonstrate that the ASWF-pw can describe a localized electronic structure, even when a completely inappropriate Slater determinant is employed. Nevertheless the quality of the results cannot be considered satisfactory, when compared to the JS-DFT trial wave function.

In conclusion, ASWF-pw provides an overall enhancement of the simple JS-pw, but the inclusion of realistic orbitals in the Slater determinant seems to be inevitable in order to attain a realistic electronic structure.

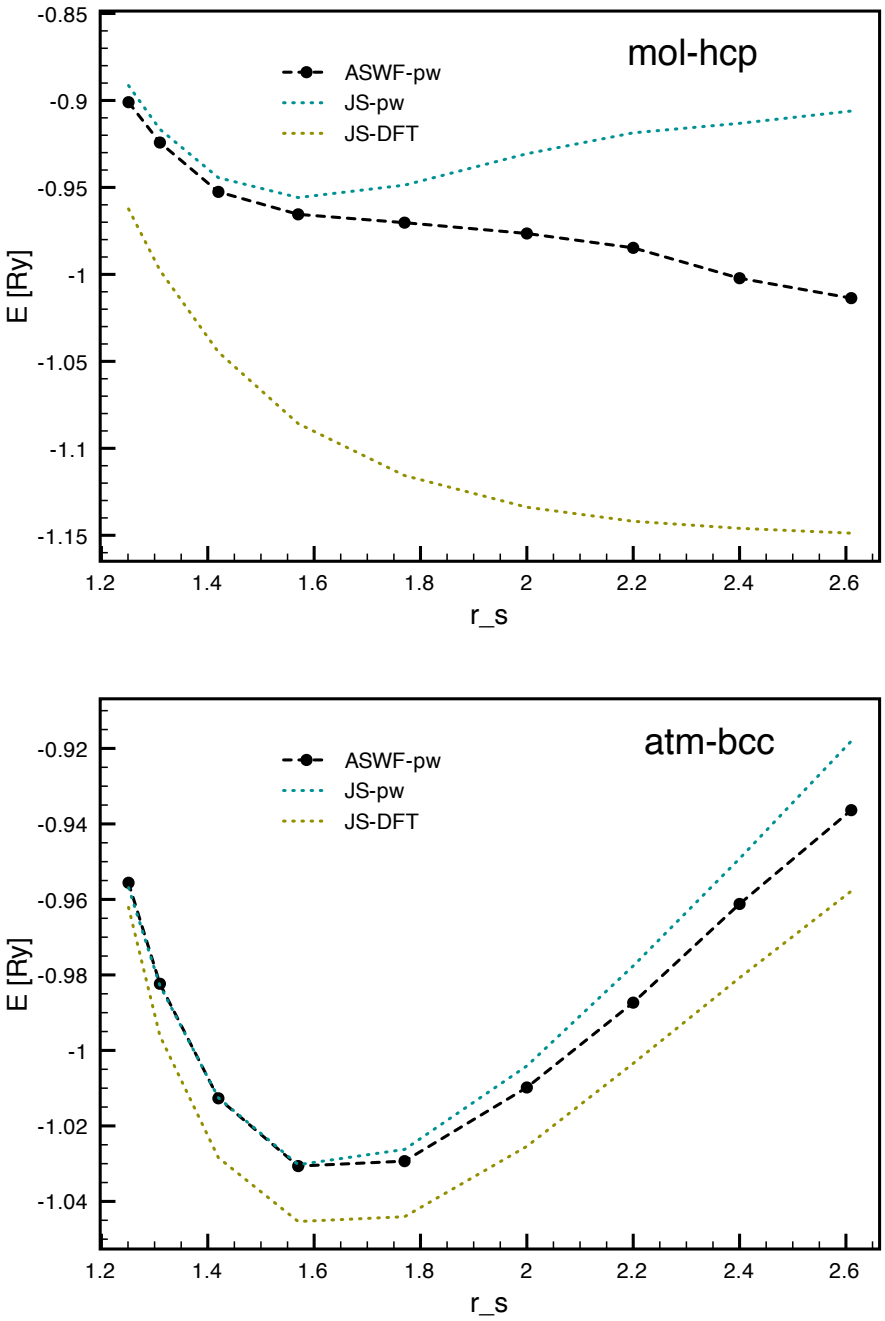


Figure 3.27: ASWF-pw variational energies for atm-bcc and mol-hcp hydrogen at various densities. All the results were obtained simulating $N = 128$ atoms.

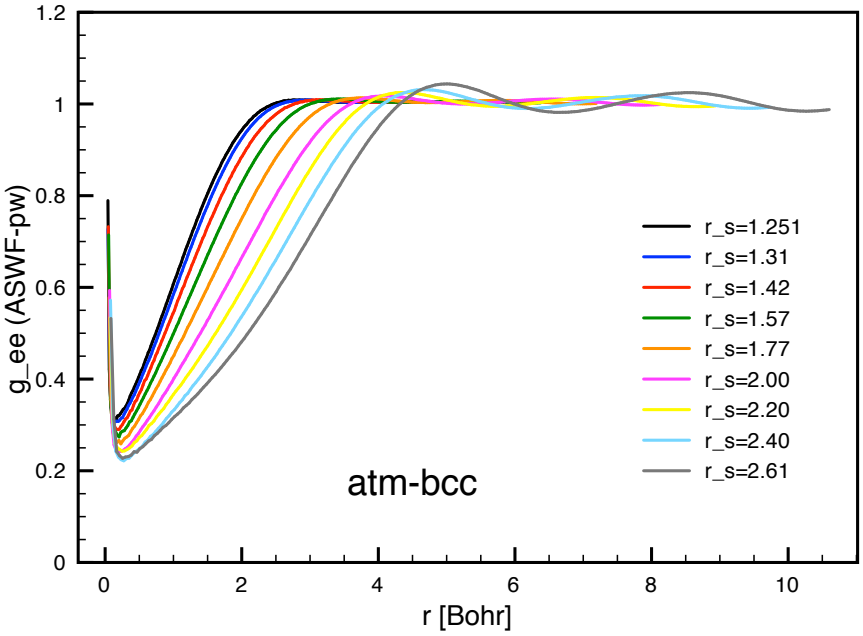


Figure 3.28: ASWF-pw electron-electron pair correlation functions in atm-bcc hydrogen for various $r_s$.

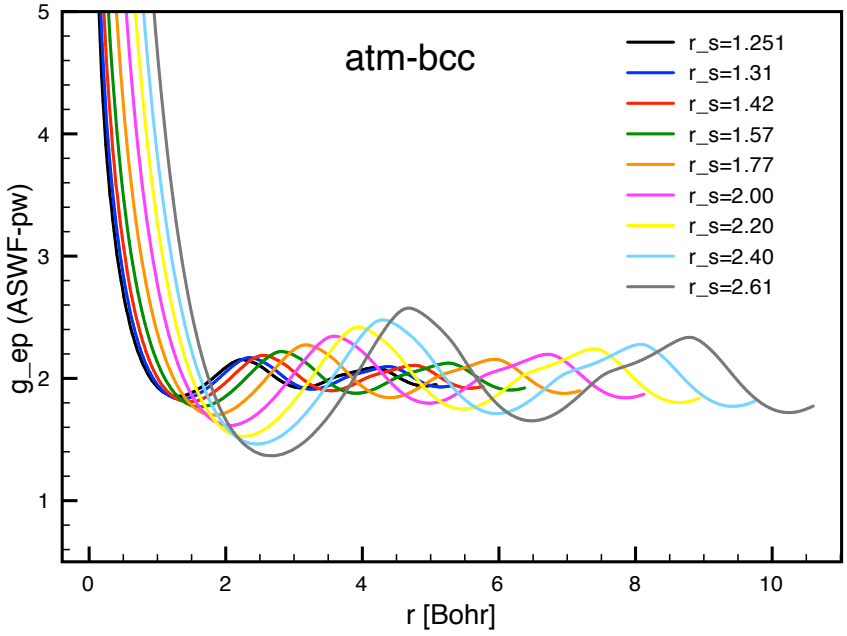


Figure 3.29: ASWF-pw electron-proton pair correlation functions in atm-bcc hydrogen for various $r_s$.

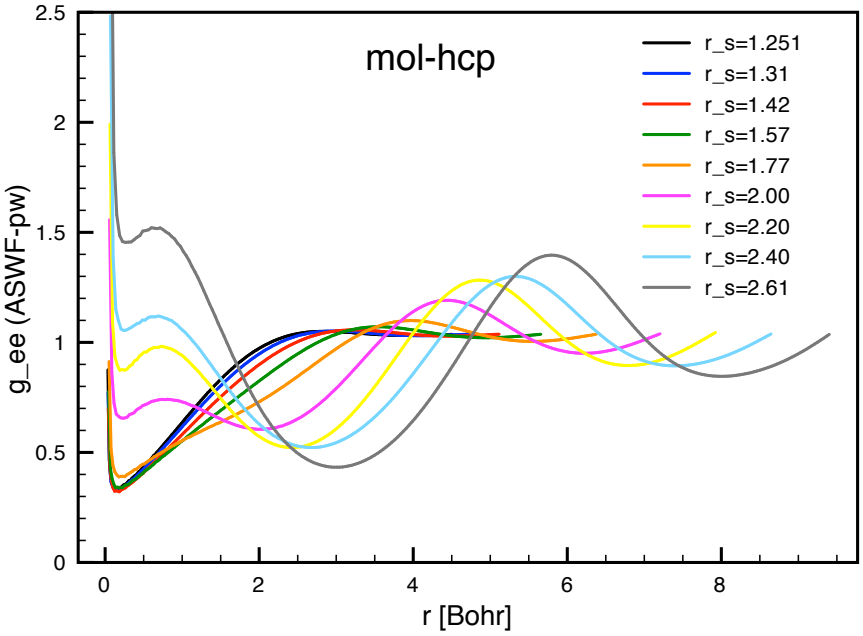


Figure 3.30: ASWF-pw electron-electron pair correlation functions in mol-hcp hydrogen for various $r_s$.

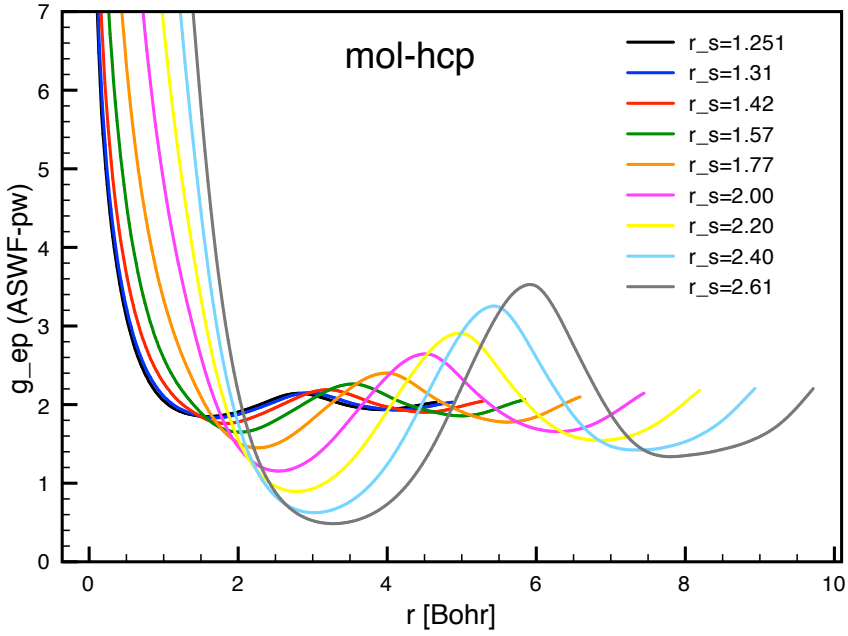


Figure 3.31: ASWF-pw electron-proton pair correlation functions in mol-hcp hydrogen for various $r_s$.

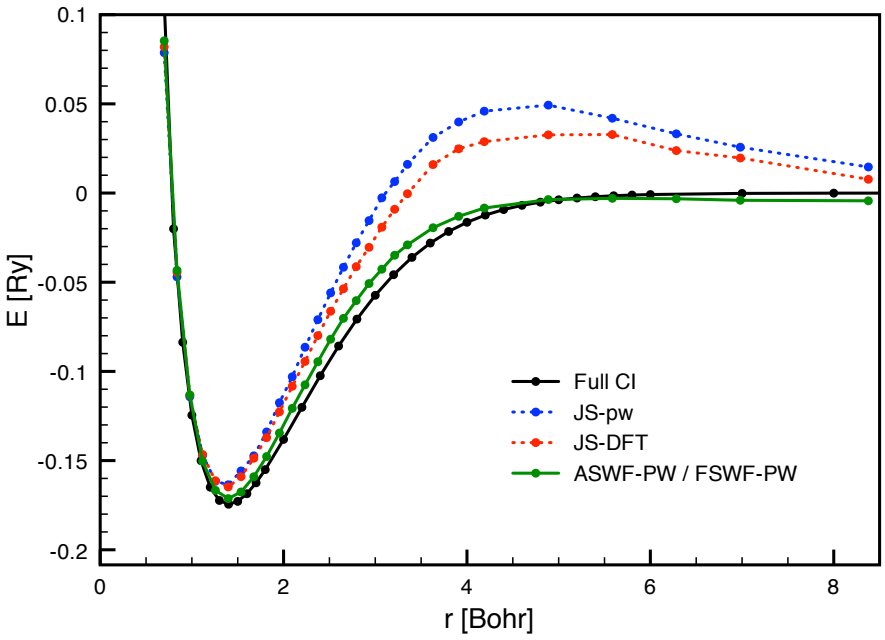


Figure 3.32: $H_2$ binding curve according to ASWF-pw. Each curve uses a different off-set, namely the variational energy obtained with the trial wave function of interest at $r = 10$ Bohr. Full Configuration Interaction (Full CI) provides exact solutions up to the sixth digit [98, 99], and for our purpose can be considered exact.

### 3.2.4 ASWF-DFT

The ASWF-DFT trial wave function is expected to compete with JS-DFT, since its construction ensures lower variational energies. In fact, either by setting the variational parameter $C$ equal to zero or by making vanish the $J_{ss\text{-YUK}}$ and $J_{sp\text{-YUK}}$ correlation factors, one recovers the JS-DFT form. However, as shown in Figure 3.33, ASWF-DFT has in general a higher variational energy than JS-DFT, especially in the low density regime. As we have already pointed out, this is impossible from a theoretical point of view, therefore the reason for such a failure must originate from the optimization algorithm. We have already seen in subsection 3.1.9 that the optimization of the SWF is somewhat troublesome, therefore this problem is not too surprising.

Considered these results, ASWF-DFT seems worthless of employment. It is less accurate than JS-DFT and it introduces several complications related to the additional integral required by the shadows.

Not too much to say about the pair correlation function in Figures 3.34, 3.35, 3.36, and 3.37, which are substantially the same as for JS-DFT.

Although ASWF-DFT seems to be disadvantageous from many points of view, it has the great merit of being able to accurately describe the $H_2$ binding curve, as can be seen in Figure 3.38. Therefore, even though its variational energy is slightly larger than for JS-DFT, its ability of correctly reproducing the dissociation energy of the covalent bond could play an essential role in a MD or MC simulation that includes the protons dynamic and aim to study the molecular dissociation of hydrogen.

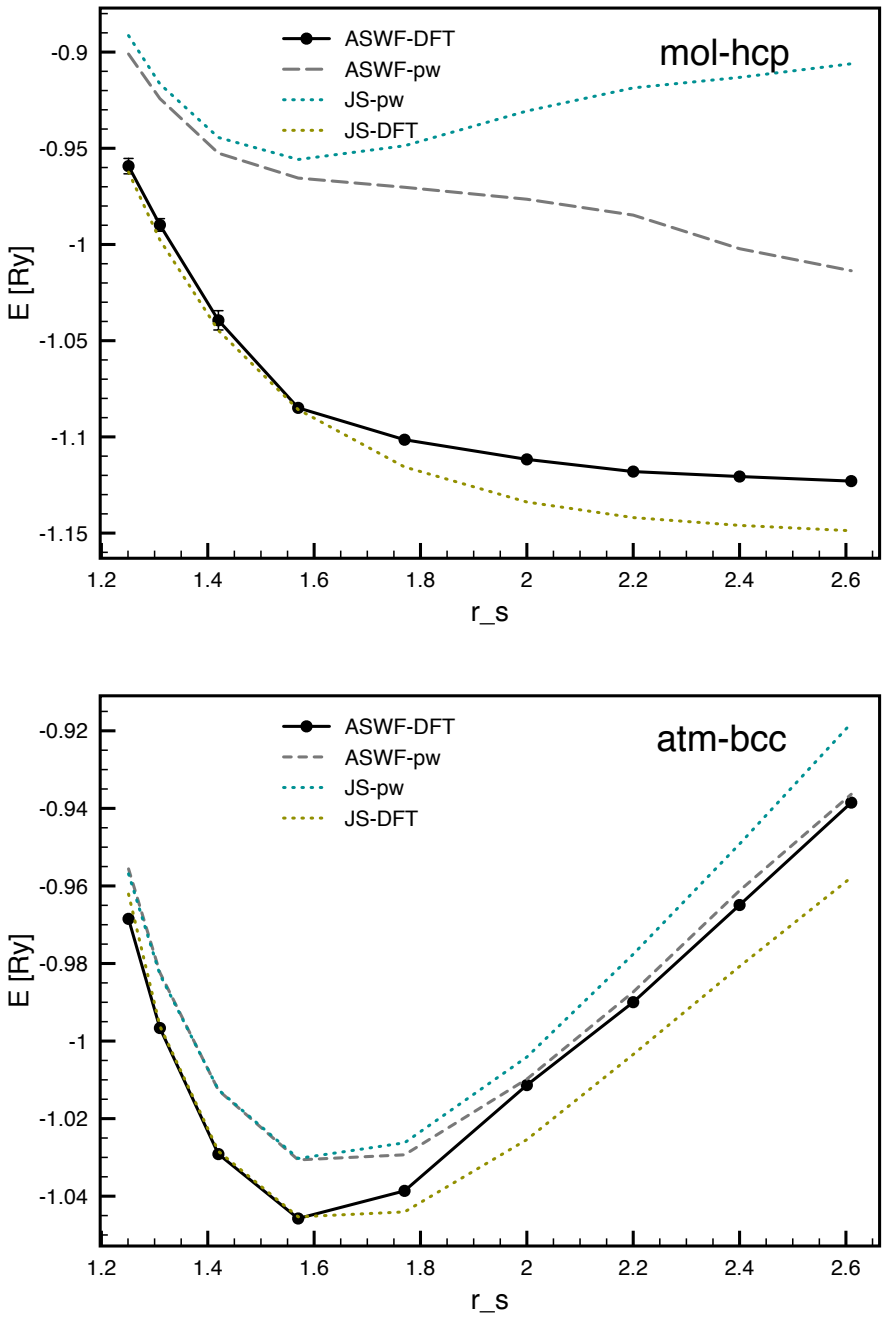


Figure 3.33: ASWF-DFT variational energies for atm-bcc and mol-hcp hydrogen at various densities. All the results were obtained simulating $N = 128$ atoms.

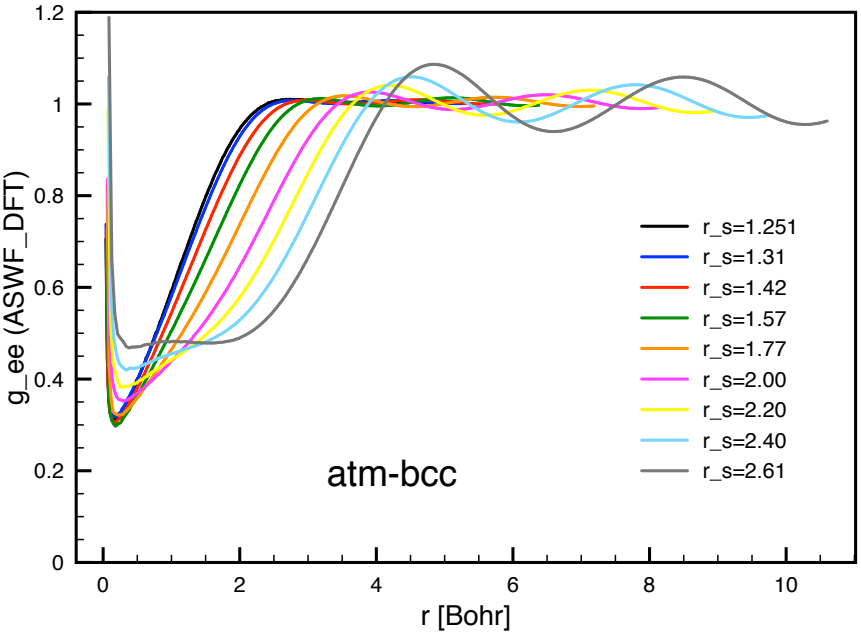


Figure 3.34: ASWF-DFT electron-electron pair correlation functions in atm-bcc hydrogen for various $r_s$.

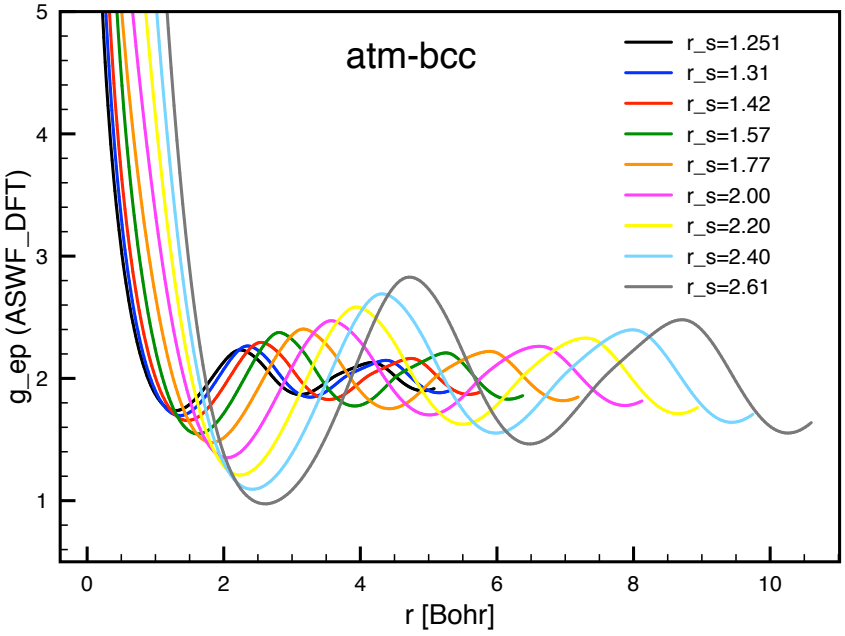


Figure 3.35: ASWF-DFT electron-proton pair correlation functions in atm-bcc hydrogen for various $r_s$.

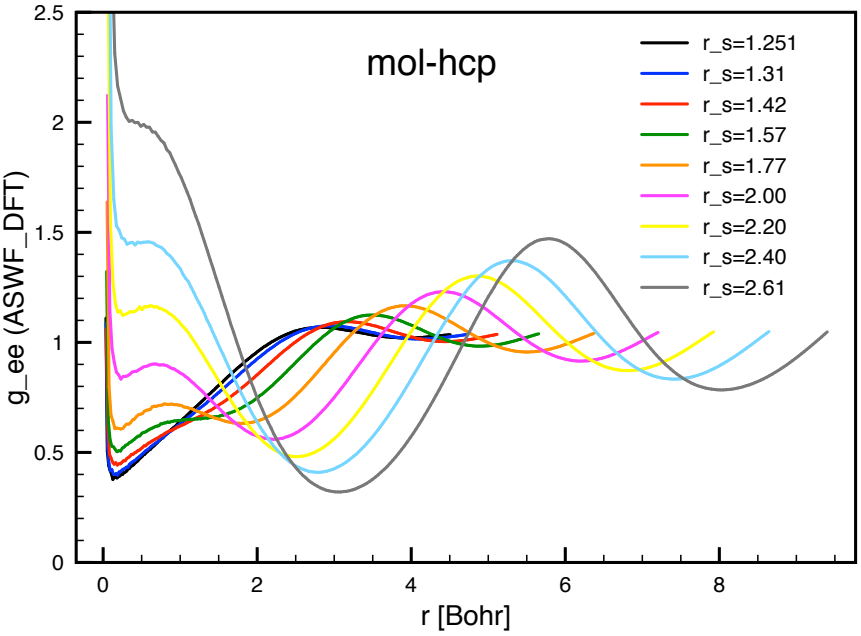


Figure 3.36: ASWF-DFT electron-electron pair correlation functions in mol-hcp hydrogen for various $r_s$.

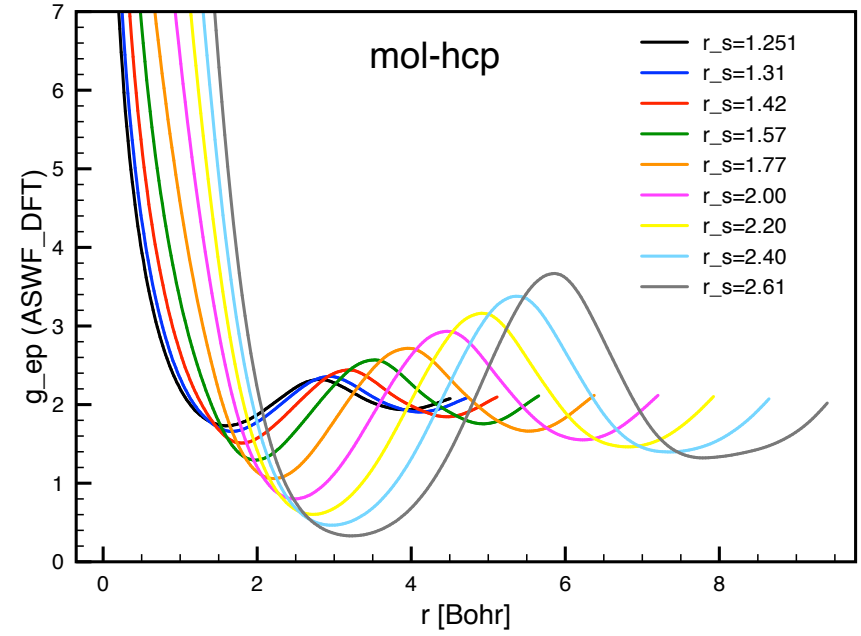


Figure 3.37: ASWF-DFT electron-proton pair correlation functions in mol-hcp hydrogen for various $r_s$.

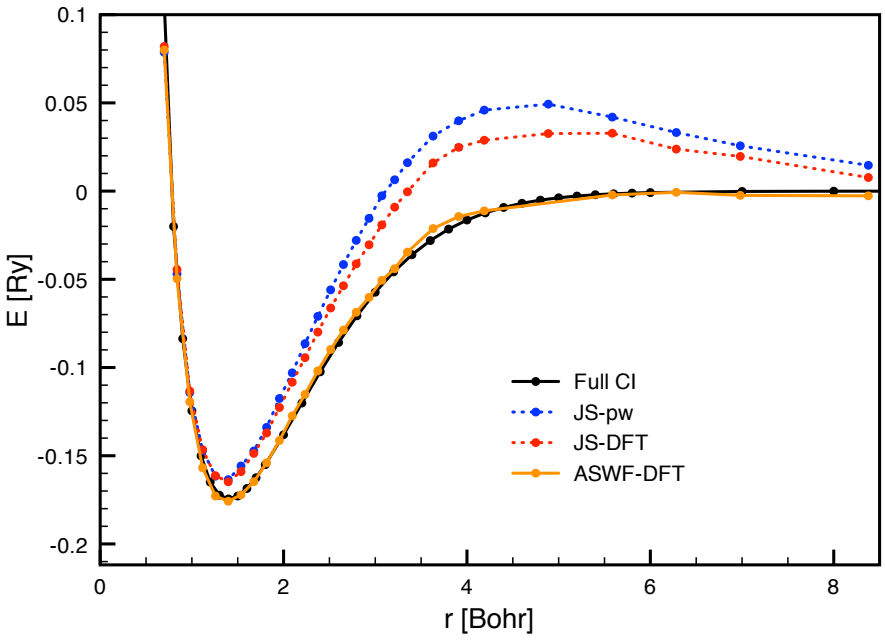


Figure 3.38: $H_2$ binding curve according to ASWF-DFT. Each curve uses a different off-set, namely the variational energy obtained with the trial wave function of interest at $r = 10$ Bohr. Full Configuration Interaction (Full CI) provides exact solutions up to the sixth digit [98, 99], and for our purpose can be considered exact.

### 3.2.5 FSWF-pw and FSWF-DFT

The FSWF sign problem prevented us from calculations with $N = 128$. In order to have a feeling of the possible improvements, we have repeated all the calculations with 16 atoms. The high computational costs involved convinced us to avoid the optimization process and to rather adopt the same variational parameters as for the ASWF. In fact, as we have discussed in subsection 3.1.9, this seemed to be a reasonably good approximation. For all the other trial wave functions, we have employed the same variational parameters as for $N = 128$.

The results are illustrated in Figure 3.39. The error bars are progressively larger for lower densities, so that for $r_s < 1.77$ the computed energies result to be not reliable any longer.

In the mol-hcp case, FSWF-pw provides a major improvement compared to the ASWF-pw. However, this enhancement is not sufficient to compete with JS-DFT, which demonstrates once again its reliability. In the atm-bcc phase, however, the assumption of using the same variational parameters as for the ASWF-pw appears to be inappropriate. In the case of FSWF-DFT such an assumption is definitely wrong, since energies are higher than for ASWF-DFT.

In summary, the employment of the FSWF is still far from a concrete application.

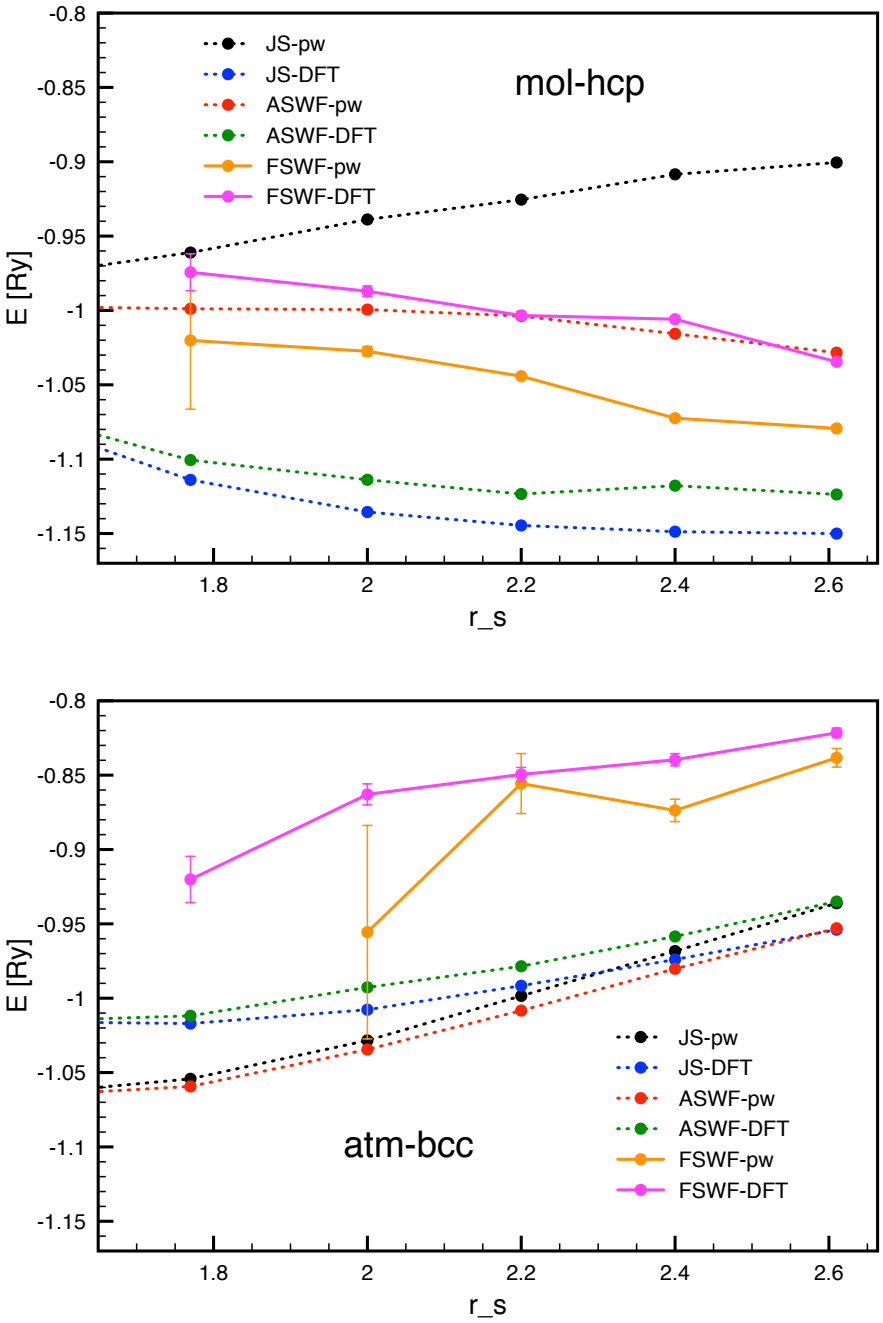


Figure 3.39: Variational energies for atm-bcc and mol-hcp hydrogen at various densities. All the results were obtained simulating $N = 16$ atoms.

## 3.3 Conclusion

By means of VMC, we have compared the variational energies and pair correlation functions of JS-pw, JS-DFT, ASWF-pw, and ASWF-DFT at various densities for two different hydrogen structures, atm-bcc and mol-hcp. Moreover, we have checked the accuracy of these trial wave functions in reproducing the $H_2$ binding curve. The sign problem of the FSWF prevented us from an analysis of the FSWF-pw and FSWF-DFT predictions.

From our results, it emerged that the employment of DFT orbitals is of fundamental importance for describing the electronic structure of hydrogen. This can be visualized in Figure 3.40, where we have summarized all the variational energies that we have computed. Both for the JS and the ASWF trial wave functions, the DFT orbitals provide substantially decreased variational energies, especially for the molecular phase.

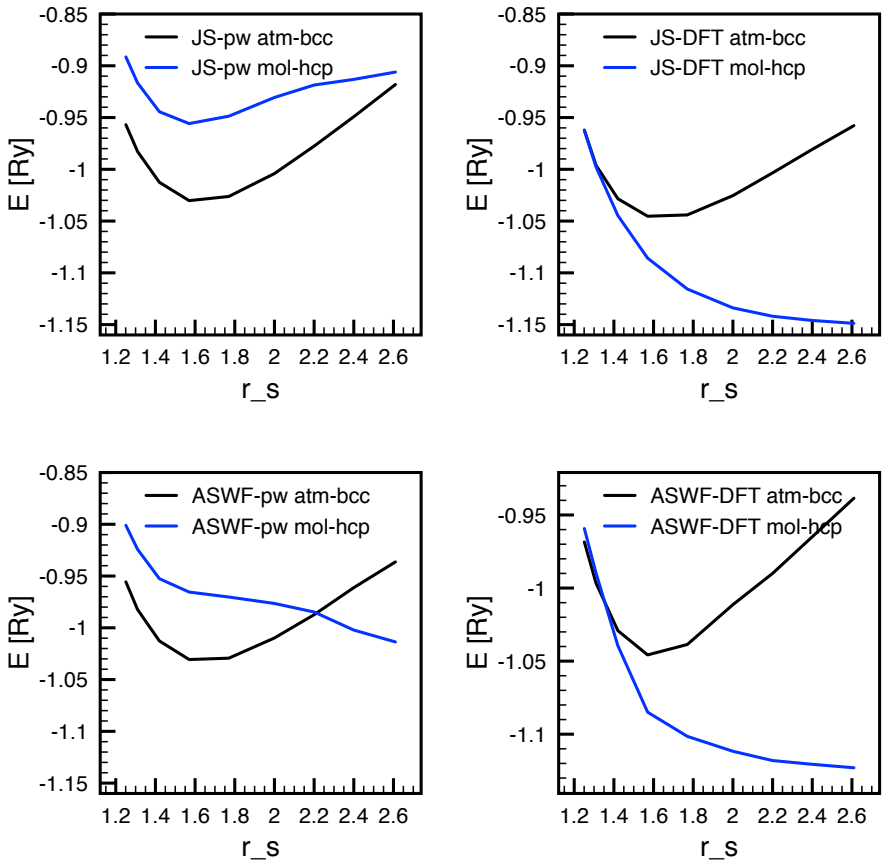


Figure 3.40: Comparison of the variational energies of JS-pw, JS-DFT, ASWF-pw, and ASWF-DFT for the atm-bcc and mol-hcp hydrogen structures at various densities.

However, the inspection of the $H_2$ binding curve revealed a fundamental

limitation of the JS-DFT trial wave function. In fact, a positive lobe appears when the molecule is stretched, as shown in Figure 3.41. However, such a misdescription can be eliminated by the shadows, which introduce high-order correlation effects. Figure 3.42 illustrates the difference between the $H_2$ VMC energies and the Full CI ones, demonstrating that ASWF-DFT obviates this problem. We impute the small fluctuations appearing for ASWF-DFT to the difficulties in the optimization process.

The aforementioned limitation of JS-DFT is not detectable by inspecting the total energy of the mol-hcp hydrogen structure. The reason for this is that in the mol-hcp structure we have assumed the $H_2$ molecules to be in the equilibrium position. In such a configuration, only the minimum of the binding curve is important, and in this respect JS-DFT performs as good as ASWF-DFT (see Figure 3.43). However, if JS-DFT is used to evaluate the Born-Oppenheimer potential in a simulation which includes the dynamic of the protons, the misdescription of the $H_2$ binding curve affects the evolution of the simulation, with unpredictable consequences.

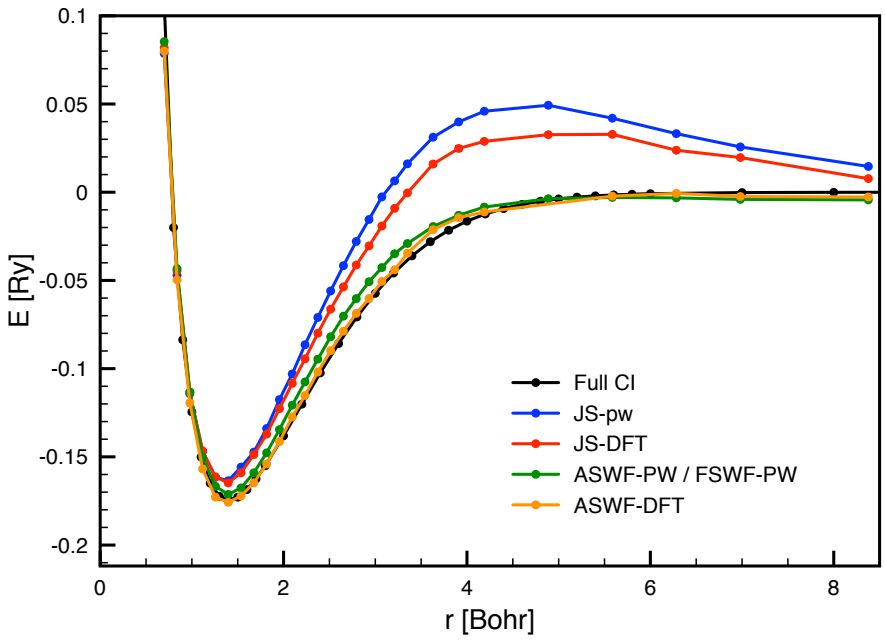


Figure 3.41: $H_2$ binding energy obtained by means of JS-pw, JS-DFT, ASWF-pw, and ASWF-DFT. For our purposes the Full CI values can be considered exact.

In conclusion, high order correlations effects are of fundamental importance for the evaluation of the Born-Oppenheimer potential of hydrogen. In this respect, ASWF-DFT represent an excellent alternative to other methods which account for such contributions, like Reptation Quantum Monte Carlo (see [48]) in the Coupled Electron-Ion Monte Carlo (CEIMC) method of Ceperley et al. [47, 51], or the VMC trial wave function used by Sorella [74]. In fact, a

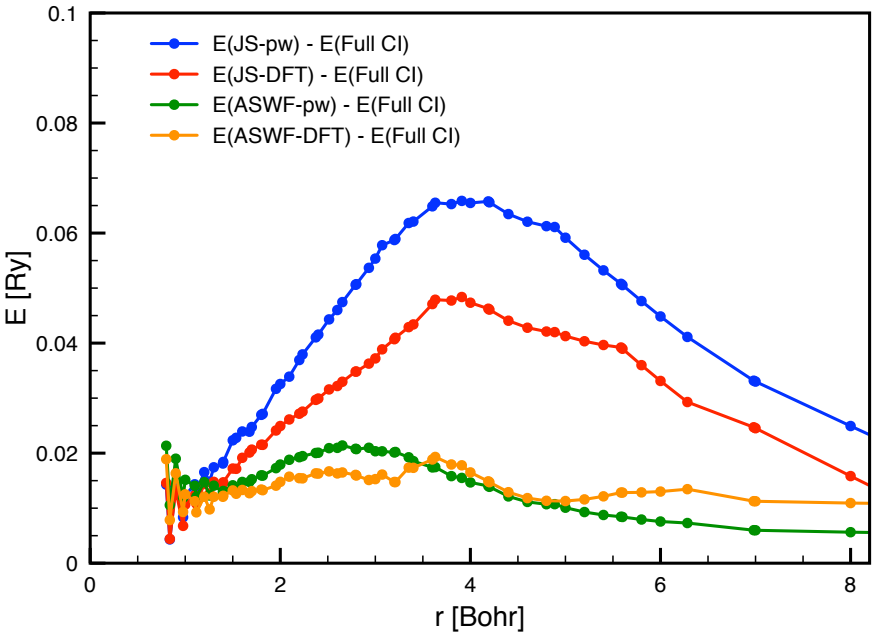


Figure 3.42: Energy difference between the $H_2$ Born-Oppenheimer energies obtained with JS-pw, JS-DFT, ASWF-pw, and ASWF-DFT, and the exact one provided by Full CI.

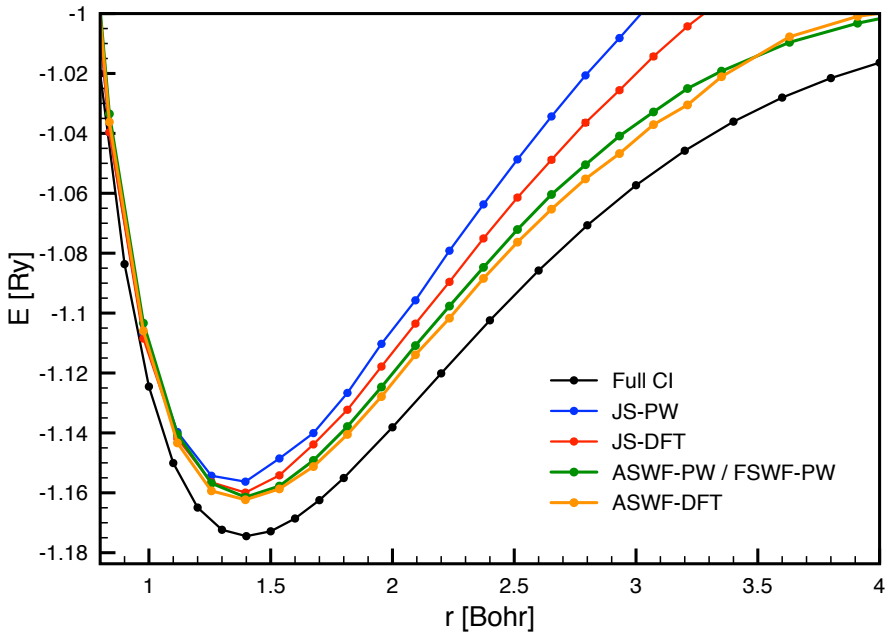


Figure 3.43: Variational Born-Oppenheimer energy of the $H_2$ molecule for various proton-proton distances by means of JS-pw, JS-DFT, ASWF-pw, and ASWF-DFT.

VMC calculation which employs ASWF-DFT is substantially less demanding in terms of computational resources than any projection method like Reptation Quantum Monte Carlo (RQMC). Moreover, it involves less variational parameters than the more sophisticated trial wave function used by Sorella. However, the employment of the ASWF-DFT is still not completely straightforward, due to the difficult optimization process, which requires an enhancement.

The natural perspective of this work is the application of the ASWF-DFT Born-Oppenheimer potential in the context of a simulation which includes the protonic dynamics. The limited computational cost required by ASWF-DFT allows for a larger number of protonic steps, and for the simulation of a larger number of atoms, within the same level of accuracy achieved by RQMC-CEIMC or the method of Sorella.

Another very interesting perspective could arise from the enhancement of the $SD_{dnfH}$ introduced in subsection 3.1.3. In fact, besides the obvious advantages related to orbitals which are less demanding in terms of computational time, a direct connection between the protons positions and the trial wave function opens for the possibility of their optimization in the context of SR. In other words, it is possible to efficiently generate new candidate crystal structures for hydrogen by optimizing the ASWF-dnfH (or JS-dnfH) trial wave function. This method would establish a significant improvement compared to the most accredited techniques, which rely on DFT [64–66].

Finally we would like to emphasize the importance of the advancements that we have accomplished for the variational optimization of the SWF. The SR method has proven to be very robust in this respect, and allowed us for the automatic minimization of the SWF variational energy. Such a systematic procedure opens for the possibility of an extensive employment of the SWF. However, the procedure that we have devised cannot be considered completely satisfactory yet. In fact, the occurrence of ASWF-DFT variational energies higher than for JS-DFT demonstrate that there are some residual problems that we have not been able to eliminate. Further investigation in this direction are in progress.

In conclusion, we believe that the ASWF-DFT trial wave function allows for a major improvement of the simulation techniques used to predict the phase diagram of hydrogen. In fact, it permits for accounting strong correlations in the Born-Oppenheimer potential, by means of an elegant formalism which requires small additional computational resources. The inclusion of such effects is essential in the context of phase transitions. Moreover, as we have shown, it is necessary for a proper description of the $H_2$ covalent bond.

# Chapter 4

# The Sign Problem of the Fermionic Shadow Wave Function

In subsection 1.3 we have anticipated that the FSWF, even though it is superior to the ASWF from many points of view, suffers from a sign problem that drastically reduces the convergence speed of the variational energy estimation. This entails a serious difficulty, since the application of the variational principle requires its computation. Moreover, the same problem affects the evaluation of any other observable.

The *sign problem of the FSWF* belongs to a general class of unsolved problems with similar characteristics, generally referred to as *sign problem*. The sign problem is one the most renowned open problems in physics, and impedes the exact determination of the fermionic ground state in the case of an interacting many-body system. More specifically, although exact algorithms exist, their computational time grows exponentially with the number of simulated particles $N$, contrary to bosonic systems, where the growth is polynomial. Until now all the efforts to find a solution[1] to the sign problem have been in vain, apart from some system-specific, not generalizable, cases (for example, see [100]).

In 2005 Troyer and Wiese provided arguments to support the conjecture that the fermionic sign problem is nondeterministic polynomial (NP) hard, so that its resolution would provide a solution to all NP problems [101]. Even though it has not been proven yet, this class of problems is so firmly believed to be unsolvable, that all the classical encryption schemes rely on this conjecture, and its negation would undermine their security. Nevertheless, attempts towards its solution can be pursued as long as a definitive proof of its impossibility has not been demonstrated. It should be remembered that also the determination of the exact bosonic ground state in polynomially growing time was considered to be impossible, until Green's Function Monte Carlo allowed for a breakthrough [102].

[1] A general algorithm that can determine numerically the exact ground state of an interacting many-body fermionic system and that scales polynomially in time with $N$, is what would be acknowledged as *solution* of the sign problem.

The FSWF sign problem, however, does not obviously fall in the category of the other QMC sign problems, and until now there has not been any report proving its exponential growth in computational time, therefore it might not fall into the class of NP hard problems. Generally, VMC does not suffer from a sign problem, and its occurrence with the FSWF is, at first sight, puzzling.

In this chapter we analyze in detail the characteristics of this particular sign problem, and the methods that we have devised to reduce its severity. The outline of this Chapter is the following:

- In section 4.1 we demonstrate the occurrence of the FSWF sign problem in our test system, liquid $^3$He;
- In section 4.2 we introduce novel techniques which take advantage of the underlying symmetries of the FSWF in order to alleviate the sign problem;
- In section 4.3 we discuss the importance of the marginal distribution of $R$, and present some actual methods to exploit it in order to alleviate the sign problem;
- In section 4.4 we compare the efficiencies of all the previously presented methods;
- Finally, in section 4.5, we introduce novel insights on the origin of the FSWF sign problem, obtaining analytical predictions of its magnitude, and reducing the complexity of the problem to a one-dimensional integral.

## 4.1 Introduction

Our test system for the FSWF sign problem is an unpolarized 3D liquid consisting of $^3$He atoms with density $\rho = 0.016588$ Å$^{-3}$, very close to the liquid equilibrium density [103]. The atomic interaction is described through the Aziz potential HFDHE2 [104, 105], and periodic boundary conditions are employed in order to mimic an essentially infinite system. The energy is always expressed in terms of Kelvin, hence referring to the energy per atom. The adopted FSWF reads

$$\Psi_{\mathrm{FSWF}}(R) = e^{-\frac{1}{2}\sum_{i<j}(\frac{b}{r_{ij}})^5} \int dS\, e^{-\sum_{i=1}^{N_p} C|\mathbf{r}_i-\mathbf{s}_i|^2}\, \mathrm{SD}_{\mathrm{pw}}(S)\, e^{-\sum_{i<j} c_1 V(c_2 s_{ij})}, \tag{4.1}$$

where $V$ is the Aziz potential, and $b$, $C$, $c_1$, and $c_2$ are variational parameters. We set them to be $b = 2.76$ Å, $C = 0.55$ Å$^{-2}$, $c_1 = 0.11$ K$^{-1}$, $c_2 = 0.88$, as reported to be optimal in [6]. Indeed, in the mentioned paper these values referred to the ASWF, but as we have shown in subsection 3.1.9, it is very reasonable to assume the same for the FSWF. When a SD of simple plane waves $e^{i\mathbf{k}\cdot\mathbf{r}}$ is used, the occurrence of a drift (i.e. $\Sigma_\beta \mathbf{k}_\beta \neq 0$) as well as anisotropy effects arise[2]. The adopted remedy was to consider only magic numbers for $N$, i.e. numbers that fill the momenta shell. For a 3D polarized system, they are

[2]The TABC obviate these problems, but we have not implemented this technique for $^3$He, since we were only interested in the analysis of the sign problem and not in its physical properties.

| **Trial wave function** | **Energy** [K] |
|---|---|
| JS | −1.08(3) [106] |
| JTS | −1.61(3) [106] |
| JBS | −1.55(4) [106] |
| JTBS | −1.91(4) [106] |
| ASWF | −1.32(2) |
| FSWF | −1.93(5) |
| Experiment | −2.47(1) [106] |

Table 4.1: Energy obtained simulating 54 atoms of liquid $^3$He. J refers to the Jastrow, T to the triplet, B to the backflow and S to the $\mathrm{SD_{pw}}$. The employed FSWF and ASWF variational parameters were the same, and are reported in the text.

1, 7, 19, 27, 33, etc., and therefore for an unpolarized system are 2, 14, 38, 54, 66, etc. Without the TABC, it is possible and computationally convenient to write $\mathrm{SD_{pw}}$ in terms of $\sin(\mathbf{k}\cdot\mathbf{r})$ and $\cos(\mathbf{k}\cdot\mathbf{r})$ instead of $e^{i\mathbf{k}\cdot\mathbf{r}}$. For this reason we assume that

$$\Psi_{\mathrm{FSWF}}(R) = \int dS\, \Gamma(R,S) \tag{4.2}$$

and its integrand $\Gamma$ are real functions. However, all the following discussions can be extended to the more general case of complex functions. Since we have not implemented the TABC, the occurrence of relevant finite size effects is to be expected.

We sampled by means of the $M(RT)^2$ algorithm, submitting single atom random steps to the acceptance/rejection process. We have not exploited the possibility of using the gradient of the trial wave function to drive the random steps. Finally, the step length was adjusted in order to obtain an acceptance rate of $\sim 50\%$.

In Table 4.1, we report the variational energies obtainable with several different trial wave functions. The ASWF introduces a significant improvement to the simple JS trial wave function, but cannot replicate the enhancements provided by the Triplet and the Backflow terms. By contrast, the FSWF has a variational energy as low as a JTBS wave function, despite its simple functional form.

In order to evaluate the mean value of the energy per atom $E$ and its unbiased error bar, we have divided the data into $n_{\mathrm{block}}$ disjoint blocks (typical values for $n_{\mathrm{block}}$ are between 4 and 50), each one with its corresponding average value $E_j^{\mathrm{block}}$. It is subsequently possible to calculate the average $\langle E\rangle_{\mathrm{block}}$ and the standard deviation $\sigma_{\mathrm{block}}$ as following:

$$\langle E\rangle_{\mathrm{block}} = \frac{1}{n_{\mathrm{block}}}\sum_{j=1}^{n_{\mathrm{block}}} E_j^{\mathrm{block}}, \tag{4.3a}$$

$$\sigma^2_{\mathrm{block}} = \frac{1}{n_{\mathrm{block}}-1}\sum_{j=1}^{n_{\mathrm{block}}} \left(E_j^{\mathrm{block}} - \langle E\rangle_{\mathrm{block}}\right)^2. \tag{4.3b}$$

This is the so-called blocking technique [24], used to estimate the variance when

| | **ASWF** | | **FSWF** | |
|---|---|---|---|---|
| $C\,[\text{Å}^{-2}]$ | $E\,[\text{K}]$ | $\eta\,[a.u.]$ | $E\,[\text{K}]$ | $\eta\,[a.u.]$ |
| 14 atoms | | | | |
| 0.32 | $-0.9134(25)$ | 23.5 | $-1.5730(662)$ | $0.005 \cdot 10^{-2}$ |
| 0.43 | $-1.1753(33)$ | 13.5 | $-1.8620(168)$ | $0.08 \cdot 10^{-2}$ |
| 0.55 | $-1.2725(30)$ | 16.3 | $-1.9664(78)$ | $0.39 \cdot 10^{-2}$ |
| 0.64 | $-1.2386(108)$ | 1.3 | $-1.9123(44)$ | $1.21 \cdot 10^{-2}$ |
| 0.75 | $-1.0976(109)$ | 1.2 | $-1.7622(52)$ | $0.86 \cdot 10^{-2}$ |
| 0.86 | $-0.9154(157)$ | 1.7 | $-1.5262(45)$ | $1.16 \cdot 10^{-2}$ |
| 0.96 | $-0.6606(137)$ | 1.3 | $-1.2472(58)$ | $0.70 \cdot 10^{-2}$ |
| 38 atoms | | | | |
| 0.32 | $-0.905(21)$ | 0.43 | $-0.634(2548)$ | $0.0003 \cdot 10^{-5}$ |
| 0.43 | $-1.202(17)$ | 0.65 | $-2.516(771)$ | $0.003 \cdot 10^{-5}$ |
| 0.55 | $-1.335(19)$ | 0.52 | $-1.835(118)$ | $0.1 \cdot 10^{-5}$ |
| 0.64 | $-1.332(24)$ | 0.33 | $-1.979(37)$ | $1. \cdot 10^{-5}$ |
| 0.75 | $-1.194(24)$ | 0.33 | $-1.774(23)$ | $3. \cdot 10^{-5}$ |
| 0.86 | $-1.045(36)$ | 0.15 | $-1.554(15)$ | $7. \cdot 10^{-5}$ |
| 0.96 | $-0.815(31)$ | 0.20 | $-1.255(17)$ | $5. \cdot 10^{-5}$ |

Table 4.2: Energy and efficiency comparison between the VMC calculations which employ the ASWF and the ones which employ the FSWF.

successive samples are correlated. In fact, its straightforward evaluation would result in an underestimation, due to serial correlations of successive data points (autocorrelations). Given that $\frac{M}{n_{\text{block}}}$ is large enough, the autocorrelations between the $E_j^{\text{block}}$ values are suppressed and thus $\sigma$ can be correctly estimated. Plotting $\sigma_{\text{block}}$ as a function of $n_{\text{block}}$, a plateau is to be observed in correspondence to the correctly estimated error.

We remark that, mathematically speaking, $\langle E \rangle_{\text{block}}$ can change for different values of $n_{\text{block}}$, but as long as $M$ is large enough, each $E_j^{\text{block}}$ will be very close to $E$, and eventually $\langle E \rangle_{\text{block}}$ will be independent from the choice of $n_{\text{block}}$.

In Figure 4.1 we report $\sigma_{\text{block}}$ and $\langle E \rangle_{\text{block}}$ for various $n_{\text{block}}$, obtained simulating 66 $^3$He atoms on 640 cpus for 41 hours. The expected plateau for $\sigma_{\text{block}}$ appears only for very small $n_{\text{block}}$, and the estimated energies $\langle E \rangle_{\text{block}}$ strongly fluctuate for $n_{\text{block}} > 10$. These results indicate that the sampling length was barely sufficient for a reliable estimate, despite the enormous computational resources employed. This is a clear evidence of what we call the sign problem of the FSWF.

In order to quantitatively address the sign problem of the FSWF, a criterion to compare different methods is necessary. In that respect, the variance (or the standard deviation $\sigma$) allows for reliable comparisons, but it is not completely satisfactory, since the different computational time is not taken into account. A more accurate parameter is the *efficiency* $\eta$, defined as

$$\eta = \frac{1}{\text{variance} \times \text{computational time}}\,. \tag{4.4}$$

This definition fulfills the intuitive requirement of invariance under changes of

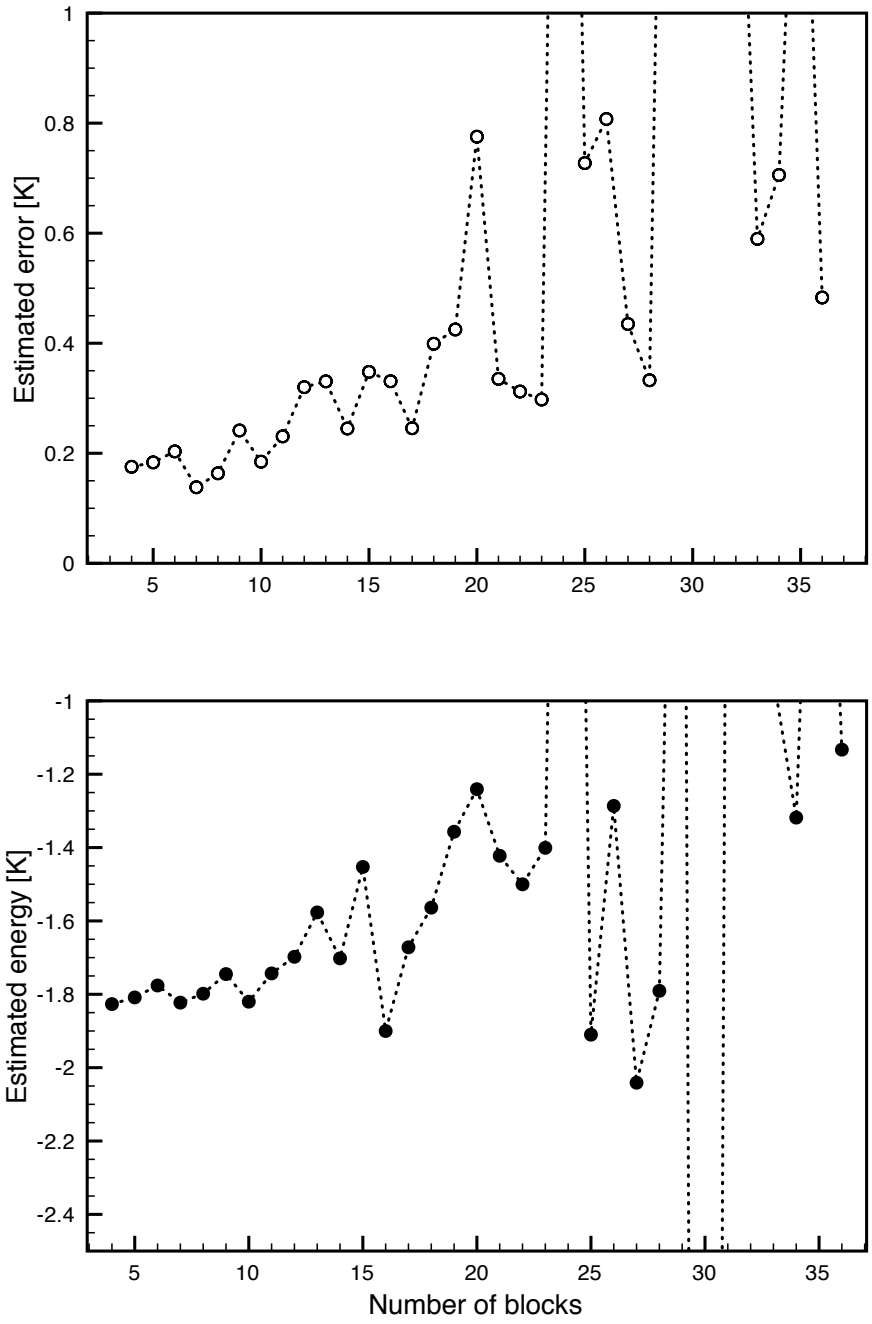


Figure 4.1: $\sigma_{\text{block}}$ (empty circles) and $\langle E \rangle_{\text{block}}$ (solid circles) versus $n_{\text{block}}$, employing the FSWF to simulate a large number of $^3$He atoms. The calculation required 41 hours on 640 cores, by seting $N = 66$ and $M = 128 \cdot 10^8$.

| N | Steps |
|---|---|
| 14 | 828(29) |
| 38 | 741(09) |
| 54 | 864(13) |
| 66 | 873(16) |
| 114 | 654(05) |
| 162 | 832(11) |

Table 4.3: Average single-particle proposed moves necessary for a sign change of the weight. In all cases the acceptance rate was around 50%.

the sampling length. In fact, the computational time is proportional to $M$, while the variance inversely decreases with $M$, thus their product is independent of $M$.

A clear demonstration of the sign problem is provided by Table 4.2, where we report an efficiency comparison between the ASWF and the FSWF. Whereas employing the ASWF $\eta$ remains substantially constant while changing the kernel variational parameter $C$, for the FSWF a drastic decrease can be observed when $C$ becomes smaller than $\sim$ 0.64 Å$^{-2}$. The same holds for the variations of $N$, which decreases the efficiency for the FSWF by orders of magnitude, whereas the results obtained employing the ASWF are scarcely affected.

In liquid $^3$He and with the adopted variational parameters, the sign problem is particularly severe, but this is not always the case. In fact, for large $C$, which physically correspond to localized systems, $S_1$ and $S_2$ are tightly packed around $R$, and hence SD($S_1$) and SD($S_2$) have the same sign, and the sign problem does not occur.

It is important to emphasize that the sign problem of the FSWF is substantially different from the infamous fermion sign problem of projection methods such as Green Function [102] or Diffusion Monte Carlo [107]. In fact, there is no fundamental reason why it should not be possible to evaluate an integral, even though the integrand strongly fluctuates, whereas in the case of projection methods there is an intrinsic difficulty to selectively sample a probability density function generated by a squared antisymmetric function.

As we have already mentioned, large values of $N$ aggravate the sign problem. For this reason, it might be expected that the rate of sign changes (equation (1.44)) would increase with increasing $N$. Unexpectedly, this is not the case, as shown by Table 4.3. Thus, it appears that the density of nodes in the one-particle projection space remains almost constant. Nevertheless, in the global $9N$-dimensional space ($3N$ for the particles, $6N$ for the shadows), the density of nodes rises with $N$.

As a last generic comment, we point out that an antisymmetric term in the integral always entails convergence problems in Monte Carlo integrators. Therefore, the sign problem that we are discussing belongs to a more general class of problems.

## 4.2 FSWF Symmetries and Antithetic Variates

A general method to reduce the variance of an integral is the introduction of the so-called *antithetic variates* (see e.g. [24], pag. 93). The antithetic variates are $n > 0$ points $a_1, \dots a_n \in V$, where $V$ is a vector space, mapped from the sampled point $X \in V$ by a set of $n$ functions

$$A_j : V \to V \qquad j \in \{1, \dots n\} \tag{4.5}$$

as

$$A_j(X) = a_j \qquad \forall j = 1, \dots n\,. \tag{4.6}$$

Assuming that the maps $A_j$ fulfill the requirement

$$\int dX\, f(A_j(X)) = \int dX\, f(X) \qquad \forall j = 1, \dots n\,, \tag{4.7}$$

the integral

$$\int dX\, f(X) \tag{4.8}$$

can then be replaced by

$$\frac{1}{n+1} \int dX \left( f(X) + \sum_{j=1}^{n} f(A_j(X)) \right) = \int dX\, S(X)\,. \tag{4.9}$$

The latter integral can be computed by means of MC as

$$\int dX\, f(X) \simeq \frac{1}{n+1} \sum_{i=1}^{M} \left( f(X_i) + \sum_{j=1}^{n} f(a_{j\,i}) \right)\,. \tag{4.10}$$

This manipulation is convenient only if the variance of $S$ is smaller than the variance of $f$, and the computational time to compute $S$ does not exceed the quantity

$$(\text{time to compute } f) \times \left( \frac{\text{variance of } f}{\text{variance of } S} \right)\,. \tag{4.11}$$

If all these conditions are satisfied, the points $\{a_1, \dots a_n\}_i$ are properly called antithetic variates. The determination of the functions $A_j$ which allow for an effective improvement is the difficulty that one has to face in order to exploit this method.

It is worth mentioning that the improvement of this method arises from the negative correlation (also known as anticorrelation) between the random points $X$ and the associated antithetic variates. A correlation is said to be negative if it relates values higher than average to values lower than average.

In the following we present several methods based on the underlying symmetry structure of the FSWF that accelerates the convergence through the introduction of appropriate antithetic variates.

### 4.2.1 Permutations

At first glance, spin-like shadow permutations appear to be the best approach to obtain an effective antithetic contributions for the FSWF. In particular, applying a pair permutation, the kernel is the only term subject to a variation in its absolute value, whereas $J_\mathrm{s}$ remains unaltered and $\mathrm{SD}_\mathrm{pw}$ changes its sign. Thus, antithetic variates based on permutations of $S$ permit for the estimation of the integral

$$\int dS\, \mathrm{SD}_\mathrm{pw}(S)\, J_\mathrm{s}(S) = 0 \tag{4.12}$$

within zero variance. However, when the kernel is introduced in the integral, it breaks the symmetry and hence the variance assumes a non-zero value. Nevertheless, when $C$ is small, the kernel variation induced by a permutation is minor, and therefore it is reasonable to expect the variance to be reduced anyway.

We now present our novel methods based on the shadow permutation symmetry.

#### Gaussian determinant

We will now show that it is possible to directly sum over all the possible permutation by introducing a determinant consisting of Gaussians. The sum over all pair permutations $\mathcal{P}_{ij}^{(2)}$ can be embedded into the FSWF as follows:

$$J_p(R) \int dS(\Xi(R,S) - \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \Xi(R, \mathcal{P}_{ij}^{(2)} S))\mathrm{SD}_\mathrm{pw}(S) J_s(S). \tag{4.13}$$

Introducing the summation over all three-body permutations leads to

$$J_p(R) \int dS(\Xi(R,S) + \sum_{i=1}^{N-2} \sum_{j=i+1}^{N-1} \sum_{k=j+1}^{N} \Xi(R, \mathcal{P}_{ijk}^{(3)} S))\mathrm{SD}_\mathrm{pw}(S) J_s(S). \tag{4.14}$$

In general, permutations of an even number of shadows imply a subtraction, whereas odd numbers imply an addition. Hence, the sum over all possible permutations can be elegantly summarized by a matrix determinant

$$J_p(R) \int dS \det(e^{-C(r_\alpha - s_\beta)^2})\, \mathrm{SD}_\mathrm{pw}(S)\, J_s(S)\,, \tag{4.15}$$

where $\alpha$ and $\beta$ label the matrix rows and columns, respectively. This expression is particularly convenient because it permits the summation of $N!$ terms within a computational cost of $N^3$. We name this term *Gaussian determinant* $\mathrm{GD}(R,S)$.

Accordingly, our first approach consists of replacing $\Xi(R,S)$ with $\mathrm{GD}(R,S)$. Figure 4.2 illustrates the results obtainable by adopting this new kernel form, demonstrating a considerable improvement.

As expected, the GD sensibly reduces the variance when small $C$ are employed. Results suggest an exponential growth of the variance magnitude as $C$ decreases, and the introduction of the GD seems to be able to diminish the exponential factor. Nevertheless, a reduction of the exponential form into polynomial is out of reach.

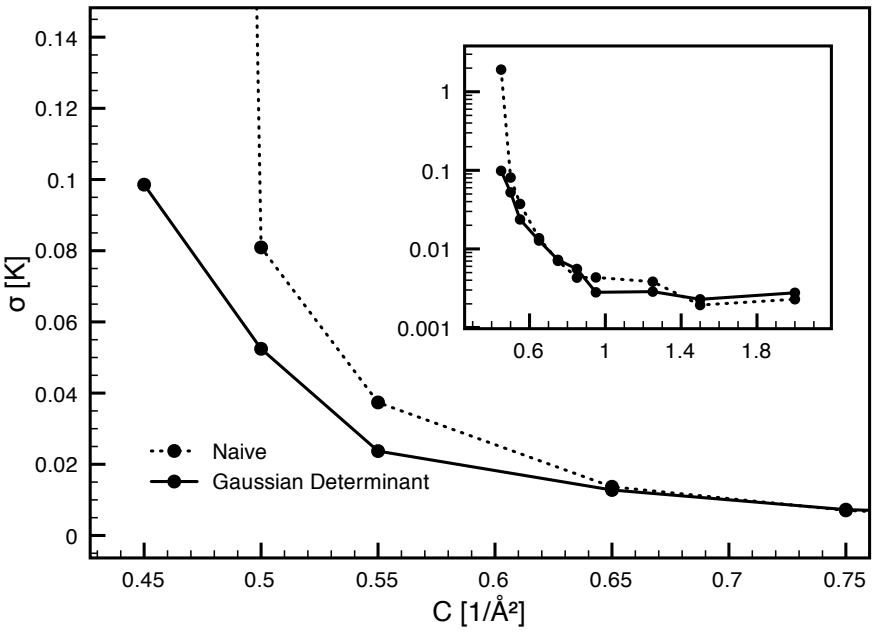


Figure 4.2: Estimated error for different values of $C$, with and without the adoption of the Gaussian determinant. Results refers to $N = 38$ and $M = 64{\cdot}10^7$. We regarded the data for $C < 0.45$ Å$^{-2}$ as unreliable, hence they are not reported. In the inset the $x$-axis range is extended and the $y$ axis is in logarithmic scale, in order to permit a view on all the data points.

An improvable aspect of the Gaussian determinant method is the sampling function of choice. In fact, by direct inspection, we have noticed that in the summation over all the permutations only one term is dominant, whereas all the others are nearly null. Thus, the effect of the antithetic variates is not as effective as it would be desirable, and a sampling that promotes a more effective cancellation is to be pursued. In the next subsection we present a technique to accomplish this outcome.

**Explicit pair permutation term: Duet and Quartet**

Antithetic contributions can be achieved by means of one simple pair permutation $\mathcal{P}_{ab}^{(2)}$

$$\begin{aligned} \Psi_{\mathrm{FSWF}}(R) &= \int dS \left(\Gamma(R,S) + \Gamma(R,\mathcal{P}_{ab}^{(2)}S)\right) \\ &= J_p(R) \int dS\, J_s(S)\, \mathrm{SD}_{\mathrm{pw}}(S) \left(\Xi(R,S) - \Xi(R,\mathcal{P}_{ab}^{(2)}S)\right) . \end{aligned} \tag{4.16}$$

This expression has the advantage over the Gaussian determinant to allow for a more balanced sampling from the original term and the permuted one, by means of

$$\rho_{\mathrm{MC}}(R,S_1,S_2) = \sqrt{\Gamma(R,S_1)\Gamma(R,\mathcal{P}_{ab}^{(2)}S_1)\Gamma(R,S_2)\Gamma(R,\mathcal{P}_{ab}^{(2)}S_2)}. \tag{4.17}$$

In fact, the sampling function in Equation (4.17) grants equal importance to the original configuration and the permuted one, so that eventually their contributions is of the same order and, therefore, there is a more effective cancellation.

The permuted term can be added to only $\Gamma(R,S_1)$ or to both $\Gamma(R,S_1)$ and $\Gamma(R,S_2)$. We will then refer to them as the *Duet* and *Quartet* techniques, due to the two and four terms in the expression for the energy. We have considered $a = 1$ and $b = 2$, since there is no reason to expect that different choices would lead to better or worse results.

We performed a calculation with 14 particles and $M = 16{\cdot}10^6$, employing the Duet technique. Fluctuations were such that we were unable to estimate the error bar. This discouraging result suggests that the presented sampling function is not optimal for integrating either $\Gamma(R,S)$ or $\Gamma(R,\mathcal{P}_{ab}S)$. Thence the gain in efficiency obtainable from the antithetic contributions is not visible because of the improper sampling function. As expected, the Quartet technique provides even lower efficiency than the Duet.

Even though this method does not provide any improvement compared to the GD, its failure grants a significant insight: simple spin-like shadow permutation do not provide effective antithetic variates. In fact, if this would not be the case, the product $\Gamma(R,S)\Gamma(R,\mathcal{P}_{ab}S)$ would possess a maximum nearly in the same regions where $\Gamma$ itself does, and the occurred sampling problem would not take place.

Therefore, other geometrical transformations than permutations have to be considered.

**Permutation as proposed move**

As it emerged in the previous subsection, even though a permutation implies a sign change, it is far from giving an optimal antithetic contribution. We can deduce that after performing a permutation, a certain translation would be necessary in order to effectively obtain an antithetic contribution. Since this translation is unknown, a stratagem should be devised in order to determine it. An interesting possibility is to let the shadows diffuse after the permutation towards the correct antithetic positions. In order to put this idea into practice, permutations can be used as proposed moves in the context of the $\mathrm{M(RT)}^2$ algorithm. In this respect, it is of extreme importance to take all the possible pair permutations into account, and consider the most favorable one in order to increase the acceptance rate. In the following we explain the details of this procedure.

First of all, $\Xi(R, \mathcal{P}_{ij}^{(2)} S)$ is evaluated for all the possible $i$ and $j$. Then, a permutation $(a, b)$ is proposed with a transition probability

$$T(S \to \mathcal{P}_{ab}^{(2)} S) = \frac{\Xi(R, \mathcal{P}_{ab}^{(2)})}{\sum_{(i,j)} \Xi(R, \mathcal{P}_{ij}^{(2)})} \,, \tag{4.18}$$

and accepted according to

$$A(S \to \mathcal{P}_{ab}^{(2)} S) = \frac{|\Gamma(R, \mathcal{P}_{ab}^{(2)} S)| T(\mathcal{P}_{ab}^{(2)} S \to S)}{|\Gamma(R, S)| T(S \to \mathcal{P}_{ab}^{(2)} S)} \,. \tag{4.19}$$

Each permutation move was attempted right after each estimation of the local energy, to allow the walkers to subsequently diffuse before the next $E_{\mathrm{loc}}$ estimate.

Following the afore-described procedure, we obtained $E = -1.997(11)$ K, using 14 particles and $M = 16{\cdot}10^6$, to be compared to $E = -1.986(18)$ K, of the naive algorithm. The acceptance rate for the permutation moves was roughly 2.5%.

Results demonstrate that using permutations as moves is a viable possibility, considered the significant acceptance rate. Moreover, they seem to be able to reduce the variance, even though more extensive calculations are necessary in order to assess it (in section 4.4 all the methods are accurately compared).

### 4.2.2 Reflections

The Gaussian kernel localizes the integrand around $R$. A small $C$ implies a significant spread of $\Gamma$, and a consequent frequent cross by the shadow walkers of the nodal surface defined by $\mathrm{SD}_{\mathrm{pw}}(S)$, causing the sign problem. This consideration leads naturally to the introduction of *reflected shadows* $S' = 2R - S$, that are expected to be located in a different nodal pocket than the one of the original shadows $S$. Figure 4.3 efficaciously motivates this conjecture.

In order to take advantage of this idea, we sampled from

$$\rho(R, S_1, S_2) = |\Gamma(R, S_1)\Gamma(R, S_2)| \tag{4.20}$$

as usual, but summed the contribution originating from the reflected shadows

in the energy estimator

$$E = \frac{\frac{1}{2}\sum_{i=1}^{M}\left\{\left[\frac{\big(\Gamma(R,S_1)+\Gamma(R,S_1')\big)\mathrm{H}\big(\Gamma(R,S_2)+\Gamma(R,S_2')\big)}{|\Gamma(R,S_1)\Gamma(R,S_2)|}\right]\right\}_i}{\sum_{i=1}^{M}\left\{\frac{\big(\Gamma(R,S_1)+\Gamma(R,S_1')\big)\big(\Gamma(R,S_2)+\Gamma(R,S_2')\big)}{|\Gamma(R,S_1)\Gamma(R,S_2)|}\right\}_i}. \tag{4.21}$$

Unfortunately, during the calculation $\Gamma(R,S')$ was typically so much smaller than $\Gamma(R,S)$ that its contribution was negligible. In fact, using 14 particles, the estimated energy and error were modified by a factor smaller than $10^{-4}$ by the introduction of the reflected shadows. The further adoption of the Gaussian determinant did not modify the result. Therefore, our initial conjecture to build antithetic contributions employing reflected shadows should be reconsidered

Finally, we remark that the method presented here suffers from the infinite variance problem, which should be rectified by removing the zero values from the sampling function. Since the contribution given by the reflected shadows is negligible anyway, we did not deal with this problem.

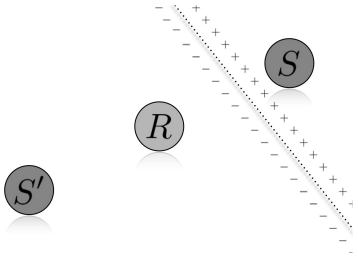


Figure 4.3: Illustration of a reflected shadow. $R$ is close to the nodal surface defined by the SD, thence $S$ and $S'$ are in different pockets.

### 4.2.3 Constrained Domain

Ultimately, it is possible to use symmetry arguments to constrain the domain of the integration.

First of all, because of the antisymmetric nature of the FSWF, $R$ can be constrained to the positive (or negative) domain of a $\mathrm{SD}(R)$. No particular form is required for the SD, it is correct independently of this choice. In our tests we employed the $\mathrm{SD}_{\mathrm{pw}}$.

A second argument is valid only in conjunction with the Gaussian determinant. In fact, when the sum over all the $S$ permutations is accomplished, $S_1$ and $S_2$ can be restricted onto the positive (or negative) domain, i.e. where $\mathrm{SD}(S_1)$ and $\mathrm{SD}(S_2)$ are positive (or negative). We stress that these restrictions do not imply that $\Gamma$ must always be positive (or negative), as the Gaussian determinant is not positive (negative) definite.

These two constraints significantly reduce the integration space.

The simulation of 38 atoms adopting the afore-described constraints provided $E = -2.5(10)$ K for $M = 60{\cdot}10^6$, whereas without any restriction the energy estimator was $E = -1.9(1)$ K. No reduction in the sign fluctuations was

hence achieved by constraining the domain. Nevertheless, the fact that the error increased by a factor of 10 is surprising, but can be explained by the fact that moves rejections due to a restrained domain reduce the efficiency of integration along the constrained domain borders.

For completeness, we report that 5% and 0.5% of the moves for $R$ and $S$, respectively, were rejected because of the constraints.

## 4.3 Grouping Technique and Marginal Distribution

We now introduce a completely different approach, devised by analyzing the expression for the energy with a special emphasis on the integrals over $S_1$ and $S_2$

$$E = \frac{\int dR \, \left(\int dS_1 \, \Gamma(R, S_1)\right) \mathrm{H} \left(\int dS_2 \, \Gamma(R, S_2)\right)}{\int dR \, \underbrace{\left(\int dS_1 \, \Gamma(R, S_1)\right)}_{\Omega_1(R)} \underbrace{\left(\int dS_2 \, \Gamma(R, S_2)\right)}_{\Omega_2(R)}} \, .$$

By definition, $\Omega_1(R) = \Omega_2(R) = \Psi_{\mathrm{FSWF}}(R)$. The knowledge of $\Psi_{\mathrm{FSWF}}(R)$, i.e. of the analytical solution of the integral over $S$, would allow for the solution of the sign problem, as $\Omega_1(R)\Omega_2(R) = \Psi^2_{\mathrm{FSWF}}(R) \geq 0$. The fact that $\Psi_{\mathrm{FSWF}}(R)$ is not known has thence two major consequences:

1. $\Omega_1(R)$ and $\Omega_2(R)$ are approximated by the sampled $\Gamma(R, S_1)$ and $\Gamma(R, S_2)$, which happen to have different signs. As a consequence the estimated local energy is weighted by the extremely noisy function

$$w(R, S_1, S_2) = \Gamma(R, S_1)\Gamma(R, S_2) = \pm 1 \, . \tag{4.22}$$

   This point has been already remarked in the introduction and has been addressed, for example, by introducing antithetic variates;

2. $R$ is not sampled efficiently from its marginal distribution $\Psi^2_{\mathrm{FSWF}}(R)$. To understand this more subtle point, we will make use of an example. Suppose that $R_0$ is on the nodal surface, i.e. $\Psi_{\mathrm{FSWF}}(R_0) = 0$. Thus, this configuration should never be sampled. Nevertheless, in reality it will be sampled with probability $\Gamma(R_0, S_1)\Gamma(R_0, S_2)$, which is fundamentally different from $\Omega_1(R_0)\Omega_2(R_0) = \Psi^2_{\mathrm{FSWF}}(R)$. This is because the integrand can be vastly different from the integral. As additional drawback, non-properly sampled $R$ further increase the weight noise. In fact, when $R$ is far from the maximum values of $\Psi_{\mathrm{FSWF}}$, it is to be expected that $\Gamma(R_0, S_1)$ and $\Gamma(R_0, S_2)$ carry a different sign, since the average is nearly zero. On the contrary, when $\Psi_{\mathrm{FSWF}}(R) \gg 1$, the average is sensibly different from zero, and thus their fluctuations are biased to have the same sign;

The first problem has been addressed in the past with the so-called *Blocking Technique* (BT), and used to study the vacancy formation energy in solid $^3$He [9]. The main idea is to sample many successive shadows in order to obtain a decent estimate of $\Omega_1(R)$ and $\Omega_2(R)$, possibly with the same sign. The adopted algorithm then reads:

| $\mathbf{M_s}$ | **Efficiency** $[1/(\mathbf{sec\ K^2})]$ |
|---|---|
| 1 | 0.63 |
| 10 | 0.36 |
| 100 | 0.36 |
| 1000 | 0.23 |

Table 4.4: Efficiency of the Blocking Technique for different blocks sizes, simulating $N = 14$ atoms.

1. Start from a configuration $R_{(0)}$, $S_{1(0,1)}$ and $S_{2(0,1)}$. $i := 1$;
2. Sample $R_{(i)}$ from
$$\rho(R) = |\Gamma(R, S_{1(i-1,1)})\Gamma(R, S_{2(i-1,1)})| \ ;$$
3. Sample $M_s$ points $\left(S_{1(i,1)}, \dots S_{1(i,M_s)}\right)$ from
$$\rho(S_1) = |\Gamma(R_{(i)}, S_1)\Gamma(R_{(i)}, S_{2(0,1)})|$$
and then $\left(S_{2(i,1)}, \dots S_{2(i,M_s)}\right)$ from
$$\rho(S_2) = |\Gamma(R_{(i)}, S_{1(i,M_s)})\Gamma(R_{(i)}, S_2)| \ ;$$
4. Evaluate $\Omega_1(R_{(i)}) = \sum_{j=1}^{M_s} \frac{\Gamma(R_{(i)}, S_{1(i,j)})}{|\Gamma(R_{(i)}, S_{1(i,j)})|}$ and $\Omega_2(R_i) = \sum_{j=1}^{M_s} \frac{\Gamma(R_i, S_{2(i,j)})}{|\Gamma(R_i, S_{2(i,j)})|}$;
5. Evaluate $\mathrm{H}_1^{\mathrm{loc}}(R_i) = \sum_{j=1}^{M_s} \frac{\mathrm{H}\Gamma(R_{(i)}, S_{1(i,j)})}{|\Gamma(R_{(i)}, S_{1(i,j)})|}$ and $\mathrm{H}_2^{\mathrm{loc}}(R_i) = \sum_{j=1}^{M_s} \frac{\mathrm{H}\Gamma(R_{(i)}, S_{2(i,j)})}{|\Gamma(R_{(i)}, S_{2(i,j)})|}$
6. $i := i + 1$;
7. Repeat $M$ times from 2 to 6;
8. Estimate the total energy
$$E = \frac{\frac{1}{2}\sum_{i=1}^{M}\left(\Omega_1(R_i)\mathrm{H}_2^{\mathrm{loc}}(R_i) + \Omega_2(R_i)\mathrm{H}_1^{\mathrm{loc}}(R_i)\right)}{\sum_{i=1}^{M}\Omega_1(R_i)\Omega_2(R_i)} \ . \tag{4.23}$$

Even though the grouping technique stabilizes the sign of $\Omega_1(R)\,\Omega_2(R)$, the required computational cost is not adequately compensated by the variance reduction. This is demonstrated by Table 4.4, where we report our results for the efficiency for different $M_s$. Although this method does not improve the efficiency for the liquid phase, it could for the solid state, because the sign problem is less severe and a small $M_s$ is sufficient to stabilize the sign.

In the next subsections we address the second problem presented at the beginning of this section, by introducing and testing two novel methods. The first one uses an analytical approximation of the marginal distribution. The second one employes a numerical estimation of it.

### 4.3.1 JS Approximation

As first approach to the aforementioned problem, $R$ can be sampled from the best available approximation of $\Psi_{\mathrm{FSWF}}(R)$. More specifically, $R$ can be sampled with a JS trial wave function, e.g.

$$\rho_{\mathrm{MC}}(R) = \left|J(R)\mathrm{SD}_{\mathrm{pw}}(R)\right|^2 , \tag{4.24}$$

whereas $S_1$ and $S_2$ can be sampled from the usual expression

$$\rho_{\mathrm{MC}}(S_1, S_2) = \left|\Gamma(R, S_1)\Gamma(R, S_2)\right| . \tag{4.25}$$

However, sampling $R$ and the shadows from different probability density functions (pdf) introduce the necessity of relaxing the latter ones after each $R$ displacement. To avoid this complication, it is convenient to employ a common sampling function. More specifically, in the algorithm presented for the blocking technique, the $\Gamma$ function within the sampling distribution can be replaced with opportune variations of it. Naturally, the modification of the sampling function causes the introduction of opportune weights in the energy estimator.

We have proposed and tested the following modification of $\Gamma$:

- $\Gamma'_1(R, S) = \Gamma(R, S)\left(\mathrm{SD}_{\mathrm{pw}}(R)^2 + \Lambda^2(R)\right)^{1/4}$;
- $\Gamma'_2(R, S) = \Gamma(R, S)\left(\mathrm{SD}_{\mathrm{pw}}(R)^2 + \Lambda^2(R)\right)^{1/2}$;
- $\Gamma'_3(R, S) = \Gamma(R, S)\left(\mathrm{SD}_{\mathrm{pw}}(R)^2 + \Lambda^2(R)\right)$;
- $\Gamma'_4(R, S) = \Gamma(R, S)\left(\mathrm{SD}_{\mathrm{pw}}(R)^4 + \Lambda^4(R)\right)^{1/4}$.

$\Lambda$ has been introduced to avoid the infinite variance problem, and its optimal value is expected to be of the same order of magnitude of the mean value of $\mathrm{SD}(R)$.

From Figure 4.4, we can conclude that by introducing a $\mathrm{SD}(R)$ in the sampling function it is possible to reduce the variance by up to a factor 3/2. Among the various possibilities, $\Gamma'_4$ shows the most convincing improvement.

The achieved improvement proves that normally $R$ is not efficiently sampled from its marginal distribution, confirming our conjecture. In particular, it demonstrates that the antisymmetric component plays an essential role in the marginal distribution, but is normally not well reproduced by $\Gamma$.

However, the JS approximation method is cumbersome and not very practical, since many preliminary calculations are required for the determination of the optimal $\Lambda$. It is anyhow a useful instrument to asses the importance of the marginal distribution of $R$.

### 4.3.2 $S$-averaged Marginal Distribution

A second possibility is a numerical estimation of $\Omega_1$ and $\Omega_2$ within the sampling function. In order to accomplish this, one can introduce $2N_S$ shadows $S_{1,1}$, $S_{1,2}, \ldots S_{1,N_S}$, $S_{2,1}$, $S_{2,2}, \ldots S_{2,N_S}$, and replace the expression

$$\int dS_1\, \Gamma(R, S_1) \qquad \text{by} \qquad \sum_{i=1}^{N_S} \int dS_{1,i}\, \Gamma(R, S_{1,i}) , \tag{4.26}$$

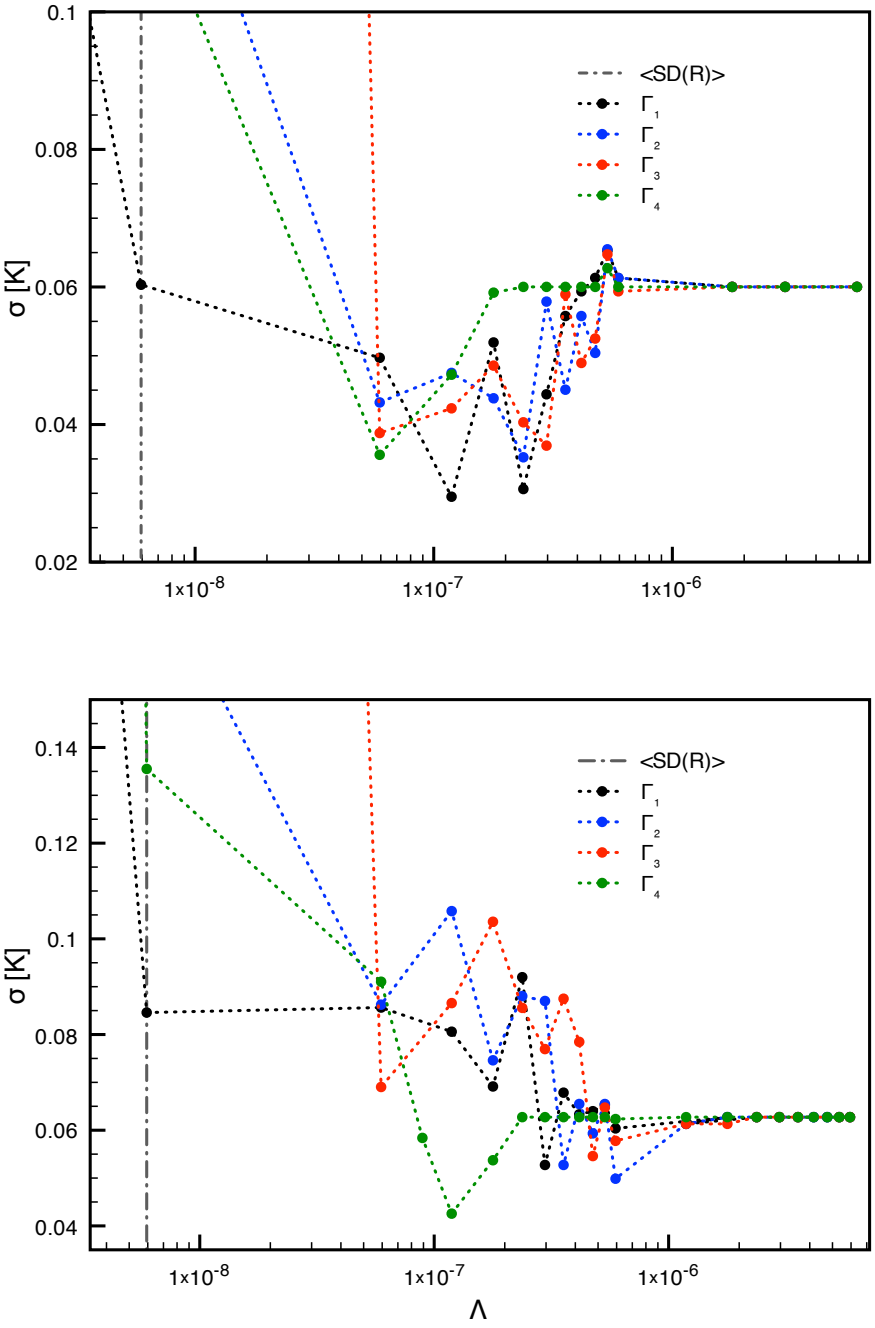


Figure 4.4: Estimated $\sigma$ obtained with the four different $\Gamma'(R,S)$ introduced in the text, for various values of the parameter $\Lambda$. In the upper graph, the Blocking Technique with $M_s = 100$ and $M = 96 \cdot 10^4$ was used.In the lower graph, the Blocking Technique was not employed, and $M = 48 \cdot 10^6$. The average value $\langle \mathrm{SD}(R) \rangle$ was evaluated by sampling from $\Gamma(R,S_1)\Gamma(R,S_2)$. In both cases we adopted the Gaussian determinant kernel form, and we simulated 38 atoms.

| $N_S$ | Error |
|---|---|
| 1 | 0.024 |
| 2 | 0.025 |
| 4 | 0.034 |
| 10 | 0.029 |

Table 4.5: Estimated $\sigma$ for the energy obtained with the $S$-averaged marginal distribution method, using different $N_S$. Results refer to simulations of 14 $^3$He atoms with $M = 8 \cdot 10^6$.

$$\text{and} \qquad \int dS_2\, \Gamma(R, S_2) \qquad \text{by} \qquad \sum_{i=1}^{N_S} \int dS_{2,i}\, \Gamma(R, S_{2,i})\,. \tag{4.27}$$

Thence, the $S$-averaged sampling function reads

$$\rho(R, S_{1,1}, \dots S_{2,1}, \dots) = \left| \left( \sum_{i=1}^{N_S} \Gamma(R, S_{1,i}) \right) \left( \sum_{i=1}^{N_S} \Gamma(R, S_{2,i}) \right) \right| . \tag{4.28}$$

In this way $R$ is sampled from an improved approximation of $\Omega_1(R)\,\Omega_2(R)$.

In Table 4.5 we report the results for the *$S$-averaged marginal distribution*. The introduction of additional shadows does not seem to reduce the variance. In fact, the small differences can be attributed to statistical fluctuations.

Figure 4.5 clarifies why the $S$-averaged marginal distribution is not effective: Along the simulation, only one of the $S_{1,i}$ and one of the $S_{2,i}$ is relevant and all the others tend to zero. Thence, the algorithm returns to the original form. This behavior is similar to what we have already encountered in the context of the Gaussian determinant (subsection 4.2.1), where only one permuted term is effectively sampled. It appears that, more generally, effective sampling from a sum of functions cannot be easily accomplished. We therefore expect this tendency not to be preventable. In conclusion, the whole algorithm requires a fundamental revision.

## 4.4 Comparison of the Methods

Table 4.6 summarizes the efficiency obtainable by the novel methods that we have presented along this chapter.

The JS approximation method competes with the original algorithm, but it does not provide a significant improvement.

Surprisingly, the Gaussian determinant method does not seem to guarantee a melioration, even though we have already shown that it reduces the variance. Evidently, the time required to evaluate the Gaussian determinants is not adequately compensated by the error reduction. Nevertheless, it is expected that for smaller values of $C$ this method proves to be valuable.

The Gaussian determinant and JS approximation methods seem to have a certain potential, and therefore we decided to review the accuracy of their efficiency estimations. In particular, we wanted to rule out the possibility that ergodicity problems affect our results. To exclude this eventuality, we performed

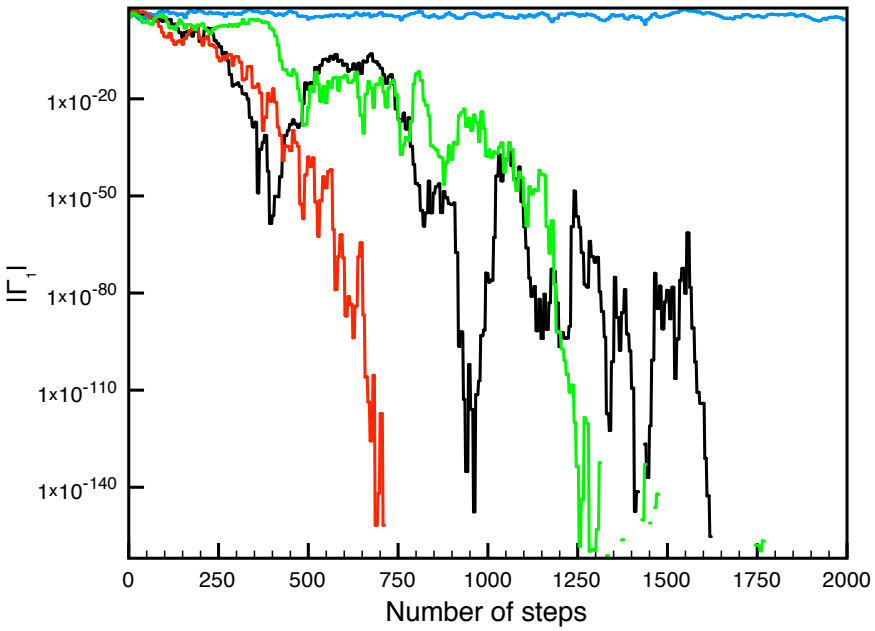


Figure 4.5: Trend of the various $\Gamma$ associated to $S_{1,1}, \ldots S_{1,4}$, during the progression of a simulation. The lines appear to be interrupted because of the double precision float limitations. The values of $\Gamma$ are so small that they are numerically equal to zero and thus no longer visualizable on a logarithmic scale.

| **Used technique** | **Energy** [K] | **Efficiency** $[10^{-2} \cdot \mathrm{sec}^{-1}\mathrm{K}^{-2}]$ |
|---|---|---|
| Naive | $-1.949 \pm 0.016$ | $30 \pm 3$ |
| Gaussian Determinant | $-1.943 \pm 0.014$ | $22 \pm 3$ |
| Duet | $2.7 \pm 4.3$ | $\approx 0$ |
| Quartet | $-2.2 \pm 1.6$ | $\approx 0$ |
| Permutation moves | $-2.012 \pm 0.016$ | $25 \pm 3$ |
| Reflections$^{\dagger}$ | $-1.945 \pm 0.016$ | $15 \pm 2$ |
| *R-S* Domain Constraint$^{\dagger}$ | $-1.974 \pm 0.028$ | $\approx 5$ |
| *R* Domain Constraint$^{\dagger}$ | $-1.975 \pm 0.017$ | $24 \pm 3$ |
| JS Approximation | $-1.953 \pm 0.015$ | $31 \pm 3$ |
| JS Approximation$^{\dagger}$ | $-1.949 \pm 0.017$ | $35 \pm 7$ |

Table 4.6: Average energy, associated error, and efficiency of all the methods presented along this chapter. The results have been obtained by averaging eight independent simulations of $N = 38$ atoms with $M = 12 \cdot 10^8$. If the estimated error of the efficiency is not reported, only one run was performed. For the J-SD approximation method we set $M_s = 1$. The symbol $^{\dagger}$ means that the technique was used in conjunction with the Gaussian determinant. The efficiencies of the Duet and Quartet methods were extremely low (smaller than $10^{-3}$) and therefore approximated with zero. *R-S* Domain Constraint refers to the algorithm described in the subsubsection Constrained Domain (see section 4.2), whereas for the *R* Domain Constraint method we limited the constraint only to *R*.

several calculations for different $M$. We fitted the resulting energy errors with the function $f(x) = A/\sqrt{M}$, which is the expected asymptotic behavior, as exemplified in Figure 4.6. If the number of steps would have not been sufficient to ensure the ergodicity, the chi-squared test would have failed. Therefore this procedure was useful both to ensure the reliability of our results and to have a more accurate estimate of the efficiency. In fact, the squared fitted parameter $A$ can be used in place of the variance for the estimation of the efficiency.

Finally, we studied the efficiency dependence on $N$ by computing $\eta$ for 14, 38 and 54 atoms. The results are reported in Table 4.7 and Figure 4.7, which clearly show that the efficiency decays exponentially with the number of particles, and that none of the presented methods alleviates this tendency.

This is the first ever reported evidence supporting the conjecture that the FSWF sign problem grows exponentially with the number of simulated particles. In the next section we present a theoretical derivation that justifies this numerical result.

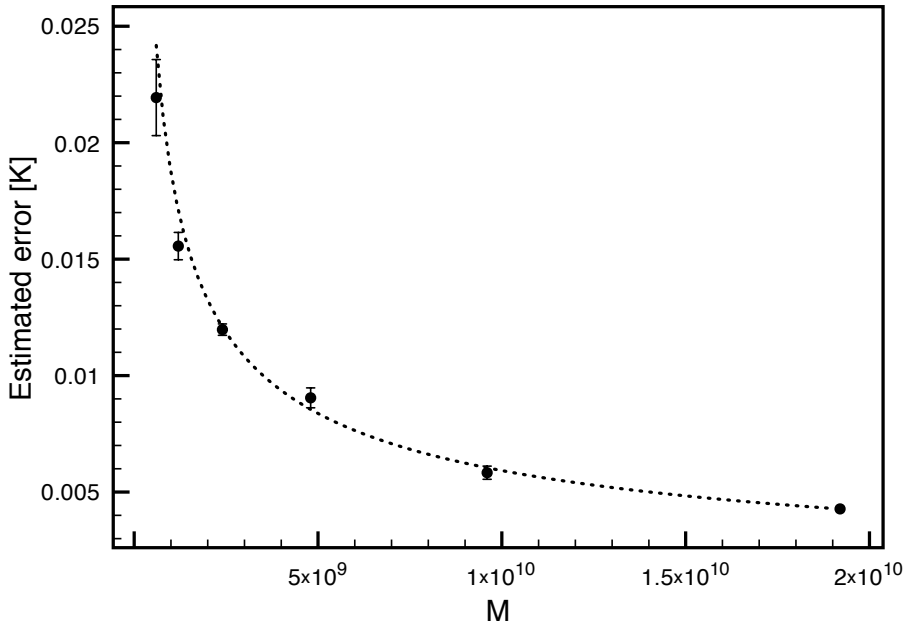


Figure 4.6: Fitting procedure used to accurately estimate the efficiency. The displayed data are obtained with the naive method, simulating 38 atoms and averaging over eight independent simulations. The dotted line represents the fitted function $f(x) = A/\sqrt{M}$. The chi-squared test is successfully fulfilled.

## 4.5 On the Nature of the FSWF Sign Problem

In this section we derive an the analytical expression for the exponential decay of the efficiency that arises from the employment of the FSWF.

This derivation happened incidentally, while exploiting a new method to

| Technique | N=16 | N=38 | N=54 |
|---|---|---|---|
| Naive | $1375 \pm 50$ | $3.03 \pm 0.06$ | $0.0210 \pm 0.0005$ |
| Gaussian Determinant | $1220 \pm 30$ | $2.54 \pm 0.14$ | $0.0216 \pm 0.0032$ |
| J-SD approximation* | - | $3.23 \pm 0.22$ | - |

Table 4.7: Accurate efficiency estimates of the most interesting techniques for different number of atoms $N$. We decided not to evaluate the efficiency of the "J-SD approximation*" method for different sizes because there are no reasons to expect significantly different results.

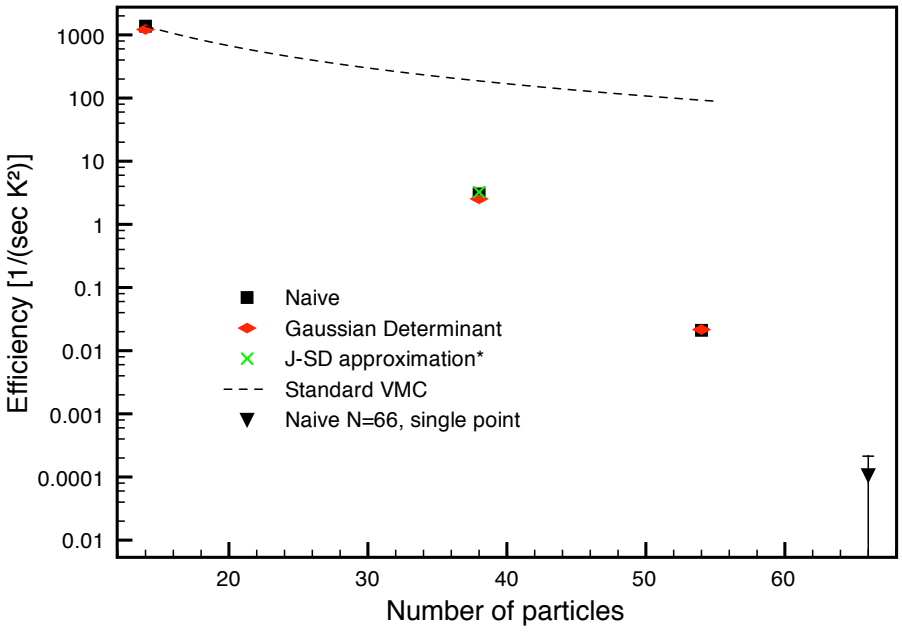


Figure 4.7: Graphic representation of the data reported in Table 4.7. The dotted lines represent the expected behavior for fermionic VMC calculations without the sign problem. The triangle represents the efficiency related to the results reported in Figure 4.1.

reduce the sign problem, and we now want to present it accordingly. We first introduce this attempt, and thereafter we point out its inability to provide substantial improvements. In the end we show that, within this framework, it is possible to account for the efficiency dependence on $N$ and $C$.

A general method to reduce the variance of MC calculations is the *control variates* technique (see [108] and [24] pag. 91). The basic principle is that if the analytical integral of a function $h(x)$ is known, and $h(x)$ is similar to the integrand of interest $f(x)$, it is convenient to subtract and add $\int dx\, h(x)$,

$$\int dx\, f(x) = \int dx\ (f(x) - h(x)) + \int dx\, h(x)\,. \tag{4.29}$$

This procedure results to be convenient only if the condition

$$\mathrm{var}\,[f(x) - h(x)] < \mathrm{var}\,[f(x)] \tag{4.30}$$

is satisfied. However, this additional term should not exceedingly increase the computational time.

In the case of the FSWF, the integral of a non-interacting $\Gamma_0$ is known analytically:

$$\begin{aligned}
\int dS\, \Gamma_0(R,S) &\equiv \int dS\, J_{\mathrm{p}}(R)\, e^{-C(R-S)^2} \det(e^{i\mathbf{k}_\alpha\cdot\mathbf{s}_\beta}) && (4.31)\\
&= J_{\mathrm{p}}(R) \left(\frac{\pi}{C}\right)^{3N/2} e^{-\frac{\sum_{i=1}^{N}\mathbf{k}_i^2}{4C}} \det(e^{i\mathbf{k}_\alpha\cdot\mathbf{r}_\beta}) \times && (4.32)\\
&\quad \prod_{\substack{\gamma=x,y,z\\ i=1,\ldots N}} \left[\mathrm{erf}\left(\frac{CL - ik_{\gamma\, i}}{2\sqrt{C}}\right) - \mathrm{erf}\left(\frac{-CL - ik_{\gamma\, i}}{2\sqrt{C}}\right)\right]\\
&\equiv \Psi_{\mathrm{FSWF}\,0}(R)\,. && (4.33)
\end{aligned}$$

The product sequence containing the error function arises from the restriction of the integral domain on a $L \times L \times L$ box with periodic boundary conditions. The FSWF can be then recast as

$$\begin{aligned}
\Psi_{\mathrm{FSWF}}(R) &= \int dS\ (\Gamma(R,S) - \Gamma_0(R,S)) + \int dS\, \Gamma_0(R,S) && (4.34)\\
&= \int dS\ (\Gamma(R,S) - \Gamma_0(R,S)) + \Psi_{\mathrm{FSWF}\,0}(R)\,. && (4.35)
\end{aligned}$$

However, this formulation is not convenient in the context of the VMC energy evaluation, because of the different integral variables ($R$ and $S$ on the left, and only $R$ on the right). This difficulty can be surmounted introducing

$$\bar{\Gamma}_0(R,S) = \Psi_{\mathrm{FSWF}\,0}(R) \left(\frac{C}{\pi}\right)^{3N/2} \left[\frac{1}{\mathrm{erf}(\frac{\sqrt{C}L}{2})}\right]^{3N} e^{-C(R-S)^2}\,, \tag{4.36}$$

which fulfills the condition

$$\int dS\, \bar{\Gamma}_0(R,S) = \Psi_{\mathrm{FSWF}\,0}(R)\,. \tag{4.37}$$

Differently from $\Gamma_0$, $\bar{\Gamma}_0$ is symmetric. Finally, we express the FSWF in its new final form

$$\Psi_{\mathrm{FSWF}}(R) = \int dS \, \left(\Gamma(R,S) - \Gamma_0(R,S) + \bar{\Gamma}_0(R,S)\right) . \tag{4.38}$$

The expression in equation (4.38) has the potential to significantly reduce the severity of the sign problem, if $\Gamma - \Gamma_0 \ll 1$. This is the case when $J_{\mathrm{s}}$ does not have large fluctuation, for example in the gas phase, when two-body correlations are negligible and the sign problem is particularly strong. In the asymptotic limit $J_{\mathrm{s}} \to 1$, the sign problem is completely suppressed. However, the imposition $J_{\mathrm{s}} = 1$ prevents any sort of benefit from the introduction of the shadows. In the following we show that in practice the presented method cannot be effective.

As reported in Equation (4.32), $\bar{\Gamma}_0$ contains the factor $e^{-\frac{\sum_{i=1}^{N} \mathbf{k}_i^2}{4C}}$, which decays exponentially with $\frac{\sum_{i=1}^{N} \mathbf{k}_i^2}{C}$. It is possible to relate $\sum_{i=1}^{N} \mathbf{k}_i^2$ to the number of particles, assuming the thermodynamic limit, and using the identity $\sum_{\mathbf{k}} = \left(\frac{L}{2\pi}\right)^3 \int d\mathbf{k}$. In fact, by integration over the Fermi sphere $\mathcal{S}_{\mathrm{F}}$, the summation reads

$$\begin{aligned}
\sum_{i=1}^{N} \mathbf{k}_i^2 &= \left(\frac{L}{2\pi}\right)^3 \int_{\mathcal{S}_{\mathrm{F}}} d\mathbf{k}\, \mathbf{k}^2 && (4.39)\\
&= \left(\frac{L}{2\pi}\right)^3 4\pi \int_0^{k_{\mathrm{F}}} dk\, k^2 && (4.40)\\
&= \frac{(3\pi)^{5/3}}{10\pi^2}\, \rho^{2/3}\, N . && (4.41)
\end{aligned}$$

In other words, given a physical system with density $\rho$, both $\bar{\Gamma}_0$ and $\Psi_{\mathrm{FSWF}\,0}$ decay exponentially according to a $e^{-N/C}$ law. By generalization, we expect $\Psi_{\mathrm{FSWF}}$ to follow the same law.

In summary, if either the number of simulated particles increases or $C$ decreases, $\bar{\Gamma}_0$, $\Psi_{\mathrm{FSWF}\,0}$, and $\Psi_{\mathrm{FSWF}}$ become small, contrary to $\Gamma$ and $\Gamma_0$, which remain essentially unchanged.

Let us suppose that the energy of a $N_0$-particles system can be estimated within a variance $\mathrm{var}_0$ and associated efficiency $\eta_0$, assuming $C = C_0$. By changing the parameters

$$\begin{aligned}
N &= N_0 \kappa_N \qquad \text{and} && (4.42)\\
C &= C_0 \kappa_C , && (4.43)
\end{aligned}$$

$\Psi_{\mathrm{FSWF}\,0}(R)$ in equation (4.35) is modified by the factor $e^{-\kappa_N/\kappa_C}$. By direct comparison, the integral

$$\int dS \, \left(\Gamma(R,S) - \Gamma_0(R,S)\right) \tag{4.44}$$

should also be determined within a standard deviation $e^{-\kappa_N/\kappa_C}$ times $\mathrm{var}_0$, in order to maintain the efficiency $\eta_0$. It should be noted that, in general, a larger

$N$ requires longer computational time, but also decreases the variance by a factor $1/\kappa_N$, as long as one is interested in the energy per particle. Hence, including this correction, the factor $e^{-\kappa_N/\kappa_C}$ has to be changed into $e^{-\kappa_N/\kappa_C}/\sqrt{\kappa_N}$. We remind the reader that MC estimates an integral within an error bar that decays as $M^{-1/2}$, where $M$ is the number of sampled points. In order reduce the variance of the integral in equation (4.44) by the aforementioned factor,

$$M = M_0 e^{2\kappa_N/\kappa_C} \kappa_N \tag{4.45}$$

points have to be sampled. This can be straightforwardly related to the computational time, leading to the efficiency

$$\eta = \eta_0 e^{-2N/C} \kappa_N^2 \,. \tag{4.46}$$

Since the factor $\kappa_N^2$ is in general marginal compared to $e^{-2N}$, it is omitted in the following discussion.

The expression in equation (4.46) cannot be considered exact, since variations of $C$ modify the efficiency in a non-trivial way. In fact, the variance is minimized near the variational ground state, and deviations from the optimal variational parameters result into a variance augmentation. Unfortunately, to the best of our knowledge a theory that predicts such a behavior is not available. However, this effect is minimal when the variational parameters are slightly changed, and can be neglected (see again Table 4.2).

The following table summarizes the efficiency dependence on $N$ and $C$.

| **Method** | $\eta$ |
|---|---|
| Standard VMC | $\propto \left(\frac{1}{N}\right)^2$ |
| VMC with FSWF | $\propto e^{-N/C}$ |

Figures 4.8 and 4.9 demonstrate that our prediction conforms to the results of our calculations. In Figure 4.9 we have restricted the data around the optimal value of $C$ because, as already pointed out, our prediction is not reliable when $C$ is considerably different from the optimal value.

In summary, by means of the control variates technique it is possible to predict the gravity of the sign problem, and to have a better understanding about its nature. The most important inference is that the sign problem can be reduced to the one-dimensional integral

$$I = \int ds\, f(s) = \int ds\, e^{-C(x-s)^2}\, e^{iks} = \sqrt{\frac{\pi}{C}}\, e^{-\frac{k^2}{4C}}\, e^{ikx} \,. \tag{4.47}$$

$J_{\mathrm{s}}$ does not contribute to the genesis of the sign problem, as well as the antisymmetrical structure of the SD($S$). In fact, it can be easily shown that

$$\frac{\mathrm{var}(f)}{|I|^2} \propto e^{k^2/C}. \tag{4.48}$$

In conclusion, the provided $f$ contains all the difficulties encountered with the FSWF.

The solution to the sign problem is connected to the possibility of computing the integral in equation (4.47) at various $k$ and $C$ values with constant (or, at

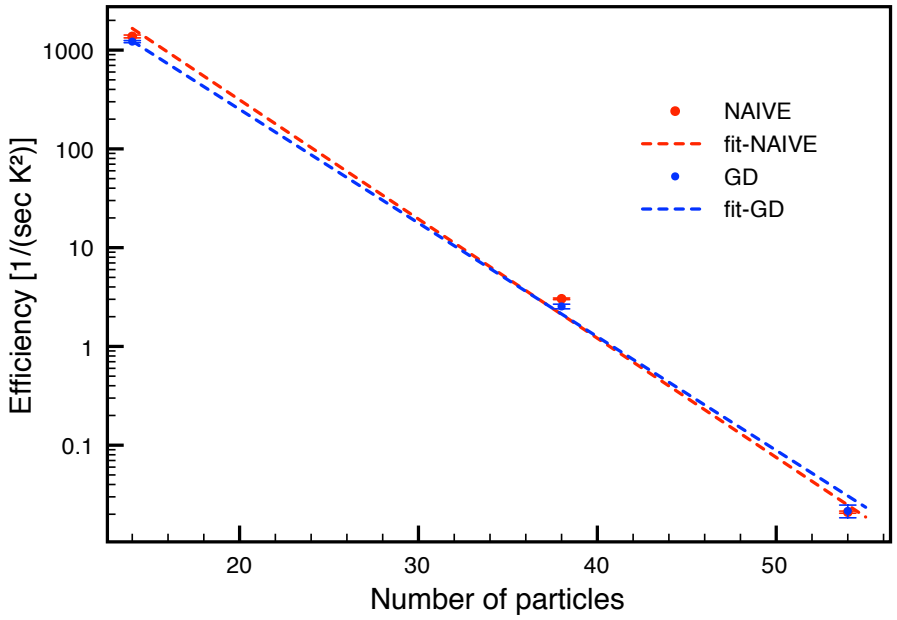


Figure 4.8: Fit of the data in Table 4.7 to our predicted law $\propto e^{-N}$.

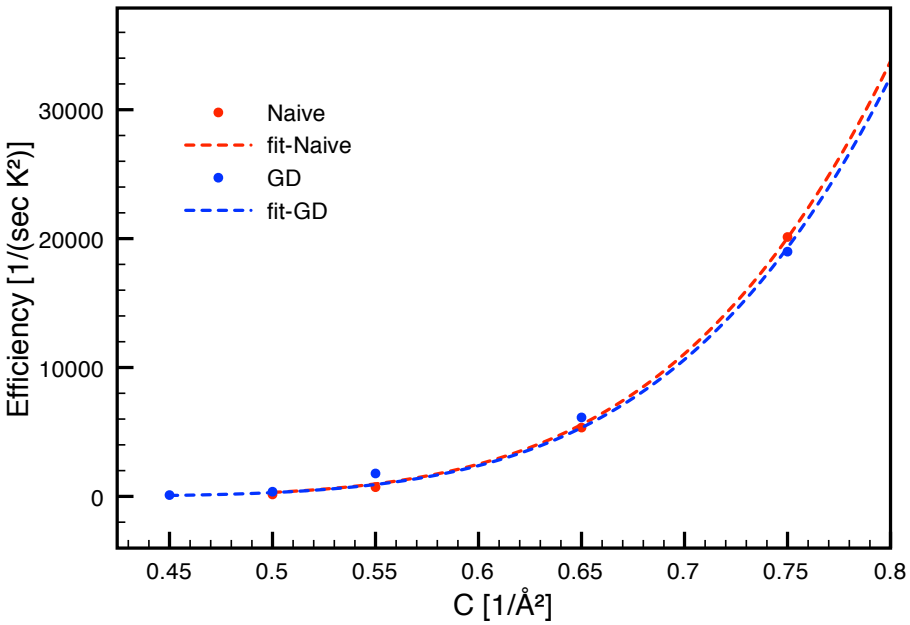


Figure 4.9: Fit of the data in Figure 4.2 to our predicted law $\propto e^{-1/C}$.

least, polynomially scaling) $\frac{\mathrm{var}}{|I|^2}$ ratio, with special focus on large $k$ and small $C$. Moreover, the procedure should be general enough to allow for its validity also when $J_\mathrm{s}$ is included.

The aforementioned conclusions exclude the possibility of any substantial progress in the sign problem by means of shadow permutations only. In fact, the sign problem would arise even replacing $\mathrm{SD}_\mathrm{pw}$ with $e^{i\mathbf{k}_1\cdot\mathbf{r}_1}e^{i\mathbf{k}_2\cdot\mathbf{r}_2}\cdots e^{i\mathbf{k}_{\frac{N}{2}}\cdot\mathbf{r}_{\frac{N}{2}}}$. Conversely, the approach outlined in section 4.3 has now a solid theoretical fundament.

To conclude, we emphasize once more on the importance of the considerations presented in this section. Understanding the origin of a problem is the first step towards its solution.

# Appendix A

# Energy Tables

## A.1 JS-pw

### A.1.1 atm-bcc hydrogen, $N = 128$

| $r_s$ | $E$ | $\sigma$ |
|---|---|---|
| 1.251 | −0.95686892108598887 | $8.9667180266457557E-005$ |
| 1.31 | −0.98283688767866662 | $9.7341274190646820E-005$ |
| 1.42 | −1.0125555272340536 | $6.2367921078935909E-005$ |
| 1.57 | −1.0302678186446712 | $6.0586387017204362E-005$ |
| 1.77 | −1.0262322076955124 | $4.5711127759539218E-005$ |
| 2.00 | −1.0041113165461855 | $4.9307591733370725E-005$ |
| 2.20 | −0.97760886156947968 | $3.8282923712390103E-005$ |
| 2.40 | −0.94918479419570190 | $2.9021376486267630E-005$ |
| 2.61 | −0.91806560611554477 | $4.6732829272644375E-005$ |

| $r_s$ | $E_{\text{pot}}$ | $\sigma$ |
|---|---|---|
| 1.251 | −2.5343485480940373 | $1.7282937687894599E-004$ |
| 1.31 | −2.4289447703824458 | $2.1048583615747020E-004$ |
| 1.42 | −2.2654474046295419 | $1.7978145017074295E-004$ |
| 1.57 | −2.1161672514949106 | $1.7280670524427827E-004$ |
| 1.77 | −1.9211489341668722 | $1.9810087745944895E-004$ |
| 2.00 | −1.7550388732410598 | $2.1508376570715710E-004$ |
| 2.20 | −1.6275226482128902 | $2.1241965671406023E-004$ |
| 2.40 | −1.5190052112659618 | $2.0094579794156217E-004$ |
| 2.61 | −1.4165099219554897 | $1.7617408787154755E-004$ |

| $r_s$ | $E_{\mathrm{kin}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.5774798438081008 | $1.4961588319500099E-004$ |
| 1.31 | 1.4461077490125964 | $2.5197061420961477E-004$ |
| 1.42 | 1.2528919233451967 | $1.9996030821164379E-004$ |
| 1.57 | 1.0858993150015495 | $1.8990710478070777E-004$ |
| 1.77 | 0.89491712818482361 | $2.0618249163101410E-004$ |
| 2.00 | 0.75092801591611524 | $2.2679635226944086E-004$ |
| 2.20 | 0.64991353886978542 | $2.4890066519825448E-004$ |
| 2.40 | 0.56982055778546037 | $2.2941548433492418E-004$ |
| 2.61 | 0.49844430970256193 | $1.6858988028767435E-004$ |

| $r_s$ | $E_{\mathrm{JF}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.5765577371668780 | $6.9025566446093709E-004$ |
| 1.31 | 1.4466120631043555 | $1.0921328336586697E-003$ |
| 1.42 | 1.2534754570741442 | $6.3440889244125223E-004$ |
| 1.57 | 1.0854920981648231 | $4.5034150896518704E-004$ |
| 1.77 | 0.89489457652079540 | $3.3106029723828563E-004$ |
| 2.00 | 0.75087821643495167 | $2.8933518156721097E-004$ |
| 2.20 | 0.64968746988933734 | $2.4916492270946500E-004$ |
| 2.40 | 0.56997455203388681 | $2.1163675890060684E-004$ |
| 2.61 | 0.49872312035043631 | $1.7605782440660692E-004$ |

### A.1.2 mol-hcp hydrogen, $N = 128$

| $r_s$ | $E$ | $\sigma$ |
|---|---|---|
| 1.251 | −0.89141810272978117 | $2.2106001404453084E-004$ |
| 1.31 | −0.91657680322010870 | $2.4819754902130522E-004$ |
| 1.42 | −0.94439653791592848 | $2.5850186767960651E-004$ |
| 1.57 | −0.95582863109008709 | $1.6979838946611732E-004$ |
| 1.77 | −0.94869290152266650 | $1.9525670968931585E-004$ |
| 2.00 | −0.93063095976782051 | $1.8802528075980104E-004$ |
| 2.20 | −0.91860525310897900 | $2.3478334735485291E-004$ |
| 2.40 | −0.91315061038941558 | $2.7852243193069484E-004$ |
| 2.61 | −0.90607466266857539 | $2.7212730410872913E-004$ |

| $r_s$ | $E_{\mathrm{pot}}$ | $\sigma$ |
|---|---|---|
| 1.251 | −2.5833569171640320 | $2.2785122333124710E-004$ |
| 1.31 | −2.5034216922225481 | $2.0709367329753921E-004$ |
| 1.42 | −2.3372970463099842 | $2.1768868202341071E-004$ |
| 1.57 | −2.1864392100616570 | $2.8133466976986956E-004$ |
| 1.77 | −2.0557275044254846 | $2.6555683219471696E-004$ |
| 2.00 | −1.9154525667843929 | $2.9996674964778746E-004$ |
| 2.20 | −1.9373827843584743 | $2.9406202366452324E-004$ |
| 2.40 | −1.9346087565452583 | $4.8203804481241090E-004$ |
| 2.61 | −1.8953290689309865 | $4.4493412156402715E-004$ |

| $r_s$ | $E_{\mathrm{kin}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.6919382560050562 | $3.1945579367975821E-004$ |
| 1.31 | 1.5868447817604385 | $3.4945462361529045E-004$ |
| 1.42 | 1.3928978089258433 | $2.7816866115021045E-004$ |
| 1.57 | 1.2306106889981578 | $3.3709372084109857E-004$ |
| 1.77 | 1.1070373324522969 | $3.1439342964851430E-004$ |
| 2.00 | 0.98482156292834877 | $2.8168157752462528E-004$ |
| 2.20 | 1.0187753769583014 | $2.8778156099290385E-004$ |
| 2.40 | 1.0214590380710920 | $3.0577117695394914E-004$ |
| 2.61 | 0.98925435375230819 | $3.0135473723599560E-004$ |

| $r_s$ | $E_{\mathrm{JF}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.6906594207862053 | $7.9966089536525362E-004$ |
| 1.31 | 1.5889076194655496 | $1.4391318038289461E-003$ |
| 1.42 | 1.3923591045109325 | $6.3862592758061375E-004$ |
| 1.57 | 1.2302648346133491 | $6.0837232624371936E-004$ |
| 1.77 | 1.1079435130206112 | $1.1772425954817347E-003$ |
| 2.00 | 0.98446113034362792 | $2.9152730541669928E-004$ |
| 2.20 | 1.0187443454915532 | $2.5118862996296353E-004$ |
| 2.40 | 1.0213900374501688 | $2.1304290201435390E-004$ |
| 2.61 | 0.98920590040978029 | $1.8738349637239878E-004$ |

## A.2 JS-DFT

### A.2.1 atm-bcc hydrogen, $N = 128$

| $r_s$ | $E$ | $\sigma$ |
|---|---|---|
| 1.251 | −0.96209469553124682 | $2.4448817388522394E-004$ |
| 1.31 | −0.99613524324398561 | $1.9977974625521729E-004$ |
| 1.42 | −1.0284239337890269 | $1.6403957734198316E-004$ |
| 1.57 | −1.0453209141238113 | $1.2937691704123110E-004$ |
| 1.77 | −1.0440543869751475 | $1.1843064901772222E-004$ |
| 2.00 | −1.0254113673795009 | $1.2690530228337915E-004$ |
| 2.20 | −1.0034476665088277 | $1.4202702736494777E-004$ |
| 2.40 | −0.98075263460535345 | $1.0585863825477774E-004$ |
| 2.61 | −0.95780002678106591 | $1.7584931671341995E-004$ |

| $r_s$ | $E_{\mathrm{pot}}$ | $\sigma$ |
|---|---|---|
| 1.251 | −2.6126229636832807 | $2.2756864076630326E-004$ |
| 1.31 | −2.5107290893228056 | $1.3605400473735082E-004$ |
| 1.42 | −2.3593457049796593 | $2.1681658667277353E-004$ |
| 1.57 | −2.2150338453193612 | $2.2961492097898123E-004$ |
| 1.77 | −2.0348193189521018 | $1.9950299288369833E-004$ |
| 2.00 | −1.9014232424289550 | $2.7167171362669595E-004$ |
| 2.20 | −1.8159360103879942 | $2.8193004073481608E-004$ |
| 2.40 | −1.7540692920822893 | $3.0842007494681254E-004$ |
| 2.61 | −1.7063601750344704 | $3.8017619411011982E-004$ |

| $r_s$ | $E_{\text{kin}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.6505288184690805 | $2.3967299529647357E-004$ |
| 1.31 | 1.5145935926956808 | $2.7600802287813472E-004$ |
| 1.42 | 1.3309223591849058 | $2.7156479265394066E-004$ |
| 1.57 | 1.1697140961180748 | $3.1320748259082692E-004$ |
| 1.77 | 0.99076531234612331 | $2.2092251748620265E-004$ |
| 2.00 | 0.87601163271675786 | $3.2399498712850950E-004$ |
| 2.20 | 0.81248765644162657 | $3.6659241401995753E-004$ |
| 2.40 | 0.77331694589157163 | $3.4111644845175417E-004$ |
| 2.61 | 0.74856022182204252 | $3.6969683398362769E-004$ |

| $r_s$ | $E_{\text{JF}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.6505471347409373 | $1.5464323995221744E-003$ |
| 1.31 | 1.5151641274767680 | $1.2784660194712666E-003$ |
| 1.42 | 1.3302392562295220 | $5.7787744056108281E-004$ |
| 1.57 | 1.1697619417562881 | $4.8636901173863096E-004$ |
| 1.77 | 0.99091186000584752 | $3.9873335118700945E-004$ |
| 2.00 | 0.87597065609530922 | $3.4662614346715460E-004$ |
| 2.20 | 0.81242910015225256 | $2.4562817288582633E-004$ |
| 2.40 | 0.77337410112766691 | $2.2683888968854615E-004$ |
| 2.61 | 0.74921911025215582 | $3.2407535635914402E-004$ |

### A.2.2 mol-hcp hydrogen, $N = 128$

| $r_s$ | $E$ | $\sigma$ |
|---|---|---|
| 1.251 | −0.96236134704692788 | $3.2133012231428624E-004$ |
| 1.31 | −0.99762230987269851 | $3.1119034703424475E-004$ |
| 1.42 | −1.0446869798500533 | $2.9820325918120317E-004$ |
| 1.57 | −1.0857880763000110 | $1.2082966744106935E-004$ |
| 1.77 | −1.1156351242329026 | $1.8519997756673653E-004$ |
| 2.00 | −1.1338552189360154 | $1.7828956084472224E-004$ |
| 2.20 | −1.1419398961413860 | $1.9681078008270661E-004$ |
| 2.40 | −1.1459954249730997 | $1.5135087761452017E-004$ |
| 2.61 | −1.1487791169337613 | $2.0468562723578514E-004$ |

| $r_s$ | $E_{\text{pot}}$ | $\sigma$ |
|---|---|---|
| 1.251 | −2.7927706193237274 | $6.2339780734902209E-004$ |
| 1.31 | −2.7297921811711809 | $4.4252393333769940E-004$ |
| 1.42 | −2.6213176255754531 | $4.7289145332330464E-004$ |
| 1.57 | −2.5423258732508254 | $5.2261090230610544E-004$ |
| 1.77 | −2.4646478245989494 | $5.0103087904760822E-004$ |
| 2.00 | −2.4227623543163554 | $4.8726121272665423E-004$ |
| 2.20 | −2.4057355290030102 | $4.9143718588539492E-004$ |
| 2.40 | −2.3898574243890720 | $5.6708772494763748E-004$ |
| 2.61 | −2.3584832655704178 | $5.2562773683647852E-004$ |

| $r_s$ | $E_{\text{kin}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.8304092452122369 | $7.3970704319091985E-004$ |
| 1.31 | 1.7321695667772510 | $5.4717249990047275E-004$ |
| 1.42 | 1.5766300237688819 | $4.4634844829068859E-004$ |
| 1.57 | 1.4565394898085957 | $5.5503580798933001E-004$ |
| 1.77 | 1.3490129499713435 | $5.9419220579704163E-004$ |
| 2.00 | 1.2889077738911161 | $5.7945986997340695E-004$ |
| 2.20 | 1.2637952401619110 | $6.3161334204602274E-004$ |
| 2.40 | 1.2438615646561617 | $5.9679742364192659E-004$ |
| 2.61 | 1.2097047286572620 | $5.7967089325080276E-004$ |

| $r_s$ | $E_{\text{JF}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.8316206801735948 | $1.4170918583502905E-003$ |
| 1.31 | 1.7312218811535125 | $1.1408848199183188E-003$ |
| 1.42 | 1.5773268754621430 | $7.4010545372809092E-004$ |
| 1.57 | 1.4560483140702605 | $6.5978221814573263E-004$ |
| 1.77 | 1.3489513025579005 | $5.3468351421307283E-004$ |
| 2.00 | 1.2881019896846102 | $2.6569495453267829E-004$ |
| 2.20 | 1.2631128819623856 | $2.1173763792930620E-004$ |
| 2.40 | 1.2429880945665772 | $1.9742463624175161E-004$ |
| 2.61 | 1.2082243605769349 | $1.3773765364597174E-004$ |

## A.3 ASWF-pw

### A.3.1 atm-bcc hydrogen, $N = 128$

| $r_s$ | $E$ | $\sigma$ |
|---|---|---|
| 1.251 | −0.95559156797175882 | $4.8673792563993530E-004$ |
| 1.31 | −0.98234579516248000 | $2.6063539777035431E-004$ |
| 1.42 | −1.0126916471447345 | $4.5910416890097025E-004$ |
| 1.57 | −1.0306420622683659 | $2.9456228850094270E-004$ |
| 1.77 | −1.0293311257132991 | $2.7227959616074279E-004$ |
| 2.00 | −1.0098184407280273 | $2.9837272544079178E-004$ |
| 2.20 | −0.98734498915742130 | $3.7147389568937807E-004$ |
| 2.40 | −0.96120292842643884 | $3.3089939751029920E-004$ |
| 2.61 | −0.93635430922644591 | $3.8679210219836566E-004$ |

| $r_s$ | $E_{\text{pot}}$ | $\sigma$ |
|---|---|---|
| 1.251 | −2.5072244090801679 | $3.8899975310125042E-004$ |
| 1.31 | −2.4167570931491822 | $2.1917096314978107E-004$ |
| 1.42 | −2.2597740505887560 | $2.8224203095058204E-004$ |
| 1.57 | −2.0833693268737528 | $1.8657717856962607E-004$ |
| 1.77 | −1.9059868715531729 | $2.0127770946056522E-004$ |
| 2.00 | −1.7554970863745867 | $1.7721969591997686E-004$ |
| 2.20 | −1.6593356840913118 | $2.8789675176102615E-004$ |
| 2.40 | −1.5715847384366339 | $2.7973146383717599E-004$ |
| 2.61 | −1.5136096854403902 | $3.4109915910915639E-004$ |

| $r_s$ | $E_{\mathrm{kin}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.5516323502745222 | $5.6473002475966984E-004$ |
| 1.31 | 1.4344096699274598 | $2.8261675799551338E-004$ |
| 1.42 | 1.2470831128648492 | $5.2697572637198378E-004$ |
| 1.57 | 1.0527266806976165 | $3.4763733278741235E-004$ |
| 1.77 | 0.87665455607091125 | $3.2939918277388204E-004$ |
| 2.00 | 0.74567949061822858 | $3.6730813004025293E-004$ |
| 2.20 | 0.67199096593566399 | $4.9006846679126208E-004$ |
| 2.40 | 0.61038256838805072 | $3.7181068467693427E-004$ |
| 2.61 | 0.57725528740458498 | $4.5457252797792959E-004$ |

| $r_s$ | $E_{\mathrm{JF}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.5504156844553210 | $1.6810010490589423E-003$ |
| 1.31 | 1.4334909046109914 | $7.2596498435022736E-004$ |
| 1.42 | 1.2472260230684027 | $1.5232740461730652E-003$ |
| 1.57 | 1.0528699152661078 | $6.5110781370119166E-004$ |
| 1.77 | 0.87707692382381064 | $6.8461398217777958E-004$ |
| 2.00 | 0.74587103559819434 | $4.9180795535859381E-004$ |
| 2.20 | 0.67110238474414974 | $4.7279085871603443E-004$ |
| 2.40 | 0.61064559619007985 | $4.7088267038798422E-004$ |
| 2.61 | 0.57661859056736919 | $5.9852938003337498E-004$ |

### A.3.2 mol-hcp hydrogen, $N = 128$

| $r_s$ | $E$ | $\sigma$ |
|---|---|---|
| 1.251 | −0.90100118887105962 | $3.5356068062335753E-004$ |
| 1.31 | −0.92419645189056476 | $3.5439885602698515E-004$ |
| 1.42 | −0.95255988803715508 | $3.3582405910446630E-004$ |
| 1.57 | −0.96547908948546002 | $3.3236994334160080E-004$ |
| 1.77 | −0.97026405808577487 | $4.2732671157623539E-004$ |
| 2.00 | −0.97648328808822393 | $4.1631479176515278E-004$ |
| 2.20 | −0.98473734608711005 | $3.8531085552217354E-004$ |
| 2.40 | −1.0021921179128761 | $6.9544028170590846E-004$ |
| 2.61 | −1.0136509706971992 | $6.6781718252286251E-004$ |

| $r_s$ | $E_{\mathrm{pot}}$ | $\sigma$ |
|---|---|---|
| 1.251 | −2.5452781027228020 | $2.2445026305251402E-004$ |
| 1.31 | −2.4315202927419630 | $2.1037152893475894E-004$ |
| 1.42 | −2.2831907257786894 | $2.7703085911424296E-004$ |
| 1.57 | −2.1461096817629177 | $2.8743922220176042E-004$ |
| 1.77 | −2.0439618083301361 | $3.6038766728879539E-004$ |
| 2.00 | −2.0496400058756374 | $6.2668027721377918E-004$ |
| 2.20 | −2.1018517024634327 | $7.6881693646036874E-004$ |
| 2.40 | −2.2153269365462043 | $1.1612613413227058E-003$ |
| 2.61 | −2.2249845854738832 | $9.4199943771434156E-004$ |

| $r_s$ | $E_{\text{kin}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.6442785789424379 | $3.7064825150093944E-004$ |
| 1.31 | 1.5073238227841341 | $4.4470595856988470E-004$ |
| 1.42 | 1.3306306775819878 | $3.8995635072585570E-004$ |
| 1.57 | 1.1806308632880345 | $3.2039518599972540E-004$ |
| 1.77 | 1.0736991069147219 | $4.5973618005639561E-004$ |
| 2.00 | 1.0731555730083211 | $6.2017514624967604E-004$ |
| 2.20 | 1.1171162014716052 | $7.0881471593669967E-004$ |
| 2.40 | 1.2131325966714090 | $7.8999762102794811E-004$ |
| 2.61 | 1.2113338709235422 | $5.5460220480993796E-004$ |

| $r_s$ | $E_{\text{JF}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.6430334723367481 | $1.0445363232583747E-003$ |
| 1.31 | 1.5071758942747806 | $1.3613672691455676E-003$ |
| 1.42 | 1.3487824921223091 | $1.7742514127482772E-002$ |
| 1.57 | 1.1800593035319220 | $7.5857333986239378E-004$ |
| 1.77 | 1.0738169330774017 | $8.0329077610040520E-004$ |
| 2.00 | 1.0736283224351937 | $6.9632779862593833E-004$ |
| 2.20 | 1.1164260739932677 | $6.5720934330260343E-004$ |
| 2.40 | 1.2132425774014501 | $7.8560707648848256E-004$ |
| 2.61 | 1.2111136118583308 | $7.1812655427914355E-004$ |

## A.4 ASWF-DFT

### A.4.1 atm-bcc hydrogen, $N = 128$

| $r_s$ | $E$ | $\sigma$ |
|---|---|---|
| 1.251 | −0.96848716668184598 | $2.2253102118554224E-004$ |
| 1.31 | −0.99664207872699118 | $2.0247427940508507E-004$ |
| 1.42 | −1.0291452961801257 | $1.7858607460696446E-004$ |
| 1.57 | −1.0457528186633238 | $1.6419941755645726E-004$ |
| 1.77 | −1.0386289335023187 | $2.0360707101457220E-004$ |
| 2.00 | −1.0113536851207507 | $2.9793776159946217E-004$ |
| 2.20 | −0.98997501845887992 | $2.9882493083079249E-004$ |
| 2.40 | −0.96491963840458483 | $2.9141897504945789E-004$ |
| 2.61 | −0.93851219957910859 | $2.1730417570400840E-004$ |

| $r_s$ | $E_{\text{pot}}$ | $\sigma$ |
|---|---|---|
| 1.251 | −2.5964246305319585 | $1.6109980254503292E-004$ |
| 1.31 | −2.5221882263574544 | $1.6802631513313708E-004$ |
| 1.42 | −2.3639744741197237 | $1.9247502692476572E-004$ |
| 1.57 | −2.2249451472045005 | $1.4417935121115478E-004$ |
| 1.77 | −1.9918775304101222 | $2.2193463382494521E-004$ |
| 2.00 | −1.8155507715894039 | $2.2181843668803798E-004$ |
| 2.20 | −1.7357063742972874 | $2.3379949369705866E-004$ |
| 2.40 | −1.6696016304817443 | $2.3483540424536952E-004$ |
| 2.61 | −1.6278261980221043 | $2.9049764458039993E-004$ |

| $r_s$ | $E_{\mathrm{kin}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.6279373643952419 | $2.7227339050038194E-004$ |
| 1.31 | 1.5255461808262858 | $2.2749987205275248E-004$ |
| 1.42 | 1.3348290803243414 | $2.2949636279401649E-004$ |
| 1.57 | 1.1791920812222798 | $2.4271264224781623E-004$ |
| 1.77 | 0.95324885447484287 | $3.3832387583159970E-004$ |
| 2.00 | 0.80419680036204844 | $3.4729975540939869E-004$ |
| 2.20 | 0.74573134246439821 | $3.3337278228049810E-004$ |
| 2.40 | 0.70468196939194039 | $3.9847723455302712E-004$ |
| 2.61 | 0.68931398964282609 | $3.2667379854811189E-004$ |

| $r_s$ | $E_{\mathrm{JF}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.6265185111815752 | $8.0737355544421299E-004$ |
| 1.31 | 1.5257683707953416 | $8.1078332345359879E-004$ |
| 1.42 | 1.3340134432225526 | $4.4661561844213196E-004$ |
| 1.57 | 1.1799661699450541 | $3.7033445046252193E-004$ |
| 1.77 | 0.95285950947375797 | $4.1448715550729606E-004$ |
| 2.00 | 0.80382681713107029 | $3.6951723331475438E-004$ |
| 2.20 | 0.74535878384637999 | $4.1378586175415193E-004$ |
| 2.40 | 0.70433895300646920 | $4.0863989352049742E-004$ |
| 2.61 | 0.68932311621466380 | $4.3963509598898929E-004$ |

### A.4.2 mol-hcp hydrogen, $N = 128$

| $r_s$ | $E$ | $\sigma$ |
|---|---|---|
| 1.251 | −0.95927608114965901 | $3.9952486414107698E-003$ |
| 1.31 | −0.98984162616192595 | $3.2700458324837218E-003$ |
| 1.42 | −1.0394218645181867 | $5.0176912160902563E-003$ |
| 1.57 | −1.0849161230759976 | $3.5821170167933219E-004$ |
| 1.77 | −1.1014696811420837 | $6.8979024574880555E-004$ |
| 2.00 | −1.1116882014628569 | $8.9830257101174756E-004$ |
| 2.20 | −1.1180312987184315 | $7.6915705655155090E-004$ |
| 2.40 | −1.1205991984417778 | $9.1579674385053062E-004$ |
| 2.61 | −1.1229806444633728 | $8.6577834919531924E-004$ |

| $r_s$ | $E_{\mathrm{pot}}$ | $\sigma$ |
|---|---|---|
| 1.251 | −2.7935512087596410 | $1.1115923438113096E-004$ |
| 1.31 | −2.7256420214710464 | $5.8466786663226506E-004$ |
| 1.42 | −2.6396570609139531 | $1.2854435486327315E-004$ |
| 1.57 | −2.5500080622934931 | $7.0502502854375803E-004$ |
| 1.77 | −2.3794195893489833 | $7.1301110897775037E-004$ |
| 2.00 | −2.2956489640514710 | $6.2490199698412999E-004$ |
| 2.20 | −2.2791930930345252 | $7.4120838126337182E-004$ |
| 2.40 | −2.2463140671015687 | $6.1496735379329946E-004$ |
| 2.61 | −2.2613557283884256 | $5.9901357859243971E-004$ |

| $r_s$ | $E_{\mathrm{kin}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.8342752063361687 | $3.9040747291041542E-003$ |
| 1.31 | 1.7357895790335733 | $3.5787462725403100E-003$ |
| 1.42 | 1.6002352989263195 | $5.0428237064802026E-003$ |
| 1.57 | 1.4650920008228412 | $8.6157837429307093E-004$ |
| 1.77 | 1.2779467555872150 | $1.0121642460421819E-003$ |
| 2.00 | 1.1839599856802641 | $1.0826815248088234E-003$ |
| 2.20 | 1.1611618549536435 | $8.6183630970965749E-004$ |
| 2.40 | 1.1257136432807981 | $1.1940916592755476E-003$ |
| 2.61 | 1.1383758366088788 | $8.1315233990237430E-004$ |

| $r_s$ | $E_{\mathrm{JF}}$ | $\sigma$ |
|---|---|---|
| 1.251 | 1.8319515052155884 | $5.5283604241262519E-004$ |
| 1.31 | 1.7284380110447122 | $1.3108867024575052E-003$ |
| 1.42 | 1.5952676762208036 | $3.8652550512599114E-004$ |
| 1.57 | 1.4661436726203130 | $1.1137797341650988E-003$ |
| 1.77 | 1.2773593015175289 | $9.0674247933610747E-004$ |
| 2.00 | 1.1833511194306885 | $6.1158124625161235E-004$ |
| 2.20 | 1.1608820794179893 | $8.6255132933272388E-004$ |
| 2.40 | 1.1251600519219793 | $7.5616896635969516E-004$ |
| 2.61 | 1.1373793410634796 | $7.3982321240189622E-004$ |

# Bibliography


[1] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, et al., Journal of Physics: Condensed Matter **21**, 395502 (19pp) (2009), URL `http://www.quantum-espresso.org`.

[2] W. L. McMillan, Phys. Rev. **138**, A442 (1965), URL `http://link.aps.org/doi/10.1103/PhysRev.138.A442`.

[3] S. Vitiello, K. Runge, and M. H. Kalos, Phys. Rev. Lett. **60**, 1970 (1988).

[4] L. Reatto and G. L. Masserini, Phys. Rev. B **38**, 4516 (1988).

[5] F. Pederiva, A. Ferrante, S. Fantoni, and L. Reatto, Phys. Rev. Lett. **72**, 2589 (1994).

[6] F. Pederiva, S. A. Vitiello, K. Gernoth, S. Fantoni, and L. Reatto, Phys. Rev. B **53**, 15129 (1996).

[7] F. Pederiva, G. V. Chester, S. Fantoni, and L. Reatto, Phys. Rev. B **56**, 5909 (1997).

[8] F. Operetto and F. Pederiva, Phys. Rev. B **69**, 024203 (2004).

[9] L. Dandrea, F. Pederiva, S. Gandolfi, and M. H. Kalos, Phys. Rev. Lett. **102**, 255302 (2009), URL `http://link.aps.org/doi/10.1103/PhysRevLett.102.255302`.

[10] S. A. Vitiello, K. J. Runge, G. V. Chester, and M. H. Kalos, Phys. Rev. B **42**, 228 (1990), URL `http://link.aps.org/doi/10.1103/PhysRevB.42.228`.

[11] S. A. Vitiello and P. A. Whitlock, Phys. Rev. B **44**, 7373 (1991), URL `http://link.aps.org/doi/10.1103/PhysRevB.44.7373`.

[12] W. Wu, S. A. Vitiello, L. Reatto, and M. H. Kalos, Phys. Rev. Lett. **67**, 1446 (1991), URL `http://link.aps.org/doi/10.1103/PhysRevLett.67.1446`.

[13] T. MacFarland, S. A. Vitiello, L. Reatto, G. V. Chester, and M. H. Kalos, Phys. Rev. B **50**, 13577 (1994), URL `http://link.aps.org/doi/10.1103/PhysRevB.50.13577`.

[14] S. A. Vitiello, L. Reatto, G. V. Chester, and M. H. Kalos, Phys. Rev. B **54**, 1205 (1996), URL http://link.aps.org/doi/10.1103/PhysRevB.54.1205.

[15] D. E. Galli, E. Cecchetti, and L. Reatto, Phys. Rev. Lett. **77**, 5401 (1996), URL http://link.aps.org/doi/10.1103/PhysRevLett.77.5401.

[16] S. Moroni, D. E. Galli, S. Fantoni, and L. Reatto, Phys. Rev. B **58**, 909 (1998), URL http://link.aps.org/doi/10.1103/PhysRevB.58.909.

[17] B. Krishnamachari and G. V. Chester, Phys. Rev. B **61**, 9677 (2000), URL http://link.aps.org/doi/10.1103/PhysRevB.61.9677.

[18] D. E. Galli and L. Reatto, Phys. Rev. Lett. **96**, 165301 (2006), URL http://link.aps.org/doi/10.1103/PhysRevLett.96.165301.

[19] T. MacFarland, S. Vitiello, and L. Reatto, Journal of Low Temperature Physics **89**, 433 (1992), ISSN 0022-2291, URL http://dx.doi.org/10.1007/BF00694057.

[20] D. Galli and L. Reatto, Journal of Low Temperature Physics **136**, 343 (2004), ISSN 0022-2291, URL http://dx.doi.org/10.1023/B%3AJOLT.0000041271.95836.2e.

[21] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964), URL http://link.aps.org/doi/10.1103/PhysRev.136.B864.

[22] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965), URL http://link.aps.org/doi/10.1103/PhysRev.140.A1133.

[23] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys **21**, 1087 (1953).

[24] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods* (Wiley-VCH, Weinheim, 2008).

[25] E. Feenberg, Annals of Physics **84**, 128 (1974), ISSN 0003-4916, URL http://www.sciencedirect.com/science/article/pii/0003491674902966.

[26] R. P. Feynman and M. Cohen, Phys. Rev. **102**, 1189 (1956), URL http://link.aps.org/doi/10.1103/PhysRev.102.1189.

[27] A. Jayaraman, Rev. Mod. Phys. **55**, 65 (1983), URL http://link.aps.org/doi/10.1103/RevModPhys.55.65.

[28] H.-k. Mao and R. J. Hemley, Rev. Mod. Phys. **66**, 671 (1994), URL http://link.aps.org/doi/10.1103/RevModPhys.66.671.

[29] J. M. McMahon, M. A. Morales, C. Pierleoni, and D. M. Ceperley, Rev. Mod. Phys. **84**, 1607 (2012), URL http://link.aps.org/doi/10.1103/RevModPhys.84.1607.

[30] URL http://encyclopedia.airliquide.com.

[31] URL http://imagine.gsfc.nasa.gov/docs/ask_astro/answers/971113i.html.

[32] C. L. Guillaume, E. Gregoryanz, O. Degtyareva, M. I. McMahon, M. Hanfland, S. Evans, M. Guthrie, S. V. Sinogeikin, and H.-K. Mao, Nature Physics (2011).

[33] E. Wigner and H. B. Huntington, The Journal of Chemical Physics **3**, 764 (1935), URL http://link.aip.org/link/?JCP/3/764/1.

[34] P. Loubeyre, R. LeToullec, D. Hausermann, M. Hanfland, R. J. Hemley, H. K. Mao, and L. W. Finger, Nature **383**, 702 (1996).

[35] H. Eshet, R. Z. Khaliullin, T. D. Kühne, J. Behler, and M. Parrinello, Phys. Rev. Lett. **108**, 115701 (2012), URL http://link.aps.org/doi/10.1103/PhysRevLett.108.115701.

[36] N. W. Ashcroft, Physical Review Letters **21**, 1748 (1968).

[37] E. Gregoryanz, O. Degtyareva, M. Somayazulu, R. J. Hemley, and H.-k. Mao, Phys. Rev. Lett. **94**, 185502 (2005), URL http://link.aps.org/doi/10.1103/PhysRevLett.94.185502.

[38] W. J. Nellis, S. T. Weir, and A. C. Mitchell, Phys. Rev. B **59**, 3434 (1999), URL http://link.aps.org/doi/10.1103/PhysRevB.59.3434.

[39] S. T. Weir, A. C. Mitchell, and W. J. Nellis, Phys. Rev. Lett. **76**, 1860 (1996), URL http://link.aps.org/doi/10.1103/PhysRevLett.76.1860.

[40] P. W. Bridgman, *The Logic of Modern Physics* (1927).

[41] I. F. Silvera and S. Deemyad, Low Temperature Physics **35**, 318 (2009).

[42] M. I. Eremets and I. A. Trojan, JETP Letters **89**, 174 (2009).

[43] N. C. Holmes, M. Ross, and W. J. Nellis, **52**, 15835 (1995).

[44] T. J. Lenosky, J. D. Kress, L. A. Collins, and I. Kwon, Phys. Rev. B **55**, R11907 (1997), URL http://link.aps.org/doi/10.1103/PhysRevB.55.R11907.

[45] A. Bunker, S. Nagel, R. Redmer, and G. Röpke, Phys. Rev. B **56**, 3094 (1997), URL http://link.aps.org/doi/10.1103/PhysRevB.56.3094.

[46] S. Scandolo, Proceedings of the National Academy of Sciences **100**, 3051 (2003), http://www.pnas.org/content/100/6/3051.full.pdf+html, URL http://www.pnas.org/content/100/6/3051.abstract.

[47] C. Pierleoni, D. M. Ceperley, and M. Holzmann, Phys. Rev. Lett. **93**, 146402 (2004), URL http://link.aps.org/doi/10.1103/PhysRevLett.93.146402.

[48] S. Baroni and S. Moroni, Phys. Rev. Lett. **82**, 4745 (1999), URL http://link.aps.org/doi/10.1103/PhysRevLett.82.4745.

[49] K. T. Delaney, C. Pierleoni, and D. M. Ceperley, Phys. Rev. Lett. **97**, 235702 (2006), URL http://link.aps.org/doi/10.1103/PhysRevLett.97.235702.

[50] I. Tamblyn and S. A. Bonev, Phys. Rev. Lett. **104**, 065702 (2010), URL http://link.aps.org/doi/10.1103/PhysRevLett.104.065702.

[51] M. A. Morales, C. Pierleoni, E. Schwegler, and D. M. Ceperley, Proceedings of the National Academy of Sciences **107**, 12799 (2010), http://www.pnas.org/content/107/29/12799.full.pdf+html, URL http://www.pnas.org/content/107/29/12799.abstract.

[52] W. Lorenzen, B. Holst, and R. Redmer, Phys. Rev. B **82**, 195107 (2010), URL http://link.aps.org/doi/10.1103/PhysRevB.82.195107.

[53] M. Hanfland, R. J. Hemley, and H.-k. Mao, Phys. Rev. Lett. **70**, 3760 (1993), URL http://link.aps.org/doi/10.1103/PhysRevLett.70.3760.

[54] R. J. Hemley and H. K. Mao, Phys. Rev. Lett. **61**, 857 (1988), URL http://link.aps.org/doi/10.1103/PhysRevLett.61.857.

[55] H. E. Lorenzana, I. F. Silvera, and K. A. Goettel, Phys. Rev. Lett. **63**, 2080 (1989), URL http://link.aps.org/doi/10.1103/PhysRevLett.63.2080.

[56] P. Loubeyre, F. Occelli, and R. LeToullec, Nature **416**, 613 (2002).

[57] M. I. Eremets and I. A. Troyan, Nature Materials **10**, 927 (2011).

[58] R. T. Howie, C. L. Guillaume, T. Scheler, A. F. Goncharov, and E. Gregoryanz, Phys. Rev. Lett. **108**, 125501 (2012), URL http://link.aps.org/doi/10.1103/PhysRevLett.108.125501.

[59] L. Vereshchagin, E. Yakovlev, and Y. A. Timofeev, JETP Lett. **21**, 85 (1975).

[60] K. K. Mon, G. V. Chester, and N. W. Ashcroft, Phys. Rev. B **21**, 2641 (1980), URL http://link.aps.org/doi/10.1103/PhysRevB.21.2641.

[61] D. M. Ceperley and B. J. Alder, Phys. Rev. B **36**, 2092 (1987), URL http://link.aps.org/doi/10.1103/PhysRevB.36.2092.

[62] V. Natoli, R. M. Martin, and D. M. Ceperley, Phys. Rev. Lett. **70**, 1952 (1993), URL http://link.aps.org/doi/10.1103/PhysRevLett.70.1952.

[63] V. Natoli, R. M. Martin, and D. Ceperley, Phys. Rev. Lett. **74**, 1601 (1995), URL http://link.aps.org/doi/10.1103/PhysRevLett.74.1601.

[64] C. J. Pickard and R. J. Needs, Nature Physics **3**, 473 (2007).

[65] J. M. McMahon and D. M. Ceperley, Phys. Rev. Lett. **106**, 165302 (2011), URL http://link.aps.org/doi/10.1103/PhysRevLett.106.165302.

[66] C. J. Pickard, M. Martinez-Canales, and R. J. Needs, Phys. Rev. B **85**, 214114 (2012), URL http://link.aps.org/doi/10.1103/PhysRevB.85.214114.

[67] S. Azadi, W. M. C. Foulkes, and T. D. Kühne, New Journal of Physics **15**, 113005 (2013), URL http://stacks.iop.org/1367-2630/15/i=11/a=113005.

[68] F. Datchi, P. Loubeyre, and R. LeToullec, Phys. Rev. B **61**, 6535 (2000), URL http://link.aps.org/doi/10.1103/PhysRevB.61.6535.

[69] D. Hohl, V. Natoli, D. M. Ceperley, and R. M. Martin, Phys. Rev. Lett. **71**, 541 (1993), URL http://link.aps.org/doi/10.1103/PhysRevLett.71.541.

[70] A. Alavi, M. Parrinello, and D. Frenkel, Science **269**, 1252 (1995), http://www.sciencemag.org/content/269/5228/1252.full.pdf, URL http://www.sciencemag.org/content/269/5228/1252.abstract.

[71] S. A. Bonev, E. Schwegler, T. Ogitsu, and G. Galli, Nature **431**, 669 (2004).

[72] S. Deemyad and I. F. Silvera, Phys. Rev. Lett. **100**, 155701 (2008), URL http://link.aps.org/doi/10.1103/PhysRevLett.100.155701.

[73] J. Chen, X.-Z. Li, Q. Zhang, M. I. J. Probert, C. J. Pickard, R. J. Needs, A. Michaelides, and E. Wang, Nature Communications (2013).

[74] G. Mazzola, S. Yunoki, and S. Sorella, Nature Communications (2014).

[75] J. Oliva and N. W. Ashcroft, Phys. Rev. B **23**, 6399 (1981), URL http://link.aps.org/doi/10.1103/PhysRevB.23.6399.

[76] J. Oliva and N. W. Ashcroft, Phys. Rev. B **25**, 223 (1982), URL http://link.aps.org/doi/10.1103/PhysRevB.25.223.

[77] E. Babaev, A. Sudbo, and N. W. Ashcroft, Nature (2004), URL http://dx.doi.org/10.1038/nature02910.

[78] E. Babaev and N. W. Ashcroft, Nature (2007), URL http://dx.doi.org/10.1038/nature02910.

[79] C. Lin, F. H. Zong, and D. M. Ceperley, Phys. Rev. E **64**, 016702 (2001), URL http://link.aps.org/doi/10.1103/PhysRevE.64.016702.

[80] C. Attaccalite, Ph.D. thesis, SISSA - Trieste (2005).

[81] S. Sorella, Phys. Rev. Lett. **80**, 4558 (1998), URL http://link.aps.org/doi/10.1103/PhysRevLett.80.4558.

[82] M. Casula and S. Sorella, The Journal of Chemical Physics **119**, 6500 (2003), URL http://link.aip.org/link/?JCP/119/6500/1.

[83] S. Sorella, Phys. Rev. B **71**, 241103 (2005), URL http://link.aps.org/doi/10.1103/PhysRevB.71.241103.

[84] D. Ceperley, G. V. Chester, and M. H. Kalos, Phys. Rev. B **16**, 3081 (1977), URL http://link.aps.org/doi/10.1103/PhysRevB.16.3081.

[85] URL http://www.tcm.phy.cam.ac.uk/~ajw29/thesis/node130.html.

[86] L. M. Fraser, W. M. C. Foulkes, G. Rajagopal, R. J. Needs, S. D. Kenny, and A. J. Williamson, Phys. Rev. B **53**, 1814 (1996), URL http://link.aps.org/doi/10.1103/PhysRevB.53.1814.

[87] M. Holzmann, private communication.

[88] P. P. Ewald, Annalen der Physik **369**, 253 (1921), ISSN 1521-3889, URL http://dx.doi.org/10.1002/andp.19213690304.

[89] M. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford Science Publications, 1987).

[90] V. Natoli and D. M. Ceperley, Journal of Computational Physics **117**, 171 (1995), ISSN 0021-9991, URL http://www.sciencedirect.com/science/article/pii/S0021999185710546.

[91] H. J. Monkhorst and J. D. Pack, Physical Review B **13**, 5188 (1976).

[92] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in Fortran* (Cambridge University Press, 1992), 2nd ed.

[93] J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer, 2006), 2nd ed.

[94] P. Komarek, Ph.D. thesis, School of Computer Science Carnegie Mellon University (2004).

[95] E. Polak, *Computational Methods in Optimization* (New York: Academic Press, 1971).

[96] M. Holzmann, D. M. Ceperley, C. Pierleoni, and K. Esler, Phys. Rev. E **68**, 046707 (2003), URL http://link.aps.org/doi/10.1103/PhysRevE.68.046707.

[97] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Science **321**, 792 (2008), http://www.sciencemag.org/content/321/5890/792.full.pdf, URL http://www.sciencemag.org/content/321/5890/792.abstract.

[98] W. Kolos and L. Wolniewicz, The Journal of Chemical Physics **49**, 404 (1968), URL http://scitation.aip.org/content/aip/journal/jcp/49/1/10.1063/1.1669836.

[99] W. Kolos, K. Szalewicz, and H. J. Monkhorst, The Journal of Chemical Physics **84**, 3278 (1986), URL http://scitation.aip.org/content/aip/journal/jcp/84/6/10.1063/1.450258.

[100] S. Chandrasekharan and U.-J. Wiese, Phys. Rev. Lett. **83**, 3116 (1999), URL http://link.aps.org/doi/10.1103/PhysRevLett.83.3116.

[101] M. Troyer and U.-J. Wiese, Phys. Rev. Lett. **94**, 170201 (2005), URL http://link.aps.org/doi/10.1103/PhysRevLett.94.170201.

[102] M. H. Kalos, Phys. Rev. **128**, 1791 (1962), URL http://link.aps.org/doi/10.1103/PhysRev.128.1791.

[103] E. C. Kerr, Phys. Rev. **96**, 551 (1954), URL http://link.aps.org/doi/10.1103/PhysRev.96.551.

[104] R. A. Aziz, V. P. S. Nain, J. S. Carley, W. L. Taylor, and G. T. McConville, J. Chem. Phys. (1979).

[105] S. A. Sofianos, S. A. Rakityansky, and S. E. Massen, Phys. Rev. A **60**, 337 (1999).

[106] K. E. Schmidt, M. A. Lee, M. H. Kalos, and G. V. Chester, Phys. Rev. Lett. **47**, 807 (1981), URL http://link.aps.org/doi/10.1103/PhysRevLett.47.807.

[107] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980), URL http://link.aps.org/doi/10.1103/PhysRevLett.45.566.

[108] H. Kahn and A. W. Marshall, Journal of the Operations Research Society of America **1**, 263 (1953), http://pubsonline.informs.org/doi/pdf/10.1287/opre.1.5.263, URL http://pubsonline.informs.org/doi/abs/10.1287/opre.1.5.263.