SISSA ISAS

SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI
INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

Auxiliary Field Quantum Monte Carlo for systems with repulsive Coulomb interactions

Thesis submitted for the degree of

"Doctor Philosophiæ"

CANDIDATE

Pier Luigi Silvestrelli

SUPERVISORS

Prof. Stefano Baroni

Prof. Roberto Car

November 1992

Index

Introduction

A reliable solution of the electronic structure problem, i.e. the quantum many-body problem of interacting electrons, is a very important goal in several areas of science. For instance, accurate calculations of molecular properties such as binding energies, bond lengths, charge distributions and potential energy surfaces are essential in *Quantum Chemistry* for both basic scientific understanding and technological applications.

In *Condensed Matter Physics* a suitable treatment of many-body effects is a crucial achievement, especially for studying the new high-temperature superconductors [1] and other strongly correlated systems such as Mott-Hubbard insulators [2] and heavy fermion metals [3]. In order to attain this result one must accurately take into account particle correlations, a very hard task when realistic and truly interesting systems are considered. For a many-electron system the *correlation energy* is defined as the difference between the Hartree-Fock energy, that is the energy obtained when each electron is considered as moving in the average field of the remaining $(N-1)$ electrons subject to Pauli principle, and the exact non-relativistic energy. We point out that the correlation energy contribution is usually very small. For example, in molecular systems, it is typically of the order of 1%. Therefore, if one is merely interested in the total molecular energy, then correlation represents only a slight correction. However it has an important effect on properties like molecular binding energies which are often characterized by an

energy scale of only a fraction of the correlation energy; for instance the O-H bond strength in water is about 50% of the molecular correlation energy. In these cases if correlation is not taken into account wrong results may be obtained even for qualitative predictions.

In recent years computer power greatly increased thus opening the possibility to attack the problem by numerical calculations. In this context various techniques have been devised.

Exact techniques (full Configuration Interaction[4], Many-Body Perturbation Theory[5]...) aim at evaluating the exact ground state of a given Hamiltonian $\hat{H}$. In practice, however, these methods are applicable to systems with a few electrons only, since the computational cost required to deal with a real many-body wave function grows exponentially with the size of the basis set and the number of electrons.

Obviously a lot of *approximate techniques* exist. We can mention the variational procedures like Hartree, Hartree-Fock[6], and Jastrow[7], the Random Phase Approximation[8] in Many-Body Perturbation Theory, and the Local Density Approximation (LDA)[9], which is a very convenient and popular approximation to the well-known Density Functional Theory[10–12].

Unfortunately all these approximations are difficult to control and to improve systematically: for example the variational technique results are too much dependent on the choice of the form of the variational wave function, since only a small number of parameters can practically be varied in order to minimize the ground state estimated energy. Current approximation methods have costs ranging from the 3^{th} power to the 7^{th} power of the number of electrons and, therefore, they are still rather cumbersome. For instance the Hartree-Fock method, which is largely

pursued by quantum chemists, is already expensive enough (the required computational cost varies between the 3^{th} and the 4^{th} power of the number of electrons) even though it completely neglects correlation. The Density Functional methods are extensively used by condensed matter physicists and are characterized by their applicability to large systems (hundreds of electrons with present computing capabilities). However the Density Functional approach is, in practice, implemented by assuming that the exchange-correlation energy density depends locally on the electron density (LDA), and, although the incorrect character of this assumption is evident, no genuine and really efficient improvement has been developed till now in the DFT framework.

Another group of methods, the *stochastic techniques*, simulate quantum systems and calculate their ground state properties by using classical statistical methods: these techniques are generally called Quantum Monte Carlo (QMC) methods[13–21]. They are both of the *variational* type in which the Monte Carlo method is used to evaluate numerically expectation values obtained from a given trial wave function ψ_T, and of the *exact* type in which the Schrödinger equation is solved. In these latter approaches various procedures are employed to stochastically sample the exact wave function of the physical system, subject only to statistical errors. Properties of interest are "measured" as the system evolves in imaginary time, under the action of the Schrödinger equation. When a stationary state is obtained, averages of the measured quantities provide the desired expectation values.

Monte Carlo algorithms are very promising because they treat correlation effects, either approximately or exactly, at a numerical cost having a size dependence similar to that of single-particle approaches (like Hartree or Hartree-Fock).

However the proper inclusion of the Pauli principle is a major difficulty in the exact QMC methods. This is so because the many-body wave function is usually described by a statistically evolving ensemble of configurations (specified by the coordinates of each particle). The Pauli principle, which enforces a spatially non-local relation between configurations which differ by the interchange of a pair of fermionic particles, is thus difficult to implement in the simple local algorithms used to evolve the ensemble.

A different Quantum Monte Carlo method is performed by introducing auxiliary fields to decouple the interaction and was successfully used by Hirsch, Scalapino et al.[22–27] to investigate finite temperature properties in the Hubbard model. It was recently suggested as an alternative method of approaching the ground state electronic problem by Sugiyama and Koonin[28].

In this technique, which was pursued[29–40] by Sorella et al., and by Hamann and Fahy et al., the *Hubbard-Stratonovich Transformation*[41,42] (HST) is applied. The ground state $|\psi_0\rangle$ of the Hamiltonian $\hat{H}$ is obtained by filtering out from an initial trial wave function $|\psi_T\rangle$ its ground state component by applying to $|\psi_T\rangle$ the imaginary time propagator $e^{-\beta\hat{H}}$ for large enough imaginary time β. In fact if $\hat{H}|\psi_i\rangle = E_i|\psi_i\rangle$ the exponential decay of the amplitude of higher energy states in the imaginary time evolution:

$$\lim_{\beta\to\infty} e^{-\beta\hat{H}}|\psi_T\rangle = \lim_{\beta\to\infty}\sum_i e^{-\beta E_i}|\psi_i\rangle\langle\psi_i|\psi_T\rangle = e^{-\beta E_0}|\psi_0\rangle\langle\psi_0|\psi_T\rangle \propto |\psi_0\rangle$$

leaves only the lowest state $|\psi_0\rangle$ in the infinite β limit, provided that $\langle\psi_0|\psi_T\rangle \neq 0$. The HST of the propagator $e^{-\beta\hat{H}}$ introduces an *auxiliary field* σ to reduce the exponential of a *two-body* operator to a functional integral over an infinite set of exponentials of *one-body* operators. In fact the imaginary time evolution $e^{-\beta\hat{H}}$ is convenient for numerical treatment when the Hamiltonian contains only one-body

operators and no interaction term. In practice HST transforms the many-body problem into a functional integral over the variables σ, since the two particle interaction term in $\hat{H}$ is replaced by one particle interactions with a set of random, time-dependent auxiliary fields; integration over a Gaussian distribution of these fields restores the physical interaction. After a suitable discretization the functional integral can be evaluated numerically by statistical methods (Monte Carlo, Langevin Dynamics, Hybrid Molecular Dynamics-Monte Carlo techniques...). Therefore this approach is especially suitable for lattice models. In fact, until now, the auxiliary field formulation has been used extensively to study electron correlations in the Hubbard model[29–40].

In comparison with other Quantum Monte Carlo techniques, this new Auxiliary Field method is attractive for condensed matter physics applications essentially because (i) it allows an easier application of the widely used one-body techniques, (ii) general observable estimators can be calculated without any particular problem, (iii) antisymmetrization for electron wave function can be enforced exactly, and (iv) the famous "fermion sign problem", which is one of the main difficulty in fermionic many-body calculations, could become less troublesome.

In this thesis the Auxiliary Field Quantum Monte Carlo (AFQMC) method is applied, for the first time, to continuous, realistic physical systems, namely to the Hydrogen molecules H_2 and H_3. For the H_2 ground state the "fermion sign problem" is not present since, in this case, the wave function is nodeless. Nevertheless the generalization from a simple and schematic Hubbard model to a molecular system with a continuous Coulomb potential r^{-1} is, in itself, a non trivial task and even a fundamental step towards the application of this technique to more interesting and complex physical systems.

In preliminary applications to a realistic Hamiltonian, for the H_2 molecule, a serious difficulty was found: huge statistical fluctuations made an accurate calculation of physical properties practically impossible. They were mainly related to the repulsive, long-range character of the Coulomb interaction and they increased dramatically by using larger and larger plane wave basis sets.

We have developed a particular AFQMC scheme to master these statistical fluctuations and we have demonstrated our technique by studying the Singlet and Triplet ground states (for the Triplet ground state the fermion sign problem could be important) of H_2. For this simple system the method really works since the statistical fluctuations have been reduced to an acceptable level. Calculations for the H_3 molecule show that, even in this case, our AFQMC method can be successfully applied. However the fermion sign problem appears and this indicates that a modified AFQMC algorithm, to cope with this difficulty, is probably required for applications to more interesting systems, containing many electrons.

The outline of this thesis is the following:

In Chap. 1 some of the most popular QMC methods, Variational Monte Carlo (VMC), Green's Function Monte Carlo (GFMC) and Diffusion Monte Carlo (DMC) are reviewed. Then the HST formalism is presented in general terms, with the introduction of auxiliary fields, and some arguments are put forward about the potential advantages of the AFQMC method in comparison with other stochastic approaches.

In Chap. 2 this technique is specifically developed in a way suitable for our particular problem with a Coulomb electron-electron interaction.

Chap. 3 contains a brief discussion about the technicalities of our practical algorithm and about the methods one can adopt to perform the functional integral

by efficiently sampling the auxiliary fields.

In Chap. 4 the serious fluctuation problem is illustrated, the basic causes which give rise to this difficulty are discussed and a possible solution to the problem, i.e. a different formulation of the AFQMC technique, is introduced, together with some technical improvements we have developed in order to make the method really efficient and practically applicable.

In Chap. 5 we give a concise but exhaustive description of many numerical tests performed to check the reliability of our algorithm, by comparing AFQMC simulations results for the H_2 molecule with exact (full CI) corresponding data. Finally, our calculations of several physical properties of H_2 and H_3 are presented.

Chapter 1

Stochastic methods for the fermionic ground state problem

1.1 Variational Monte Carlo

The variational method has proved to be a very useful way of computing ground state properties of many-body systems. Conceptually it is quite simple. The variational principle tells us that, for any many-body trial function $\psi_T(\mathbf{R})$ (here $\mathbf{R} = \{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N\}$ refers to the coordinates of the N particles), the variational energy E_T defined as:

$$E_T = \frac{\int d\mathbf{R}\psi_T(\mathbf{R})\hat{H}\psi_T(\mathbf{R})}{\int d\mathbf{R}|\psi_T(\mathbf{R})|^2}, \tag{1.1}$$

will be a minimum when $|\psi_T\rangle$ is the ground state solution of the Schrödinger equation $\hat{H}|\psi_0\rangle = E_0|\psi_0\rangle$. The variational method then consists of constructing a family of functions $\psi_T(\mathbf{R}, \mathbf{a})$ and optimizing the parameters $\mathbf{a}$ so that the energy (1.1) is minimized for $\mathbf{a} = \mathbf{a}_0$. The variational energy is a rigorous upper bound to the ground state energy and, if the family of functions is chosen well, then $\psi_T(\mathbf{R}, \mathbf{a}_0)$ will be a good approximation to the ground state wave function. For example, for Fermi liquids, the following Jastrow form for the trial wave function

is widely used:

$$\psi_T(\mathbf{R}) = \psi_D(\mathbf{R}) \prod_{\alpha \neq \beta} f(|\mathbf{r}_\alpha - \mathbf{r}_\beta|) = \psi_D(\mathbf{R}) e^{-\frac{1}{2}\sum_{\alpha \neq \beta} J(|\mathbf{r}_\alpha - \mathbf{r}_\beta|)}, \tag{1.2}$$

where $\psi_D(\mathbf{R})$, the ideal Fermi gas wave function, i.e. a determinant of plane waves, is multiplied by a product of two-particle correlation functions $f(|\mathbf{r}_\alpha - \mathbf{r}_\beta|)$. This wave function was introduced by Jastrow[7] et al.[43,44] who generalized an expression originally developed by Bijl[45] as a good trial function for a Bose liquid at zero temperature. We note that the square of the term $e^{-\frac{1}{2}\sum_{\alpha \neq \beta} J(|\mathbf{r}_\alpha - \mathbf{r}_\beta|)}$ is completely equivalent to the Boltzmann distribution of a classical system with $J(r)$ replaced by the particle-particle interaction potential over $K_B T$. However, in contrast to the classical situation, now the "potential" $J(r)$ is varied to minimize the energy in (1.1). Optimization of a Jastrow wave function (1.2), with respect to some free parameters in the the "potential" $J(r)$, by variational methods often produces a good approximation for ground state wave functions. In practice the main task is to evaluate multidimensional integrals to get expectation values and, in particular, to calculate the variational energy (1.1).

The Monte Carlo algorithm[46], which was invented to calculate properties of classical statistical systems, is an extremely powerful way to compute multi-dimensional integrals. In particular, for quantum systems, an algorithm which produces configurations with a probability proportional to the square of the wave function, is required. Then any measurable quantity can be written as an average over such configurations. Let us suppose $\hat{O}$ is an operator and we wish to compute its expectation value defined as

$$\langle \hat{O} \rangle = \frac{\int d\mathbf{R} \psi_T(\mathbf{R}) \hat{O} \psi_T(\mathbf{R})}{\int d\mathbf{R} |\psi_T(\mathbf{R})|^2}. \tag{1.3}$$

Let $\{\mathbf{R}_i\}$ be a set of points drawn from the probability distribution:

$$p(\mathbf{R}) = \frac{|\psi_T(\mathbf{R})|^2}{\int d\mathbf{R}|\psi_T(\mathbf{R})|^2} , \tag{1.4}$$

where the integral in the denominator serves here merely to normalize $p(\mathbf{R})$. Then, for any function $f(\mathbf{R})$, the "central limit" theorem [47] of probability gives that:

$$\lim_{M\to\infty} \frac{1}{M} \sum_{i=1}^{M} f(\mathbf{R}_i) = \frac{\int d\mathbf{R} f(\mathbf{R})|\psi_T(\mathbf{R})|^2}{\int d\mathbf{R}|\psi_T(\mathbf{R})|^2} . \tag{1.5}$$

In particular one has:

$$\lim_{M\to\infty} \frac{1}{M} \sum_{i=1}^{M} \psi_T^{-1}(\mathbf{R}_i) O(\mathbf{R}_i) \psi_T(\mathbf{R}_i) = \langle \hat{O} \rangle . \tag{1.6}$$

If $\hat{O} = \hat{H}$ then

$$\langle \hat{H} \rangle = \lim_{M\to\infty} \frac{1}{M} \sum_{i=1}^{M} E_L(\mathbf{R}_i) , \tag{1.7}$$

where we have introduced the so-called "local energy": $E_L(\mathbf{R}_i) \equiv \hat{H}\psi_T/\psi_T$. We observe that a good approximate trial function, containing whatever information is known about the exact wave function, yields averages with low statistical uncertainties. In fact, if ψ_T is a good approximation to ψ_0, then $E_L(\mathbf{R}) \to E_0$, that is it becomes nearly independent on $\mathbf{R}$.

The Monte Carlo algorithm is a biased random walk in configuration space. Usually it is implemented by moving one particle at a time to a new position. That move is either accepted or rejected depending on the magnitude of the trial function at the *new* position $\mathbf{R}'$ compared with the magnitude at the *old* position $\mathbf{R}$: if $|\psi_T(\mathbf{R}')|^2 \geq |\psi_T(\mathbf{R})|^2$ the new point $\mathbf{R}'$ is accepted. Otherwise it is accepted with a probability q given by:

$$q = \frac{|\psi_T(\mathbf{R}')|^2}{|\psi_T(\mathbf{R})|^2} . \tag{1.8}$$

In general the VMC algorithm [48,49] is quite simple to program and test, and follows very closely a Monte Carlo simulation of a classical system. VMC calculations typically recover about 15-80% of the correlation energy. Through this method it is possible to optimize at most a few wave function parameters, since a single calculation of the variational energy (1.1) takes a significant amount of computer time and this must be repeated several times before an optimal set of parameters is found. Recently [50] a different variational procedure was developed, wherein the variance of the local energy $E_L(\mathbf{R}_i)$ is minimized. In this way one can optimize wave functions with a relatively large number (~ 100) of variational parameters: in fact the configurations over which the optimization is performed are fixed and so a correlated sampling is used to arrive at an optimal set of parameters. Even though this procedure allows a more efficient optimization (for instance 99% of the correlation energy can be obtained for the Be atom), obviously a suitable choice of the trial wave function form remains a critical ingredient in this VMC approach too.

1.2 Diffusion and Green's Function Monte Carlo

The history of "exact" Quantum Monte Carlo calculations to "solve" the Schrödinger equation goes back more years than might be expected. Metropolis [51] reported the first such computation in 1949 and gave credit to Fermi for suggesting that the similarity between the Schrödinger equation and the classical diffusion equation could be exploited to produce a solution of the Schrödinger equation. Anyway the available vacuum tube computers were not up to doing anything more substantial that a harmonic oscillator, so the field marked time until 1962.

A publication by Kalos[52] signaled the development of what is now known as *Green's Function Monte Carlo* (GFMC). The *Diffusion Monte Carlo* (DMC) method was first described by Anderson[14] and developed in parallel by Anderson[53], Kalos[54] and Ceperley[15] in the early 1980's. Diffusion Monte Carlo and Green's Function Monte Carlo are essentially equivalent. Both techniques use a Green's function to solve the Schrödinger equation with an integral calculated by Monte Carlo. Basically the ground state properties are evaluated by performing the imaginary time propagation $e^{-\beta H}\psi_T$. This can be obtained in different ways.

The DMC method[18,21] uses the time-dependent Schrödinger equation:

$$i\frac{\partial\psi}{\partial t} = \left[-\frac{1}{2}\nabla^2 + V(\mathbf{R})\right]\psi = \hat{H}\psi \;, \tag{1.9}$$

where a.u. are used and $V(\mathbf{R})$ is the potential energy of the system. Eq. (1.9) can be represented in imaginary time ($\tau \equiv it$):

$$-\frac{\partial\psi}{\partial\tau} = \left[-D\nabla^2 + V(\mathbf{R}) - E_T\right]\psi \;, \tag{1.10}$$

with $D = 1/2$ and a constant energy offset, E_T, was introduced for convenience to alter the zero of energy without affecting the properties calculated from the solution of the Schrödinger equation. Eq. (1.10) has solutions of the form:

$$\psi(\mathbf{R},\tau) = \sum_\alpha \phi_\alpha(\mathbf{R}) e^{-(E_\alpha - E_T)\tau} \;. \tag{1.11}$$

So, for positive real τ, the exponential factor causes the states with the larger eigenvalues to decay away, leaving only the state with the smallest eigenvalue, after long τ (see also Introduction).

Solving Eq. (1.10) by a random-walk process with branching is inefficient, because the branching rate (which is proportional to the Coulomb potential $V(\mathbf{R})$)

can diverge to $\pm\infty$. This leads to large fluctuations in the number of diffusers, and to slow convergence when calculating averages such as $\langle V(\mathbf{R}) \rangle$ and hence the ground state energy. However the fluctuations and the consequent statistical uncertainties, can be greatly reduced by the technique of *importance sampling*[13]. Essentially the exact wave function is multiplied by a trial wave function ϕ_T to obtain a new function f:

$$f(\mathbf{R},\tau) \equiv \phi_T(\mathbf{R})\psi(\mathbf{R},\tau) \ . \tag{1.12}$$

Substituting f/ϕ_T for ψ in Eq. (1.10), we obtain the following Fokker-Planck equation for $f(\mathbf{R},\tau)$:

$$-\frac{\partial f}{\partial \tau} = -D\nabla^2 f + (E_L(\mathbf{R}) - E_T) f + D\nabla\left[f\mathbf{F}_Q(\mathbf{R})\right] \ , \tag{1.13}$$

where $E_L(\mathbf{R}) \equiv \hat{H}\phi_T(\mathbf{R})/\phi_T(\mathbf{R})$ is the local energy obtained from the trial function, and

$$\mathbf{F}_Q(\mathbf{R}) \equiv \nabla \ln |\phi_T(\mathbf{R})|^2 = \frac{2\nabla\phi_T(\mathbf{R})}{\phi_T(\mathbf{R})} \ , \tag{1.14}$$

plays the role of a "quantum force". The terms on the R.H.S of Eq. (1.13) may be identified as a *diffusion* term, a *branching* (source/sink) term and a *drift* (advection) term, respectively. As usual, the Monte Carlo simulation of the Fokker-Planck Eq. (1.13) is carried out by representing the "density" f by particles that take random steps to simulate the diffusion, take directed steps to simulate the advection (the quantum force pushes the particles toward regions of higher importance, that is higher ϕ_T), and are multiplied or eliminated to model sources and sinks. The asymptotic solution to Eq. (1.13) is:

$$f(\mathbf{R},\tau) = \phi_T(\mathbf{R})\psi_0(\mathbf{R},\tau)\, e^{-(E_0 - E_T)\tau} \ . \tag{1.15}$$

Then the ground state energy may be calculated[18] by using the average value of the local energy, at large imaginary times:

$$\langle \hat{H} \rangle = E_0 = \lim_{\beta\to\infty} \frac{\int d\mathbf{R} f(\mathbf{R},\beta) E_L(\mathbf{R})}{\int d\mathbf{R} f(\mathbf{R},\beta)} = \lim_{M\to\infty} \frac{1}{M} \sum_{i=1}^{M} E_L(\mathbf{R}_i) , \qquad (1.16)$$

where M is the number of points $\{\mathbf{R}_i\}$ distributed according to $f(\mathbf{R},\beta)$. Again statistical accuracy is greatly improved if a good trial wave function is chosen. Obtaining the expectation value for a generic operator $\hat{O}$, which does not commute with $\hat{H}$, is somewhat more difficult and less accurate. By assuming that the trial function ϕ_T is close to the ground state wave function ψ_0, one can consider the approximate estimate[15,55]:

$$\langle \hat{O} \rangle = \frac{2}{M} \left[\sum_{i=1}^{M} O(\mathbf{R}_i) \right]_f - \frac{1}{M} \left[\sum_{i=1}^{M} O(\mathbf{R}_i) \right]_{\phi_T^2} + o(\varepsilon^2) , \qquad (1.17)$$

where the subscripts indicate that the points $\{\mathbf{R}_i\}$ are distributed according to f or to ϕ_T^2, and $\varepsilon \propto (\psi_0 - \phi_T)$.

As far as the GFMC[13,15] method is concerned, one considers the time-independent Schrödinger equation:

$$\left[-\frac{1}{2}\nabla^2 + V(\mathbf{R}) \right] \psi(\mathbf{R}) = E\psi(\mathbf{R}) . \qquad (1.18)$$

This may be rewritten in its integral form:

$$\psi(\mathbf{R}) = E \int d\mathbf{R}' G(\mathbf{R},\mathbf{R}') \psi(\mathbf{R}') , \qquad (1.19)$$

where $G(\mathbf{R},\mathbf{R}')$, the Green's function, satisfies the equation:

$$\left[-\frac{1}{2}\nabla^2 + V(\mathbf{R}) \right] G(\mathbf{R},\mathbf{R}') = \delta(\mathbf{R}-\mathbf{R}') , \qquad (1.20)$$

and the boundary conditions of the problem. Let a succession of functions be defined for some initial $\psi^{(0)}(\mathbf{R})$ by:

$$\psi^{(n+1)}(\mathbf{R}) = E \int d\mathbf{R}' G(\mathbf{R}, \mathbf{R}') \psi^{(n)}(\mathbf{R}') . \tag{1.21}$$

When the spectrum of the Hamiltonian is discrete near the ground state ψ_0 of the Schrödinger equation, then $\psi_0(\mathbf{R})$ is the limiting value of $\psi^{(n)}(\mathbf{R})$ for large n. It is possible to devise a Monte Carlo method (in the general sense of a random sampling algorithm) which produces populations drawn in turn from the successive $\psi^{(n)}$. In practice, in Eq. (1.21), the exact eigenvalue of the ground state is substituted by a trial eigenvalue E_T and the main technical problem lies in constructing a method [15] for efficiently sampling $G(\mathbf{R}, \mathbf{R}')$, which, in general, does not exhibit an analytic expression. In this case too, some importance sampling technique has to be introduced to get small statistical errors.

In applying the DMC or GFMC approaches a serious problem arises when fermions are considered. In fact, in this case, the wave function must be anti-symmetric. This requirement is troublesome if we wish to solve the Schrödinger equation as a diffusion equation, because the density of diffusers, or "walkers", must be positive. There are at least two ways of handling this problem.

One possibility is to fix the nodes by selecting the best possible trial wave function and to perform a guided random walk in which no particles are allowed to cross the nodal surfaces of the trial wave function. Within each nodal region the problem is effectively bosonic. This "fixed-node" approximation [16,18] is variational in nature, i.e. the energy thus derived is an upper bound to the true energy.

A second possibility, the "release-node" approach [20], is to perform a transient estimate, starting with the best fixed-node distribution and releasing the

fixed-node restrictions. As points cross nodes cancellations will give rise to statistical errors which increase exponentially with imaginary time (the *fermion sign problem*). In fact the lowest energy solution of the diffusion equation, on the space of configurations of the system, is nodeless: this solution is the boson ground state. Therefore the number of "positive" walkers and the number of "negative" walkers increase exponentially compared to the difference between them, which is the number we need to calculate. The rate of this exponential increase equals the difference between the boson and fermion ground state energies or, alternatively, the difference between the energy of the nodeless solution of the diffusion equation and that of the lowest antisymmetric solution. In practice, in the release-node method, only the early stages of the relaxation can be calculated before numerical instability spoils the quality of the simulation.

Other techniques[56,57] have been devised to treat fermions by using the DMC or GFMC methods, but all of them seem to contain some drawbacks, or are very complex and expensive, so they are actually feasible only for few-body systems.

1.3 The Auxiliary Field method

1.3.1 General description

In the following we introduce a functional integral formulation for interacting fermions. We describe a formalism to calculate ground state properties of a many-body system by using a classical statistical method.

Let us consider the generic Hamiltonian $\hat{H} = \hat{K} + \hat{V}$, where $\hat{K}$ is a *one-body*

operator (generally consisting of kinetic and external potential terms), and $\hat{V}$ is a *two-body* interaction term due to electron-electron correlation. We can write $\hat{H}$ in second quantization (for the sake of clearness we use the coordinate representation with basis functions given by delta functions, but, obviously, the procedure holds for a generic representation):

$$\hat{H} = \sum_{i,j} K_{ij} c_i^\dagger c_j + \frac{1}{2} \sum_{i,j} V_{ij} c_i^\dagger c_j^\dagger c_j c_i \ , \tag{1.22}$$

where $V_{ij} = V(|\mathbf{r}_i - \mathbf{r}_j|)$, and $c_i^\dagger$, c_i are the creation and annihilation operators of a particle at position i (for the moment we omit to explicitly write spin indices). As we are only interested in *ground state* properties, instead of considering the thermodynamic partition function $Z = \mathrm{Tr}\left(e^{-\beta H}\right)$, where β indicates the inverse temperature, we consider a *pseudo partition function*[28,58]:

$$Q = \langle \psi_T | \, e^{-\beta H} \, | \psi_T \rangle \ , \tag{1.23}$$

where ψ_T is a trial wave function and β can be thought of as an imaginary time. We have already observed (see Introduction) that, if ψ_0 , the ground state of $\hat{H}$, has a non-vanishing overlap with the trial wave function ψ_T, the imaginary time propagator $e^{-\beta \hat{H}}$, for $\beta \to \infty$, projects from ψ_T its component along ψ_0; therefore Q behaves asymptotically as the *true* partition function Z:

$$\lim_{\beta \to \infty} Q = |\langle \psi_0 | \psi_T \rangle|^2 \, e^{-\beta E_0} \, . \tag{1.24}$$

Then, in terms of Q, the ground state energy is given by

$$E_0 = \lim_{\beta \to \infty} \left(-\frac{1}{\beta} \ln Q \right) . \tag{1.25}$$

More general expressions for other ground state expectation values can be obtained by differentiating Eq. (1.25) with respect to appropriate external fields coupled to

the quantity of interest. In fact, by using the Hellmann-Feynman theorem, we can calculate the ground state expectation values of a general operator $\hat{O}$, $\langle\psi_0|\,\hat{O}\,|\psi_0\rangle$, by differentiating, with respect to λ, the ground state energy of the corresponding perturbed Hamiltonian $\hat{H}+\lambda\hat{O}$:

$$\langle\psi_0|\,\hat{O}\,|\psi_0\rangle = \left(\frac{\partial}{\partial\lambda}E_0(\lambda)\right)_{\lambda=0} = -\frac{1}{\beta}\lim_{\beta\to\infty}\left(\frac{\partial}{\partial\lambda}\ln Q_\lambda\right)_{\lambda=0}. \tag{1.26}$$

Hence, in this scheme, the pseudo partition function Q can be considered as the "generator" of all the ground state correlation functions. However, a direct evaluation of Q is actually a difficult task, since $\hat{H}$ contains two-body contributions and ψ_T is a many-body wave function. Now we show that, by using a suitable transformation and introducing auxiliary fields σ, Q can be rewritten as:

$$Q = \frac{1}{C}\int d\sigma\, G(\sigma)\,\langle\psi_T|\,\hat{U}(\sigma)\,|\psi_T\rangle\ , \tag{1.27}$$

where $G(\sigma)$ is a Gaussian weight and $\hat{U}(\sigma)$ is a one-body, auxiliary field-dependent, operator. For each auxiliary field configuration σ, propagation $\hat{U}(\sigma)\,|\psi_T\rangle$ may be easily performed because $\hat{U}$ does not contain two-body terms. Therefore, in principle, Q can be explicitly computed, together with all ground state expectation values, via Eq. (1.26).

We start by splitting the total imaginary time propagator into a product of P short time propagators and by applying the Trotter approximation [59] to each of them:

$$e^{-\beta\hat{H}} = \left(e^{-\Delta\tau\hat{H}}\right)^P = \left(e^{-\frac{\Delta\tau}{2}\hat{T}}e^{-\Delta\tau\hat{V}_{\rm tot}}e^{-\frac{\Delta\tau}{2}\hat{T}}\right)^P + o(\Delta\tau^3)\,, \tag{1.28}$$

with $\Delta\tau=\beta/P$ and, usually, Trotter decomposition separates the kinetic term $\hat{T}$, from $\hat{V}_{\rm tot}=\hat{H}-\hat{T}$, the remaining component consisting of external and electron-electron potentials. The short time propagator $e^{-\frac{\Delta\tau}{2}\hat{T}}e^{-\Delta\tau\hat{V}_{\rm tot}}e^{-\frac{\Delta\tau}{2}\hat{T}}$ is clearly

hermitian and positive definite. Then we rewrite the approximated short time propagator of the Hamiltonian $\hat{H}$ as an *exact* short time propagator of an *equivalent* Hamiltonian $\tilde{H}$:

$$e^{-\Delta\tau\tilde{H}} = e^{-\frac{\Delta\tau}{2}T} e^{-\Delta\tau V_{\rm tot}} e^{-\frac{\Delta\tau}{2}T} , \tag{1.29}$$

with $\tilde{H} = \hat{H} + o(\Delta\tau^2)$. In conclusion all the calculations obtained by using the Trotter approximation give exact ground state properties of the effective Hamiltonian $\tilde{H}$ which differ at most by $o(\Delta\tau^2)$ from the desired ground state properties of the true Hamiltonian $\hat{H}$. As we have already said, the evaluation of the propagation performed by operator (1.29) is numerically tractable when $\hat{H}$ contains *only one-body* operators. Therefore we can introduce the *Hubbard-Stratonovich Transformation*[41,42] (HST) which exactly aims at reducing the exponential of a two-body operator (e.g. the term involving V_{ij} in Eq. (1.22)) to a functional integral, over an auxiliary field, where only exponentials of one-body operators are present.

First of all we can rewrite Hamiltonian (1.22) by anticommuting the creation and annihilation operators in the normal-ordered two-body interaction:

$$\hat{H} = \sum_{i,j} K_{ij} c_i^\dagger c_j + \frac{1}{2}\sum_{i,j} V_{ij}\hat{\rho}_i\hat{\rho}_j - \frac{1}{2}\sum_i V_{ii}\hat{\rho}_i , \tag{1.30}$$

where $\hat{\rho}_i$ is the fermion density operator and the last term is an unphysical self-interaction contribution which can be temporarily grouped with the one-body term K and which will be removed later: $K'_{ij} = K_{ij} - \frac{1}{2}V_{ij}\delta_{ij}$,

$$\hat{H} = \sum_{i,j} K'_{ij} c_i^\dagger c_j + \frac{1}{2}\sum_{i,j} V_{ij}\hat{\rho}_i\hat{\rho}_j . \tag{1.31}$$

Now we apply the HST to the exponential of the two-body operator, which appears in the Trotter decomposition (1.28), $e^{-\frac{\Delta\tau}{2}\sum_{i,j} V_{ij}\hat{\rho}_i\hat{\rho}_j}$ (see Appendix A):

$$e^{-\frac{\Delta\tau}{2}\sum_{i,j} V_{ij}\hat{\rho}_i\hat{\rho}_j} = e^{\frac{\Delta\tau}{2}\sum_{i,j} W_{ij}\hat{\rho}_i\hat{\rho}_j}$$
$$= \left(\Delta\tau \det W_{ij}^{-1}\right)^{\frac{1}{2}} \int \prod_i \frac{d\tilde{\sigma}_i}{\sqrt{2\pi}} e^{-\frac{\Delta\tau}{2}\sum_{i,j} W_{ij}^{-1}\tilde{\sigma}_i\tilde{\sigma}_j} e^{\Delta\tau\sum_i \tilde{\sigma}_i\hat{\rho}_i}, \tag{1.32}$$

where $W_{ij} \equiv -V_{ij}$ and, for simplicity, we have assumed that V_{ij} is a definite negative matrix and, therefore, W_{ij} a positive definite one. In fact (see Appendix A) the previous transformation can be directly applied only for *negative definite* two-body operators contained in the Hamiltonian $\hat{H}$. While this property is certainly not true in general (for example it does not hold for the repulsive electron-electron interaction), nevertheless HST can always be performed for quite general two-body fermionic operators, as we will show in the following Chapter.

Eq. (1.32) introduces $\tilde{\sigma}$ variables as auxiliary fields with dimensions of potentials. In order to avoid using the inverse matrix W_{ij}^{-1} it is possible to change integration variables, by defining:

$$\sigma_i \equiv \sum_j W_{ij}^{-1}\tilde{\sigma}_j . \tag{1.33}$$

Therefore σ assumes the dimensions of densities and Eq. (1.32) becomes:

$$e^{\frac{\Delta\tau}{2}\sum_{i,j} W_{ij}\hat{\rho}_i\hat{\rho}_j} = \left(\Delta\tau \det W_{ij}\right)^{\frac{1}{2}} \int \prod_i \frac{d\sigma_i}{\sqrt{2\pi}} e^{-\frac{\Delta\tau}{2}\sum_{i,j} W_{ij}\sigma_i\sigma_j} e^{\Delta\tau\sum_{i,j} W_{ij}\sigma_i\hat{\rho}_j} . \tag{1.34}$$

Obviously we have to perform the transformation (1.34) at each time step of the imaginary time propagation (1.28). Hence we introduce a discrete time index l in σ variables and we are able to write $e^{\frac{\beta}{2}\sum_{i,j} W_{ij}\hat{\rho}_i\hat{\rho}_j}$ as:

$$\frac{1}{C}\int \prod_{l=1}^{P}\prod_i d\sigma_i(l) e^{-\frac{\Delta\tau}{2}\sum_{l=1}^{P}\sum_{i,j} W_{ij}\sigma_i(l)\sigma_j(l)} e^{\Delta\tau\sum_{l=1}^{P}\sum_{i,j} W_{ij}\sigma_i(l)\hat{\rho}_j} , \tag{1.35}$$

where C is a normalization constant given by

$$\frac{1}{C} = \prod_{l=1}^{P} \frac{(\Delta\tau \det W_{ij})^{\frac{1}{2}}}{\prod_i (2\pi)^{\frac{1}{2}}} . \tag{1.36}$$

Now if we consider again the complete Hamiltonian $\hat{H} = \hat{T} + \hat{V}_{\rm tot}$, with:

$$\hat{T} = \sum_{i,j} T_{ij} c_i^\dagger c_j \ , \tag{1.37}$$

and:

$$\begin{aligned} \hat{V}_{\rm tot} &= \sum_i V_i^{\rm ext} \hat{\rho}_i + \frac{1}{2} \sum_{i,j} V_{ij} \hat{\rho}_i \hat{\rho}_j - \frac{1}{2} \sum_i V_{ii} \hat{\rho}_i \\ &= \sum_i \left(V_i^{\rm ext} - \frac{1}{2} V_{ii} \right) \hat{\rho}_i - \frac{1}{2} \sum_{i,j} W_{ij} \hat{\rho}_i \hat{\rho}_j \ , \end{aligned} \tag{1.38}$$

where $V^{\rm ext}$ is a generic external potential, then we can rewrite the propagation (1.28) by using the previous HST relations:

$$\begin{aligned} e^{-\beta \hat{H}} &\simeq \left(e^{-\frac{\Delta\tau}{2}\hat{T}} e^{-\Delta\tau \hat{V}_{\rm tot}} e^{-\frac{\Delta\tau}{2}\hat{T}} \right)^P \\ &= \frac{1}{C} \int \prod_{l=1}^{P} \prod_i d\sigma_i(l) e^{-\frac{\Delta\tau}{2} \sum_{l=1}^{P} \sum_{i,j} W_{ij}\, \sigma_i(l) \sigma_j(l)} \\ &\times \prod_{l=1}^{P} \left\{ e^{-\frac{\Delta\tau}{2}\hat{T}} e^{-\Delta\tau \left(\sum_i V_i^{\rm ext} \hat{\rho}_i - \sum_{i,j} W_{ij}\, \sigma_i(l) \hat{\rho}_j \right)} e^{-\frac{\Delta\tau}{2}\hat{T}} \right\} . \end{aligned} \tag{1.39}$$

Thus, the evolution operator is the functional integral, over auxiliary fields σ, of the evolution operators for *one-body, time-dependent* Hamiltonians, whose potential energy contribution is given by

$$\hat{V}_\sigma(l) = \sum_i V_i^{\rm ext} \hat{\rho}_i - \sum_{i,j} W_{ij} \sigma_i(l) \hat{\rho}_j \ , \tag{1.40}$$

and which are weighted by a Gaussian factor. In the last relation we omitted the *self-energy* term, $\frac{1}{2}\sum_i V_{ii}\hat{\rho}_i$, since it acts as a constant potential in the one-body

propagation and this does not affect the computed estimators of ground state properties (see the following Section).

The main potential advantages of the AFQMC method over other stochastic techniques (DMC or GFMC) are the following:

i) – Two-body interaction terms are replaced with random auxiliary fields which act as external potentials on the particles, therefore the interacting problem is replaced by a sum over an ensemble of non-interacting systems in a set of random time-varying external fields and one can easily apply all typical *one-body* techniques (for instance local and non-local pseudopotentials) in a natural way. The situation is quite different for the ordinary QMC methods in which, for example, using non-local pseudopotentials (that destroy the diffusive character of the Schrödinger equation) introduces serious difficulties[60–62], even though, very recently, a significant progress was made in this direction[63].

ii) – GFMC or DMC methods accurately describe boson systems but for fermions a difficulty arises. In fact, as we have seen in Section 1.2, the calculation is unstable because, through a long imaginary time propagation, the fermionic component becomes undetectable from a numerical point of view (the fermion sign problem). On the contrary, in the AFQMC method, the *antisymmetric* property of the fermion wave function is preserved at any time of field evolution. In fact a Slater determinant trial wave function evolves into another Slater determinant for each auxiliary field configuration sampled. In this approach observables are calculated as averages over the set of auxiliary field configurations. As we will discuss in Section 1.3.4, the difficulty here is that the quantity to be averaged is not always positive (the statistical weight may be not positive definite). In practice the *fermion sign problem* reappears since, as $\beta \to \infty$, the number of positive terms

can nearly equal the number of negative terms and the difference, which is the quantity we are interested in, becomes very small compared to the total number of terms. Nevertheless, in this situation, the fermion sign problem seems to be less dramatic than in GFMC or DMC approaches. In fact one can show (this aspect was extensively studied by Sorella et al.[31,38], by Fahy and Hamann[36,37], and by Loh et al.[35]), by theoretic arguments and numerical evidence, that, in many non trivial cases, the AFQMC method is stable for arbitrarily large imaginary time and that the fermion sign problem can be often circumvented with negligible or small errors in the calculated physical quantities.

iii) – In the AFQMC method the estimator calculation of general observables (see next Section and Section 2.2) is a simple task, while, in the other QMC schemes, as we discussed in Section 1.2, the estimate of properties involving operators which do not commute with the total Hamiltonian $\hat{H}$, cannot be done without considerably degrading the statistical efficiency of the algorithm, or introducing further approximations through the so-called *extrapolated estimate*[15,55].

Finally, the AFQMC method is exact, apart from statistical errors we can, in principle, reduce as small as we like, but, compared with a full Configuration Interaction approach, we expect it requires computer resources which do not grow exponentially. In fact, in the AFQMC procedure, in practice, only the calculation of determinants depends on the number of electrons. Therefore a cost increasing with the 3^{rd} power (or less) of the number of electrons (see Chap. 2) has to be foreseen.

1.3.2 Calculation of ground state properties

In the previous Section we used HST to rewrite the pseudo partition function Q as a multidimensional integral over classical auxiliary fields:

$$Q = \langle \psi_T | \, e^{-\beta \hat{H}} \, | \psi_T \rangle = \frac{1}{C} \int d\sigma \, G(\sigma) \, \langle \psi_T | \, \hat{U}(\sigma) \, | \psi_T \rangle \ , \tag{1.41}$$

where $G(\sigma)$ indicates the Gaussian weighting factor:

$$G(\sigma) = e^{-\frac{\Delta\tau}{2} \sum_l \sum_{i,j} W_{ij} \, \sigma_i(l) \sigma_j(l)} \ , \tag{1.42}$$

and $\hat{U}(\sigma)$ is the one-body, field-dependent, propagator:

$$\begin{aligned} \hat{U}(\sigma) &= \prod_{l=1}^{P} \hat{U} \left[\sigma(l) \right] \\ &= \prod_{l=1}^{P} \left\{ e^{-\frac{\Delta\tau}{2} \hat{T}} e^{-\Delta\tau \left(\sum_i V_i^{\mathrm{ext}} \hat{\rho}_i - \sum_{i,j} W_{ij} \, \sigma_i(l) \hat{\rho}_j \right)} e^{-\frac{\Delta\tau}{2} \hat{T}} \right\}. \end{aligned} \tag{1.43}$$

At this point the quantum problem would be solved if an exact numerical evaluation of the multidimensional integral (1.41) were possible. Unfortunately this is not the case because the functional Q contains a prohibitively large number of variables. Nevertheless we can use a statistical approach by interpreting the functional Q as a *classical* partition function of the variables σ, and apply a statistical method for evaluating Q and related physical quantities. In fact we can write:

$$Q = \frac{1}{C} \int d\sigma e^{-V(\sigma)/K_B T} \ , \tag{1.44}$$

by considering Q as a classical partition function of a system of classical fields σ, that "interact", at an *effective temperature* $K_B T = 1$, through a potential:

$$\begin{aligned} V(\sigma) &= -\ln G(\sigma) - \ln \langle \psi_T | \, \hat{U}(\sigma) \, | \psi_T \rangle + \text{const.} \\ &= \frac{\Delta\tau}{2} \sum_l \sum_{i,j} W_{ij} \sigma_i(l) \sigma_j(l) - \ln \langle \psi_T | \, \hat{U}(\sigma) \, | \psi_T \rangle + \text{const.} \, . \end{aligned} \tag{1.45}$$

Here we assume that $\langle \psi_T | \hat{U}(\sigma) | \psi_T \rangle$ is always positive. In general this is not true (it is the practical manifestation of the *fermion sign problem* - see Section 1.3.4) because the propagated many-body wave function, $\hat{U}(\sigma) | \psi_T \rangle$, may have a negative overlap with the initial trial wave function.

Now the ground state expectation value of a given operator $\hat{O}$ is defined as:

$$\begin{aligned} \langle \hat{O} \rangle &\equiv \frac{\langle \psi_0 | \hat{O} | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle} \\ &= \lim_{\beta \to \infty} \frac{\langle \psi_T | e^{-\frac{\beta}{2}\hat{H}} \hat{O} e^{-\frac{\beta}{2}\hat{H}} | \psi_T \rangle}{\langle \psi_T | e^{-\beta \hat{H}} | \psi_T \rangle} . \end{aligned} \tag{1.46}$$

By applying HST into the previous expression, one easily obtains:

$$\langle \hat{O} \rangle = \frac{\int d\sigma E_{\hat{O}}(\sigma) e^{-V(\sigma)}}{\int d\sigma e^{-V(\sigma)}} , \tag{1.47}$$

with the *estimator* $E_{\hat{O}}(\sigma)$ given by:

$$E_{\hat{O}}(\sigma) = \frac{\langle \psi_T | \hat{U}_\sigma(\beta, \beta/2) \hat{O} \hat{U}_\sigma(\beta/2, 0) | \psi_T \rangle}{\langle \psi_T | \hat{U}_\sigma(\beta, 0) | \psi_T \rangle} . \tag{1.48}$$

In practice the ground state expectation value of the quantum operator $\hat{O}$ can be computed by calculating a well defined *classical average*.

An equivalent expression, which, in principle, should give better statistics for measuring $\langle \hat{O} \rangle$ when the auxiliary field integration is carried out by stochastic methods, may be obtained by averaging over all possible ways of inserting the $\hat{O}$ operator:

$$\langle \hat{O} \rangle = \frac{\int_0^\beta d\tau \langle \psi_T | e^{-(\beta-\tau)\hat{H}} \hat{O} e^{-\tau \hat{H}} | \psi_T \rangle}{\beta \langle \psi_T | e^{-\beta \hat{H}} | \psi_T \rangle} . \tag{1.49}$$

In this more general case the estimator, in Eq. (1.47), becomes:

$$E_{\hat{O}}(\sigma) = \frac{1}{\beta} \int_0^\beta d\tau \frac{\langle \psi_T | \hat{U}_\sigma(\beta, \tau) \hat{O} \hat{U}_\sigma(\tau, 0) | \psi_T \rangle}{\langle \psi_T | \hat{U}_\sigma(\beta, 0) | \psi_T \rangle} . \tag{1.50}$$

The same result can be derived using directly Eq. (1.26):

$$\frac{\langle\psi_0|\,\hat{O}\,|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle} = \lim_{\beta\to\infty}\left(-\frac{1}{\beta}\frac{\partial}{\partial\lambda}\ln Q_\lambda\right)_{\lambda=0} = \quad Q^{-1}\frac{1}{C}\int d\sigma E_{\hat{O}}(\sigma)e^{-V(\sigma)}\ , \quad (1.51)$$

with:

$$E_{\hat{O}}(\sigma) = \lim_{\beta\to\infty}\left(-\frac{1}{\beta}\frac{\partial}{\partial\lambda}\ln\langle\psi_T|\,\hat{U}^\lambda(\sigma)\,|\psi_T\rangle\right)_{\lambda=0}\ , \quad (1.52)$$

where the λ-modified propagator $\hat{U}^\lambda(\sigma)$ is obtained by adding to the Hamiltonian (in practice one can add it to the kinetic term) a perturbation $\lambda\hat{O}$. Then $E_{\hat{O}}(\sigma)$ can be calculated, by performing the differentiation with respect to the external perturbation in Eq. (1.52), as an imaginary time average of independent measurements computed at a fixed imaginary time τ:

$$\begin{aligned} E_{\hat{O}}(\sigma) &= \frac{\Delta\tau}{\beta}\frac{\sum_{l=1}^{P} w_l\,\langle\psi_T|\,\hat{U}_\sigma(P,l)\,\hat{O}\hat{U}_\sigma(l,0)\,|\psi_T\rangle}{\langle\psi_T|\,\hat{U}(\sigma)\,|\psi_T\rangle} \quad (1.53)\\ &= \frac{1}{P}\frac{\sum_{l=1}^{P} w_l\,\langle\psi_T|\,\hat{U}_\sigma(P,l)\,\hat{O}\hat{U}_\sigma(l,0)\,|\psi_T\rangle}{\langle\psi_T|\,\hat{U}_\sigma(P,0)\,|\psi_T\rangle} \longrightarrow \frac{1}{\beta}\int_0^\beta d\tau E^\tau_{\hat{O}}(\sigma)\ , \end{aligned}$$

where the last relation holds in the limit $\Delta\tau\to 0$, and:

$$E^\tau_{\hat{O}}(\sigma) = \frac{\langle\psi_T|\,\hat{U}_\sigma(\beta,\tau)\,\hat{O}\hat{U}_\sigma(\tau,0)\,|\psi_T\rangle}{\langle\psi_T|\,\hat{U}_\sigma(\beta,0)\,|\psi_T\rangle}\ , \quad (1.54)$$

with:

$$\hat{U}_\sigma(l,0) = \prod_{l'=0}^{l}\hat{U}\left[\sigma(l')\right]\ , \quad (1.55)$$

$$\hat{U}_\sigma(P,l) = \prod_{l'=0}^{P-l-1}\hat{U}\left[\sigma(P-l')\right]\ , \quad (1.56)$$

$\hat{U}_\sigma(0,0) = \hat{U}_\sigma(P,P) = \hat{I}$, and the weights w_l are $w_l = 1/2$ for $l = 0$, or $l = P$, and $w_l = 1$ otherwise. Actually, when one evaluates the estimator with the imaginary time average (1.53), the contributions coming from imaginary time measurements

close to the initial $\tau = 0$ and the final $\tau = \beta$ imaginary time produce a slow convergence of the physical quantities with respect to the inverse temperature. In fact such measurements are too close to the trial wave function and give a contribution which vanishes as β^{-1}. In order to improve systematically such convergence in β, one can consider an average over an interval which is far from the initial $\tau = 0$ and final time $\tau = \beta$, instead of averaging over all the imaginary time slices. Therefore one can use ($n > 2$):

$$E_{\hat{O}}(\sigma) = \frac{1}{\beta\,(1-2/n)} \int_{\beta/n}^{\beta(1-1/n)} d\tau\, E^{\tau}_{\hat{O}}(\sigma)\,. \tag{1.57}$$

This kind of estimator can be formally obtained by taking the logarithmic derivative of the partition function $Q_{\lambda_T} = \langle \psi_T |\, e^{-\beta \hat{H}_{\lambda_T}}\, | \psi_T \rangle$:

$$-\frac{1}{\beta\,(1-2/n)} \left(\frac{\partial}{\partial \lambda} \ln \langle \psi_T |\, e^{-\beta \hat{H}_{\lambda_T}}\, | \psi_T \rangle \right)_{\lambda=0}\,, \tag{1.58}$$

where $\hat{H}_{\lambda_T} = \hat{H} + \lambda_T \hat{O}$, and now λ_T acts as a time-dependent perturbation:

$$\lambda_T = \begin{cases} \lambda, & \text{if } \beta/n \le \tau \le \beta\,(1-1/n); \\ 0, & \text{otherwise.} \end{cases}$$

This gives, for $\beta \to \infty$, the correct ground state expectation value with an exponential convergence in β. A good value for n could be $n = 4$.

In practical calculations the expectation value of the operator $\hat{O}$ is obtained by evaluating the multidimensional integral (1.51) through a statistical method:

$$\langle \hat{O} \rangle \equiv \frac{\langle \psi_0 |\, \hat{O}\, | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle} = \frac{\int d\sigma E_{\hat{O}}(\sigma) e^{-V(\sigma)}}{\int d\sigma e^{-V(\sigma)}}\,. \tag{1.59}$$

To this end some *importance sampling* scheme must be used, by generating σ configurations according to the probability function $e^{-V(\sigma)}$. For this purpose either a Monte Carlo algorithm or a Molecular Dynamics strategy can be used: one

simply considers the system with σ degrees of freedom in the fictitious classical potential $V(\sigma)$. The variables σ are taken as functions of a formal, continuous time variable s (the *fictitious* time). In this way the statistical evaluation of classical expectation values of estimators, depending on the variables σ, can be expressed as a *temporal* average:

$$\langle \hat{O} \rangle = \frac{\int d\sigma E_{\hat{O}}(\sigma) e^{-V(\sigma)}}{\int d\sigma e^{-V(\sigma)}} = \lim_{s \to \infty} \frac{1}{(s - s_0)} \int_{s_0}^{s} ds' E_{\hat{O}}[\sigma(s')] \ , \tag{1.60}$$

where s_0 is the time needed to reach equilibrium for the Molecular Dynamics Equations, or the Monte Carlo scheme. For infinite fictitious time s, Eq. (1.60) would lead to zero statistical uncertainty. Actually this is not feasible and one has always to consider statistical errors. Obviously a reliable computation of these statistical errors is a crucial ingredient of the algorithm. A naive estimate would give:

$$\Delta \hat{O} \simeq \frac{\left(\langle \hat{O}^2 \rangle - \langle \hat{O} \rangle^2 \right)^{\frac{1}{2}}}{\sqrt{N_c}} \ , \tag{1.61}$$

where N_c is the number of sampled configurations. However this is not exact since, usually, strong correlation exists between successive configurations. In order to correct for this, one could measure the average interval N_i between statistically independent configurations and this should modify (1.61) in the form:

$$\Delta \hat{O} \simeq \frac{\left(\langle \hat{O}^2 \rangle - \langle \hat{O} \rangle^2 \right)^{\frac{1}{2}}}{\sqrt{N_c / N_i}} \ . \tag{1.62}$$

The underlying hypothesis is that, at equilibrium, $E_{\hat{O}}(\sigma)$ is Gaussianly distributed. Since this is not always the case it is practically convenient to measure statistical error by dividing the measure into segments of sufficient length and comparing the averages obtained in these intervals. If the segments are long compared to the

correlation time of the simulation, then the *sub-averages* are roughly Gaussianly distributed, due to the "central limit" theorem [47]. An estimate of the error using Eq. (1.61), where N_c now represents the number of sub-averages, is therefore correct and ensures the 68% of probability of finding the exact value of $\langle \hat{O} \rangle$, within the calculated uncertainty.

1.3.3 Jastrow Auxiliary Field method

The previous technique can also be used, within the VMC approach, to evaluate a multidimensional integral with, for instance, a Jastrow trial wave function. In fact we show that variational Jastrow estimates of ground state properties may be computed by an Auxiliary Field method analogous to that we illustrated in Section 1.3.1 for the exact many-body problem.

First of all let us define a Jastrow operator which, again, can be written as the exponential of a two-body operator:

$$\hat{U}_J \equiv e^{-\frac{1}{2}\hat{J}} = e^{-\frac{1}{2}\sum_{ij} J_{ij}\hat{\rho}_i\hat{\rho}_j} , \tag{1.63}$$

such that:

$$|\psi_J\rangle = \hat{U}_J |\psi_T\rangle , \tag{1.64}$$

where $J_{ij} = J(|\mathbf{r}_i - \mathbf{r}_j|)$, ψ_J is the Jastrow many-body wave function and ψ_T is a trial determinant. Now, by using Jastrow approximation, the expectation value of a generic operator $\hat{O}$ can be obtained as:

$$\langle \hat{O} \rangle_J = \frac{\langle \psi_J | \hat{O} | \psi_J \rangle}{\langle \psi_J | \psi_J \rangle} . \tag{1.65}$$

According to variational principle, if $\hat{O} = \hat{H}$, then $E_J \equiv \langle \hat{H} \rangle_J \geq E_0$. By definition

(1.64), relation (1.65) can be rewritten as:

$$\langle \hat{O} \rangle_J = \frac{\langle \psi_T | e^{-\frac{1}{2}\hat{J}} \hat{O} e^{-\frac{1}{2}\hat{J}} | \psi_T \rangle}{\langle \psi_T | e^{-\hat{J}} | \psi_T \rangle} . \tag{1.66}$$

Then it is easy to show that:

$$\langle \hat{O} \rangle_J = - \left(\frac{\partial}{\partial \lambda} \ln Q_J^\lambda \right)_{\lambda=0} , \tag{1.67}$$

where the Jastrow "partition function" Q_J^λ is defined as:

$$Q_J^\lambda \equiv \langle \psi_T | e^{-\frac{1}{2}\hat{J}} e^{-\lambda \hat{O}} e^{-\frac{1}{2}\hat{J}} | \psi_T \rangle . \tag{1.68}$$

We note that this procedure is formally identical to the scheme previously developed by introducing the pseudo partition function $Q = \langle \psi_T | e^{-\beta \hat{H}} | \psi_T \rangle$. Here the $\beta \hat{H}$ term is simply substituted by $\hat{J}$.

Now the HST can be performed in order to reduce the Jastrow two-body operator to a functional integral, over two auxiliary fields σ_1,σ_2, of one-body, field-dependent operators. Therefore:

$$Q_J^\lambda = \frac{1}{C} \int d\sigma_1 d\sigma_2 \, G(\sigma_1, \sigma_2) \, \langle \psi_T | \hat{U}_J(\sigma_1) e^{-\lambda \hat{O}} \hat{U}_J(\sigma_2) | \psi_T \rangle . \tag{1.69}$$

Hence, by applying (1.67), we obtain:

$$\langle \hat{O} \rangle_J = Q_J^{-1} \frac{1}{C} \int d\sigma_1 d\sigma_2 \, e^{-V_J(\sigma_1, \sigma_2)} E_{\hat{O}_J}(\sigma_1, \sigma_2) , \tag{1.70}$$

where:

$$V_J(\sigma_1, \sigma_2) = -\ln G(\sigma_1, \sigma_2) - \ln \langle \psi_T | \hat{U}_J(\sigma_1) \hat{U}_J(\sigma_2) | \psi_T \rangle , \tag{1.71}$$

and the Jastrow estimator is given by:

$$E_{\hat{O}_J}(\sigma_1, \sigma_2) = \frac{\langle \psi_T | \hat{U}_J(\sigma_1) \hat{O} \hat{U}_J(\sigma_2) | \psi_T \rangle}{\langle \psi_T | \hat{U}_J(\sigma_1) \hat{U}_J(\sigma_2) | \psi_T \rangle} . \tag{1.72}$$

The introduction of Jastrow auxiliary fields is a useful technique, not only to perform a variational calculation, but also to reduce the computer CPU time required by an exact AFQMC procedure. In fact, if the trial many-body wave function ψ_T is not the usual single Slater determinant obtained by a previous self-consistent Hartree-Fock calculation, but a Jastrow trial wave function ψ_J, which already accounts for some particle correlations, we expect that a smaller β will be necessary, in the imaginary time propagation $e^{-\beta \hat{H}} |\psi_J\rangle$, in order to attain ground state properties. Hence we can obtain an efficient algorithm by introducing two different kinds of auxiliary fields: the former, really time-dependent, is connected with imaginary time propagation $e^{-\beta \hat{H}}$ performed in the usual Trotter decomposition, and the latter is due to the presence of Jastrow propagator $e^{-\hat{J}}$. This procedure can save a relevant amount of computer time since, roughly speaking, we introduce only *two* (Jastrow) propagation steps when, generally, *a lot of* imaginary time propagation steps would be necessary to recover the same correlation already present in ψ_J.

1.3.4 The fermion sign problem

In Section 1.3.2 we have seen that HST allows us to write the pseudo partition function Q as a multidimensional integral over classical auxiliary fields. This integral has the schematic form:

$$Q = \int d\sigma \cdot G \cdot D \,, \tag{1.73}$$

where G represents the Gaussian weight and $D = \langle \psi_T | \hat{U}(\sigma) | \psi_T \rangle$ is the so-called *fermion determinant* (if $|\psi_T\rangle$ is a trial determinant then $\hat{U}(\sigma) |\psi_T\rangle$ continues to be a determinant and the scalar product of two Slater determinants is the determinant

of the corresponding overlap matrix). Then the ground state expectation value of a generic operator can be expressed as:

$$\langle \hat{O} \rangle = \frac{\int d\sigma \cdot G \cdot D \cdot E}{\int d\sigma \cdot G \cdot D} , \tag{1.74}$$

where the estimator E in the integrand (as well as G and D, of course) is understood to be a functional of the auxiliary field σ.

In Section 1.3.2 we have assumed that $G \cdot D$ is positive and we have considered it as a probability; in this way we can compute $\langle \hat{O} \rangle$ as a well defined classical average. Anyway, in general, this assumption is not true. In fact D is positive only in special cases because the many-body wave function $\hat{U}(\sigma)\,|\psi_T\rangle$, after the time-dependent propagation, can have a negative overlap with the initial trial wave function. The standard approach[31,36] is to use $G \cdot |D|$ as the probability to be sampled, so:

$$\langle \hat{O} \rangle = \frac{\int d\sigma \cdot G \cdot |D| \cdot S \cdot E}{\int d\sigma \cdot G \cdot |D| \cdot S} = \frac{\int d\sigma \cdot G \cdot |D| \cdot S \cdot E}{\langle S \rangle} , \tag{1.75}$$

where $S = D/|D|$ is the *sign* and $\langle S \rangle$ is its average over the probability distribution. Since our functional integrals are to be evaluated by a statistical sampling approach, it is clear that it will be difficult to obtain accurate statistics if $\langle S \rangle$ is small. In fact, in this case, there will be large cancellations in the numerator and in the denominator of Eq. (1.75) and statistical fluctuations will make an accurate evaluation of $\langle \hat{O} \rangle$ extremely difficult. This is the *fermion sign problem* which can be present in AFQMC calculations too. Obviously the crucial point is the behaviour of $\langle S \rangle$ as $\beta \to \infty$.

S. Sorella has rigorously shown[38] that the average sign $\langle S \rangle$ is either bounded by a constant (depending on the trial wave function) or it vanishes exponentially in the asymptotic $\beta \to \infty$ limit. Whenever this exponential instability occurs the

method faces difficulties which are similar to problems found in other Quantum Monte Carlo schemes (see Section 1.2). There are a limited number of models for which one can prove[38] that the average sign is always bounded by a constant equal to $|\langle\psi_0|\psi_T\rangle|^2$. Some examples, by considering the Hubbard model on a bipartite lattice, are the half-filled case with a repulsive Coulomb interaction, the one-dimensional case and the situation in which there is an attractive Coulomb interaction with an equal number of spin-up and spin-down particles. In all cases the fact that $\langle S\rangle$ remains finite seems to follow from a discrete symmetry of the model (for example the electron-hole symmetry in the half-filled Hubbard model on a bipartite lattice[40]).

If $\langle S\rangle$ does not go to zero exponentially at large β but approaches a constant, then the ground state expectation values can be computed ignoring the minus signs. In other words the quantity derived from Eq. (1.75) by setting $S = 1$:

$$\langle \hat{O} \rangle^* = \frac{\int d\sigma \cdot G \cdot |D| \cdot E}{\int d\sigma \cdot G \cdot |D|} , \tag{1.76}$$

converges to $\langle \hat{O} \rangle$ as $\beta \to \infty$. This can be derived, for the total energy, as follows (see ref. [36] for a general and strictly consistent derivation). First we can rewrite $\langle \hat{H} \rangle$ as (assuming $\beta \to \infty$):

$$\begin{aligned} \langle \hat{H} \rangle &= -\frac{d}{d\beta} \ln \langle \psi_T | \, e^{-\beta \hat{H}} \, |\psi_T\rangle \\ &= -\frac{d}{d\beta} \ln \left(\int d\sigma \cdot G \cdot D \right) . \end{aligned} \tag{1.77}$$

Multiplying and dividing by the correspondent expression, with D replaced by $|D|$, gives:

$$\langle \hat{H} \rangle = -\frac{d}{d\beta} \ln \left(\int d\sigma \cdot G \cdot |D| \frac{\int d\sigma \cdot G \cdot D}{\int d\sigma \cdot G \cdot |D|} \right)$$
$$= \langle \hat{H} \rangle^* - \frac{d}{d\beta} \ln \langle S \rangle . \qquad (1.78)$$

Now, if $\langle S \rangle$ goes to a constant as $\beta \to \infty$, we obtain that $\langle \hat{H} \rangle = \langle \hat{H} \rangle^*$.

On the other end, if $\langle S \rangle$ vanishes exponentially with β (this certainly happens, for instance, in numerical studies of the Hubbard model away from half-filling[35,38]), Eq. (1.78) indicates that one cannot expect to obtain exact results from calculations that ignore the signs. There are numerical examples[31,35] in which $\langle \hat{O} \rangle^*$ gives a good approximation to some ground state physical quantities, but it should be emphasized that, in general, Eq. (1.76) is an uncontrolled approximation which could lead to incorrect results[35].

If $\langle S \rangle$ goes to zero exponentially, then a *variational* approximation can be used to compute ground state properties: it is the "Positive Projection Method", introduced by Fahy and Hamann (see ref. [37] and Appendix E). This technique was derived by recasting the AFQMC method as a diffusion problem and does not exhibit the poor statistical behaviour due to vanishing sign. So far it was only applied to two-dimensional Hubbard models[37], but it could be of considerable advantage in extending auxiliary field methods to atoms, molecules and solid state systems where the effects of the fermion sign problem are, to a large extent, unknown.

Chapter 2

Interacting Electrons

In the following we specialize our discussion to the case of a realistic physical system: a *molecule* with positively charged nuclei and electrons interacting with a repulsive Coulomb potential. Our Hamiltonian is given by (a.u. are employed throughout this thesis):

$$\hat{H} = -\frac{1}{2}\sum_{\alpha=1}^{N}\nabla_{\alpha}^{2} + \sum_{\alpha=1}^{N} V^{\text{ext}}(\mathbf{r}_{\alpha}) + \frac{1}{2}\sum_{\alpha\neq\beta}\frac{1}{|\mathbf{r}_{\alpha}-\mathbf{r}_{\beta}|} + E_{\text{ion}}\ , \tag{2.1}$$

with:

$$V^{\text{ext}}(\mathbf{r}_{\alpha}) = -\sum_{I}\frac{Z_I}{|\mathbf{r}_{\alpha}-\mathbf{R}_I|}\ , \tag{2.2}$$

$$E_{\text{ion}} = \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J}{|\mathbf{R}_I-\mathbf{R}_J|}\ , \tag{2.3}$$

and where Greek indices are used for electrons, Latin capital letters for nuclei and Z_I indicates the charge of the I^{th} nucleus.

2.1 The AFQMC method for realistic systems

In practice, in order to apply the AFQMC formalism to a continuous system, we have first to discretize it. For example this may be carried out by introducing a

spatial lattice. In second quantization Hamiltonian (2.1) becomes:

$$\hat{H} = -\frac{1}{2} \sum_{i,j,\mu,\mu'} \langle i\mu | \nabla^2 | j\mu' \rangle \, c^{\dagger}_{i\mu} c_{j\mu'} + \sum_{i,\mu} V_i^{\text{ext}} \hat{\rho}_{i\mu} + \frac{1}{2} \sum_{i,j,\mu,\mu'} V_{ij} \hat{\rho}_{i\mu} \hat{\rho}_{j\mu'} \, , \quad (2.4)$$

where V_{ij} is the Coulomb potential matrix:

$$V_{ij} = V\left(|\mathbf{r}_i - \mathbf{r}_j|\right) = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \, , \quad (2.5)$$

and, for simplicity, we omitted the unphysical self-interaction contribution (see previous Chapter) and the constant ion-ion repulsion term. Now we have explicitly introduced the spin variable $\mu = \uparrow, \downarrow$, and the following relations hold for the fermion density operator $\hat{\rho}_{i\mu} = c^{\dagger}_{i\mu} c_{i\mu}$:

$$\hat{\rho}^2_{i\mu} = \hat{\rho}_{i\mu} \quad , \quad (2.6)$$

$$\sum_{i,\mu} \hat{\rho}_{i\mu} = \hat{N} \, . \quad (2.7)$$

Here $\hat{N}$ is the total particle number operator.

In this case the difficulty is given by the fact that V_{ij} is a positive definite matrix. Therefore, when we consider the exponential of two-body term in the Trotter decomposition (1.28), we have an expression of the form:

$$e^{\Delta\tau \sum_{i,j} (-V_{ij}) \hat{\rho}_i \hat{\rho}_j} \, . \quad (2.8)$$

Hence we cannot directly apply the HST, since $(-V_{ij})$ is a negative definite matrix (see Appendix A). In order to solve this problem two different approaches can be used [28]. The first is based on the same formulation described in Chapter 1, by adding an appropriate two-body interaction term to the Hamiltonian to ensure that the eigenvalues of the resulting effective potential matrix have the right sign

to perform the HST. The second technique uses a slightly different formulation involving a complex HST.

2.1.1 The Real HST

In order to use the HST with real auxiliary fields, let us consider a modified interaction matrix:

$$W_{ij} \equiv \lambda \delta_{ij} - V_{ij} \,. \tag{2.9}$$

Obviously, for sufficiently large λ, W_{ij} is a positive definite matrix. In practice it is convenient to choose λ equal to the maximum eigenvalue of the matrix V_{ij}. Now, if we define the *total local density* and *magnetization* operators:

$$\hat{d}_i = \hat{\rho}_{i\uparrow} + \hat{\rho}_{i\downarrow} \,, \tag{2.10}$$

$$\hat{m}_i = \hat{\rho}_{i\uparrow} - \hat{\rho}_{i\downarrow} \,, \tag{2.11}$$

we can write the two-body term:

$$\frac{1}{2} \sum_{i,j,\mu,\mu'} V_{ij} \hat{\rho}_{i\mu} \hat{\rho}_{j\mu'} = \frac{1}{2} \sum_{i,j} V_{ij} \hat{d}_i \hat{d}_j = \frac{\lambda}{2} \sum_i \hat{d}_i^2 - \frac{1}{2} \sum_{i,j} W_{ij} \hat{d}_i \hat{d}_j \,. \tag{2.12}$$

Then, by using relations (2.6) and (2.7), it is easy to show that:

$$\sum_i \hat{\rho}_{i\uparrow} \hat{\rho}_{i\downarrow} = \frac{1}{2} \left(\hat{N} - \sum_i \hat{m}_i^2 \right) \,. \tag{2.13}$$

Therefore:

$$\sum_i \hat{d}_i^2 = \hat{N} + 2 \sum_i \hat{\rho}_{i\uparrow} \hat{\rho}_{i\downarrow} = 2\hat{N} - \sum_i \hat{m}_i^2 \,. \tag{2.14}$$

And finally:

$$\frac{1}{2} \sum_{i,j,\mu,\mu'} V_{ij} \hat{\rho}_{i\mu} \hat{\rho}_{j\mu'} = -\frac{1}{2} \sum_{i,j} W_{ij} \hat{d}_i \hat{d}_j - \frac{\lambda}{2} \sum_i \hat{m}_i^2 + \lambda \hat{N} \,. \tag{2.15}$$

In this way:

$$e^{-\frac{\Delta\tau}{2}\sum_{i,j,\mu,\mu'} V_{ij}\,\hat{\rho}_{i\mu}\hat{\rho}_{j\mu'}} = e^{\frac{\Delta\tau}{2}\left(\sum_{i,j} W_{ij}\,\hat{d}_i\hat{d}_j + \lambda\sum_i \hat{m}_i^2 - 2\lambda\hat{N}\right)}$$
$$= e^{\frac{\Delta\tau}{2}\sum_{i,j} W_{ij}\,\hat{d}_i\hat{d}_j}\, e^{\frac{\Delta\tau}{2}\lambda\sum_i \hat{m}_i^2}\, e^{-\Delta\tau\lambda\hat{N}} . \qquad (2.16)$$

In conclusion we have obtained two quadratic contributions, $e^{\frac{\Delta\tau}{2}\sum_{i,j} W_{ij}\,\hat{d}_i\hat{d}_j}$, $e^{\frac{\Delta\tau}{2}\lambda\sum_i \hat{m}_i^2}$, and therefore two distinct Hubbard-Stratonovich transformations are necessary. Essentially, the presence of spin variables and the strategy used for reducing the negative definite matrix $(-V_{ij})$ to a positive definite form, implied introduction of two different, time-dependent auxiliary fields: one, σ^d, coupled to total local density and the other, σ^m, to local magnetization. Now we can apply the HST:

$$e^{\frac{\Delta\tau}{2}\sum_{i,j} W_{ij}\,\hat{d}_i\hat{d}_j} = \text{const.} \int \prod_i d\sigma_i^d\, e^{-\frac{\Delta\tau}{2}\sum_{i,j} W_{ij}\,\sigma_i^d\sigma_j^d}\, e^{\Delta\tau\sum_{i,j} W_{ij}\,\sigma_i^d\hat{d}_j} \qquad (2.17)$$

$$e^{\frac{\Delta\tau}{2}\lambda\sum_i \hat{m}_i^2} = \text{const.} \int \prod_i d\sigma_i^m\, e^{-\lambda\frac{\Delta\tau}{2}\sum_i \sigma_i^{m\,2}}\, e^{\lambda\Delta\tau\sum_i \sigma_i^m\hat{m}_i} . \qquad (2.18)$$

Using these expressions, the two-body contribution to the imaginary time propagation reads:

$$e^{-\frac{\Delta\tau}{2}\sum_{i,j,\mu,\mu'} V_{ij}\,\hat{\rho}_{i\mu}\hat{\rho}_{j\mu'}} = \text{const.} \int \prod_i d\sigma_i^m\, d\sigma_i^d\, e^{-\frac{\Delta\tau}{2}\left(\sum_{i,j} W_{ij}\,\sigma_i^d\sigma_j^d + \lambda\sum_i \sigma_i^{m\,2}\right)}$$
$$\times\, e^{\Delta\tau\left(\sum_{i,j} W_{ij}\,\sigma_i^d\hat{d}_j + \lambda\sum_i \sigma_i^m\hat{m}_i\right)}\, e^{-\lambda\Delta\tau\hat{N}} . \qquad (2.19)$$

Finally, by considering all time slices in the Trotter decomposition, our pseudo partition function becomes:

$$Q = \langle\psi_T|\, e^{-\beta\hat{H}}\,|\psi_T\rangle = \text{const.} \int d\sigma\, G(\sigma)\,\langle\psi_T|\,\hat{U}(\sigma)\,|\psi_T\rangle\ , \qquad (2.20)$$

where the Gaussian weighting factor is given by:

$$G(\sigma) = e^{-\frac{\Delta\tau}{2}\sum_{l=1}^{P}\left(\sum_{i,j} W_{ij}\,\sigma_i^d(l)\sigma_j^d(l) + \lambda\sum_i \sigma_i^{m\,2}(l)\right)}\ , \qquad (2.21)$$

and the single particle propagator is:

$$\hat{U}(\sigma) = \prod_{l=1}^{P} \hat{U}\left[\sigma(l)\right] \tag{2.22}$$
$$= \prod_{l=1}^{P} \left[e^{-\frac{\Delta\tau}{2}\hat{T}} e^{-\Delta\tau\left(\sum_i V_i^{\text{ext}}\hat{d}_i - \sum_{i,j} W_{ij}\sigma_i^d(l)\hat{d}_j - \lambda\sum_i \sigma_i^m(l)\hat{m}_i\right)} e^{-\frac{\Delta\tau}{2}\hat{T}} \right] ,$$

where we have omitted the $e^{-\lambda\Delta\tau N}$ term which is inessential in the propagation since, in our calculation, the total number of particles is fixed. We note that $\hat{U}(\sigma)$ can be written as the product of two propagators, $\hat{U}(\sigma) = \hat{U}^{\uparrow}(\sigma)\cdot\hat{U}^{\downarrow}(\sigma)$, each one acting on different spin subspaces. Therefore, since we consider, for the trial many-body wave function, a single Slater determinant made up of $N^{\uparrow}$ spin-up and $N^{\downarrow}$ spin-down orbitals, $\hat{U}(\sigma)$ acts independently on spin-up states $\varphi_{p\uparrow}$ and on spin-down states $\varphi_{p\downarrow}$. Hence, by observing that the scalar product of two Slater determinants is equal to the determinant of the corresponding overlap matrix, we can write:

$$\langle\psi_T|\,\hat{U}(\sigma)\,|\psi_T\rangle = \det\langle\varphi_{p\uparrow}|\,\hat{U}^{\uparrow}(\sigma)\,|\varphi_{q\uparrow}\rangle\cdot\det\langle\varphi_{p\downarrow}|\,\hat{U}^{\downarrow}(\sigma)\,|\varphi_{q\downarrow}\rangle\,. \tag{2.23}$$

As far as the non-kinetic part of the propagator $\hat{U}(\sigma)$ is concerned, we observe that the one-body propagated wave function is given by:

$$\varphi'_{p\uparrow\downarrow}(\mathbf{r}_i) = \langle\mathbf{r}_i|\,e^{-\Delta\tau\left(\sum_j V_j^{\text{ext}}\hat{d}_j - \sum_{j,k} W_{jk}\sigma_j^d(l)\hat{d}_k - \lambda\sum_j \sigma_j^m(l)\hat{m}_j\right)}\,|\varphi_{p\uparrow\downarrow}\rangle$$
$$= e^{-\Delta\tau V_{\uparrow\downarrow}^{\text{eff}}(\mathbf{r}_i,l)}\varphi_{p\uparrow\downarrow}(\mathbf{r}_i)\,. \tag{2.24}$$

In conclusion HST allows us to use a simple one-body formalism, but with an effective external potential which becomes time-dependent, spin-dependent and which is a function of σ variables:

$$V_{\uparrow}^{\text{eff}}(\mathbf{r}_i,l) = V^{\text{ext}}(\mathbf{r}_i) - \sum_j W_{ji}\sigma_j^d(l) - \lambda\sigma_i^m(l)\,, \tag{2.25}$$

$$V_{\downarrow}^{\text{eff}}(\mathbf{r}_i,l) = V^{\text{ext}}(\mathbf{r}_i) - \sum_j W_{ji}\sigma_j^d(l) + \lambda\sigma_i^m(l)\,. \tag{2.26}$$

2.1.2 The Complex HST

The second method for treating the repulsive Coulomb potential makes use of a complex version of the HST (see Appendix A). In fact, if we consider the two-body interaction, expressed in terms of total local density operators (see Eqs. (2.10) and (2.12)), we can write:

$$e^{-\frac{\Delta\tau}{2}\sum_{i,j} V_{ij}\hat{d}_i\hat{d}_j} = \text{const.} \int \prod_i d\sigma_i^d e^{-\frac{\Delta\tau}{2}\sum_{i,j} V_{ij}\sigma_i^d\sigma_j^d} e^{i\Delta\tau\sum_{i,j} V_{ij}\sigma_i^d\hat{d}_j} . \quad (2.27)$$

In this case we introduce only one auxiliary field σ^d, coupled to total local density, and an imaginary factor "i" is present inside the one-body, field-dependent propagator. By performing this new HST the Gaussian weighting factor and the single particle propagator of Eq. (2.20) become, respectively:

$$G(\sigma) = e^{-\frac{\Delta\tau}{2}\sum_{l=1}^{P}\left(\sum_{i,j} V_{ij}\sigma_i^d(l)\sigma_j^d(l)\right)} , \quad (2.28)$$

$$\hat{U}(\sigma) = \prod_{l=1}^{P}\left[e^{-\frac{\Delta\tau}{2}\hat{T}} e^{-\Delta\tau\left(\sum_i V_i^{\text{ext}}\hat{d}_i - i\sum_{i,j} V_{ij}\sigma_i^d(l)\hat{d}_j\right)} e^{-\frac{\Delta\tau}{2}\hat{T}}\right] . \quad (2.29)$$

Therefore we obtain the following effective potential to be used in the one-body propagation:

$$V_{\uparrow\downarrow}^{\text{eff}}(\mathbf{r}_i, l) = V^{\text{ext}}(\mathbf{r}_i) - i\sum_j V_{ji}\sigma_j^d(l) . \quad (2.30)$$

Comparing the previous expression with Eqs. (2.25) and (2.26) we observe that, now, V^{eff} is a complex, spin-independent, effective potential. This second method turns out to be much more convenient (see Chapter 4) from a numerical point of view.

We conclude this Section by introducing another possible approach to deal with repulsive Coulomb interaction. It is based on a different decomposition (sug-

gested by Fahy and Hamann[64]) of the usual two-body term:

$$\frac{1}{2}\sum_{i,j,\mu,\mu'} V_{ij}\hat{\rho}_{i\mu}\hat{\rho}_{j\mu'} = \frac{1}{2}\sum_{i,j} V_{ij}\left(\hat{\rho}_{i\uparrow}+\hat{\rho}_{i\downarrow}\right)\left(\hat{\rho}_{j\uparrow}+\hat{\rho}_{j\downarrow}\right) \tag{2.31}$$
$$= \sum_{i,j} V_{ij}\hat{\rho}_{i\uparrow}\hat{\rho}_{j\uparrow} + \sum_{i,j} V_{ij}\hat{\rho}_{i\downarrow}\hat{\rho}_{j\downarrow} + \frac{1}{2}\sum_{i,j}(-V_{ij})\hat{m}_i\hat{m}_j\,.$$

In this way the third term on the R.H.S. of Eq. (2.31) contains an interaction matrix $(-V_{ij})$ which is negative definite, so it has the right sign to perform a standard HST (see Eq. (1.34)):

$$e^{-\frac{\Delta\tau}{2}\sum_{i,j}(-V_{ij})\hat{m}_i\hat{m}_j} = e^{\frac{\Delta\tau}{2}\sum_{i,j}V_{ij}\hat{m}_i\hat{m}_j} \tag{2.32}$$
$$= \text{const.}\int\prod_i d\sigma_i^m\, e^{-\frac{\Delta\tau}{2}\sum_{i,j}V_{ij}\sigma_i^m\sigma_j^m}\, e^{\Delta\tau\sum_{i,j}V_{ij}\sigma_i^m\hat{m}_j}\,.$$

In the other two terms, on the R.H.S. of Eq. (2.31), the interaction matrix is positive definite, therefore the HST can be applied only by using one of the two techniques (the modification of the two-body interaction or the introduction of a complex HST) we have previously described. If we consider the complex HST we obtain:

$$e^{-\Delta\tau\sum_{i,j}V_{ij}\hat{\rho}_{i\mu}\hat{\rho}_{j\mu}} = \text{const.}\int\prod_i d\sigma_i^\mu\, e^{-\Delta\tau\sum_{i,j}V_{ij}\sigma_i^\mu\sigma_j^\mu}\, e^{2i\Delta\tau\sum_{i,j}V_{ij}\sigma_i^\mu\hat{\rho}_{j\mu}}\,. \tag{2.33}$$

Hence, on the whole, we have:

$$e^{-\frac{\Delta\tau}{2}\sum_{i,j,\mu,\mu'}V_{ij}\hat{\rho}_{i\mu}\hat{\rho}_{j\mu'}} = \text{const.}\int\prod_i d\sigma_i^m\, d\sigma_i^\uparrow\, d\sigma_i^\downarrow\, e^{-\frac{\Delta\tau}{2}\left(\sum_{i,j}V_{ij}\sigma_i^m\sigma_j^m+2\sum_{i,j,\mu}V_{ij}\sigma_i^\mu\sigma_j^\mu\right)}$$
$$\times\, e^{\Delta\tau\left(\sum_{i,j}V_{ij}\sigma_i^m\hat{m}_j+2i\sum_{i,j,\mu}V_{ij}\sigma_i^\mu\hat{\rho}_{j\mu}\right)}\,. \tag{2.34}$$

Using this approach three kinds of auxiliary fields are to be introduced: σ^m coupled to local magnetization, $\sigma^\uparrow$ coupled to spin-up density and $\sigma^\downarrow$ coupled to spin-down density. By considering all Trotter time slices the Gaussian weighting factor becomes:

$$G(\sigma) = e^{-\frac{\Delta\tau}{2}\sum_{l=1}^{P}\sum_{i,j}V_{ij}\left(\sigma_i^m(l)\sigma_j^m(l)+2\sum_\mu\sigma_i^\mu(l)\sigma_j^\mu(l)\right)}\,, \tag{2.35}$$

and the single particle propagator is

$$\hat{U}(\sigma) = \prod_{l=1}^{P} \left[e^{-\frac{\Delta\tau}{2}T} e^{-\Delta\tau\left(\sum_i V_i^{\rm ext} d_i - \sum_{i,j} V_{ij}\sigma_i^m(l)m_j - 2i\sum_{i,j,\mu} V_{ij}\sigma_i^\mu(l)\hat{\rho}_{j\mu}\right)} e^{-\frac{\Delta\tau}{2}T} \right] . \tag{2.36}$$

Finally we obtain an effective, time-dependent, spin-dependent, one-body potential given by:

$$V_\uparrow^{\rm eff}(\mathbf{r}_i, l) = V^{\rm ext}(\mathbf{r}_i) - \sum_j V_{ji}\left(2i\sigma_j^\uparrow(l) + \sigma_j^m(l)\right) , \tag{2.37}$$

$$V_\downarrow^{\rm eff}(\mathbf{r}_i, l) = V^{\rm ext}(\mathbf{r}_i) - \sum_j V_{ji}\left(2i\sigma_j^\downarrow(l) - \sigma_j^m(l)\right) . \tag{2.38}$$

As far as the correlation energy is concerned we expect that, in this approach, the σ^m auxiliary field is the most important one, since it is introduced by decoupling the interaction term which couples electrons of opposite spins, through the local magnetization operators. The other two auxiliary fields, $\sigma^\uparrow$ and $\sigma^\downarrow$, should be less important to recover the correlation energy, in fact they are obtained by decoupling two-body interaction terms taking into account particles with parallel spins, on which correlation should have only a weak effect, due to Pauli exclusion principle.

We point out that the different HST's we have introduced, although strictly equivalent from a formal point of view, exhibit a quite different numerical behaviour, when actually implemented (see Chapter 4).

2.2 Calculation of the estimators

The ground state expectation value of an arbitrary operator $\hat{O}$ can be calculated using the fundamental relations introduced in Section 1.3.2. Evidently the main task is the calculation of the auxiliary field-dependent estimator $E^{\tau}_{\hat{O}}(\sigma)$, at a fixed imaginary time τ. We start by deriving it for the operator $\hat{O} = c_{i\mu}c^{\dagger}_{j\mu}$. Eq. (1.54) can be easily written as:

$$E^{\tau}_{\hat{O}}(\sigma) = \frac{\left\langle c^{\dagger}_{i\mu}\hat{U}_{\sigma}(\tau,\beta)\,\psi_T \middle| c^{\dagger}_{j\mu}\hat{U}_{\sigma}(\tau,0)\,\psi_T \right\rangle}{\left\langle \hat{U}_{\sigma}(\tau,\beta)\,\psi_T \middle| \hat{U}_{\sigma}(\tau,0)\,\psi_T \right\rangle}. \tag{2.39}$$

If ψ_T is a N-state single Slater determinant, then Eq. (2.39) involves the scalar product of two $(N+1)$-particle determinants, due to the presence of creation operators. Now the scalar product of two Slater determinants is the determinant of the corresponding overlap matrix, that is:

$$E^{\tau}_{\hat{O}}(\sigma) = \frac{\det \tilde{A}^{\mu}(i,j)}{\det A^{\mu}}, \tag{2.40}$$

where A^{μ} is the $(N^{\mu} \times N^{\mu})$ overlap matrix:

$$A^{\mu}_{pq} = \langle \varphi_p | \hat{U}(\sigma) | \varphi_q \rangle = \langle \varphi^{<}_p(\tau) | \varphi^{>}_q(\tau) \rangle, \tag{2.41}$$

in which single particle orbitals have the same spin μ, and $\tilde{A}^{\mu}(i,j)$ is the $(N^{\mu}+1) \times (N^{\mu}+1)$ matrix defined by:

$$\tilde{A}^{\mu}_{pq} = \begin{pmatrix} \delta_{ij} & \cdots & \varphi^{>}_q(\mathbf{r}_j) & \cdots \\ \vdots & & & \\ \varphi^{<}_p(\mathbf{r}_i) & & A^{\mu}_{pq} & \\ \vdots & & & \end{pmatrix}. \tag{2.42}$$

Here we have introduced the *forward* and *backward* propagated wave function for integer times ($l = \tau/\Delta\tau$), defined by (see Eqs. (1.55) and (1.56)):

$$\varphi_q^{>}(\tau) = \hat{U}_\sigma(\tau, 0)\,\varphi_q = \varphi_q^{>}(l \cdot \Delta\tau) = \hat{U}_\sigma(l, 0)\,\varphi_q\ , \tag{2.43}$$

$$\varphi_p^{<}(\tau) = \hat{U}_\sigma(\tau, \beta)\,\varphi_p = \varphi_p^{<}(l \cdot \Delta\tau) = \hat{U}_\sigma(l, \beta)\,\varphi_p\ . \tag{2.44}$$

At first sight the full calculation of this estimator seems to be very expensive. In fact, by letting the indices i and j assume all possible values, corresponding to all possible N_a lattice sites, we obtain N_a^2 matrices of the form (2.42). Therefore we should evaluate N_a^2 determinants of order $(N^\mu + 1)$ for each spin value $\mu = \uparrow, \downarrow$. However we can simplify the problem by following a procedure[32] suggested by S. Sorella.

The N_a^2 matrices, corresponding to different values of i and j, differ one from the other simply by the exchange of one row and one column. Hence it is convenient to introduce the quantities:

$$B^\mu(i,j) \equiv \sum_{p,q} \varphi_p^{>}(\mathbf{r}_j)\,(A^\mu)^{-1}_{pq}\,\varphi_q^{<}(\mathbf{r}_i)\,. \tag{2.45}$$

Obviously a determinant remains unchanged if one adds to a column any linear combination of the others. Hence we may add to the first column of the matrix $\tilde{A}^\mu$ a linear combination of the other columns in order to make vanishing all the elements of the first column but the one in the *first* row:

$$\det\tilde{A}^\mu = \det\begin{pmatrix} \delta_{ij} - \sum_q b_q \varphi_q^{>}(\mathbf{r}_j) & \cdots & \varphi_q^{>}(\mathbf{r}_j) & \cdots \\ \vdots & & & \\ \varphi_p^{<}(\mathbf{r}_i) - \sum_q b_q A^\mu_{pq} & & A^\mu_{pq} & \\ \vdots & & & \end{pmatrix}. \tag{2.46}$$

Now, if we choose:

$$b_q = \sum_{q'} (A^\mu)^{-1}_{qq'}\,\varphi_{q'}^{<}(\mathbf{r}_i)\ , \tag{2.47}$$

we obtain the desired result:

$$\varphi_p^<(\mathbf{r}_i) - \sum_q b_q A_{pq}^\mu = 0\,. \tag{2.48}$$

Then, by using definition (2.45) and relations (2.46) – (2.48), we can write:

$$\det\tilde{A}^\mu = \det\begin{pmatrix} \delta_{ij} - B^\mu(i,j) & \cdots & \varphi_q^>(\mathbf{r}_j) & \cdots \\ \vdots & & & \\ 0 & & A_{pq}^\mu & \\ \vdots & & & \end{pmatrix} = \det A^\mu \cdot [\delta_{ij} - B^\mu(i,j)]\,. \tag{2.49}$$

Therefore the factor $\det A^\mu$ cancels out in Eq. (2.40) and finally we obtain:

$$E^\tau_{c_{i\mu}c^\dagger_{j\mu}}(\sigma) \equiv \langle\, c_{i\mu}c^\dagger_{j\mu}\,\rangle = \delta_{ij} - B^\mu(i,j)\,. \tag{2.50}$$

Now the computation of the $(N_a \times N_a)$ matrices $B^\uparrow$ and $B^\downarrow$ requires the inversion of two $(N^\mu \times N^\mu)$ matrices A^μ, amounting to $\approx \left(N^{\uparrow\,3} + N^{\downarrow\,3}\right)$ operations, a change of basis $(A^\mu)^{-1}\varphi^<$, that is $\approx \left(N^{\uparrow\,2} + N^{\downarrow\,2}\right) N_a$ operations, plus $\left(N^\uparrow + N^\downarrow\right)$ multiplications for each different couple of lattice sites for which the matrices are defined. Altogether we have $\sum_\mu \left(N^{\mu\,3} + N^{\mu\,2}N_a + N^\mu N_a^2\right) \simeq N_a^2\left(N^\uparrow + N^\downarrow\right)$ operations, since, usually, $N_a \gg N^\mu$. This is to be compared with the $N_a^2 \sum_\mu (N^\mu + 1)^3$ operations required by a direct evaluation of $\det\tilde{A}^\mu(i,j)$. Then, obviously, by omitting the spin variable μ:

$$\langle\, c_i^\dagger c_j\,\rangle = \delta_{ij} - \langle\, c_j c_i^\dagger\,\rangle = B(j,i)\,, \tag{2.51}$$

where brackets mean the quantum expectation value calculated over a given configuration of σ fields and at a fixed time τ, according to definition (2.39).

By using standard properties of determinants it is easy to derive (see Appendix B) another useful relation:

$$\langle\, c_i^\dagger c_j c_m^\dagger c_n\,\rangle = \langle\, c_i^\dagger c_j\,\rangle\langle\, c_m^\dagger c_n\,\rangle + \langle\, c_i^\dagger c_n\,\rangle\langle\, c_j c_m^\dagger\,\rangle\,. \tag{2.52}$$

Now, by means of Eqs. (2.49) – (2.52), we can compute all physically interesting ground state estimators, simply by expressing their corresponding operators in terms of c, $c^\dagger$. In the following we give explicit formulae for some of them:

a) *Particle density* $\hat{\rho}_{i\mu} = c^\dagger_{i\mu} c_{i\mu}$:

$$E^\tau_{\hat{\rho}_{i\mu}}(\sigma) = B^\mu(i,i) = \sum_{p,q} \varphi^>_p(\mathbf{r}_i)\,(A^\mu)^{-1}_{pq}\,\varphi^<_q(\mathbf{r}_i)\,. \tag{2.53}$$

b) *Kinetic energy* $\hat{T} = -\frac{1}{2}\sum_{i,j,\mu} T_{ij} c^\dagger_{i\mu} c_{j\mu}$, with $T_{ij} = \langle i|\nabla^2|j\rangle$:

$$E^\tau_{\hat{T}}(\sigma) = -\frac{1}{2}\sum_{i,j,\mu} T_{ij} B^\mu(j,i) = -\frac{1}{2}\sum_{i,j,\mu} T_{ij} \sum_{p,q} \varphi^>_p(\mathbf{r}_i)\,(A^\mu)^{-1}_{pq}\,\varphi^<_q(\mathbf{r}_j)\,. \tag{2.54}$$

c) *External energy* $\hat{V}^{\rm ext} = \sum_{i,\mu} V^{\rm ext}(\mathbf{r}_i) c^\dagger_{i\mu} c_{i\mu}$:

$$E^\tau_{\hat{V}^{\rm ext}}(\sigma) = \sum_{i,\mu} V^{\rm ext}(\mathbf{r}_i) B^\mu(i,i) = \sum_{i,\mu} V^{\rm ext}(\mathbf{r}_i) \sum_{p,q} \varphi^>_p(\mathbf{r}_i)\,(A^\mu)^{-1}_{pq}\,\varphi^<_q(\mathbf{r}_i)\,. \tag{2.55}$$

d) *Electron-electron energy* $\hat{V} = \frac{1}{2}\sum_{i,j,\mu,\mu'} V_{ij} c^\dagger_{i\mu} c^\dagger_{j\mu'} c_{j\mu'} c_{i\mu}$:

Evidently:

$$\begin{aligned}\langle\, c^\dagger_{i\mu} c^\dagger_{j\mu'} c_{j\mu'} c_{i\mu}\,\rangle &= \langle\, c^\dagger_{i\mu} c_{i\mu} c^\dagger_{j\mu'} c_{j\mu'}\,\rangle - \delta_{ij}\delta_{\mu\mu'}\langle\, c^\dagger_{i\mu} c_{j\mu'}\,\rangle \\ &= \langle\, c^\dagger_{i\mu} c_{i\mu}\,\rangle\langle\, c^\dagger_{j\mu'} c_{j\mu'}\,\rangle - \langle\, c^\dagger_{i\mu} c_{j\mu'}\,\rangle\langle\, c^\dagger_{j\mu'} c_{i\mu}\,\rangle\,. \end{aligned}\tag{2.56}$$

Hence, by observing that $\langle\, c^\dagger_{i\mu} c_{j\mu'}\,\rangle = \delta_{\mu\mu'}\langle\, c^\dagger_{i\mu} c_{j\mu}\,\rangle$:

$$\begin{aligned} E^\tau_{\hat{V}}(\sigma) &= \frac{1}{2}\sum_{i,j,\mu,\mu'} V_{ij} B^\mu(i,i)\, B^{\mu'}(j,j) - \frac{1}{2}\sum_{i,j,\mu} V_{ij} B^\mu(j,i)\, B^\mu(i,j) \qquad (2.57)\\ &= \frac{1}{2}\sum_{i,j,\mu,\mu'} V_{ij} \sum_{p,q,p',q'} (A^\mu)^{-1}_{pq} (A^{\mu'})^{-1}_{p'q'} \varphi^<_{q\mu}(\mathbf{r}_i)\varphi^>_{p\mu}(\mathbf{r}_i)\varphi^<_{q'\mu'}(\mathbf{r}_j)\varphi^>_{p'\mu'}(\mathbf{r}_j) \\ &\quad - \frac{1}{2}\sum_{i,j,\mu} V_{ij} \sum_{p,q,p',q'} (A^\mu)^{-1}_{pq} (A^\mu)^{-1}_{p'q'} \varphi^<_{q\mu}(\mathbf{r}_j)\varphi^>_{p\mu}(\mathbf{r}_i)\varphi^<_{q'\mu}(\mathbf{r}_i)\varphi^>_{p'\mu}(\mathbf{r}_j)\,, \end{aligned}$$

where, for clarity, we have pointed out the spin component of the orbitals.

We note that Eq. (2.57) is a generalization of the usual electron-electron energy, obtained in the Hartree-Fock approximation, with a "direct" and an "exchange" term. In fact, in the AFQMC approach, the ground state electron-electron energy estimator is given by the matrix element of $\hat{V}$ between two different, auxiliary field-dependent, determinants.

Chapter 3

The practical algorithm

3.1 Field updating techniques

In the AFQMC method ground state expectation values can be obtained by calculating functional integrals over auxiliary fields:

$$\langle \hat{O} \rangle = \frac{1}{Q} \int d\sigma E_{\hat{O}}(\sigma) e^{-V(\sigma)} . \tag{3.1}$$

Therefore choosing a method which efficiently samples σ variables is a crucial step. In order to attain this purpose a lot of strategies are available. Obviously the simplest one is a conventional Monte Carlo (MC) method [28,46] in which parallel updates of the auxiliary fields are allowed (due to the large number of auxiliary field variables, σ_i, the standard Monte Carlo procedure of updating a single degree of freedom, at each step, surely would give rise to an extremely slow sampling). In this approach the σ fields, for all space points (if a spatial discretization is used), and at all imaginary time slices are updated:

$$\sigma^{\text{new}} = \sigma^{\text{old}} + \delta \Delta\sigma , \tag{3.2}$$

where δ is a uniformly distributed random number between -1 and 1, and $\Delta\sigma$ is a constant factor used to fix the size of the random steps. Then a Metropolis acceptance/rejection test [46] is applied: if $V(\sigma^{\text{new}}) \leq V(\sigma^{\text{old}})$ the new configuration is

accepted, otherwise it is accepted with a probability given by $e^{-[V(\sigma^{new})-V(\sigma^{old})]}$. Surely this is not the most efficient way to proceed since all the field variables are updated by completely random movements and, usually, a very small value for $\Delta\sigma$ has to be chosen to obtain a good Metropolis acceptance ratio (between 30 and 70%). Therefore large amounts of computer time are required to correctly sample the auxiliary fields.

In Section 1.3.2 we anticipated that evaluation of the integrals over σ variables can be performed as a temporal average:

$$\langle \hat{O} \rangle = \lim_{s\to\infty} \frac{1}{(s - s_0)} \int_{s_0}^{s} ds' E_{\hat{O}} [\sigma(s')] \, . \tag{3.3}$$

In this context σ fields have to be suitably updated in the fictitious time s. An average over all σ can be replaced by an average over a fictitious time evolution at least in two ways.

One can introduce a Gaussian "white noise" function $\eta(s)$ and define the time dependence of σ by the *Langevin* equation:

$$\frac{d\sigma}{ds} = -\frac{\partial V(\sigma)}{\partial \sigma} + \eta(s) \, , \tag{3.4}$$

with:

$$\langle \eta(s)\eta(s') \rangle = 2\delta (s - s') \, . \tag{3.5}$$

From a physical point of view Langevin equation governs the Brownian motion of particles. The rationale behind this approach, first suggested by Parisi [65], and applied by Sorella et al. [29-34], is that a Fokker-Planck equation is associated with the stochastic evolution described by Eqs. (3.4) and (3.5) :

$$\frac{dP(\sigma)}{ds} = \nabla^2 P + \nabla \cdot [\nabla V(\sigma)] P \, , \tag{3.6}$$

where $P(\sigma)$ is the probability that the stochastic trajectory, determined by the Langevin equation, generates a configuration $\{\sigma\}$. In the limit $s \to \infty$ then $P(\sigma) \to e^{-\beta V(\sigma)}$ and one can use this property to sample the Boltzmann factor.

Another approach (see also following Section) is a *microcanonical* method. From a Molecular Dynamics (MD) viewpoint $V(\sigma)$ is considered as the potential energy for a classical dynamics (with unit mass), governed by Newton's law, so that:

$$\frac{d^2\sigma}{ds^2} = -\frac{\partial V(\sigma)}{\partial \sigma} . \tag{3.7}$$

This system conserves total energy, and time averages will agree (for $s \to \infty$) with functional averages, provided the system is ergodic and the initial conditions are arranged to satisfy the constraint that the kinetic energy should average to 1/2 per degree of freedom. In practice, in numerical simulations, the continuous dependence on time is replaced by a finite difference approximation, introducing a fictitious time step Δs. Then, in principle, results have to be extrapolated to $\Delta s \to 0$, since, obviously, a finite time step introduces errors [66].

A more recent class of simulation techniques, the so-called *Smart Monte Carlo* methods [67–69] promises to be very efficient. In essence one can only approximately integrate the (Langevin or MD) equations of motion, taking some discrete sequence of Δs steps. Then this entire trajectory is accepted or rejected by a Metropolis test. It is just this global acceptance/rejection step that makes the algorithm exact. In comparison with a simple MC approach this strategy is certainly much more expensive (due to force computation), nevertheless it should represent a main improvement, since the σ variables are no longer randomly updated, but the force term tends to guide the sampling of the auxiliary fields along the trajectory of the natural motion of the system. On the other hand, due to the presence of Metropolis

test, the truncation errors (associated with a finite time step Δs), affecting a pure MD or Langevin procedure, disappear. Therefore Δs may be chosen as large as possible while keeping the Monte Carlo acceptance rate satisfactorily high.

In the following Section a MD technique, together with its Smart Monte Carlo improvement, the "Hybrid MC" method, is described in detail. It was extensively used in previous applications [25,36] of AFQMC algorithm to Hubbard model and was also employed in our preliminary numerical simulations. However we have verified that, at least for the physical systems we have taken into account, a simpler and less expensive technique, a straightforward modification of the standard MC scheme, is considerably more efficient in sampling the auxiliary fields. It will be introduced in Section 3.1.2 .

3.1.1 The Hybrid MC method

The *Hybrid Monte Carlo* approach (HMC) [25,68] derives from similar algorithms that are being used in the study of lattice gauge theory [70]. In the AFQMC method Q is interpreted as a classical partition function of the variables σ (see Section 1.3.2), with a potential energy $V(\sigma)$ given by:

$$V(\sigma) = -\ln G(\sigma) - \ln \langle \psi_T | \hat{U}(\sigma) | \psi_T \rangle + \text{const.}\,. \tag{3.8}$$

Therefore it is convenient to introduce a momentum variable p, conjugate to each auxiliary field variable σ, and rewrite the partition function in the form:

$$Z = \text{const.} \int dp d\sigma e^{-E^{\text{cl}}(\sigma,p)/K_B T}\,, \tag{3.9}$$

where $K_B T = 1$, and $E^{\text{cl}}(\sigma, p)$ is the total classical energy for the auxiliary field σ dynamics. In the following we specify our formulae for the basic HST

defined through Eqs. (1.34) – (1.39) and consider the fermion determinant $\langle\psi_T|\hat{U}(\sigma)|\psi_T\rangle \geq 0$; generalization to other HST's, suitable for studying repulsive Coulomb interaction systems, and for the situation in which $\langle\psi_T|\hat{U}(\sigma)|\psi_T\rangle$ can be negative, is quite straightforward. Then

$$\begin{aligned} E^{\rm cl}(\sigma,p) &= E_{\rm kin}(p) + V(\sigma) \\ &= \frac{1}{2}\sum_{l=1}^{P}\sum_{i} {p_i}^2(l) \\ &+ \frac{\Delta\tau}{2}\sum_{l=1}^{P}\sum_{ij} W_{ij}\sigma_i(l)\sigma_j(l) - \ln\langle\psi_T|\hat{U}(\sigma)|\psi_T\rangle + \text{const.}\,. \end{aligned} \tag{3.10}$$

Since ground state properties are computed by estimators which are functions of σ fields only, the introduction of the p fields has no effect on physical results. Our task is to obtain a set of configurations of σ and p fields distributed as $e^{-E^{\rm cl}(\sigma,p)}$. Therefore we can adopt a microcanonical approach with the classical dynamics governed by familiar Hamilton's equations:

$$\dot{\sigma} = \frac{\partial E^{\rm cl}(\sigma,p)}{\partial p} = p\,, \tag{3.11}$$

$$\dot{p} = -\frac{\partial E^{\rm cl}(\sigma,p)}{\partial \sigma} = -\frac{\partial V(\sigma)}{\partial \sigma}\,, \tag{3.12}$$

where the dots over p and σ signify differentiation with respect to simulation fictitious time s. Obviously non-linear Hamilton's equations must be integrated numerically and this requires the introduction of a finite size, Δs, in fictitious simulation time. The integration can be carried out by the *leap frog* method[71] :

$$p(s+\Delta s/2) = p(s-\Delta s/2) - \Delta s\left(\frac{\partial V(\sigma)}{\partial \sigma}\right)_s\,, \tag{3.13}$$

$$\sigma(s+\Delta s) = \sigma(s) + p(s+\Delta s/2)\,\Delta s\,, \tag{3.14}$$

with initial and final half-steps:

$$p(\Delta s/2) = p(0) - \frac{\Delta s}{2}\left(\frac{\partial V(\sigma)}{\partial \sigma}\right)_0 , \tag{3.15}$$

$$p(s_f) = p(s_f - \Delta s/2) - \frac{\Delta s}{2}\left(\frac{\partial V(\sigma)}{\partial \sigma}\right)_{s_f} . \tag{3.16}$$

The half-steps differ from the exact integration by errors of order Δs^2, whereas the intermediate steps are affected by errors of order Δs^3. At the beginning each $p(0)$ field is given by a Gaussian random number distributed as $\exp(-p^2/2)$.

There are two things wrong with this procedure. First one cannot numerically integrate the equations of motion exactly, and thus one cannot conserve $E^{\rm cl}(\sigma, p)$ exactly. Second, if one could integrate them exactly, σ would be confined to one region of its configuration space by the zeroes of $\langle \psi_T | \hat{U}(\sigma) | \psi_T \rangle$, since the potential $V(\sigma)$ of Eq. (3.8) would have logarithmic infinities at these locations; this would invalidate the sample. These problems are both elegantly solved by incorporating a Metropolis rejection step based on:

$$q = \exp\left[E^{\rm cl}(\sigma^{\rm old}, p^{\rm old}) - E^{\rm cl}(\sigma^{\rm new}, p^{\rm new})\right] . \tag{3.17}$$

If $q \geq 1$ we accept $\{\sigma^{\rm new}\}$ as a new sample distribution, and if $q < 1$ we accept it with probability q. If we reject it, we reuse $\{\sigma^{\rm old}\}$ as a sample configuration, choose a new set of random momenta, and integrate again. By repeating the preceding procedure a set of configurations, which correctly samples the desired distribution[68], is generated. One has considerable latitude in applying this scheme since one can choose both the integration step length Δs and the number of Molecular Dynamics steps n_{MD}, between Monte Carlo rejection steps. The effective potential for the classical problem tends to have the form of isolated favourable "valleys" separated by large unfavourable regions, so increasing error in

the integration almost always makes $\left(E^{\rm cl}_{\rm new} - E^{\rm cl}_{\rm old}\right)$ more positive, and lowers the acceptance rate. The error increases when either Δs or n_{MD} is increased, while the statistical independence of $\{\sigma^{\rm old}\}$ and $\{\sigma^{\rm new}\}$ increases when the product $n_{MD}\Delta s$ is increased. In addition, this approach can help to overcome the second difficulty too. In fact the logarithmic barriers are usually very "thin" on the scale of σ "displacements" and therefore, if a finite, relatively large Δs may be used (the only practical limitation being a large enough acceptance rate), the procedure should make the (advantageous for the sampling) "error" of going through them.

Now we explicitly derive formulae for the *forces*, $-\partial V(\sigma)/\partial\sigma$, which have to be calculated for integrating the classical equations of motion. Their computation is usually the most time-consuming phase of the auxiliary field sampling, in the Hybrid MC scheme. Incidentally we observe[68] that, in the HMC algorithm, there is nothing which requires the form of the potential energy $V(\sigma)$, which determines the σ dynamics through Eqs. (3.13) – (3.16), to be equal to expression (3.8) that must be present in $E^{\rm cl}(\sigma,p)$, when the Metropolis test is performed. This, in principle, can introduce some further scope for optimizing the procedure, the only condition to be satisfied being, again, to get a reasonably high acceptance rate.

As far as the Gaussian part, of the potential (3.8), is concerned, derivatives may be performed in a straightforward way. On the contrary, for the remaining part, $-\ln\langle\psi_T|\,\hat{U}(\sigma)\,|\psi_T\rangle$, the task is not so trivial. If we define the force:

$$F_i(l) \equiv -\frac{\partial V(\sigma)}{\partial\sigma_i(l)}\,, \tag{3.18}$$

then its non-Gaussian part is given by:

$$-\frac{\partial}{\partial\sigma_i(l)}\left(-\ln\langle\psi_T|\,\hat{U}(\sigma)\,|\psi_T\rangle\right) \tag{3.19}$$
$$=\frac{\partial}{\partial\sigma_i(l)}\left(\ln\det\langle\varphi^{\uparrow}_p|\,\hat{U}^{\uparrow}(\sigma)\,|\varphi^{\uparrow}_q\rangle + \ln\det\langle\varphi^{\downarrow}_p|\,\hat{U}^{\downarrow}(\sigma)\,|\varphi^{\downarrow}_q\rangle\right)$$

that is the problem of computing the forces is completely decoupled in spin space. The derivative, in the last expression, affects only the propagator $\hat{U}(\sigma)$ at the time slice l and, in practice, can be performed in the following way. Let us consider the usual overlap matrix $A_{pq} = \langle \varphi_p | \hat{U}(\sigma) | \varphi_q \rangle$. Then determinants appearing in Eq. (3.19) can be formally written as:

$$\det A = e^{\mathrm{tr} \ln A} . \tag{3.20}$$

Therefore:

$$\frac{\partial}{\partial \sigma_i(l)} \ln \det A = \frac{\partial}{\partial \sigma_i(l)} \mathrm{tr} \ln A = \mathrm{tr} \left[\frac{\partial A}{\partial \sigma_i(l)} A^{-1} \right] = \sum_{pq} \left[\frac{\partial A}{\partial \sigma_i(l)} \right]_{pq} A^{-1}_{qp} . \tag{3.21}$$

Now the derivative $\partial A / \partial \sigma_i(l)$ can be explicitly calculated by introducing the back and forth propagated orbitals, at intermediate times:

$$\tilde{\varphi}^{>}_p(l) = e^{\Delta\tau T/2} \varphi^{>}_p (l \cdot \Delta\tau) , \tag{3.22}$$

$$\tilde{\varphi}^{<}_q(l) = e^{-\Delta\tau T/2} \varphi^{<}_q (l \cdot \Delta\tau) , \tag{3.23}$$

where this formulation depends on our particular Trotter decomposition (1.28), and $\varphi^{>}$, $\varphi^{<}$ are the back and forth propagated orbitals for integer times, defined in Eqs. (2.43) and (2.44). Obviously $\varphi^{>}_p(0)$ and $\varphi^{<}_q(P \cdot \Delta\tau) = \varphi^{<}_q(\beta)$ are the one particle orbitals corresponding to the trial Slater determinant ψ_T. Then we can write:

$$\left[\frac{\partial A^{\mu}}{\partial \sigma_i(l)} \right]_{pq} = \Delta\tau \sum_j W_{ij} \tilde{\varphi}^{<}_{p\mu}(\mathbf{r}_j, l) \tilde{\varphi}^{>}_{q\mu}(\mathbf{r}_j, l) . \tag{3.24}$$

Finally we give explicit formulae for the forces by considering Gaussian contributions too:

$$F_i(l) = -\Delta\tau \sum_j W_{ij} \sigma_j(l) \tag{3.25}$$

$$+ \Delta\tau \sum_j W_{ij} \left[\sum_{q\uparrow} \varphi^{<A^{-1}}_{q\uparrow}(\mathbf{r}_j, l) \tilde{\varphi}^{>}_{q\uparrow}(\mathbf{r}_j, l) + \sum_{q\downarrow} \varphi^{<A^{-1}}_{q\downarrow}(\mathbf{r}_j, l) \tilde{\varphi}^{>}_{q\downarrow}(\mathbf{r}_j, l) \right] ,$$

with:

$$\varphi_{q\mu}^{<A^{-1}}(\mathbf{r}_i, l) = \sum_p (A^\mu)_{qp}^{-1} \, \tilde{\varphi}_{p\mu}^{<}(\mathbf{r}_i, l) \, . \tag{3.26}$$

3.1.2 A simple MC algorithm

Let us consider again our pseudo partition function:

$$Q = \text{const.} \int d\sigma G(\sigma) D(\sigma) = \text{const.} \int d\sigma G(\sigma) |D(\sigma)| S(\sigma) \, , \tag{3.27}$$

where, as in Section 1.3.4, we have defined:

$$D(\sigma) = \langle \psi_T | \, \hat{U}(\sigma) \, | \psi_T \rangle \, , \tag{3.28}$$

$$S(\sigma) = D(\sigma)/|D(\sigma)| \, . \tag{3.29}$$

Essentially we need an algorithm to produce a set of auxiliary field $\{\sigma\}$ configurations distributed according to probability $G(\sigma) \cdot |D(\sigma)|$. This can be accomplished in a simple way by using the following procedure. First of all we generate an initial configuration σ_0, with a probability $G(\sigma_0)$, that is taking into account the Gaussian part only. Gaussianly distributed random variables can be easily obtained numerically, for instance, by adopting the Box-Muller method [47]. Then we generate a new configuration σ_1, using again the Gaussian probability $G(\sigma_1)$. At this point we compute the quantity:

$$q = \frac{|D(\sigma_1)|}{|D(\sigma_0)|} \, , \tag{3.30}$$

and we perform a standard Metropolis acceptance/rejection test, that is we accept σ_1 as a new auxiliary field configuration with probability q; in practice only if $q \geq \xi$, where ξ is a random number, uniformly distributed on the interval $(0, 1)$.

Then we repeat the previous procedure by generating the desired set of σ configurations. One can easily verify that this algorithm really samples the distribution $G(\sigma) \cdot |D(\sigma)|$. In fact, if $P(\sigma_0 \to \sigma_1)$ denotes the probability to reach σ_1, starting from σ_0, in this case we have:

$$P(\sigma_0 \to \sigma_1) = G(\sigma_1) \cdot \min\left(1, \frac{|D(\sigma_1)|}{|D(\sigma_0)|}\right) . \tag{3.31}$$

Now the required detailed balance condition:

$$\frac{P(\sigma_0 \to \sigma_1)}{P(\sigma_1 \to \sigma_0)} = \frac{G(\sigma_1) \cdot |D(\sigma_1)|}{G(\sigma_0) \cdot |D(\sigma_0)|} , \tag{3.32}$$

is certainly satisfied. For instance, if $|D(\sigma_1)| > |D(\sigma_0)|$, then:

$$\frac{P(\sigma_0 \to \sigma_1)}{P(\sigma_1 \to \sigma_0)} = \frac{G(\sigma_1)}{G(\sigma_0) \cdot |D(\sigma_0)|/|D(\sigma_1)|} = \frac{G(\sigma_1) \cdot |D(\sigma_1)|}{G(\sigma_0) \cdot |D(\sigma_0)|} . \tag{3.33}$$

Verification for the case $|D(\sigma_1)| \leq |D(\sigma_0)|$ is analogous.

Substantially our technique differs from conventional MC scheme because we update auxiliary fields using a Gaussian distribution in place of performing small random displacements (see Eq. (3.2)). The advantages, with respect to previously described Hybrid MC approach, are essentially two. First, our algorithm is much faster (at least a factor 3 in practical calculations) since no forces are to be computed. Secondly, apart from the unavoidable correlation induced by repeating the Metropolis test, which selects the accepted configurations, successive auxiliary field configurations are completely independent. In addition, obviously, the logarithmic barriers, corresponding to zeroes of $D(\sigma)$, are crossed without any problem, in this scheme.

On the contrary, using the Hybrid MC method, we have often observed a relatively large correlation between successive configurations. Therefore very long

simulation runs were necessary in order to obtain small enough statistical errors in the estimated physical quantities. Finally we have verified that, with HMC technique, crossing logarithmic barriers is surely allowed but, sometimes, it happens in a rather inefficient way.

In our calculations we updated the auxiliary fields at all time slices, always having an acceptance rate greater than 50%. Possibly, considering longer imaginary time propagations, with a larger number of Trotter time slices, or more complex physical systems, the acceptance rate could become too low, since, in that case, the left and right propagated trial determinants can easily be orthogonal, giving rise to a very small value for $|D(\sigma)|$. If this difficulty were present, it could be overcome updating the σ variables not at all time slices but, for example, only at one or a few of them, before the Metropolis acceptance/rejection test is applied.

3.2 Calculations and sampling in Fourier space

Until now we have considered a space discretization, that is, in our formulae, we have defined all the relevant physical quantities (wave functions, external potentials, auxiliary fields) on the lattice sites of a spatial mesh. Anyway, in actual calculations, when realistic systems are to be studied, working in Fourier component space is more convenient.

A practical motivation to introduce Fourier components, as the basic quantities, is the following. The kinetic term propagation, $e^{-\frac{\Delta\tau}{2}\hat{T}}$, is conveniently performed in reciprocal space, where $\hat{T}$ is a diagonal operator (then a Fast Fourier Transform[72] can be used to come back to direct space and perform the one-body,

effective potential propagation). Therefore it seems to be natural to expand the one particle wave functions $\varphi_p(\mathbf{r})$, which trial Slater determinant consists of, in plane waves $e^{i\mathbf{kr}}$, that form a complete set of functions. This is a well-known procedure in Solid State calculations where the periodicity of a Bravais lattice is exploited. In fact, if $\varphi_p(\mathbf{r})$ has the periodicity of a Bravais lattice, that is $\varphi_p(\mathbf{r}+\mathbf{R}) = \varphi_p(\mathbf{r})$ for all $\mathbf{r}$, and all $\mathbf{R}$ in the Bravais lattice, then only plane waves with the periodicity of the Bravais lattice can occur in the expansion. Since the set of wave vectors for plane waves with the periodicity of the lattice is just the reciprocal lattice, a function periodic in the direct lattice will have a plane wave expansion of the form:

$$\varphi_p(\mathbf{r}) = \sum_{\mathbf{G}} C_p(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} , \tag{3.34}$$

where the sum is over all reciprocal lattice vectors $\mathbf{G}$. The Fourier coefficients $C_p(\mathbf{G})$ are given by:

$$C_p(\mathbf{G}) = \frac{1}{\Omega} \int d\mathbf{r} \varphi_p(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} , \tag{3.35}$$

where the integral is over any direct lattice primitive cell, and Ω is the volume of the primitive cell. By a formal point of view this approach means that we are studying a periodically repeated physical system. Nevertheless, if an isolated molecule has to be considered, we can, equally well, use this method, provided that our direct primitive cell is large enough. By choosing, for example, a cubic box with volume $\Omega = a^3$, if a is sufficiently large and the system inside the cell is neutral, then we substantially obtain an infinite number of non-interacting copies of the same system and, therefore, we are able to recover the properties of an isolated molecule.

The number of plane waves involved in expansion (3.34) is determined by the kinetic energy cutoff:

$$\frac{1}{2} G^2 \leq E_{\text{cut}} . \tag{3.36}$$

Obviously the accuracy of the expansion can be improved simply by increasing the value of the cutoff energy $E_{\rm cut}$ and, consequently, the CPU computer time requested by numerical calculation. Working in Fourier space makes the AFQMC algorithm significantly faster, by eliminating the *convolution* integrals appearing in the effective potential. For instance, in order to compute single particle wave function propagation, we have to evaluate (see Section 2.1) convolution integrals which have the basic form:

$$V^{\rm eff}(\mathbf{r}_i) = \sum_j \sigma_j V_{ij} \longrightarrow \int d\mathbf{r}\sigma(\mathbf{r})V(\mathbf{r},\mathbf{r}_i)\,. \tag{3.37}$$

Certainly, by a numerical point of view, it is convenient to calculate (3.37) in reciprocal space:

$$V^{\rm eff}(\mathbf{G}) = \Omega V(\mathbf{G})\sigma(\mathbf{G})\,. \tag{3.38}$$

Then $V^{\rm eff}(\mathbf{r})$ can be quickly recovered by using the Fast Fourier Transform.

Furthermore computational efficiency should be improved by performing the calculations in Fourier space as far as the auxiliary field sampling is concerned[28]. In fact, if the σ variables are updated by using a spatial mesh ($\sigma_i = \sigma(\mathbf{r}_i)$), extremely irregular auxiliary fields are obtained. This is caused not by insufficient resolution in space discretization, but rather is due to the Monte Carlo algorithm itself which produces random, uncorrelated changes of the σ field at every lattice site. The physical system is not expected to fluctuate on this scale. Moreover, the integrability of such extremely erratic functions becomes rather questionable. Therefore it is convenient to update Fourier components $\sigma(\mathbf{G})$ of σ, rather than their values at individual spatial mesh points, since this is equivalent to perform "correlated" changes of the field at *all* space points (at a given time slice).

In conclusion it is generally profitable to use the Fourier coefficients of the various quantities as the fundamental variables and compute ground state prop-

erties and auxiliary field updating directly in reciprocal space. In Appendix C the main expressions, for calculating ground state estimators and sampling the auxiliary fields, are given in terms of Fourier components and these formulae were actually used by our algorithm.

3.3 Gram-Schmidt orthonormalization

By performing single particle wave function propagation, as shown in Eqs. (2.43), (2.44), a problem about numerical stability may occur. In fact, since imaginary time propagation is not unitary, the orthonormality conditions, initially satisfied by the orbitals, are not preserved during such propagation. Therefore, after repeated applications of the propagator over an elementary time step, an orthonormal basis set $\{\varphi_p(\mathbf{r}_i)\}$ (with $p = 1, \ldots, N$ and $i = 1, \ldots, N_a$) will no longer remain orthonormal and the algorithm will become numerically instable. This can be understood since, due to auxiliary field introduction, the orbitals are independently propagated through an imaginary time, one-body propagator. Thus a direct application of Eqs. (2.43), (2.44) results in a Slater determinant with a great deal of linear dependence among the wave functions. In this way the numerical information about the fermionic ground state is gradually lost. Sugiyama and Koonin [28] were the first to recognize this problem.

In order to have a stable propagation, one can apply the *Gram-Schmidt orthonormalization* [73] every few time slices. In fact the τ-time Slater determinant:

$$\psi^\tau = \det\left[\varphi_p^\tau(\mathbf{r}_i)\right] \ , \tag{3.39}$$

can be rewritten in terms of an orthonormal basis set by introducing a transfor-

mation:

$$\varphi_p^\tau = \sum_q U_{pq} \varphi'_q \ , \tag{3.40}$$

where the matrix U_{pq} is chosen in such a way that $\langle \varphi'_p | \varphi'_q \rangle = \delta_{pq}$. The matrix U_{pq} is not univocally determined by Eq. (3.40). A convenient choice is to use the Gram-Schmidt orthonormalization procedure. In this case, U_{pq} is a triangular matrix and ψ^τ can be written as:

$$\psi^\tau = \det \left[\sum_q U_{pq} \varphi'_q(\mathbf{r}_i) \right] = \det(U) \cdot \det \left[\varphi'_q(\mathbf{r}_i) \right] \ , \tag{3.41}$$

where the latter equality simply follows by expanding the determinant of the product of two square matrices: U_{pq} and $\varphi'_q(\mathbf{r}_i)$. Therefore, once more, ψ^τ can be expressed by means of orthogonal orbitals. Hence we have again to propagate a Slater determinant, made up of orthogonal orbitals, and one can proceed as usual, until the numerical stability will require another orthonormalization. With such a strategy we can propagate for a long imaginary time any function without any numerical problem, even though the computation time increases, due to orthonormalization of the orbitals which costs $N^2 N_a$ operations.

Chapter 4

Mastering the fluctuations

As usual, even though the theoretical formulation of the AFQMC method is well established, developing a really efficient algorithm, suitable for numerical computation, is not a trivial task.

The AFQMC technique was extensively applied to study electron correlations in the Hubbard model (see Appendix D). The development of the methods for carrying out those calculations has been dominated by a tendency to take advantage of the simplicity of the Hubbard model to maximize computational efficiency. In fact, in that case, various simplifications and variable transformations may be exploited which, unfortunately, do not hold any more when realistic, continuous systems are to be considered. Therefore a large amount of numerical tests was necessary in order to achieve a reasonable understanding of the technique and to develop a reliable algorithm.

In this Chapter one of the main technical problems, which can affect the AFQMC method, the presence of large statistical fluctuations, is described. The basic causes which give rise to huge fluctuations are illustrated, together with a possible solution to the problem. A selection of the results we have obtained will be presented in Chapter 5.

4.1 A serious difficulty: large fluctuations

At the beginning we carried out AFQMC simulations using the HST presented in Section 2.1.1, where a λ parameter is introduced to make the two-body interaction matrix negative definite As we have illustrated in Chapter 2, this is not the only way to apply the AFQMC technique to systems with a repulsive interaction, but surely it is the simplest one (no imaginary factors are introduced) and, in fact, it was suggested as the best choice in ref. [28].

Our preliminary numerical tests were applications to the Hydrogen molecule ground state, using a very small energy cutoff: $G^2/2 = E_{\rm cut} = 0.5$ Ry. With this $E_{\rm cut}$ value, expansion (3.34), for the one-body wave functions, contains only 7 plane waves (corresponding to the (0,0,0) and (1,0,0) shells in reciprocal space); therefore a low accuracy can be obtained and the physical system actually becomes a "toy model". In this way the main technical features of the AFQMC method can be readily analyzed, even though the quantitative results for ground state properties are meaningless as far as their absolute values are concerned. In addition, in this small cutoff situation, the "exact" (with respect to the fixed energy cutoff) ground state properties can be quickly obtained by an *exact diagonalization* procedure and this is very useful since we can immediately verify the correctness and precision of our numerical results. In this particular case, due to small energy cutoff, the correlation energy would be much smaller than real H_2 correlation energy, so its estimate would be difficult. Therefore, to increase the correlation energy contribution, we multiplied electron-electron interaction by a factor 4.

In Fig. 1 a typical AFQMC simulation is shown. Hydrogen nuclei were placed at equilibrium experimental distance for H_2 molecule, $R = 1.401$ a.u.[74].

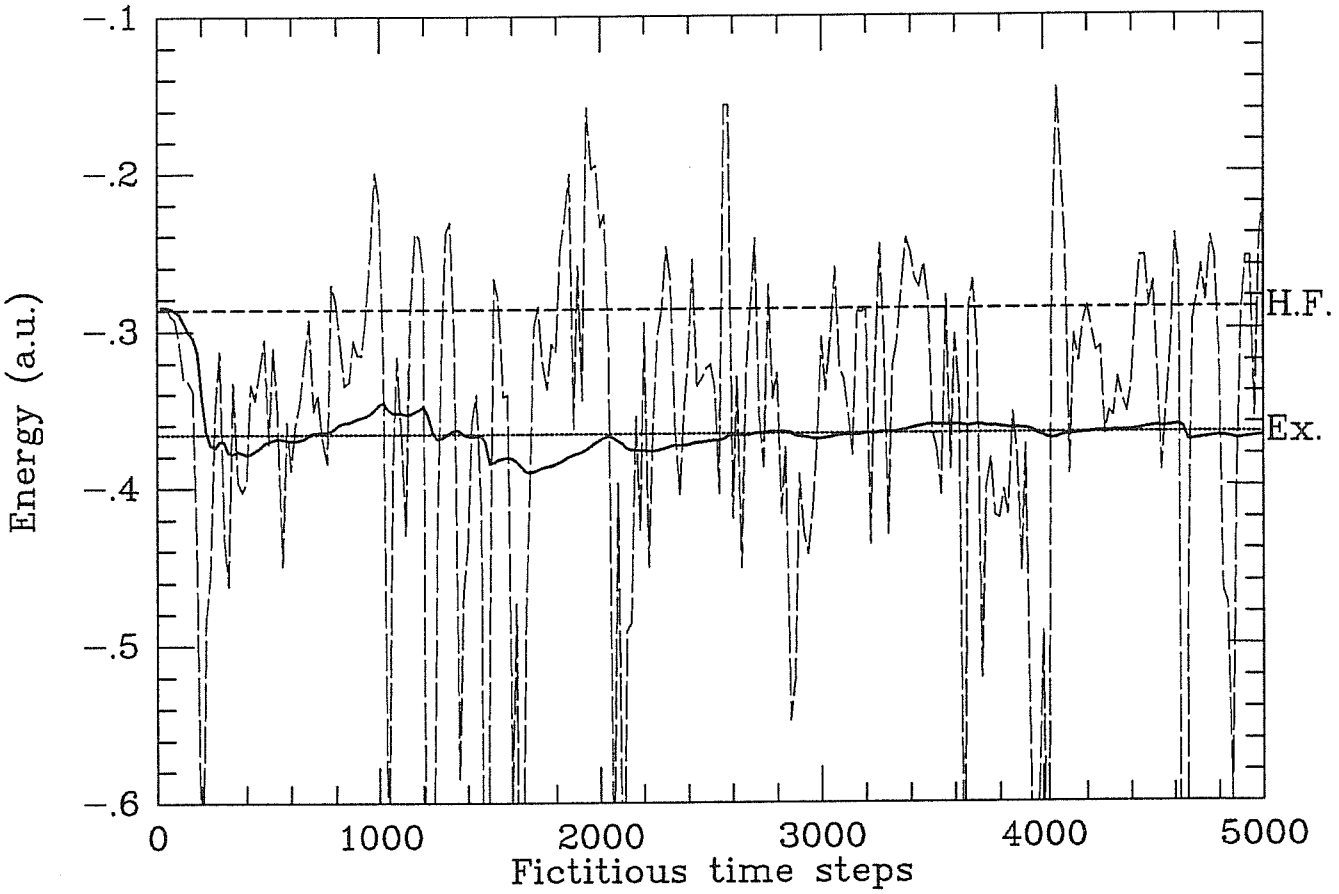

Figure 1. Fictitious time evolution for H_2 at E_{cut}=0.5 Ry in the Hybrid MC scheme. The fluctuating dashed line represents the energy estimator evolution while the continuous line is the cumulative average. As a reference Hartree-Fock and Exact results are shown.

The trial wave function was the self-consistent solution of a previous Hartree-Fock calculation, and the imaginary propagation time was $\beta = 2$ a.u.. For the auxiliary field sampling a Hybrid MC scheme was used. The H_2 ground state total energy estimator is plotted against the fictitious simulation time. The average value of this estimator, over all field configurations gives an estimate of the ground state energy. The Hartree-Fock energy (-0.287 a.u.) and the exact diagonalization

energy (-0.366 a.u.) are also shown: their difference is, by definition, the correlation energy contribution. To appreciate the convergence of the energy estimator average to the exact ground state value, the cumulative average (the solid line) is drawn. With 5000 simulation steps we have obtained $E_0 = -0.368 \pm 0.011$ a.u., which correctly estimates the exact value, even though statistical error is rather large due to considerable estimator fluctuations, that are evident in figure. Obviously a better precision can be achieved simply by performing longer simulations, by taking into account that statistical error is inversely proportional to the square root of the total number of sampled, independent, auxiliary field configurations. Fig. 1 clearly indicates that AFQMC technique really works for the simple system we have considered.

Unfortunately, when one uses an higher energy cutoff, in order to obtain meaningful quantitative results, a serious difficulty arises. In fact, with $E_{\rm cut}$ greater than about 1 Ry, very large fluctuations, on the scale of correlation energy, in ground state property estimators, occur. They make the statistical error huge, actually preventing this AFQMC scheme from getting acceptable results in a reasonable computer time. Just to realize the severe effect of these fluctuations we show in Fig. 2 the behaviour of the total energy estimator when $E_{\rm cut} = 1.5$ Ry is used. We must point out that this energy cutoff is still quite small in comparison with the values we usually need to get realistic results for molecules and Solid State systems. In this case, not only the local, field-dependent, estimator, but also the cumulative average shows enormous fluctuations on the scale of the correlation energy. Obviously, in such a situation, one cannot obtain any meaningful result. To study more precisely the energy cutoff dependence of the statistical fluctuations the following test has been performed.

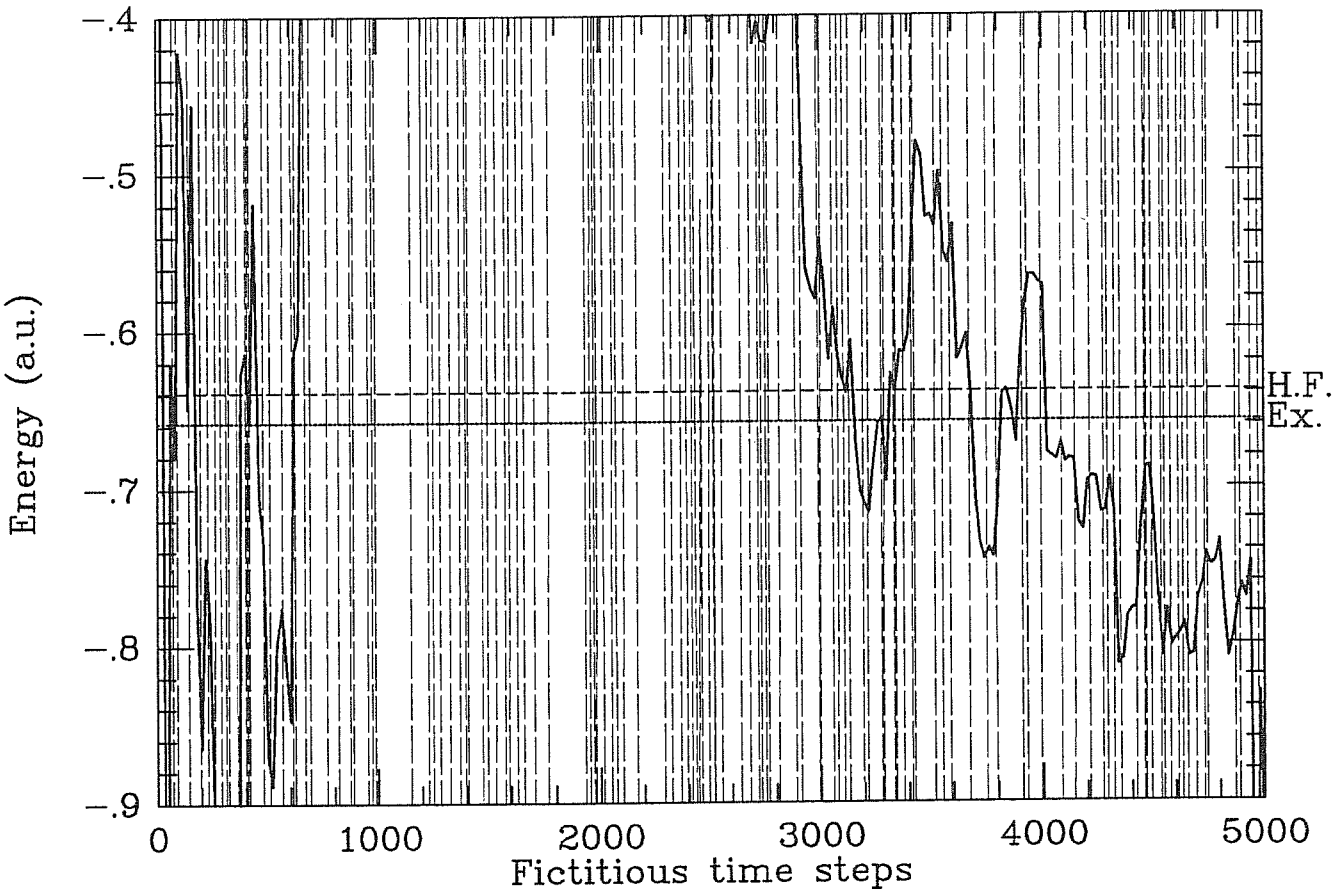

Figure 2. Fictitious time evolution for H_2 at E_{cut}=1.5 Ry in the Hybrid MC scheme. The fluctuating dashed line represents the energy estimator evolution while the continuous line is the cumulative average. As a reference Hartree-Fock and Exact results are shown.

We have used the variational Jastrow approach of the AFQMC method (see Section 1.3.3), considering $E_{\rm cut} = 6$ Ry, but only with the $N_G^{\rm eff}$ lowest G Fourier components which are effectively taken into account in the calculation. In Fig. 3 a), b), c), $N_G^{\rm eff}$ was equal to 20, 60, 200 respectively and the corresponding total energy statistical errors (which can be chosen as a measure of the fluctuations) were $\Delta E_0 = 0.002$ a.u., $\Delta E_0 = 0.017$ a.u., $\Delta E_0 = 0.308$ a.u..

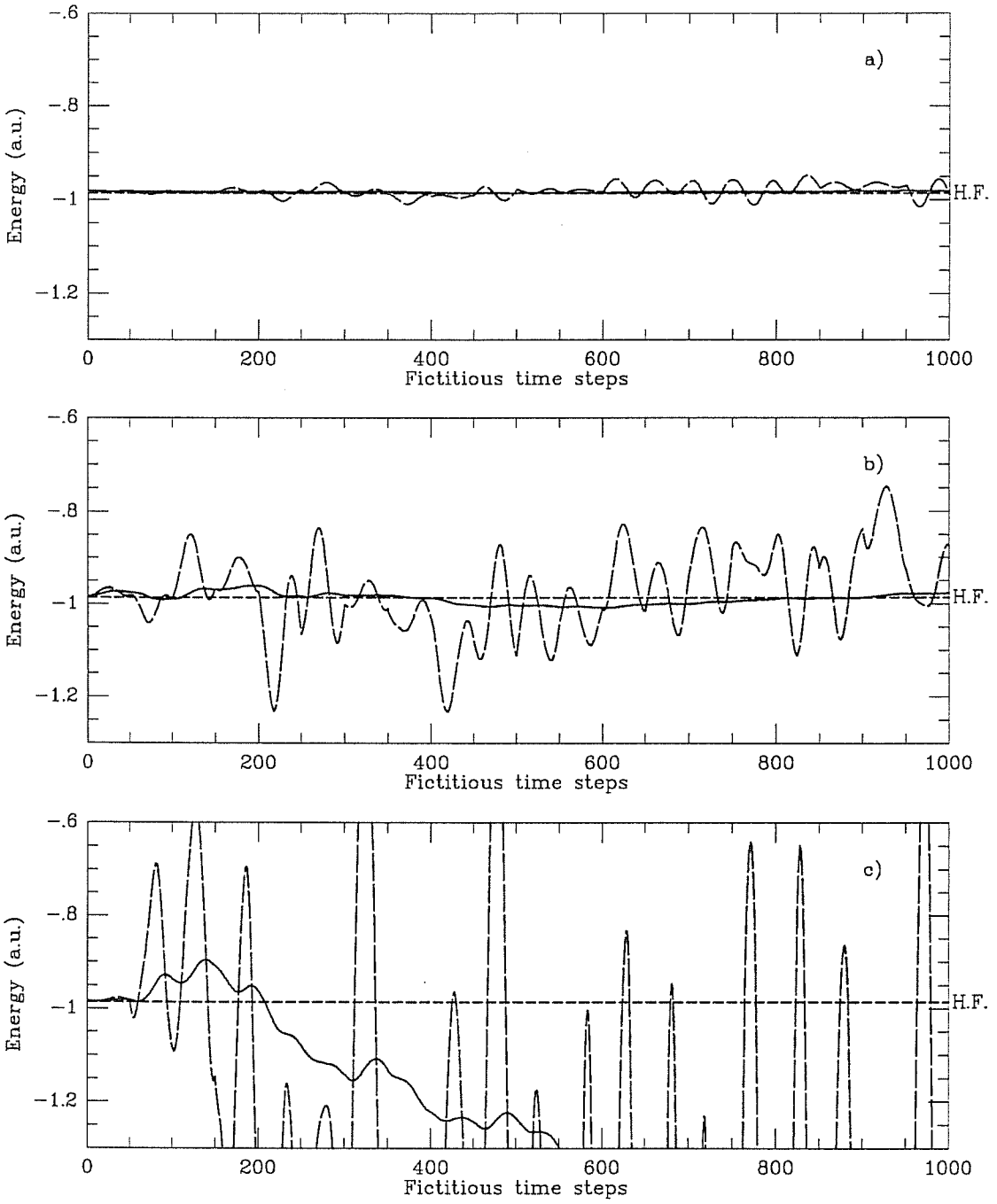

Figure 3. Variational Jastrow tests at E_{cut}=6.0 Ry, with only the N_G^{eff} lowest G Fourier components used. In Fig. 3a N_G^{eff}=20 and ΔE_0=0.002 a.u.; in Fig. 3b N_G^{eff}=60 and ΔE_0=0.017 a.u.; in Fig. 3c N_G^{eff}=200 and ΔE_0=0.308 a.u..

As one can see, the fluctuations are still acceptable for $N_G^{\rm eff} = 60$ but they make the results meaningless already for $N_G^{\rm eff} = 200$ (the $E_{\rm cut} = 6$ Ry energy cutoff would require about 2000 σ Fourier components). The reported errors show that fluctuations are proportional, at least, to $\left(N_G^{\rm eff}\right)^2$. Now $N_G \propto E_{\rm cut}^{3/2}$, so we can conclude empirically that, as a consequence of statistical fluctuations, our errors grow at least as $E_{\rm cut}^3$. Since the statistical error is inversely proportional to the square root of the number of independent configurations, our previous estimate indicates that, at least $4 \cdot 10^6$ configurations should be sampled, in the situation illustrated in Fig. 2 ($E_{\rm cut} = 1.5$ Ry), to obtain the same precision, in computed ground state quantities, of the simulation represented in Fig. 1 ($E_{\rm cut} = 0.5$ Ry)! Evidently this AFQMC scheme cannot be applied, except that in cases of very small energy cutoff. Therefore its practical utility is rather limited.

4.2 A possible solution to the problem

4.2.1 A convenient HST

A considerable effort was spent in order to understand what the presence of huge fluctuations depends on. We found that the main source of large statistical fluctuations is the introduction of the λ parameter, used in Section 2.1.1 to perform a real HST, suitable for repulsive interaction systems. In principle this λ factor represents only a convenient mathematical trick, which should not affect the final results, but, actually, its presence does give rise to a serious fluctuation problem. This can be understood by considering (see Eqs. (2.25) and (2.26)) the effec-

tive, auxiliary field-dependent, one-body potential, which governs wave function propagation, at a fixed time slice:

$$V^{\rm eff}_{\uparrow\downarrow}(\mathbf{r}_i) = V^{\rm ext}(\mathbf{r}_i) - \sum_j W_{ji}\sigma^d_j \mp \lambda\sigma^m_i \,. \tag{4.1}$$

The corresponding equation for the Fourier components is:

$$V^{\rm eff}_{\uparrow\downarrow}(\mathbf{G}) = V^{\rm ext}(\mathbf{G}) - \Omega W(\mathbf{G})\sigma^d(\mathbf{G}) \mp \lambda\sigma^m(\mathbf{G}) \,. \tag{4.2}$$

Now, if we remember that W_{ij} was defined in Eq. (2.9) as $W_{ij} = \lambda\delta_{ij} - V_{ij}$, then Eq. (4.2) can be written as:

$$V^{\rm eff}_{\uparrow\downarrow}(\mathbf{G}) = V^{\rm ext}(\mathbf{G}) - \left(\lambda - \frac{4\pi}{G^2}\right)\sigma^d(\mathbf{G}) \mp \lambda\sigma^m(\mathbf{G}) \,. \tag{4.3}$$

We observe that $V^{\rm ext}(\mathbf{G})$ is proportional to $1/G^2$ and, $\lambda \geq 4\pi/G^2_{\rm min}$ to make W_{ij} positive definite. Therefore, for $G \gg G_{\rm min}$, the effective potential Fourier coefficients become:

$$V^{\rm eff}_{\uparrow\downarrow}(\mathbf{G}) \simeq -\lambda\left(\sigma^d(\mathbf{G}) \pm \sigma^m(\mathbf{G})\right) \,. \tag{4.4}$$

As far as the auxiliary field dynamics is concerned, the Gaussian factor contribution to the classical potential energy (see Eqs. (2.21) and (3.10)), which is usually the dominant one, is (always at a fixed time slice):

$$V_{\rm Gauss}(\sigma) = \Delta\tau\left(\sum_{ij} W_{ij}\sigma^d_i\sigma^d_j + \lambda\sum_i \sigma^{m\,2}_i\right) \,, \tag{4.5}$$

and in terms of Fourier components:

$$V_{\rm Gauss}(\sigma) = \Delta\tau\Omega\sum_{\mathbf{G}}\left(\Omega W(\mathbf{G})|\sigma^d(\mathbf{G})|^2 + \lambda|\sigma^m(\mathbf{G})|^2\right) = \sum_{\mathbf{G}} V^{\sigma}_{\rm Gauss}(\mathbf{G}) \,. \tag{4.6}$$

Again, for the $G \gg G_{\min}$ terms, we can write:

$$V^{\sigma}_{\text{Gauss}}(\mathbf{G}) = \Delta\tau\Omega\left[\left(\lambda - \frac{4\pi}{G^2}\right)|\sigma^d(\mathbf{G})|^2 + \lambda|\sigma^m(\mathbf{G})|^2\right]$$
$$\simeq \Delta\tau\Omega\lambda\left(|\sigma^d(\mathbf{G})|^2 + |\sigma^m(\mathbf{G})|^2\right) . \tag{4.7}$$

Eq. (4.7) shows that, most of the auxiliary field components, evolve in a way that is practically independent on the physical system one is considering. Their dynamics depends essentially on the λ parameter. As a consequence, in Eq. (4.4), the effective one-body potential, for large G values, instead of decaying as $1/G^2$ (as we would expect for a Coulomb interaction), assumes a completely unphysical behaviour (the same for all G components), governed by the λ parameter. We must also point out that, usually, λ is a quite large constant. In fact $\lambda_{\min} = 4\pi/G^2_{\min} = a^2/\pi$, with a, the length of our cubic box side, which is to be chosen sufficiently long to make the system really isolate (we have typically used $a = 10$ a.u.).

One can easily realize that the longer the range of the interaction is, the larger $\lambda_{\min}$ becomes and, as a consequence, the larger the statistical fluctuations we must expect from our numerical calculation. Let us consider, for example, a Yukawa potential:

$$V(\mathbf{r}) = \frac{e^{-\mu r}}{r} . \tag{4.8}$$

Its Fourier components are:

$$V(\mathbf{G}) = \frac{4\pi}{\Omega\left(\mu^2 + G^2\right)} . \tag{4.9}$$

In this case the minimum possible value for λ is:

$$\lambda_{\min} = \frac{4\pi}{\mu^2 + G^2_{\min}} . \tag{4.10}$$

Hence the smaller the μ parameter is chosen, the longer the range of the $V(\mathbf{r})$ potential is, the larger $\lambda_{\min}$ is to be introduced. In particular, if $\mu \to 0$, we recover the usual Coulomb interaction. If, on the contrary, μ is quite large we have a small range interaction and the λ parameter becomes rather small too. For instance, in the standard Hubbard model, with a zero range, on site, interaction, we expect that $\lambda_{\min} \to 0$. In fact (see Appendix D), in that case, a particular, real HST can be employed, without introducing any λ parameter, even though a repulsive interaction is considered.

In conclusion this λ factor, which affects both the one-body propagation and the auxiliary field evolution, introduces a considerable "white noise" in numerical calculations when repulsive, long range potentials are taken into account. Obviously this negative effect increases as the energy cutoff grows.

To confirm our previous assertions we have performed two kinds of tests, always using the variational Jastrow approach, with $E_{\text{cut}} = 6$ Ry, considered at the end of Section 4.1.

In the first test all the parameters were the same, but the sign of the Jastrow factor was "wrong" (the electrons were forced to come closer to each other) and this made the introduction of a λ factor useless. We found that, when λ was not present, the statistical error in the ground state energy estimate was about an order of magnitude smaller than in the other case.

Then we used again the "correct" sign, for the Jastrow factor, but we introduced a λ parameter which was 5 times greater than the minimum allowed one, $\lambda_{\min}$. In principle this should be of no consequence in final results, since a correct, real HST only requires to make use of a sufficiently large ($\lambda \geq \lambda_{\min}$) λ factor. Nevertheless the statistical error was about 4 times larger.

At this point is clear that an AFQMC formulation which avoids using the λ parameter is to be employed in order to apply the method to realistic systems. In Section 2.1.2 we have introduced two different HST's which can be performed without introducing any λ factor. In particular we have seen, in Eq. (2.31), that the two-body interaction term can be rewritten as:

$$\frac{1}{2}\sum_{i,j,\mu,\mu'} V_{ij}\hat{\rho}_{i\mu}\hat{\rho}_{j\mu'} = -\frac{1}{2}\sum_{i,j} V_{ij}\hat{m}_i\hat{m}_j + \sum_{i,j} V_{ij}\left(\hat{\rho}_{i\uparrow}\hat{\rho}_{j\uparrow} + \hat{\rho}_{i\downarrow}\hat{\rho}_{j\downarrow}\right) . \tag{4.11}$$

For the H_2 Singlet ground state the previous expression is physically equivalent to:

$$\frac{1}{2}\sum_{i,j,\mu,\mu'} V_{ij}\hat{\rho}_{i\mu}\hat{\rho}_{j\mu'} = -\frac{1}{2}\sum_{i,j} V_{ij}\hat{m}_i\hat{m}_j , \tag{4.12}$$

with $\hat{m}_i = \hat{\rho}_{i\uparrow} - \hat{\rho}_{i\downarrow}$. In fact the second term on the R.H.S. of Eq. (4.11) involves only interactions between equal spin particles and, therefore, it represents only an unphysical self-interaction contribution, since our system contains two electrons with opposite spins. Now the usual HST can be applied to Eq. (4.12), which has the "right" sign (see Eq. (2.32)). In this special case only one auxiliary field is required to decouple the interaction, through a real HST. This alternative scheme generates an effective one-body potential (see Eqs. (2.37) and (2.38)) given by (at a fixed time slice):

$$V^{\text{eff}}_{\uparrow\downarrow}(\mathbf{r}_i) = V^{\text{ext}}(\mathbf{r}_i) \mp \sum_j V_{ji}\sigma^m_j , \tag{4.13}$$

and, for the Fourier components:

$$V^{\text{eff}}_{\uparrow\downarrow}(\mathbf{G}) = V^{\text{ext}}(\mathbf{G}) \mp \frac{4\pi}{G^2}\sigma^m(\mathbf{G}) , \tag{4.14}$$

while the Gaussian contribution to the classical potential energy, for the auxiliary field dynamics, is:

$$V^{\sigma}_{\text{Gauss}} = \Delta\tau \sum_{ij} V_{ij}\sigma_i^m\sigma_j^m \qquad (4.15)$$
$$= \Delta\tau\Omega^2 \sum_{\mathbf{G}} V(\mathbf{G})|\sigma^m(\mathbf{G})|^2 = \sum_{\mathbf{G}} V^{\sigma}_{\text{Gauss}}(\mathbf{G}) \,,$$

with:

$$V^{\sigma}_{\text{Gauss}}(\mathbf{G}) = \Delta\tau\Omega\frac{4\pi}{G^2}|\sigma^m(\mathbf{G})|^2 \,. \qquad (4.16)$$

In this case, to be compared with Eqs. (4.4) – (4.7), the auxiliary field evolution is connected to the real Coulomb interaction and the effective one-body potential of Eq. (4.14) exhibits a more physical behaviour for large G values. One can expect a small increasing of statistical fluctuations by using larger and larger energy cutoffs, but surely it should be much less dramatic than adopting the previous AFQMC formulation with λ factor.

We applied this new scheme to the usual variational Jastrow approach, with $E_{\text{cut}} = 6$ Ry, for the H_2 ground state, and the results were quite satisfactory. In fact Fig. 4 shows that the situation is much better than using the old formulation (see Fig. 2 where the energy cutoff was 4 times smaller!). Here the fluctuations of the total energy estimator are still relatively large on the scale of the correlation energy, anyway they are much smaller than in Fig. 2. As a consequence, the cumulative average, after a short, transient period, becomes a smooth curve which eventually converges to the exact value. From a quantitative point of view the statistical error was at least an order of magnitude smaller; obviously this improvement becomes more and more relevant by increasing the energy cutoff.

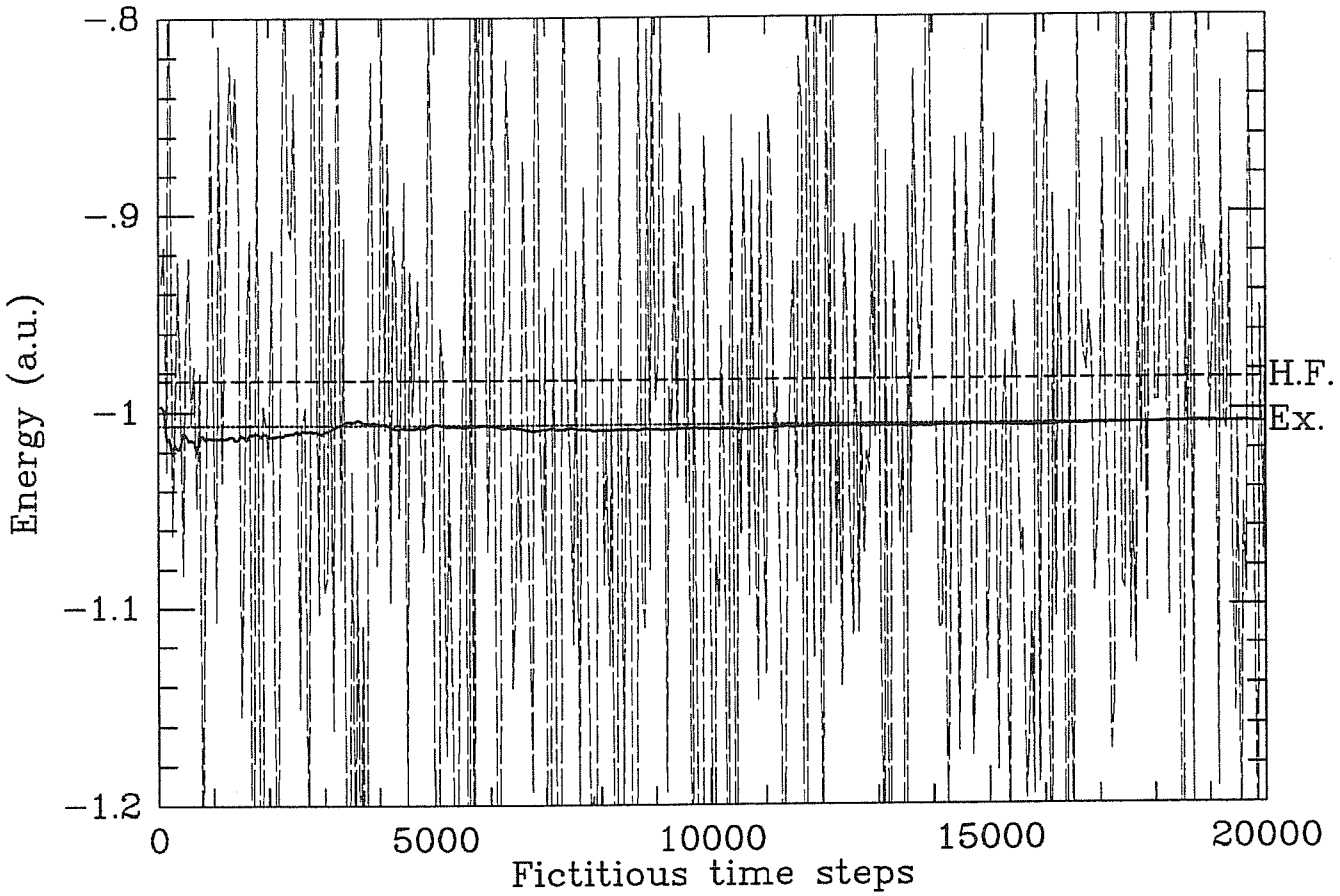

Figure 4. Fictitious time evolution for H_2 at E_{cut}=6.0 Ry in the Jastrow AFQMC scheme. The fluctuating dashed line represents the energy estimator evolution while the continuous line is the cumulative average. As a reference Hartree-Fock and Exact results are shown.

Hence, after theoretical considerations, numerical evidence too indicates that an AFQMC scheme which does not require the introduction of the λ parameter is much more advantageous for dealing with realistic physical systems. Surely the H_2 ground state is a rather special and favourable case. In general it is not possible to avoid using the λ factor through a real HST, and complex transformations are to be considered (see Section 2.1). Making use of complex HST's is generally regarded

as a bad choice, since they imply a Monte Carlo sampling of a complex function and this introduces a possible sign problem also in a case (the H_2 ground state) where the usual fermion sign problem is certainly absent. In fact we have applied to H_2 ground state the AFQMC method, described by Eqs. (2.27) – (2.30) and based on a single auxiliary field, coupled to the local total density operator, with the field-dependent part of the one-body potential which is completely imaginary. Using again the variational Jastrow approach, we obtained correct results, even though the convergence was slower because the statistical errors were 2-3 times larger than employing the real HST. Anyway the situation was much better than introducing the λ parameter. So we conclude that, applying the AFQMC technique to general systems, with a repulsive, long-range electron-electron interaction, the introduction of a complex HST seems to be the only practical solution to the fluctuation problem.

4.2.2 A further improvement

Although eliminating the λ parameter is a crucial step towards the reduction of statistical fluctuations further improvements are possible. A source of spurious fluctuations can be easily detected in the following way. Let us rewrite, schematically, the basic, single step of the imaginary time propagation (for instance that one corresponding to the first Trotter time slice), using the HST and considering only the potential term:

$$\begin{aligned} e^{-\Delta\tau\hat{V}}\left|\psi_T\right\rangle &= \int_{-\infty}^{+\infty} d\sigma G(\sigma) e^{-\Delta\tau\left(\hat{V}^{\mathrm{ext}}+\hat{V}(\sigma)\right)}\left|\psi_T\right\rangle \\ &= e^{-\Delta\tau\hat{V}^{\mathrm{ext}}}\int_{-\infty}^{+\infty} d\sigma G(\sigma)\left|\psi_T^{\sigma}\right\rangle \ , \end{aligned} \tag{4.17}$$

where $G(\sigma)$ is the usual Gaussian weight, $\hat{V}(\sigma)$ is the field-dependent part of the one-body potential, and:

$$\left|\psi_T^\sigma\right\rangle \equiv e^{-\Delta\tau\hat{V}(\sigma)}\left|\psi_T\right\rangle \ . \tag{4.18}$$

Now, by expanding the exponential in the previous expression:

$$\left|\psi_T^\sigma\right\rangle = \left(\hat{1} - \Delta\tau\hat{V}(\sigma) + \frac{1}{2}\Delta\tau^2\hat{V}^2(\sigma) + \cdots\right)\left|\psi_T\right\rangle \ , \tag{4.19}$$

one can immediately realize that the odd terms give no contribution when inserted in Eq. (4.17), as they integrate to zero. In particular this is true for the linear term, which we expect to be the most important one ($\Delta\tau$ is usually small):

$$\int_{-\infty}^{+\infty} d\sigma\, G(\sigma)\left(-\Delta\tau\hat{V}(\sigma)\right) = 0 \ . \tag{4.20}$$

Anyway, since we cannot compute analytically the functional integrals and we have to perform a stochastic evaluation, the linear term in (4.19) gives rise to spurious one-body excitations which produce harmful fluctuations. Surely their average will be zero, but, in practical calculations, they will increase the statistical fluctuations in a relevant way.

One solution to this difficulty is rather simple. In fact we observe (see Eqs. (4.1) and (4.13)) that the potential $\hat{V}(\sigma)$ depends linearly on the auxiliary fields σ, that is $\hat{V}(\pm\sigma) = \pm\hat{V}(\sigma)$. Therefore, by changing the sign of the integration variable in Eq. (4.17), one easily obtains:

$$\begin{aligned}\int_{-\infty}^{+\infty} d\sigma G(\sigma)\left|\psi_T^{+\sigma}\right\rangle &= \int_{-\infty}^{+\infty} d\sigma G(\sigma)\left|\psi_T^{-\sigma}\right\rangle \\ &= \int_{-\infty}^{+\infty} d\sigma G(\sigma)\left(\frac{\left|\psi_T^{+\sigma}\right\rangle + \left|\psi_T^{-\sigma}\right\rangle}{2}\right) \ .\end{aligned} \tag{4.21}$$

If now we expand the quantity:

$$\frac{|\psi_T^{+\sigma}\rangle + |\psi_T^{-\sigma}\rangle}{2} = \left(\hat{1} + \frac{1}{2}\Delta\tau^2\hat{V}^2(\sigma) + \cdots\right)|\psi_T\rangle \ , \tag{4.22}$$

the advantage of this new scheme is evident, since the undesirable odd terms, particularly the linear one, are eliminated. In practice, what we have to do is, for each time slice, to compute not only the propagated wave function $|\psi_T^{\sigma}\rangle$, but also the wave function obtained by changing the sign of the auxiliary field σ, $|\psi_T^{-\sigma}\rangle$.

We have applied this new "$\pm\sigma$" scheme, using again the variational Jastrow algorithm (with $E_{\rm cut} = 6$ Ry), to the H_2 ground state. In this simple Jastrow case (see Eq. (1.72)) only two propagation steps, one acting "on the left" and one "on the right", are present, so that, by employing our "$\pm\sigma$" algorithm, the local, field-dependent, estimator becomes:

$$E_{\hat{O}_J}(\sigma) = \frac{\left(\langle\psi_T^{+\sigma_1}| + \langle\psi_T^{-\sigma_1}|\right)\hat{O}\left(|\psi_T^{+\sigma_2}\rangle + |\psi_T^{-\sigma_2}\rangle\right)}{\left(\langle\psi_T^{+\sigma_1}| + \langle\psi_T^{-\sigma_1}|\right)\left(|\psi_T^{+\sigma_2}\rangle + |\psi_T^{-\sigma_2}\rangle\right)}. \tag{4.23}$$

We performed a simulation, generating 1500 auxiliary field configurations, and we found that total energy statistical error was $\Delta E_0 = 0.005$ a.u. without introducing the "$\pm\sigma$" scheme, while it was lowered to $\Delta E_0 = 0.001$ a.u. using this technique. Obviously the "$\pm\sigma$" algorithm is more expensive, since a double propagation is required in place of a single one, however, observing that the statistical error is inversely proportional to the square root of the number of independent configurations, this new scheme seems to be really convenient (approximately $1500 \cdot 25 = 37500$ configurations would be necessary to reduce the $5 \cdot 10^{-3}$ a.u. error to $1 \cdot 10^{-3}$ a.u.).

If we consider a long imaginary time propagation, where a lot of (not only two) time slices, and therefore many propagation steps, are taken into account, the situation is more involved. In fact one should apply the "$\pm\sigma$" technique at each

time slice of the propagation and this would lead to an exponential increase of the number of propagated determinants to be considered (the number of left and the number of right determinants should be equal to $2^{P/2}$, P being the total number of time slices). Certainly computation of estimators in such a situation would become really a too expensive task. What we actually did, in our calculations, was to employ a "restricted" version of the general "$\pm\sigma$" algorithm, by computing the local estimator in the following way:

$$E_{\hat{O}}(\sigma) = \frac{\left(\left\langle \psi^{+\sigma}_{\beta/2}\right| + \left\langle \psi^{-\sigma}_{\beta/2}\right|\right) \hat{O} \left(\left|\psi^{+\sigma}_{\beta/2}\right\rangle + \left|\psi^{-\sigma}_{\beta/2}\right\rangle\right)}{\left(\left\langle \psi^{+\sigma}_{\beta/2}\right| + \left\langle \psi^{-\sigma}_{\beta/2}\right|\right) \left(\left|\psi^{+\sigma}_{\beta/2}\right\rangle + \left|\psi^{-\sigma}_{\beta/2}\right\rangle\right)} , \tag{4.24}$$

with:

$$\left|\psi^{\pm\sigma}_{\beta/2}\right\rangle \equiv \prod_{l=1}^{P/2} \hat{U}\left[\pm\sigma(l)\right] \left|\psi_T\right\rangle , \tag{4.25}$$

$$\left\langle \psi^{\pm\sigma}_{\beta/2}\right| \equiv \left\langle \psi_T\right| \prod_{l=P/2+1}^{P} \hat{U}\left[\pm\sigma(l)\right] . \tag{4.26}$$

In practice we consider, for the initial trial determinant, two different one-body propagations, which differ in the sign of the auxiliary fields. Expression (4.24) reduces to the exact "$\pm\sigma$" scheme for $P = 2$ (only two Trotter time slices, or a simple Jastrow approach, see Eq. (4.23)). The larger the number of time slices is, the less efficient our "restricted $\pm\sigma$" scheme becomes, in order to decrease spurious statistical fluctuations. In fact we verified that, for example, using $\beta = 2$ a.u., with $\Delta\tau = 0.1$ a.u., that is 10 time slices on the left and 10 on the right, the advantage of employing Eq. (4.24), in place of the ordinary expression, was practically negligible. Therefore the "restricted $\pm\sigma$" technique is actually convenient only when one performs a simple variational Jastrow calculation (in that case the spurious one-body excitations are completely eliminated), or when a rather small number of propagation steps is used. Surely "intermediate" algorithms, in which the

"$\pm\sigma$" scheme is applied only a small number P' ($P/2 > P' > 1$) of times, during the imaginary time propagation from $\tau = 0$ to $\tau = \beta/2$, could be implemented. Anyway, in this case, a careful investigation would be necessary to find a suitable compromise between the reduction of statistical fluctuations one can achieve and the growing computer resources which are to be employed, by increasing the number of "$\pm\sigma$" applications.

Finally we observe that, if one makes use of a complex HST, where complex propagation potentials are generated, then the propagated wave functions become, in general, complex. However, at the end of our calculation, we want to get real quantities, corresponding to physical properties of the system we are studying. This can be accomplished simply by considering the real part ($\Re$) of the left and right propagated wave functions, for each sampled auxiliary field configuration. In fact:

$$\begin{aligned} e^{-\frac{\beta}{2}\hat{H}}\left|\psi_T\right\rangle &= \Re\left(e^{-\frac{\beta}{2}\hat{H}}\left|\psi_T\right\rangle\right) \\ &= \Re\left(\int_{-\infty}^{+\infty} d\sigma G(\sigma)\left|\psi^{\sigma}_{\beta/2}\right\rangle\right) \\ &= \int_{-\infty}^{+\infty} d\sigma G(\sigma)\Re\left(\left|\psi^{\sigma}_{\beta/2}\right\rangle\right) . \end{aligned} \tag{4.27}$$

Therefore:

$$\begin{aligned} \langle\hat{O}\rangle &= \frac{\langle\psi_T|\, e^{-\frac{\beta}{2}\hat{H}}\hat{O}e^{-\frac{\beta}{2}\hat{H}}\left|\psi_T\right\rangle}{\langle\psi_T|\, e^{-\frac{\beta}{2}\hat{H}}e^{-\frac{\beta}{2}\hat{H}}\left|\psi_T\right\rangle} = \frac{\displaystyle\int_{-\infty}^{+\infty} d\sigma G(\sigma)\Re\left(\left\langle\psi^{\sigma}_{\beta/2}\right|\right)\hat{O}\,\Re\left(\left|\psi^{\sigma}_{\beta/2}\right\rangle\right)}{\displaystyle\int_{-\infty}^{+\infty} d\sigma G(\sigma)\Re\left(\left\langle\psi^{\sigma}_{\beta/2}\right|\right)\Re\left(\left|\psi^{\sigma}_{\beta/2}\right\rangle\right)} \\ &= \frac{\displaystyle\int_{-\infty}^{+\infty} d\sigma G(\sigma)E_{\hat{O}}(\sigma)\left|\Re\left(\left\langle\psi^{\sigma}_{\beta/2}\right|\right)\Re\left(\left|\psi^{\sigma}_{\beta/2}\right\rangle\right)\right|}{\displaystyle\int_{-\infty}^{+\infty} d\sigma G(\sigma)S(\sigma)\left|\Re\left(\left\langle\psi^{\sigma}_{\beta/2}\right|\right)\Re\left(\left|\psi^{\sigma}_{\beta/2}\right\rangle\right)\right|} , \end{aligned} \tag{4.28}$$

where:

$$E_{\hat{O}}(\sigma) = \frac{\Re\left(\left\langle\psi^{\sigma}_{\beta/2}\right|\right)\hat{O}\,\Re\left(\left|\psi^{\sigma}_{\beta/2}\right\rangle\right)}{\left|\Re\left(\left\langle\psi^{\sigma}_{\beta/2}\right|\right)\Re\left(\left|\psi^{\sigma}_{\beta/2}\right\rangle\right)\right|}\,, \tag{4.29}$$

$$S(\sigma) = \frac{\Re\left(\left\langle\psi^{\sigma}_{\beta/2}\right|\right)\Re\left(\left|\psi^{\sigma}_{\beta/2}\right\rangle\right)}{\left|\Re\left(\left\langle\psi^{\sigma}_{\beta/2}\right|\right)\Re\left(\left|\psi^{\sigma}_{\beta/2}\right\rangle\right)\right|}\,, \tag{4.30}$$

and $\left|\psi^{\sigma}_{\beta/2}\right\rangle$ is to be substituted by $1/2\left(\left|\psi^{+\sigma}_{\beta/2}\right\rangle + \left|\psi^{-\sigma}_{\beta/2}\right\rangle\right)$ if the "$\pm\sigma$" scheme is used. In particular, if one employs the HST described by Eqs. (2.27) – (2.30), with a completely imaginary, auxiliary field-dependent, one-body potential, then our "restricted $\pm\sigma$" scheme is equivalent to consider the real part of the propagated determinant; that is, in this special case:

$$\frac{1}{2}\left(\left|\psi^{+\sigma}_{\beta/2}\right\rangle + \left|\psi^{-\sigma}_{\beta/2}\right\rangle\right) = \Re\left(\left|\psi^{+\sigma}_{\beta/2}\right\rangle\right)\,. \tag{4.31}$$

Clearly, in this approach, only "measuring" the $\hat{O}$ operator at the midpoint is well-defined and we give up the possibility (see Section 1.3.2) to compute an estimator at each time slice.

We must point out that Eqs. (4.27) – (4.30) do not represent the only way to recover real quantities from complex propagated wave functions. Another possibility, suggested in ref. [28], is to write:

$$\begin{aligned}\langle\hat{O}\rangle &= \frac{\int_{-\infty}^{+\infty} d\sigma G(\sigma)\Re\left(\left\langle\psi^{\sigma}_{\beta/2}\right|\hat{O}\left|\psi^{\sigma}_{\beta/2}\right\rangle\right)}{\int_{-\infty}^{+\infty} d\sigma G(\sigma)\Re\left(\left\langle\psi^{\sigma}_{\beta/2}\mid\psi^{\sigma}_{\beta/2}\right\rangle\right)} \\ &= \frac{\int_{-\infty}^{+\infty} d\sigma G(\sigma)E_{\hat{O}}(\sigma)\left|\left\langle\psi^{\sigma}_{\beta/2}\mid\psi^{\sigma}_{\beta/2}\right\rangle\right|}{\int_{-\infty}^{+\infty} d\sigma G(\sigma)W(\sigma)\left|\left\langle\psi^{\sigma}_{\beta/2}\mid\psi^{\sigma}_{\beta/2}\right\rangle\right|}\,,\end{aligned} \tag{4.32}$$

where:

$$E_{\hat{O}}(\sigma) = \frac{\Re\left(\left\langle \psi^{\sigma}_{\beta/2}\right| \hat{O} \left|\psi^{\sigma}_{\beta/2}\right\rangle\right)}{\left|\left\langle \psi^{\sigma}_{\beta/2} \mid \psi^{\sigma}_{\beta/2}\right\rangle\right|}\,, \tag{4.33}$$

$$W(\sigma) = \frac{\Re\left(\left\langle \psi^{\sigma}_{\beta/2} \mid \psi^{\sigma}_{\beta/2}\right\rangle\right)}{\left|\left\langle \psi^{\sigma}_{\beta/2} \mid \psi^{\sigma}_{\beta/2}\right\rangle\right|}\,. \tag{4.34}$$

We note that, while in Eq. (4.28) $S(\sigma) = \pm 1$, here $-1 \leq W(\sigma) \leq 1$, since $\left\langle \psi^{\sigma}_{\beta/2} \mid \psi^{\sigma}_{\beta/2}\right\rangle$ is generally a complex number $Z_\sigma = \rho e^{i\theta}$, so:

$$W(\sigma) = \frac{\Re(Z_\sigma)}{|Z_\sigma|} = \frac{\rho \Re(e^{i\theta})}{\rho} = \cos\theta\,. \tag{4.35}$$

Obviously, considering the averages over the probability distributions $G(\sigma)\,\Re\left(\left\langle \psi^{\sigma}_{\beta/2}\right|\right)\Re\left(\left|\psi^{\sigma}_{\beta/2}\right\rangle\right)$, and $G(\sigma)\left|\left\langle \psi^{\sigma}_{\beta/2} \mid \psi^{\sigma}_{\beta/2}\right\rangle\right|$, respectively, we have:

$$\langle W(\sigma)\rangle < \langle S(\sigma)\rangle\,. \tag{4.36}$$

In particular, even though $\langle S(\sigma)\rangle = 1$, that is the fermion sign problem is completely absent, nevertheless $\langle W(\sigma)\rangle < 1$. Therefore, from a statistical point of view, we expect to obtain worse results by this last scheme for which the fermion sign problem should be more troublesome. This was confirmed by some numerical tests, hence we have not used this approach.

Chapter 5

Results

The determination of energies of molecular systems is a problem of general interest in chemistry and physics. We have chosen the Hydrogen molecule to test our algorithm for various reasons. It is a very simple molecular system, the fermion sign problem is not present, in its Singlet ground state, because we have two electrons with opposite spins, and accurate theoretical predictions, together with high quality experimental measurements about ground state properties (in particular the dissociation energy), are available.

Although H_2 has been the object of numerous calculations over the years, the interest in this system, with its prototype chemical bond, remains high[75]. The history of accurate calculations of energies for H_2 begins with the 1933 paper of James and Coolidge[76]. Their work represented one of the first success in solving the Schrödinger equation for molecules. In the 1960's, more accurate results for the Hydrogen molecule were obtained by Kolos and Roothaan[77] and by Kolos and Wolniewicz[78], who established the foundation for future calculations. They implemented a variational approach in which the wave function is expressed in elliptic coordinates, and, using a method of Born[79], the Hamiltonian is separated into two parts, $\hat{H} = \hat{H}_0 + \hat{H}'$, where $\hat{H}_0$ is the electronic Hamiltonian including nuclear repulsion, and $\hat{H}'$ is the Hamiltonian for the nuclear motion including coupling between the electrons and the nuclei. The adiabatic approximation is

made by neglecting the off-diagonal contributions of $\hat{H}'$. Their calculations have been outlined in detail by Fischer [80]. Improvements by Wolniewicz [81], including a more flexible wave function, a variational-perturbation method to take into account the off-diagonal contributions to the exact nonrelativistic Hamiltonian, and the relativistic and radiative corrections, gave a better dissociation energy. In addition, Bishop and Cheung [82] calculated the energy of H_2 by treating the full four-body problem as a non-adiabatic variational problem.

In 1990 a very expensive calculation was performed by Traynor et al. [83], using a massively parallel supercomputer. They obtained the ground state energy of the Hydrogen molecule by the Quantum Monte Carlo method of solving the Schrödinger equation, without the use of the Born-Oppenhaimer or any other adiabatic approximation. The wave function sampling was carried out in the full 12-dimensional configuration space of the four particles (two electrons and two protons). Both a DMC and a GFMC algorithm were used. Their result is in close agreement with the best, experimentally determined dissociation energy of McCormack and Eyler [84], that is 36118.1 ± 0.2 cm^{-1}. Recently Sanders and Banyard [85] have illustrated the correlation effects in H_2, in particular the correlation influence on bonding and correlation-induced changes in the two-particle density.

We applied our AFQMC technique by expanding single particle wave functions in plane waves. Probably this basis set is not the optimum choice for the Hydrogen molecule, anyway we adopted it since we are more interested in testing an algorithm useful to study general Solid State systems (for which plane wave expansion is the standard implementation) than in obtaining very accurate data that are largely available in literature.

In Section 5.1 we used a "discretized" form (on a sparse lattice) of the H_2

Hamiltonian to accurately test the reliability of our algorithm, by comparing simulation results with exact, corresponding data.

In Section 5.2 several interesting physical quantities are shown and plotted. They were obtained employing a plane wave expansion with realistic energy cutoffs and give a sufficient, qualitative description of the correlation effects in the H_2 molecule.

Finally, in Section 5.3, an application of the AFQMC method to a more complex physical system, the H_3 molecule, is presented.

5.1 H_2: discretized Hamiltonian

After developing an algorithm which makes the statistical fluctuations acceptable, the first thing to do is to verify its actual accuracy in computing physical properties. To this end a comparison with abundant and precise literature data, for H_2, is impractical. In fact, if one uses the previously described AFQMC scheme, where an expansion in N_{pw} plane waves is performed and a finite imaginary time β (divided into P time slices $\Delta\tau$) is introduced, a meaningful comparison with really exact values would be possible only by extrapolating simulation results to values corresponding to $N_{\mathrm{pw}} \to \infty, \beta \to \infty, \Delta\tau \to 0$. In practice this would require a large number of rather long simulations making the computer cost very high.

Actually such a quite expensive approach is not necessary. In fact, due to the H_2 molecule simplicity (only two electrons are present), it is possible (in the sense that it needs an acceptable amount of computer time) to develop an algorithm, described in Appendix F, which is able to generate exact results, corresponding to fixed, finite N_{pw}, $\Delta\tau$, β values. Therefore, in the following, by "exact" we will

always mean: "results obtained by exactly propagating the trial wave function, using the same $N_{\rm pw}$, $\Delta\tau$, β parameters employed in the AFQMC simulation".

We have considered a $(8 \times 8 \times 8)$ spatial mesh in a cubic, periodically repeated cell, with volume $\Omega = a^3$ (a was chosen equal to 10 a.u.). The N_l^3 lattice points are given by:

$$\mathbf{r}_i = \Delta\,(i_x, i_y, i_z)\ , \tag{5.1}$$

where $\Delta = a/N_l$ ($N_l = 8$ in our case), and $i_x, i_y, i_z = 0, \ldots, N_l - 1$. The corresponding, discretized, cubic cell, in reciprocal space, is made up of G-vectors expressed as:

$$\mathbf{G}_n = \frac{2\pi}{a}\,(n_x, n_y, n_z)\ , \tag{5.2}$$

with $n_x, n_y, n_z = -N_l/2+1, \ldots, N_l/2$. All the relevant physical quantities (single particle wave functions, auxiliary field variables) are defined on the lattice sites and can be expanded in plane waves as:

$$f(\mathbf{r}_i) = \sum_{\mathbf{G}_n} \tilde{f}(\mathbf{G}_n) e^{i\mathbf{G}_n\cdot\mathbf{r}_i}\ , \tag{5.3}$$

with their Fourier coefficients given by:

$$\tilde{f}(\mathbf{G}_n) = \frac{1}{N_l^3} \sum_{\mathbf{r}_i} f(\mathbf{r}_i) e^{-i\mathbf{G}_n\cdot\mathbf{r}_i}\ . \tag{5.4}$$

Using this scheme no spherical cutoff in reciprocal space was introduced (all the G-vectors in the first Brillouin zone were taken into account).

In our AFQMC simulations the auxiliary fields were updated according to the MC sampling scheme of Section 3.1.2; the MC acceptance rate was about 60 - 70% in all the calculations. The "$\pm\sigma$" technique was also implemented. For the Jastrow factor we have chosen the form:

$$J(r) = \alpha e^{-r^2/r_{\rm J}^2}\ , \tag{5.5}$$

where α and r_j are two parameters that are to be optimized to minimize the ground state energy estimate. The Jastrow function (5.5) is different from the most used one[18], $J(r) = ar/(1+br)$; in particular it does not satisfy the "cusp condition"[18], but this fact is not important in our calculations where a finite, plane wave expansion is used and, therefore, the singularity in the Coulomb potential, $1/r$, is never, really, taken into account. Expression (5.5) has a simple analytical Fourier transform:

$$\tilde{J}(k) = \frac{\alpha}{\Omega}\pi^{3/2} r_j^3 e^{-k^2 r_j^2/4} . \tag{5.6}$$

In the practical implementation we used $\tilde{J}(\mathbf{G}_n)$ defined by Eq. (5.6) on reciprocal lattice G-vectors and, then, $J(\mathbf{r}_i)$ was obtained through Eq. (5.3).

First of all we show results for the H_2 molecule ground state (Singlet state). As we have pointed out in Section 4.2.1, in this particular case, a real HST (see Eqs. (4.11) - (4.16)) can be adopted, introducing an auxiliary field coupled to the local magnetization operator. We performed our simulations considering two different ionic configurations. In the first case the Hydrogen nuclei were placed at a relative distance of $R_1 = 1.401$ a.u. (the experimental equilibrium distance), while, in the second case, we used $R_2 = 4.0$ a.u.. Simulations were performed, both considering a simple variational Jastrow approach (see Section 1.3.3), and applying, to the trial determinant, a Jastrow operator plus a $\beta/2$ long imaginary time propagation. For the Trotter time slice $\Delta\tau$ we used $\Delta\tau = 0.1$ a.u., and the maximum β value, employed in our calculations, was $\beta = 2.0$ a.u.. In order to allow a comparison with "exact" data we have been obliged (see Appendix F) to use a quite sparse lattice with a relatively small ($8^3 = 512$) number of expansion plane waves. Hence the absolute values of our computed physical quantities are not particularly meaningful, the really interesting data, we have actually reported, being their difference

with respect to initial results, obtained using the trial, Hartree-Fock (H.F.) wave function. In particular, adopting this convention, our "total energy" results give directly the correlation energies of the systems we have studied.

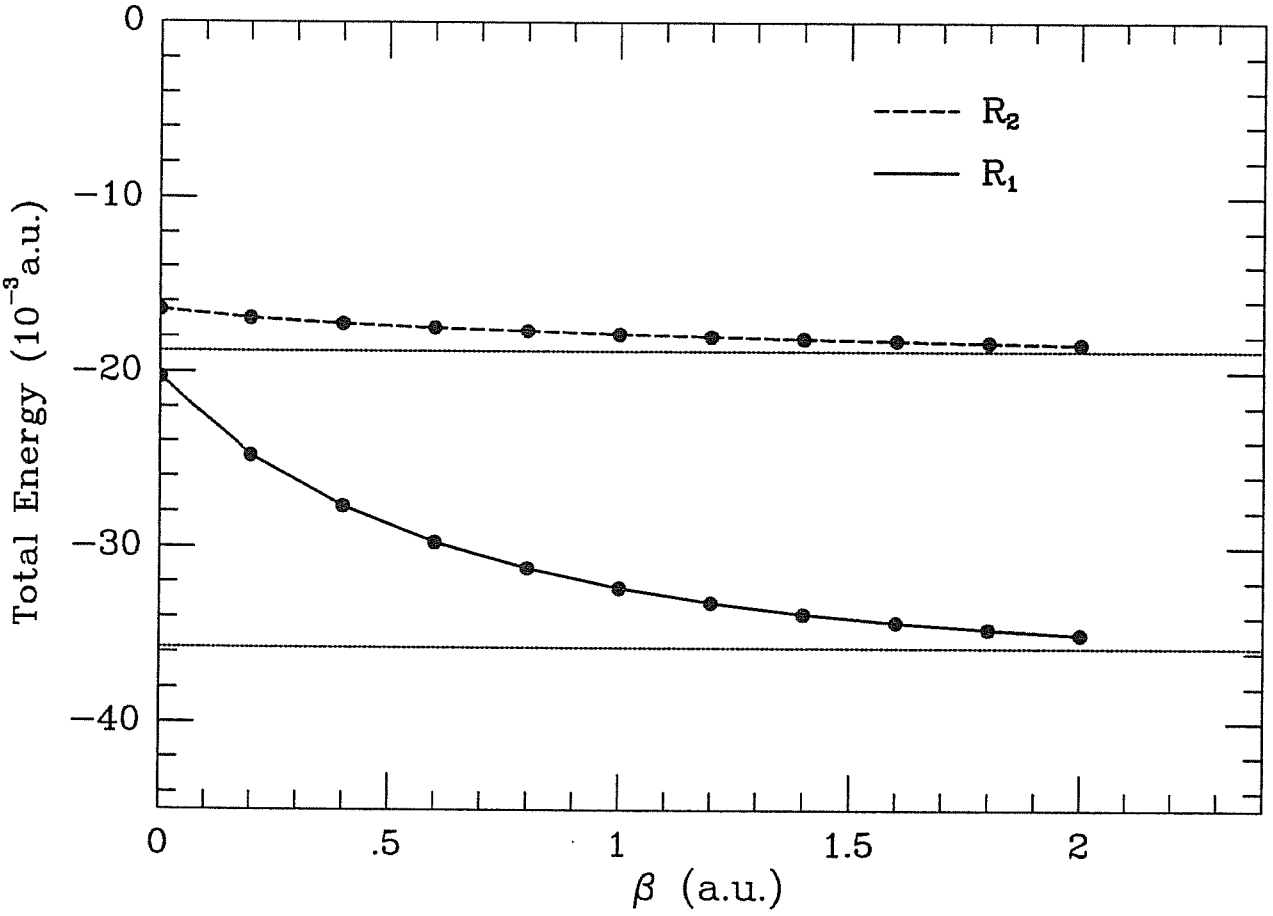

Figure 5. H_2 - Singlet state. Total energy behaviour (exact results), as a function of the imaginary time propagation β, for R_1=1.401 a.u. and R_2=4.0 a.u.; the β=0 values correspond to Jastrow variational estimates. The dotted, straight lines represent extrapolations for $\beta \rightarrow \infty$.

In Fig. 5 the behaviour of the total energies has been plotted, for the two different ionic configurations, by increasing the length β of the imaginary time propagation.

The energy zero value corresponds to initial H.F. result, while, for $\beta = 0$, we have reported the total energy obtained by a simple variational Jastrow scheme. Here we employed the "exact" algorithm (see Appendix F), just to check the convergence of the propagated determinant to the ground state. As one can see, at $\beta = 2.0$ a.u., making use of a Jastrow correlation function too, the total energies seem to have attained an acceptable level of convergence. The dotted, straight lines represent an estimate for the $\beta \to \infty$ value of the total energy, obtained by extrapolating from our available data (from $\beta = 0$ to $\beta = 2$ a.u.) and considering an exponential decay. Even though our Jastrow correlation functions were not accurately optimized, in order to minimize the total energy, the variational Jastrow approach was already able to recover a large amount of correlation energy ($\approx 60\%$ for the case with the R_1 distance, and $\approx 90\%$ for R_2). Regarding the Jastrow parameters α, r_j, we have used $\alpha = 0.45$ a.u., $r_j = 1.75$ a.u., and $\alpha = 0.80$ a.u., $r_j = 3.1$ a.u. for the R_1 and R_2 configurations, respectively. In practice the r_j parameter (see Eq. (5.5)) gives the range of the Jastrow correlation function. Therefore the fact that one has to employ a larger r_j value, when the nuclei are farther from each other, to minimize the total energy, is clearly understandable.

In Table I some H_2 ground state properties, total energy, kinetic energy, external energy (due to electron-ion interaction) and electron-electron energy, computed by AFQMC simulations and compared with corresponding exact values, are reported for both R_1 and R_2 configurations, using $\beta = 2.0$ a.u.. Only for the R_1 case, simulation results, for total and kinetic energies, employing a simple variational Jastrow scheme ($\beta = 0$) and an intermediate β value ($\beta = 0.6$ a.u.) are also shown. In order to establish credible statistical errors and to avoid problems due to possible correlation between successive sampled configurations, we divided

the whole set of simulation data (a typical run being made up of 10000 - 20000 auxiliary field configurations) into long enough segments, and then we estimated errors by using as independent data the averages computed within these intervals.

TABLE I. H_2 - Singlet state: AFQMC simulation results, using two different internuclear separations, R_1=1.401 a.u. and R_2=4.0 a.u., for some β values; β=0 means a simple variational Jastrow procedure. Statistical errors, in the last digit, are shown in square bracket. The "exact" data are also reported (in parenthesis). All the values are given in a.u. and are referred to H.F. corresponding data.

β	E_{tot}	E_{kin}	E_{ext}	E_{el}
0	-0.0207[5]	-0.1179[6]	——	——
(R_1)	(-0.02032)	(-0.11809)		
0.6	-0.0294[8]	-0.074[2]	——	——
(R_1)	(-0.02973)	(-0.07246)		
2.0	-0.034[1]	-0.017[2]	0.063[2]	-0.0805[4]
(R_1)	(-0.03496)	(-0.01854)	(0.06427)	(-0.08069)
2.0	-0.0187[9]	0.003[1]	0.0095[4]	-0.0315[1]
(R_2)	(-0.01840)	(0.00335)	(0.00972)	(-0.03148)

Table I results clearly indicate a good agreement (within the statistical errors) between simulation and exact data. We must point out that, although a relatively small number (8^3) of plane waves was used, the exact algorithm is much more expensive than the AFQMC technique (see also Appendix F). In fact, while a typical AFQMC simulation (10000 auxiliary field configurations) required about one hour

of CPU time, on a CRAY-2 supercomputer, a corresponding exact calculation took, at least, ten hours of computer time.

Next we report AFQMC applications to the H_2 Triplet, ${}^3\Sigma_u^+$, state, that is the state at the lowest energy, which can be obtained when the two electrons are forced to have parallel spins. In this case the two-body interaction term is given by $1/2\sum_{ij} V_{ij}\hat{\rho}_i^{\uparrow}\hat{\rho}_j^{\uparrow}$ and, therefore, one auxiliary field, coupled to up-up interaction, is to be introduced through a complex HST (see Eq. (2.33)). For the Triplet state two different calculations were carried out, always keeping the internuclear distance fixed at $R = 1.401$ a.u.. First we have performed a simulation using only the Jastrow factor ($\beta = 0$), that is a simple variational approach (with $\alpha = 0.2$ a.u., $r_j = 2.25$ a.u.). Then we have deliberately chosen an initial Slater determinant quite different from the H.F. one (the corresponding, reference total energy being much higher than H.F. value). In this case a $\beta = 2.0$ a.u. imaginary time propagation considerably lowers the total energy, as one can see in Fig. 6, where the exact data are reported.

In Table II a comparison between AFQMC simulation results and exact values is shown, considering both the Triplet calculations. In the Jastrow case ($\beta = 0$) the small $E_{\rm tot}$ value clearly indicates that, for the Triplet state, the H.F. approximation, which we started from, is rather good, as one can expect, since correlation between parallel spin electrons is weak. For this simple variational Jastrow AFQMC scheme the practical advantage of using the "$\pm\sigma$" algorithm (see Section 4.2.2) is remarkable. In fact we got a total energy statistical error equal to $8\cdot10^{-5}$ a.u., while the corresponding error, obtained without the "$\pm\sigma$" procedure, was much larger: $4\cdot10^{-4}$ a.u..

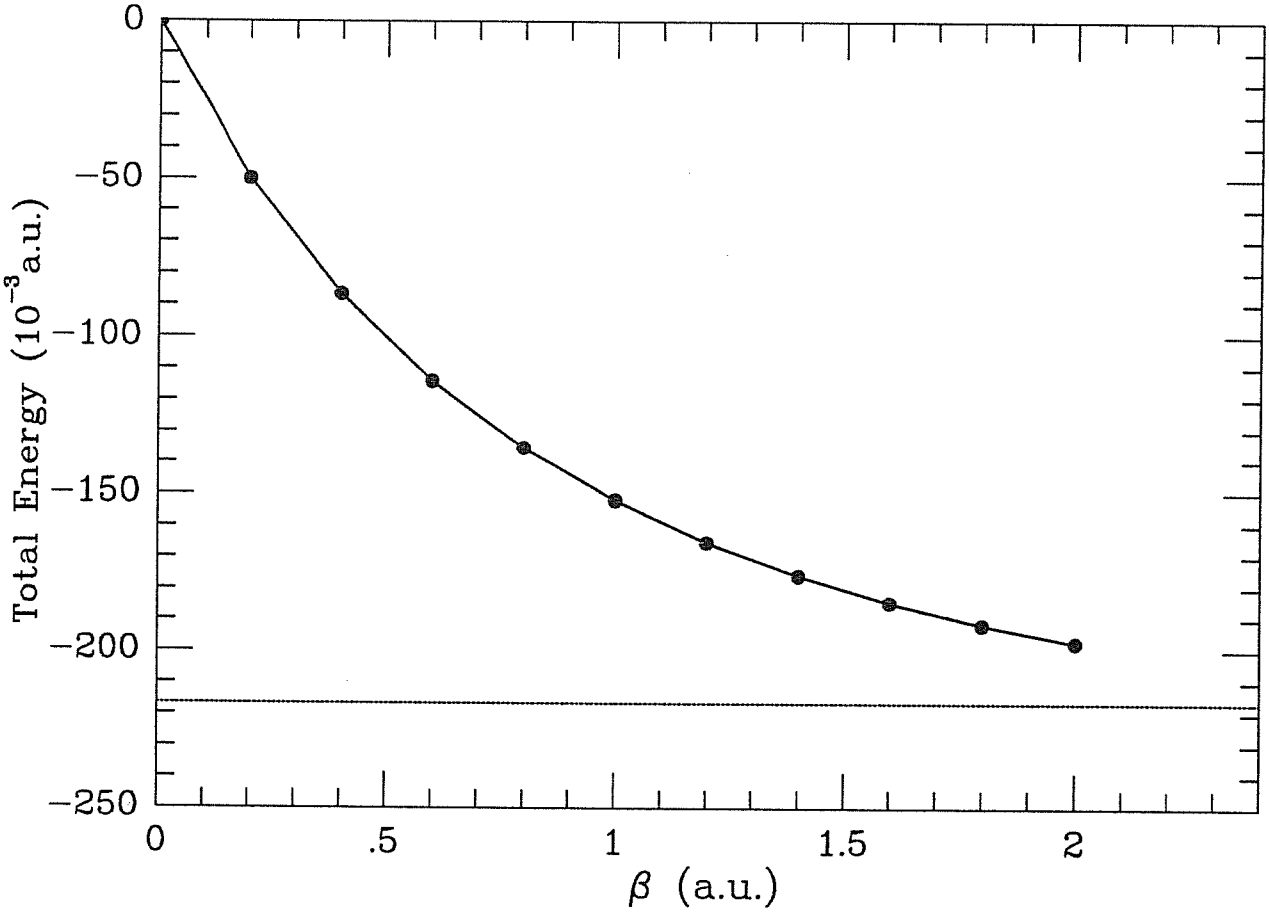

Figure 6. H_2 - Triplet state. Total energy behaviour (exact results), as a function of the imaginary time propagation β, for R=1.401 a.u.; the β=0 value corresponds to the initial trial determinant (different from H.F. solution) total energy. The dotted, straight line represents an extrapolation for $\beta \rightarrow \infty$.

In the Triplet simulations, in principle (see Section 1.3.4), the fermion sign problem could be present, since the fermion determinant can become negative. Anyway, in our H_2 tests, its average sign (the values in the last column of Table II) remained very far from zero, so a relatively accurate estimate of physical properties was certainly possible. Again a good agreement is found between simulation and exact results, by taking into account the AFQMC statistical errors, which can be lowered

TABLE II. H_2 - Triplet state: AFQMC simulation results, using R=1.401 a.u.; β=0 refers to a simple variational Jastrow procedure, while the β=2.0 a.u. propagation started from an initial Slater determinant, different from the H.F. one. Statistical errors, in the last digit, are shown in square bracket. The "exact" data are also reported (in parenthesis). All the values are given in a.u. and are referred to the trial determinant corresponding data. $\langle S \rangle$ represents the average sign of fermion determinant.

β	E_{tot}	E_{kin}	E_{ext}	E_{el}	$\langle S \rangle$
0	-0.00043[8]	-0.0023[6]	0.0045[6]	-0.00254[3]	1.0
	(-0.00040)	(-0.00253)	(0.00469)	(-0.00255)	
2.0	-0.200[4]	0.17[1]	-0.390[9]	0.018[1]	0.83
	(-0.19750)	(0.16648)	(-0.38227)	(0.01829)	

merely increasing the number of sampled auxiliary field configurations.

Hence, these calculations confirm that, at least for the simple H_2 system, our AFQMC scheme really works, the statistical fluctuations being reduced to an acceptable level.

5.2 H_2: continuous Hamiltonian

After the previous, preliminary tests we performed some more realistic H_2, AFQMC calculations. We gave up the possibility of an accurate comparison with exact data and we came back to usual Solid State approach (see Section 3.2). We no longer discretized the H_2 Hamiltonian. Simply we expanded single particle orbitals, using plane wave basis sets whose completeness can be monitored by a spherical kinetic energy cutoff, $E_{\rm cut}$ (all the plane waves with a squared wave vector smaller than $E_{\rm cut}$ are included in the basis set); correspondingly the Fourier transforms of the external (nuclear) and fluctuating (auxiliary) fields are truncated at a wave vector whose squared modulus is four times as large. Apart from this, the same AFQMC technique, used in the preceding Section, was employed.

TABLE III H_2 - Singlet state: AFQMC simulation results with a 6 Ry energy cutoff. β=0 refers to a simple variational Jastrow procedure. Statistical errors, in the last digit, are shown in square bracket. Initial H.F. determinant corresponding data are also reported. All the values are given in a.u..

β	E_{tot}	E_{kin}
0	-1.0067[4]	0.8453[6]
0.2	-1.0080[5]	0.8469[7]
0.4	-1.0072[5]	0.8515[7]
0.6	-1.0088[3]	0.8525[4]
0.8	-1.0087[6]	0.8552[9]
1.0	-1.0095[3]	0.8558[5]
H.F.	-0.98427	0.83890

First we report some results obtained with $E_{\rm cut} = 6$ Ry (the corresponding basis set included 251 plane waves). In Table III H_2 Singlet state simulation data, for some β values, together with the total and kinetic energies of the initial H.F. solution, are shown, using the experimental, equilibrium internuclear separation $R = 1.401$ a.u.. Again $\beta = 0$ means a simple variational Jastrow approach. The Jastrow parameters were $\alpha = 0.45$ a.u., $r_{\rm j} = 1.75$ a.u.. Now data are given as absolute values and they are no longer referred to initial H.F. data, which are also reported in the last row.

TABLE IV H_2 - Triplet state: AFQMC simulation results with a 6 Ry energy cutoff. β=0 refers to a simple variational Jastrow procedure. Statistical errors, in the last digit, are shown in square bracket. The initial Slater determinant was made up of LSDA single particle orbitals. All the values are given in a.u..

β	$E_{\rm tot}$	$E_{\rm kin}$	$\langle S \rangle$
0	-0.6706[2]	0.6267[7]	1.0
0.6	-0.6712[9]	0.6358[2]	0.99
1.0	-0.6725[4]	0.6355[9]	0.98
$\psi_T = \psi_{\rm LSDA}$	-0.66889	0.63783	

AFQMC simulation results, for the H_2 Triplet state, are shown in Table IV. In this case the initial wave function was a Slater determinant made up of single particle orbitals, which were the solutions of a previous LSDA calculation. $\beta = 0$ corresponds to a simple Jastrow approach, with parameters $\alpha = 0.35$ a.u., $r_{\rm j} = 2.75$ a.u.. By increasing the β value the total energy shows only a little decrease; again, this indicates that correlation energy is very small for the Triplet state. The average sign of the fermion determinant remained equal to 1, using the

simple variational Jastrow scheme, and it diminished very slowly, by considering longer and longer imaginary time propagations.

In the AFQMC method all the interesting physical properties can be computed without any particular difficulty. For instance, by considering the H_2 Singlet state, and performing a Jastrow AFQMC approach, we have obtained the electron charge density distribution $\rho(x)$ and the two-body, squared wave function $|\psi(x',x)|^2$, plotted along the H_2 molecular axis, in Fig. 7. As far as $\rho(x)$ is concerned, although the energy cutoff, $E_{\rm cut}$, is too small to distinguish the two cusps we expect to find on top of the two nuclei, the qualitative correlation effect is evident. In comparison with H.F. distribution, the AFQMC electronic density is slightly decreased in the centre of the molecule, while it is correspondingly increased near the nuclei. By fixing x', the $|\psi(x',x)|^2$ quantity represents the probability of finding one electron in position x if the other one is placed in x'. Again one can readily observe the correlated behaviour of the AFQMC result. In particular (see lower panel), if x' is chosen as the same position of one of the two nuclei, then the H.F. curve remains symmetric (in the H.F. approximation no correlation is taken into account between antiparallel spin particles), while the AFQMC one is evidently asymmetric; in fact the other electron has a tendency to stay near the opposite nucleus. In practice the $\rho(x)$ and $|\psi(x',x)|^2$ functions were computed at 100 selected positions on the molecular axis, and then, the resulting points, were connected by a continuous line. The corresponding statistical errors were too small to be drawn in the figure, being of the same order of the line width.

Then we considered a quite greater energy cutoff, $E_{\rm cut} = 14$ Ry (895 plane waves were included in the single particle wave function expansion).

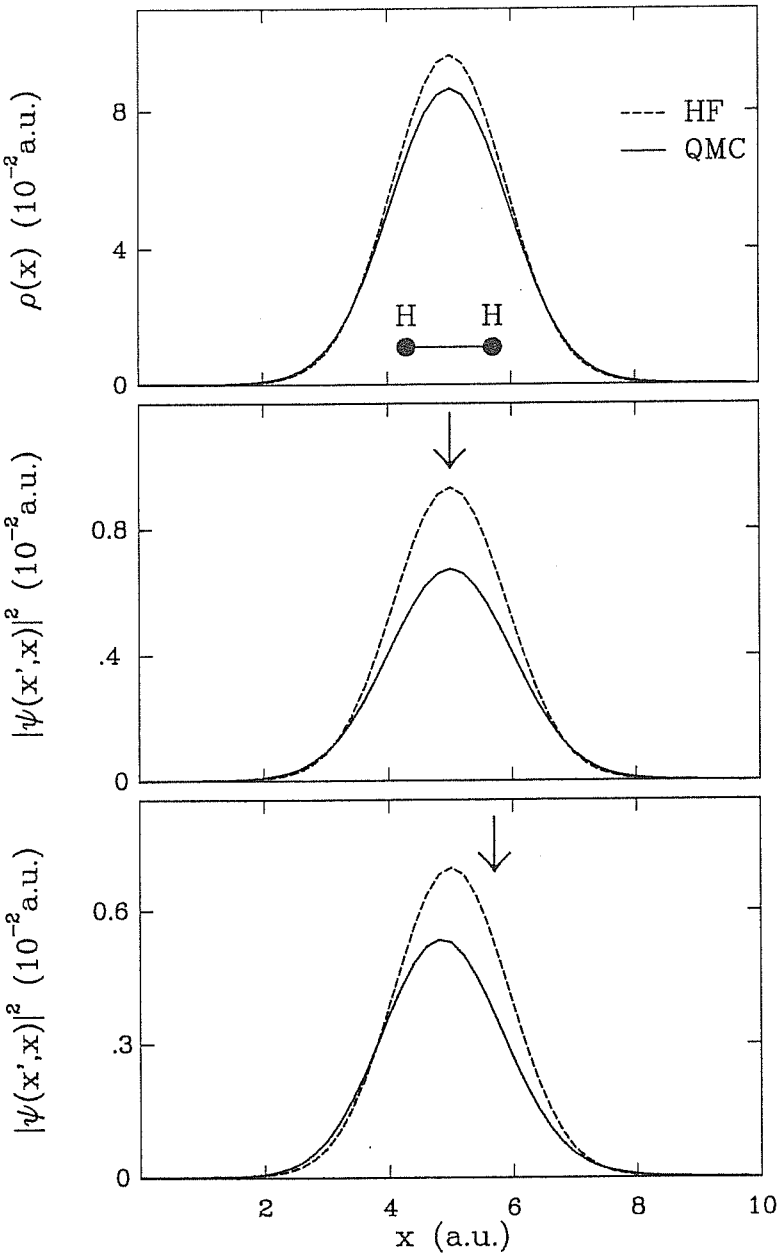

Figure 7. Electron charge density distribution ($\rho(x)$, upper panel) and squared wave function ($|\psi(x',x)|^2$, lower panels) of H_2 molecule, calculated along the molecular axis at the experimental internuclear separation (R=1.401 a.u.), with a 6 Ry energy cutoff. The continuous line indicates results from Jastrow AFQMC simulations, while Hartree-Fock results are denoted by a dashed line. The value of one of the two electronic coordinates (x') is indicated by an arrow in the lower two panels.

Again we studied the Singlet (at two different internuclear separations,

TABLE V. H_2 - Singlet and Triplet states: AFQMC simulation results, using a 14 Ry energy cutoff. The values in parenthesis refer to corresponding, initial, H.F. data. All the values are given in a.u..

	E_{tot}	E_{kin}	E_{ext}	E_{el}	$\langle S \rangle$
Singlet (R_1)	-1.092[3]	0.983[5]	-2.263[3]	0.2865[9]	1.0
H.F.	(-1.05838)	(0.99938)	(-2.32237)	(0.36387)	
Singlet (R_2)	-0.954[2]	0.524[6]	-1.007[5]	0.0546[3]	1.0
H.F.	(-0.90957)	(0.50128)	(-1.00710)	(0.12157)	
Triplet (R_1)	-0.740[2]	0.756[4]	-1.380[3]	-0.0169[2]	0.96
H.F.	(-0.73919)	(0.74392)	(-1.36832)	(-0.01553)	

$R_1 = 1.401$ a.u. and $R_2 = 4.0$ a.u.) and Triplet states.

In Table V numerical results for various physical quantities are reported. In all the cases we performed a $\beta = 2.0$ a.u. (with a Trotter time slice $\Delta\tau = 0.1$ a.u.) imaginary time propagation, starting from an initial H.F. solution. For the Singlet state configurations only, also a Jastrow correlation function (with parameters $\alpha = 0.45$ a.u., $r_j = 1.75$ a.u. for R_1, and $\alpha = 0.50$ a.u., $r_j = 2.75$ a.u. for R_2) was introduced. Comparing these simulation results with the previous, $E_{cut} = 6$ Ry, data, we can observe that statistical errors do not grow dramatically by increasing the energy cutoff. The calculated correlation energy, for the Singlet configuration, with R_1, is $E_{corr} = 0.034[3]$ a.u., to be compared with the exact value, quoted in literature[77], $E_{corr}^{exact} = 0.04081$ a.u..

Even by taking into account the statistical error there is a little discrepancy, due to

finite energy cutoff we imposed. Again the Triplet state correlation energy is very small, of the same order of magnitude, 10^{-3} a.u., of the statistical uncertainty affecting the total energy value (by performing a typical simulation with about 10000 auxiliary field configurations). In fact the initial H.F. determinant is a very good approximation for the exact Triplet wave function. In this case the average sign of the fermion determinant was $\langle S \rangle = 0.96$, therefore no numerical instability, due to the fermion sign problem, occurred. In Figures 8 and 9 we have plotted, for the two Singlet state configurations, the electron charge density distribution and the two-body squared wave function, along the H_2 molecular axis. The qualitative behaviours are similar to the previous ones, obtained with $E_{\rm cut} = 6$ Ry. The correlation effect is particularly evident when the internuclear separation is large ($R_2 = 4.0$ a.u.). In fact, in this Heitler-London regime, the H.F. solution is obviously a rather poor approximation for the molecule ground state.

Finally, for the H_2 molecule binding energy, defined as

$$E_{\rm bind} = E_{\rm tot}(H_2) - 2E_{\rm tot}(H) \ , \tag{5.7}$$

where $E_{\rm tot}(H_2)$ is our calculated Singlet state total energy and $E_{\rm tot}(H)$ is the single Hydrogen atom total energy (obtained using the same $E_{\rm cut} = 14$ Ry value), we found $E_{\rm bind} = -0.136[3]$ a.u., smaller than the exact result[77], $E_{\rm bind}^{\rm exact} = -0.17444$ a.u.. Probably this is mainly due to the fact that, using a finite energy cutoff, the equilibrium distance is different from the exact, experimental one. In fact, H_2 Local Density calculations, we have previously performed, indicate that, at $E_{\rm cut} = 14$ Ry, $R_{\rm eq}$ was about 10% larger than $R_{\rm eq}^{\rm exact} = 1.401$ a.u..

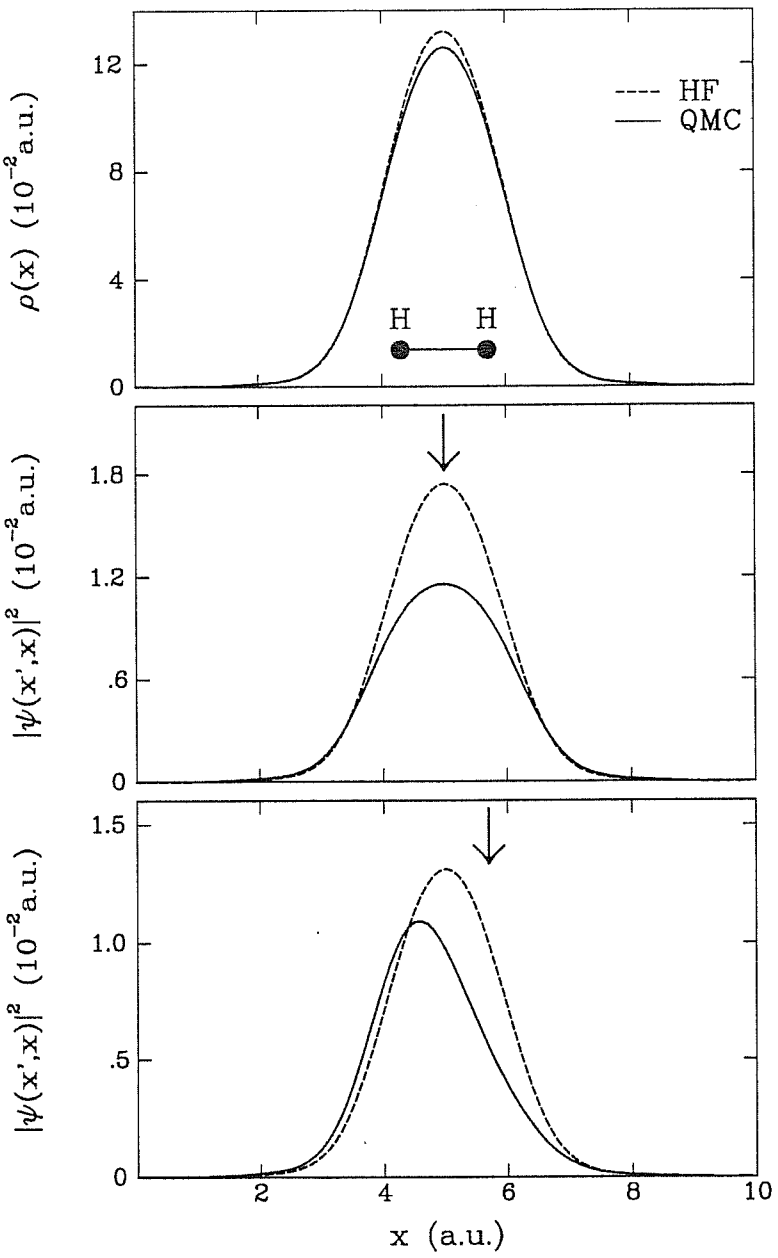

Figure 8. Electron charge density distribution ($\rho(x)$, upper panel) and squared wave function ($|\psi(x',x)|^2$, lower panels) of H_2 molecule, calculated along the molecular axis at the experimental internuclear separation (R_1=1.401 a.u.), with a 14 Ry energy cutoff. The continuous line indicates results from AFQMC simulations, while Hartree-Fock results are denoted by a dashed line. The value of one of the two electronic coordinates (x') is indicated by an arrow in the lower two panels.

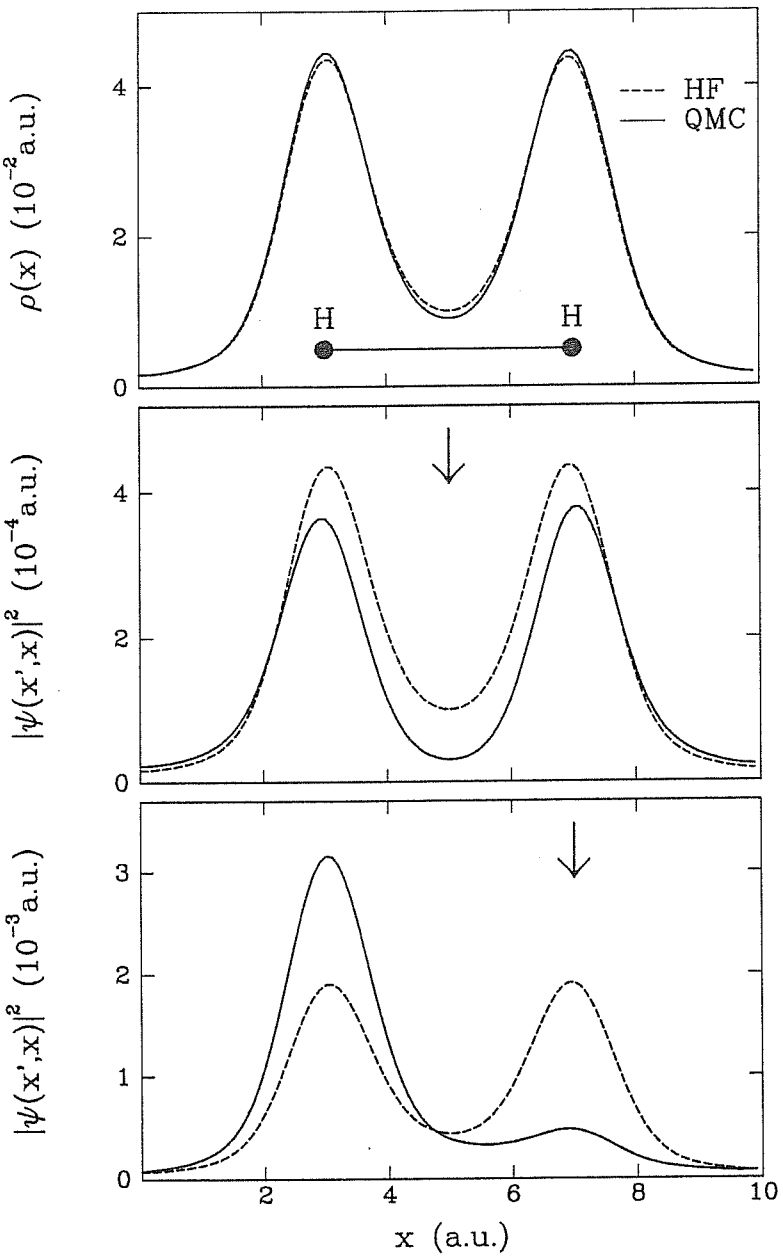

Figure 9. Same as Fig. 8, but for an internuclear separation representative of the Heitler-London regime (R_2=4.0 a.u.).

5.3 H_3

The H_3 system is of considerable interest in Quantum Chemistry, essentially because, from a dynamical point of view, the Hydrogen exchange reaction, $H + H_2 \rightarrow H_2 + H$, is probably the simplest of chemical reactions. There have been several calculations of high accuracy for this system. They include: analytic variational calculations by Siegbahn and Liu [86] (who obtained an accurate three-dimensional potential energy surface for H_3), and Liu [87]; fixed-node Quantum Monte Carlo simulations by Mentch and Anderson [88], and Barnett, Reynolds and Lester [89]; and release-node QMC calculations by Ceperley and Alder [20], and (using an improved algorithm) Anderson, Traynor and Boghosian [57].

For our AFQMC simulations we have chosen a triangular geometry (see Fig. 10) without any particular symmetry, for which previous calculations are available [20,86]. The coordinates (in a.u.) of the three Hydrogen nuclei are, respectively, $\mathbf{R}_1 = (3.17, 1.36, 0)$, $\mathbf{R}_2 = (-0.81, 0, 0)$, $\mathbf{R}_3 = (0.81, 0, 0)$. We employed the complex HST described in Eqs. (2.27) - (2.30), with an auxiliary field, σ^d, coupled to the total local density operator, and the simple MC algorithm introduced in Section 3.1.2. In Table VI our AFQMC numerical results are reported, using $E_{\rm cut} = 14$ Ry. We started from the initial H.F., Slater determinant, we applied a Jastrow operator (with the same parameters used for the H_2 Singlet state: $\alpha = 0.45$ a.u., $r_{\rm j} = 1.75$ a.u.), and then we performed a $\beta = 2.0$ a.u. imaginary time propagation (with $\Delta\tau = 0.1$ a.u. for the Trotter time slice). To reduce the statistical errors to the same level of our previous H_2 results we had to perform a very long AFQMC simulation, made up of 65000 auxiliary field configurations. The H_3 total energy values, which are available in literature [20,86], for the particular geometry we have chosen, are not directly comparable with our simulation

results, since we have used a plane wave expansion with a finite energy cutoff.

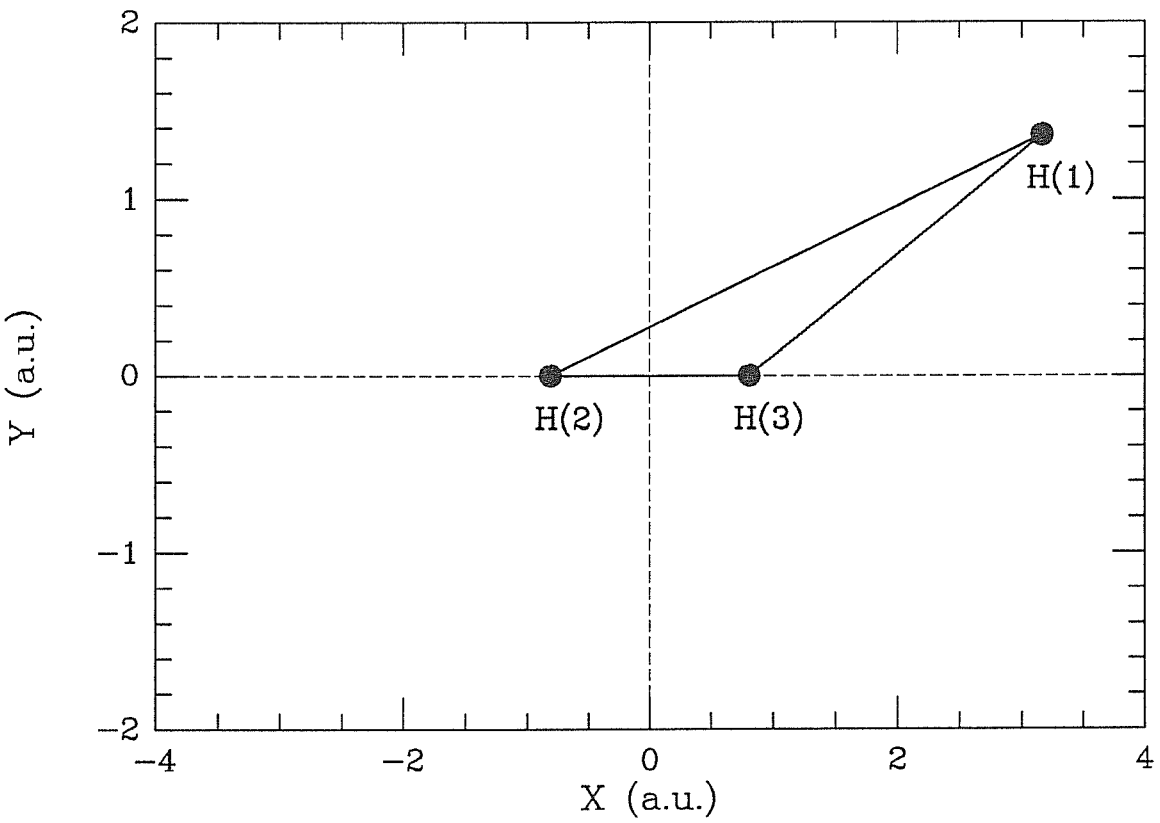

Figure 10. The geometry of the H_3 molecule studied.

A faster convergence, with respect to the energy cutoff, could be achieved by replacing the Coulomb interaction with a suitable pseudopotential which reproduces the correct electronic properties. In fact, if such a pseudopotential, $V_{\rm ps}(\mathbf{r})$, does not diverge as r^{-1}, for $r \to 0$, its Fourier transform, $V_{\rm ps}(\mathbf{G})$, decays more rapidly (by increasing G) than the $1/G^2$ function, which characterizes the Coulomb potential. In this way using very large energy cutoffs would be no longer necessary

TABLE VI. H_3 - AFQMC simulation results, using a 14 Ry energy cutoff. The values in parenthesis refer to the corresponding, initial, H.F. data. All the values are given in a.u..

β	E_{tot}	E_{kin}	E_{ext}	E_{el}	$\langle S \rangle$
2.0	-1.336[4]	0.960[6]	-2.094[4]	0.145[3]	0.38
H.F.	(-1.29321)	(0.85663)	(-2.00062)	(0.19750)	

to perform realistic calculations.

We notice that our total energy statistical error is an order of magnitude larger than the corresponding ones, obtained by the best H_3, QMC simulations [20]. We have also estimated other ground state properties, with about the same accuracy achieved for the total energy. We recall that the calculation of expectation values of operators which do not commute with the Hamiltonian is generally performed in an approximate way with other QMC schemes (see Section 1.2). Therefore the estimates of quantities different from the total energy are usually affected by much larger statistical errors.

The average sign of the fermion determinant was found to be $\langle S \rangle = 0.38$. This value is still not too small, although it is much smaller than in our previous H_2 Triplet state simulations. This indicates that, considering longer imaginary time propagations or more complex physical systems, the fermion sign problem could become really troublesome. In fact we expect that, whenever the average sign $\langle S \rangle$ becomes smaller than ≈ 0.1, one has to resort to approximations like the Positive-Projection method (see Appendix E) to get reasonable statistical errors in practical calculations.

We have performed some H_3 test simulations, making use of the complex HST

introduced at the end of Section 2.1.2, with three different auxiliary fields, σ^m, $\sigma^\uparrow$, $\sigma^\downarrow$. The results were substantially worse than adopting the HST with one auxiliary field, σ^d, only. In fact the average sign of the fermion determinant was $\langle S \rangle = 0.09$ and, correspondingly, the statistical errors were larger than in the previous case. Therefore this last HST seems to be less convenient.

For the H_3 molecule we have also computed the pair correlation function $g(\mathbf{r}', \mathbf{r})$. This quantity is the generalization of the function $|\psi(\mathbf{r}', \mathbf{r})|^2$ used for the H_2 Singlet state in Section 5.2. It gives the probability that, if one particle is observed at some point $\mathbf{r}'$, another particle will be found at point $\mathbf{r}$. By considering normalized quantities and taking spin variables into account, we can express the up-up and up-down pair correlation function operators as:

$$\hat{g}_{i\uparrow,j\uparrow} = \frac{1}{N^\uparrow(N^\uparrow - 1)} \left(\hat{\rho}_{i\uparrow}\hat{\rho}_{j\uparrow} - \delta_{ij}\hat{\rho}_{i\uparrow}\right) , \tag{5.8}$$

$$\hat{g}_{i\uparrow,j\downarrow} = \frac{1}{N^\uparrow N^\downarrow}\hat{\rho}_{i\uparrow}\hat{\rho}_{j\downarrow} , \tag{5.9}$$

where $\hat{\rho}_{i\mu} = c^\dagger_{i\mu}c_{i\mu}$ and $N^{\uparrow(\downarrow)}$ represents the total number of spin-up (down) electrons. In Eq. (5.8), we have properly subtracted the "autocorrelation", i.e. the correlation of a particle with itself. Obviously $\hat{g}_{i\uparrow,i\uparrow} = \hat{0}$, since $\hat{\rho}^2_{i\uparrow} = \hat{\rho}_{i\uparrow}$. Now the field-dependent estimators, for the pair correlation functions, can be readily obtained using the definitions (5.8) – (5.9) and the formulae introduced in Section 2.2:

$$E^\tau_{\hat{g}_{i\uparrow,j\uparrow}}(\sigma) = \frac{1}{N^\uparrow(N^\uparrow - 1)} \sum_{p,q,p',q'} (A^\uparrow)^{-1}_{pq}(A^\uparrow)^{-1}_{p'q'} \left[f^\uparrow_{qp}(\mathbf{r}_i)f^\uparrow_{q'p'}(\mathbf{r}_j) - f^\uparrow_{q'p}(\mathbf{r}_i)f^\uparrow_{qp'}(\mathbf{r}_j)\right]$$

$$E^\tau_{\hat{g}_{i\uparrow,j\downarrow}}(\sigma) = \frac{1}{N^\uparrow N^\downarrow} \sum_{p,q,p',q'} (A^\uparrow)^{-1}_{pq}(A^\downarrow)^{-1}_{p'q'} f^\uparrow_{qp}(\mathbf{r}_i)f^\downarrow_{q'p'}(\mathbf{r}_j) , \tag{5.10}$$

where:

$$f^\mu_{qp}(\mathbf{r}) = \varphi^<_{p\mu}(\mathbf{r})\varphi^>_{q\mu}(\mathbf{r}) . \tag{5.11}$$

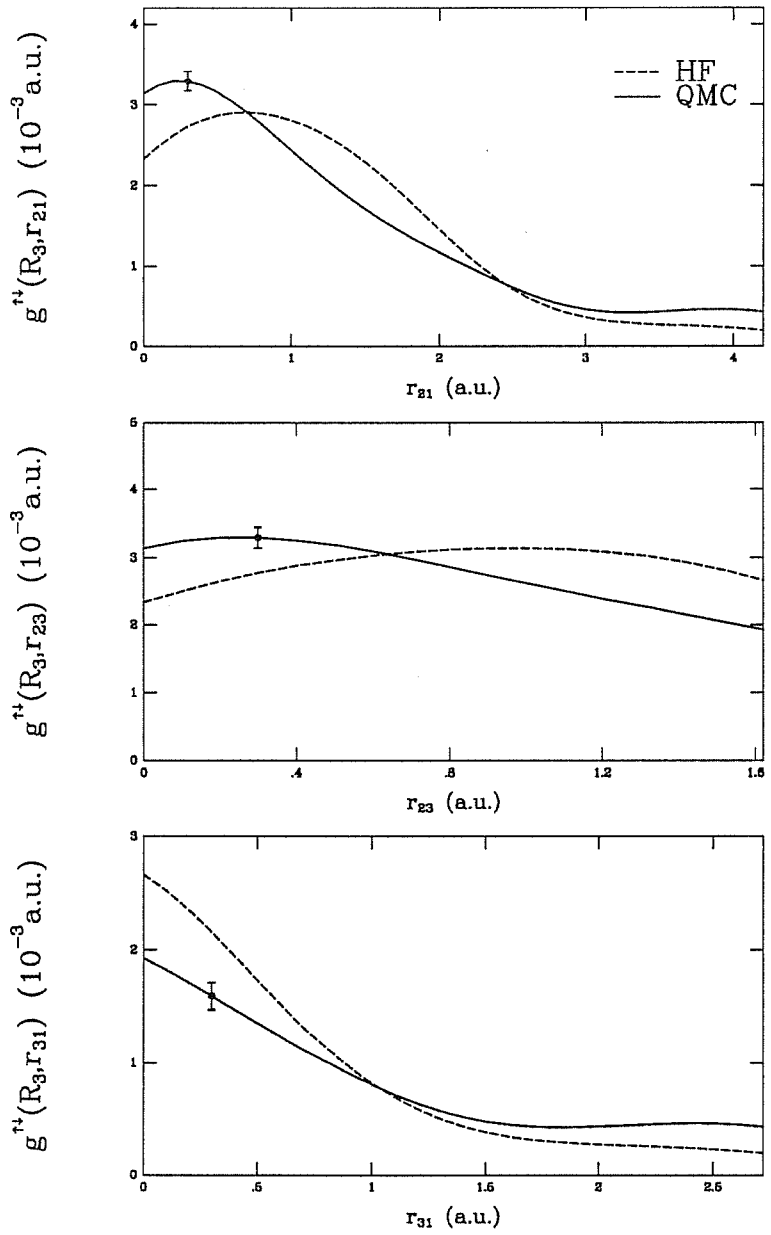

Figure 11. Electron pair correlation function for H_3 molecule, calculated along the internuclear axes, with a 14 Ry energy cutoff. The geometrical configuration is drawn in Fig. 10. One of the two pair correlation function variables was set equal to R_3 and the r_{21}, r_{23}, r_{31} coordinates are defined in Eq. (5.12). The continuous lines refer to our AFQMC simulations, while Hartree-Fock results are denoted by dashed lines. Typical error bars of the simulation are also reported.

In Fig. 11 we have reported the behaviour of the up-down pair correlation function $g_{\uparrow\downarrow}(\mathbf{R}_3, \mathbf{r}_{IJ})$. In practice we place a spin-up particle on top of the Hydrogen nucleus at $\mathbf{R}_3$, and we compute the probability of finding a spin-down particle at various positions, along the internuclear axes. The coordinates $\mathbf{r}_{IJ}$ are defined as:

$$\mathbf{r}_{IJ} = \mathbf{R}_I + (\mathbf{R}_J - \mathbf{R}_I) \frac{\mathrm{r}_{IJ}}{|\mathbf{R}_J - \mathbf{R}_I|} , \tag{5.12}$$

where $0 \leq \mathrm{r}_{IJ} \leq |\mathbf{R}_J - \mathbf{R}_I|$. Some typical statistical errors are shown in the Figure. The correlation effects are evident, by taking into account the H_3 geometry, illustrated in Fig. 10. In all cases the probability of finding an electron of opposite spin is reduced, with respect to H.F., in proximity to $\mathbf{R}_3$ and it is augmented far from it. The corresponding up-up pair correlation function, $g_{\uparrow\uparrow}(\mathbf{R}_3, \mathbf{r}_{IJ})$, is not too different from the H.F. one and it has not been reported. In fact up-up correlations are dominated by exchange effects which are already taken into account at the H.F. level.

Conclusions and Outlook

We have described, in detail, a method for the simulation of realistic many-electron systems. In this approach the Hubbard-Stratonovich Transformation allows us to replace direct electron-electron interactions with couplings to external auxiliary fields. Then, sums over these fields are performed statistically, using the fermion determinant and the Gaussian weight to guide the importance sampling by means of a stochastic (Monte Carlo) algorithm.

As in any other statistical method the most important problem is the minimization of the statistical error. In preliminary tests a serious difficulty occurred, by adopting the standard procedure to deal with repulsive Coulomb potentials, i.e. a modification of the two-body interaction. In fact huge fluctuations made a practical application of the AFQMC technique impossible, except that in situations of very small energy cutoff.

We have found a possible solution to this problem, by introducing a scheme which substantially reduces the statistical fluctuations. It is based on a different, generally complex, Hubbard-Stratonovich Transformation, that avoids introducing a suitably modified two-body interaction.

We have accurately tested our technique by applying it to the H_2 molecule and studying various physical properties of its Singlet and Triplet states. The Singlet state is a particularly favourable case, since a convenient real HST can be

employed. Our results, for H_2, are quite satisfactory: the method works, and an acceptable level of statistical accuracy can be achieved. We have also verified that the statistical fluctuations grow slowly by increasing the energy cutoff, used for the plane wave expansion of the single particle wave functions.

Applications to the H_3 molecule show that our technique still works, since the statistical errors are not much larger than those affecting H_2 numerical results. However the average sign of the fermion determinant tends to become rather small, by using a relatively long imaginary time propagation. This suggests that, for more complex systems, some suitable procedure (like the Positive-Projection technique), which is able to cope with the fermion sign problem, will be probably necessary.

Surely further improvements are possible. In particular, making use of appropriate pseudopotentials, both for the electron-electron and for the electron-ion interaction, would avoid employing very large energy cutoffs to get realistic physical results. Moreover a more sophisticated version of the "$\pm\sigma$" scheme, introduced in Section 4.2.2, could be applied to obtain an additional reduction of the statistical errors. At present, in comparison with other QMC techniques, our AFQMC method is significantly less accurate in the calculation of the total ground state energy. However it could offer some advantages, especially for what concerns the estimate of general observables and the introduction of non-local pseudopotentials.

The next crucial step, to establish the actual utility of our scheme, will be testing whether it may be really applied to more interesting and complex systems. This work is currently in progress, together with a more accurate investigation about useful technical improvements that can be introduced and with a careful analysis of the fermion sign problem.

Appendix A

Derivation of Hubbard-Stratonovich Transformation

We must prove the following identity for multidimensional integrals over real variables:

$$[\det(\beta A)]^{-\frac{1}{2}} e^{\frac{\beta}{2}\sum_{ij} A^{-1}_{ij}\rho_i\rho_j} = \frac{1}{(2\pi)^{\frac{n}{2}}}\int dx_1 \ldots dx_n\, e^{-\frac{\beta}{2}\sum_{ij} A_{ij}x_ix_j+\beta\sum_i x_i\rho_i}\,, \tag{A.1}$$

where A is a real, symmetric, *positive definite*, $(n \times n)$ matrix. This identity is straightforwardly established by changing variables to reduce it to diagonal form and using the familiar Gaussian integral:

$$\sqrt{\frac{\pi}{a}} = \int_{-\infty}^{+\infty} dx e^{-ax^2}\,. \tag{A.2}$$

First of all let us perform the transformation:

$$y_i = x_i - \sum_j A^{-1}_{ij}\rho_j\,. \tag{A.3}$$

By using the relation $\sum_k A_{ik}A^{-1}_{kj} = \delta_{ij}$ and the symmetry of A we can rewrite the integral on the R.H.S. of Eq. (A.1):

$$\begin{aligned}&\int dx_1 \ldots dx_n e^{-\frac{\beta}{2}\sum_{ij} A_{ij}x_ix_j+\beta\sum_i x_i\rho_i}\\ &= \int dy_1 \ldots dy_n e^{-\frac{\beta}{2}\sum_{ij} A_{ij}y_iy_j+\frac{\beta}{2}\sum_{ij} A^{-1}_{ij}\rho_i\rho_j}\,.\end{aligned} \tag{A.4}$$

Then we introduce the variables $z_k = \sum_i O_{ki} y_i$, where O is the orthogonal transformation which diagonalizes A:

$$\sum_{ij} O_{ki}\, A_{ij}\, O^{-1}_{jm} = \alpha_m \delta_{km} \ . \tag{A.5}$$

With this transformation the R.H.S. of Eq. (A.4) becomes:

$$\begin{aligned} & e^{\frac{\beta}{2}\sum_{ij} A^{-1}_{ij}\rho_i\rho_j} \int dy_1 \ldots dy_n e^{-\frac{\beta}{2}\sum_{ij} A_{ij} y_i y_j} \\ =& e^{\frac{\beta}{2}\sum_{ij} A^{-1}_{ij}\rho_i\rho_j} \int dz_1 \ldots dz_n e^{-\frac{\beta}{2}\sum_{l} z_l^2 \alpha_l} \ , \end{aligned} \tag{A.6}$$

where we have used the fact that O is a unitary transformation, that is $\sum_k O_{ki} O_{kj} = \sum_k O^{-1}_{ik} O_{kj} = \delta_{ij}$, and α_l are the eigenvalues of A. Now we can directly apply the well-known Eq. (A.2):

$$e^{\frac{\beta}{2}\sum_{ij} A^{-1}_{ij}\rho_i\rho_j} \int dz_1 \ldots dz_n e^{-\frac{\beta}{2}\sum_{l} z_l^2 \alpha_l} = e^{\frac{\beta}{2}\sum_{ij} A^{-1}_{ij}\rho_i\rho_j} \prod_l \sqrt{\frac{2\pi}{\alpha_l \beta}} \, . \tag{A.7}$$

Finally we can write:

$$\prod_l \sqrt{\frac{2\pi}{\alpha_l \beta}} = \frac{(2\pi)^{\frac{n}{2}}}{[\det(\beta A)]^{\frac{1}{2}}} \ , \tag{A.8}$$

and therefore we have shown that:

$$\int dx_1 \ldots dx_n e^{-\frac{\beta}{2}\sum_{ij} A_{ij} x_i x_j + \beta \sum_i x_i \rho_i} = [\det(\beta A)]^{-\frac{1}{2}} (2\pi)^{\frac{n}{2}} e^{\frac{\beta}{2}\sum_{ij} A^{-1}_{ij}\rho_i\rho_j} \ , \tag{A.9}$$

and hence identity (A.1) is proved. Note that the positivity of all eigenvalues α_l of A, that is the fact that A is a positive definite matrix, is essential for convergence of Gaussian integrals and, therefore, a crucial condition for the validity of the proof.

If we consider a real, symmetric, *negative definite*, $(n \times n)$ matrix B, then the Hubbard-Stratonovich Transformation becomes:

$$[(-1)^n \det(\beta B)]^{-\frac{1}{2}} e^{\frac{\beta}{2}\sum_{ij} B^{-1}_{ij}\rho_i\rho_j} = \frac{1}{(2\pi)^{\frac{n}{2}}} \int dx_1 \ldots dx_n \, e^{\frac{\beta}{2}\sum_{ij} B_{ij} x_i x_j + i\beta \sum_i x_i \rho_i} \tag{A.10}$$

where an imaginary factor "i" must be introduced. This can be demonstrated in the following way. Let us consider again the orthogonal transformation O, which diagonalizes B, and the new variables $z_k = \sum_i O_{ki} x_i$. Then:

$$\begin{aligned} &\int dx_1 \dots dx_n e^{\frac{\beta}{2}\sum_{ij} B_{ij} x_i x_j + i\beta \sum_i x_i \rho_i} \\ &= \int dz_1 \dots dz_n e^{\frac{\beta}{2}\sum_k b_k z_k^2 + i\beta \sum_k c_k z_k} , \end{aligned} \tag{A.11}$$

where b_k are the (negative) eigenvalues of B, and $c_k \equiv \sum_i O_{ik}^{-1} \rho_i$. Now:

$$\begin{aligned} &\int dz_1 \dots dz_n e^{\frac{\beta}{2}\sum_k b_k z_k^2 + i\beta \sum_k c_k z_k} \\ &= \prod_k \left(\int_{-\infty}^{+\infty} dz_k e^{\frac{\beta}{2} b_k z_k^2 + i\beta c_k z_k} \right) \\ &= \prod_k \left(\int_{-\infty}^{+\infty} dz_k e^{\frac{\beta}{2} b_k z_k^2} \cos\left(\beta c_k z_k\right) \right) . \end{aligned} \tag{A.12}$$

At this point we use the basic formula (for $b < 0$):

$$\sqrt{\frac{2\pi}{(-b)}}\, e^{c^2/2b} = \int_{-\infty}^{+\infty} dz\, e^{1/2bz^2} \cos(cz) , \tag{A.13}$$

to write:

$$\begin{aligned} \prod_k \left(\int_{-\infty}^{+\infty} dz_k e^{\frac{\beta}{2} b_k z_k^2} \cos\left(\beta c_k z_k\right) \right) &= \prod_k \left(\sqrt{\frac{2\pi}{(-\beta b_k)}}\, e^{\beta c_k^2/2b_k} \right) \\ &= \frac{(2\pi)^{\frac{n}{2}}}{\left[(-1)^n \det\left(\beta B\right)\right]^{\frac{1}{2}}} e^{\frac{\beta}{2}\sum_k c_k^2/b_k} . \end{aligned} \tag{A.14}$$

Now:

$$\sum_k \frac{c_k^2}{b_k} = \sum_{ijk} \frac{O_{ik}^{-1} O_{jk}^{-1}}{b_k} \rho_i \rho_j = \sum_{ij} B_{ij}^{-1} \rho_i \rho_j . \tag{A.15}$$

Therefore:

$$e^{\frac{\beta}{2}\sum_k c_k^2/b_k} = e^{\frac{\beta}{2}\sum_{ij} B_{ij}^{-1} \rho_i \rho_j} , \tag{A.16}$$

and, considering Eq. (A.14), identity (A.10) is proved.

We note that (A.10) could be formally obtained from identity (A.1), simply by using the transformation $\rho_i \rightarrow i\rho_i$, and observing that $B_{ij} = -A_{ij}$, A being a positive definite matrix.

Appendix B

Higher order correlation function estimator

Here we want to show that (see Section 2.2):

$$\langle c_i^\dagger c_j c_m^\dagger c_n \rangle = \langle c_i^\dagger c_j \rangle \langle c_m^\dagger c_n \rangle + \langle c_i^\dagger c_n \rangle \langle c_j c_m^\dagger \rangle \,. \tag{B.1}$$

Let us consider the quantity:

$$\langle c_j c_n c_i^\dagger c_m^\dagger \rangle \equiv \frac{\left\langle c_n^\dagger c_j^\dagger \hat{U}_\sigma(\tau,\beta)\,\psi_T \middle| c_i^\dagger c_m^\dagger \hat{U}_\sigma(\tau,0)\,\psi_T \right\rangle}{\left\langle \hat{U}_\sigma(\tau,\beta)\,\psi_T \middle| \hat{U}_\sigma(\tau,0)\,\psi_T \right\rangle} \,. \tag{B.2}$$

This involves the scalar product of two $(N+2)$-state determinants and we can write:

$$\langle c_j c_n c_i^\dagger c_m^\dagger \rangle = \frac{\det \tilde{A}(jn,im)}{\det A} \,, \tag{B.3}$$

where A is the usual overlap matrix:

$$A_{pq} = \langle \varphi_p^< | \varphi_q^> \rangle \,, \tag{B.4}$$

and:

$$\tilde{A}_{pq}(jn,im) = \begin{pmatrix} \delta_{jm} & \delta_{nm} & \dots & \varphi_q^>(\mathbf{r}_m) & \dots \\ \delta_{ji} & \delta_{ni} & \dots & \varphi_q^>(\mathbf{r}_i) & \dots \\ \vdots & \vdots & & & \\ \varphi_p^<(\mathbf{r}_j) & \varphi_p^<(\mathbf{r}_n) & & A_{pq} & \\ \vdots & \vdots & & & \end{pmatrix} . \tag{B.5}$$

Let us introduce the quantity:

$$B(j,m) = \sum_{pq} \varphi_p^{>}(\mathbf{r}_m) A_{pq}^{-1} \varphi_q^{<}(\mathbf{r}_j) . \tag{B.6}$$

Now, since a determinant remains unchanged if one adds to a column any linear combination of the others, we may add to the *first column* of the matrix $\tilde{A}$ a linear combination of the other columns, in order to make vanishing all the elements of the first column but the ones in the *first two rows*:

$$\det\tilde{A} = \det \begin{pmatrix} \delta_{jm} - b_0\delta_{nm} - \sum_q b_q \varphi_q^{>}(\mathbf{r}_m) & \dots & \dots \\ \delta_{ji} - b_0\delta_{ni} - \sum_q b_q \varphi_q^{>}(\mathbf{r}_i) & \dots & \dots \\ \vdots & & \\ \varphi_p^{<}(\mathbf{r}_j) - b_0\varphi_p^{<}(\mathbf{r}_n) - \sum_q b_q A_{pq} & \dots & \dots \\ \vdots & & \end{pmatrix} . \tag{B.7}$$

We choose $b_0 = 0$ and $b_q = \sum_{q'} A_{qq'}^{-1} \varphi_{q'}^{<}(\mathbf{r}_j)$. Then we obtain:

$$\det\tilde{A} = \det \begin{pmatrix} \delta_{jm} - B(j,m) & \delta_{nm} & \dots & \varphi_q^{>}(\mathbf{r}_m) & \dots \\ \delta_{ji} - B(j,i) & \delta_{ni} & \dots & \varphi_q^{>}(\mathbf{r}_i) & \dots \\ \vdots & \vdots & & & \\ 0 & \varphi_p^{<}(\mathbf{r}_n) & & A_{pq} & \\ \vdots & \vdots & & & \end{pmatrix} . \tag{B.8}$$

Now we can repeat the same procedure in order to make vanishing all the elements of the *second column* but the ones in the *first two rows* and we have:

$$\det\tilde{A} = \det \begin{pmatrix} \delta_{jm} - B(j,m) & \delta_{nm} - B(n,m) & \dots & \varphi_q^{>}(\mathbf{r}_m) & \dots \\ \delta_{ji} - B(j,i) & \delta_{ni} - B(n,i) & \dots & \varphi_q^{>}(\mathbf{r}_i) & \dots \\ \vdots & \vdots & & & \\ 0 & 0 & & A_{pq} & \\ \vdots & \vdots & & & \end{pmatrix} , \tag{B.9}$$

that is:

$$\det\tilde{A} = \det \begin{pmatrix} B_{(2\times 2)} & C_{(2\times N)} \\ 0_{(N\times 2)} & A_{(N\times N)} \end{pmatrix} = \det B \cdot \det A . \tag{B.10}$$

Therefore, by using Eqs. (B.3) and (2.50):

$$\begin{aligned}\langle c_j c_n c_i^\dagger c_m^\dagger \rangle &= \det B \\ &= [\delta_{jm} - B(j,m)] \cdot [\delta_{ni} - B(n,i)] - [\delta_{nm} - B(n,m)] \cdot [\delta_{ji} - B(j,i)] \\ &= \langle c_j c_m^\dagger \rangle \langle c_n c_i^\dagger \rangle - \langle c_n c_m^\dagger \rangle \langle c_j c_i^\dagger \rangle \,. \end{aligned} \tag{B.11}$$

Finally, by usual anticommutation rules for fermion operators c, $c^\dagger$, we obtain:

$$\begin{aligned}\langle c_i^\dagger c_j c_m^\dagger c_n \rangle &= -\langle c_j c_n c_i^\dagger c_m^\dagger \rangle + \delta_{in} \langle c_j c_m^\dagger \rangle - \delta_{ij} \langle c_n c_m^\dagger \rangle + \delta_{nm} \langle c_i^\dagger c_j \rangle \\ &= \langle c_i^\dagger c_j \rangle \langle c_m^\dagger c_n \rangle + \langle c_i^\dagger c_n \rangle \langle c_j c_m^\dagger \rangle \,. \end{aligned} \tag{B.12}$$

Appendix C

Formulae in reciprocal space

Let us consider the plane wave expansion:

$$\varphi_p(\mathbf{r}) = \sum_{\mathbf{G}} C_p(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} , \tag{C.1}$$

with Fourier coefficients given by:

$$C_p(\mathbf{G}) = \frac{1}{\Omega} \int d\mathbf{r} \varphi_p(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} . \tag{C.2}$$

The estimator of total ground state energy, evaluated at a particular auxiliary field configuration σ and at a fixed imaginary time τ, is:

$$E_H^\tau(\sigma) = E_{\hat{T}}^\tau(\sigma) + E_{\hat{V}^{\rm ext}}^\tau(\sigma) + E_{\hat{V}^{\rm ee}}^\tau(\sigma) + E_{\rm ion} . \tag{C.3}$$

Then, by using definitions (C.1), (C.2) and relations (2.54) – (2.57), we can write the various components in the following way:

$$E_{\hat{T}}^\tau(\sigma) = \frac{1}{2} \sum_{\mu} \sum_{pq} (A^\mu)^{-1}_{pq} \sum_{\mathbf{G}} G^2 C_q^{*<}(\mathbf{G}) C_p^{>}(\mathbf{G}) , \tag{C.4}$$

$$E_{\hat{V}^{\rm ext}}^\tau(\sigma) = \sum_{\mu} \sum_{pq} (A^\mu)^{-1}_{pq} \sum_{\mathbf{G}\neq 0} V^{\rm ext}(\mathbf{G}) f^*_{qp}(\mathbf{G}) , \tag{C.5}$$

$$\begin{aligned} E_{\hat{V}^{\rm ee}}^\tau(\sigma) = {} & \frac{2\pi}{\Omega} \sum_{\mu\mu'} \sum_{pqp'q'} (A^\mu)^{-1}_{pq} (A^{\mu'})^{-1}_{p'q'} \sum_{\mathbf{G}\neq 0} \frac{1}{G^2} f^{*\mu}_{qp}(\mathbf{G}) f^{*\mu'}_{q'p'}(\mathbf{G}) - \\ & - \frac{2\pi}{\Omega} \sum_{\mu} \sum_{pqp'q'} (A^\mu)^{-1}_{pq} (A^{\mu})^{-1}_{p'q'} \sum_{\mathbf{G}\neq 0} \frac{1}{G^2} f^{*\mu}_{q'p}(\mathbf{G}) f^{*\mu}_{qp'}(\mathbf{G}) + \\ & + N \left(\frac{2\pi}{\Omega} \sum_{\mathbf{G}\neq 0} \frac{1}{G^2} - \frac{1}{\pi} G_{\rm cut} \right) , \end{aligned} \tag{C.6}$$

$$E_{\rm ion} = 2\pi\Omega \sum_{\mathbf{G}\neq 0} \frac{|\rho_{\rm Gauss}(\mathbf{G})|^2}{G^2} + \frac{1}{2}\sum_{I\neq J} Z_I Z_J \frac{\mathrm{erfc}\left(|\mathbf{R}_I - \mathbf{R}_J|\sqrt{\frac{2}{\alpha}}\right)}{|\mathbf{R}_I - \mathbf{R}_J|} - \sqrt{\frac{\alpha}{2\pi}}\sum_I Z_I^2 \tag{C.7}$$

where:

$$A_{pq} \equiv \sum_{\mathbf{G}} C_p^{*<}(\mathbf{G}) C_q^{>}(\mathbf{G}) \,, \tag{C.8}$$

$$f_{qp}(\mathbf{G}) \equiv \frac{1}{\Omega}\int d\mathbf{r}\varphi_q^{<}(\mathbf{r})\varphi_p^{>}(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} \,, \tag{C.9}$$

$$V^{\rm ext}(\mathbf{G}) \equiv \frac{1}{\Omega}\int d\mathbf{r} V^{\rm ext}(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} \,. \tag{C.10}$$

The last term in Eq. (C.6), with $G_{\rm cut} = \sqrt{2E_{\rm cut}}$, is due to the fact that a periodically repeated system is considered. Therefore, in order to obtain electron-electron interaction per unit cell, one has to formally compute interaction between *all* the electrons of *all* the cells, by avoiding to take self-interaction into account, and then divide by the (infinite) number of cells. The ion-ion interaction energy, produced by repulsion of unit point charges, is conveniently computed by adding and subtracting the interaction between Gaussianly shaped charge distributions:

$$\rho_{\rm Gauss}(\mathbf{r}) = \left(\frac{\alpha}{\pi}\right)^{\frac{3}{2}} \sum_I Z_I e^{-\alpha(\mathbf{r}-\mathbf{R}_I)^2} \,, \tag{C.11}$$

where the α parameter, which determines Gaussian charge radius $R_{\rm Gauss} = 1/\sqrt{\alpha}$, has to be suitably chosen. We note that, in principle, $E^{\tau}_{\hat{V}^{\rm ext}}(\sigma), E^{\tau}_{\hat{V}^{\rm ee}}(\sigma), E_{\rm ion}$ all contain $\mathbf{G} = 0$ terms and, therefore, they are separately divergent, due to the form of the Coulomb potential Fourier transform. Anyway, all these divergent contributions cancel each other as we expect since the entire system is neutral.

As far as auxiliary field dynamics is concerned, by adopting the complex HST introduced in Eqs. (2.27) – (2.30), we can write the classical potential energy $V(\sigma)$, considering real and imaginary parts of Fourier coefficients:

$$\sigma(\mathbf{G}) = \frac{1}{\Omega}\int d\mathbf{r}\sigma(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} \,, \tag{C.12}$$

as independent variables:

$$V(\sigma) = \Delta\tau\Omega \sum_{l=1}^{P} \sum_{G>0} \frac{4\pi}{G^2} |\sigma^d(\mathbf{G})|^2 - \ln \langle \psi_T | \hat{U} | \psi_T \rangle + \text{const.} . \tag{C.13}$$

We regard only the positive G Fourier coefficients as independent variables since we keep our σ fields real and therefore:

$$\sigma(-\mathbf{G}) = \sigma^*(\mathbf{G}) . \tag{C.14}$$

Hence, N_d, the total number of degrees of freedom, is given by:

$$N_d = 2 \cdot P \cdot N_G . \tag{C.15}$$

It is obtained by multiplying the number N_G of positive G vectors, by the number P of imaginary time slices, by a 2 factor for the real and imaginary components of complex Fourier coefficients.

Appendix D

The AFQMC method applied to Hubbard model

The AFQMC method has been used extensively to study electron correlations in the Hubbard model[22–40]. The Hubbard model[90] is described by the Hamiltonian:

$$\hat{H} = -t \sum_{<ij>,\mu} c^{\dagger}_{i\mu} c_{j\mu} + U \sum_{i} \hat{\rho}_{i\uparrow} \hat{\rho}_{i\downarrow} \,, \tag{D.1}$$

where $\sum_{<ij>}$ indicates nearest neighbours sum, the indices run over the N_a lattice sites, $c^{\dagger}_{i\mu}$ ($c_{i\mu}$) are the usual creation (annihilation) operators at site i with spin μ, $\hat{\rho}_{i\mu} = c^{\dagger}_{i\mu} c_{i\mu}$, and $U > 0$. The first term, on the R.H.S. of Eq. (D.1), is the kinetic term which allows hopping of electrons between nearest neighbours sites, while the second one represents the on site Coulomb repulsion.

Despite its apparent simplicity the Hubbard model embodies many features (and difficulties) of a truly interacting system. Without interaction ($U = 0$) one obtains a pure band behaviour, typical of the kinetic term; in the atomic limit ($t = 0$) the particles are localized. Of course the intermediate regime ($t \sim U$) is of particular interest, as, in this range of parameters, the competition between band effects and localization due to correlation is most important. A major development, in understanding the physics of the Hubbard model, has been provided by the use of computer simulation techniques. In particular, the AFQMC method

can be easily applied to this model. In fact we observe that:

$$\sum_i \hat{\rho}_{i\uparrow}\hat{\rho}_{i\downarrow} = -\frac{1}{2}\sum_i \hat{m}_i^2 + \frac{\hat{N}}{2} , \tag{D.2}$$

where $\hat{m}_i = \hat{\rho}_{i\uparrow} - \hat{\rho}_{i\downarrow}$ is the usual local magnetization operator, while $\hat{N}$ is the total electron number operator. Therefore the HST can be readily performed, without introducing suitably modified interactions or imaginary factors (see Section 2.1). In this case, due to the special form of the electron-electron interaction (only electrons with opposite spins, on the same lattice site, can interact), although it is repulsive, it can be expressed as a sum of negative, quadratic terms, plus a constant, irrelevant contribution. Hence (see Eq. (1.34)):

$$\begin{aligned} e^{-\Delta\tau U\sum_i \hat{\rho}_{i\uparrow}\hat{\rho}_{i\downarrow}} &= e^{-\Delta\tau U\frac{\hat{N}}{2}}\, e^{\Delta\tau\frac{U}{2}\sum_i \hat{m}_i^2} \\ &= \text{const.} \int d\sigma e^{-\frac{\Delta\tau}{2}U\sum_i \sigma_i^2}\, e^{\Delta\tau U\sum_i \sigma_i \hat{m}_i} . \end{aligned} \tag{D.3}$$

In this AFQMC approach the σ variables are continuous auxiliary fields which can range from $-\infty$ to $+\infty$. Anyway a useful feature of spin and lattice fermion systems is the possibility of using discrete, rather than continuous, auxiliary fields. In fact, if some operator can only assume a finite set of discrete values, it is always possible to replace the integral over a continuous auxiliary variable, in the HST, by a discrete sum. In our case the local magnetization operator $\hat{m}_i$ can only assume the three values $\{-1, 0, 1\}$, then a single sum over an Ising variable suffices:

$$e^{\Delta\tau\frac{U}{2}\sum_i \hat{m}_i^2} = \frac{1}{2}\sum_{\sigma_i=\pm 1} e^{\sigma_i a \hat{m}_i} , \tag{D.4}$$

where:

$$\tanh^2\left(\frac{a}{2}\right) = \tanh\left(\frac{\Delta\tau U}{4}\right) . \tag{D.5}$$

We observe that, if $\Delta\tau \to 0$, then Eq. (D.5) becomes:

$$a^2 \simeq \Delta\tau U \ , \tag{D.6}$$

therefore:

$$e^{\Delta\tau\frac{U}{2}\sum_i \dot{m}_i^2} \simeq \frac{1}{2}\sum_{\sigma_i=\pm 1} e^{\sigma_i\sqrt{\Delta\tau U}\dot{m}_i} \ . \tag{D.7}$$

This discrete HST was introduced by Hirsch [22,23] and it is particularly used to study the Hubbard model, since Ising variables are more convenient to sample, even though alternative, intermediate choices, with continuous but bounded auxiliary fields [38] can be adopted too.

Numerical simulations of the Hubbard model have been performed, both at finite temperature, within the grand canonical ensemble [22–27], and at zero temperature, for a fixed number of electrons [29–40]. Applying the AFQMC method to the Hubbard model a careful analysis of the fermion sign problem was also performed [35–40], and some possible techniques, for coping with this difficulty, were suggested.

Appendix E

The Positive-Projection method

The Positive-Projection (PP) technique was developed by Fahy and Hamann[37] to avoid the fermion sign problem which, in some cases, can make an accurate evaluation of physical properties extremely difficult (see Section 1.3.4). Essentially it is a *variational* form of the AFQMC method, for fermions, which does not exhibit the exponentially vanishing sign. Let us consider the usual HST:

$$|\psi_0\rangle = \lim_{\beta\to\infty} |\psi_\beta\rangle = \lim_{\beta\to\infty} e^{-\beta\hat{H}} |\psi_T\rangle = \int d\sigma G(\sigma)\hat{U}(\sigma)\,|\psi_T\rangle \ . \tag{E.1}$$

Since $|\psi_T\rangle$ remains a Slater determinant, as it propagates through the σ fields, the Gaussian distribution $G(\sigma)$ of all the σ's, from 0 to β, generates a distribution $f\,(\psi;\beta)$ of determinants $|\psi\rangle$, and:

$$|\psi_\beta\rangle = \int f\,(\psi;\beta)\,|\psi\rangle\,d\psi \ , \tag{E.2}$$

where $\int d\psi$ represents integration over all normalized Slater determinants[40]. Fahy and Hamann have shown[40] that f obeys a diffusion equation with drift and branching:

$$-\frac{\partial f}{\partial \tau} = -\frac{1}{2}D(\psi)f - [\nabla_\psi V_1(\psi)]\,\nabla_\psi f + V_2(\psi)f \ , \tag{E.3}$$

where the second-order-derivative diffusion operator D, and the drift and branching potentials, V_1 and V_2, can be explicitly written in terms of the operators of the

initial many-body Hamiltonian. Eq. (E.3) is to be solved with the initial condition $f(\psi;0) = \delta(\psi - \psi_T)$ and its solution can, in principle, be expanded in terms of eigenfunctions f^i and eigenvalues ε_i of its R.H.S., as:

$$f(\psi;\beta) = \sum_i c_i e^{-\varepsilon_i \beta} f^i(\psi)\,. \tag{E.4}$$

Given the ordinary diffusion form of Eq. (E.3), the lowest eigenvalue must belong to an even-parity eigenfunction f^+, which therefore dominates the sum in Eq. (E.4) at large β. However, even-parity terms in f cancel in the integral of Eq. (E.2). These *integrals* are dominated by the lowest odd-parity eigenfunction f^-, while their *integrands* are dominated by f^+. Therefore statistical evaluation of the integrals requires sufficient accuracy to extract the exponentially small physical contribution.

The PP approximation, which was motivated by the previous analysis, consists of supplementing Eq. (E.3) with the boundary condition $f(0) = 0$, which is equivalent to setting $V_2(\psi) = \infty$, for $\psi < 0$. Hence the eigenfunction with the lowest eigenvalue becomes f^-. Let us denote as $f^*(\psi;\beta)$ the solution of Eq. (E.3) subject to the usual initial condition and the new PP nodal condition. Then, using f^*, the average sign (see Section 1.3.4) of the overlap term $\langle\psi'|\psi\rangle$, which is present in the estimator computation, need not be positive, however it is bounded from below[37] and usually stays near 1. Substituting f^* for f has a variational character and gives an upper bound for the total energy. As with conventional variational calculations, the approximate placement (in general one does not know the exact position of the f^- distribution node) of the f^* node obviates the possibility of converging to a truly exact solution. Unconstrained simulation is always to be preferred where the rate of sign decay permits an acceptable combination of large β and small statistical errors. The PP method has a *formal* resemblance

to the fixed-node approximation [18] in GFMC and DMC approaches (see Section 1.2), but the context and physical content are quite different [37,40]. In particular the PP method removes f^+ without imposing any direct constraint on the nodal structure of the many-body wave function, which may adjust with nearly complete freedom according to the Hamiltonian.

A practical way to define the PP constraint is to set the nodal surface equal to the hyperplane perpendicular to a "constraint" wave function $|\psi_C\rangle$, which could be one or a small sum of Slater determinants. Heuristically, we expect that, the better $|\psi_C\rangle$ approximates the ground state, the closer the hyperplane will come to the exact nodal surface of f^-. If one imposes the P ($P = \beta/\Delta\tau$) conditions:

$$\langle\psi_C|\,\hat{U}(\sigma_l)\hat{U}(\sigma_{l-1})\cdots\hat{U}(\sigma_2)\hat{U}(\sigma_1)\,|\psi_T\rangle > 0\,, \quad l = 1,\ldots,P/2\,, \tag{E.5}$$

$$\langle\psi_T|\,\hat{U}(\sigma_P)\hat{U}(\sigma_{P-1})\cdots\hat{U}(\sigma_{l+1})\hat{U}(\sigma_l)\,|\psi_C\rangle > 0\,, \quad l = P/2+1,\ldots,P\,, \tag{E.6}$$

on the stochastic process used to sample the fields σ, one achieves the equivalent of setting V_2 to $+\infty$ for all ψ on the wrong side of the constraint surface. In practice [37] the constraining state $|\psi_C\rangle$ can be chosen as a single determinant, the same used as the trial state $|\psi_T\rangle$. The PP approach was tested [37] considering several two-dimensional Hubbard models, with periodic boundary conditions. It is known [35] that the sign problem is worst when the Fermi level falls in a degenerate set of states and such cases were deliberately chosen [37]: 2×2 with 3 electrons, 3×3 with 8, and 4×4 with 14. The results were quite encouraging [37]: while the average value of the sign was exponentially decaying, as a function of β, for conventional simulations, it was essentially β independent with PP. The variational improvement of the total energy, with an improved PP constraint, was also verified [37].

In conclusion, even though there is no guarantee that PP will give superior

results to a conventional variational calculation with a cleverly chosen wave function, however it should be of considerable advantage in extending applicability of the AFQMC method.

Appendix F

The "Exact" approach

In order to test the accuracy of our AFQMC simulation results, obtained using a fixed set of parameters (the number $N_{\rm pw}$ of expansion plane waves, the single Trotter time slice $\Delta\tau$, the total imaginary propagation time β), an algorithm which can produce exact data, by taking into account the same parameter values, is required. For the simple H_2 molecule, if $N_{\rm pw}$ is not too large, such a scheme can be easily implemented. In fact the basic step is to perform the short time propagation:

$$|\psi'\rangle = e^{-\Delta\tau\hat{H}}\,|\psi\rangle \simeq \left(e^{-\Delta\tau\frac{\hat{T}}{2}}e^{-\Delta\tau\hat{V}_{\rm tot}}e^{-\Delta\tau\frac{\hat{T}}{2}}\right)|\psi\rangle\ , \tag{F.1}$$

where $\hat{H} = \hat{T} + \hat{V}_{\rm tot}$ is the usual H_2 Hamiltonian $\left(\hat{V}_{\rm tot} = \hat{V}^{\rm ext} + \hat{V}\right)$ and $|\psi\rangle$ is a generic two-body wave function, we can expand in plane waves:

$$|\psi\rangle = \sum_{\mathbf{G}_m\mathbf{G}_n} C(\mathbf{G}_m,\mathbf{G}_n)\,|\mathbf{G}_m\mathbf{G}_n\rangle\ . \tag{F.2}$$

Now the short time kinetic term propagation is trivial since the kinetic operator is diagonal in Fourier space:

$$e^{-\Delta\tau\frac{\hat{T}}{2}}\,|\psi\rangle = \sum_{\mathbf{G}_m\mathbf{G}_n} e^{-\frac{\Delta\tau}{4}\left(\mathrm{G}_m^2+\mathrm{G}_n^2\right)}C(\mathbf{G}_m,\mathbf{G}_n)\,|\mathbf{G}_m\mathbf{G}_n\rangle\ . \tag{F.3}$$

As far as the potential term propagation is concerned, it is convenient to consider the following expansion:

$$e^{-\Delta\tau\hat{V}_{\rm tot}} = \hat{1} - \Delta\tau\hat{V}_{\rm tot} + \frac{1}{2!}\left(\Delta\tau\hat{V}_{\rm tot}\right)^2 - \frac{1}{3!}\left(\Delta\tau\hat{V}_{\rm tot}\right)^3 + \cdots\ . \tag{F.4}$$

In this way we reduce the problem of applying the exponential operator $e^{-\Delta\tau\hat{V}_{\rm tot}}$ to repeated applications of $\Delta\tau\hat{V}_{\rm tot}$. Then it is easily to verify that:

$$\Delta\tau\hat{V}_{\rm tot}\,|\psi\rangle = \Delta\tau \sum_{\mathbf{G}_m\mathbf{G}_n\mathbf{G}_l} \Big\{ V^{\rm ext}(\mathbf{G}_l)\,[C\,(\mathbf{G}_m-\mathbf{G}_l,\mathbf{G}_n)+C\,(\mathbf{G}_m,\mathbf{G}_n-\mathbf{G}_l)] \\ +V(\mathbf{G}_l)\,C\,(\mathbf{G}_m-\mathbf{G}_l,\mathbf{G}_n+\mathbf{G}_l)\Big\}\,|\mathbf{G}_m\mathbf{G}_n\rangle\ , \qquad \text{(F.5)}$$

where $V^{\rm ext}(\mathbf{G}_l)$ and $V(\mathbf{G}_l)$ are the Fourier coefficients of the external (electron-ion) and electron-electron interaction potentials, respectively. In this way the basic (F.1) propagation can be practically carried out. In actual calculation one has to increase the number of expansion terms, in (F.4), which are taken into account, until the stability of final results has reached the desired level.

By considering Eq. (F.5) we observe that the computation cost of this algorithm grows as $N_{\rm pw}^3$, therefore it increases quite rapidly with the size of our plane wave basis set.

Bibliography

[1] See, for instance, K. C. Hass, Solid State Phys. **42**, 213 (1989).

[2] B. H. Brandow, Adv. Phys. **26**, 651 (1977).

[3] P. A. Lee, T. M. Rice, J. W. Serene, L. J. Sham and J. W. Wilkins, Comments Cond. Mat. Phys. **12**, 99 (1986).

[4] I. Shavitt, *Methods of Electronic Structure Theory*, edited by H. F. Schaefer (Univ. of California, Berkeley, Plenum Press, 1977), Chap. 6, p. 189.

[5] R. J. Bartlett, Annu. Rev. Phys. Chem. **32**, 359 (1981).

[6] See, for instance, J. C. Slater, *Quantum Theory of Matter* (R. E. Krieger Publishing Company, 1977), p. 326.

[7] R. Jastrow, Phys. Rev. **98**, 1479 (1955).

[8] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, 1971).

[9] For a review see, e.g., *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N. H. March (New York-Plenum, 1983).

[10] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964).

[11] W. Kohn and L. J. Sham, Phys. Rev. A **140**, 1133 (1965).

[12] R. M. Dreizler and E. K. V. Gross, *Density Functional Theory* (Springer Verlag, 1990).

[13] M. H. Kalos, D. Levesque and L. Verlet, Phys. Rev. A **9**, 2178 (1974).

[14] J.B. Anderson, J. Chem. Phys. **63**, 1499 (1975).

[15] D. M. Ceperley and M. H. Kalos, *Monte Carlo Methods in Statistical Physics*, edited by K. Binder (Springer Verlag, New York, 1979).

[16] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[17] D. M. Ceperley, *Recent Progress in Many Body Theories*, edited by J. G. Zabolitzky, M. De Llano, M. Forbes and J. W. Clark (Springer Verlag, Berlin, 1981), p. 262.

[18] P. J. Reynolds, D. M. Ceperley, B. J. Alder, W. A. Lester Jr., J. Chem. Phys. **77**, 5593 (1982).

[19] M. H. Kalos, *Monte Carlo Method in Quantum Physics* (NATO, ASI Series, D. Reidel Publ. Co., Dordrecht, 1984), p. 19.

[20] D. M. Ceperley and B. J. Alder, J. Chem. Phys. **81**, 5833 (1984).

[21] V. Mohan and J. B. Anderson, Chem. Phys. Lett. **156**, 520 (1989).

[22] J. E. Hirsch, Phys. Rev. B **28**, 4059 (1983).

[23] J. E. Hirsch, Phys. Rev. B **29**, 4159 (1984).

[24] J. E. Hirsch, Phys. Rev. B **31**, 4403 (1985).

[25] S. R. White, R. L. Sugar, and R. T. Scalettar, Phys. Rev. B **38**, 11665 (1988).

[26] S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, R. T. Scalettar, Phys. Rev. B **40**, 506 (1989).

[27] J. E. Hirsch and S. Tang, Phys. Rev. Lett. **62**, 591 (1989).

[28] G. Sugiyama and S. E. Koonin, Ann. Phys. **168**, 1 (1986).

[29] S. Sorella, E. Tosatti, S. Baroni, R. Car and M. Parrinello, Int. J. Mod. Phys. B **1**, 993 (1988).

[30] S. Sorella, A. Parola, M. Parrinello and E. Tosatti, Int. J. Mod. Phys. B **3**, 1875 (1989).

[31] S. Sorella, S. Baroni, R. Car and M. Parrinello, Europhys. Lett. **8**, 663 (1989).

[32] S. Sorella, Ph.D. thesis, SISSA-Trieste, 1989 (unpublished).

[33] A. Parola, S. Sorella, S. Baroni, R. Car and E. Tosatti, Physica C **162-164**, 771 (1989).

[34] S. Sorella, A. Parola, M. Parrinello and E. Tosatti, Europhys. Lett. **12**, 721 (1990).

[35] E. Y. Loh, J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino and R. L. Sugar, Phys. Rev. B **41**, 9301 (1990).

[36] D. R. Hamann and S. B. Fahy, Phys. Rev. B **41**, 11352 (1990).

[37] S. B. Fahy and D. R. Hamann, Phys. Rev. Lett. **65**, 3437 (1990).

[38] S. Sorella, Int. J. Mod. Phys. B **5**, 937 (1991).

[39] S. Sorella, A. Parola, and E. Tosatti, Int. J. Mod. Phys. B **5**, 143 (1991).

[40] S. Fahy and D. R. Hamann, Phys. Rev. B **43**, 765 (1991).

[41] R. L. Stratonovich, Dokl. Akad. Nauk. SSSR **1115**, 1097 (1957); Soviet Phys. Doklady **2**, 416 (1957).

[42] J. Hubbard, Phys. Rev. Lett. **3**, 77 (1959).

[43] R. B. Dingle, Phil. Mag. **40**, 573 (1949).

[44] N. F. Mott, Phil. Mag. **40**, 61 (1949).

[45] A. Bijl, Phisica **7**, 869 (1940).

[46] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. M. Teller, E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[47] See, for instance, J. W. Negele and H. Orland, *Quantum Many-Particle Systems* (Addison-Wesley Publishing Company, 1988), Chap. 8.

[48] W. L. McMillan, Phys. Rev. A **138**, 442 (1965).

[49] J. W. Moskowitz, K. E. Schmidt, M. A. Lee and M. H. Kalos, J. Chem. Phys. **76**, 1064 (1982).

[50] C. J. Umrigar, K. G. Wilson and J. W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).

[51] N. Metropolis and S. Ulam, J. Am. Stat. Assoc. **44**, 247 (1949).

[52] M. H. Kalos, Phys. Rev. **128**, 1791 (1962).

[53] F. Mentch and J. B. Anderson, J. Chem. Phys. **63**, 1499 (1975).

[54] J. W. Moskowitz, K. E. Schmidt, M. A. Lee and M. H. Kalos, J. Chem. Phys. **77**, 349 (1982).

[55] P. A. Whitlock, D. M. Ceperley, G. V. Chester, M. H. Kalos, Phys. Rev. B **19**, 5598 (1979).

[56] D. M. Arnow, M. H. Kalos, M. A. Lee, K. E. Schmidt, J. Chem. Phys. **77**, 5562 (1982).

[57] J. B. Anderson, C. A. Traynor, B. M. Boghosian, J. Chem. Phys. **95**, 7418 (1991).

[58] S. E. Koonin, G. Sugiyama and H. Friedrich, *Proceedings of the International Symposium Bad Honnef*, edited by K. Goeke and P. G. Greinhard (Springer Verlag, Berlin, 1982), p. 214.

[59] H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959); M. Suzuki, Commun. Math. Phys. **51**, 183 (1976).

[60] M. M. Hurley and P. A. Christiansen, J. Chem. Phys. **86**, 1069 (1987); B. L. Hammond, P. J. Reynolds and W. A. Lester Jr., *ibid.* **87**, 1130 (1987).

[61] G. B. Bachelet, D. M. Ceperley and M. G. B. Chiocchetti, Phys. Rev. Lett. **62**, 1631 (1988).

[62] B. L. Hammond, P. J. Reynolds and W. A. Lester Jr., Phys. Rev. Lett. **61**,

2312 (1988).

[63] L. Mitáš, to appear in: *Computer Simulation in Condensed Matter Physics* IV, edited by D. P. Landau, K. K. Mon and H. B. Schüttler (Springer, 1992).

[64] S. B. Fahy and D. R. Hamann (unpublished).

[65] G. Parisi and W. Yougshi, Scientia Sinica **24**, 483 (1981).

[66] S. Duane, Nucl. Phys. B **257**, 652 (1985); S. Duane and J. B. Kogut, Phys. Rev. Lett. **55**, 2774 (1985).

[67] P. J. Rossky, J. Doll and H. L. Friedman, J. Chem. Phys. **69**, 4630 (1978).

[68] S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth, Phys. Lett. B **195**, 216 (1987).

[69] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1987).

[70] R. T. Scalettar, D. J. Scalapino and R. L. Sugar Phys. Rev. B **34**, 7911 (1986).

[71] R. W. Hockney and J. W. Eastwood, *Computer Simulation using Particles* (New York, 1981).

[72] M. D. Fest, J. A. Fleck and A. Steiger, J. Comp. Phys. **47**, 412 (1982).

[73] See, for instance, H. R. Schwarz, H. Rutishauser and E. Stiefel, *Numerical Analysis of Symmetric Matrices* (Prentice Hall, Englewood Cliffs).

[74] G. Herzberg and L. L. Howe, Can. J. Phys. **37**, 636 (1959).

[75] R. Moszynski and K. Szalewicz, J. Phys. B **20**, 4347 (1987).

[76] H. M. James and A. S. Coolidge, J. Chem. Phys. **1**, 825 (1933).

[77] W. Kolos, C. C. J. Roothaan, Rev. Mod. Phys. **32**, 219 (1960).

[78] W. Kolos, L. Wolniewicz, Rev. Mod. Phys. **35**, 473 (1963); J. Chem. Phys. **43**, 2429 (1965).

[79] W. Kolos, L. Wolniewicz J. Chem. Phys. **41**, 3663 (1964).

[80] G. Fischer, *Vibronic Coupling* (Academic Press Inc. Ltd., London, 1984).

[81] L. Wolniewicz, J. Chem. Phys. **78**, 6173 (1983).

[82] D. M. Bishop and L. M. Cheung, Phys. Rev. A **18**, 1846 (1978).

[83] C. A. Traynor, J. B. Anderson, B. M. Boghosian, J. Chem. Phys. **94**, 3657 (1991).

[84] E. McCormack, E. E. Eyler, Bull. Am. Phys. Soc. **32**, 1279 (1987).

[85] J. Sanders and K. E. Banyard, J. Chem. Phys. **96**, 4536 (1992).

[86] P. Siegbahn and B. Liu, J. Chem. Phys. **68**, 2457 (1978).

[87] B. Liu, J. Chem. Phys. **80**, 581 (1984).

[88] F. Mentch and J. B. Anderson, J. Chem. Phys. **80**, 2675 (1984).

[89] R. N. Barnett, P. J. Reynolds and W. A. Lester, J. Chem. Phys. **82**, 2700 (1985).

[90] J. Hubbard, Proc. R. Soc. London A **276**, 238 (1963); **277**, 237 (1964); **281**, 401 (1964).

Acknowledgements

First of all I would like to express a deep gratitude to my parents for their continuous and precious encouragement.

I thank Prof. Stefano Baroni and Prof. Roberto Car who, not only introduced me to the research field of this thesis, but also suggested the smartest ideas used throughout my work. In particular I am grateful to Prof. Roberto Car for his support during my repeated visits to IRRMA, in Lausanne.

I wish to express my thanks to Dr. Sandro Sorella: part of my thesis is based on his excellent research about Auxiliary Field techniques and I benefited a lot from fruitful discussions with him.

I am extremely indebted to Pasquale Pavone for his invaluable and essential help to a better use of TEX facilities.

Finally I would like to thank Andrea Ferrante and Giorgio Mazzeo, who heroically spent many hours per day together with me, working in the same office, and all my colleagues at SISSA for their friendship and sympathy.