

**Electronic Structure Quantum Monte Carlo****Michal Bajdich, Lubos Mitas<sup>1</sup>***Department of Physics, North Carolina State University, Raleigh, 27695*

Quantum Monte Carlo (QMC) is an advanced simulation methodology for studies of many-body quantum systems. The QMC approaches combine analytical insights with stochastic computational techniques for efficient solution of several classes of important many-body problems such as the stationary Schrödinger equation. QMC methods of various flavors have been applied to a great variety of systems spanning continuous and lattice quantum models, molecular and condensed systems, BEC-BCS ultracold condensates, nuclei, etc. In this review, we focus on the electronic structure QMC, i.e., methods relevant for systems described by the electron-ion Hamiltonians. Some of the key QMC achievements include direct treatment of electron correlation, accuracy in predicting energy differences and favorable scaling in the system size. Calculations of atoms, molecules, clusters and solids have demonstrated QMC applicability to real systems with hundreds of electrons while providing 90-95% of the correlation energy and energy differences typically within a few percent of experiments. Advances in accuracy beyond these limits are hampered by the so-called fixed-node approximation which is used to circumvent the notorious fermion sign problem. Many-body nodes of fermion states and their properties have therefore become one of the important topics for further progress in predictive power and efficiency of QMC calculations. Some of our recent results on the wave function nodes and related nodal domain topologies will be briefly reviewed. This includes analysis of few-electron systems and descriptions of exact and approximate nodes using transformations and projections of the highly-dimensional nodal hypersurfaces into the 3D space. Studies of fermion nodes offer new insights into topological properties of eigenstates such as explicit demonstrations that generic fermionic ground states exhibit the minimal number of two nodal domains. Recently proposed trial wave functions based on pfaffians with pairing orbitals are presented and their nodal properties are tested in calculations of first row atoms and molecules. Finally, backflow “dressed” coordinates are introduced as another possibility for capturing correlation effects and for decreasing the fixed-node bias.

PACS: 02.70.Ss, 05.30.Fk, 03.65.Ge, 71.15.Nc, 71.10.-w, 71.15.-m, 31.10.+z, 31.25.Qm, 31.25.-v

---

<sup>1</sup>E-mail address: lmitas@ncsu.edu

## Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	Many-Body Schrödinger Equation . . . . .	7
1.2	Hartree–Fock Theory . . . . .	8
1.3	Post Hartree–Fock Methods . . . . .	9
1.4	Density Functional Theory . . . . .	10
<b>2</b>	<b>Quantum Monte Carlo Methods</b>	<b>13</b>
2.1	Introduction . . . . .	13
2.1.1	Metropolis Sampling . . . . .	13
2.2	Variational Monte Carlo . . . . .	15
2.2.1	Correlated Sampling . . . . .	15
2.3	Diffusion Monte Carlo . . . . .	16
2.3.1	Imaginary Time Schrödinger Equation . . . . .	16
2.3.2	DMC Importance Sampling . . . . .	17
2.3.3	Fixed-Node Approximation . . . . .	19
2.3.4	Beyond The Fixed-Node Approximation . . . . .	23
2.4	Trial Variational Wave Functions . . . . .	25
2.4.1	Basic Properties . . . . .	25
2.4.2	Trial Wave Functions Forms . . . . .	26
2.4.3	Basis Functions for the Jastrow Correlation Factor . . . . .	29
2.5	Pseudopotentials . . . . .	30
2.6	Optimization of Variational Wave Functions . . . . .	33
2.6.1	Variance Minimization . . . . .	33
2.6.2	Energy Minimization . . . . .	34
<b>3</b>	<b>Nodal Properties of Fermionic Wave Functions</b>	<b>36</b>
3.1	Introduction . . . . .	36
3.2	General Properties . . . . .	37
3.3	Exact Nodal Surfaces . . . . .	42
3.3.1	Two-Electron Atomic $^3P(p^2)$ , $^3S(1s2s)$ and Molecular $^3\Sigma_g(\pi^2)$ Triplet States . . . . .	42
3.3.2	Three-Electron Quartet $^4S(2p^3)$ State . . . . .	43
3.4	Approximate Hartree–Fock Nodes . . . . .	45
3.4.1	Hartree–Fock Node of $^6S(d^5)$ State . . . . .	45
3.4.2	Hartree–Fock Nodes of the $^5S(sp^3)$ Ion . . . . .	48
3.4.3	Hartree–Fock Nodes of Spin-Polarized $p^3d^5$ and $sp^3d^5$ Shells with $S$ Symmetry . . . . .	48
3.5	Changes in Nodal Topologies From Interactions and Spin Correlations . . . . .	50
3.5.1	HF and CI Nodes for $N_2$ Dimer State $^1\Sigma_g^+$ . . . . .	53
3.6	Conclusions . . . . .	54

<b>4</b>	<b>Pfaffian Pairing Wave Functions</b>	<b>56</b>
4.1	Introduction . . . . .	56
4.2	Algebra of Pfaffians . . . . .	56
4.2.1	Definitions . . . . .	56
4.2.2	Calculation of a Pfaffian . . . . .	57
4.2.3	Gradient and Hessian of Pfaffian . . . . .	59
4.3	Pairing Wave Functions . . . . .	60
4.3.1	Singlet Pairing Wave Function . . . . .	60
4.3.2	Triplet Pairing Wave Function . . . . .	61
4.3.3	Generalized Pairing Wave Function . . . . .	62
4.4	Applications of Pairing Wave Functions in QMC . . . . .	63
4.4.1	Single Pfaffian Calculations . . . . .	64
4.4.2	Multi-Pfaffian Calculations . . . . .	65
4.4.3	Nodal Properties . . . . .	66
4.5	Conclusions . . . . .	68
<b>5</b>	<b>Backflow Correlations in Slater and Pfaffian Wave Functions</b>	<b>69</b>
5.1	Introduction . . . . .	69
5.2	Backflow Wave Function Form . . . . .	70
5.3	Applications of Backflow Correlated Wave Functions . . . . .	71
5.3.1	Homogeneous electron gas . . . . .	71
5.3.2	Carbon atom and dimer . . . . .	71
5.4	Conclusions . . . . .	72
<b>6</b>	<b>Example of Large Scale Application: FeO Solid</b>	<b>75</b>
<b>7</b>	<b>QWalk: A Quantum Monte Carlo Program for Electronic Structure</b>	<b>78</b>
7.1	Organization and Implementation . . . . .	78
7.2	Optimization Methods in QWalk . . . . .	79
7.3	Conclusion . . . . .	80
<b>8</b>	<b>Summary</b>	<b>83</b>

## 1 Introduction

Properties and behavior of quantum many-body systems are determined by the laws of quantum physics which have been known since the 1930s. The time-dependent and stationary Schrödinger equations describe molecules, condensed matter systems, surfaces and interfaces, but also artificial structures such as quantum dots, nanowires, ultracold condensates and so forth. However, the task of solving the Schrödinger equation for systems of electrons and ions and predicting the key quantities of interest such as cohesion and binding energies, optical gaps, chemical reaction barriers, types of magnetic order or presence of quantum condensates seems to be beyond formidable. Indeed, Paul Dirac recognized this already in 1929: “The general theory of quantum mechanics is now almost complete. The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.” Let us therefore try to understand the key obstacles. The first complication comes from the fact that a typical system of interest involves large number of particles. Even moderate size molecules or periodic supercells of solids contain from tens to thousands or more quantum particles. Second, the particles interact and for interesting problems the interactions are strong and influence the systems in profound ways. Third, the solutions have to conform to the required quantum symmetries such as fermionic antisymmetry and others. This is a fundamental departure from classical systems and makes the problem particularly challenging. Fourth, the required accuracy for sensible predictions, which are comparable with experiments, is very high.

Past studies of quantum systems were often based on ideas of reducing the interactions into various types of effective mean-fields. These approaches gradually developed into a high level of sophistication and despite limitations, they are still very useful and important on their own or as starting points for building new theories. At the same time, advent of computer technology has offered us a new window of opportunity for studies of quantum (and many other) problems. It spawned a “third way” of doing science which is based on simulations, in contrast to analytical approaches and experiments. In a broad sense, by simulations we mean computational models of reality based on fundamental physical laws. Such models have value when they enable to make predictions or to provide new information which is otherwise impossible or too costly to obtain otherwise. In this respect, QMC methods represent an illustration and an example of what is the potential of such methodologies.

Some of the ideas used in QMC methods go back to the times before the invention of electronic computers. Already in 1930s Fermi noticed similarities between the imaginary time Schrödinger equation and stochastic processes in statistical mechanics. Apparently, scientists at the Los Alamos National Laboratory attempted to calculate the hydrogen molecule by a simple version of QMC in the early 1950s, around the same time when a seminal work on the first Monte Carlo study of classical systems was published by Metropolis and coworkers. In the late 1950s Kalos initiated QMC simulations for a few-particle systems and laid down the statistical and mathematical background of the Green’s function Monte Carlo method. Pioneering QMC calculations of the correlation energy of homogeneous electron gas by Ceperley and Alder in 1980 started a new era of stochastic methods applied to electronic structure problems. QMC methods were then further developed for quantum lattice models, quantum liquids, nuclear and other systems with contributions from many scientists.

Currently, the term “quantum Monte Carlo” covers several related stochastic methodologies

adapted to determine ground or excited states or finite-temperature equilibrium properties for a variety of quantum systems. Note that the word “quantum” is important since QMC approaches differ in many ways from the Monte Carlo methods for classical systems. In this overview we present QMC methods which enable to solve the stationary Schrödinger equation for interacting Hamiltonians in continuous space. QMC reaches the high accuracy by a direct treatment of many-body effects and by employing stochastic computational techniques which enable to capture many-body effects well beyond the scope of mean-field or analytical methods. Conceptually very straightforward is the variational Monte Carlo (VMC) which involves construction of *trial (variational) wave functions* and evaluations of corresponding expectation values by stochastic integration. More advanced approaches are based on projection techniques such as the diffusion Monte Carlo (DMC) which projects out the ground state from a given trial wave function. The method is formally exact, however, for fermions it suffers from a well-known fermion sign problem. The reason is that in using statistical methodology we tacitly assume that the sampled distributions are nonnegative. Unfortunately, for fermions this is not true and in order to circumvent this obstacle many practical implementations of DMC employ the so-called fixed-node approximation. This introduces the *fixed-node error* which is one of the focal points of this overview.

Elimination or, at least, alleviation of the fixed-node errors is therefore of interest from both fundamental and practical point of views—since this is essentially the *only* approximation which is not deeply understood and also because its full control requires inordinate computational resources. At first, the nodal problem looks intractable since many-body wave function nodes are complicated high-dimensional hypersurfaces determined by the full solution of the Schrödinger equation. However, study of nodal structures has revealed a number of important properties of fermionic states such as nodal domain topologies and their relationships to ideas in geometry and spectral theory of operators. Therefore, some of the latest research developments on fermionic wave function nodes will be presented and explained.

Along the same line of the research, we outline current developments in construction of efficient and accurate wave functions which explore ideas beyond the traditional determinants of single-particle orbitals. Recently proposed *Pfaffian* wave functions built from pairing orbitals, which provide higher accuracy and also correct nodal topologies, will be explained [1, 2]. In the subsequent part, a fermion coordinate transformation of the *backflow type* which was recently tested for inhomogeneous systems is introduced. We generalize its application to Pfaffian wave functions and illustrate the results on some of the first benchmarks of these ideas.

Fixed-node QMC simulations are computationally more demanding by two to three orders of magnitude when compared to the mainstream electronic structure methods which rely on mean-fields treatments of electron-electron interactions. However, traditional mean-field theories often lack the needed accuracy and also require frequent re-tuning, sometimes just to qualitatively reconcile the results with experiments. Although at present, building of accurate many-body wave functions is far from being automatic and demands very significant effort as well, the results often provide high accuracy and become benchmarks for other methods. For cohesive and binding energies, band gaps, and other energy differences the agreement with experiments is typically within 1-3% [3–5]. The computational cost of QMC increases basically as a third power of the system size, making calculations with hundreds of electrons tractable and large clusters [6] and solids [7] up to 1000 electrons have been already studied. Producing accurate “numbers” has its own merit and the predictive power is particularly valuable for experiments and for practical purposes. However, even more important are new insights into quantum phenomena and

detailed analysis of theoretical ideas. Indeed, QMC is very much in the line of “it from bit” paradigm, alongside, for example, of substantial computational efforts in quantum chromodynamics which not only predict hadron masses but, at the same time, contribute to the validation of the fundamental theory. Similar simulations efforts exist in other areas of physics as well. Just a few decades ago it was almost unthinkable that one would be able to solve Schrödinger equation for hundreds of electrons in an explicit, many-body wave function framework. Today, such calculations are feasible using available computational resources. At the same time, much more remains to be done, of course, to make the methods more insightful, more efficient and their application less laborious. We hope this overview will contribute to the growing interest in this rapidly developing field of research.

The review is organized as follows:

- The remainder of this Section provides brief introduction into several traditional methods of solving the stationary Schrödinger problem for electron-ion systems.
- The next section follows with the description of the methodology behind the VMC and DMC methods.
- Results of studies on exact and approximate nodes of fermionic wave functions are summarized in the Sec. 3.
- Section 4 discusses generalized pairing wave functions and their properties.
- Overview of preliminary results on the backflow wave functions is given in the section 5.
- Section 6 shows the example of large scale QMC application: FeO solid.
- A QMC package Qwalk is introduced in Sec. 7.
- The last section concludes with comments and ideas on current developments and outlooks.

### 1.1 Many-Body Schrödinger Equation

Let us consider a system of quantum particles such as electrons and ions with Coulomb interactions. Since the masses of nuclei and electrons differ by three orders of magnitude or more, let us simplify the problem by the *Born-Oppenheimer approximation* which allows to separate the electrons from the slowly moving system of ions. The electronic part of non-relativistic Born-Oppenheimer Hamiltonian is given by

$$\mathcal{H} = T + V = T + V_{ext} + V_{ee} = -\frac{1}{2} \sum_i \nabla_i^2 + \sum_{i,I} \frac{-Z_I}{|r_i - R_I|} + \sum_{j>i} \frac{1}{|r_i - r_j|}, \quad (1)$$

where  $i$  and  $j$  label the electrons and while  $I$  sums over the ions with charges  $Z_I$ . (Throughout the review we employ the atomic units, a.u., with  $\hbar = m_e = e = 1$ .) The basic properties of such systems can be found from eigenstates  $\Psi_n$  of the stationary Schrödinger equation

$$\mathcal{H}\Psi_n = E_n \Psi_n. \quad (2)$$

Colloquially, we call such solutions (exact or approximate) and derived properties as electronic structure.

The history of electronic structure *ab initio* methods (see, for example, Ref. [8]) started soon after the discovery of quantum mechanics by pioneering works of Heitler and London [9], Hartree [10], and Hylleraas [11, 12]. An important step forward was made by Fock [13] who established the simplest antisymmetric wave function and formulated the Hartree–Fock (HF) theory which takes into account the Pauli *exclusion principle* [14]. The theory was further developed by Slater [15] and others and it has become a starting point of many sophisticated approaches not only in condensed matter physics or quantum chemistry but also in other areas such as nuclear

physics. The HF theory replaces the hard problem of many interacting electrons with a system of independent electrons in an effective, self-consistent field (SCF).

For periodic systems the effective free electron theory and the *band theory* of Bloch [16] were the first critical steps towards understanding real crystals. In 1930s, the foundation of the basic classification of solids into metals, semiconductors and insulators was laid down and later Wigner and Seitz [17, 18] performed the first quantitative calculation of electronic states in sodium metal. Building upon homogeneous electron gas model, the density functional theory (DFT) was invented by Hohenberg and Kohn [19] and further developed by Kohn and Sham [20]. Subsequent developments proved to be very successful and DFT has become the mainstream method for many applications which cover not only condensed systems but also molecules, surfaces, even nuclear and other systems. DFT together with HF and post-HF methods are relevant for our discussion of quantum Monte Carlo, which uses the results from these approaches as a reference and also for building wave functions. What follows is their brief overview.

## 1.2 Hartree–Fock Theory

Due to the Pauli exclusion principle, any solution of the Schrödinger equation with Hamiltonian (1) has to be antisymmetric under coordinate exchange of any two electrons as

$$\Psi(\dots, i, \dots, j, \dots) = -\Psi(\dots, j, \dots, i, \dots), \quad (3)$$

where  $i, j$  denote both spatial and spin degrees of freedom. The Hartree–Fock theory [13, 15] is based on the simplest antisymmetric wave function, sometimes called the Slater determinant,

$$\begin{aligned} \Psi(1, 2, \dots, N) &= \frac{1}{\sqrt{N!}} \sum_P (-1)^P \tilde{\varphi}_{i_1}(1) \tilde{\varphi}_{i_2}(2) \dots \tilde{\varphi}_{i_N}(N) \\ &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \tilde{\varphi}_1(1) & \tilde{\varphi}_1(2) & \dots & \tilde{\varphi}_1(N) \\ \tilde{\varphi}_2(1) & \tilde{\varphi}_2(2) & \dots & \tilde{\varphi}_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\varphi}_N(1) & \tilde{\varphi}_N(2) & \dots & \tilde{\varphi}_N(N) \end{vmatrix} \equiv \det[\tilde{\varphi}_1(1), \dots, \tilde{\varphi}_N(N)], \end{aligned} \quad (4)$$

to approximate the state of  $N$ -electron system.  $\{\tilde{\varphi}_i(j)\}$  are one-particle spin-orbitals, each being a product of the spatial  $\varphi_i^\sigma(j)$  and spin  $\alpha_i(\sigma_j)$  orbitals. (Note that  $\varphi_i^\sigma(j)$  is independent of spin  $\sigma$  in closed-shell cases. In open-shell systems, this assumption corresponds to the *spin-restricted Hartree–Fock approximation*). The Hartree–Fock total energy can be found from the following variational condition

$$\langle \Psi | \mathcal{H} | \Psi \rangle = \min; \quad \langle \varphi_j | \varphi_j \rangle = 1, \quad j = 1, \dots, N. \quad (5)$$



If the Hamiltonian is spin-independent or it is diagonal in the spin basis  $\sigma = \{|\uparrow\rangle, |\downarrow\rangle\}$ , the expectation value of Hamiltonian (1) with the wave function (4) is given by

$$\begin{aligned}
 E_{HF} = & \sum_{i,\sigma} \int \varphi_i^{\sigma*}(\mathbf{r}) \left[ -\frac{1}{2} \nabla^2 + V_{ext} \right] \varphi_i^\sigma(\mathbf{r}) d\mathbf{r} \\
 & + \frac{1}{2} \sum_{i,j,\sigma_i,\sigma_j} \iint \frac{\varphi_i^{\sigma_i*}(\mathbf{r}) \varphi_j^{\sigma_j*}(\mathbf{r}') \varphi_i^{\sigma_i}(\mathbf{r}) \varphi_j^{\sigma_j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \\
 & - \frac{1}{2} \sum_{i,j,\sigma} \iint \frac{\varphi_i^{\sigma*}(\mathbf{r}) \varphi_j^{\sigma*}(\mathbf{r}') \varphi_j^\sigma(\mathbf{r}) \varphi_i^\sigma(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \tag{6}
 \end{aligned}$$

The expression for the HF energy contains three different terms. The first one represents one-particle contributions, the second is the Coulomb interaction also called the Hartree term and the last contribution is the *exchange* or Fock term. Note that for  $i = j$  (self-interaction) the last two terms explicitly cancel each other. Variation of the HF total energy leads to one-particle HF equations with the resulting spectrum  $\{\varphi_i^\sigma, \epsilon_i\}$  of occupied and virtual orbitals obtained by solving

$$\left[ -\frac{1}{2} \nabla^2 + V_{ext} + V_{HF}(\{\varphi_j^\sigma\}) \right] \varphi_i^\sigma(\mathbf{r}) = \epsilon_i \varphi_i^\sigma(\mathbf{r}) \tag{7}$$

The effective field operator  $V_{HF}$  contains both Coulomb and exchange terms and due to the self-consistency depends also on the set of occupied orbitals. The resulting operator is both nonlinear and nonlocal as manifested by the fact that  $V_{HF}$ , in general, differs for each state  $i$ . This complicated dependence makes the solution of HF equations rather involved. The solutions are usually expanded in an appropriate basis sets and iterated to self-consistency. Sophisticated and powerful tools such as well-developed and extensive quantum chemical or solid state packages are necessary for such solutions. For solids the HF equations were out of reach for a long time and only in late eighties the first package had emerged which was able to find reasonably accurate and converged solutions.

### 1.3 Post Hartree-Fock Methods

The HF energy is variational, i.e, above the exact eigenvalue, and whatever is left out is referred to as *electronic correlation*. For a given state, the correlation energy is therefore defined as a difference of the HF energy and the exact eigenvalue,

$$E_{corr} = E_{HF} - E. \tag{8}$$

In electron-ion systems  $E_{corr}$  is only a small fraction of the total energy, of the order of  $\approx 1-3\%$ , however, correlations impact almost all quantities in a major way. For example, correlation accounts for a large portion of cohesions, excitations, barrier heights and other important energy differences. Many other effects such as magnetism or superconductivity are driven by correlations so that its accurate description becomes crucial. Revealing the nature of correlation effects and how they influence the behavior of quantum systems is therefore one of the central challenges of modern electronic structure research.

The missing correlation in the HF theory can be accounted for by expanding the wave function into linear combination of Slater determinants with one-, two-, ...,  $N$ -particle excitations. Providing the one-particle set of states is complete and all possible excitations are covered one can formally reach the exact solution. Unfortunately, the convergence of such methods is slow and the computational demands grow exponentially with the system size. Nevertheless, for small systems such methods are effective and provide valuable checks for other approaches. The most frequently used post-HF approaches are configuration interaction (CI), multi-configurational self-consistent field (MCSCF), and coupled cluster (CC) methods. The idea behind CI is to diagonalize the  $N$ -electron Hamiltonian on the space of orthogonal Slater determinants. These determinants are typically constructed as few-particle excitations from the reference determinant (usually the HF solution) using the virtual HF orbitals. The CI method thus provides the most optimal expansion on such determinantal space. The most commonly used is CI with all the single and double excitations (CISD). Besides the slow convergence in one-particle orbitals the method becomes problematic with increasing  $N$  since the recovered correlation is proportional to  $\sqrt{N}$  (instead of  $N$ ), thus leading to the well-known *size-consistency problem* (i.e., a wrong scaling of total energy with  $N$ ).

The MCSCF is a generalization of the expansion in excited determinants and differs from CI by optimization of both of the orbitals and the determinantal coefficients. Such optimizations are quite difficult and the demands often limit the size of the problem which can be studied. The special case of the MCSCF is the complete active space SCF (CASSCF), which is size consistent, at the price of exponential scaling in the number of particles.

The size-consistency problem can be overcome by going beyond linear expansions as it is done in the coupled cluster method [21]. The wave function ansatz is based on exponentiated excitation operators and the corresponding optimized parameters are found by solving a set of coupled nonlinear equations. The computational cost for CCSD (with singles and doubles) scales as  $N^6$  and therefore constitutes a formidable challenge for larger systems. The commonly used CCSD(T) (CC with single, double and perturbative triple excitations) is often used as a standard in accuracy for small systems. For an overview and detailed introduction into the above mentioned quantum chemical methods we refer the reader, for example, to the book by A. Szabo and N. S. Ostlund [22].

## 1.4 Density Functional Theory

Previous methods were examples of the wave function based theories. For the density functional theory (DFT) the primary object is the *one-particle electron density*, i.e., the diagonal part of one-particle density matrix. DFT is formally an exact method based on Hohenberg and Kohn theorem [19] that states that the ground state properties of a many-electron system can be obtained by minimizing the total energy functional

$$E[\rho(\mathbf{r})] = \int V_{ext}(\mathbf{r})\rho(\mathbf{r}) d\mathbf{r} + F[\rho(\mathbf{r})], \quad (9)$$

where the  $F[\rho(\mathbf{r})]$  is some unknown, universal functional of the electronic density  $\rho(\mathbf{r})$ . Assuming such functional would be known, the total energy  $E$  would exhibit a minimum for the exact ground state density  $\rho(\mathbf{r})$ .

Kohn and Sham [20] introduced an *ansatz* for the density in terms of one-electron orbitals of an auxiliary non-interacting system as

$$\rho(\mathbf{r}) = \sum_i^N |\varphi_i(\mathbf{r})|^2. \quad (10)$$

The total energy of electronic system can be then expressed as

$$E[\rho(\mathbf{r})] = -\frac{1}{2} \sum_i \int \varphi_i^*(\mathbf{r}) \nabla^2 \varphi_i(\mathbf{r}) d\mathbf{r} + \int \rho(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{xc}[\rho(\mathbf{r})], \quad (11)$$

where the first two terms represent the familiar energy of a non-interacting system in the external potential, the third term being the Hartree term while the rest is an unknown universal *exchange-correlation functional*. By minimization of the total energy as a functional of normalized occupied orbitals one can derive the Kohn-Sham equations

$$\left[ -\frac{1}{2} \nabla^2 + V_{ext} + V_{DFT}(\rho(\mathbf{r})) \right] \varphi_i^\sigma(\mathbf{r}) = \epsilon_i \varphi_i^\sigma(\mathbf{r}). \quad (12)$$

The structure of the equations appears similar to the HF equations, the key difference is that the effective exchange-correlation operator is nonlinear yet local, since it depends only on the density but not on individual states. This feature of the DFT theory simplifies the calculations considerably and it has enabled to expand the DFT applicability to very large systems. If the functional  $E_{xc}[\rho(\mathbf{r})]$  would be precisely known, we would have one to one mapping between difficult many-electron system and an effective one-particle problem. The fundamental difficulty is that the exact exchange-correlation functional is unknown and therefore we have to rely on approximations.

The simplest is the local density approximation (LDA),

$$E_{xc}[\rho(\mathbf{r})] = \int \epsilon_{xc}^{heg}[\rho(\mathbf{r})] \rho(\mathbf{r}) d\mathbf{r}, \quad (13)$$

where  $\epsilon_{xc}^{heg}$  is the exchange-correlation energy of the homogeneous electron gas of the same density. It is interesting that basically in all practical implementations of LDA, the correlation portion of this energy functional comes from the pioneering QMC calculations of the homogeneous electron gas [23].

It is intuitively suggestive to generalize the exchange-correlation functional to include also the density derivatives, i.e., to express  $\epsilon_{xc}[\rho(\mathbf{r}), \nabla\rho(\mathbf{r})]$  as a function of the local density and its gradient. This was the essential idea behind the generalized gradient approximation (GGA) [24–27], which provide improved results for some quantities and has enabled wide-spread use both in condensed matter physics and quantum chemistry.

There are many flavors and generalizations of the density functional theory, e.g., hybrid-functionals [24, 26, 27], time-dependent DFT (TD-DFT) [28, 29] for calculations of excited states, extensions to non-local density functionals (averaged density approximation (ADA) [30]

and weighted density approximation (WDA) [30]), or orbital-dependent functionals (e.g. self-interaction corrections (SIC) [31] and LDA+U [32]). For some systems the DFT methods are very effective and provide reasonable accuracy at very low cost. Unfortunately, there are many cases where DFT requires further elaboration: introduction of ad hoc parameters such as in LDA+U approach or perturbative corrections such as in the so-called GW method. Even then for many systems such as transition metal compounds, weakly interacting van der Waals systems and other cases the DFT has profound difficulties in reaching the required accuracy.

## 2 Quantum Monte Carlo Methods

### 2.1 Introduction

Quantum Monte Carlo (QMC) is a suite of approaches based on two complementary strategies for solving quantum many-body problems. The first key component is an analysis of many-body effects at the wave function level: wave function analytical structure and impact of potential singularities, efficient functional forms, handling of symmetries and boundary conditions, explicit multi-particle correlations and construction of intermediate effective Hamiltonians which enable to build-up more accurate many-body treatment. The second key ingredient is based on employing stochastic techniques whenever it is impractical or inefficient to use analytical and explicit constructions. This is done on several fronts: evaluations of high-dimensional integrals, solving of quantum equations by mapping onto equivalent stochastic processes, development of algorithms and numerical approaches for high accuracy and efficiency.

Over the last three decades, the impact of QMC has been steadily growing as scientists explored new algorithms and produced results which were otherwise impossible to obtain by traditional methods. Currently, QMC covers a great variety of specialized approaches tuned to particular Hamiltonians and applications. In this review we will briefly describe the variational and diffusion quantum Monte Carlo methods.

In the variational Monte Carlo (VMC) the expectation values are evaluated via stochastic integration over  $3N$ -dimensional space using trial (variational) wave functions constructed to capture the desired many-body phenomena. The variational theorem ensures that the expectation value of the Hamiltonian is, within the statistical error bar, a true upper bound to the exact ground-state energy. The results are determined by the choice and quality of the trial wave function construction and therefore are subject to a *variational bias*. The second described method, the diffusion Monte Carlo, removes most of the variational bias by projecting out the ground state component from any trial function with nonzero overlap within the given symmetry class. The fixed-node approximation enables to avoid the fermion sign problem and makes the method scalable to large sizes.

#### 2.1.1 Metropolis Sampling

Basic Monte Carlo multi-variate integration approaches are based on the following simple idea. Assume that we are interested in finding an expectation value of  $A(\mathbf{R})$  for a known non-negative distribution  $\mathcal{P}(\mathbf{R})$

$$\langle A \rangle = \frac{\int A(\mathbf{R})\mathcal{P}(\mathbf{R})d\mathbf{R}}{\int \mathcal{P}(\mathbf{R})d\mathbf{R}}, \quad (14)$$

where  $\mathbf{R}$  is a position in  $d$ -dimensional space. This is a common expression for evaluation of observables encountered in classical or quantum statistical mechanics. If the dimension of space is large and the integral cannot be reduced to sums and/or products of low-dimensional integrals, its straightforward evaluation scales exponentially in the dimension  $d$ . This is easy to see by considering a simple rectangular integration rule with  $N_g$  grid points in each dimension. The total number of evaluations will be  $(N_g)^d$  and therefore not feasible even for moderate values of  $N_g$  and  $d$ . The key advantage of stochastic methods is that they are immune to the

dimensionality problem under reasonable conditions which are easy to verify (see below). Let us first consider how to generate *samples* of the distribution  $\mathcal{P}(\mathbf{R})$ . This is most conveniently done by the Metropolis algorithm [33] which generates a Markov chain of samples by a stochastic walk in the configuration  $\mathbf{R}$ -space. The stochastic samples are sometimes colloquially referred to as random *walkers*. The algorithm enables us to sample any desired multi-dimensional probability distribution  $\mathcal{P}(\mathbf{R})$  without a prior knowledge of its normalization. Given some transition rule  $P(\mathbf{R} \rightarrow \mathbf{R}')$  from  $\mathbf{R}$  to  $\mathbf{R}'$ , which satisfies *ergodicity* and *detailed balance*,

$$P(\mathbf{R} \rightarrow \mathbf{R}')\mathcal{P}(\mathbf{R}) = P(\mathbf{R}' \rightarrow \mathbf{R})\mathcal{P}(\mathbf{R}'), \quad (15)$$

the desired probability density  $\mathcal{P}(\mathbf{R})$  will converge to an equilibrium state given by

$$\int P(\mathbf{R} \rightarrow \mathbf{R}')\mathcal{P}(\mathbf{R}) d\mathbf{R} = \mathcal{P}(\mathbf{R}'), \quad (16)$$

see, for example, [34]. The transition rule  $P(\mathbf{R} \rightarrow \mathbf{R}')$  can be further written as a product of the sampling distribution  $T(\mathbf{R} \rightarrow \mathbf{R}')$  (typically Gaussian distribution centered around  $\mathbf{R}$ ) and the probability of an acceptance  $A(\mathbf{R} \rightarrow \mathbf{R}')$  of the proposed step

$$P(\mathbf{R} \rightarrow \mathbf{R}') = T(\mathbf{R} \rightarrow \mathbf{R}')A(\mathbf{R} \rightarrow \mathbf{R}'). \quad (17)$$

The conditions of detailed balance are satisfied by choosing the  $A(\mathbf{R} \rightarrow \mathbf{R}')$  to be

$$A(\mathbf{R} \rightarrow \mathbf{R}') = \min \left[ 1, \frac{T(\mathbf{R}' \rightarrow \mathbf{R})\mathcal{P}(\mathbf{R}')}{T(\mathbf{R} \rightarrow \mathbf{R}')\mathcal{P}(\mathbf{R})} \right]. \quad (18)$$

In this manner, the Markov chain initialized at some random state evolves towards an equilibrium. After some initial convergence and equilibration period we can start to collect the statistics. Assuming that we can apply the *central limit theorem*, the usual estimators for the mean, variance and error of the mean are given by

$$\begin{aligned} \langle A \rangle &= \frac{1}{M} \sum_m^M A(\mathbf{R}_m) + \mathcal{O}(1/\sqrt{M}), \\ \langle \sigma_A^2 \rangle &\approx \frac{1}{M-1} \sum_m^M (A(\mathbf{R}_m) - \langle A \rangle)^2, \\ \epsilon_A &\approx \frac{\sigma_A}{\sqrt{M}}, \end{aligned} \quad (19)$$

where  $M$  represents the number of statistically independent sampling points.

The computational demands depend on the desired error  $\epsilon_A$  and on the variance  $\sigma_A^2$ . Seemingly, the estimators and the error bar do not depend on the dimensionality  $d$  of the problem. In general, however, one needs to take into account the impact of increasing  $d$  for the following: i) amount of calculation needed to obtain statistically independent sample of  $\mathcal{P}(\mathbf{R})$  and for evaluation of  $A(\mathbf{R})$ ; and ii) growth of variance  $\sigma_A^2$ . Providing all of these grow as some polynomials of  $d$ , the method guarantees the overall polynomial scaling of the algorithm. In its basic form the expression for computational demands is then

$$T_{comp} = \frac{\sigma_A^2}{\epsilon_A^2} T_{sample}, \quad (20)$$

where  $T_{sample}$  lumps the time needed to obtain statistically independent sample and to evaluate required quantities.

We note that it is often possible to decrease the variance by the importance sampling methods and thus increase the efficiency as explained, for example, in Ref. [34].

## 2.2 Variational Monte Carlo

Let us consider a *trial variational wave function*  $\Psi_T(\mathbf{R})$  which approximates the lowest eigenstate of  $\mathcal{H}$  for a given symmetry. We will elaborate on constructions and properties of  $\Psi_T$  in Sec. 2.4. In variational Monte Carlo (VMC) the expectation values for the given  $\Psi_T(\mathbf{R})$  are evaluated by stochastic sampling. For example, the total energy is given by

$$E_{VMC} = \frac{\int \Psi_T^*(\mathbf{R}) \mathcal{H} \Psi_T(\mathbf{R}) d\mathbf{R}}{\int \Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R}) d\mathbf{R}} = \frac{\int |\Psi_T(\mathbf{R})|^2 E_L(\mathbf{R}) d\mathbf{R}}{\int |\Psi_T(\mathbf{R})|^2 d\mathbf{R}} \geq E_0, \quad (21)$$

where we have introduced the so-called local energy

$$E_L(\mathbf{R}) = \frac{\mathcal{H} \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}. \quad (22)$$

The VMC estimator  $E_{VMC}$  is then given as

$$E_{VMC} = \frac{1}{M} \sum_m^M E_L(\mathbf{R}_m) + \mathcal{O}(1/\sqrt{M}), \quad (23)$$

assuming the set of  $M$  configurations is distributed according to  $\mathcal{P}(\mathbf{R}) = |\Psi_T(\mathbf{R})|^2$ .

### 2.2.1 Correlated Sampling

Correlated sampling exploits the fact that the variance of difference of correlated random variables  $X$  and  $Y$  decreases as their correlation increases. For example, consider variational energy difference for two close wave functions  $\Psi_T^{(1)}$  and  $\Psi_T^{(2)}$ , e.g., for different ionic positions or with small external field when evaluating responses. The energy difference can be written as

$$\begin{aligned} E_1 - E_2 &= \frac{\int |\Psi_T^{(1)}(\mathbf{R})|^2 E_L^{(1)}(\mathbf{R}) d\mathbf{R}}{\int |\Psi_T^{(1)}(\mathbf{R})|^2 d\mathbf{R}} - \frac{\int |\Psi_T^{(2)}(\mathbf{R})|^2 E_L^{(2)}(\mathbf{R}) d\mathbf{R}}{\int |\Psi_T^{(2)}(\mathbf{R})|^2 d\mathbf{R}} \\ &= \int |\Psi_T^{(1)}(\mathbf{R})|^2 \left[ \frac{E_L^{(1)}(\mathbf{R})}{\int |\Psi_T^{(1)}(\mathbf{R})|^2 d\mathbf{R}} - \frac{w(\mathbf{R}) E_L^{(2)}(\mathbf{R})}{\int w(\mathbf{R}) |\Psi_T^{(1)}(\mathbf{R})|^2 d\mathbf{R}} \right] d\mathbf{R}, \end{aligned} \quad (24)$$

where the re-weighting factor is given by  $w(\mathbf{R}) = |\Psi_T^{(2)}(\mathbf{R})|^2 / |\Psi_T^{(1)}(\mathbf{R})|^2$ . Therefore, we can obtain the above energy difference as a sum over weighted local energy differences. The estimations can be carried out using the set of samples drawn, for example, from the distribution  $(\Psi_T^{(1)})^2$ . There are other possibilities such as sampling  $|\Psi_T^{(1)} \Psi_T^{(2)}|$  and modifying the expressions accordingly. By taking the difference, the noise mostly cancels out, since for very similar wave functions the random fluctuations will be very similar and the weights will be close to

unity. Consequently, the variance of the difference can be much smaller than the variance of each individual quantity alone. The actual decrease in variance depends on particular application but gains can be two orders of magnitude or more. The algorithm can be further generalized for more complicated estimators and has been developed also for the diffusion Monte Carlo method presented below [35].

## 2.3 Diffusion Monte Carlo

### 2.3.1 Imaginary Time Schrödinger Equation

Diffusion Monte Carlo method belongs to a class of projection and Green's function QMC approaches. Let us consider the imaginary time Schrödinger equation

$$-\frac{\partial \Psi(\mathbf{R}, \tau)}{\partial \tau} = (\mathcal{H} - E_T) \Psi(\mathbf{R}, \tau), \quad (25)$$

where  $\tau$  is a real variable (imaginary time) and  $E_T$  is an energy offset. The effect of imaginary time is an exponentially fast projection of the lowest state of a given symmetry from any trial function with non-zero overlap. This can be easily seen by expanding  $\Psi(\mathbf{R}, \tau)$  in eigenstates  $\{\Phi_n\}$  of  $\mathcal{H}$  so that for time  $\tau \rightarrow \infty$  one finds

$$\begin{aligned} \lim_{\tau \rightarrow \infty} \Psi(\mathbf{R}, \tau) &= \lim_{\tau \rightarrow \infty} \sum_n \langle \Phi_n | \Psi(0) \rangle e^{-\tau(E_n - E_T)} \Phi_n(\mathbf{R}) \\ &= \lim_{\tau \rightarrow \infty} \langle \Phi_0 | \Psi(0) \rangle e^{-\tau(E_0 - E_T)} \Phi_0(\mathbf{R}). \end{aligned} \quad (26)$$

Although the absolute ground states of the Hamiltonians we consider are bosonic, by keeping the wave function in the antisymmetric sector, one can find the lowest antisymmetric state. The parameter  $E_T$  can be adjusted to  $E_0$  so that the exponential weight of  $\Psi$  becomes asymptotically constant. It is convenient to rewrite the equation (25) into an integral form using the Green's function

$$G(\mathbf{R} \leftarrow \mathbf{R}', \tau) = \langle \mathbf{R} | e^{-\tau(\mathcal{H} - E_T)} | \mathbf{R}' \rangle. \quad (27)$$

Then after propagation by a time interval  $\Delta\tau$  the wave function is given by

$$\Psi(\mathbf{R}, \tau + \Delta\tau) = \int G(\mathbf{R} \leftarrow \mathbf{R}', \Delta\tau) \Psi(\mathbf{R}', \tau) d\mathbf{R}'. \quad (28)$$

and the propagation can be continued by iterations.

Let us first consider the Schrödinger equation with non-interacting particles

$$\frac{\partial \Psi(\mathbf{R}, \tau)}{\partial \tau} = \frac{1}{2} \nabla^2 \Psi(\mathbf{R}, \tau). \quad (29)$$

The equation has a familiar form of diffusion in  $\mathbf{R}$ -dimensional space. Assuming that we have  $N$  particles in a 3D space the corresponding Green's function is a  $3N$ -dimensional Gaussian with variance  $\sigma^2 = \tau$ . The equation describes a well-known process of diffusing particles in



a random noise medium with the diffusion constant equal to  $1/2$ . The key departure from traditional methods of solving such differential equations is in the solution representation. The quantum amplitude  $\Psi(\mathbf{R}, \tau)$  is described by sampling, i.e., as a density of random walkers (delta functions in  $\mathbf{R}$ -space)

$$\Psi(\mathbf{R}, \tau) = \text{dens} \left[ \sum_{m=1}^M \delta(\mathbf{R} - \mathbf{R}_m(\tau)) \right] + \mathcal{O}(1/\sqrt{M}), \quad (30)$$

where by dens we denote the density estimator, which can be carried out by appropriate binning of the walker ensemble. The function is therefore represented only statistically so that the values are known up to the statistical uncertainty term as indicated above. If we substitute the ensemble of delta functions into the integral Eq.(28), we obtain Gaussians of width  $\tau$  centered at the locations  $\{\mathbf{R}_m\}$ . In order to restore the walker representation back we then *sample* the Gaussians, i.e., each walker evolves to a new position as given by

$$\mathbf{R}_m(\tau + \Delta\tau) = \mathbf{R}_m(\tau) + \sqrt{\Delta\tau} \chi, \quad (31)$$

where  $\chi$  is a  $3N$ -dimensional vector of independent random numbers with Gaussian distribution and unit variance. The process is iterated until the long-time limit  $\tau \rightarrow \infty$  is reached.

Considering now more interesting Hamiltonian with external potential and interaction terms represented by  $V$ , one can find the following short-time approximation [36] of the Green's function using the Trotter-Suzuki formula as

$$\begin{aligned} G(\mathbf{R} \leftarrow \mathbf{R}', \tau) &= \langle \mathbf{R} | e^{-\tau(T+V-E_T)} | \mathbf{R}' \rangle \\ &\approx e^{-\tau(V(\mathbf{R})-E_T)/2} \langle \mathbf{R} | e^{-\tau T} | \mathbf{R}' \rangle e^{-\tau(V(\mathbf{R}')-E_T)/2} \\ &\approx (2\pi\tau)^{-3N/2} \exp \left[ -\frac{(\mathbf{R} - \mathbf{R}')^2}{2\tau} \right] \exp \left[ -\tau \left( \frac{V(\mathbf{R}) + V(\mathbf{R}')}{2} - E_T \right) \right]. \end{aligned} \quad (32)$$

The approximation is of the order of  $\mathcal{O}(\tau^3)$ . For sufficiently small time step we can propagate the walkers even for Hamiltonians with complicated local interactions. The last term shows that the propagation involves two processes: Gaussian diffusion caused by the kinetic energy operator, as in the case of free particles, and reweighting, which originates from the presence of the potential. Note that the weights will vary very strongly whenever the potential varies. This is true, in particular, close to any potential singularities and makes such process very noisy and inefficient. In order to make the sampling more efficient we need to smooth out the terms entering the weights and to modify the the algorithm.

### 2.3.2 DMC Importance Sampling

Let us introduce *importance sampling* [36–38] Instead of sampling the wave function  $\Psi(\mathbf{R}, \tau)$  we will sample a mixed distribution  $f(\mathbf{R}, \tau) = \Psi_T(\mathbf{R})\Psi(\mathbf{R}, \tau)$ , where  $\Psi_T(\mathbf{R})$  is some trial function. Multiplying the Schrödinger equation with  $\Psi_T$  and rearranging the terms leads to the following equation for the mixed distribution  $f$

$$\begin{aligned} -\frac{\partial f(\mathbf{R}, \tau)}{\partial \tau} &= -\frac{1}{2} \nabla^2 f(\mathbf{R}, \tau) + \nabla \cdot [\mathbf{v}_D(\mathbf{R}) f(\mathbf{R}, \tau)] \\ &\quad + (E_L(\mathbf{R}) - E_T) f(\mathbf{R}, \tau), \end{aligned} \quad (33)$$

where  $\mathbf{v}_D(\mathbf{R}) = \nabla \ln |\Psi_T(\mathbf{R})|$  represents a drift velocity term and  $E_L(\mathbf{R})$  is the local energy given by Eq. (22). The corresponding short-time approximation of the Green's function is given by

$$\begin{aligned} \tilde{G}(\mathbf{R} \leftarrow \mathbf{R}', \Delta\tau) &\approx (2\pi\Delta\tau)^{-3N/2} \exp \left[ -\frac{(\mathbf{R} - \mathbf{R}' - \Delta\tau \mathbf{v}_D(\mathbf{R}'))^2}{2\Delta\tau} \right] \\ &\times \exp \left[ -\Delta\tau \left( \frac{E_L(\mathbf{R}) + E_L(\mathbf{R}')}{2} - E_T \right) \right] \\ &= \tilde{G}_D(\mathbf{R} \leftarrow \mathbf{R}', \Delta\tau) \times \tilde{G}_B(\mathbf{R} \leftarrow \mathbf{R}', \Delta\tau). \end{aligned} \quad (34)$$

Here, we can again interpret the action of the short-time Green's function as a diffusion-drift process evolved according to  $\tilde{G}_D$  and the re-weighting process  $\tilde{G}_B$  controlled by the average of local energies at  $\mathbf{R}$ ,  $\mathbf{R}'$  and  $E_T$ . This transformation has several implications. First, due to the drift, the density of walkers is enhanced in the regions where  $\Psi_T(\mathbf{R})$  is large. Second, the re-weighting term contains the average of local energies, which for  $\Psi_T$  approximating  $\Psi_0$ , is close to a constant and therefore much better behaved than possibly unbounded potential  $V(\mathbf{R})$ . By a proper choice and construction of  $\Psi_T$  one can eliminate large amount of statistical noise, potential singularities, etc. For any large-scale calculation accurate  $\Psi_T$  is important and improves the efficiency by two orders of magnitude or more.

In actual simulations, the walkers are initially distributed according to  $f(\mathbf{R}, 0) = |\Psi_T(\mathbf{R})|^2$  with the weights of all walkers set to unity. The propagation then proceeds by iterative application of the Green's function. First, the position of each walker evolves according to the diffusion-drift  $\tilde{G}_D$  from Eq. (34). However, accuracy of  $\tilde{G}_D$  is not uniform and decreases in regions with strongly varying drift. In order to correct for this finite time step bias the walker move is accepted with the Metropolis probability

$$A(\mathbf{R} \leftarrow \mathbf{R}') = \min \left( 1, \frac{\tilde{G}_D(\mathbf{R}' \leftarrow \mathbf{R}) |\Psi_T(\mathbf{R}')|^2}{\tilde{G}_D(\mathbf{R} \leftarrow \mathbf{R}') |\Psi_T(\mathbf{R})|^2} \right). \quad (35)$$

The time step  $\Delta\tau$  is chosen to be small so that the acceptance in the DMC method is high, typically 99% or so. The Metropolis step improves the Green's function approximation appreciably and decreases the time step bias due to the fact that it enforces the detailed balance condition which is fulfilled by the exact Green's function.

After evolving the walker position we need to update the walker weight according to

$$w_m(\tau + \Delta\tau) = \tilde{G}_B(\mathbf{R}_m \leftarrow \mathbf{R}'_m, \Delta\tau) w_m(\tau). \quad (36)$$

Note that over the course of the simulation some weights will grow very rapidly due to the exponential nature of the weight product. As a result, in the long-time limit, the weight of a single walker will exponentially dominate the rest. It is therefore necessary to introduce importance sampling also for the weights and adjust the walker population to improve the efficiency of the reweighting. One possible way how to accomplish this is stochastic *branching* and *elimination* of walkers. Consider a walker evolving from  $\mathbf{R}'_m$  to  $\mathbf{R}_m$  according to  $G_D(\mathbf{R}_m \leftarrow \mathbf{R}'_m)$ . The weight for the step becomes  $w_m = G_B(\mathbf{R}_m \leftarrow \mathbf{R}'_m)$ . We can now calculate  $n_m = \text{int}[\eta + w]$ , where  $\eta$  is a uniform random number from the interval  $[0, 1]$ , and  $\text{int}$  is the integer part.

The resulting integer  $n_m$  is the *the number of walker copies* placed at the position  $\mathbf{R}_m$  and subsequently evolved as independent walker(s). In case  $n_m = 0$ , the walker is eliminated from the ensemble. In this manner, the branching/elimination step enables to keep all the weights  $w_m = 1, m = 1, \dots, M$  throughout the simulation. More elaborated approaches for sampling the weights can be found elsewhere [39].

After an equilibration period, we can start to collect the statistics needed for the calculation of the projected ground state energy

$$\begin{aligned} E_{DMC} &= \lim_{\tau \rightarrow \infty} \frac{\int \Psi^*(\mathbf{R}, \tau) \mathcal{H} \Psi_T(\mathbf{R}) d\mathbf{R}}{\int \Psi^*(\mathbf{R}, \tau) \Psi_T(\mathbf{R}) d\mathbf{R}} \\ &= \lim_{\tau \rightarrow \infty} \frac{\int f(\mathbf{R}, \tau) E_L(\mathbf{R}) d\mathbf{R}}{\int f(\mathbf{R}, \tau) d\mathbf{R}} \\ &= \frac{1}{M} \sum_m E_L(\mathbf{R}_m) + \mathcal{O}(1/\sqrt{M}) \end{aligned} \quad (37)$$

or other quantities.

### 2.3.3 Fixed-Node Approximation

For simplicity, we assume that the Hamiltonian has the time reversal symmetry (i.e., no magnetic field, induced currents, etc) so that the states can be chosen real. For bosonic systems, the ground state wave function has no nodes and is positive everywhere. However, the fermionic wave function is by definition antisymmetric, and therefore can no longer be represented by the walker distribution, which has to be non-negative everywhere. Clearly, for fermions, the wave function will have both positive and negative amplitude domains defined by the nodes—subsets of the configuration space where the wave function vanishes, i.e.,  $\Psi(\mathbf{R}) = 0$ . (The nodal properties of wave functions will be studied in detail in the subsequent section.)

We could try to overcome this problem by decomposing  $\Psi(\mathbf{R}, \tau)$  into positive functions  $\Psi^+(\mathbf{R}, \tau)$  and  $\Psi^-(\mathbf{R}, \tau)$  and assign each walker a positive or a negative weight accordingly

$$\Psi(\mathbf{R}, \tau) = \Psi^+(\mathbf{R}, \tau) - \Psi^-(\mathbf{R}, \tau). \quad (38)$$

However, since the Schrödinger equation is linear, both  $\Psi^+(\mathbf{R}, \tau)$  and  $\Psi^-(\mathbf{R}, \tau)$  independently converge to the same bosonic distribution, hence making the fermionic component to decay in time (see Fig. 1).

This is a well-known *fermion sign problem* where the fermionic “signal” disappears exponentially quickly in the projection time. The method is possible to use providing one can estimate quantities of interest before the efficiency deteriorates, nevertheless, for large systems the method becomes very inefficient. The inefficiency can be eliminated at the price of introducing the so-called *fixed-node approximation* [36, 40–42], which restricts the evolution of positive and negative walkers within the separate regions defined by the sign domains of the trial wave function  $\Psi_T(\mathbf{R})$ . This is achieved by enforcing the mixed distribution  $f(\mathbf{R}, \tau) = \Psi_T(\mathbf{R})\Psi(\mathbf{R}, \tau)$  to be non-negative everywhere

$$f(\mathbf{R}, \tau) \geq 0. \quad (39)$$

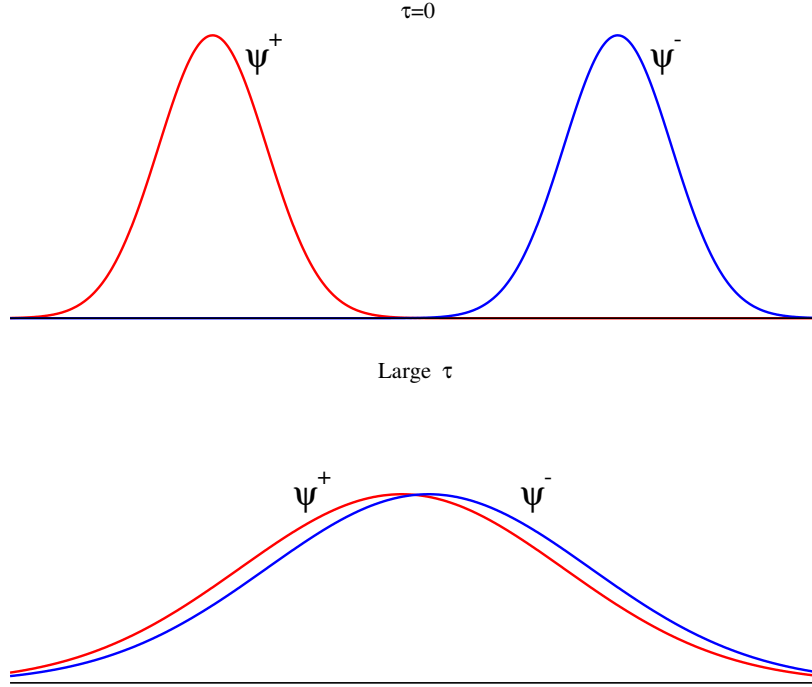


Fig. 1. Imaginary time behavior of the walker distributions  $\Psi^+$  (red) and  $\Psi^-$  (blue). Upper figure: at the beginning of the propagation ( $\tau = 0$ ). Lower figure: after some imaginary time evolution. The fermionic signal decays as  $(\Psi^+ - \Psi^-) \propto \exp[-\tau(E_0^F - E_0^B)]$ , where  $E_0^F - E_0^B$  is the difference of fermionic and bosonic ground state energies.

In other words,  $\Psi(\mathbf{R}, \tau)$  shares the same zeros and sign structure as  $\Psi_T(\mathbf{R})$  for all  $\tau$ , so that Eq. (39) holds. This is an important conceptual change in the sampling process: the fixed-node approximation replaces the (nonlocal) antisymmetric condition by the local boundary condition in Eq. (39). Note that the drift term diverges at the node  $\Psi_T(\mathbf{R}) = 0$  and therefore repels the walkers away from the nodal regions. The fixed-node approximation has several appealing properties. The change of symmetries into boundaries effectively *bosonizes* the fermionic problem and enables us to apply the statistical machinery to solve the evolution equation. The sketch of the fixed-node DMC in Fig. 2 reminds us Feynman path integrals and the walker paths really are closely related to quantum paths except that they happen in imaginary instead of real time.

The fixed-node solution provides the lowest possible energy within the given boundary conditions and as such is a rigorous upper bound to the true ground state energy [36, 42]. This result also implies that resulting fixed-node energy bias is proportional to the square of the nodal displacement error and in actual applications the method has proved to be very successful as further elaborated later.

The basic form of the fixed-node DMC method is conceptually rather straightforward. Let us therefore sketch a simple version of the algorithm. We assume that  $\Psi_T(\mathbf{R})$  approximates the

desired ground state and conforms to the required symmetries. The algorithm is as follows.

- (1) Generate a set of walkers  $\{\mathbf{R}_m\}_{m=1}^M$  distributed according to  $\Psi_T^2$ , for example, by using the VMC method.
- (2) Evaluate the drift  $\mathbf{v}(\mathbf{R}) = \nabla_{\mathbf{R}} \ln |\Psi_T(\mathbf{R})|$  for each walker.
- (3) Evolve each walker from its position  $\mathbf{R}'$  to a new position  $\mathbf{R}$  according to

$$\mathbf{R} \leftarrow \mathbf{R}' + \Delta\tau \mathbf{v}(\mathbf{R}) + \chi, \quad (40)$$

where  $\chi$  is a 3N-dimensional vector with components drawn from the normal distribution with zero mean and variance  $\Delta\tau$ .

- (4) Accept the move with the probability

$$P_{acc}(\mathbf{R} \leftarrow \mathbf{R}') = \min \left[ 1, \frac{G_D(\mathbf{R} \leftarrow \mathbf{R}') \Psi_T^2(\mathbf{R})}{G_D(\mathbf{R}' \leftarrow \mathbf{R}) \Psi_T^2(\mathbf{R}')} \delta_{ss'} \right], \quad (41)$$

where  $s = \text{sign}[\Psi_T(\mathbf{R})]$  and  $s' = \text{sign}[\Psi_T(\mathbf{R}')].$

- (5) For each walker calculate the number of copies positioned at  $\mathbf{R}$  as

$$n(\mathbf{R}) = \text{int}\{\eta + \exp[-\tau(E_L(\mathbf{R}') + E_L(\mathbf{R}) - 2E_T)/2]\}, \quad (42)$$

where  $\eta$  is drawn from a uniform random distribution on the interval  $[0, 1]$ .

- (6) Evaluate averages of interest over the ensemble of walkers and update  $E_T$  as

$$E_T(\tau + \Delta\tau) = E_T(\tau) - C_{adj} \ln(M_{act}/M_{ave}), \quad (43)$$

where  $M_{act}$  is the actual walker population,  $M_{ave}$  is the desired walker population and  $C_{adj}$  is a positive constant, which tunes how quickly the population would reach the desired value. The constant is typically chosen to rebalance the population to the desired number in a few steps, say, for 10 steps its value is  $\approx 10/\tau$ .

- (7) Repeat the steps (2) to (6) until the error bars of the calculated quantities reach the desired values.

Note that in the step 4) we correct the short-time Green's function approximation in two ways. Besides the acceptance step explained previously, the fixed node approximation is enforced for the case when a walker violates the nodal boundary. This occasionally happens, again, due to the Green's function inaccuracy. Since the evolution is discretized, the inherent fluctuations in the diffusion step can cause the walker to cross the nodal boundary—although for the exact Green's function the probability of such an event vanishes. As mentioned above, the time step  $\Delta\tau$  is usually chosen to be small enough so that the acceptance is 99% or higher. Any remaining time step bias can be eliminated by extrapolating  $\Delta\tau \rightarrow 0$  from repeated simulations with decreasing  $\Delta\tau$ .

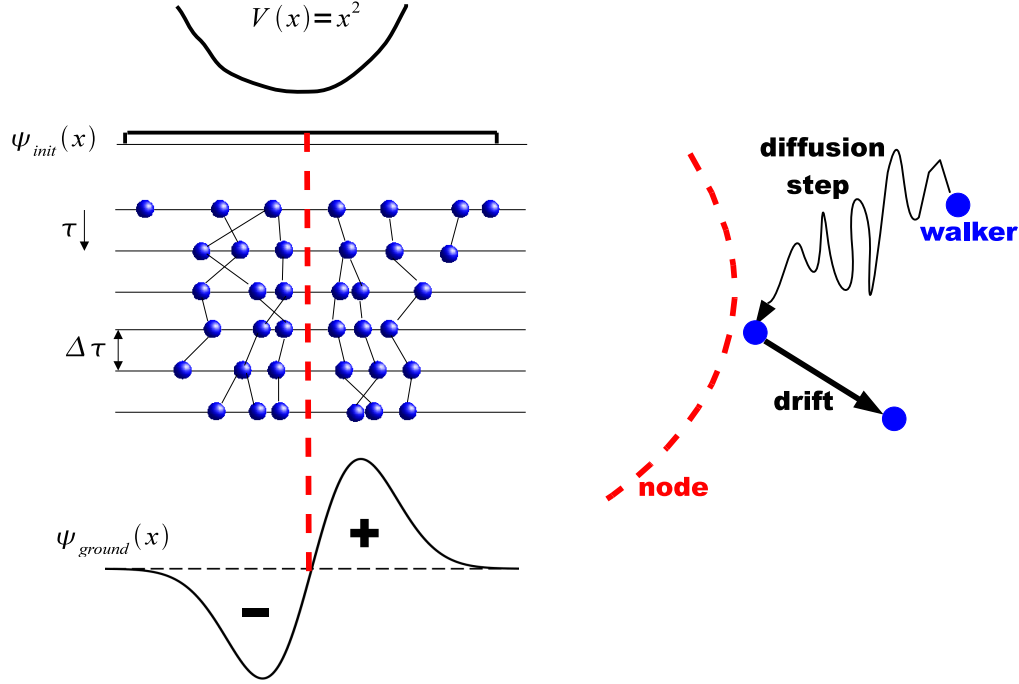


Fig. 2. On the left is a sketch of the fixed-node diffusion Monte Carlo method. The first excited state of 1D harmonic oscillator serves as a fermionic ground state toy model. Walkers are generated from a constant initial distribution and are divided into “plus” and “minus” populations according to the trial function (in 1D the node is a point). Enforcing the fixed node condition implies that the two populations never mix. The plot on the right shows walker diffusion and drift steps close to the node which is depicted by the dashed line.

In order to clarify the fixed-node DMC scaling of the computational time in the system size we invoke the previously given formula Eq.(20). We assume a requirement of the fixed error bar  $\epsilon$  for the total energy. The growth of variance  $\sigma^2$  is basically proportional to the number of electrons  $N$ , since for large systems fluctuations from different electrons will be approximately independent, so that  $\sigma^2 \propto N$ . The time  $T_{sample}$  then includes evaluations of the wave function and operators—such as the kinetic energy, etc—and we can write for  $T_{sample} = c_2 N^2 + c_3 N^3$ . The quadratic component comes from the pair terms in the Jastrow factor and also from filling the Slater matrices, while the cubic part is from evaluating the determinant. For systems with up to a few thousand electrons  $c_2 \gg c_3$  so that the quadratic term dominates. There is also some cost associated with an increase in statistical correlation of samples which we estimate as being of the order of  $N$ . The overall scaling then looks like  $N^4$ . However, this appears to be a rather conservative estimate. For larger systems one can recast the one-particle orbitals into the Wannier orbitals which are localized and also impose sufficiently loose cut-off on the range of the Jastrow correlations. That makes the Slater matrix sparse [6] and basically eliminates one power of  $N$ . If we are interested in intensive quantities such as cohesion (energy per particle) that

eliminates another factor of  $N$ . For some energy differences correlated sampling shows similar improvements and so forth. The algorithms and implementation has a significant impact here and in practice the fixed-node DMC calculations scale approximately as  $N^p$  where  $2 \leq p \leq 3$ .

### 2.3.4 Beyond The Fixed-Node Approximation

The fixed-node approximation has been an important step on the way to overcome the fundamental difficulties with fermion signs. Besides improving the nodes with more accurate trial wave functions (see Sec. 4 and 5) or attempts to understand them (see Sec. 3), several ideas how to reach beyond the limits of fixed-node approximation were proposed and tested.

Ceperley and Alder [43] developed the *released-node method*. As the name suggests, the walkers are allowed to sample the regions of opposite sign by using the guiding distribution, which is positive everywhere including the nodal regions. The method suffers from the fermion sign problem but enables to somewhat better control the loss of the fermion signal by appropriate choice of the guiding distribution and other tune-ups. The algorithm produces transient estimators so that the decay of fermionic excitations has to be faster than the decay of the fermionic ground state component. Let us explain the basic idea of the release node approaches. The fixed-node condition is eliminated and instead of  $\Psi_T$  the importance sampling is carried out with another importance distribution  $\Psi_G$  which is sometimes called the *guiding* wave function. Unlike  $\Psi_T$ , which is antisymmetric,  $\Psi_G$  is *symmetric* and positive in the whole configurations space, allowing thus the walkers to reach any point. The DMC simulation then runs according to the importance sampled Eq.(33), where  $\Psi_T$  is replaced by  $\Psi_G$ . The resulting release node  $\Psi_{RN}$  solution is found from  $f(R, \tau) = \Psi_{RN}\Psi_G$ . The release node solution is transient and can be qualitatively expressed as

$$\Psi_{RN} \approx c(\tau)\Psi_F(\tau) + (1 - c(\tau))\Psi_B(\tau), \quad (44)$$

where  $c(\tau) \approx \exp[-(E_F - E_B)\tau]$  and indices  $F$  and  $B$  correspond to fermions and bosons, respectively. The energy of the fermionic component can be evaluated from

$$\begin{aligned} E_{RN}(\tau) &= \frac{\int \Psi_G \Psi_{RN} \frac{\Psi_T}{\Psi_G} \frac{\mathcal{H}\Psi_T}{\Psi_T} d\mathbf{R}}{\int \Psi_G \Psi_{RN} \frac{\Psi_T}{\Psi_G} d\mathbf{R}} \\ &= \frac{\int f(\mathbf{R}, \tau) \frac{\Psi_T}{\Psi_G} E_L(\mathbf{R}) d\mathbf{R}}{\int f(\mathbf{R}, \tau) \frac{\Psi_T}{\Psi_G} d\mathbf{R}} \\ &= \frac{\sum_m s_m \frac{\Psi_T(\mathbf{R}_m)}{\Psi_G(\mathbf{R}_m)} E_L(\mathbf{R}_m)}{\sum_m s_m \frac{\Psi_T(\mathbf{R}_m)}{\Psi_G(\mathbf{R}_m)}} + \mathcal{O}(e^{\alpha\tau}/\sqrt{M}), \end{aligned} \quad (45)$$

where  $\alpha \propto (E_F - E_B)$ . Each weight  $s_m$  carries a *sign* which is assigned initially as  $w_m(\tau = 0) = \text{sign}[\Psi_T(\mathbf{R}_m)]$  and kept constant during the simulation. A simple guiding function can be written as  $\Psi_G = \sqrt{\Psi_T^2 + \alpha}$ , where the parameter  $\alpha$  controls the average rate of walkers passing into the opposite sign region. More sophisticated forms are necessary for inhomogeneous systems. The transient sampling efficiency can be measured by the ratio

$$\kappa = \frac{1}{M} \sum_m s_m \frac{\Psi_T(\mathbf{R}_m)}{\Psi_G(\mathbf{R}_m)} \quad (46)$$

which, ideally, should be close to unity (as is the case for the fixed-node algorithm when  $\Psi_G = |\Psi_T\rangle$ ). As the simulation proceeds, the ratio decreases and the fermionic signal decays while the statistical noise grows. Despite this, the method can be useful providing one can reach the fermionic ground state estimator before the error bars become too large. In fact, the release node method was used for the famous calculations of the homogeneous electron gas correlation energies [23] and for small atoms and molecules later [43]. Note that the simplicity of the homogeneous electron gas enabled to apply this method to hundreds of electrons already in 1980. Unfortunately, applications to inhomogeneous systems are much more difficult and have progressed at much slower pace.

Another idea advanced by Kalos and collaborators is based on the *cancellation of paired walkers* with opposite signs [44]. This method has been subsequently improved to work with fully interacting ensemble of walkers [45]. The method delays the fermionic decay and enhances the fermion solution component, however, the current understanding is that the scaling appears to be exponential. An important improvement is, however, that the exponent is smaller than the difference between the bosonic and fermionic energies and therefore the technique has a potential for efficiency gains in some applications. [46].

Anderson and Traynor [47] combined some of the ideas in the fixed-node, released-node and cancellation methods in an algorithm which employed improvements such as relocation after the node crossing, self and multiple cancellations, and maximum use of symmetry. Among the systems of interest, the authors calculated the excited state of  $\text{H}_2$   $^3\Sigma_u^+$  and the barrier height for the simplest possible chemical reaction  $\text{H} + \text{H}_2 \rightarrow \text{H}_2 + \text{H}$  [48]. The accuracy of this essentially exact calculation was exceptional with resulting statistical uncertainty of only 0.01 kcal/mol.

Among other alternative attempts, let us also mention the method of *adaptive nodes* (see, e.g., Ref. [49]). The key idea is based on representing the approximate nodal function directly by the walker ensemble (e.g., by positioning a Gaussian on each walker). This results in the adaptive description of the nodes that does not depend upon *a priori* knowledge of the wave function. The main difficulty of the method, however, comes from scaling to larger systems since one basically has to populate the whole configurations space with walkers.

In simulations of quantum problems the fermion sign problem is pervasive. It appears in various incarnations and determines computational complexity of the underlying problem. It has been recently shown [50] explicitly, that QMC simulations of random coupling Ising model with fermion sign problem belongs to the class of *NP-hard* problems, i.e., with non-polynomial computational complexity. There are exceptions which enable to eliminate the fermion sign problem exactly such as separable Hamiltonians or problems with certain symmetries (the half-filled Hubbard model, for example). However, all such known cases are “special” having some underlying symmetries which enable to eliminate the fermion signs. It is therefore possible that any *exact* QMC approach with broad applicability belongs to the class of NP-hard problems and therefore finding such method might be an elusive and exceedingly challenging, if not impossible goal. The next best goal should be therefore based on a more pragmatic view on what does it mean to “solve” a given many-body quantum problem. Perhaps the right question to ask can be formulated as follows: What is the most optimal scaling of a method which can calculate expectations within a prescribed error? (say, energy differences with errors below 0.1 eV, 0.01 eV and so forth). Here by the error we mean the total error which contains both *systematic biases* from approximations and *statistical uncertainties*. The answer to this question is a lot easier, at



least in a qualitative sense, since already at present the electronic structure QMC methods can achieve accuracies of 0.1-0.2 eV for energy differences with low-order polynomial scaling in the system size. Similar progress has been achieved also in simulations in other areas of physics.

## 2.4 Trial Variational Wave Functions

The great advantage of QMC methods is the capability to use explicitly correlated trial wave functions. Their quality is the most important factor which influences both the achieved accuracy and efficiency. On the other hand, repeated evaluations of  $\Psi_T$  (and  $\nabla\Psi_T$ ,  $\nabla^2\Psi_T$ , etc) is the most costly part of QMC calculation and therefore restrict possible alternatives for  $\Psi_T$  to forms which can be evaluated rapidly and scale with the system size.

### 2.4.1 Basic Properties

$\Psi_T$  has to be antisymmetric with respect to exchange of both spin and spatial coordinates

$$\Psi_T(P\mathbf{R}, P\mathbf{\Sigma}) = (-1)^P \Psi_T(\mathbf{R}, \mathbf{\Sigma}), \quad (47)$$

where  $\mathbf{\Sigma} = (\sigma_1, \dots, \sigma_N)$  are discrete spin variables with values  $\pm\frac{1}{2}$  and  $P$  is an arbitrary permutation with the sign equal to  $(-1)^P$ . Assuming that the Hamiltonian (or other quantities of interest) does not involve spin-dependent terms, any expectation value requires only the spatial antisymmetry since the summation over spin variables becomes just a permutation of indices inside the same spatial integral. That enables to assign the spins to particular electrons and requires imposing antisymmetry only with respect to spatial interchanges between the electrons of the same spin. The spins of all electrons are therefore fixed to the total spin projection equal to  $\frac{1}{2}(N_\uparrow - N_\downarrow)$ . The first  $N_\uparrow$  particles are labeled as spin-up and the rest  $N - N_\uparrow$  as spin-down. Throughout this paper we assume that the spin variables are factored out by this labeling scheme. Using an appropriate expansion in determinants the wave function can be made an eigenstate of the total spin  $S^2$  and possibly other integrals of motion such as the total angular momentum  $L^2$  and  $L_z$  in atoms.

We demand that  $\Psi_T$  approximates a given bound state so that  $\int \Psi_T^* \Psi_T$ ,  $\int \Psi_T^* \mathcal{H} \Psi_T$  and  $\int \Psi_T^* \mathcal{H}^2 \Psi_T$  do exist. The last condition is necessary for having a finite variance. Further,  $\Psi_T$  and  $\nabla\Psi_T$  have to be continuous, wherever the potential is finite. In addition, the continuity of the potential implies also the continuity of  $\nabla^2\Psi_T$ . Important corollary is that as  $\Psi_T$  approaches an exact solution, the local energy becomes constant everywhere. We can express this also as  $\sigma^2 = \langle (\mathcal{H} - \langle \mathcal{H} \rangle)^2 \rangle \rightarrow 0$  for  $\Psi_T \rightarrow \Phi_0$ , and it is often called as *zero variance property*. Note that both VMC and DMC estimators have the zero variance property. (This is not automatically true for many QMC approaches, in particular, this does not apply whenever there are non-localities in the Hamiltonian. For example, many quantum lattice models such as the Hubbard model are nonlocal due to hopping terms.)

We wish that the local energy is close to the eigenvalue and also that the variance as small as possible. Therefore we demand that any singularity from the potential is to be canceled by the opposite sign singularity in the kinetic energy, what gives rise to Kato *cusp conditions* [51, 52]. As the distance of electron to nucleus vanishes,  $r_{Ii} \rightarrow 0$ , the potential energy divergence cancels

out when

$$\left. \frac{1}{\Psi_T} \frac{\partial \Psi_T}{\partial r_{Ii}} \right|_{r_{Ii}=0} = -Z_I. \quad (48)$$

This electron-nucleus cusp condition [Eq. (48)] can be satisfied by a proper choice of orbitals, removed by pseudopotentials (see Sec. 2.5) or enforced by Jastrow factor. Of course, the hydrogenic solutions such as, for example, 1s and 2p orbitals given as  $\phi_{1s}(r) = \exp(-Z_I r)$  and  $\phi_{2p}(r) = \exp(-Z_I r/2) Y_{1m}$ , have the correct cusp. In addition, it is clear that any linear combination of orbitals with the correct cusp has the same property. In particular, if all orbitals in a Slater determinant have correct cusps then the same is true for the determinant what is easy to verify by an expansion in cofactors.

Similarly, as the distance  $r_{ij}$  of any two electrons vanishes, there is a divergence in the electron-electron Coulomb potential. Let us first assume that electrons  $i$  and  $j$  have unlike spins. In general, the wave function does not vanish at such a collision point. Close to the singularity we can write it as an s-wave in  $r_{ij}$  times a smooth function which depends on the rest of the variables  $\tilde{\mathbf{R}}$ , i.e.,  $\lim_{r_{ij} \rightarrow 0} \Psi(\mathbf{R}) = \exp(cr_{ij}) F(\tilde{\mathbf{R}})$ . Since there are two contributing Laplacians  $\Delta_i + \Delta_j$  we get

$$\left. \frac{1}{\Psi(\mathbf{R})} \frac{\partial \Psi(\mathbf{R})}{\partial r_{ij}} \right|_{r_{ij}=0} = c = \frac{1}{2}. \quad (49)$$

For two like spins electrons with  $r_{ij} \rightarrow 0$  the wave function vanishes. The behavior around the collision point has a character of a p-wave. Let us transform the coordinates  $\mathbf{r}_i$  and  $\mathbf{r}_j$  to  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$  and  $\mathbf{r}_{ij}^+ = \mathbf{r}_i + \mathbf{r}_j$ . The wave function can be then written as  $\lim_{r_{ij} \rightarrow 0} \Psi(\mathbf{R}) = f(\mathbf{r}_{ij}) \exp(cr_{ij}) F(\tilde{\mathbf{R}})$ . By an appropriate coordinate rotation the linear behavior around  $r_{ij} \rightarrow 0$  can be expressed as  $f(\mathbf{r}_{ij}) = (x_i - x_j) + \mathcal{O}((x_i - x_j)^2)$ . The like spins cusp condition therefore modifies to

$$\left. \frac{1}{\Psi} \frac{\partial \Psi}{\partial r_{ij}} \right|_{r_{ij}=0} = c = \frac{1}{4}. \quad (50)$$

#### 2.4.2 Trial Wave Functions Forms

It is convenient to express  $\Psi_T(\mathbf{R})$  as a product of an antisymmetric part  $\Psi_A$  and an exponential of a symmetric correlation function  $U_{corr}(\mathbf{R})$ , often called the Jastrow factor, as given by

$$\Psi_T(\mathbf{R}) = \Psi_A \times \exp[U_{corr}(\mathbf{R})]. \quad (51)$$

Let us describe the forms for  $\Psi_A(\mathbf{R})$ ,  $U_{corr}(\mathbf{R})$  and their generalizations as commonly used in calculations (and also implemented in our QMC code QWalk [53]). The Jastrow part [36, 54–56] explicitly depends on electron-nucleus and electron-electron distances

$$U_{corr}(\{r_{ij}\}, \{r_{iI}\}, \{r_{jI}\}) = \sum_{iI} \chi(r_{iI}) + \sum_{i \neq j} u(r_{ij}) + \sum_{i \neq j, I} w(r_{ij}, r_{iI}, r_{jI}) \quad (52)$$

and is written as the sum of one-body (electron-nucleus), two-body (electron-electron) and three-body (electron-electron-nucleus) terms. These correlation functions describe several effects.

First, the form captures the electron-electron cusp conditions [Eqs. (49) and (50)]. As derived above, these are satisfied providing

$$\left. \frac{\partial u_{\sigma_i \sigma_j}}{\partial r_{ij}} \right|_{r_{ij}=0} = \begin{cases} 1/4 & \sigma_i = \sigma_j, \\ 1/2 & \sigma_i = -\sigma_j, \end{cases} \quad (53)$$

and while all the remaining functions have zero derivatives for  $r_{ij} = 0$ . Second, the terms which depend on electron-ion distances enable to vary the electron density which needs to be adjusted once the electron-electron terms are included. Third, the triple electron-electron-ion correlations enable to decrease the correlation effects and decrease the local energy fluctuations. Expanded in the basis of one dimensional functions the components of the Jastrow factor take form

$$\chi(r) = \sum_k c_k^{en} a_k(r), \quad (54)$$

$$u(r) = f_{cusp}(r) + \sum_k c_k^{ee} b_k(r), \quad (55)$$

$$w(r_{ij}, r_{iI}, r_{jI}) = \sum_{klm} c_{klm}^{een} [a_k(r_{iI}) a_l(r_{jI}) + a_k(r_{jI}) a_l(r_{iI})] b_m(r_{ij}). \quad (56)$$

The above Jastrow factor has proved to be very efficient in describing significant portion of correlation effects with small number of variational parameters. More details about the implemented basis functions can be found below or in the QWalk documentation [53].

The simplest antisymmetric functional form is the determinant so that one can expand the wave function in spin-factorized Slater determinants as

$$\Psi_A^{Slater}(\mathbf{R}) = \sum_k d_k \det[\dots, \varphi_k^\uparrow(\mathbf{r}_i), \dots] \det[\dots, \varphi_l^\downarrow(\mathbf{r}_j), \dots] \quad (57)$$

where  $i = 1, \dots, N_\uparrow$  and  $j = N_\uparrow + 1, \dots, N$ . Each spin-up and spin-down Slater determinant is a function of one-particle orbitals typically found by (post-)HF or DFT methods. In the introductory part we have touched upon some of the properties of  $\Psi_A^{Slater}$ . Although formally the Slater determinants provide a complete set of antisymmetric functions, in practice, we are limited to use only finite expansions, very often to a single determinant. It is remarkable that the *single determinant Slater-Jastrow* form proved to be very effective for many applications, perhaps beyond expectations. It provides surprisingly accurate energy differences such as cohesion and binding, excitations, etc, due to partial cancellations of the fixed-node errors.

On the other hand, multi-determinant expansions which are used to improve the fixed-node energies require rather involved and technically difficult processing such as the following tasks:

- (a) Large determinantal expansions often need computationally expensive reoptimizations of the determinantal coefficients. At present, with such wave functions we are thus bound to small or intermediate sizes of chemical systems.
- (b) For very high accuracy calculations one-particle orbitals need to be reoptimized. Typically, orbitals come from methods which do not incorporate the explicit correlations and therefore are not optimal with Jastrow correlation factor included. However, full optimization of orbitals in many-particle systems is still a formidable challenge. What has been

successfully tried is to parameterize the effective one-particle Hamiltonian such as hybrid DFT functionals and find optimal orbitals in such restricted parameter space.

The above limitations motivate to think about possibilities beyond  $\Psi_T^{S-J}$ . Since the determinant is the simplest antisymmetric form constructed from one-particle orbitals, it is interesting to ask what is the simplest antisymmetric object built from two, three or  $n$ -particle orbitals? For two particle orbitals such antisymmetric form is called *Pfaffian* and the resulting *Pfaffian pairing wave function* is written as

$$\text{pf}_i(1, \dots, N) = \mathcal{A}[\tilde{\phi}(1, 2), \tilde{\phi}(3, 4), \dots, \tilde{\phi}(N-1, N)] \quad (58)$$

$$= \text{pf}[\tilde{\phi}(1, 2), \tilde{\phi}(3, 4), \dots, \tilde{\phi}(N-1, N)], \quad (59)$$

where  $\mathcal{A}$  is the antisymmetrizing operator and the arguments of the pairing spin-orbital  $\tilde{\phi}(i, j)$  denote both spatial and spin variables. Note that in the most basic form the Pfaffian contains only *one* pair orbital unlike the Slater determinant which requires as many independent orbitals as is the number of particles.

Similarly to multi-determinantal expansions, we can expand the wave function in a linear combination of Pfaffians

$$\Psi_A^{PF}(\mathbf{R}) = \sum_m w_m \text{pf}_m(\dots, \tilde{\phi}_m(i, j), \dots). \quad (60)$$

Note that each Pfaffian in the expansion (60) contains different pairing orbital. The form will be later simplified to spatial antisymmetries only such as in the case of determinants by employing spin-singlet and spin-triplet pairing orbitals. The description of Pfaffians together with applications to several systems and possible extensions is discussed in more detail in Sec. 4.

Another way how to boost the variational freedom of trial functions is based on the so-called backflow coordinates. The idea goes back to Feynman and has been employed for description of homogeneous systems [55, 57–62]. Feynman suggested that correlations in quantum liquids can be approximately captured by replacing the actual coordinates  $\mathbf{R}$  in  $\Psi_A(\mathbf{R})$  by quasi-particle coordinates  $\mathbf{X}$  which take into account the influence from the rest of the particles. As a consequence, the nodes and other properties of  $\Psi_A(\mathbf{X})$  will be, in general, different from those of  $\Psi_A(\mathbf{R})$ . Recently, some progress was also reported in applying this idea to chemical (inhomogeneous) systems [63, 64]. In this review we report some of the results we obtained with these new developments. The quasi-particle coordinate of the  $i$ -th electron at position  $\mathbf{r}_i$  is given as

$$\begin{aligned} \mathbf{x}_i &= \mathbf{r}_i + \boldsymbol{\xi}_i(\mathbf{R}) \\ &= \mathbf{r}_i + \boldsymbol{\xi}_i^{en}(\mathbf{R}) + \boldsymbol{\xi}_i^{ee}(\mathbf{R}) + \boldsymbol{\xi}_i^{een}(\mathbf{R}), \end{aligned} \quad (61)$$

where we have again divided the contributions to one-body (electron-nucleus), two-body (electron-electron) and three-body (electron-electron-nucleus) backflow terms. They can be further ex-

pressed as

$$\xi_i^{en}(\mathbf{R}) = \sum_I \chi(r_{iI}) \mathbf{r}_{iI} \quad (62)$$

$$\xi_i^{ee}(\mathbf{R}) = \sum_{j \neq i} u(r_{ij}) \mathbf{r}_{ij} \quad (63)$$

$$\xi_i^{een}(\mathbf{R}) = \sum_I \sum_{j \neq i} [w_1(r_{ij}, r_{iI}, r_{jI}) \mathbf{r}_{ij} + w_2(r_{ij}, r_{iI}, r_{jI}) \mathbf{r}_{iI}], \quad (64)$$

where  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ .  $\chi$ ,  $u$  and  $w_1$  with  $w_2$  are similar to one, two and three-body Jastrow terms. The calculational and implementation details together with some results are further discussed in Sec. 5.

### 2.4.3 Basis Functions for the Jastrow Correlation Factor

The one-particle orbitals and also the correlations has to be represented in an appropriate *basis* functions. There is a vast amount of literature on basis sets for expressing the one-particle orbitals. Typically, the localized basis sets are based on Gaussians, while for periodic systems plane waves are frequently used. Other options exist such as numerical orbitals, mixed basis, etc. We have little to add to this technically involved subject except that QWalk is able to import several types of these. We refer the reader to the literature or manuals for the actual packages.

However, we briefly outline the basis which is used for correlations in the Jastrow factor. The exact electron-electron cusp conditions are satisfied by the choice of following cusp function [see Fig. (3)] with variable  $x = r/r_{cut}$ , where  $r_{cut}$  is some cutoff radius and  $\gamma$  is the curvature as

$$f_{cusp}(r) = f_{cusp}(x, \gamma) = c \left( \frac{x - x^2 + x^3/3}{1 + \gamma(x - x^2 + x^3/3)} - \frac{1}{\gamma + 3} \right) \quad (65)$$

The cusp constant is  $c = \frac{1}{4}$  for electrons with like and  $c = \frac{1}{2}$  for electrons with unlike spins.

Polynomial Padé functions for the same variable  $x = r/r_{cut}$  and curvature  $\beta$  had proved to be a convenient choice for describing the electron-electron and electron-nucleus correlation. In calculations, we use the form

$$a_k(r), b_k(r) = f_{poly-Pade}(x, \beta) = \frac{1 - x^2(6 - 8x + 3x^2)}{1 + \beta x^2(6 - 8x + 3x^2)} \quad (66)$$

The  $f_{poly-Pade}(0) = 1$  with derivative  $f'_{poly-Pade}(0) = 0$  and also goes smoothly to zero as  $r \rightarrow r_{cut}$  [see Fig. (4)]. These conditions are necessary for preserving the cusp conditions already fixed by the cusp functions and by the choice of orbitals or pseudopotentials. Polynomial Padé functions are then used as basis functions in expansions for correlation part in the Jastrow factor, for example, for the electron-nucleus part we have

$$\chi(r) = \sum_k c_k^{en} a_k(r), \quad (67)$$

where  $c_k$  is an expansion coefficient and basis functions  $a_k(r)$  with different  $k$  differ in curvatures  $\beta_k$ . Similarly for other components as given by Eq. (67).

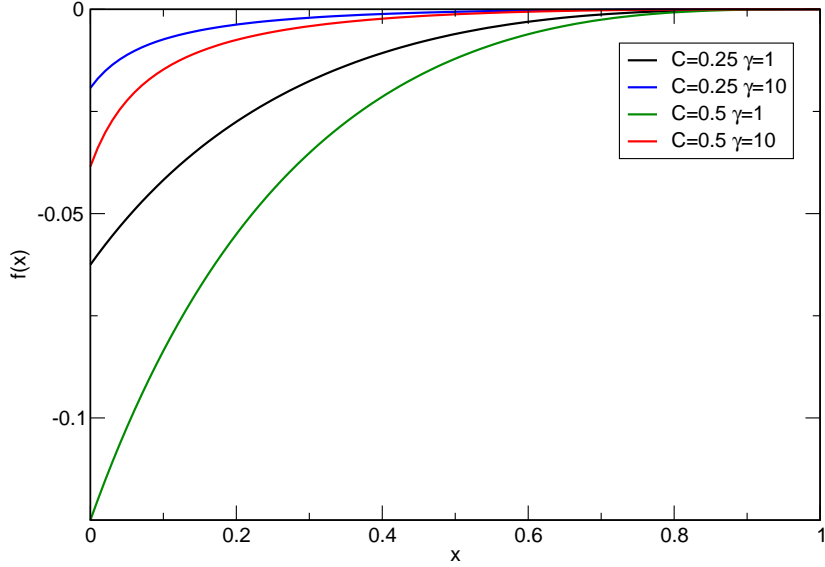


Fig. 3. Cusp functions with smooth cut-off for two different values of the curvature parameter ( $\gamma = 1$  and  $\gamma = 10$ ) for like ( $c = \frac{1}{4}$ ) and unlike ( $c = \frac{1}{2}$ ) spins.

## 2.5 Pseudopotentials

To a high degree of accuracy, most properties of real systems are determined by the states in valence electronic sub-shells only. Atomic cores are also expensive in QMC simulations since the computational cost increases with the atomic number  $Z$  as  $\approx Z^6$  (see Ref. [65, 66]) and makes the all-electron calculations unnecessarily difficult for heavy atoms. The presence of core electrons complicates calculations for several reasons. The short length scales associated with atomic cores states require decrease in the QMC time step. Even more difficult problem is the increase in fluctuations of the local energy as  $Z$  gets higher due strongly varying kinetic and potential terms close to the nucleus. This implies also basic efficiency issue since almost all computational time would go to sampling energy fluctuations in the core instead of the important valence effects.

In analogy with theories such as Hartree–Fock and DFT, in QMC the atomic core states can be eliminated by introducing an *pseudopotential (or effective core potential)* operator which represents the impact of core on valence states. The action of pseudopotential is typically different for states with different angular momenta. It is customary to divide the pseudopotential operator into local (common to all angular momenta) and nonlocal (different for each angular momentum) parts. For a given ion we can therefore eliminate the core states and replace the ionic term as follows

$$\frac{-Z}{r} \rightarrow U^{ps} = V^{ps}(r) + W^{nloc}, \quad (68)$$

where  $V^{ps}(r)$  is the local part while

$$W^{nloc} = \sum_l^{l_{max}} V_l^{ps}(r) \sum_m |lm\rangle\langle lm| \quad (69)$$

is the nonlocal pseudopotential with projectors over  $lm$  states of angular momentum. The non-local portion of the local energy  $E_L(\mathbf{R})$  from the electron  $i$  in the field of the ion  $I$  is given by

$$\frac{W^{nloc}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} = \sum_l V_l^{ps}(r_{iI}) \frac{2l+1}{4\pi} \int_{4\pi} P_l[\cos(\theta'_{iI})] \times \frac{\Psi_T(\mathbf{r}_1, \dots, \mathbf{r}'_i, \dots, \mathbf{r}_N)}{\Psi_T(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N)} d\Omega'_{iI}, \quad (70)$$

where  $P_l$  denotes the Legendre polynomial of degree  $l$ . The angular integral above is over the surface of the sphere passing through the  $i$ -th electron and centered on the ion. The integral is evaluated numerically using Gaussian spherical quadrature grids. When the orientation of axes of a given quadrature are chosen randomly, the result is an unbiased Monte Carlo estimation of the given integral (70). For more details, see original Ref. [67].

In the DMC method, the application of the nonlocal operator  $W^{nloc}$  causes the matrix element  $\langle \mathbf{R} | \exp[-\Delta\tau W^{nloc}] | \mathbf{R}' \rangle$ , and therefore also the propagator, to have a fluctuating sign. This would create its own fermion sign problem even for a single electron. Consequently, the imaginary time Schrödinger equation for the function  $f(\mathbf{R}, \tau) = \Psi_T(\mathbf{R})\Psi(\mathbf{R}, \tau)$  is rearranged

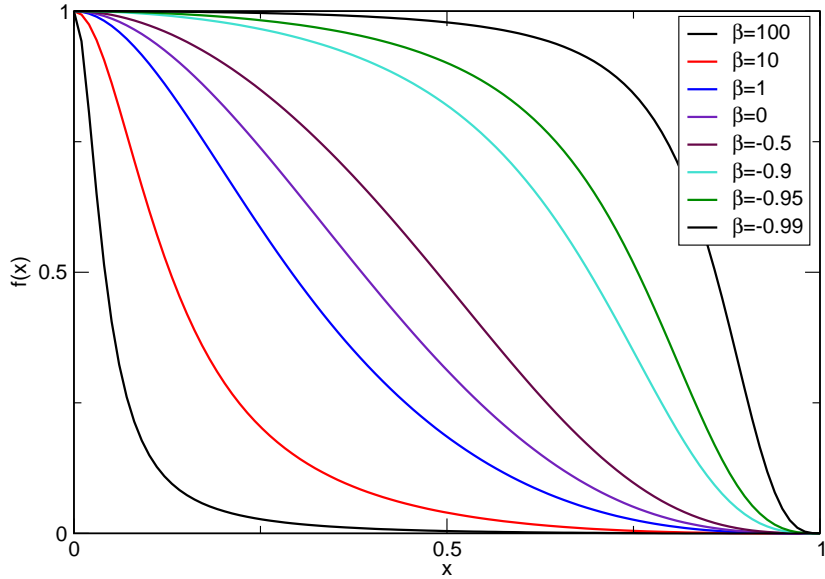


Fig. 4. Polynomial Padé functions with curvatures  $\beta$  ranging from -0.99 to 100.

as given by

$$-\frac{\partial f(\mathbf{R}, \tau)}{\partial \tau} = -\frac{1}{2}\nabla^2 f(\mathbf{R}, \tau) + \nabla \cdot [\mathbf{v}_D(\mathbf{R})f(\mathbf{R}, \tau)] + [E_L(\mathbf{R}) - E_T]f(\mathbf{R}, \tau) + \left\{ \frac{V_{nloc}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} - \frac{V_{nloc}\Psi(\mathbf{R}, \tau)}{\Psi(\mathbf{R}, \tau)} \right\} f(\mathbf{R}, \tau), \quad (71)$$

where  $V_{nloc}$  denotes the sum of all nonlocal pseudopotential components from all ions. Note that, unfortunately, we do not know how to calculate the last term since it contains the solution which is a priori unknown. We therefore introduce the *localization approximation* by neglecting the last term in Eq. (71). It has been shown [67] that the error in energy from the localization approximation  $\propto (\Psi_T - \Phi_0)^2$ . However, since the Hamiltonian is modified due to the neglected term, the obtained energy is not necessarily an upper bound to the true eigenenergy of the original Hamiltonian. Fortunately, accurate Jastrow-Slater trial wave functions enable to carry out calculations without appreciable impact from the localization error. Calculations have shown that the localization approximation bias is typically smaller than the fixed-node bias [68, 69]. Therefore this approximation proved to be less of a problem in practice since the localization error is overshadowed by the fixed-node error. The issue of the upper bound has been resolved some years ago once ten Haaf, van Bemmelen and co-workers [70, 71] suggested modification of Hamiltonians with nonlocal terms which does not violate the variational principle. Let us write the matrix element of the nonlocal part as

$$\langle \mathbf{R} | V_{nloc} | \mathbf{R}' \rangle = \langle \mathbf{R} | V_{nloc}^+ | \mathbf{R}' \rangle + \langle \mathbf{R} | V_{nloc}^- | \mathbf{R}' \rangle \quad (72)$$

where  $\langle \mathbf{R} | V_{nloc}^+ | \mathbf{R}' \rangle$  are matrix elements for which  $\langle \mathbf{R} | V_{nloc}^+ | \mathbf{R}' \rangle \Psi_T(\mathbf{R}) \Psi_T(\mathbf{R}') > 0$  and *vice versa* for  $\langle \mathbf{R} | V_{nloc}^- | \mathbf{R}' \rangle$ . For small times we can write

$$\langle \mathbf{R} | -\Delta\tau V_{nloc} | \mathbf{R}' \rangle = \delta(\mathbf{R} - \mathbf{R}') - \langle \mathbf{R} | V_{nloc} | \mathbf{R}' \rangle \Delta\tau + \mathcal{O}[(\Delta\tau)^2] \quad (73)$$

so that it is clear that the *negative* matrix elements do not flip the walker sign and can be sampled without introducing any sign problems. The sign flipping (positive) part is evaluated by the projection onto the trial function very much like in the localization approximation, i.e.,

$$V_{eff}^+ = \Psi_T(\mathbf{R})^{-1} \int \langle \mathbf{R} | V_{nloc}^+ | \mathbf{R}' \rangle \Psi_T(\mathbf{R}') d\mathbf{R}' \quad (74)$$

so that the new modified Hamiltonian has a form

$$\tilde{\mathcal{H}} = \mathcal{H}_{loc} + V_{eff}^+ + \int \langle \mathbf{R} | V_{nloc}^- | \mathbf{R}' \rangle d\mathbf{R}' + V_{node}^\infty(\Psi_T(\mathbf{R}) = 0) \quad (75)$$

where  $\mathcal{H}_{loc}$  contains all the local parts such as kinetic energies and local interactions including the local components of pseudopotentials. The node term enforces the fixed-node approximation explicitly (otherwise  $V_{eff}^+$  would be unbounded at the nodes). Interestingly, one can explicitly show that

$$\langle \Psi_T | \tilde{\mathcal{H}} | \Psi_T \rangle \geq \langle \Psi_T | (\mathcal{H}_{loc} + V_{nloc}) | \Psi_T \rangle \quad (76)$$



for any  $\Psi_T$  while the equality takes place when  $\Psi_T = \Psi_0$ , i.e., the energy of the modified Hamiltonian is always an upper bound to the original  $\mathcal{H} = \mathcal{H}_{loc} + V_{nloc}$ . New algorithmic developments by Casula and coworkers [72] enabled to implement this treatment of nonlocal operators into the DMC methods as the so-called  $T$ -moves.

Pseudopotentials are also very useful for heavy elements which require inclusion of relativistic effects. The scalar relativistic corrections can be built-in into the pseudopotentials and for very heavy elements this is basically the only choice if one wants to avoid the full relativistic treatment.

Although this section seem somewhat specialized and technical, the results have wider implications. Note that the last development enables us to work with *any type of nonlocal operators* using the decomposition from Eq. (72) and the effective upper-bound Hamiltonian from Eq. (75). Clearly, this opens possibilities for simulations with more complicated Hamiltonians, beyond just electrons and ions.

## 2.6 Optimization of Variational Wave Functions

The quality of trial wave functions has a significant impact on the efficiency of VMC and DMC methods as well as on the accuracy of the final results. The presented functional forms typically depend on a number of linear and non-linear parameters which has to be optimized. The optimization of functions is a classic numerical problem and many approaches can be found in the literature. However, in combination with random fluctuations and objective functions defined by the finite ensemble of walkers, the problem becomes more complicated. We will therefore outline a few methods used for this purpose.

### 2.6.1 Variance Minimization

Let us denote the set of variational parameters  $\{c\}$  of some real valued trial wave function  $\Psi_{\{c\}}$ . The mean value of the local energy with respect to  $\Psi_{\{c\}}$  is given by

$$E_{VMC}^{\{c\}} = \frac{\int \Psi_{\{c\}}^2 E_L^{\{c\}} d\mathbf{R}}{\int \Psi_{\{c\}}^2 d\mathbf{R}} = \langle E_L^{\{c\}} \rangle \equiv \bar{E}^{\{c\}}. \quad (77)$$

The variance of the local energy can be written as

$$\sigma_{\{c\}}^2 = \frac{\int \Psi_{\{c\}}^2 (E_L^{\{c\}} - \bar{E}^{\{c\}})^2 d\mathbf{R}}{\int \Psi_{\{c\}}^2 d\mathbf{R}} = \langle (E_L^{\{c\}} - \bar{E}^{\{c\}})^2 \rangle. \quad (78)$$

Since the variance directly determines the efficiency one possible approach for wave function optimizations is minimize the variance [56] in Eq. (78) using a finite set of fixed configurations, where the walkers are distributed according to  $\Psi_{\{c_0\}}^2$ . The set of starting variational parameters is denoted as  $\{c_0\}$ . The variance optimization is a rather robust strategy since it is always positive and bounded from below. In practice, the mean value of the local energy is not *a priori* known, so we replace  $\bar{E}^{\{c\}}$  in Eq. (78) by a reference energy  $E_{ref}$ .

The method can be further improved. One possibility is to use weights which account for the wave function change after the parameter updates. The *re-weighted variance* with the new set

of parameters  $\{c_{new}\}$  is then given as

$$\sigma_{\{c_{new}\}}^2 = \frac{\int \Psi_{\{c_0\}}^2 W_{\{c_{new}\}}(\mathbf{R}) (E_L^{\{c_{new}\}} - E_{ref})^2 d\mathbf{R}}{\int \Psi_{\{c_0\}}^2 W_{\{c_{new}\}}(\mathbf{R}) d\mathbf{R}}, \quad (79)$$

where we have introduced the weighting function

$$W_{\{c_{new}\}}(\mathbf{R}) = \frac{\Psi_{\{c_{new}\}}^2(\mathbf{R})}{\Psi_{\{c_0\}}^2(\mathbf{R})}. \quad (80)$$

The advantage of the reweighting scheme is a more accurate value of variance at each optimization step. However, once the wave function departs significantly from its original distribution one has to update the walker positions as well to reflect the updated wave function otherwise the method becomes unreliable.

For minimization algorithms we use two different methods. The first one is a modified version of a *quasi-Newton method* [73]. It uses only the values and numerical gradients of the objective function (in this case the variance) and builds up the inverse Hessian from updates along the optimization trajectory. Second algorithm is based on the general Levenberg-Marquardt (LM) method. It can be smoothly tuned between *Newton* and *steepest descent* approaches with a built-in stabilization parameter. However, it requires both gradient and Hessian and is therefore more costly to calculate. The gradient of the variance with respect to  $i$ -th parameter is readily obtained by differentiating Eq. (78) as

$$\sigma_i^2 = 2 \left[ \langle E_{L,i} (E_L - \bar{E}) \rangle + \left\langle \frac{\Psi_i}{\Psi} E_L^2 \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_L^2 \rangle - 2\bar{E} - \left\langle \frac{\Psi_i}{\Psi} (E_L - \bar{E}) \right\rangle \right], \quad (81)$$

where the subscript  $i$  denotes  $\frac{\partial}{\partial c_i}$ . Since the variance minimization method can be viewed as a fit of the local energy for a fixed set of configurations [56], by ignoring the change of the wave function an alternative expression for the same gradient follows as

$$\sigma_i^2 = 2 \langle E_{L,i} (E_L - \bar{E}) \rangle. \quad (82)$$

The Hessian derived from the gradient (82) is then given as

$$\sigma_{ij}^2 = 2 \langle (E_{L,i} - \bar{E})(E_{L,j} - \bar{E}) \rangle. \quad (83)$$

An important property of this approximate Hessian (83) is that it is symmetric in  $i$  and  $j$  and positive definite even for a finite set of configurations what is not automatically true when using the formally exact expression.

## 2.6.2 Energy Minimization

In general, a straightforward minimization of the energy in Eq. (77) can be quite unstable. This is somewhat counterintuitive and has its origin in the finite sample. The reason is that for a sufficiently flexible variational wave function it is possible to seemingly lower the energy for the fixed set of configurations, while, in fact, raising the true expectation value. As we have

mentioned earlier, the variance is bounded from below and therefore this problem is not present even for the finite sample [56]. Although one can frequently re-sample the wave function so that the method is always close to the actual minimum one can do better with some modifications.

We can improve the method convergence by employing the LM method and information about gradient and Hessian of the local energy mean. The corresponding gradient can be readily obtained from Eq. (77) as

$$\begin{aligned}\bar{E}_i &= \left\langle \frac{\Psi_i}{\Psi} E_L + \frac{H\Psi_i}{\Psi} - 2\bar{E} \frac{\Psi_i}{\Psi} \right\rangle \\ &= 2 \left\langle \frac{\Psi_i}{\Psi} (E_L - \bar{E}) \right\rangle\end{aligned}\quad (84)$$

where we used the fact that  $H$  is Hermitian. Note, that the expression in the last step of Eq. (84) has a favorable property of zero fluctuations as  $\Psi$  approaches the true eigenstate. By differentiating the Eq. (84) we get the Hessian as

$$\begin{aligned}\bar{E}_{ij} &= 2 \left[ \left\langle \left( \frac{\Psi_{ij}}{\Psi} + \frac{\Psi_i \Psi_j}{\Psi^2} \right) (E_L - \bar{E}) \right\rangle \right. \\ &\quad \left. - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \bar{E}_j - \left\langle \frac{\Psi_j}{\Psi} \right\rangle \bar{E}_i + \left\langle \frac{\Psi_i}{\Psi} E_{L,j} \right\rangle \right].\end{aligned}\quad (85)$$

It is clear that the above Hessian is not symmetric in  $i$  and  $j$  when approximated by a finite sample and the expression leads to quite large noise. Umrigar and Filippi [74] have recently demonstrated that the fully symmetric Hessian written entirely in the terms of covariances ( $\langle ab \rangle - \langle a \rangle \langle b \rangle$ ) has much smaller fluctuations than the Hessian from Eq. (85). The modified Hessian from Ref. [74] is given by

$$\begin{aligned}\bar{E}_{ij} &= 2 \left[ \left\langle \left( \frac{\Psi_{ij}}{\Psi} + \frac{\Psi_i \Psi_j}{\Psi^2} \right) (E_L - \bar{E}) \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \bar{E}_j - \left\langle \frac{\Psi_j}{\Psi} \right\rangle \bar{E}_i \right] \\ &\quad + \left[ \left\langle \frac{\Psi_i}{\Psi} E_{L,j} \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_{L,j} \rangle \right] + \left[ \left\langle \frac{\Psi_j}{\Psi} E_{L,i} \right\rangle - \left\langle \frac{\Psi_j}{\Psi} \right\rangle \langle E_{L,i} \rangle \right],\end{aligned}\quad (86)$$

where we have added three additional terms with vanishing expectation value (for proof, see e.g. Ref. [75]). The terms make the Hessian symmetric while at the same time decreasing the variance. Hence, addition of terms with zero expectations helps to cancel most of the finite sample fluctuations and makes the minimization method vastly more efficient.

Another useful rearrangement of the Hessian (86) is

$$\begin{aligned}\bar{E}_{ij} &= 2 \left[ \left\langle \left( \frac{\Psi_{ij}}{\Psi} - \frac{\Psi_i \Psi_j}{\Psi^2} \right) (E_L - \bar{E}) \right\rangle \right. \\ &\quad \left. + 2 \left\langle \left( \frac{\Psi_i}{\Psi} - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \right) \left( \frac{\Psi_j}{\Psi} - \left\langle \frac{\Psi_j}{\Psi} \right\rangle \right) (E_L - \bar{E}) \right\rangle \right] \\ &\quad + \left[ \left\langle \frac{\Psi_i}{\Psi} E_{L,j} \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_{L,j} \rangle \right] + \left[ \left\langle \frac{\Psi_j}{\Psi} E_{L,i} \right\rangle - \left\langle \frac{\Psi_j}{\Psi} \right\rangle \langle E_{L,i} \rangle \right].\end{aligned}\quad (87)$$

If the minimization procedure involves only the linear parameters in the exponential (i.e., of the form  $\exp[c(i)f]$ ) such as linear coefficients in the Jastrow factor, the two terms on the first line of Eq. (87) cancel out, eliminating thus the need for evaluation of the  $\Psi_{ij}$  term. For details on Levenberg-Marquardt method and other minimization see, for example, Ref. [76].

### 3 Nodal Properties of Fermionic Wave Functions

#### 3.1 Introduction

For a long time, nodes of fermionic wave functions and also related nodes of density matrices were considered uninteresting, since they are very complicated and their usefulness or connection to physical effects were unclear. For a given real state  $\Psi(\mathbf{R})$ , the node is defined as the subset of  $\mathbf{R}$ -space for which  $\Psi(\mathbf{R}) = 0$ . The node divides the configuration space into *nodal domains* with constant wave function sign, “+” or “-”. Interestingly, the first ideas on nodal properties of eigenstates go back to Hilbert and later to Courant. In particular, Courant proved the famous theorem that the  $n$ -th eigenstate of the 2D Laplacian on simply connected convex 2D regions with Dirichlet or Neumann boundary conditions have  $n$  or less nodal domains. Later this has been further expanded into higher dimensions and more complicated Hamiltonians. In the context of physics problems, Breit analyzed the nodes of two-electron states in 1930 and he actually found the first exact node for the atomic  $^3P(2p^2)$  state [77]. More recently, the behavior of nodes for highly excited (quasi-classical) states were studied and various bounds on nodal (Hausdorff) volumes for  $n \rightarrow \infty$  were established. As we have seen, the nodes play an important role in eliminating the fermion sign problem and knowledge of exact nodes would enable us to obtain fermionic ground states essentially exactly.

In relation to QMC methods, the nodes were perhaps for the first time encountered and identified as a “problem” in the paper by Anderson in 1976 [41]. The fermion nodes of small systems, mostly atoms, were investigated in several other papers [12, 40, 78–80]. The fermion nodes for degenerate and excited states were further studied by Foulkes and co-workers [81]. Bressanini, Reynolds and Ceperley demonstrated differences in the nodal surface topology between Hartree–Fock and correlated wave functions for the Be atom and explained thus the large impact of the  $2s, 2p$  near-degeneracy on the fixed-node DMC energy [82]. More recently, improvements in fixed-node DMC energies of small systems were studied by using CI expansions [83, 84], and pairing wave functions [1, 2, 85, 86].

Some of the general properties of fermion nodes were analyzed in an extensive study by Ceperley [87], which included proof of the tiling and connectivity properties as well as generalizations of the notion of nodes for density matrices. From few-electron systems as well as from numerical studies of free particle electron gas in 2D and 3D [87] came the conjecture that fermionic ground states have only two nodal cells (one “+”, one “-”). The paper laid down some of the ideas, however, explicit demonstrations were missing even for non-interacting systems. Explicit proofs of this conjecture for several important cases was found later [88, 89], when one of us used the property of connectivity to show analytically that a number of spin-polarized non-interacting and mean-field systems (non-interacting homogeneous electron gas, HF atomic states, harmonic fermions and fermions on a sphere surface) have ground state wave functions with *the minimal number of two nodal cells* [88, 89]. Furthermore, using the Bardeen–Cooper–Schrieffer wave function, it was shown that for in spin-unpolarized systems (i.e., with both spin channels occupied) an arbitrarily weak interaction reduces the four non-interacting nodal cells again to the minimal two. Finally, in the same paper was also demonstrated that the minimal number of nodal cells property extends to the temperature density matrices.

From the point of view of QMC methods, the two key issues are the *nodal topologies* and the *nodal shapes* and their efficient description. The topologies are important since they determine

how the configuration space domains are connected, property crucial for the correct sampling process. In addition, the topologies are also related to physical effects such as multiple particle exchanges and pairing. The shapes are important for the accuracy and so far our ability to describe these with sufficient flexibility is quite limited. Some recent advances on this aspect will be presented later in the review.

This section is organized as follows. In Sec. 3.2, we outline some of the general properties of fermion nodes. In Sec. 3.3, *exact* fermion nodes are described for few-electron spin-polarized atomic states, including the only known three-electron case discovered very recently. In Sec. 3.4, we categorize the nodal surfaces for the several half-filled sub-shells relevant for atomic and molecular states. In Sec. 3.5, we show how opposite spin correlations eliminate the redundant nodal structure of HF wave functions for cases of spin-unpolarized states.

### 3.2 General Properties

Let us consider a system of  $N$  spin-polarized fermions in  $d$  dimensions described by a real antisymmetric wave function  $\Psi(\mathbf{R})$ . We assume that  $\Psi(\mathbf{R})$  is an eigenstate of a Hamiltonian which, for simplicity, contains only local interactions. Consequently, there exists a subset of electron configurations called a fermion node, for which the wave function vanishes so that we can specify the nodal subset by an implicit equation

$$\Psi(\mathbf{R}) = 0. \quad (88)$$

Assuming that  $\Psi(\mathbf{R})$  describes the ground state, from the nodal subset we omit regions where the wave function might vanish because of other reasons (e.g., external potential, additional symmetries, etc). Note that when we talk about the nodes, we have in mind the nodes of many-body wave functions *in the space of all particles*. These nodes have nothing to do with familiar nodes of one-particle orbitals such as the angular nodes of spherical harmonics  $Y_{lm}$  or radial nodes of radial one-particle orbitals. In general, the many-particle state node is an  $(Nd - 1)$ -dimensional hypersurface defined by the implicit Eq. (88). Whenever positions of any two electrons coincide (i.e.,  $\mathbf{r}_i = \mathbf{r}_j$ ), the antisymmetry ensures that the  $\Psi(\mathbf{R}) = 0$ . For  $d > 1$ , these coincidences do not fully specify the nodes, but only define  $(Nd - d)$ -dimensional subspaces of *coincidence planes*. As we will see, the coincidence planes will play an important role for the nodes in 1D systems.

Another definition which we will need is the *nodal domain*. The node divides the configuration space into subsets in which the wave function sign is constant and which are called nodal domains (or cells). The nodal domain definition can be reformulated also in terms of *paths in the configuration space*. A nodal domain  $\Omega(\mathbf{R}_t)$  is a subset of configurations which can be reached from the point  $\mathbf{R}_t$  by a continuous path without crossing the node.

Let us now introduce the basic properties of wave function nodes.

- (a) *Manifold almost everywhere*—Uhlenbeck [90] proved that the node is generically a manifold, i.e., it can be locally mapped onto a  $(dN - 1)$ -dimensional plane except, possibly, for a subset of zero measure. This assumes reasonable Hamiltonians and a few other general conditions. Note that this *does not preclude* two or more hypersurfaces to cross since crossing is automatically lower dimensional, i.e., of zero measure.

- (b) *Nodal crossing angles*—If two nodal surfaces cross each other, they are orthogonal at the crossing. If  $n$  nodal surfaces cross each other, the crossing angles are all equal to  $\pi/n$ .
- (c) *Symmetry of the node*—The symmetry of the node cannot be lower than the symmetry of the state. Typically, these symmetries are the same, however, there are exceptions. There are known exact cases with the node symmetry being higher than the state symmetry.
- (d) *Tiling property for non-degenerate ground states*—The tiling property says that by applying all possible particle permutations to an arbitrary nodal cell of a ground state wave function one covers the complete configuration space  $\mathfrak{R}$  (i.e.,  $\sum_P \Omega(P\mathbf{R}_t) + \Omega(\mathbf{R}_t) = \mathfrak{R}$ ). Note that this does not specify how many nodal cells are there: the number of nodal domains can be any even number between two and  $N!$ . We will use this property later.

Let us analyze the nodes for several models. It is instructive to start with 1D systems for which the exact ground state nodes can be found explicitly and for arbitrary interaction. The reason is that in 1D the coincidence and the nodal hypersurfaces are identical due to geometric constraints of 1D space. Let us first consider  $N$  spin-polarized fermions in a 1D harmonic oscillator well. The wave function is given by the Slater determinant

$$\Psi_{1D}(1, \dots, N) = \det[\phi_k(\mathbf{r}_i)] = A \prod_i e^{-x_i^2/2} \det[1, 2x, \dots, H_{N-1}(x)], \quad (89)$$

where  $H_n(x)$  is a Hermite polynomial of degree  $n$  and  $A$  is the normalization. We omit the prefactors and transform the Slater matrix to monomials so that the wave function is given by the Vandermonde determinant

$$\Psi_{1D}(1, \dots, N) = \det[1, x, x^2, \dots, x^{N-1}] = \prod_{i < j} (x_j - x_i). \quad (90)$$

The node is encountered whenever two fermions pass through each other and the wave function has  $N!$  nodal cells, since any permutation requires at least one node-crossing. In general, the derived node is exact for other 1D models including systems with arbitrary *local* interactions (which are not too singular). For periodic boundary conditions one can find a similar expression. Let us consider  $N = (2k_F + 1)$  particles in a 1D periodic box  $(-\pi, \pi)$ , so that the coordinates are dimensionless. In this setting the Fermi momentum becomes an integer,  $k_F = 1, 2, \dots$ . The one-particle occupied states are written as  $\phi_n(x) = e^{inx}$ ,  $n = 0, \pm 1, \dots, \pm k_F$  and the spin-polarized ground state is given by

$$\Psi_{1D}(1, \dots, N) = \det[\phi_n(x_j)], \quad (91)$$

where  $x_j$  is the  $j$ -th particle coordinate and  $j = 1, \dots, N$ . The term  $\exp(-ik_F \sum_j x_j)$  can be factorized out of the determinant so that the Slater matrix elements become powers of  $z_j = e^{ix_j}$ . This leads to the Vandermonde determinant and after some rearrangements we find

$$\Psi_{1D}(1, \dots, N) = e^{-ik_F \sum_j x_j} \prod_{j > k} (z_j - z_k) = A \prod_{j > k} \sin(x_{jk}/2) \quad (92)$$

where  $x_{jk} = x_j - x_k$  and  $A$  is an unimportant constant prefactor. The periodic boundary conditions result in the same 1D nodal hypersurface which is also manifestly periodic, as expected. Again, for 1D periodic ground states these nodes are exact regardless of interactions.

Properties of nodes for  $d > 1$  are very different and also far more difficult to analyze than in one dimension. The key difference is that the node dimensionality is higher than the dimensionality of the coincidence hyperplanes. Clearly, for  $d > 1$  particles have enough space to “avoid each other” and that has implications on the nodal topologies and the number of nodal cells. In addition, this spatial or dimensional flexibility also means that the nodal hypersurfaces can be and are affected by interactions.

We first consider spin-polarized non-interacting fermions in a 2D harmonic well. The one-particle states are simply  $\phi_{nm} = C_{nm}H_n(x)H_m(y)$ ,  $n, m = 0, 1, \dots$  where  $C_{nm}$  includes the Gaussian and normalization which are absorbed into a common prefactor and omitted. The Slater matrix elements can be rearranged to monomials as follows

$$\Psi_{2D}(1, \dots, N) = \det[1, x, y, \dots, x^n y^m, \dots]. \quad (93)$$

The closed-shell states and the system size are conveniently labeled by  $M = 1, 2, \dots$  where  $n + m \leq M$ , with the corresponding number of fermions given by  $N = (M + 1)(M + 2)/2$ .

Let us illustrate the qualitative difference in nodal topologies between higher dimensions and 1D systems on the following example of three harmonic particles. For  $M = 1$ , the three-particle wave function is given by  $\Psi_{2D}(1, 2, 3) = \det[1, x, y]$ . In order to understand the nodal domain topology, it is convenient to extend the 2D coordinates by a “dummy” third dimension as  $\mathbf{r}_i = (x_i, y_i, 0)$ . Then  $\Psi_{2D}(1, 2, 3) = \mathbf{z}_0 \cdot (\mathbf{r}_{21} \times \mathbf{r}_{31})$  where  $\mathbf{z}_0$  is the unit vector in the third dimension and  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ . Note that while in 1D three particle state exhibits  $3! = 6$  nodal cells, the result we just derived implies something very different. There are only *two nodal domains* since the set of vectors  $\mathbf{z}_0, \mathbf{r}_{21}, \mathbf{r}_{31}$  is either left- or right-handed, and the node is encountered whenever the three particles become collinear, i.e.,  $\mathbf{r}_{21} \times \mathbf{r}_{31} = 0$ . Another important property which will be explored in detail below, is the fact that it is possible to carry out a triple exchange  $1 \rightarrow 2, 2 \rightarrow 3, 3 \rightarrow 1$  without node-crossing (e.g., rotate an equilateral triangle).

The triple exchanges are important for proving the two-nodal cell property in general cases for arbitrary number of particles. In order to explain this aspect, we introduce a notion of *particles connected by triple exchanges*. The three particles  $i, j, k$  are called *connected* if there exist a cyclic exchange path  $i \rightarrow j, j \rightarrow k, k \rightarrow i$ , which does not cross the node. Let us now suppose that there exist a point  $\mathbf{R}_t$  inside a nodal domain such that all particles are connected into a single cluster of triple exchanges (see Fig. 5). Then  $\Psi(\mathbf{R})$  has only two nodal cells. In other words, the whole configuration space is covered by only one positive and one negative nodal cell (i.e., the nodal cells are *maximal*). Let us try to understand why this is the case. First, any triple exchange is equal to two pair exchanges. The existence of the point  $\mathbf{R}_t$  with all particles connected then means that all positive (even) permutations  $P_+ \mathbf{R}_t$  can be carried out within the same nodal cell. Second, the tiling property implies that once all particles are connected for  $\mathbf{R}_t$  the same is true for entire cell  $\Omega(\mathbf{R}_t)$ . Therefore, there will be only one maximal cell per each sign. More details on this construction can be found in Ref. [87].

We are now ready to show that six particles ( $M = 2$ ) in the harmonic 2D well also have the ground state with the minimal nodal count of two. We position particles 1, 2 and 3 so that they have common  $y$ -coordinate,  $y_0$ . For  $M = 2$  the wave function reads

$$\Psi_{2D}(1, \dots, 6) = \det[1, x, y, x^2, xy, y^2] =$$

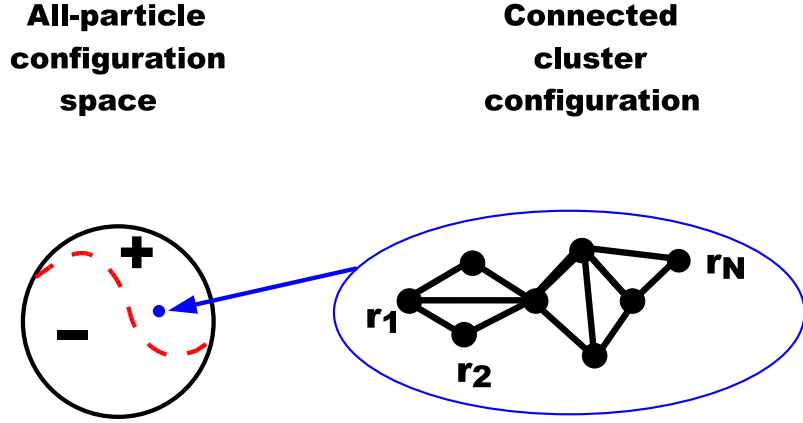


Fig. 5. Sketch of a connected cluster of triple exchanges. Each triangle means that there is a triple exchange path which does not cross the node.

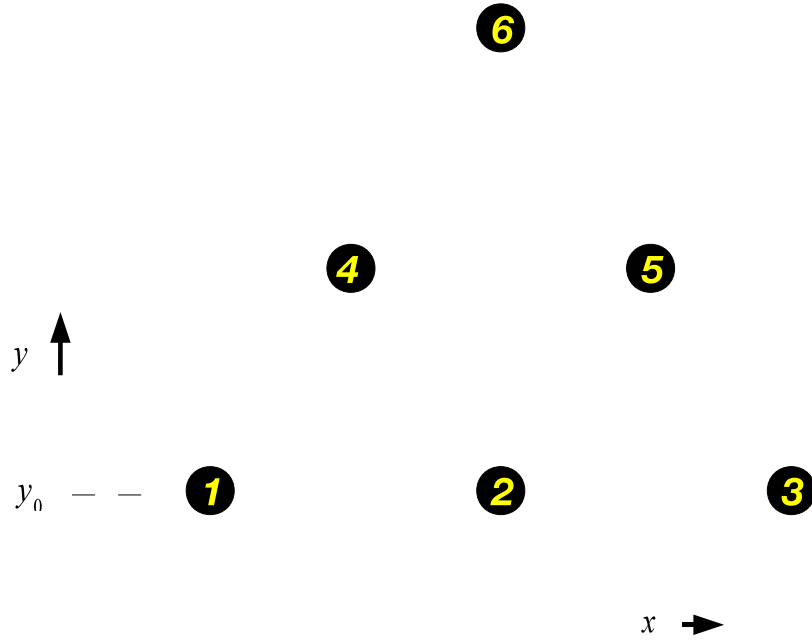


Fig. 6. Positions of particles for proving the connectivity of six 2D harmonic fermions.

$$= \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ y_0 & y_0 & y_0 & y_4 & y_5 & y_6 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 & x_5^2 & x_6^2 \\ x_1 y_0 & x_2 y_0 & x_3 y_0 & x_4 y_4 & x_5 y_5 & x_6 y_6 \\ y_0^2 & y_0^2 & y_0^2 & y_4^2 & y_5^2 & y_6^2 \end{vmatrix}, \quad (94)$$



where the particle configuration which will enable to figure out the connectivity is given in Fig.6. All elements containing  $y_0$  can be eliminated by appropriate rearrangements and the determinant factorizes as

$$\Psi_{2D}(1, \dots, 6) = \prod_{i=4}^6 (y_i - y_0) \Psi_{1D}(1, 2, 3) \Psi_{2D}(4, 5, 6). \quad (95)$$

It is now clear that we can rotate the triangle of particles 4, 5, 6 without crossing the node since  $\Psi_{2D}(4, 5, 6)$  remains constant and the wave function does not change the sign. That implies that particles 4, 5, 6 are connected. However, the wave function is rotationally invariant so the particles 1, 2, 4 must be connected, as well as the particles 2, 3, 5. Therefore all six particles are connected and the wave function has only two nodal cells.

With some effort it is possible to generalize these types of constructions and show that harmonic particles are connected for any closed-shell state with arbitrary  $M \rightarrow \infty$  and for arbitrary  $d$  as given elsewhere [88].

It seems that the harmonic oscillator might be somewhat special due to the simple structure of the wave function, since the nodes are zeros of antisymmetric polynomials [89]. In fact, *the minimal two-nodal cells property appears to be generic*. Let us present an example which is more realistic and does not reduce to simple polynomials. Consider an atomic spin-polarized state  ${}^6S(1s2s3p^3)$ , i.e., 5 electrons in the Coulomb  $-Z/r$  potential for both non-interacting and HF wave-functions. The wave function can be written as

$$\Psi_{at}(1, \dots, 5) = \det[\rho_{1s}^*(r), \rho_{2s}^*(r), x, y, z], \quad (96)$$

where  $\rho_{1s}^*(r) = \rho_{1s}(r)/\rho_{2p}(r)$  and  $\rho_{2s}^*(r) = \rho_{2s}(r)/\rho_{2p}(r)$  since non-negative  $\rho_{2p}(r)$  is factorized out. The coordinates become dimensionless by rescaling with the atomic number  $Z$  as  $\mathbf{r} \leftarrow Z\mathbf{r}$ . Let us position particle 1 at the origin and particles 2 to 5 on the surface of a sphere with the radius  $\eta_0$  equal to the radial node of  $\rho_{2s}(r)$  orbital, i.e.,  $\rho_{2s}(\eta_0) = 0$ . For such configurations we obtain

$$\Psi_{at}(1, \dots, 5) = \rho_{1s}^*(\eta_0) \rho_{2s}^*(0) \mathbf{r}_{32} \cdot (\mathbf{r}_{42} \times \mathbf{r}_{52}) \quad (97)$$

Three-particle exchanges 2, 3, 4, 5 easily avoid node-crossings by appropriate positioning and rotations on the sphere, e.g., using positions of the tetrahedron vertices. We can also show that particle 1 is connected by the exchange  $123 \rightarrow 312$  which is parameterized as  $\mathbf{r}_1(t) = \eta_0[t, 0, 0]$ ;  $\mathbf{r}_2(t) = \eta_0[c(t), s(t), 0]$  and  $\mathbf{r}_3(t) = \eta_0[0, 1-t, 0]$  where  $t = 0$  ( $t = 1$ ) corresponds to the beginning (end) point of the exchange path while  $c(t) = \cos(\pi t/2)$  and  $s(t) = \sin(\pi t/2)$ . Setting  $\mathbf{r}_4 = [0, 0, \eta_0]$  and  $\mathbf{r}_5 = [0, 0, -\eta_0]$  we find that  $\Psi_{at}$  is proportional to

$$\rho_{2s}^*(t\eta_0)c(t)(1-t) + \rho_{2s}^*[(1-t)\eta_0]s(t)t > 0 \quad (98)$$

The inequality holds for the whole path  $0 \leq t \leq 1$  since  $\rho_{2s}^*(t\eta_0) > 0$  for  $0 \leq t < 1$  for both non-interacting and HF cases. Note that the demonstration is non-trivial since the HF wave function reflects the mean-field interactions and the orbitals can be found only by a numerical solution. Nevertheless, qualitative features of the HF and non-interacting orbitals (symmetries, structure of radial nodes, etc) are basically identical and enable to verify the nodal cell count for

both cases. Note also that the wave functions is not a multi-variate polynomial like for the case of harmonic potentials. The proof can be further extended to more shells with more electrons such as  $1s2s2p^33s3p^3$  and  $1s2s2p^33s3p^33d^5$  [89].

Similar two-nodal cell demonstrations can be carried out for many paradigmatic quantum models such as 2D and 3D homogeneous fermion gas, fermions on a sphere, etc [89]. In addition, the same can be shown also for the temperature density matrices, for which one can define the nodes in similar way as for the wave functions.

### 3.3 Exact Nodal Surfaces

We assume the usual electron-ion Hamiltonian and we first investigate a few-electron ions focusing on fermion nodes for sub-shells of one-particle states with  $s, p, d, \dots$  symmetries using variable transformations, symmetry operations and explicit expressions for the nodes.

#### 3.3.1 Two-Electron Atomic $^3P(p^2)$ , $^3S(1s2s)$ and Molecular $^3\Sigma_g(\pi^2)$ Triplet States

Apparently, the exact node of the atomic  $^3P(p^2)$  state was derived in a different context by Breit in 1930 [77, 83, 91]. Here we offer an independent proof which enables us to apply the analysis to molecular states with the same symmetries. The state is the lowest within its symmetry class. It has even parity and cylindric symmetry, say, around  $z$ -axis. It is also odd under rotation by  $\pi$  around any axis orthogonal to the cylindric axis. We denote the rotation by  $\pi$  around, say, the  $x$ -axis, as  $R(\pi x)$ . In order to uncover the *exact* node it is convenient to define new coordinates. Let us denote  $\mathbf{r}_{12}^+ = \mathbf{r}_1 + \mathbf{r}_2$ ,  $r_{12}^+ = |\mathbf{r}_{12}^+|$ , together with the customary  $\mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2$ ,  $r_{12} = |\mathbf{r}_{12}|$ . We can now introduce the following coordinate map

$$(\mathbf{r}_1, \mathbf{r}_2) \rightarrow (r_{12}^+, r_{12}, \cos \omega, \cos \beta, \varphi, \varphi'), \quad (99)$$

where  $\cos \omega = \mathbf{z}_0 \cdot (\mathbf{r}_1 \times \mathbf{r}_2) / |\mathbf{r}_1 \times \mathbf{r}_2|$  with  $\mathbf{z}_0$  being the unit vector along the  $z$ -axis,  $\cos \beta = \mathbf{r}_{12}^+ \cdot \mathbf{r}_{12} / (r_{12}^+ r_{12})$ , and  $\varphi'$  being the azimuthal angle of  $\mathbf{r}_1 \times \mathbf{r}_2$ . Further,  $\varphi$  is the azimuthal angle of  $\mathbf{r}_{12}^+$  in the relative coordinate system with the  $x$ -axis unit vector given by a projection of  $\mathbf{z}_0$  into the plane defined by  $\mathbf{r}_1, \mathbf{r}_2$ , i.e.,  $\mathbf{e}_x = \mathbf{z}_{0p} / |\mathbf{z}_{0p}|$ ,  $\mathbf{e}_z = (\mathbf{r}_1 \times \mathbf{r}_2) / |\mathbf{r}_1 \times \mathbf{r}_2|$  and  $\mathbf{e}_y = \mathbf{e}_z \times \mathbf{e}_x$ . The coordinate  $\varphi'$  can be omitted from further considerations right away due to the cylindric symmetry so that the wave-function dependence simplifies to  $\Psi(r_{12}^+, r_{12}, \cos \beta, \cos \omega, \varphi)$ . We introduce an operator  $Q$  which acts on coordinates and is given by  $Q = P_I P_{12} R(\pi x)$  where  $P_I$  inverts around the origin and  $P_{12}$  is the pair exchange of particles 1 and 2. Considering the state symmetries, action of  $Q$  reveals that the wave function is invariant in the simultaneous change  $(\cos \beta, \varphi) \rightarrow (-\cos \beta, -\varphi)$ . Combining action of  $Q$  and  $P_{12}$  we find that

$$\Psi(\dots, -\cos \omega, \dots) = -\Psi(\dots, \cos \omega, \dots) \quad (100)$$

with the rest of variables unchanged. The node is therefore given by  $\cos \omega = 0$ , i.e.,  $\mathbf{z}_0 \cdot (\mathbf{r}_1 \times \mathbf{r}_2) = 0$ , and it is encountered when both electrons happen to be on a plane which contains the  $z$ -axis. This exact node therefore agrees with the node of non-interacting or HF wave function  $\Psi = \det[\rho(r)x, \rho(r)y]$ .

The derived node equation is applicable to any cylindric potential with  $D_{\infty h}$  symmetry, e.g., equidistant homo-nuclear dimer, trimer, etc, with one-particle orbitals  $\pi_x, \pi_y$  which couple into the triplet state  $^3\Sigma_g(\pi_x \pi_y)$ .

Note that the parametrization given above leads to the exact node for the lowest two-electron triplet state,  $^3S(1s2s)$  which has been known for some time [12, 78]. The spherical symmetry makes angles  $\omega$  and  $\varphi$  irrelevant and simplifies the wave function dependence to distances  $r_1, r_2, r_{12}$  or, alternatively, to  $r_{12}, r_{12}^+, \cos \beta$ . Applying  $P_{12}$  leads to

$$-\Psi(r_{12}, r_{12}^+, \cos \beta) = \Psi(r_{12}, r_{12}^+, -\cos \beta), \quad (101)$$

so that the node is given by the condition  $\cos \beta = 0$ , i.e.,  $r_1 - r_2 = 0$ , or, equivalently,  $r_1^2 = r_2^2$ . The last equation clearly shows that the node is generically a quadratic manifold, namely, a 5D hyperboloid in 6D space (one recalls that in 3D space the 2D hyperboloid is given as  $z^2 \pm \text{const} = x^2 - y^2$ ). Again, the nodal domain count is two and the domains correspond to  $r_1 > r_2$  or  $r_1 < r_2$ .

For the presented states the exact nodes are determined solely by symmetry and are independent of interactions. That further implies that all the excited states with the same symmetries will contain the ground state node although they will exhibit also additional nodal hypersurfaces. We will come back to this point in the following subsection.

### 3.3.2 Three-Electron Quartet $^4S(2p^3)$ State

Recently, we have found another case of exact node for a three-electron atomic state, namely, the lowest spin quartet of  $S$  symmetry and odd parity,  $^4S(2p^3)$ . It is actually very easy to find the node by transforming it to the previously analyzed cylindric symmetry. We can always rotate the system so that one of the electrons, say, the electron 3, is positioned on the  $z$ -axis so that the remaining symmetry is only cylindric. However, then the wave function for the remaining two electrons is identical to the previous  $^3P(2p^2)$  state and we can directly write the nodal condition as  $\mathbf{r}_3 \cdot (\mathbf{r}_1 \times \mathbf{r}_2) = 0$ . The wave function vanishes when all three electrons appear on a plane which contains the origin.

One can demonstrate the same by defining an alternative coordinate map  $\cos \alpha = \mathbf{r}_3 \cdot (\mathbf{r}_1 \times \mathbf{r}_2) / r_3 |\mathbf{r}_1 \times \mathbf{r}_2|$ ,  $\cos \beta = [(\mathbf{r}_1 \times \mathbf{r}_2) \times \mathbf{r}_{12}^+] \cdot \mathbf{r}_{12} / [|\mathbf{r}_1 \times \mathbf{r}_2| |\mathbf{r}_{12}^+|]$  and  $\gamma$  being the azimuthal angle of  $\mathbf{r}_3$  in the relative coordinate system given by the following unit vectors,  $\mathbf{e}_x = \mathbf{r}_{12}^+ / r_{12}^+$ ,  $\mathbf{e}_z = [(\mathbf{r}_1 \times \mathbf{r}_2) \times \mathbf{r}_{12}^+] / [|\mathbf{r}_1 \times \mathbf{r}_2| |\mathbf{r}_{12}^+|]$  and  $\mathbf{e}_y = \mathbf{e}_z \times \mathbf{e}_x$ . Besides these relative coordinates, the three Euler angles fix the orientation in the original coordinate system, however, the  $S$  symmetry makes them irrelevant. In the new coordinates the node comes out by the action of  $P_{12}$  on  $\Psi(r_{12}^+, r_{12}, r_3, \cos \alpha, \cos \beta, \gamma)$  as

$$\Psi(\dots, -\cos \alpha, \cos \beta, \gamma) = -\Psi(\dots, \cos \alpha, \cos \beta, \gamma). \quad (102)$$

It is clear that  $P_{23}$  and  $P_{13}$  would lead to the same nodal condition, assuming one would redefine the relative coordinate map accordingly. The node of non-interacting and HF wave function  $\Psi_{HF} = \det[\rho(r)x, \rho(r)y, \rho(r)z]$  again agrees with the exact condition. As mentioned before, the combination of symmetries is so restrictive that interactions leave the node invariant and this is in fact the case also for the excited states. Therefore *all* the excited states must contain the ground state fermionic node. This is illustrated in Fig. 7 which shows the ground state (planar) node together with additional nodes, as expected for excitations.

The excitations have therefore special symmetry structure which enables to show that the one-particle  $s$  channel is absent from the whole spectrum in this symmetry sector. One can

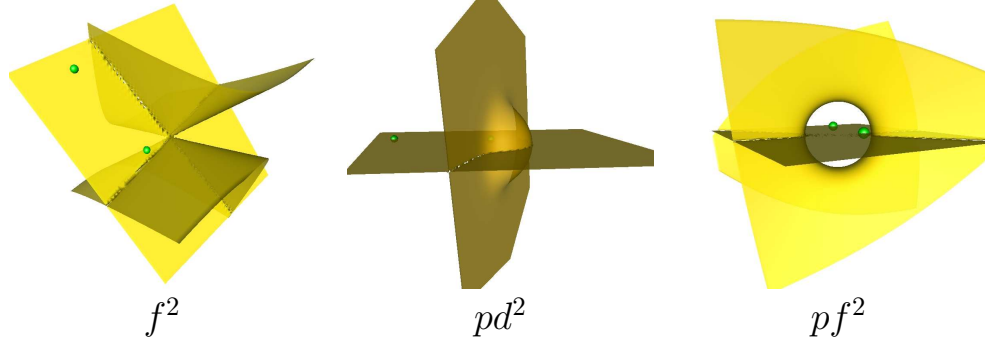


Fig. 7. The 3D projected nodes of a few selected excitations for the symmetry adopted CI expansion of the  $^4S(p^3)$  ground state. The exact planar node of the quartet ground state is also possessed by all the excitations. The small spheres show the fixed positions of two electrons while the third one is scanning the nodal surface. Labels indicate the types of excitations.

demonstrate this as follows. We expand any state of this symmetry in linear combinations of excited determinants

$$|^4S(p^3)\rangle = \sum_{i,j,k} c_{ijk} \det[|n_i l_i m_i\rangle |n_j l_j m_j\rangle |n_k l_k m_k\rangle], \quad (103)$$

where the sum is over the complete one-particle spectrum (including scattering states). Due to the selection rules coefficients  $c_{ijk}$  vanish for any excitation which contains one or more  $l = 0$  orbitals. Consider first states with  $l_i = 0$ . That implies that all determinants in this configuration must have  $l_j = l_k$  otherwise the Clebsch–Gordan coefficients for  $S$  symmetry vanish. However, then the excitation would be of even parity and therefore such configurations drop off. The only remaining possibility is that  $l_i = l_j = 0$ . Such configurations then require that also  $l_k = 0$ , however, such states with odd parity do not exist. Therefore the one-particle channel with  $s$  symmetry is “switched-off” completely. This property was useful for showing that the fixed-node calculation of this state provided the exact energy even when using nonlocal pseudopotential with  $s$  non-locality [92].

In order to illustrate how exact and approximate nodal surfaces affect the fixed-node DMC results we studied the valence-electron states of the spin-polarized nitrogen cation, some of which were analyzed above. The core electrons were eliminated by pseudopotentials [93]. The trial wave function was of the commonly used form with single HF determinant times the Jastrow correlation factor [5]. As we explained above, the pseudopotential nonlocal  $s$ –channel does not couple to either the odd parity  $S$  state or the even parity  $P(p^2)$  state so that the nonlocal contribution to these states vanishes exactly.

In order to compare the fixed-node QMC calculations with an independent method we have carried out also configuration interaction calculations with ccpV6Z basis [94] (with up to three  $g$  basis functions), which generates more than 100 virtual orbitals in total. In the CI method the wave function is expanded in excited determinants and we have included all single, double and triple excitations. Since the doubles and triples include two- and three-particle correlations

Tab. 1. Total energies (in Hartrees) of  $N^+$ ,  $N^{+2}$  and  $N^{+3}$  ions with core electrons eliminated by pseudopotentials [93]. The energies are calculated by HF, configuration interaction (CI), variational (VMC) and fixed-node diffusion Monte Carlo (DMC) methods.

State	HF	CI	VMC	DMC
$^3P(p^2)$	-5.58528	-5.59491	-5.59491(2)	-5.59496(3)
$^4S(p^3)$	-7.24716	-7.27566	-7.27577(1)	-7.27583(2)
$^5S(sp^3)$	-8.98570	-9.02027	-9.01819(4)	-9.01962(5)

exactly, the accuracy of the CI results is limited only by the size of the basis set. By comparison with other two- and three-electron CI calculations we estimate that the order of magnitude of the basis set CI bias is  $\approx 0.01$  mH for two electrons and  $\approx 0.1$  mH. and for three electrons (despite the large number of virtuals the CI expansion converges relatively slowly [95] in the maximum angular momentum of the basis functions, in our case  $l_{max} = 4$ ). The pseudopotentials we used were identical in both QMC and CI calculations.

The first two rows of Tab. 1 show the total energies of variational and fixed-node DMC calculations with the trial wave functions with HF nodes together with results from the CI calculations. For  $^3P(p^2)$  the energies agree within a few hundredths of mH with the CI energy being slightly higher but within two standard deviations from the fixed-node QMC result. For  $^4S(p^3)$  the CI energy is clearly above the fixed-node DMC by about 0.17 mH which is the effect of the finite basis. In order to illustrate the fixed-node bias in the case when the HF node is *not* exact we have also included calculations for four electron state  $^5S(sp^3)$  (for further discussion of this Hartree–Fock node see Sec 3.4.2 below). For this case, we estimate that the CI energy is above the exact value by  $\approx 0.3$  mH so that the fixed-node energy is significantly *higher* than both the exact and CI energies. Using these results we estimate that the fixed-node error is  $\approx 1$  mH, i.e., almost 3% of the correlation energy.

### 3.4 Approximate Hartree–Fock Nodes

It is quite instructive to investigate the nodes of half-filled sub-shells of one-particle states with higher angular momentum. We will explore the nodal hypersurface projections into the 3D space to understand what an individual electron “sees” during the imaginary time evolution. The study also helps to classify nodes for various cases of atomic and molecular systems. For the next few subsections we will closely follow the derivations from our paper [92].

#### 3.4.1 Hartree–Fock Node of $^6S(d^5)$ State

The HF determinant wave function for  $^6S(d^5)$  is given by

$$\Psi_{HF} = \prod_{i=1}^5 \rho(r_i) \det[2z^2 - x^2 - y^2, x^2 - y^2, xz, yz, xy], \quad (104)$$

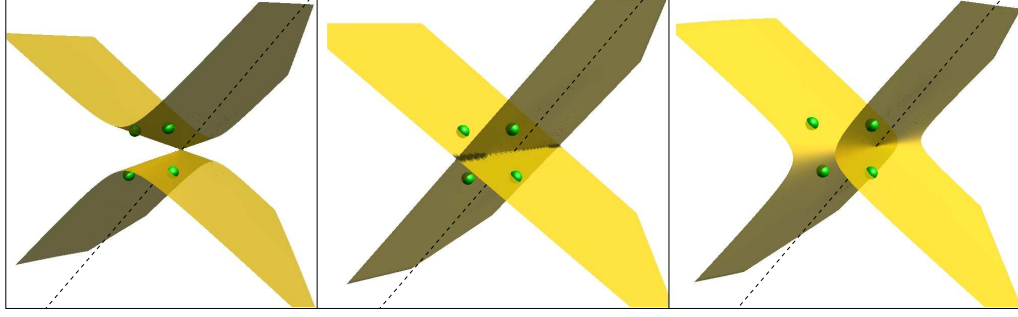


Fig. 8. The 3D projected Hartree-Fock node of  ${}^6S(d^5)$  state, which is an elliptic cone (left and right pictures). The picture in the center shows a configuration of two pairs of electrons lying on two orthogonal planes which contain the origin. This particular node is of lower dimension because of the additional constraint on positions of the electrons. It appears as a crossover between the elliptic cones with different orientation (left and right pictures). The small spheres show the positions of the four electrons while the line denotes the  $z$ -axis.

where  $\rho(r_i)$  is the radial part of the  $3d$ -orbital. Since the radial prefactor is irrelevant it can be omitted. The  $S$  symmetry allows to rotate the system so that, say, electron 1 is on the  $z$ -axis, and then the corresponding column in the Slater matrix becomes  $(2z_1^2, 0, 0, 0, 0)$ . Assuming that  $z_1 \neq 0$  we can then write the nodal condition as

$$\det[x^2 - y^2, xz, yz, xy] = 0. \quad (105)$$

Using one of the electrons as a *probe* (i.e., looking at the node from the perspective of one of the electrons) we can find the projection of the node to the 3D space. We denote the probe electron coordinates simply as  $(x, y, z)$  and expand the determinant so that we get

$$(x^2 - y^2)m_1 + xzm_2 + yzm_3 + xym_4 = 0 \quad (106)$$

where  $m_i$  are the corresponding cofactors. We divide out  $m_1$  assuming that it is non-zero (not a crucial assumption as clarified below). We get

$$(x^2 - y^2) + axz + byz + cxy = 0, \quad (107)$$

where  $a = m_2/m_1$ ,  $b = m_3/m_1$ ,  $c = m_4/m_1$ . By completing the square this can be further rearranged to

$$(x - k_1y)(x - k_2y) + z(ax + by) = 0 \quad (108)$$

with  $k_{1,2} = (-c \pm \sqrt{c^2 + 4})/2$ . Let us define rotated and rescaled coordinates

$$u^* = -(ak_2 - b)(x - k_1y)/(k_1 - k_2) \quad (109)$$

$$v^* = (ak_1 - b)(x - k_2y)/(k_1 - k_2) \quad (110)$$

$$w^* = z[(ak_1 - b)(ak_2 - b)]/(k_1 - k_2)^2 \quad (111)$$

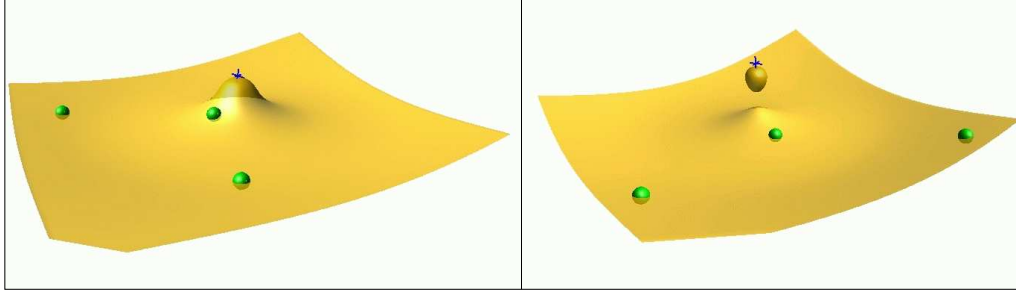


Fig. 9. The 3D projection of the nitrogen cation  ${}^5S(sp^3)$  Hartree-Fock node (the core electrons are eliminated by pseudopotentials). The projected node exhibits two topologies. It is either a planar surface deformed by the radial orbital functions at the nucleus or, in certain configurations, the deformation forms a small bubble detached from the surface (the picture on the right). The small cross is the location of the ion while the small spheres denote positions of electrons.

so we can write the Eq. (107) as

$$u^*v^* + w^*u^* + w^*v^* = 0. \quad (112)$$

Note that this equation has a form which is identical to Eq. (106) with  $m_1 = 0$  so this representation is correct for any  $m_1$ . After some effort one finds that Eq. (112) is a cone equation (i.e.,  $d_{z^2}$  orbital) as can be easily verified by using the following identity

$$(2u^2 - v^2 - w^2)/8 = u^*v^* + w^*u^* + w^*v^*, \quad (113)$$

where  $u = u^* + v^* + 2w^*$ ,  $v = (-u^* + v^* + 2w^*)$ ,  $w = (u^* - v^* + 2w^*)$ . The 3D projected node is therefore an elliptic cone. Note that its equations is the same as rotated and rescaled angular part of the  $d_{z^2}$  orbital,  $2z^2 - x^2 - y^2$ .

Therefore the 3D projection of the  $d^5$  HF node is a family of cones given by the homogeneous second-order polynomial in three variables with coefficients determined by the positions of the four electrons. From the theory of quadrics [96], one finds that a general elliptic cone can possibly “fit” up to five 3D points/electrons. However, in our case there is an additional constraint of one electron being fixed on the  $z$ -axis. That implies that the  $z$ -axis lies on the cone so that the cone always cuts the  $xy$  plane in two lines, which are orthogonal to each other. The orthogonality can be verified by imposing  $z = 0$  in Eq. (108) and checking that  $k_1k_2 = -1$ . In addition, one can find “degenerate” configurations with two pairs of two electrons lying on orthogonal planes (Fig. 8). This corresponds to the “opening” of the cone with one of the elliptic radii becoming infinite and the resulting node having a form of two orthogonal planes (Fig. 8). The condition is equivalent to  $A_{44} = b^2 - a^2 - abc = 0$ , where  $A_{44}$  is one of the quadratic invariants [96]. There are more special cases of nodes with additional restrictions: (a) when two electrons lie on a straight line going through the origin; (b) when three electrons lie on a plane going through the origin; (c) when four electrons lie on a single plane.

### 3.4.2 Hartree–Fock Nodes of the $^5S(sp^3)$ Ion

The HF node for this spin-polarized state with  $1s^2$  (He-core) states replaced by pseudopotentials can be studied in a similar way as studied previously [92]. We can expand the determinant as in the previous cases with a new feature that the radial parts will modify the shapes of the hypersurfaces. The system therefore corresponds, e.g., to the  $N^+$  ion from Tab. 1. Note that due to pseudized core the one-particle  $s$  orbital has no radial node although it corresponds to the physical  $2s$  state. By expanding the determinant in the column of the probe electron with position  $x, y, z$  the 3D node projection is simply given by

$$x + b'y + c'z + d'\eta(r) = 0, \quad (114)$$

where  $b', c', d'$  depend on ratios of cofactors and  $\eta(r) = \rho_s(r)/\rho_p(r)$  is the ratio of radial parts of  $s$  and  $p$  orbitals. The probe electron will see a plane with a approximately bell-shape deformation in the area of the nucleus (see Fig. 9). The shape of deformation depends on the ratio of  $s$  and  $p$  radial parts and the magnitudes and signs of the cofactors. For certain configurations the deformation in a detached, separate ellipsoid-like bubble. The bubble is caused by the radial dependence of  $\eta(r)$  which for pseudized core is not a monotonic function and therefore can create new topologies. Note that despite the fact that the 3D projection shows a separated region of space (the bubble) the complete node which is 11-dimensional has again the minimal number of two nodal cells. This can be checked by placing the four particles at the vertices of a regular tetrahedron with the center at the ionic center. The wave function does not vanish for such configuration as can be easily verified. Rotations by  $\pi/3$  around of each of the four three-fold symmetry axes of the  $T_d$  group then shows that all the particles are connected.

### 3.4.3 Hartree–Fock Nodes of Spin-Polarized $p^3d^5$ and $sp^3d^5$ Shells with $S$ Symmetry

Let us for a moment assume a model wave function in which the radial parts of  $s, p, d$  orbitals are identical. Then, using the arrangements similar to  $d^5$  case, we can expand the determinant of  $p^3d^5$  in one column and for the 3D node projection we then get

$$2u^2 - v^2 - w^2 + \alpha u + \beta v + \gamma w = 0, \quad (115)$$

where  $u, v, w$  are appropriate linear combinations of  $x, y, z$ . This can be further rewritten as

$$2(u + \alpha/4)^2 - (v - \beta/2)^2 - (w - \gamma/2)^2 + \delta_0 = 0, \quad (116)$$

where  $\delta_0 = (-\alpha^2/2 + \beta^2 + \gamma^2)/4$ . It is clear that the quadratic surface is offset from the origin (nucleus) by a vector normal to  $\alpha u + \beta v + \gamma w = 0$  plane. Using the properties of quadratic surfaces one finds that for  $(\alpha^2/(\alpha^2 + \beta^2 + \gamma^2)) < 2/3$  the node is a single-sheet hyperboloid with the radius  $\sqrt{\delta_0}$ ; otherwise it has a shape of a double-sheet hyperboloid. The double-sheet hyperboloid forms when there is an electron located close to the origin. A special case is a cone which corresponds to  $(\delta_0 = 0)$ . The case of  $sp^3d^5$  is similar, but with different  $\delta_0$ , which now has a contribution from the  $s$ -orbital (see Fig. 10). Once we include also the correct radial parts of orbitals in the  $s, p, d$  channels the coefficients of the quadratic form depend on both cofactors and orbital radial functions. The resulting nodal surface is deformed beyond an ideal quadric and shows some more complicated structure around the nucleus (see Fig. 11) as illustrated on HF nodes of the majority spin electrons in  $Mn^{+2}$  ion (note that the Ne-core electrons were eliminated by pseudopotentials).



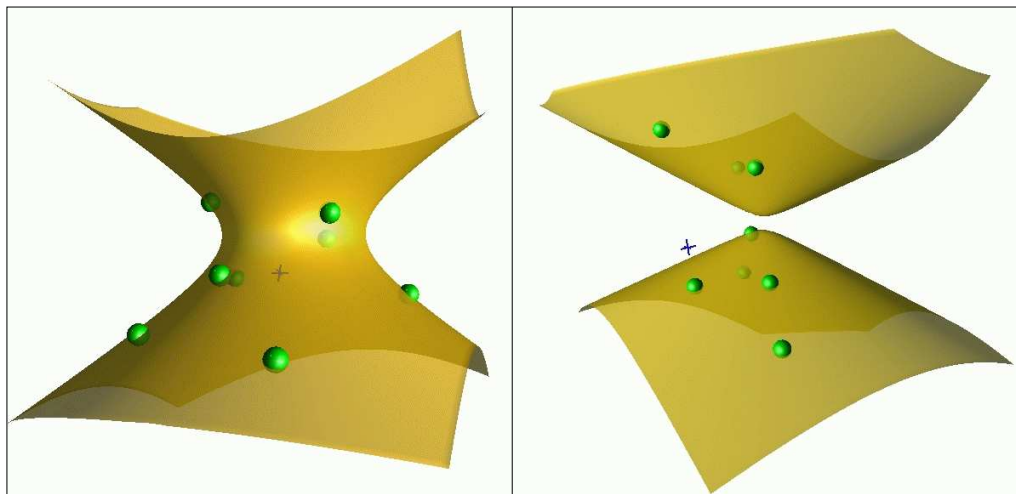


Fig. 10. The 3D projection of the angular part of the  $^{10}S(sp^3d^5)$  state Hartree-Fock node (with radial parts of orbitals identical for all  $spd$  orbitals). The projection has a topology of a single-sheet or double-sheet hyperboloid. The small cross shows the location of the nucleus while the spheres illustrate the electron positions.

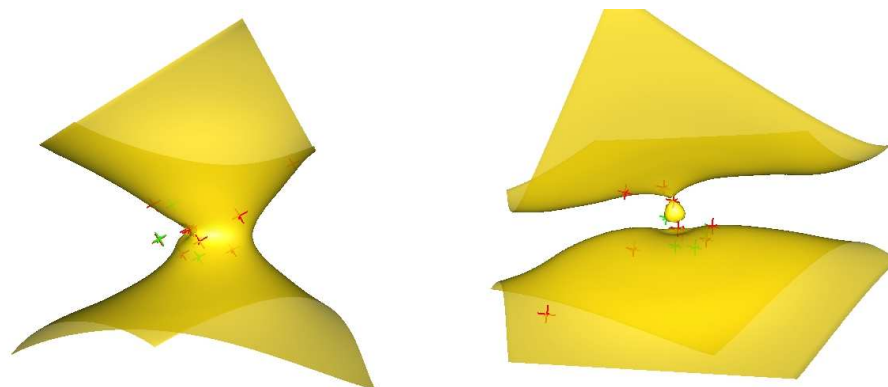


Fig. 11. Projected Hartree-Fock node of  $^{10}S(sp^3d^5)$  of the majority spin valence electrons in  $Mn^{+2}$  ion. The Ne-core electrons are eliminated by pseudopotentials. Note the deformations from the radial parts of orbitals, including a small bubble detached from the rest of the surface (the right picture). The green cross indicates the nucleus and the red crosses the electron positions.

### 3.5 Changes in Nodal Topologies From Interactions and Spin Correlations

We conjecture that a non-degenerate ground state of any given symmetry possesses only two maximal nodal cells. It was demonstrated for the considered fully spin-polarized systems that the corresponding HF wave functions have the desired topology, i.e., two maximal nodal cells. On the other hand, for partially spin-polarized and unpolarized systems the corresponding HF wave functions of the form

$$\Psi_{HF} = \det[\varphi_{\alpha}^{\uparrow}(\mathbf{r}_i)] \det[\varphi_{\beta}^{\downarrow}(\mathbf{r}_j)] \quad (117)$$

lead to at least *four* nodal cells due to the product of two determinants each with two cells. As it turned out, this is correct only for special cases such as non-interacting systems with occupations in both spin channels and some separable models. In general, such structure is an artefact. It appears since two different spins and corresponding antisymmetry requirement on configuration subspaces is essentially a special non-uniform symmetry requirement on the wave function. We will see below that requirements of additional or special symmetries could lead to the increase of nodal counts even in the spin-polarized cases. Interestingly, generic interactions have the effect of smoothing out the artificial and redundant nodal structures so that the minimal two nodal cell property is restored. This is obvious since it is not clear right away that interactions would drive the nodal topology changes, in particular, the changes of this type. We will illustrate this on a few examples.

Let us first analyze the simplest spin-polarized case of three electrons in a central, radially symmetric potential. Consider the lowest atomic quartet of  $S$  symmetry and even parity is  $^4S(1s2s3s)$ . In the non-interacting limit this state exhibits *six nodal cells*. The reason is that all the single-particle orbitals depend only on radii and therefore the nodes have effectively one-dimensional character. We will sketch how interactions change the topology and the number of nodal cells decreases to minimal two. In the recent paper [88] we showed this numerically for the Coulomb potential. For simplicity, we assume a simpler model of harmonically confined electrons while using the same symmetries of states and one-particle orbitals as for the Coulomb case. The harmonic oscillator has the advantage that the exponential Gaussian prefactor is the same for all states and one-particle orbitals irrespective of the occupied sub-shells. As before, we omit the prefactors in all expressions since they are irrelevant for the nodes. The non-interacting (or HF) wave function then becomes simply

$$\Psi_{HF}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = (r_2^2 - r_1^2)(r_3^2 - r_1^2)(r_3^2 - r_2^2), \quad (118)$$

since  $\phi_{2s}$  and  $\phi_{3s}$  are first and second order polynomials in  $r^2$ , respectively. It is clear that there are *six* nodal domains, because  $\Psi_{HF}$  vanishes whenever  $r_i = r_j$  and there are six permutations of  $r_1, r_2, r_3$ . Now, we can switch-on the interactions and describe the resulting correlations by expanding the wave function in excitations. The excitations into  $s$ -orbitals with higher quantum numbers do not change the nodal structure and their contributions to the exact wave function are actually very small. The dominant excited configuration is  $^4S(1s, 2p, 3p)$ , which captures angular correlations in the outer shells pair. We write the correlated, two-configuration wave function as

$$\begin{aligned} \Psi = & \det[\phi_{1s}, \phi_{2s}, \phi_{3s}] \\ & + \beta \{ \det[\phi_{1s}, \phi_{2px}, \phi_{3px}] + (x \rightarrow y) + (x \rightarrow z) \}, \end{aligned} \quad (119)$$

where  $\beta$  is the expansion coefficient. (This effect can be equivalently formulated also as a Pfaffian with triplet pairing). Let us now try to carry out the triple exchange. For this purpose we position particles as follows:  $\mathbf{r}_1 = [0, r_a, 0]$ ,  $\mathbf{r}_2 = [0, 0, 0]$  and  $\mathbf{r}_3 = [r_b, 0, 0]$  where  $r_a > r_b > 0$ . During the exchange the particles first move as follows:  $1 \rightarrow 2$ ,  $2 \rightarrow 3$  and  $3 \rightarrow [r_a, 0, 0]$ . The exchange is then finished by rotating particle 3 by  $+\pi/2$  around the  $z$ -axis to the position  $[0, r_a, 0]$ , completing thus the repositioning  $3 \rightarrow 1$ . The particle 1 moves to the origin along the  $y$ -axis while particles 2, 3 move along the  $x$ -axis while their radii are such that  $r_3 > r_2$  is maintained during the exchange. Clearly, during the exchange the node of non-interacting wave function will be crossed twice: when  $r_1 = r_3$  and again when  $r_1 = r_2$ . We will now show that for the correlated wave function the exchange can be carried out without crossing the node. Using the expressions for the states of harmonic oscillator  $\phi_{2px} = x$  and  $\phi_{3px} = x(r^2 - 1)$  and substituting the coordinates as outlined above, one finds that the wave function becomes

$$\Psi = (r_3^2 - r_2^2)[(r_2^2 - r_1^2)(r_3^2 - r_1^2) + \beta x_2 x_3] \quad (120)$$

We assume that  $\beta > 0$ . (In the case  $\beta < 0$ , it is straightforward to modify the positions and exchange path accordingly: the initial position of the particle 3 should be  $[-r_b, 0, 0]$ , it moves to  $[-r_a, 0, 0]$  and it is then rotated by  $-\pi/2$  around the  $z$ -axis.) Typical value of  $\beta$  is of the order of  $10^{-2}$  which reflects the fact that the same spins (triplet) correlations are usually very small. After a little bit of analysis one finds that this function does not change the sign provided the difference between the radii  $r_2$  and  $r_3$  when  $r_1 = (r_2 + r_3)/2$  is sufficiently small. Then the following inequality holds

$$|r_2 - r_3 + \mathcal{O}[(r_3 - r_2)^2]| < \beta, \quad (121)$$

which is easy to guarantee with some care when moving the particles along the exchange path. The minimum of the radial part appears essentially at the point  $r_1 = (r_2 + r_3)/2$  and the value is negative. However, the other term keeps the wave function positive providing the negative term is not very large. Since the term  $|r_2 - r_3 + \mathcal{O}[(r_3 - r_2)^2]|$  can be made arbitrarily small by adjusting radii  $r_2, r_3$  to be sufficiently close to each other one can easily fulfill the inequality. The wave function then remains positive along the whole exchange path. The nodes of uncorrelated and correlated wave function are plotted in Fig. 12.

Similar topological change occurs in spin-unpolarized systems. The simplest possible case is the Be atom ground state with the electronic state  $^1S[1s^2 2s^2]$ . The wave-function can be written as a sum of HF and correlation components, i.e.,  $\Psi = \Psi_{HF} + \alpha \Psi_{corr}$ , where  $\alpha$  is an expansion coefficient. The HF part is given as

$$\Psi_{HF} = \det^\uparrow[\phi_{1s}, \phi_{2s}] \det^\downarrow[\phi_{1s}, \phi_{2s}], \quad (122)$$

while the dominant correlating configuration is the  $2s^2 \rightarrow 2p^2$  excitation

$$\Psi_{corr} = \det^\uparrow[\phi_{1s}, \phi_{2px}] \det^\downarrow[\phi_{1s}, \phi_{2px}] + (x \rightarrow y) + (x \rightarrow z). \quad (123)$$

Clearly, the wave function has the required  $S$  symmetry and it is straightforward to show that the wave function has only two nodal cells. Consider a plane passing through the origin and configure the two spin-up electrons at antipodal points on a circle with radius  $r_a > 0$ . Place the spin-down electrons on the same plane at antipodal points of a circle with different radius, say,

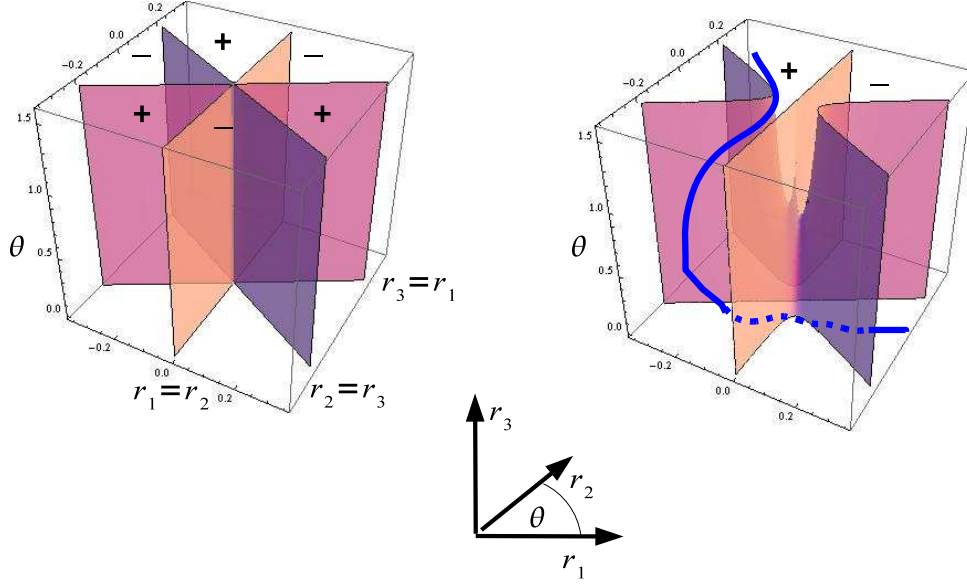


Fig. 12. The HF node and correlated node for the  $^4S(1s2s3s)$  central potential state. Note the change in topology after the inclusion of the correlation through the dominant additional configuration. The figure on the left plots the *complete* HF node since it depends only on the three radii. The figure on the right shows the node only on 3D subspace, there are more nodal “openings” in other dimensions.

$r_b > 0$  and  $r_b \neq r_a$ . In this configuration the HF wave function vanishes since particles lie on the Hartree–Fock node. In general, however, the correlated wave function does not vanish since we have

$$\Psi = C \mathbf{r}_{21} \cdot \mathbf{r}_{43}, \quad (124)$$

where  $C$  is a symmetric non-zero factor. We can now rotate *all four particles* in the common plane by  $\pi$  and the wave function remains constant. The rotation exchanges the two spin-up electrons and the two spin-down electrons simultaneously. This shows that the spin-up and -down nodal *domains are connected* otherwise we would have encountered zeros. Note that doing the rotation in one of the spin-channels only, i.e., the two-particle exchange for that spin, would necessarily result in hitting the node. The correlation therefore enables the four-particle exchange to avoid the node and the example provides an interesting insight into the many-particle concerted exchange effects. Similar behavior can be found in homogeneous periodic gas and again can be generalized to arbitrary sizes [89]. In the periodic systems one can show that these effects enable to wind singlet pairs around the periodic cell without encountering the node providing the path is chosen appropriately and also that similar effects is impossible for HF wave functions [89].

Multi-particle exchanges are closely related to quantum condensates such as Bose-Einstein condensation in bosonic systems or superconductivity in fermionic systems. The interpretation of condensates as systems which exhibit macroscopic chains of quantum exchanges goes back

to Feynman. In bosonic systems this does not help much and even looks “trivial” since there are no nodes to begin with and the emergence of condensates is driven by bosonic correlations. In fermionic systems the correct nodal topologies also appear only as a *necessary* but *not sufficient* condition for the condensation to appear. However, from this point of view it is important that the Bardeen–Cooper–Schrieffer wave functions, which will be described later, do exhibit the correct nodal topologies as shown elsewhere [88, 89].

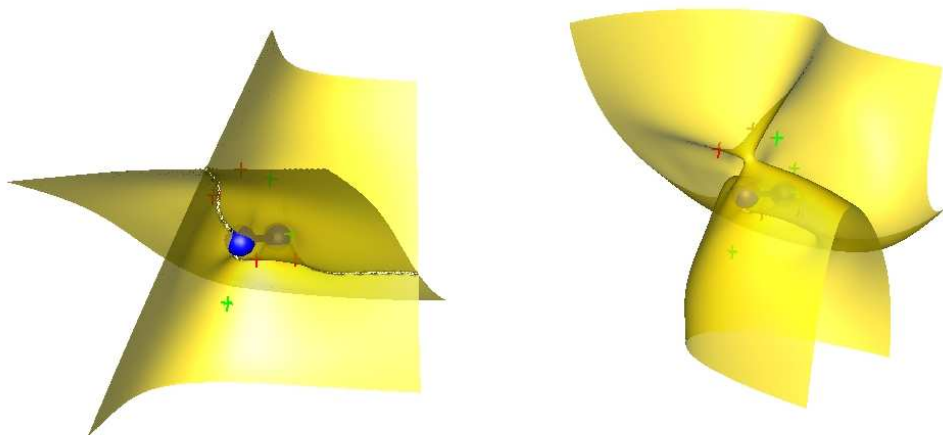


Fig. 13. The 3D projected nodes for  $N_2$  dimer. The HF nodes (left) are bended and distorted due to spin correlations into CI nodal surface (right) with just two nodal cells. Sampling of nodes is performed with a pair of electrons with opposite spins, which are close to each other; positions of other electrons (crosses) are fixed; small spheres indicate ions.

### 3.5.1 HF and CI Nodes for $N_2$ Dimer State $^1\Sigma_g^+$

The projected nodal structure of the spin-unpolarized ground state  $^1\Sigma_g^+$  for  $N_2$  dimer is shown in Fig. 13. The He-core electrons are eliminated by pseudopotentials as in previous calculations [93]. The HF wave function with the separation of electrons into independent spin-up and spin-down subspaces forms four nodal cells. In the CI case we can see that the HF nodes have been distorted and bended to build up channels, which connect regions of the same sign with resulting two maximal nodal cells.

In order to illustrate the accuracy of the constructed CI nodes we have carried out the fixed-node DMC simulations with both HF and CI trial wave functions. For the HF trial wave function we have obtained the total energy  $-19.8395(7)$  H, which recovers  $\approx 94\%$  of correlation energy. For the CI trial function with more than 5000 determinants from double excitations into 45 virtual orbitals with ccpV6Z basis (with up to  $f$  basis functions) the total energy reaches  $-19.870(5)$  H, recovering thus  $\approx 98\%$  of correlation energy. The estimated exact energy is  $-19.8822$  H. It is clear that despite the extensive determinantal expansion the fixed-node error is still visible. It is

just another demonstration that convergence of such correlated wave functions is slow and more sophisticated approaches are needed to reach beyond the limits of the determinantal expansions.

### 3.6 Conclusions

We reviewed some of the recent advances in understanding the general properties of nodes of fermionic wave functions. One of the key findings is that the nodal topologies of generic fermionic ground states appear to be remarkably simple. For  $d > 1$  the node *bisects* the configurations space with the resulting nodal domains being connected. Intuitively, this is not completely unexpected since presence of any node raises the wave function curvature, i.e., the kinetic energy. Therefore the energy minimization leads to decreasing both the nodal “volume” and also the corresponding curvatures. We have also seen that by imposing additional symmetries one can increase the nodal domain counts. What is less obvious that even in such cases the interactions tend to smooth out the nodal surfaces and fuse the multiple domains into the minimal number. This happens for both spin-polarized and unpolarized systems as we have shown on few-particle examples with correlated wave functions.

Comparison of bosonic and fermionic ground states from the viewpoint of nodal topologies reveals an interesting perspective. It is well-known that generic bosonic ground states are nodeless and so that one can say that they represent “global S-waves”. On the other hand, the bisection of the configuration space suggests that generic fermionic ground states are “global P-waves”. This is a remarkably beautiful result of the interplay between continuous space and symmetries: it shows that Nature has chosen the simplest topologies to fulfill the fundamental symmetries.

Furthermore, analysis of QMC data also reveals that much of the correlation in fermionic systems has actually bosonic nature, i.e., particle correlations within the nodal domains such two-electron cusps, multi-particle even parity exchanges, etc. This gives a rather prominent role to the fixed-node DMC method which recovers such bosonic correlations exactly. Comparison of fixed-node results with estimations of exact total energies (from other methods for small number of electrons and from experiments for extended systems) show that this is indeed the case. Calculations of atoms, molecules, clusters, solids, quantum liquids, etc, shows that the amount of recovered correlation is similar essentially in all cases and reaches 90-95% of the total correlation energy irrespective of the type of the system. This is very high when compared with other approaches based on explicitly correlated wave functions. In fact, for extended systems, there is no other method which can achieve such accuracy. Indeed, this is a significant outcome of the QMC efforts over the years and it is also one of the key reasons why these approaches are becoming more common. The achieved level of accuracy now enables to predict many “robust” quantities, i.e., which are larger than 0.1-0.2 eV, such as cohesions, optical gaps, etc.

On the other hand, the effects which are beyond the fixed-node approximation still leave many important and exciting problems unsolved or unreachable. Subtle effects such as magnetic ordering, presence of new quantum phases, happen at scales of tens of meV or smaller and, consequently, can be significantly affected by the fixed-node bias. A limited understanding of many of these aspects is still the main reason why the progress has been very gradual albeit steady over the years. Although the origin of these errors is exclusively in *inaccuracies of the nodal shapes* at present there are no clear-cut solutions which would enable to describe these in an efficient and general manner. Also how much and which types of nodal errors can influence interesting quantum phenomena is not yet understood. One possible path forward is to get better handle of the

nodal variational freedom using more effective nodal representations or new constructions which will address this particular challenge. It seems that this is out of reach of traditional approaches: common correlated wave function constructions, such as expansion in excited determinants, appear to be very slowly converging and are therefore inefficient, limiting the applications to small sizes. Such constructions are essentially non-existent for extended systems such as solids or surfaces. Some of the recently explored ideas in this direction will be presented below. Another possibility to overcome this fundamental obstacle might come from better use of statistical methods to deal with the nodal inaccuracies. In this respect, we mentioned some approaches in the previous section and there are ongoing efforts by us and others in exploring ideas in this research area as well. Finally, “brute force” approaches, such as release node methods, are not necessarily useless and could be useful for benchmarks in smaller and intermediate size systems.

## 4 Pfaffian Pairing Wave Functions

### 4.1 Introduction

As we mentioned, commonly used trial wave functions are based on one-particle orbitals and Slater-Jastrow forms. In order to overcome this one-particle based constructions, pair orbitals have been previously tried in various frameworks. In condensed systems one such example is the Bardeen–Cooper–Schrieffer (BCS) wave function, which is an antisymmetrized product of singlet pairs. In quantum chemistry, the pair orbital is sometimes referred to as geminal and the resulting wave function as the antisymmetrized geminal product (AGP). It has been recently used to calculate several atoms and molecules as well as superfluid ultracold Fermi gases [85, 86, 97]. The results show promising gains, nevertheless, for partially spin-polarized systems the improvements are less systematic due to the lack of pair correlations in the spin-polarized subspace [85, 86]. The spin-polarized (triplet) pairing wave functions naturally lead to Pfaffians, which have been previously applied to model systems [98–100].

In this section, we further develop ideas from Sec. 2.4.2, in which we have introduced the Pfaffian form. Sec. 4.2 presents mathematical identities and expressions associated with Pfaffians, some of them derived for the first time. In Sec. 4.3 we establish relationships between the Pfaffian, BCS and HF wave functions. These forms are tested on atomic and molecular systems in variational and fixed-node DMC OA methods as described in Sec. 4.4.1. In Sec. 4.4.2 we investigate generalizations to linear combinations of Pfaffians and compare the results with other methods in regard to wave functions compactness. the wave functions. In Sec. 4.4.3 we analyze the topological differences between HF, Pfaffian and an essentially exact wave functions for a given test example.

### 4.2 Algebra of Pfaffians

Here we will closely follow our exposition of Pfaffian algebra from Ref. [2].

#### 4.2.1 Definitions

First introduced by Cayley in 1852 [101], the Pfaffian is named after German mathematician Johann Friedrich Pfaff. Given a  $2n \times 2n$  skew-symmetric matrix  $A = [a_{i,j}]$ , the Pfaffian of  $A$  is defined as antisymmetrized product

$$\begin{aligned} \text{pf}[A] &= \mathcal{A}[a_{1,2}a_{3,4} \dots a_{2n-1,2n}] \\ &= \sum_{\alpha} \text{sgn}(\alpha) a_{i_1,j_1} a_{i_2,j_2} \dots a_{i_n,j_n}, \end{aligned} \tag{125}$$

where the sum runs over all possible  $(2n-1)!!$  pair partitions  $\alpha = \{(i_1, j_1), (i_2, j_2), \dots, (i_n, j_n)\}$  of  $\{1, 2, \dots, 2n\}$  with  $i_k < j_k$ . The sign of permutation associated with the partition  $\alpha$  is denoted as  $\text{sgn}(\alpha)$ . The Pfaffian for a matrix of odd order equals to zero. The following example



gives Pfaffian of a  $A(4 \times 4)$  skew-symmetric matrix

$$\text{pf} \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ -a_{14} & -a_{24} & -a_{34} & 0 \end{bmatrix} = a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23}. \quad (126)$$

It can be also evaluated recursively as

$$\begin{aligned} \text{pf}[A] &= \sum_{j=2}^{2n} a_{1,j} \sum_{\alpha_{1,j}} \text{sgn}(\alpha_{1,j}) a_{i_1,j_1} a_{i_2,j_2} \cdots a_{i_{n-1},j_{n-1}} \\ &\equiv \sum_{j=2}^{2n} a_{1,j} P_c(a_{1,j}), \end{aligned} \quad (127)$$

where  $\alpha_{1,j}$  is partition with  $i_k, j_k \neq 1, j$  and  $P_c(a_{1,j})$  is defined as Pfaffian cofactor of  $a_{1,j}$ . The cofactor for an element  $a_{j,k}$  is given by a formula,

$$P_c(a_{j,k}) = (-1)^{j+k+1} \text{pf}[A(j, k; j, k)] \quad (128)$$

where the matrix  $A(j, k; j, k)$  has the rank  $2(n-1) \times 2(n-1)$  and is obtained from  $A$  by eliminating  $j$  and  $k$  rows and columns.

#### 4.2.2 Calculation of a Pfaffian

There exist several identities involving Pfaffians and determinants. For any  $2n \times 2n$  skew-symmetric matrix  $A$  and arbitrary matrices  $B(2n \times 2n)$  and  $M(n \times n)$  we have the following relations:

$$\text{pf}[A^T] = (-1)^n \text{pf}[A] \quad (129a)$$

$$\text{pf}[A]^2 = \det[A] \quad (129b)$$

$$\text{pf} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} = \text{pf}[A_1] \text{pf}[A_2] \quad (129c)$$

$$\text{pf}[BAB^T] = \det[B] \text{pf}[A] \quad (129d)$$

$$\text{pf} \begin{bmatrix} 0 & M \\ -M^T & 0 \end{bmatrix} = (-1)^{\frac{n(n-1)}{2}} \det[M]. \quad (129e)$$

Let us comment on these identities and point out to respective proofs as appropriate:

(129a) Each permutation contains product of  $n$  pairs resulting in an overall  $(-1)^n$  factor.

(129b) This is a well-known Cayley's relationship between the Pfaffian and the determinant of a skew-symmetric matrix. This is a well-known relationship which has been proved many times in variety of ways [102–104].

(129c) Use the expansion by Pfaffian cofactors.

(129d) By squaring (4d), using Eq. (129b), and taking the square root one finds  $\text{pf}[BAB^T] = \pm \det[B] \text{pf}[A]$ . Substituting the identity matrix  $I$  for  $B$  one finds  $+$  to be the correct sign.

(129e) Assume

$$B = \begin{pmatrix} M & 0 \\ 0 & I \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

in Eq. (129d). The overall sign is given by value of  $\text{pf}[A]$ .

The identities listed above imply several important properties. First, Eqs. (129d) and (129e) show that every determinant can be written as a Pfaffian, but on the contrary, only the absolute value of Pfaffian can be given by determinant [Eq. (129b)]. The Pfaffian is therefore a generalized form of the determinant. Second, by substituting suitable matrices [105] for  $M$  in Eq. (129d) one can verify the following three properties of Pfaffians [106], similar to the well-known properties of determinants.

- (a) Multiplication of a row and a column by a constant is equivalent to multiplication of the Pfaffian by the same constant.
- (b) Simultaneous interchange of two different rows and corresponding columns changes the Pfaffian sign.
- (c) A multiple of a row and a corresponding column added to another row and corresponding column leaves the Pfaffian value unchanged.

It is also clear that any skew-symmetric matrix can be brought to the block-diagonal form by an orthogonal transformation. Recursive evaluation [Eq. (127)] implies that the Pfaffian of block-diagonal matrix is directly given by

$$\text{pf} \begin{bmatrix} 0 & \lambda_1 & & & & \\ -\lambda_1 & 0 & & & & \\ & & 0 & \lambda_2 & & \\ & & -\lambda_2 & 0 & & \\ & & & & \ddots & \\ & 0 & & & & 0 & \lambda_n \\ & & & & & -\lambda_n & 0 \end{bmatrix} = \lambda_1 \lambda_2 \dots \lambda_n. \quad (130)$$

Therefore by employing a simple Gaussian elimination technique with row pivoting (see Ref. [107]) we can transform any skew-symmetric matrix into block-diagonal form and obtain its Pfaffian value in  $O(n^3)$  time.

However, in QMC applications, one often needs to evaluate the wave function after a single electron update. Since Cayley [104] showed (for the proof see [2]) that

$$\begin{aligned} \det \begin{bmatrix} 0 & b_{12} & b_{13} & \dots & b_{1n} \\ -a_{12} & 0 & a_{23} & \dots & a_{2n} \\ -a_{13} & -a_{23} & 0 & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & -a_{3n} & \dots & 0 \end{bmatrix} \\ = \text{pf} \begin{bmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ -a_{12} & 0 & a_{23} & \dots & a_{2n} \\ -a_{13} & -a_{23} & 0 & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & -a_{3n} & \dots & 0 \end{bmatrix} \text{pf} \begin{bmatrix} 0 & b_{12} & b_{13} & \dots & b_{1n} \\ -b_{12} & 0 & a_{23} & \dots & a_{2n} \\ -b_{13} & -a_{23} & 0 & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -b_{1n} & -a_{2n} & -a_{3n} & \dots & 0 \end{bmatrix}, \end{aligned} \quad (131)$$

we can relate the Pfaffian of original matrix  $\text{pf}[A]$  to the Pfaffian of a matrix with updated first row and column  $\text{pf}[B]$  using the inverse matrix  $A^{-1}$  in only  $O(n)$  operations by

$$\text{pf}[B] = \frac{\det[A] \sum_j b_{1j} A_{j1}^{-1}}{\text{pf}[A]} = \text{pf}[A] \sum_j b_{1j} A_{j1}^{-1}. \quad (132)$$

The second part of Eq. (132) was obtained by taking advantage of the identity in Eq. (129b). Slightly more complicated relation between  $\text{pf}[A]$  and  $\text{pf}[B]$  can be derived if one considers simultaneous change of two separate rows and columns, which represents the two electron update of a wave function.

#### 4.2.3 Gradient and Hessian of Pfaffian

In optimization routines, it is often favorable to calculate the gradient and Hessian of wave function with respect to its variational parameters. Given the matrix elements  $A$  dependent on a set of parameters  $\{c\}$ , one can derive the following useful relations:

$$\frac{1}{\text{pf}[A]} \frac{\partial \text{pf}[A]}{\partial c_i} = \frac{1}{2} \text{tr} \left[ A^{-1} \frac{\partial A}{\partial c_i} \right] \quad (133)$$

and

$$\begin{aligned} \frac{1}{\text{pf}[A]} \frac{\partial^2 \text{pf}[A]}{\partial c_i \partial c_j} = & \frac{1}{4} \text{tr} \left[ A^{-1} \frac{\partial A}{\partial c_i} \right] \text{tr} \left[ A^{-1} \frac{\partial A}{\partial c_j} \right] \\ & - \frac{1}{2} \text{tr} \left[ A^{-1} \frac{\partial A}{\partial c_i} A^{-1} \frac{\partial A}{\partial c_j} \right] + \frac{1}{2} \text{tr} \left[ A^{-1} \frac{\partial^2 A}{\partial c_i \partial c_j} \right] \end{aligned} \quad (134)$$

where  $A^{-1}$  is again the inverse of  $A$ .

### 4.3 Pairing Wave Functions

Let us now consider the generalization of the one-particle orbital to a two-particle (or pair) orbital  $\tilde{\phi}(i, j)$ , where tilde again denotes dependence on both spatial and spin variables. The simplest antisymmetric wave function for  $2N$  electrons constructed from the pair orbital is a *Pfaffian*

$$\Psi = \mathcal{A}[\tilde{\phi}(1, 2), \tilde{\phi}(3, 4) \dots \tilde{\phi}(2N - 1, 2N)] = \text{pf}[\tilde{\phi}(i, j)]. \quad (135)$$

The antisymmetry is guaranteed by the definition (125), since the signs of pair partitions alternate depending on the parity of the corresponding permutation. The important difference from Slater determinant is that in the simplest case only *one* pair orbital is necessary. (This can be generalized, of course, as will be shown later.) If we further restrict our description to systems with collinear spins, the pair orbital  $\tilde{\phi}(\mathbf{r}_i, \sigma_i; \mathbf{r}_j, \sigma_j)$  for two electrons in positions  $\mathbf{r}_i$  and  $\mathbf{r}_j$  and with spins projections  $\sigma_i$  and  $\sigma_j$  can be expressed as

$$\begin{aligned} \tilde{\phi}(\mathbf{r}_i, \sigma_i; \mathbf{r}_j, \sigma_j) = & \phi(i, j) \langle \sigma_i \sigma_j | [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] / \sqrt{2} \\ & + \chi^{\uparrow\uparrow}(i, j) \langle \sigma_i \sigma_j | \uparrow\uparrow \rangle \\ & + \chi^{\uparrow\downarrow}(i, j) \langle \sigma_i \sigma_j | [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle] / \sqrt{2} \\ & + \chi^{\downarrow\downarrow}(i, j) \langle \sigma_i \sigma_j | \downarrow\downarrow \rangle. \end{aligned} \quad (136)$$

Here  $\phi(i, j) = \phi(\mathbf{r}_i, \mathbf{r}_j)$  is even while  $\chi^{\uparrow\uparrow}$ ,  $\chi^{\uparrow\downarrow}$  and  $\chi^{\downarrow\downarrow}$  are odd functions of spatial coordinates. In the rest of this section we will discuss special cases of the wave function (135).

#### 4.3.1 Singlet Pairing Wave Function

Let us consider the first  $1, 2, \dots, N$  electrons to be spin-up and the rest  $N + 1, \dots, 2N$  electrons to be spin-down and allow only  $\phi(\mathbf{r}_i, \mathbf{r}_j)$  in  $\tilde{\phi}(\mathbf{r}_i, \sigma_i; \mathbf{r}_j, \sigma_j)$  to be non-zero. Using the Pfaffian identity [Eq. (129e)], we can write the wave function for  $N$  singlet pairs, also known as the Bardeen–Cooper–Schrieffer (BCS) wave function (or AGP), in the following form

$$\Psi_{BCS} = \text{pf} \begin{bmatrix} 0 & \Phi^{\uparrow\downarrow} \\ -\Phi^{\uparrow\downarrow T} & 0 \end{bmatrix} = \det[\Phi^{\uparrow\downarrow}], \quad (137)$$

which is simply a determinant of the  $N \times N$  matrix  $\Phi^{\uparrow\downarrow} = [\phi(i, j)]$  as was shown previously [108, 109].

Let us comment on the BCS wave function which has its origin on the theory of superconductivity [110]. The original form is actually for variable number of particles, since the theory is formulated in the grand canonical ensemble—the number of particles adjusts to the chemical potential. Hence, if the grand canonical form is projected onto the state with *fixed* number of particles, its form is then given by Eq. (137).

It is straightforward to show that the BCS wave function contains the restricted HF wave function as a special case. Let us define the Slater matrix  $C = [\varphi_i(j)]$  where  $\{\varphi_i\}$  is a set of HF occupied orbitals. Then we can write

$$\Psi_{HF} = \det[C] \det[C] = \det[CC^T] = \det[\Phi_{HF}^{\uparrow\downarrow}], \quad (138)$$

where

$$(\Phi_{HF}^{\uparrow\downarrow})_{i,j} = \phi_{HF}(i,j) = \sum_{k=1}^N \varphi_k(i) \varphi_k(j). \quad (139)$$

On the other hand, we can think of the BCS wave function as a natural generalization of the HF one. To do so we write the singlet pair orbital as

$$\phi(i,j) = \sum_{k,l}^{M>N} S_{k,l} \varphi_k(i) \varphi_l(j) = \varphi(i) \mathbf{S} \varphi(j), \quad (140)$$

where the sum runs over all  $M$  (occupied and virtual) single-particle orbitals and  $\mathbf{S}$  is some symmetric matrix. Therefore, we can define one-particle orbitals, which diagonalize this matrix and call them *natural orbitals of a singlet pair*.

The BCS wave function is efficient for describing systems with single-band correlations such as Cooper pairs in conventional BCS superconductors, where pairs form from one-particle states close to the Fermi level.

### 4.3.2 Triplet Pairing Wave Function

Let us assume, in our system of  $2N$  electrons, that the first  $M_1$  electrons are spin-up and remaining  $M_2 = 2N - M_1$  electrons are spin-down. Further, we restrict  $M_1$  and  $M_2$  to be even numbers. Then by allowing only  $\chi^{\uparrow\uparrow}(i,j)$  and  $\chi^{\downarrow\downarrow}(i,j)$  in (136) to be non-zero, we obtain from expression (135) by the use of Eq. (129c)

$$\Psi_{TP} = \text{pf} \begin{bmatrix} \xi^{\uparrow\uparrow} & 0 \\ 0 & \xi^{\downarrow\downarrow} \end{bmatrix} = \text{pf}[\xi^{\uparrow\uparrow}] \text{pf}[\xi^{\downarrow\downarrow}], \quad (141)$$

where we have introduced  $M_1 \times M_1 (M_2 \times M_2)$  matrices  $\xi^{\uparrow\uparrow(\downarrow\downarrow)} = [\chi^{\uparrow\uparrow(\downarrow\downarrow)}(i,j)]$ . This result was previously found in a weaker form in Ref. [108, 109].

The connection to a restricted HF wave function for the above state can be again established as follows. In accord with what we defined above,  $\det[(C)^{\uparrow(\downarrow)}]$  are spin-up(-down) Slater determinants of some HF orbitals  $\{\varphi_i\}$ . Then, by taking advantage of Eq. (129e) we can write

$$\begin{aligned} \Psi_{HF} &= \det[C^{\uparrow}] \det[C^{\downarrow}] \\ &= \frac{\text{pf}[C^{\uparrow} A_1 C^{\uparrow T}] \text{pf}[C^{\downarrow} A_2 C^{\downarrow T}]}{\text{pf}[A_1] \text{pf}[A_2]}, \end{aligned} \quad (142)$$

given  $A_1$  and  $A_2$  are some skew-symmetric non-singular matrices. In the simplest case, when  $A_1$  and  $A_2$  have block-diagonal form (130) with all values  $\lambda_i = 1$ , one gets

$$\Psi_{HF} = \text{pf}[\xi_{HF}^{\uparrow\uparrow}] \text{pf}[\xi_{HF}^{\downarrow\downarrow}]. \quad (143)$$

The pair orbitals can be then expressed as

$$\begin{aligned} (\xi_{HF}^{\uparrow\uparrow(\downarrow\downarrow)})_{i,j} &= \chi_{HF}^{\uparrow\uparrow(\downarrow\downarrow)}(i,j) \\ &= \sum_{k=1}^{M_1(M_2)/2} (\varphi_{2k-1}(i) \varphi_{2k}(j) - \varphi_{2k-1}(j) \varphi_{2k}(i)). \end{aligned} \quad (144)$$

Similarly to the singlet pairing case, the triplet pairing wave function appears as a natural generalization of the HF one. We can write the triplet pair orbitals as

$$\begin{aligned}\chi(i, j)^{\uparrow\uparrow(\downarrow\downarrow)} &= \sum_{k, l}^{M > M_1(M_2)} A_{k, l}^{\uparrow\uparrow(\downarrow\downarrow)} \varphi_k(i) \varphi_l(j) \\ &= \varphi(i) \mathbf{A}^{\uparrow\uparrow(\downarrow\downarrow)} \varphi(j),\end{aligned}\quad (145)$$

where again the sum runs over all  $M$  (occupied and virtual) single-particle orbitals and  $\mathbf{A}^{\uparrow\uparrow(\downarrow\downarrow)}$  are some skew-symmetric matrices. Therefore we can define one-particle orbitals which block-diagonalize these matrices and call them *natural orbitals of a triplet spin-up-up (down-down) pair*.

### 4.3.3 Generalized Pairing Wave Function

Let us now consider a partially spin-polarized system with unpaired electrons. In order to introduce both types of pairing we allow  $\chi^{\uparrow\uparrow}(i, j)$ ,  $\chi^{\downarrow\downarrow}(i, j)$  and  $\phi(i, j)$  in (136) to be non-zero. However, we omit the  $\chi^{\uparrow\downarrow}(i, j)$  term. Then our usual ordered choice of electrons labels with all spin-up electrons first and remaining electrons spin-down enables us to directly write from (135) the singlet-triplet-unpaired (STU) orbital Pfaffian wave function [1]

$$\Psi_{STU} = \text{pf} \begin{bmatrix} \boldsymbol{\xi}^{\uparrow\uparrow} & \boldsymbol{\Phi}^{\uparrow\downarrow} & \boldsymbol{\varphi}^{\uparrow} \\ -\boldsymbol{\Phi}^{\uparrow\downarrow T} & \boldsymbol{\xi}^{\downarrow\downarrow} & \boldsymbol{\varphi}^{\downarrow} \\ -\boldsymbol{\varphi}^{\uparrow T} & -\boldsymbol{\varphi}^{\downarrow T} & 0 \end{bmatrix}, \quad (146)$$

where the bold symbols are block matrices or vectors of corresponding orbitals as defined in Sections 4.3.1 and 4.3.2 and  $T$  denotes transposition. For a spin-restricted STU wave function the pair and one-particle orbitals of spin-up and -down channels would be identical.

The Pfaffian form can accommodate both singlet and triplet pairs as well as one-particle unpaired orbitals into a single, compact wave function. The correspondence of STU, Pfaffian wave function to HF wave function can be established in a similar way for the pure singlet and triplet pairing cases.

It is instructive to write down some illustrative examples of the introduced form for a few-electron atomic states. For example, for the Li atom with  $^2S(1s^2 2s)$  we can write

$$\begin{aligned}\Psi_{STU} &= \text{pf} \begin{bmatrix} 0 & \xi^{\uparrow\uparrow}(1, 2) & \phi^{\uparrow\downarrow}(1, 3) & \varphi_{2s}^{\uparrow}(1) \\ \xi^{\uparrow\uparrow}(2, 1) & 0 & \phi^{\uparrow\downarrow}(2, 3) & \varphi_{2s}^{\uparrow}(2) \\ \phi^{\uparrow\downarrow}(3, 1) & \phi^{\uparrow\downarrow}(3, 2) & 0 & \varphi_{1s}^{\downarrow}(3) \\ -\varphi_{2s}^{\uparrow}(1) & -\varphi_{2s}^{\uparrow}(2) & -\varphi_{1s}^{\downarrow}(3) & 0 \end{bmatrix} = \\ &= \xi^{\uparrow\uparrow}(1, 2) \varphi_{1s}^{\downarrow}(3) - \phi^{\uparrow\downarrow}(1, 3) \varphi_{2s}^{\uparrow}(2) + \phi^{\uparrow\downarrow}(2, 3) \varphi_{2s}^{\uparrow}(1).\end{aligned}\quad (147)$$

It is easy to check that the expression has the correct antisymmetry. Also note that one can get the HF wave function by substitution

$$\phi^{\uparrow\downarrow}(i, j) = \varphi_{1s}^{\uparrow}(i) \varphi_{1s}^{\downarrow}(j), \quad \xi^{\uparrow\uparrow}(i, j) = 0 \quad (148)$$

or, alternatively, by setting

$$\phi^{\uparrow\downarrow}(i, j) = 0, \quad \xi^{\uparrow\uparrow}(i, j) = \varphi_{1s}^{\uparrow}(i)\varphi_{2s}^{\uparrow}(j) - \varphi_{1s}^{\uparrow}(j)\varphi_{2s}^{\uparrow}(i). \quad (149)$$

For the valence part of the carbon atom (we skip the He-core), we can write the state as  $^3P(2s^22p^2)$  with the corresponding wave function

$$\Psi_{STU} = \text{pf} \begin{bmatrix} 0 & \xi^{\uparrow\uparrow}(1, 2) & \xi^{\uparrow\uparrow}(1, 3) & \phi^{\uparrow\downarrow}(1, 4) \\ \xi^{\uparrow\uparrow}(2, 1) & 0 & \xi^{\uparrow\uparrow}(2, 3) & \phi^{\uparrow\downarrow}(2, 4) \\ \xi^{\uparrow\uparrow}(3, 1) & \xi^{\uparrow\uparrow}(3, 2) & 0 & \phi^{\uparrow\downarrow}(3, 4) \\ -\phi^{\uparrow\downarrow}(4, 1) & -\phi^{\uparrow\downarrow}(4, 2) & -\phi^{\uparrow\downarrow}(4, 3) & 0 \end{bmatrix}. \quad (150)$$

Finally, for the nitrogen atom valence electrons the quartet state given by  $^4S(2s^22p^3)$  and the wave function can be written as

$$\Psi_{STU} = \text{pf} \begin{bmatrix} 0 & \xi^{\uparrow\uparrow}(1, 2) & \xi^{\uparrow\uparrow}(1, 3) & \xi^{\uparrow\uparrow}(1, 4) & \phi^{\uparrow\downarrow}(1, 5) & \varphi^{\uparrow}(1) \\ \xi^{\uparrow\uparrow}(2, 1) & 0 & \xi^{\uparrow\uparrow}(2, 3) & \xi^{\uparrow\uparrow}(2, 4) & \phi^{\uparrow\downarrow}(2, 5) & \varphi^{\uparrow}(2) \\ \xi^{\uparrow\uparrow}(3, 1) & \xi^{\uparrow\uparrow}(3, 2) & 0 & \xi^{\uparrow\uparrow}(3, 4) & \phi^{\uparrow\downarrow}(3, 5) & \varphi^{\uparrow}(3) \\ \xi^{\uparrow\uparrow}(4, 1) & \xi^{\uparrow\uparrow}(4, 2) & \xi^{\uparrow\uparrow}(4, 3) & 0 & \phi^{\uparrow\downarrow}(4, 5) & \varphi^{\uparrow}(4) \\ -\phi^{\uparrow\downarrow}(5, 1) & -\phi^{\uparrow\downarrow}(5, 2) & -\phi^{\uparrow\downarrow}(5, 3) & -\phi^{\uparrow\downarrow}(5, 4) & 0 & \varphi^{\downarrow}(5) \\ -\varphi^{\uparrow}(1) & -\varphi^{\uparrow}(2) & -\varphi^{\uparrow}(3) & -\varphi^{\uparrow}(4) & -\varphi^{\downarrow}(5) & 0 \end{bmatrix}. \quad (151)$$

#### 4.4 Applications of Pairing Wave Functions in QMC

In the remainder of this section, we would like to illustrate the applications of Pfaffian pairing wave functions in QMC methods. As we have mentioned before in Sec. 2.4.2, the trial wave function used in QMC calculations by variational and fixed-node diffusion Monte Carlo is a product of an antisymmetric part  $\Psi_A$  times the Jastrow correlation factor:

$$\Psi_T(\mathbf{R}) = \Psi_A(\mathbf{R}) \exp[U_{corr}(\{r_{ij}\}, \{r_{iI}\}, \{r_{jI}\})], \quad (152)$$

where  $U_{corr}$  depends on electron-electron, electron-ion and, possibly, on electron-electron-ion combinations of distances. The antisymmetric part can be represented as  $\Psi_A = \Psi_{HF}$  and  $\Psi_A = \Psi_{STU}$  or  $\Psi_A = \Psi_{BCS}$ . Further, the pair orbitals  $\chi(i, j)$  and  $\phi(i, j)$  are expanded in products of a one-particle orbital basis [85] according to Eqs. (140) and (145). This approach was particularly convenient due to availability of one-particle atomic and molecular orbitals from either Hartree-Fock or CI correlated calculations. However, expansion of this type might not be the most optimal for homogeneous or extended periodic systems and other forms were also successfully employed [97]. For the applications below, we typically used about 10 virtual orbitals. The natural orbitals (see e.g. Ref. [22]) produced better and more systematic results than the HF ones. The pair orbital expansion coefficients were then optimized in VMC by minimizations of energy, variance or a combination of energy and variance (for details, see Sec. 7.2). The optimization procedure required the calculation of gradient and the Hessian of the wave function according to Eqs. (133) and (134).

#### 4.4.1 Single Pfaffian Calculations

The performance of single Pfaffian pairing wave functions in the STU form [Eq. (146)] was tested on several first row atoms and dimers [1]. The pseudopotentials [93, 111] were used to eliminate the atomic core electrons, except in Be atom case.

The most important result is that STU wave functions systematically achieve higher percentage of recovered correlation energy than single determinant wave functions in the fixed-node DMC method (see Table 2 and Fig. 14). Another significant result is that in general the triplet contribution for these single Pfaffian STU wave functions are small, perhaps the only exception being the nitrogen atom, where we see a gain of additional 1% in correlation energy when compared to a trial wave function without triplet pairs. We believe, this is due to the fact, that ground state of nitrogen atom has a quartet spin state and therefore the highest spin polarization from all studied cases. Overall, the single Pfaffian form is capable of capturing near-degeneracies and mixing of excited states for both spin-polarized and unpolarized systems.

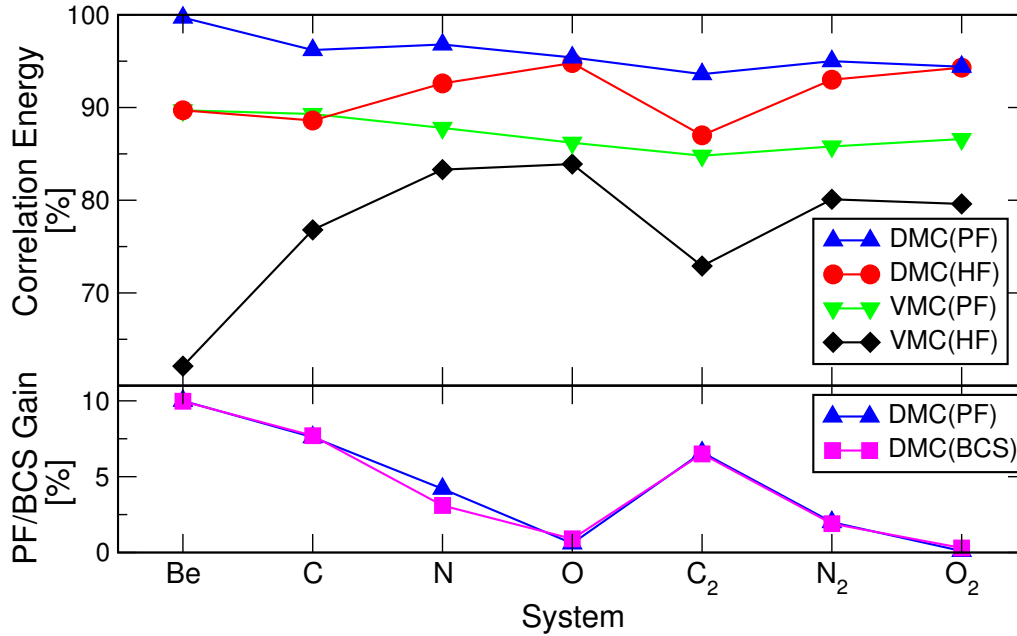


Fig. 14. Correlation energies obtained by QMC methods with the different trial wave functions: VMC and fixed-node DMC with HF nodes (HF) and STU Pfaffian nodes (PF). The lower plot shows the fixed-node DMC correlation energy gains over HF nodes for BCS and STU Pfaffian wave functions. The statistical error bars are of the symbol sizes or smaller. Except for the Be atom all the calculations used the same pseudopotentials [93, 111].



Tab. 2. Total energies for C, N and O atoms and their dimers with amounts of the correlation energy recovered in VMC and DMC methods with wave functions as discussed in the text. Unless noted otherwise, the numbers in parentheses are the statistical errors in the last digit from corresponding QMC calculation. Energies are in Hartree atomic units. For C, N, O atoms we used the correlation energies by Dolg [112] (0.1031, 0.1303, 0.1937 H). For the estimation of correlation energies of dimers we needed accurate HF energies at experimental distances [113] and the estimated exact total energies. Each exact total energy was estimated as a sum of total energies of constituent atoms minus experimental binding energy [113–115] adjusted for experimental zero-point energy [115].

Method/WF	C	$E_{corr}[\%]$	N	$E_{corr}[\%]$	O	$E_{corr}[\%]$
HF	-5.31471	0	-9.62892	0	-15.65851	0
VMC/HF	-5.3939(4)	76.8(4)	-9.7375(1)	83.3(1)	-15.8210(6)	83.9(3)
VMC/BCS	-5.4061(2)	88.6(2)	-9.7427(3)	87.3(2)	-15.8250(3)	86.0(2)
VMC/STU	-5.4068(2)	89.3(2)	-9.7433(1)	87.8(1)	-15.8255(3)	86.2(2)
DMC/HF	-5.4061(3)	88.6(2)	-9.7496(2)	92.6(2)	-15.8421(2)	94.8(1)
DMC/BCS	-5.4140(2)	96.3(2)	-9.7536(2)	95.7(2)	-15.8439(4)	95.7(2)
DMC/STU	-5.4139(2)	96.2(2)	-9.7551(2)	96.8(1)	-15.8433(3)	95.4(2)
Est./Exact	-5.417806	100	-9.759215	100	-15.85216	100
Method/WF	C <sub>2</sub>	$E_{corr}[\%]$	N <sub>2</sub>	$E_{corr}[\%]$	O <sub>2</sub>	$E_{corr}[\%]$
HF	-10.6604	0	-19.4504	0	-31.3580	0
VMC/HF	-10.9579(4)	72.9(1)	-19.7958(5)	80.0(1)	-31.7858(6)	79.6(1)
VMC/BCS	-11.0059(4)	84.7(1)	-19.8179(6)	85.0(1)	-31.8237(4)	86.7(1)
VMC/STU	-11.0062(3)	84.8(1)	-19.821(1)	85.8(2)	-31.8234(4)	86.6(1)
DMC/HF	-11.0153(4)	87.0(1)	-19.8521(3)	93.0(1)	-31.8649(5)	94.3(1)
DMC/BCS	-11.0416(3)	93.5(1)	-19.8605(6)	94.9(1)	-31.8664(5)	94.6(1)
DMC/STU	-11.0421(5)	93.6(1)	-19.8607(4)	95.0(1)	-31.8654(5)	94.4(1)
Est./Exact <sup>c</sup>	-11.068(5) <sup>a</sup>	100.0(10)	-19.8825(6) <sup>b</sup>	100.0(1)	-31.8954(1) <sup>b</sup>	100.0(1)

<sup>a</sup>There is rather large discrepancy in the experimental values of C<sub>2</sub> binding energy (141.8(9) [113], 143(3) [115] and 145.2(5) kcal/mol [114]). For the estimation of exact energy we have taken the average value of 143(3) kcal/mol.

<sup>b</sup>Experimental binding energies taken from ref. [113].

<sup>c</sup>The error bars on estimated exact total energies are due to experiment.

#### 4.4.2 Multi-Pfaffian Calculations

In order to capture the correlation energy missing from the single STU Pfaffian wave function and to test the limits of the Pfaffian functional form, we have also proposed to expand the  $\Psi_A$  in the linear combination of STU Pfaffian wave functions [1, 2].

Following the approach adopted from the CI correlated calculations, accurate  $\Psi_A$  can be expressed as a linear combination of reference state  $\Psi_0$  and single  $\Psi_i^k$  and double  $\Psi_{ij}^{kl}$  excitations:

$$\Psi_{CISD} = c_0 \Psi_0 + \sum_i^N \sum_k^M c_i^k \Psi_i^k + \sum_{ij}^N \sum_{kl}^M c_{ij}^{kl} \Psi_{ij}^{kl}, \quad (153)$$

where  $i$ -th and  $j$ -th electrons are being excited into  $k$  and  $l$  virtual orbitals. Note that the number of determinants in the expansion will be of the order of  $N^2 \times M^2$ . Analogously, we

Tab. 3. Percentages of correlation energies recovered for C, N and O atoms by VMC and DMC methods with wave functions as discussed in the text. The corresponding number of Pfaffians or determinants  $n$  for each wave function is also shown. For details, see caption of Table 2.

Method/WF	$n$	C	$n$	N	$n$	O
VMC/MPF	3	92.3(1)	5	90.6(1)	11	92.6(3)
VMC/CI <sup>a</sup>	98	89.7(4)	85	91.9(2)	136	89.7(4)
DMC/MPF	3	98.9(2)	5	98.4(1)	11	97.2(1)
DMC/CI <sup>a</sup>	98	99.3(3)	85	98.9(2)	136	98.4(2)

<sup>a</sup>The determinantal weights were taken directly from CI calculation without re-optimization in VMC.

postulate the multi-Pfaffian (MPF) wave function as a linear combination of STU Pfaffians:

$$\Psi_{MPF} = c_0 \Psi_0 + \sum_i^N \text{pf}[\tilde{\phi}_i] + \sum_{ij}^N \text{pf}[\tilde{\phi}_{ij}], \quad (154)$$

where each  $\tilde{\phi}_{ij}$  is the generalized pairing orbital [Eq. (136)] containing all possible  $M^2$  excitations of  $i$  and  $j$  electrons. The resulting wave function will in general consist of only  $N^2$  Pfaffians. Further, if the reference state  $\Psi_0$  is the most dominant state (i.e.,  $c_0 \gg c_{ij}^{kl}$ ), it is possible to show by expanding  $\tilde{\phi}_{ij}$  into orders of  $c_{ij}^{kl}/c_0$  that

$$\Psi_{MPF} = \Psi_{CISD} + \mathcal{O}((c_{ij}^{kl}/c_0)^2). \quad (155)$$

In fact, the only difference between the  $\Psi_{MPF}$  and  $\Psi_{CISD}$  wave functions is in presence of the higher order excitations, which are approximately present in the  $\Psi_{MPF}$ .

The mapping of the MPF wave functions onto equivalent CISD wave functions was also verified numerically in the variational and FN-DMC methods using the above first row atoms and molecules [1, 2]. The results in Table 3 show that for the atomic systems the MPF wave functions are able to recover close to 99% of correlation energy—very similarly to CISD wave functions, while requiring order less terms. The results for diatomic cases (see Table 4) show very similar behavior. The correlation energy recovered is on the order 98% with MPF wave functions closely matching the CISD wave functions—despite much richer electronic structure than in atomic cases. The comparison with the CI results therefore demonstrates that it is possible to obtain similar quality wave functions with corresponding improvements of the fermion nodes at much smaller calculational cost.

#### 4.4.3 Nodal Properties

As we have already extensively discussed in Sec. 3, the fermion node is a complicated hypersurface defined by an implicit equation  $\Psi_A(R) = 0$ . Due to its complexity, we approximate the exact fermion node with the node of the best available trial wave function. However, any deviations from the exact node lead to the fixed-node DMC errors. It is therefore quite valuable to compare the nodes of a different accuracy.

Tab. 4. Total energies for  $C_2$  and  $N_2$  dimers with amounts of correlation energy recovered in VMC and DMC methods with wave functions as discussed in the text. Energies are in Hartree atomic units. The corresponding number of Pfaffians or determinants  $n$  for each wave function is also shown. For details, see caption of Table 2.

Method/WF	$n$	$C_2$	$E_{corr}[\%]$	$n$	$N_2$	$E_{corr}[\%]$
VMC/MPF <sup>c</sup>	11	-11.0402(1)	93.1(1)	16	-19.8413(6)	90.5(1)
VMC/CI <sup>a</sup>	404	-11.0409(3)	93.3(1)	535	-19.8487(6)	92.2(1)
DMC/MPF <sup>c</sup>	11	-11.0574(5)	97.3(1)	16	-19.8670(8)	96.4(2)
DMC/CI <sup>a</sup>	404	-11.0593(6) <sup>b</sup>	97.8(2)	535	-19.8713(5)	97.4(1)

<sup>a</sup>The determinantal weights were re-optimized in the VMC method.

<sup>b</sup>Recently, Umrigar *et.al.* [84] published very accurate DMC result for fully optimized CI wave function with up to 500 determinants for  $C_2$  molecule. The resulting well-depth of his calculation is 6.33(1) eV, which is only 0.03 eV from estimated exact value of Ref. [116]. The well-depth resulting from our DMC/CI energy of  $-11.0593(6)$  H equals to 6.08(3) eV.

<sup>c</sup>one type of triplet-like excitation not included.

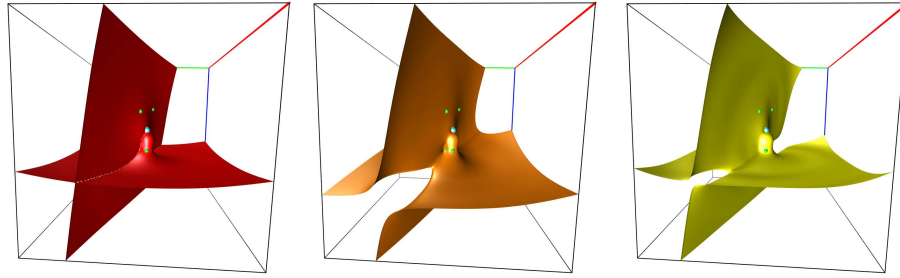


Fig. 15. A three-dimensional cut through the fermion node hypersurface of oxygen atom obtained by scanning the wave function with a spin-up and -down (singlet) pair of electrons at the equal positions, while keeping the rest of electrons at a given VMC snapshot positions (small green spheres). Nucleus is depicted in the center of the cube by the blue sphere. The three colors (from left to right) show nodes of: Hartree-Fock (red/dark gray); Multi-Pfaffian nodes (orange/medium gray); and the nodes of the CI wave function (yellow/light gray). The CI nodal surface is very close to the exact one (see text). The HF node clearly divides the space into four nodal cells while Pfaffian and CI wave functions partitioning leads to the minimal number of two nodal cells. The changes in the nodal topology occur on the appreciable spatial scale of the order of 1 a.u.

The Fig. 15 shows the example of nodal structure of oxygen atom ( $1s^2$  electrons were removed by pseudopotential). Here we compare the nodal surfaces of HF (no pairing), MPF Pfaffian (STU pairing) and a high accuracy CI wave function with more than 3000 determinants, which gives essentially exact fermion nodes (i.e., 99.8(3)% of correlation energy in fixed-node DMC). From the figure, it is clear that the changes in the nodal surfaces are significant—the most important one being the elimination of artificial four nodal cells resulting from the independence of spin-up and -down channels in HF. The Pfaffian smooths-out the crossings and fuses the compartments of the same sign into the single ones. These topology changes therefore lead to the

minimal number of two nodal cells (for details see Sec. 3). However, the MPF Pfaffian nodes are noticeably different close to nucleus and in the tails and the amount of missing correlation energy is still non-negligible.

#### **4.5 Conclusions**

The proposed Pfaffians with singlet and triplet pairs together with unpaired orbitals provides a variationally rich and compact wave functions. They offer significant and systematic improvements over commonly used Slater determinant-based wave functions and can be calculated as easily as determinants using the presented formulas. We have also shown relationships between HF, BCS(AGP) and Pfaffian wave functions. Clearly, Pfaffian pairing wave functions are able to capture a large fraction of missing correlation energy with consistent treatment of both spin-polarized and unpolarized pairs. We have explored Pfaffian expansions which enabled us to capture additional correlation and map them onto the equivalent CISD wave functions. Finally, the gains in correlation energy for Pfaffians come from improved fermion nodes, which are very close to the exact ones and exhibit the correct topology with the minimal number of two nodal cells.

## 5 Backflow Correlations in Slater and Pfaffian Wave Functions

### 5.1 Introduction

Another approach for improvement of the nodal accuracy of variational trial wave functions is to employ backflow correlations. They were first introduced by Feynman and Cohen [57] as a simple ansatz to describe correlations and excitations in liquid  $^4\text{He}$ . Since then, a number of authors [55, 58–62] showed that the backflow correlations are useful also in fermionic systems. Recently, the backflow transformation of electron coordinates has been applied to chemical (e.i., inhomogeneous) systems [63, 64, 117, 118]. This section contains detailed description of inhomogeneous backflow transformation for multi-determinantal and Pfaffian wave functions and their applications.

Let us try to briefly elucidate where the motivation for the backflow comes from. Consider the trial function of the familiar form which we used a number of times before  $\Psi_T(\mathbf{R}) = \det[\dots]e^{U_{corr}}$ , and let us apply the projector  $\exp(-\tau\mathcal{H})$  as given by

$$\exp(-\tau\mathcal{H})\Psi_T(\mathbf{R}) = \Psi_T(\mathbf{R}) - \tau\mathcal{H}\Psi_T(\mathbf{R}) + \mathcal{O}(\tau^2) \quad (156)$$

In order to analyze the first order term we re-arrange the terms in the Hamiltonian

$$\mathcal{H} = T + V_{ext} + V_{ee} = T + (V_{ext} + V_{mf}) + (V_{ee} - V_{mf}), \quad (157)$$

where  $V_{mf}$  is a one-particle mean-field approximation for  $V_{ee}$  which corresponds to the theory used for generation of orbitals. A closer look on the first order term then leads to the following

$$\begin{aligned} \mathcal{H}\det[\dots]e^{U_{corr}} = & e^{U_{corr}}(T + V_{ext} + V_{mf})\det[\dots] + \det[\dots](T + V_{ee} - V_{mf})e^{U_{corr}} \\ & - \nabla e^{U_{corr}} \cdot \nabla \det[\dots]. \end{aligned} \quad (158)$$

Note that the first two terms should be close to  $\Psi_T(\mathbf{R})E_{loc}(\mathbf{R})$ , where  $E_{loc}(\mathbf{R})$  is some function which is not very far from a constant. The reason is that in the first term the determinant solves the corresponding one-particle Hamiltonian, while in the second term the correlation is optimized to cancel out the difference  $V_{ee} - V_{mf}$ . The third term appears to be “spurious” since it is not compensated by any potential. It has a form of the dot product of two “fluxes”, one resulting from  $V_{ext} + V_{mf}$  while the other from the electron-electron interactions. These terms need to be “cancelled” out what requires modifying the wave function or including additional terms. Which types of terms would be effective? We consider two limiting cases. The first one corresponds to

$$|\nabla \det[\dots]| \gg |\nabla e^{U_{corr}}| \quad (159)$$

with that the “flux” from the determinant having much larger amplitude so that the system is strongly inhomogeneous—typical behaviour for chemical systems and solid state materials. For these cases the excited determinants, Pfaffians, etc, appear as the most obvious, if not the most efficient, way to cancel out the spurious component. On the other hand, if the inequality goes the other way

$$|\nabla \det[\dots]| \ll |\nabla e^{U_{corr}}| \quad (160)$$

then the system has density close to a constant or only mildly varying so that the electron-electron term dominates. This is typical for homogeneous electron gas, quantum liquids such  $^4\text{He}$  and  $^3\text{He}$ , and similar systems. Feynman and Cohen showed that in this case the backflow coordinates can partially compensate for the spurious term, basically by slightly displacing a given particle to reflect the influence of particles which happen to be nearby.

## 5.2 Backflow Wave Function Form

As already discussed in Sec. 2.4.2, our trial wave function has the form

$$\Psi_T(\mathbf{R}) = \Psi_A(\mathbf{X}) \times \exp[U_{corr}(\mathbf{R})], \quad (161)$$

where  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$  represents some quasi-particle coordinates dependent on all  $N$  electron positions  $\mathbf{R}$ . Further,  $\Psi_A$  is either (multi)-determinantal or Pfaffian wave function and  $U_{corr}$  is the Jastrow correlation factor both defined in the previous sections.

The quasi-coordinate of  $i$ th electron at position  $\mathbf{r}_i$  is given as

$$\begin{aligned} \mathbf{x}_i &= \mathbf{r}_i + \boldsymbol{\xi}_i(\mathbf{R}) \\ &= \mathbf{r}_i + \boldsymbol{\xi}_i^{en}(\mathbf{R}) + \boldsymbol{\xi}_i^{ee}(\mathbf{R}) + \boldsymbol{\xi}_i^{een}(\mathbf{R}), \end{aligned} \quad (162)$$

where  $\boldsymbol{\xi}_i$  is the  $i$ th electron's backflow displacement divided to the contributions from one-body (electron-nucleus), two-body (electron-electron) and three-body (electron-electron-nucleus) terms. They can be further expressed as

$$\boldsymbol{\xi}_i^{en}(\mathbf{R}) = \sum_I \chi(r_{iI}) \mathbf{r}_{iI} \quad (163)$$

$$\boldsymbol{\xi}_i^{ee}(\mathbf{R}) = \sum_{j \neq i} u(r_{ij}) \mathbf{r}_{ij} \quad (164)$$

$$\boldsymbol{\xi}_i^{een}(\mathbf{R}) = \sum_I \sum_{j \neq i} [w_1(r_{ij}, r_{iI}, r_{jI}) \mathbf{r}_{ij} + w_2(r_{ij}, r_{iI}, r_{jI}) \mathbf{r}_{iI}], \quad (165)$$

where  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$  and  $\mathbf{r}_{iI} = \mathbf{r}_i - \mathbf{r}_I$ . The  $\chi$ ,  $u$  and  $w_1$  with  $w_2$  terms are similar to one, two and three-body Jastrow terms and are further expanded as

$$\chi(r) = \sum_k c_k a_k(r), \quad (166)$$

$$u(r) = \sum_k d_k b_k(r), \quad (167)$$

$$w_{1,2}(r_{ij}, r_{iI}, r_{jI}) = \sum_{klm} g_{klm} a_k(r_{iI}) a_l(r_{jI}) b_m(r_{ij}). \quad (168)$$

The one dimensional basis functions  $\{a\}$  and  $\{b\}$  are chosen as Gaussians with the center in origin or as polynomial Padé functions (Sec. 2.4.3) to preserve the electron-electron [Eqs. (49) and (50)] and electron-nucleus [Eq. (48)] cusp conditions. The set of variational parameters  $\{c\}$ ,  $\{d\}$  and  $\{g\}$  is minimized with respect to mixture of energy and variance (for details, see Sec. 7.2). In addition, all electron-electron coefficients ( $\{d_k\}$  and  $\{g_{klm}\}$  with fixed  $k$  and  $l$ ) are allowed to be different for spin-like and for spin-unlike electron pairs.

### 5.3 Applications of Backflow Correlated Wave Functions

In the remainder of this section we present VMC and DMC results obtained with above implementation of backflow correlations for determinant and Pfaffian wave functions.

#### 5.3.1 Homogeneous electron gas

We benchmark our implementation of the backflow correlations on the homogeneous electron gas (HEG) system. For comparison purposes, we choose to study the HEG system of 54 unpolarized electrons in the simple cubic simulation cell with periodic boundary conditions (the system was studied before several times [64, 119, 120]). Due to the homogeneity of the problem, the backflow displacement in Eq. (164) has only one non-zero component:  $\xi_i^{ee}$ . We conveniently choose  $u(r)$  to be expanded in the basis of several polynomial Padé functions to preserve cusp conditions with cutoff equal to half of the simulation cell. Finally, our Jastrow correlation factor contains only electron-electron terms.

In Table 5 we show our results for the following three electron densities of  $r_s = 1, 5$  and 20. The First important result is that at the normal density ( $r_s = 1$ ), the size of correlation (e.i.,  $E_{DMC} - E_{HF}$ ) is relatively small, but at lower densities (larger  $r_s$ ), the correlation energy is larger fraction of total energy and becomes more important. Second, we observe that the HF and Slater–Jastrow (SJ) fixed-node DMC energies are in very good agreement with previous results [64, 119, 120]. However, due to the omission of higher order correlations from Jastrow factor and also from backflow displacement, we obtain the expected higher VMC energies and variances for SJ and backflow displaced SJ (SJBF) trial wave functions. Nevertheless, our fixed-node DMC energies for SJBF trial wave functions closely match the results of Kwon *et al.* [119], and only slightly deviate at  $r_s = 20$  from results of Rios *et al.* [64].

#### 5.3.2 Carbon atom and dimer

The backflow correlations in single-determinant Slater–Jastrow (SJBF) [63, 64, 117] as well as in multi-determinant Slater–Jastrow (CIBF) [118] trial wave functions were recently applied to chemical systems. We extend these studies by including the backflow into the Pfaffian–Jastrow (PFBF) pairing wave functions. Below is a brief discussion of our implementation and results for carbon atom and dimer systems.

We employ the inhomogeneous backflow given by Eq. (163), with the functions  $u$  and  $w_{1,2}$  are allowed to be spin-dependent. The  $\chi$  and  $u$  functions are expanded in the 11 Gaussian basis functions, while the three-body functions  $w_{1,2}$  were limited to a product of  $4 \times 4 \times 4$  Gaussians.

The numerical results are summarized in Tables 6 and 7. The backflow correlations are able to capture additional few percent of the correlation energy for both Slater–Jastrow and Pfaffian–Jastrow wave functions. Another important feature of backflow is 20-30% decrease in variances of local energy with respect to the wave functions without backflow correlations. We find that for the fully optimized backflow, the spin-unlike electron-electron functions are almost order of magnitude larger than spin-like ones as well as electron-nucleus functions. The gains are systematic with increasing number of parameters, however we do not find the three-body terms as important as reported in previous study [64]. This difference can be attributed to two main reasons—we use a different basis to expand the three-body functions  $w_{1,2}$  and we also eliminate

Tab. 5. VMC and fixed-node DMC energies per electron and variances of local energies for various trial wave functions (S, Slater; SJ, Slater–Jastrow; SJBF, backflow correlated SJ) for 3D unpolarized HEG of 54 electrons.

$r_s$	Method	WF	$E/N[\text{a.u./electron}]$	$\sigma^2[\text{a.u.}^2]$
1.0	HF	S	0.56925(2)	19.3(1)
	VMC	SJ	0.53360(4)	1.26(4)
		SJBF	0.53139(4)	0.81(4)
	DMC	SJ	0.53087(4)	-
		SJBF	0.52990(4)	-
5.0	HF	S	-0.056297(7)	0.776(4)
	VMC	SJ	-0.075941(6)	0.0636(1)
		SJBF	-0.078087(4)	0.0236(1)
	DMC	SJ	-0.07862(1)	-
		SJBF	-0.07886(1)	-
20.0	HF	S	-0.022051(2)	0.0482(1)
	VMC	SJ	-0.031692(2)	0.000711(4)
		SJBF	-0.031862(1)	0.000437(1)
	DMC	SJ	-0.031948(2)	-
		SJBF	-0.032007(2)	-

atomic cores by pseudopotentials. It is plausible that for systems with core electrons the three-body correlations are more important due to the strong variations of orbitals close to the nucleus.

Finally, let us discuss the difference between the two systems with respect to missing correlation energy. For the C atom, we have shown previously [1] that less than 100 determinants gives more than 99% of correlation energy ( $E_{corr}$ ). The C dimer’s fixed node errors are more pronounced, since the 148 determinants with re-optimized weights give only 97.5(1)% in a close agreement with recent calculations by Umrigar *et al.*[84]. Employing backflow correlations for our 148 determinant CI-Jastrow wave function gives no apparent gain in  $E_{corr}$  except for decrease in the variance of local energy. The improvement for the Pfaffian-Jastrow wave function is also very modest (less than 1%). Our results suggest that to reach beyond 99% of correlation one still needs complicated multi-reference wave functions, even after including quite general forms of the backflow correlations.

## 5.4 Conclusions

We have benchmarked our backflow correlated wave functions on the HEG system and achieved high level of agreement with previous results. Further, we have applied the Pfaffian pairing wave functions to chemical systems. The results for to testing cases of carbon pseudo atom and its dimer show promising gains in correlations energies, decreases in variances and improvements in the nodal structure.



Tab. 6. VMC and DMC energies and variances of local energy for Slater–Jastrow (SJ), Pfaffian–Jastrow (PF) and CI–Jastrow (CI) trial wave functions with backflow (BF) correlations for C atom. Notation for backflow parameters: 2B, electron-nucleus and electron-electron terms; 3B, all electron-electron-nucleus terms; 23B for all terms together.

Method	WF	$N_\mu$	$N_\eta$	$N_{\theta_1}$	$N_{\theta_2}$	$N_p$	E [a.u.]	$\sigma^2$ [a.u. <sup>2</sup> ]	$E_{corr}$ [%]
HF	S	-	-	-	-	-	-5.31471	-	0.0
VMC	SJ	-	-	-	-	-	-5.3990(1)	0.0677	81.8(1)
	SJBF2B	11	22	-	-	33	-5.4011(2)	0.0544	83.8(2)
	SJBF3B	-	-	128	128	256	-5.4023(3)	0.0504	85.0(3)
	SJBF23B	4	8	128	128	268	-5.4020(2)	0.0498	84.7(2)
	PF	-	-	-	-	-	-5.4083(2)	0.0626	90.8(2)
	PFBF2B	11	22	-	-	33	-5.4097(1)	0.0427	92.1(1)
	PFBF23B	4	8	128	128	268	-5.4107(1)	0.0411	93.1(1)
	CI <sup>a</sup>	-	-	-	-	-	-5.4127(1)	0.0447	95.0(1)
	CIBF2B	11	22	-	-	33	-5.4131(3)	0.0427	95.4(3)
	CIBF23B	4	8	128	128	268	-5.4140(1)	0.0342	96.3(1)
DMC	SJ	-	-	-	-	-	-5.4065(3)	-	89.0(3)
	SJBF2B	11	22	-	-	33	-5.4090(3)	-	91.5(3)
	SJBF3B	-	-	128	128	256	-5.4085(3)	-	91.0(3)
	SJBF23B	4	8	128	128	268	-5.4094(3)	-	91.8(3)
	PF	-	-	-	-	-	-5.4137(3)	-	96.0(3)
	PFBF2B	11	22	-	-	33	-5.4145(3)	-	96.8(3)
	PFBF23B	4	8	128	128	268	-5.4152(3)	-	97.5(3)
	CI	-	-	-	-	-	-5.4178(1)	-	100.0(1)
	CIBF2B	11	22	-	-	33	-5.4177(3)	-	99.9(3)
	CIBF23B	4	8	128	128	268	-5.4174(2)	-	99.6(2)
Est.	Exact	-	-	-	-	-	-5.417806	-	100.0

<sup>a</sup>Wave function consists of 100 determinants re-optimized in VMC.

Tab. 7. Slater–Jastrow (SJ), Pfaffian–Jastrow (PF) and CI–Jastrow (CI) wave functions with backflow (BF) correlations for C dimer. Notation is the same as in Table 6.

Method	WF	$N_\mu$	$N_\eta$	$N_{\theta_1}$	$N_{\theta_2}$	$N_p$	E [a.u.]	$\sigma^2$ [a.u. <sup>2</sup> ]	$E_{corr}$ [%]
HF	S	-	-	-	-	-	-10.6604	-	0.0
VMC	SJ <sup>a</sup>	-	-	-	-	-	-10.9936(4)	0.179	81.7(1)
	SJBF2B	11	22	-	-	33	-11.0012(3)	0.144	83.5(1)
	SJBF23B	4	8	128	128	268	-11.0014(2)	0.141	83.6(1)
	PF <sup>b</sup>	-	-	-	-	-	-11.0171(2)	0.160	87.4(1)
	PFBF2B	11	22	-	-	33	-11.0223(3)	0.123	88.7(1)
	PFBF23B	4	8	128	128	268	-11.0223(2)	0.128	88.7(1)
	CI <sup>c</sup>	-	-	-	-	-	-11.0420(4)	0.112	93.6(1)
	CIBF2B	11	22	-	-	33	-11.0440(3)	0.100	94.0(1)
	CIBF23B	4	8	128	128	268	-11.0438(3)	0.123	94.0(1)
DMC	SJ	-	-	-	-	-	-11.0227(2)	-	88.8(1)
	SJBF2B	11	22	-	-	33	-11.0269(4)	-	89.9(1)
	SJBF23B	4	8	128	128	268	-11.0280(3)	-	90.1(1)
	PF	-	-	-	-	-	-11.0419(9)	-	93.5(2)
	PFBF2B	11	22	-	-	33	-11.0443(6)	-	94.1(2)
	PFBF23B	4	8	128	128	268	-11.0447(3)	-	94.2(1)
	CI	-	-	-	-	-	-11.0579(5)	-	97.5(1)
	CIBF2B	11	22	-	-	33	-11.0580(4)	-	97.5(1)
	CIBF23B	4	8	128	128	268	-11.0585(5)	-	97.7(1)
Est.	Exact	-	-	-	-	-	-11.068(5)	-	100.0

<sup>a</sup>Slater determinant contains PBE DFT orbitals.

<sup>b</sup>Same PBE DFT orbitals are used also in PF wave function.

<sup>c</sup>Uses natural orbitals with weights of the 148 determinants re-optimized in VMC.

## 6 Example of Large Scale Application: FeO Solid

As an example of calculations, which are feasible at present using available both software and hardware resources, we briefly report our recent study of the FeO solid.

FeO solid is a very interesting system which has been calculated a number of times before by several approaches. Nevertheless, it is difficult to claim that the problems and research questions have been all “solved” and, in fact, our understanding of this and many other transition metal compounds is still rather limited. FeO belongs to the group of other paradigmatic transition metal (TM) oxides such as MnO, CoO and NiO. All of these have a rather simple rocksalt (NaCl type) structure denoted as B1, with very small deformations at lower temperatures resulting from effects such as magnetostriction. Since the transition elements in these systems have open d-shells, such solids should be nominally metals. However, for a long time it has been known and experiments have shown that these systems are *large gap insulators*. In addition, they exhibit antiferromagnetic ordering with the Neel temperatures of the order of a few hundred Kelvins. Interestingly, the gap remains present even at high temperatures when there is no long-range magnetic order. How is it possible? Using qualitative classification introduced by Zaanen and coworkers[121], there are basically two relevant and somewhat competing mechanisms: Mott-Hubbard and charge transfer. In the Mott-Hubbard picture, the gap opens because of large Coulomb repulsion associated with the double occupancy of the strongly localized d-states. The spin minority bands are pushed up in the energy, leading to gap opening. In the charge transfer mechanism, 4s electrons of the transition metal atom fill the unoccupied p-states of oxygen. The electronic structure has an almost ionic  $\text{TM}^{+2}\text{O}^{-2}$  character, with resulting gap opening as well. In the real materials both of these mechanisms are present and therefore the systems exhibit insulating behavior. The antiferromagnetism happens on the top of this and is caused by weak super-exchange interactions between the moments of neighbouring transition metal atoms mediated by bridging oxygens. We will not go into further details, suffice it to say that the electron correlations are very large and crucially affect many properties of these materials.

If one applies standard DFT approaches to FeO, the results are highly unsatisfactory. First, the atomic structure which comes out is not correct. Instead of the rocksalt antiferromagnet another structure, so-called iB8, appears to have the lowest energy at equilibrium conditions. Interestingly, iB8 is actually a high pressure phase of FeO. This discrepancy is surprising since DFT methods typically provide correct equilibrium structures and very reasonable geometric parameters such as lattice constants and others. Another problem appears in the electronic structure since for the correct geometry DFT predicts a *metallic* state. More sophisticated methods beyond DFT have been applied to this system in order to reconcile some of the results with experiments, nevertheless, a number of questions remain unanswered. For example, at high pressures, FeO undergoes a structural transition into the iB8 phase, however, the value of the transition pressure which agrees with experiment is difficult to obtain using the mainstream approaches. Similar transition appears in MnO, where recent benchmarking of several DFT approaches provided transition pressure estimates between 65 and 220 GPa, i.e., more than 300% spread in the predictions [122]. Large discrepancies with experiment exist also for the CoO crystal and more complicated transition metal compounds.

We have carried out QMC calculations of FeO using supercells with periodic boundary conditions to model the infinite solid [123]. Several supercell sizes were calculated and k-point sampling of the Brillouin zone was carried out by the so-called twist averaging with the purpose

Tab. 8. Comparison of the calculated structural properties of FeO solid in DFT and in the fixed-node DMC with experimental data. The energy difference  $E_{iB8}-E_{B1}$  is evaluated at the experimental lattice constant 4.334 Å.

Method/quantity	DFT/PBE	FN-DMC	Exper.
$E_{iB8}-E_{B1}$ [eV]	-0.2 [126]	0.5(1)	> 0
cohesion energy [eV]	~ 11	9.66(4)	9.7 [127]
lattice constant [Å]	4.28 [124]	4.324(6)	4.334[128]
bulk modulus [GPa]	180 [124]	170(10)	~ 180 [125]
band gap [eV]	~ 0	2.8(3)	2.4 [129]

of eliminating finite size effects. The core electrons were replaced by pseudopotentials for both Fe (Ne-core) and O (He-core). The largest simulated supercells had more than 300 valence electrons and the total energies were sizeable due to the presence of “semicore” 3s and 3p states of Fe in the valence space, what have made the calculations rather demanding. The wave function had the Slater-Jastrow form and the orbitals were obtained from unrestricted (spin-polarized) calculations within DFT with hybrid functionals and HF.

In Tab. 8 are shown the QMC calculated equilibrium parameters. Note that QMC identifies the correct equilibrium structure and provides very accurate value of the cohesive energy. Cohesive energies are very difficult to calculate since the DFT methods show typical bias of 15-30%. At present, there is basically no other method besides QMC which can get this level of accuracy. Note also good agreement with experiments for the other quantities including the band gap. In QMC the band gap is calculated as a difference of two total energies— ground state and excited state, where the excited state is formed by promoting an electron from the top valence band into the conduction band.

In addition, the equations of state have been calculated for both the equilibrium structure and also for the high pressure phase, see Fig. 16. The estimated transition pressure is 65(5) GPa, at the lower end of the experimental range 70-100 GPa [124, 125].

Finally, considering that only a simple trial wave function was employed, the results are remarkable and very encouraging. Note that the calculations do not have any free non-variational parameters. It is simply the best possible solution within the trial function nodes and the given Hamiltonian.

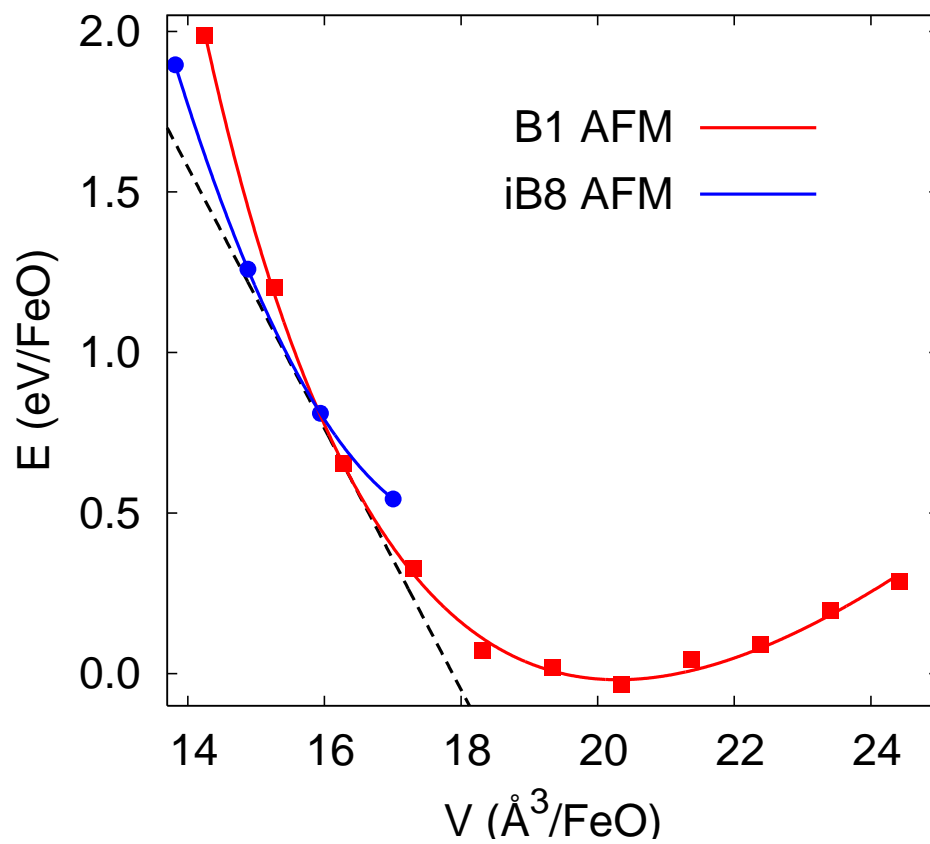


Fig. 16. FN-DMC Energy as a function of volume for FeO for B1 (red squares) and iB8 (blue circles) phases. Lines are fits with Murnaghan equation of state.

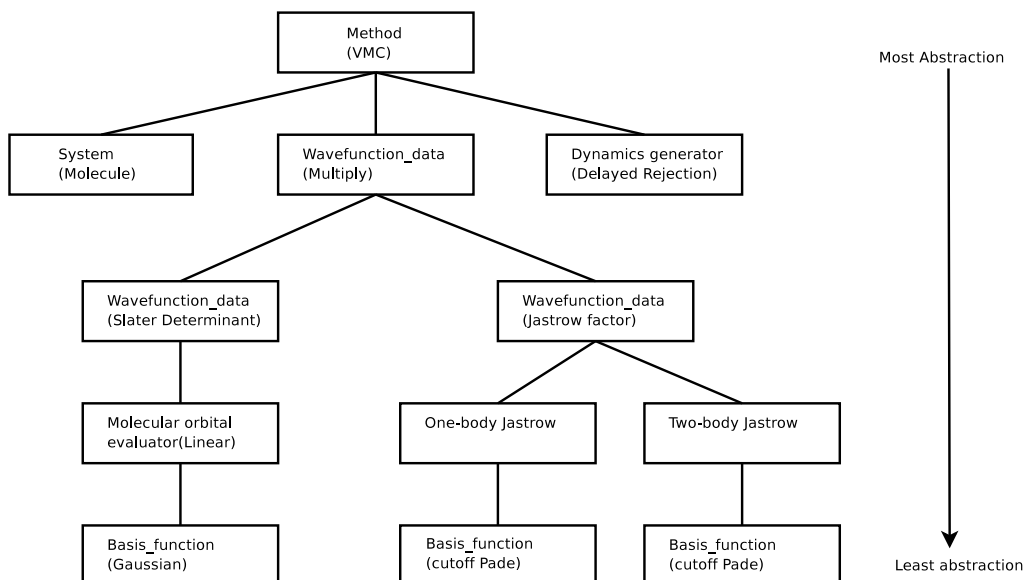


Fig. 17. Calculation structure for the VMC method on a molecule using a Slater-Jastrow wave function.

## 7 QWalk: A Quantum Monte Carlo Program for Electronic Structure

The solution of the electronic structure of many electron systems in quantum Monte Carlo requires the use of computers and appropriate software tools. For the practitioners of the QMC methods, the computational aspects such as the use of efficient algorithms, flexible code organization and parallel efficiency are important, since they influence which types of calculations and which types of physical effects are feasible to study. This section is therefore aimed to introduce a reader to one of the existing QMC packages, Open Source QWalk code available under the GPL licence [53] ([www.qwalk.org](http://www.qwalk.org)).

This section is organized as follows. In Sec. 7.1 we overview the program's organization and structure. The description of the implementation of wave function optimization is in Sec. 7.2.

### 7.1 Organization and Implementation

Qwalk ("Quantum Walk") is a computational package for QMC calculations of electronic structure systems such as molecules and solids. However, others systems can be calculated such as homogeneous electron gas, BCS-BEC condensates, 1D rings with persistent currents, etc. The code is written in a combination of object-oriented and procedural techniques. The object-oriented approach is coarse-grained, creating independent sections of code that are written efficiently in a procedural fashion. It is extremely modular; almost every piece can be removed and replaced with another. For example, a contributor of a new module only has to change one line in the main code to allow its use. This allows for flexibility while keeping the code base relatively simple and separable. The modular structure also allows for partial rewrites of the code without

Tab. 9. The central objects of the code and their physical correspondents

Module name	Mathematical object
System	parameters and form of the Hamiltonian
Sample point	$\mathbf{R}$ , the integration variables
Wave function type	Wave function ansatz
Wave function	$\Psi_T(\mathbf{R})$ , $\nabla \Psi_T(\mathbf{R})$ , $\nabla^2 \Psi_T(\mathbf{R})$
Dynamics generator	Metropolis trial move (Green's function)

worrying about other parts. In fact, each major module has been rewritten several times in this manner as we add new features and re-optimize the code flexibility and performance.

The modules form a tree of successive abstractions (Fig. 17). At the top of the tree is the QMC method, VMC in this case. It works only in terms of the objects directly below it, which are the concepts of System, Wave function data, etc. (see Table 9). These in turn may have further abstractions below them. The highest wave function object is of type “Multiply”, which uses two wave function types to create a combined wave function. In this case, it multiplies a Slater determinant with a Jastrow correlation factor to form a Slater-Jastrow function. Since various wave functions can be plugged-in, the Slater determinant can be replaced with any antisymmetric function. The same applies to the Jastrow factor. The type is listed along with the specific instant of that type in parenthesis. At each level, the part in parenthesis could be replaced with another module of the same type. The example of the flowchart of the whole program is given in Fig. 19.

## 7.2 Optimization Methods in QWalk

Historically, the first implementation of minimization method in QWalk code was the OPTIMIZE method, based on the quasi-Newton minimizer. It has proved effective for variance minimization of linear and nonlinear Jastrow parameters, however, the energy minimizations require very large configuration samples.

This deficiency was partially fixed by the OPTIMIZE2 method, which uses modified Hessians and gradients of variance and energy of Ref. [74] together with the Levenberg-Marquardt minimization algorithm. The analytical gradients and Hessians with respect to selected variational parameters were implemented (determinantal and orbital coefficients, but also Pfaffian and pair orbital coefficients). OPTIMIZE2 has proved to be very effective method and has been used for most of recent calculations.

In an attempt to remove the dependence of OPTIMIZE2 method on the fixed MC sample, the NEWTON\_OPT method was implemented as well. Its principal advantage is that the energy and variance are obtained from a small MC run performed after each optimization step.

Besides the variance and energy minimizations, it is possible to minimize several other *objective or cost functions*. Their incomplete list can be found in Table 10. As was recently pointed out by Umrigar and Filippi [74], the wave functions optimized by minimization of a mix of energy and variance produce almost as low energy as energy optimized wave functions but with

Tab. 10. List of some useful cost functions for wave function optimizations with indicated implementations in QWalk [53].

Cost Function	Minimized quantity	QWalk implementation		
		OPTIMIZE	OPTIMIZE2	NEWTON_OPT
Variance	$\langle (E_L - \bar{E})^2 \rangle$	✓	✓	✓
Energy	$\bar{E}$	✓	✓	✓
Mixed	$x\bar{E} + (1-x)\langle (E_L - \bar{E})^2 \rangle$	✓	✓	✓
Absolute value	$\langle  E_L - \bar{E}  \rangle$	✓		
Lorentz	$\langle \ln(1 + (E_L - \bar{E})^2/2) \rangle$	✓		
Ratio	$\frac{\langle (E_L - \bar{E})^2 \rangle}{\bar{E}}$			
Overlap	$\frac{\int \Psi_T \Psi(\tau \rightarrow \infty)}{\int \Psi_T^2}$			

lower variances.

We illustrate this for the ground state wave function of the  $N_2$  molecule (see Fig 18). The employed NEWTON\_OPT method minimized 23 Jastrow parameters (the curvatures of correlation basis functions and all linear coefficients) of single determinant Slater-Jastrow wave function on set of 56000 walkers. The minimized energy from mixed objective function is almost as good as from the energy optimization, while the corresponding dispersion  $\sigma$  falls in between  $\sigma$  from energy minimization and variance minimization. The wave functions obtained from mixed objective functions therefore appear as the the most efficient compromise for further DMC calculations.

### 7.3 Conclusion

QWalk represents a state of the art, usable, and extensible program for performing quantum Monte Carlo calculations on several types of quantum systems. It is able to handle medium to large systems of electrons; the maximum size is mostly limited by the available computer time. It works in parallel very efficiently, so it can take advantage of large clusters, multi-core computers, etc. Since QWalk is available without charge and under the GNU Public license, it is hoped that it will help bring both development and use of quantum Monte Carlo methods to a wide audience. Due to its modular form it is straightforward to expand the QWalk's applicability to quantum systems beyond the electron-ion Hamiltonians in continuous space. With some knowledge of C++ it is easy to modify the system module to incorporate other types of interactions and to expand the one-particle and pair orbitals using the coded basis functions.



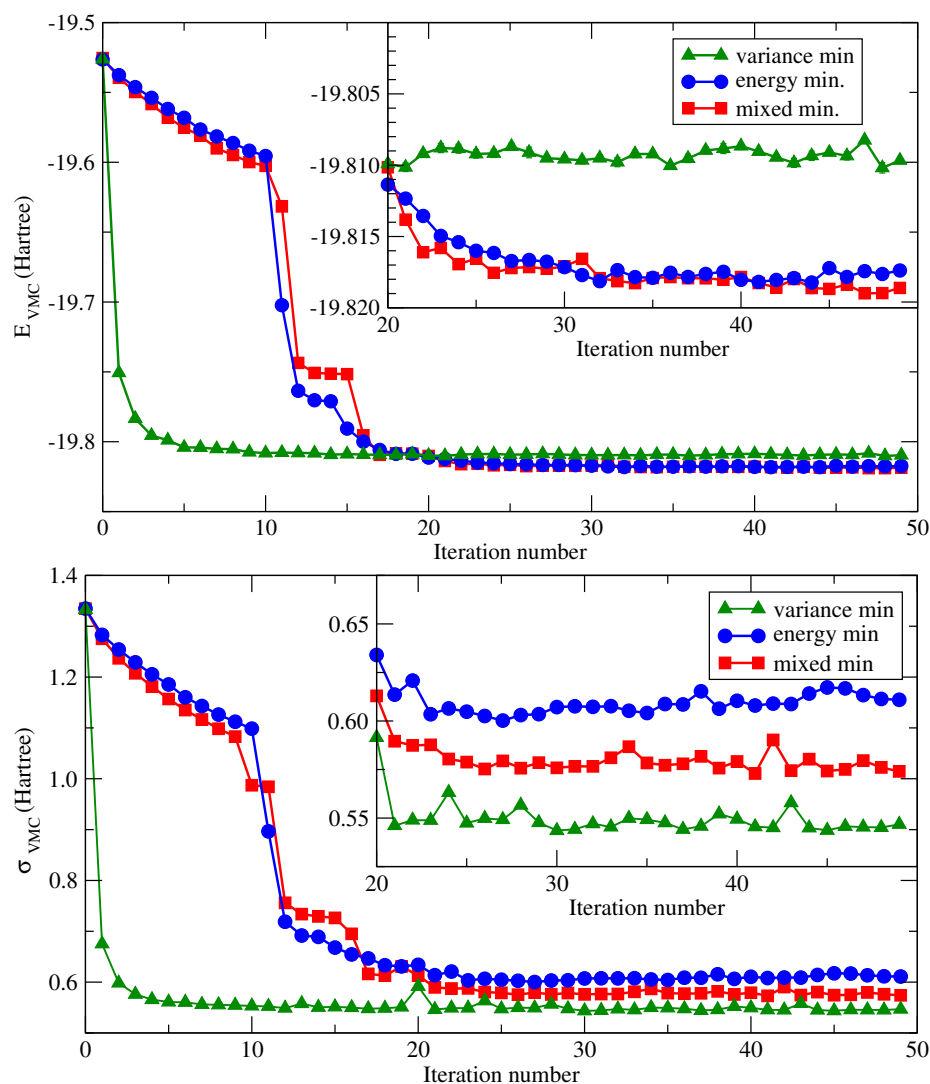


Fig. 18. Energy  $E$  and dispersion of the local energy  $\sigma$  of the  $\text{N}_2$  molecule versus the minimization step number. Notation: minimization of variance (green triangles), energy (blue circles) and 95% mixture of energy and 5% of variance (red squares). Upper figure: Energy versus the iteration number. The error bars are of the symbol size or smaller. Inset: the later iterations on expanded scale. Note the variance minimized energy being higher by almost 10 mH. Lower figure: Same as the upper figure but for the dispersion of local energy  $\sigma$  rather than energy. Inset shows that the mixture minimized  $\sigma$  falls in between  $\sigma$  from energy minimization and variance minimization.

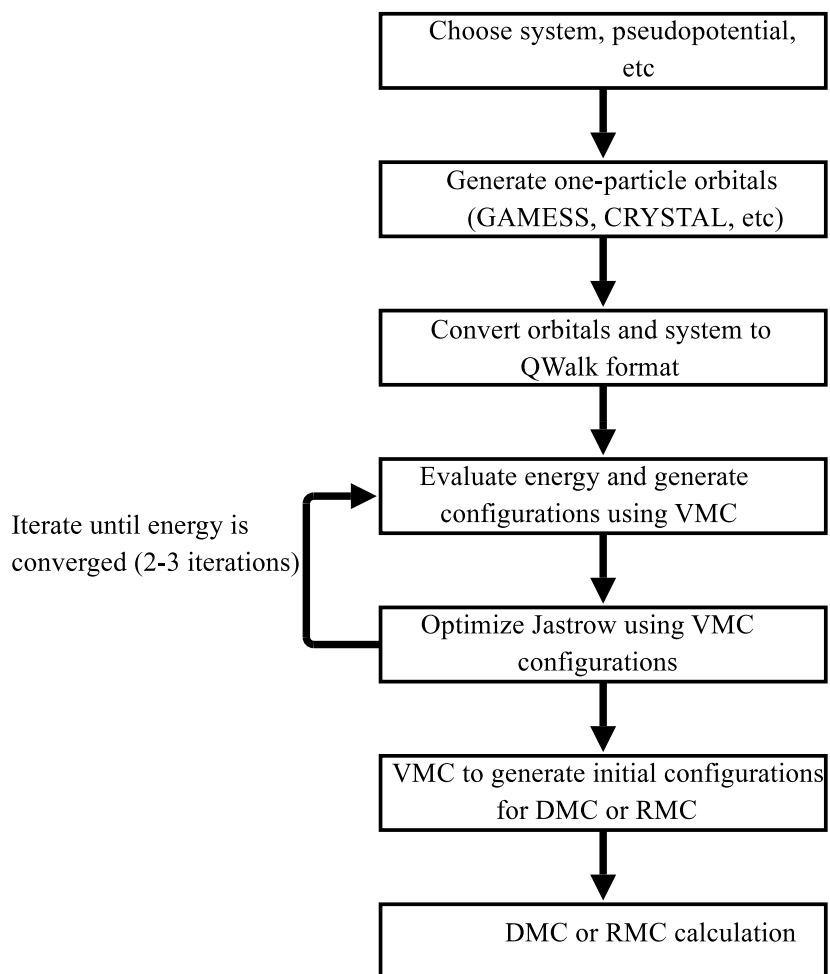


Fig. 19. Flow of a QMC calculation

## 8 Summary

We have presented an overview of QMC methods and some of the recent developments in the analysis of the fermion nodes, constructions of pairing wave functions and explorations of backflow correlations. Most of the applications were using small systems both because of educational purposes and in some cases because these were the first benchmarks of the new developments. QMC methodology has proved to be a very powerful technique for studies of quantum systems and also real materials as illustrated on the FeO solid calculations. In essence, QMC has a number of advantages when compared with other approaches:

- direct and explicit many-body wave function framework for solving the stationary Schrödinger equation;
- favorable scaling with the systems size;
- wide range of applicability;
- the fixed-node approximation which enables to obtain 90-95% of the correlation effects;
- scalability on parallel machines;
- new insights into the many-body phenomena.

The remaining fixed-node errors account for small, but still important fraction of correlations. As we have shown, further improvement calls for better understanding of the nodes and their efficient descriptions. It is encouraging that studies of the nodal properties brought new understanding such as the two nodal cell property for generic fermionic states and revealed new research directions which are yet to be explored.

We have shown that the pairing orbitals and Pfaffian wave functions show significant improvements in providing more systematic results and higher accuracy. The key virtue of the pairing functions is that they exhibit the correct nodal topologies and as such are capable to describe quantum condensates and pairing effects. As we explained, the backflow has perhaps more limited scope of usefulness and appears to be effective especially for homogeneous or marginally inhomogeneous systems. Many generalizations of QMC exist. For example, for systems with magnetic fields and stationary currents the fixed-node methods can be generalized to the fixed-phase QMC [130, 131]. For finite temperature equilibrium properties Path Integral Monte Carlo has been used very successfully especially for bosonic quantum liquids.

The overall impact of the electronic structure QMC methods has been so far mainly in providing high accuracy and benchmark results which were used by other methods as a reference. We believe that the impact of QMC methods is on the rise and their use will gradually become more routine. Already at present, there are cases for which developing elaborate mean-fields take more human time than direct solutions by many-body stochastic approaches. However, as we stated in the introduction, much needs to be done to make the methods more insightful, efficient and easy to use.

Acknowledgements. Support by NSF EAR-05301110 and DMR-0804549 grants and by DOE DE-FG05-08OR23336 Endstation grant is gratefully acknowledged. We acknowledge also allocations at ORNL through INCITE and CNMS initiatives as well as allocations at NSF NCSA and TACC centers. We would also like to thank L.K. Wagner and J. Kolorenč for discussions.

## References

- [1] M. Bajdich, L. Mitás, G. Drobný, L. K. Wagner, and K. E. Schmidt, Phys. Rev. Lett. **96**, 130201 (2006).
- [2] M. Bajdich, L. Mitás, L. K. Wagner, and K. E. Schmidt, Phys. Rev. B **77**, 115112 (2008).
- [3] J. Grossman, J Chem. Phys. **117**, 1434 (2002).
- [4] S. Manten and A. Lüchow, J. Chem. Phys. **115**, 5362 (2001).
- [5] W. M. C. Foulkes, L. Mitás, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).
- [6] A. J. Williamson, R. Q. Hood, and J. C. Grossman, Phys. Rev. Lett. **87**, 246406 (2001).
- [7] A. J. Williamson, G. Rajagopal, R. J. Needs, L. M. Fraser, W. M. C. Foulkes, Y. Wang, and M.-Y. Chou, Phys. Rev. B **55**, R4851 (1997).
- [8] R. M. Martin, *Electronic Structure, Basic theory and practical methods* (Cambridge University Press, Cambridge, UK, 2004).
- [9] W. Heitler and F. London, **44**, 455 (1927).
- [10] D. R. Hartree, Proc. Cambridge Philos. Soc. **24**, 89 (1928).
- [11] E. A. Hylleraas, Z. Phys. **65**, 209 (1930).
- [12] E. A. Hylleraas, Z. Phys. **54** (1929).
- [13] V. Fock, Z. Phys **61** (1930).
- [14] W. Pauli, Z. Phys. **41**, 81 (1927).
- [15] J. C. Slater, Phys. Rev. **35**, 210 (1930).
- [16] F. Bloch, Z. Phys. **52**, 555 (1929).
- [17] E. Wigner and F. Seitz, Phys. Rev. **43**, 804 (1933).
- [18] E. Wigner and F. Seitz, Phys. Rev. **46**, 509 (1934).
- [19] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
- [20] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
- [21] J. Čížek, Adv. Chem. Phys. **14**, 35 (1969).
- [22] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover, 1996).
- [23] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
- [24] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
- [25] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
- [26] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
- [27] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
- [28] T. Ando, Phys. Rev. B **13**, 3468 (1976).
- [29] A. Zangwill and P. Soven, Phys. Rev. A **21**, 1561 (1980).
- [30] O. Gunnarsson, M. Jonson, and B. I. Lundqvist, Phys. Rev. B **20**, 3136 (1979).
- [31] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
- [32] V. I. Anisimov, J. Zaanen, and O. K. Andersen, Phys. Rev. B **44**, 943 (1991).
- [33] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
- [34] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods* (Wiley, New York, 1986).
- [35] C. Filippi and C. J. Umrigar, Phys Rev B **61**, R16291 (2000).

- [36] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester Jr., *J. Chem. Phys.* **77**, 5593 (1982).
- [37] R. C. Grimm and R. G. Storer, *J. Comp. Phys.* **7**, 134 (1971).
- [38] D. M. Ceperley and M. H. Kalos, in *Monte Carlo Methods in Statistical Physics*, edited by K. Binder (Springer, Berlin, 1979), pp. 145–194.
- [39] C. Umrigar, M. Nightingale, and K. Runge, *J. Chem. Phys.* **99**, 2865 (1993).
- [40] J. B. Anderson, *J. Chem. Phys.* **63**, 1499 (1975).
- [41] J. B. Anderson, *J. Chem. Phys.* **65**, 4121 (1976).
- [42] J. W. Moskowitz, K. E. Schmidt, M. A. Lee, and M. H. Kalos, *J. Chem. Phys.* **77**, 349 (1982).
- [43] D. M. Ceperley and B. J. Alder, *J. Chem. Phys.* **81**, 5833 (1984).
- [44] D. M. Arnaw, M. H. Kalos, M. A. Lee, and K. E. Schmidt, *J. Chem. Phys.* **77**, 5562 (1982).
- [45] S. Zhang and M. H. Kalos, *Phys. Rev. Lett.* **67**, 3074 (1991).
- [46] R. Assaraf, M. Caffarel, and A. Khelif, *J. Phys. A-Math. Theor.* **40**, 1181 (2007).
- [47] J. B. Anderson, C. A. Traynor, and B. M. Boghosian, *J. Chem. Phys.* **95**, 7418 (1991).
- [48] D. L. Diedrich and J. B. Anderson, *Science* **258**, 786 (1992).
- [49] R. Bianchi, D. Bressanini, P. Cremaschi, and G. Morosi, *J. Chem. Phys.* **98**, 7204 (1993).
- [50] M. Troyer and U.-J. Wiese, *Phys. Rev. Lett.* **94**, 170201 (2005).
- [51] T. Kato, *Commun. Pure Appl. Math.* **10**, 151 (1957).
- [52] R. T. Pack and W. B. Brown, *J. Chem. Phys.* **45**, 556 (1966).
- [53] L. K. Wagner, M. Bajdich, and L. Mitás, *J. Comp. Phys.* **228**, 3390 (2009).
- [54] R. Jastrow, *Phys. Rev.* **98**, 1479 (1955).
- [55] J. W. Moskowitz and K. E. Schmidt, *J. Chem. Phys.* **97**, 3382 (1992).
- [56] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, *Phys. Rev. Lett.* **60**, 1719 (1988).
- [57] R. P. Feynman and M. Cohen, *Phys. Rev.* **102**, 1189 (1956).
- [58] K. E. Schmidt, M. A. Lee, M. H. Kalos, and G. V. Chester, *Phys. Rev. Lett.* **47**, 807 (1981).
- [59] R. M. Panoff and J. Carlson, *Phys. Rev. Lett.* **62**, 1130 (1989).
- [60] Y. Kwon, D. M. Ceperley, and R. M. Martin, *Phys. Rev. B* **48**, 12037 (1993).
- [61] Y. Kwon, D. M. Ceperley, and R. M. Martin, *Phys. Rev. B* **50**, 1684 (1994).
- [62] Y. Kwon, D. M. Ceperley, and R. M. Martin, *Phys. Rev. B* **53**, 7376 (1996).
- [63] N. D. Drummond, P. L. Rios, A. Ma, J. R. Trail, G. G. Spink, M. D. Towler, and R. J. Needs, *J. Chem. Phys.* **124**, 224104 (2006).
- [64] P. L. Rios, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, *Phys. Rev. E* **74**, 066701 (2006).
- [65] D. M. Ceperley, *Proceedings of the Metropolis Symposium on The Frontiers of Quantum Monte Carlo*, *J. of Stat. Phys.* **43** (1986).
- [66] B. L. Hammond, P. J. Reynolds, and J. William A. Lester, *J. Chem. Phys.* **87**, 1130 (1987).
- [67] L. Mitás, E. L. Shirley, and D. M. Ceperley, *J. Chem. Phys.* **95**, 3467 (1991).
- [68] L. Mitás, *Phys. Rev. A* **49**, 4411 (1994).
- [69] H.-J. Flad and M. Dolg, *J. Chem. Phys.* **107**, 7951 (1997).
- [70] H. J. M. van Bemmelen, D. F. B. ten Haaf, W. van Saarloos, J. M. J. van Leeuwen, and G. An, *Phys. Rev. Lett.* **72**, 2442 (1994).

- [71] D. F. B. ten Haaf, H. J. M. van Bemmelen, J. M. J. van Leeuwen, W. van Saarloos, and D. M. Ceperley, Phys. Rev. B **51**, 13039 (1995).
- [72] M. Casula, C. Filippi, and S. Sorella, Phys. Rev. Lett. **95**, 100201 (2005).
- [73] R. Fletcher, Report R7125 AERE, Harwell, ENGLAND (1972), harwell routine VA10A.
- [74] C. J. Umrigar and C. Filippi, Phys Rev Lett **94**, 150201 (2005).
- [75] X. Lin, H. Zhang, and A. M. Rappe, J. Chem. Phys. **112**, 2650 (2000).
- [76] K. Madsen, H. B. Nielsen, and O. Tingleff, *Methods for non-linear least squares problems* (Informatics and Mathematical Modeling, Technical University of Denmark, 2004), 2nd ed.
- [77] G. Breit, Phys. Rev. **35**, 569 (1930).
- [78] E. A. Hylleraas, Z. Phys. **48** (1928).
- [79] J. B. Anderson, Phys. Rev. A **35**, 3550 (1987).
- [80] W. A. Glauser, W. R. Brown, W. A. Lester, Jr., D. Bressanini, B. L. Hammond, and M. L. Kozykowski, J. Chem. Phys. **97**, 9200 (1992).
- [81] W. M. C. Foulkes, R. Q. Hood, and R. J. Needs, Phys. Rev. B **60**, 4558 (1999).
- [82] D. Bressanini, D. M. Ceperley, and P. Reynolds, in *Recent Advances in Quantum Monte Carlo Methods 2*, edited by S. T. W. A. Lester Jr., S. M. Rothstein (World Scientific, Singapore, 2002).
- [83] D. Bressanini, G. Morosi, and S. Tarasco, J. Chem. Phys. **123**, 204109 (2005).
- [84] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, cond-mat/0611094 (2006).
- [85] M. Casula and S. Sorella, J. Chem. Phys. **119**, 6500 (2003).
- [86] M. Casula, C. Attaccalite, and S. Sorella, J. Chem. Phys. **121**, 7110 (2004).
- [87] D. M. Ceperley, J. Stat. Phys. **63** (1991).
- [88] L. Mitas, Phys. Rev. Lett. **96**, 240402 (2006).
- [89] L. Mitas, cond-mat/0605550 (2006).
- [90] K. Uhlenbeck, Amer. J. Math. **98**, 1059 (1976).
- [91] D. Bressanini and P. J. Reynolds, Phys. Rev. Lett. **95**, 110201 (2005).
- [92] M. Bajdich, L. Mitas, G. Drobný, and L. K. Wagner, Phys. Rev. B **72**, 075131 (2005).
- [93] I. Ovcharenko, A. Aspuru-Guzik, and W. A. Lester, Jr., J. Chem. Phys. **114**, 7790 (2001).
- [94] A. K. Wilson, T. V. Mourik, and T. H. Dunning, Jr., J. Mol. Struct.: THEOCHEM **388**, 339 (1997).
- [95] W. Kutzelnigg, in *Explicitly Correlated Wave Functions in Chemistry and Physics*, edited by J. Rychlewski (Kluwer, Dordrecht, 2003), pp. 44–82, and references therein.
- [96] K. Rektorys, *Survey of Applicable Mathematics* (Kluwer, Dordrecht, 1994).
- [97] J. Carlson, S.-Y. Chang, V. R. Pandharipande, and K. E. Schmidt, Phys. Rev. Lett. **91**, 050401 (2003).
- [98] J. P. Bouchaud and C. L’huillier, Europhys. Lett. **3**, 1273 (1987).
- [99] M. Bhattacharjee, Z. Phys. B **82** (1991).
- [100] A. Fabrocini, S. Fantoni, A. Y. Illarionov, and K. E. Schmidt, Phys. Rev. Lett. **95**, 192501 (2005), also in arXiv.org, nucl-th/0607034.
- [101] A. Cayley, Cambridge and Mathematical Journal **7**, 40 (1852), reprinted in *The collected*

- mathematical papers of Arthur Cayley*, Cambridge [Eng.] The University Press, Cambridge, vol. 2, pp. 19 (1889).
- [102] M. Nakahara, *Geometry, Topology and Physics* (Institute of Physics, Bristol, 2003).
  - [103] J. R. Stembridge, *Adv. Math.* **83** (1990).
  - [104] A. Cayley, *J. Reine Angew. Math.* **38**, 93 (1849), reprinted in *The collected mathematical papers of Arthur Cayley*, Cambridge [Eng.] The University Press, Cambridge, vol. 1, pp. 410-413 (1889).
  - [105] F. R. Gantmacher, *The Theory of Matrices*, vol. 1 (Chelsea, New York, 1959).
  - [106] G. Galbiati and F. Maffioli, *Discrete Applied Mathematics* **51**, 269 (1994).
  - [107] M. Bajdich, Ph.D. thesis, North Carolina State University (2007).
  - [108] J. P. Bouchaud, A. S. George, and C. L'huillier, *J. Phys. France(Paris)* **49** (1988).
  - [109] J. P. Bouchaud and C. L'huillier, *Physica C: Superconductivity* **153**, 1169 (1988).
  - [110] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, *Phys. Rev.* **108**, 1175 (1957).
  - [111] W. J. Stevens, H. Basch, and M. Krauss, *J. Chem. Phys.* **81**, 6026 (1984).
  - [112] M. Dolg, *Chemical Physics Letters* **250**, 75 (1996).
  - [113] M. W. Chase, Jr., *NIST-JANAF Tables 4*, vol. 9 (AIP, 1998).
  - [114] Y. B. Randall S. Urdahl and W. M. Jackson, *Chemical Physics Letters* **178**, 425 (1991).
  - [115] K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure: Constants of Diatomic Molecules*, vol. 4 (Van Nostrand Reinhold, New York, 1979).
  - [116] L. Bytautas and K. Ruedenberg, *J. Chem. Phys.* **122**, 154110 (2005).
  - [117] I. G. Gurtubay and R. J. Needs, *J. Chem. Phys.* **127**, 124306 (2007).
  - [118] M. D. Brown, J. R. Trail, P. L. Rios, and R. J. Needs, *J. Chem. Phys.* **126**, 224110 (2007).
  - [119] Y. Kwon, D. M. Ceperley, and R. M. Martin, *Phys. Rev. B* **58**, 6800 (1998).
  - [120] M. Holzmann, D. M. Ceperley, C. Pierleoni, and K. Esler, *Phys. Rev. E* **68**, 046707 (2003).
  - [121] J. Zaanen, G. A. Sawatzky, and J. W. Allen, *Phys. Rev. Lett.* **55**, 418 (1985).
  - [122] D. Kasinathan, J. Kunes, K. Koepnik, C. V. Diaconu, R. L. Martin, I. Prodan, G. Scuse-ria, N. Spaldin, L. Petit, T. C. Schulthess, et al., *Phys. Rev. B* **74**, 195110 (2006).
  - [123] J. Kolorenč and L. Mitas, *Phys. Rev. Lett.* **101**, 185502 (2008).
  - [124] Z. Fang, I. V. Solovyev, H. Sawada, and K. Terakura, *Phys. Rev. B* **59**, 762 (1999).
  - [125] J. Zhang, *Phys. Rev. Lett.* **84**, 507 (2000).
  - [126] Z. Fang, K. Terakura, H. Sawada, T. Miyazaki, and I. Solovyev, *Phys. Rev. Lett.* **81**, 1027 (1998).
  - [127] in *CRC Handbook of Chemistry and Physics*, edited by D. R. Linde (CRC Press, Boca Raton, FL, 2007).
  - [128] C. A. McCammon and L. Liu, *Phys. Chem. Minerals* **10**, 106 (1984).
  - [129] H. K. Bowen, D. Adler, and B. H. Auken, *J. Solid State Chem.* **12**, 355 (1975).
  - [130] G. Ortiz, D. M. Ceperley, and R. M. Martin, *Phys. Rev. Lett.* **71**, 2777 (1993).
  - [131] P. Vagner, M. Mosko, R. Nemeth, L. K. Wagner, and L. Mitas, *Physica E* **32**, 350 (2005).