Improved wave functions for quantum Monte Carlo

Priyanka Seth
Corpus Christi College, Cambridge

Dissertation submitted for the degree of Doctor of Philosophy
at the University of Cambridge
August 2012

To mum, dad, દાદા and the memory of બા.

Preface

THIS DISSERTATION describes work carried out between October 2008 and August 2012 in the Theory of Condensed Matter group at the Cavendish Laboratory, Cambridge under the supervision of Prof. R. J. Needs. The following chapters are included in work that has either been published or is to be published:

Chapter 3: P. Seth, P. López Ríos and R. J. Needs, "Quantum Monte Carlo study of the first-row atoms and ions", J. Chem. Phys. **134**, 084105 (2011).

Chapter 4: P. López Ríos, P. Seth, N. D. Drummond and R. J. Needs, "Framework for constructing generic Jastrow correlation factors", Phys. Rev. E **86**, 036703 (2012).

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration with others, except where specifically indicated in the text. This dissertation has not been submitted in whole or in part for any other degree or diploma at this or any other university. This dissertation does not exceed the word limit of 60,000 words.

Priyanka Seth

Cambridge, August 2012

Abstract

QUANTUM MONTE CARLO (QMC) methods can yield highly accurate energies for correlated quantum systems. QMC calculations based on many-body wave functions are considerably more accurate than density functional theory methods, and their accuracy rivals that of the most sophisticated quantum chemistry methods. This thesis is concerned with the development of improved wave function forms and their use in performing highly-accurate quantum Monte Carlo calculations.

All-electron variational and diffusion Monte Carlo (VMC and DMC) calculations are performed for the first-row atoms and singly-positive ions. Over 98% of the correlation energy is retrieved at the VMC level and over 99% at the DMC level for all the atoms and ions. Their first ionization potentials are calculated within chemical accuracy. Scalar relativistic corrections to the energies, mass-polarization terms, and one- and two-electron expectation values are also evaluated. A form for the electron and intracule densities is presented and fits to this form are performed.

Typical Jastrow factors used in quantum Monte Carlo calculations comprise electron-electron, electron-nucleus and electron-electron-nucleus terms. A general Jastrow factor capable of correlating an arbitrary of number of electrons and nuclei, and including anisotropy is outlined. Terms that depend on the relative orientation of electrons are also introduced and applied. This Jastrow factor is applied to electron gases, atoms and molecules and is found to give significant improvement at both VMC and DMC levels.

Similar generalizations to backflow transformations will allow useful additional variational freedom in the wave function. In particular, the use of different backflow functions for different orbitals is expected to be important in systems where the orbitals are qualitatively different. The modifications to the code necessary to accommodate orbital-dependent backflow functions are described and some systems in which they are expected to be important are suggested.

Acknowledgements

Contents

Chapter 1

INTRODUCTION

1.1 ELECTRONIC STRUCTURE CALCULATIONS

THE BEHAVIOUR OF MATTER around us is determined by the behaviour of the particles that constitute it. In condensed matter systems, these elementary building blocks are the nuclei that create the backbone of the material and the electrons that bind the nuclei together. *Ab initio* electronic structure methods give an understanding of the qualitative behaviour of a broad range of quantum systems and allow quantitative predictions of their properties to be made without any prior knowledge of the system. Besides being of tremendous importance for systems where obtaining experimental results is difficult, theoretical methods also give a deeper insight into the physics of such systems.

By the first postulate of quantum mechanics, the wave function $\Psi(\mathbf{R}, t)$ contains all the information specifying the state of the system. One only needs to solve the time-dependent Schrödinger equation which, in atomic units, is

$$i\frac{\partial}{\partial t}\Psi(\mathbf{R}, t) = \hat{H}\Psi(\mathbf{R}, t),\tag{1.1}$$

subject to symmetry constraints imposed by the nature of the particles, to understand the structure of the material of interest. The $3N$-dimensional vector $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ denotes the positions of the N particles. For systems comprising N electrons and M nuclei, the non-relativistic Hamiltonian reads

$$\hat{H} = -\sum_{i=1}^{N}\frac{\nabla_i^2}{2} - \sum_{I=1}^{M}\frac{\nabla_I^2}{2m_I} + \sum_{i<j}^{N}\frac{1}{r_{ij}} + \sum_{I<J}^{M}\frac{Z_I Z_J}{r_{IJ}} - \sum_{i=1}^{N}\sum_{I=1}^{M}\frac{Z_I}{r_{iI}},\tag{1.2}$$

where m_I and Z_I are the nuclear mass and charge, respectively, $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ and $r_{iI} = |\mathbf{r}_i - \mathbf{r}_I|$.

Solving Schrödinger's equation exactly is generally impossible as the size of the Hilbert space in which the eigenvectors reside grows exponentially with the number of particles.

To be able to tackle this problem at all, approximations that allow the coupled degrees of freedom to be separated are made.

1.1.1 BORN-OPPENHEIMER APPROXIMATION

The Born-Oppenheimer approximation [1] allows the separation of the $(N + M)$-body problem into an N-body problem and an M-body problem. As nuclei are over 10^3 times heavier than electrons, the electronic motion can be decoupled from the nuclear motion by assuming that the electrons relax instantaneously in response to any nuclear motion. The separation of time scales on which electronic and nuclear motion occurs allows nuclear motion to be viewed as a perturbation to the electronic Hamiltonian; the Born-Oppenheimer approximation is thus also known as the adiabatic approximation[1]. The wave function can then be written as a product of an electronic and a nuclear wave function. The time-dependence of the electronic wave function can be ignored as it depends parametrically on the instantaneous nuclear coordinates only. If the dynamics of a system are of interest, the nuclear degrees of motion can be solved for once the electronic configuration is found for a fixed nuclear geometry.

The wave function of a stationary state ϕ_n is obtained by solving the time-independent Schrödinger equation

$$\hat{H}\phi_n(\mathbf{R}) = E_n\phi_n(\mathbf{R}), \tag{1.3}$$

where $\hat{H}$ is the Hamiltonian operator of which the wave function is an eigenfunction with energy eigenvalue E_n. The non-relativistic Born-Oppenheimer Hamiltonian for the electronic degrees of freedom is then

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 + \frac{1}{2}\sum_{i\neq j}^{N}\frac{1}{r_{ij}} - \sum_{i=1}^{N}\sum_{I=1}^{M}\frac{Z_I}{r_{iI}}. \tag{1.4}$$

As the positions of the nuclei are fixed, the nucleus-nucleus Coulomb term only contributes a constant energy offset which has been omitted.

1.1.2 VARIATIONAL PRINCIPLE

The variational principle is indispensable in electronic structure theory as it provides an upper bound on the exact ground-state energy.

As the exact normalized eigenstates ϕ_i of the Hamiltonian form a complete orthogonal basis, a normalized trial wave function may be written $\Psi = \sum_i c_i\phi_i$, where the expansion

---

[1]The condition that the nuclei move much more slowly than the electrons is most likely to be violated for light nuclei. It may not be possible to decouple nuclear and electronic degrees of freedom when electronic states are degenerate, or at finite temperatures when the increased kinetic energy of the nuclei may be able to induce electronic excitations between electronic states close in energy. In these cases, adiabaticity is not maintained and care must taken in making the Born-Oppenheimer approximation.

coefficients satisfy $\sum_i |c_i|^2 = 1$. The expectation value of the Hamiltonian[2] for this wave function is then

$$
\begin{aligned}
E[\Psi] &= \langle \Psi | \hat{H} | \Psi \rangle \\
&= \langle \sum_i c_i \phi_i | \hat{H} | \sum_j c_j \phi_j \rangle \\
&= \sum_{i,j} c_i^* c_j \langle \phi_i | \hat{H} | \phi_j \rangle \\
&= \sum_i |c_i|^2 E_i,
\end{aligned}
\tag{1.5}
$$

where $E_i = \langle \phi_i | \hat{H} | \phi_i \rangle$. By definition, $E_i \geq E_0$ where E_0 is the ground-state energy, and thus

$$
E[\Psi] \geq E_0.
\tag{1.6}
$$

The expectation value of the Hamiltonian for any wave function is therefore always greater than the true ground-state energy. In a search over all possible wave functions, that which gives the lowest energy corresponds to the exact ground state.

In practical terms, the variational principle means that a wave function with a given parametric form can be optimized by varying the parameters to minimize the energy. This forms the basic operating principle of most electronic structure methods.

1.1.3 ELECTRONIC CORRELATION

Solving the Schrödinger equation analytically proves challenging for all but the simplest systems even with the Born-Oppenheimer approximation. The form of the electron-electron Coulomb interactions prevent further decoupling of the degrees of freedom. To solve this equation without further approximations requires solving $3N$-dimensional coupled equations for a system of N electrons.

The problem of finding an exact solution to the Schrödinger equation still scales exponentially with the number of electrons. The greatest challenge for electronic structure methods is the representation of electron correlation and devising a treatment that allows the problem to be solved accurately in polynomial time. Various such methods have been developed to solve Schrödinger's equation numerically and an overview of some of them is given in Sec. 1.2.

---

[2]To guarantee that an ordered set of real eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ exists, the Hamiltonian must be Hermitian.

1.2 COMMON ELECTRONIC STRUCTURE METHODS

A natural approach to simplifying the problem of $3N$ coupled partial differential equations is to approximate the electron-electron interactions by an average single-particle potential. In this mean-field approximation, the N-body Schrödinger equation separates into N one-electron equations which can be easily solved. Such mean-field methods give rise to the orbitals that scientists use as a visualization tool and a means to a basic understanding of electronic structure including the explanation of the periodic table.

The Hartree-Fock method is one such mean-field method. It is conceptually simple and computationally affordable with a scaling of $\mathcal{O}(N^4)$. However, it does not account for the Coulomb correlation between the electrons. Post-Hartree-Fock methods improve on Hartree-Fock method by constructing more complicated wave functions either by using a larger expansion of determinants or including an explicit correlation factor, or both. As a result of the additional complexity, these methods suffer from poorer scaling with system size and typically scale as $\mathcal{O}(N^4 - N^7)$. Density functional theory is another mean-field method that scales favourably, as $\mathcal{O}(N^3)$.

1.2.1 HARTREE-FOCK THEORY

In the Hartree-Fock (HF) approximation [2, 3], the electrons are assumed to be independent. The Hamiltonian is written as the sum of single-particle Hamiltonians $\hat{h}_i$ and an effective potential $\hat{V}$ resulting from the electron-electron interaction term,

$$\hat{H} = \sum_i \hat{h}_i + \hat{V} = \sum_i \left[-\frac{1}{2}\nabla_i^2 - \sum_I \frac{Z_I}{r_{iI}} \right] + \left[\sum_{i<j} \frac{1}{r_{ij}} \right]. \tag{1.7}$$

To satisfy the Pauli exclusion principle, the wave function is written as an antisymmetrized product of one-electron orbitals known as a Slater determinant,

$$\Psi(\mathbf{X}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_1(\mathbf{x}_2) & \cdots & \psi_1(\mathbf{x}_N) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & \cdots & \psi_2(\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(\mathbf{x}_1) & \psi_N(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix}, \tag{1.8}$$

where $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$ and $\mathbf{x}_i$ includes both spatial and spin coordinates: $\mathbf{x}_i = \{\mathbf{r}_i, \sigma_i\}$.

In accordance with the variational principle, the best orbitals can be found by minimizing the energy with respect to the orbitals, subject to the orthonormality condition $\langle \psi_i | \psi_j \rangle = \delta_{ij}$ imposed by the Lagrange multipliers ϵ_{ij},

$$\frac{\delta}{\delta\psi} \left[\langle \Psi | \hat{H} | \Psi \rangle - \sum_{i,j} \epsilon_{ij} \left(\langle \psi_i | \psi_j \rangle - \delta_{ij} \right) \right] = 0. \tag{1.9}$$

Making a unitary transformation results in the canonical Hartree-Fock equations,

$$\hat{h}_i \psi_i(\mathbf{r}) + \sum_j \int \frac{|\psi_j(\mathbf{r'})|^2}{|\mathbf{r} - \mathbf{r'}|} \psi_i(\mathbf{r}) \, d\mathbf{r'} - \sum_j \delta_{\sigma_i \sigma_j} \int \frac{\psi_j^*(\mathbf{r'}) \psi_i(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|} \psi_j(\mathbf{r}) \, d\mathbf{r'} = \epsilon_i \psi_i(\mathbf{r}). \quad (1.10)$$

The second and third terms in the Hartree-Fock equation account for the electron-electron interaction. The second term is the Hartree term (also known as the Coulomb or direct term) and is the mean potential an electron feels due to the charge density $n(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2$ of the N electrons, including itself. The third term is known as the exchange term and it stems from the antisymmetric form of the wave function[3]. As a consequence of the Pauli exclusion principle, each electron is surrounded by an exchange, or Fermi, hole of positive charge that electrons of like-spin avoid. In this sense, the exchange hole is exactly equivalent to the absence of electron. The self-interaction in the Hartree term is cancelled out by the corresponding $i = j$ exchange term if i and j have the same spin. The exchange term is a non-local operator in the sense that the value of the exchange operator acting on ψ_i depends not only on the value of ψ_i locally at $\mathbf{r}$, as is true for the Hartree term, but on the value of ψ_i everywhere. The Hartree-Fock equations are non-linear as the Hartree and exchange operators depend functionally on ψ_i and they need to be solved iteratively until self-consistency is reached.

The exchange (or Fermi correlation) effects arising from the Pauli repulsion are wholly accounted for as a result of the antisymmetric form of the wave function. Electronic correlation is neglected by construction and is only included when higher-order corrections are included, as done in Møller-Plesset perturbation theory [2, 3]. The electron correlation energy is thus defined as the difference between the Hartree-Fock energy and the exact non-relativistic Born-Oppenheimer energy. As the Hartree-Fock energy is an upper bound to the exact energy, the correlation energy is always negative: $E_{\mathrm{corr}} = E_{\mathrm{exact}} - E_{\mathrm{HF}} \leq 0$. The Hartree-Fock approximation is a good first approximation for weakly-correlated atoms and molecules, and it forms the basis for more accurate post-Hartree-Fock methods such as configuration interaction and coupled-cluster methods.

1.2.2 CONFIGURATION INTERACTION

Configuration interaction (CI) methods [2, 3] are a simple and elegant extension to HF. In principle, the full CI (FCI) wave function is the exact solution to the Schrödinger equation for a given basis set. The FCI wave function is written as a linear combination of all the determinants D_i that span the N-body Hilbert space and that can be constructed from one-electron Hartree-Fock orbitals, $|\Psi_{\mathrm{FCI}}\rangle = \sum_i c_i |D_i\rangle$. The linear coefficients c_i are

---

[3]If the wave function is not made antisymmetric, and is simply a product of one-electron orbitals, the Hartree-Fock approximation reduces to the Hartree approximation. The exchange term vanishes and the mean-field becomes a purely local quantity.

obtained variationally by minimizing the energy, and the correlation problem effectively becomes a matrix eigenvalue problem.

In practice, the wave function is expanded in a finite basis set. For a basis set containing P orbitals, $|\Psi_{\text{FCI}}\rangle$ contains $\binom{P}{N} = \frac{P!}{N!(P-N)!}$ determinants. The cost of FCI thus scales as $\mathcal{O}(P!)$ and the use of the FCI method is limited to the benchmarking of the smallest atomic and molecular systems.

To make this problem more tractable, the FCI expansion is truncated to include determinants with a few excitations (typically single and double excitations) from the Hartree-Fock ground state. The number of excitations that need to be considered increases quickly even for truncated expansions. CI wave functions do not give a compact description of electronic correlation, and such methods cannot be used to study solids. Additionally, the truncated CI expansion suffers from lack of size-extensivity, which requires that the energy of a system be proportional to the number of particles N as $N \to \infty$. This leads to difficulties in comparing systems of different sizes as the errors do not cancel out.

A very promising method, the full configuration interaction quantum Monte Carlo (FCIQMC) method [4, 5], has recently been developed to maintain the accuracy of the FCI while reducing the computational burden of the traditional formulation. It uses an algorithm similar to that of diffusion Monte Carlo to stochastically integrate the Schrödinger equation using a set of evolving walkers. To date, the method has obtained some very accurate results [6, 7, 8]. The method is still being developed and efforts are being made at better understanding its strengths and weaknesses [9].

1.2.3 MULTI-CONFIGURATIONAL SELF-CONSISTENT FIELD

The multi-configurational self-consistent field (MCSCF) theory [2, 3] can be regarded as a generalization of HF to systems that cannot be adequately described by a single configuration, such as systems involving degenerate or nearly-degenerate configurations. Beginning with a truncated CI expansion, the linear expansion coefficients and orbitals are simultaneously optimized in an iterative process analogous to that in HF theory[4]. As the orbitals are variationally optimized in addition to the expansion coefficients, the MCSCF energy is lower than the CI energy using the same truncated expansion, but higher than the FCI energy. In practice, MCSCF expansions are much smaller than CI expansions as they are restricted by the difficulty of the non-linear problem of optimizing orbitals and coefficients simultaneously.

---

[4] The MCSCF method is hence also referred to as the multi-configurational Hartree-Fock (MCHF) method.

1.2.4 COUPLED CLUSTER

Starting with a non-interacting picture, correlation of an electron pair can be considered as a two-electron scattering process that excites two electrons from the ground-state one-electron orbitals into unoccupied virtual orbitals. The coupled-cluster method (CC) [2, 3] generalizes this to N-body excitations. The wave function is written $|\Psi_{CC}\rangle = e^{\hat{T}}|\Psi_{HF}\rangle$, where $|\Psi_{HF}\rangle$ is the HF ground-state determinant and $\hat{T} = \hat{T}_1 + \hat{T}_2 + \ldots + \hat{T}_N = \sum_\mu t_\mu \hat{\tau}_\mu$ is the cluster operator. $\hat{T}_n$ is the operator of all n^{th}-order excitations[5] and $\hat{\tau}_\mu$ is the excitation operator corresponding to a given excited-state determinant in the FCI expansion.

A similarity transformation is made to retain the form of an eigenvalue problem:

$$\hat{H}_{CC}|\Psi_{HF}\rangle = e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_{HF}\rangle = E_{CC}|\Psi_{HF}\rangle. \tag{1.11}$$

As the coupled-cluster Hamiltonian $\hat{H}_{CC}$ is not Hermitian, the problem cannot be solved variationally. Instead, the excitation amplitudes t_μ are calculated by projecting the Schrödinger equation onto excited-state determinants $|D_i\rangle$ and solving

$$\langle D_i|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_{HF}\rangle = \langle D_i|E_{CC}|\Psi_{HF}\rangle = 0. \tag{1.12}$$

The coupled-cluster energy is given by projecting onto the HF ground state, $E_{CC} = \langle\Psi_{CC}|\hat{H}|\Psi_{CC}\rangle = \langle\Psi_{HF}|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_{HF}\rangle$.

The FCI and CC wave functions contain the same number of parameters and so it seems that the non-linearity of the CC model only adds unnecessary complexity. However, when the cluster operator is truncated at $\hat{T}_n$, the CC state continues to include higher-excitation determinants with approximate coefficients and so is able to describe correlation more compactly than the CI wave function. Coupled cluster with singles and doubles and perturbative triples, abbreviated CCSD(T), is regarded as the 'gold standard' of quantum chemistry and scales as $\mathcal{O}(N^7)$. The exponential treatment of the cluster operator restores size-extensivity even for a truncated cluster operator, in contrast to truncated CI methods. The CC method allows a more compact wave function and is size-extensive; as such, it addresses the shortcomings of the CI method for the price of the variational property.

1.2.5 EXPLICITLY-CORRELATED METHODS

The slow convergence of the CI energy can be attributed to the absence of terms with odd powers of r_{ij} in the wave function, and in particular, to the absence of a term linear in r_{ij}. Without such terms, the wave function does not have the cusps to describe coalescing electrons. A large basis set must be used to capture the high-frequency oscillations close

---

[5]The single excitations described by $\hat{T}_1$ do not capture any electron correlation on their own. However, they serve as relaxations to the one-electron orbitals. These orbitals change in response to the modification to the HF mean-field by the many-body excitations.

to the nuclei and the slower decay at large distances. Hence, an expansion in determinants does not give a compact wave function form.

In explicitly-correlated methods the wave function is written as a product of a correlating function and either a single Slater determinant or a sum of determinants. Several such methods improve on the truncated CI and CC formalisms [10].

1.2.5.1 TRANSCORRELATED METHOD

One interesting method based on an explicitly-correlated wave function is the transcorrelated (TC) method. This method uses a correlating Jastrow factor $e^{J(\mathbf{R})}$ that is an explicit function of two-body separations r_{ij}. The correlated wave function is then a product of the Jastrow factor and the HF Slater determinant, $\Psi(\mathbf{R})$. The Schrödinger equation is given by $\hat{H}[e^{J(\mathbf{R})}\Psi(\mathbf{R})] = E[e^{J(\mathbf{R})}\Psi(\mathbf{R})]$. A similarity transformation is made such that the Coulomb singularities that necessitate cusps are absorbed into the Hamiltonian, and as a result, the wave function can be written more simply as a Slater determinant,

$$\hat{H}_{\mathrm{TC}}\Psi(\mathbf{R}) = [e^{-J(\mathbf{R})}\hat{H}e^{J(\mathbf{R})}]\Psi(\mathbf{R}) = E_{\mathrm{TC}}\Psi(\mathbf{R}). \tag{1.13}$$

Boys and Handy [11, 12, 13, 14] first proposed this method for small atoms and molecules. However, the variational principle does not hold as the more complicated similarity-transformed Hamiltonian H_{TC} is not Hermitian. Consequently, interest in the method faded.

Following a recent revival, the transcorrelated method has been applied to a variety of systems including solids [15], for which traditional quantum chemistry methods struggle. In the approach of Ten-no and coworkers [16], the Jastrow factor is fixed to satisfy the cusp conditions and improve the convergence of a wave function originating from perturbation and coupled-cluster methods.

Another approach is to minimize the variance of the energy [17, 18] which satisfies a zero-variance principle. A self-consistent field equation can be derived for the orbitals and solved in a manner similar to that of the HF method in what is referred to as the transcorrelated self-consistent field (TC-SCF) method. The Jastrow factor parameters are optimized in the variational Monte Carlo framework as difficult high-dimensional integrals must be performed in minimizing the variance. In principle, such a formalism allows the orbitals to be optimized in the presence of a Jastrow factor resulting in a better trial wave function without the inherent statistical errors in the orbital optimization associated with the Monte Carlo method.

1.2.6 DENSITY FUNCTIONAL THEORY

Density functional theory (DFT) [19] is distinct from all the other electronic structure approaches discussed thus far. The fundamental principle of DFT is to describe properties

of a system of interacting particles as a functional of the ground-state particle density $n(\mathbf{r})$. As such, the basic variable is defined in a simpler 3-dimensional space instead of the $3N$-dimensional space that the many-body wave function resides in.

The Hohenberg-Kohn theorems [20] lay the groundwork for DFT. Their first theorem states that the external potential $V_{\text{ext}}(\mathbf{r})$, to within an additive constant, is a unique functional of the ground-state density $n(\mathbf{r})$. Thus, the density rather than the wave function can be viewed as the quantity of interest. The second Hohenberg-Kohn theorem gives a variational principle for the energy functional[6]. For any given external potential, the ground-state density $n(\mathbf{r})$ minimizes the energy functional

$$E[n(\mathbf{r})] = F[n(\mathbf{r})] + \int V_{\text{ext}}(\mathbf{r})n(\mathbf{r})\, d\mathbf{r} \geq E_0, \tag{1.14}$$

where E_0 is the exact ground-state energy and F is a universal functional of the density. However, these theorems do not define the universal functional F.

Kohn and Sham [23] devised a practical method within which to apply the formal theory of Hohenberg and Kohn; this method continues to be the basis of DFT codes today. As was done in Hartree-Fock theory, the density can be varied to minimize the energy functional subject to the density normalization constraint $\int n(\mathbf{r})\, d\mathbf{r} = N$, giving

$$\frac{\delta}{\delta n}\left[E[n(\mathbf{r})] - \mu\left(\int n(\mathbf{r})\, d\mathbf{r} - N\right)\right] = 0. \tag{1.15}$$

The energy functional E is defined by Kohn and Sham as

$$E[n(\mathbf{r})] = T_s[n(\mathbf{r})] + \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}' + E_{\text{xc}}[n(\mathbf{r})] + \int V_{\text{ext}}(\mathbf{r})n(\mathbf{r})\, d\mathbf{r}, \tag{1.16}$$

where $T_s[n(\mathbf{r})]$ is the kinetic energy functional of the non-interacting electron gas with N electrons and density $n(\mathbf{r})$,

$$T_s[n(\mathbf{r})] = -\frac{1}{2}\sum_i^N \int \psi_i^*(\mathbf{r})\nabla^2 \psi_i(\mathbf{r})\, d\mathbf{r} \tag{1.17}$$

and $E_{\text{xc}}[n(\mathbf{r})]$ is the exchange-correlation energy into which all the exchange and correlation not captured by T_s and the potential energy operator is swept. The variational condition then becomes

$$\frac{\delta E[n(\mathbf{r})]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n(\mathbf{r})]}{\delta n(\mathbf{r})} + V_{\text{KS}}(\mathbf{r}) = \mu, \tag{1.18}$$

---

[6]An alternative and more insightful formulation is given by Levy and Lieb [21, 22] in what is known as the constrained-search method.

where the Kohn-Sham potential $V_{\mathrm{KS}}(\mathbf{r})$ is

$$V_{\mathrm{KS}}(\mathbf{r}) = V_{\mathrm{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|} \, d\mathbf{r'} + V_{\mathrm{xc}}(\mathbf{r}) \tag{1.19}$$

and

$$V_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}[n(\mathbf{r})]}{\delta n(\mathbf{r})} \tag{1.20}$$

is the exchange-correlation potential.

Consider an auxiliary system of N non-interacting electrons with the same density as the interacting system. The Kohn-Sham potential gives a set of single-particle Schrödinger equations,

$$\left[-\frac{1}{2}\nabla^2 + V_{\mathrm{KS}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}), \tag{1.21}$$

that can be solved self-consistently using the density $n(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2$ and the definitions of $V_{\mathrm{KS}}(\mathbf{r})$ and $V_{\mathrm{xc}}(\mathbf{r})$ (Eqs. 1.19 and 1.20). Note that the eigenfunctions $\psi_i(\mathbf{r})$ do not have physical meaning apart from the connection between $\psi_i(\mathbf{r})$ and the true density $n(\mathbf{r})$. The ground-state energy is then given by

$$E[n(\mathbf{r})] = \sum_i^N \epsilon_i + E_{\mathrm{xc}}[n(\mathbf{r})] - \int V_{\mathrm{xc}}(\mathbf{r}) n(\mathbf{r}) \, d\mathbf{r} - \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|} \, d\mathbf{r} \, d\mathbf{r'}. \tag{1.22}$$

Knowing the exact form of the exchange-correlation functional E_{xc} allows all many-body effects to be included. However, finding the exact form of the exchange-correlation functional is a great challenge, and perhaps even impossible. Instead, approximations to the functional are used with varying success. One such example is the local density approximation (LDA) functional, which is parametrized using accurate quantum Monte Carlo results [24]. This choice of exchange-correlation functional remains the main uncontrolled approximation of DFT. Despite this, DFT has been successfully applied to a large variety of systems that are much too large for the more expensive wave function-based methods.

1.2.7 QUANTUM MONTE CARLO METHODS

The class of quantum Monte Carlo (QMC) techniques rely on stochastic integration of equations resulting from various formulations of Schrödinger's equation. This class of techniques includes (but is not limited to) path-integral quantum Monte Carlo [25], auxiliary-field quantum Monte Carlo [26] and reptation Monte Carlo [27]. The two methods used in this thesis are the variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC) methods [28, 29, 30]. They are applied to continuum correlated electron systems at zero

temperature[7].

VMC is the simpler technique. Expectation values of operators such as the Hamiltonian are calculated with an approximate many-body trial wave function and the integrals are evaluated using a Monte Carlo technique which converges more quickly than numerical quadrature methods for high dimensions. The functional form of the wave function is chosen to contain a number of parameters whose values are obtained by stochastic optimization. Higher accuracy is obtained in the DMC method by evolving the wave function in imaginary time so that it decays towards the ground state, while the fixed-node approximation is made to maintain fermionic symmetry. DMC is exact in principle but in practice the method is limited by the fixed-node error. Both the VMC and DMC methods are variational, and the DMC energy is bounded from above by the VMC energy and from below by the exact energy. These methods are intrinsically parallel and scale favourably with system size, normally as $\mathcal{O}(N^2 - N^4)$, depending the form of the orbitals and wave function.

One of the unique features of VMC and DMC is that the methods are inherently independent of the form of the Hamiltonian and wave function. This is in contrast to traditional quantum chemistry methods that are generally limited to using only an expansion of Slater determinants comprising one-electron orbitals for easy integration. Additionally, systems with any mixture of fermionic and bosonic particles, custom potentials and external fields can be studied with appropriate wave function forms that are best parametrized to describe the physics of the system.

1.2.8 Discussion

The simplification of an exponentially-difficult problem into one that can be solved in polynomial time fundamentally implies the existence of limitations and a loss in accuracy. These limitations manifest themselves differently in the various systems and the challenge is to find the set of approximations optimizing the accuracy with respect to cost. CI and CCSD are limited by the truncation order (the former more so than the latter), which reflects the need for a large basis set to describe the electron-electron cusps in the wave function. In DFT, all the complexity arising from many-body interactions is swept into the exchange-correlation functional whose exact form is unlikely to be ever known. In addition to the statistical error due to finite computational resources, VMC has a systematic error arising from the incomplete parametrization of the trial wave function. While DMC is in principle exact, the fixed-node approximation results in an uncontrolled nodal error.

---

[7]At non-zero temperatures, the system is no longer in a pure state. The many-body wave function is an insufficient descriptor of the system and density matrices must be considered. This is done in path-integral Monte Carlo [25] and auxiliary-field Monte Carlo [26].

1.3　THIS THESIS

The aim of the work presented in this thesis is to make systematic improvements to trial wave functions in QMC, allowing higher-accuracy calculations to be performed more efficiently.

The VMC and DMC methods are outlined in Chapter 2. The results from a study of the first-row atoms is given in Chapter 3. In Chapter 4, a generalized form of the Jastrow factor is presented and results for a variety of benchmark systems are given. Preliminary work on the development of generalized backflow transformations is described in Chapter 5. Finally, conclusions are drawn in Chapter 6.

All of the calculations described in this thesis were performed using the CASINO package [31]. Hartree atomic units (a.u.) in which $\hbar = |e| = m_e = 4\pi\epsilon_0 = 1$, are used throughout except where otherwise indicated.

Chapter 2

QUANTUM MONTE CARLO

2.1 MONTE CARLO METHODS

ERRORS IN COMMON quadrature methods scale poorly with the dimensionality of the integral. For example, the error in the Newton-Cotes method [32] with 2-point linear fitting, commonly known as the trapezoidal rule, has an error that scales as $M^{-2/d}$. The error in Simpson's Rule, the Newton-Cotes method with 3-point quadratic fitting, scales as $M^{-4/d}$. The great strength of Monte Carlo integration lies in the scaling of the error as $M^{-1/2}$, where M is the number of sampling points used, irrespective of the dimensionality of the integral. The weak dependence of the error on dimensionality results from the central limit theorem. While grid-based methods are suited to problems of low dimensionality, Monte Carlo is the only practical technique for evaluating the high-dimensional integrals encountered in realistic electronic structure calculations.

Naïve Monte Carlo integration involves averaging the value of the integrand at a random set of points within the region of integration. Each of these sets of points $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)$ will be henceforth referred to as an electron configuration. Monte Carlo integration can be made more efficient by using importance sampling rather than generating uniformly-distributed configurations. Under the importance sampling transformation, the integrand is decomposed into a product of two functions, $g(\mathbf{R}) = f(\mathbf{R})P(\mathbf{R})$, such that

$$I = \int g(\mathbf{R}) \, d\mathbf{R} = \int f(\mathbf{R})P(\mathbf{R}) \, d\mathbf{R}, \qquad (2.1)$$

where $f(\mathbf{R}) \equiv g(\mathbf{R})/P(\mathbf{R})$ and $P(\mathbf{R})$ is chosen to satisfy the conditions for a probability density function, namely it is everywhere non-negative and normalized. The integral I can be estimated as the average of $f(\mathbf{R})$ at a finite number M of configurations $\mathbf{R}_i$ distributed as $P(\mathbf{R})$,

$$I = \langle f \rangle \approx \frac{1}{M} \sum_{i=1}^{M} f(\mathbf{R}_i). \qquad (2.2)$$

This estimate is exact in the limit $M \to \infty$. The error of the estimate is

$$\sigma = \frac{\sigma_f}{\sqrt{M}}, \tag{2.3}$$

where σ_f^2 is the variance of $f(\mathbf{R})$,

$$\begin{aligned}
\sigma_f^2 &= \int P(\mathbf{R})(f(\mathbf{R}) - \bar{f})^2 \, d\mathbf{R} \\
&\approx \frac{1}{M-1} \sum_{i=1}^{M} (f(\mathbf{R}_i) - \langle f \rangle)^2.
\end{aligned} \tag{2.4}$$

Here, $\bar{f}$ denotes the population mean of f.

Performing the integral efficiently is now a matter of choosing a probability density $P(\mathbf{R})$ such that $f(\mathbf{R})$ is as close to $\langle f \rangle = I$ as possible, resulting in a small variance. This is effectively smoothing out the function f and hiding the complexity of the task in P, which will have a greater density where g is large and vice versa. Selecting $P(\mathbf{R})$ for most applications is non-trivial as these multi-dimensional probability densities are complicated and cannot be directly sampled, and additionally, their normalizations are unknown.

2.1.1 THE METROPOLIS ALGORITHM

The Metropolis algorithm [33] overcomes these difficulties by using a random walk to generate a Markov chain of configurations that sample the phase space[1]. It generates the sequence of configurations according to the following algorithm:

1. A configuration $\mathbf{R}'$ is randomly generated.

2. A trial move to $\mathbf{R}$ with transition probability $T(\mathbf{R}' \to \mathbf{R})$ is proposed.

3. The move to the new configuration $\mathbf{R}$ is accepted with probability

$$A(\mathbf{R}' \to \mathbf{R}) = \min\left(1, \frac{P(\mathbf{R})T(\mathbf{R} \to \mathbf{R}')}{P(\mathbf{R}')T(\mathbf{R}' \to \mathbf{R})}\right). \tag{2.5}$$

4. Step (2) onward is repeated for $\mathbf{R}$ if the move was accepted and for $\mathbf{R}'$ if the move was rejected.

The initial configurations generated are correlated with the starting configuration and are discarded. Once the walk has equilibrated, configurations are distributed according to P and detailed balance is satisfied such that the probability of moving in a given direction

---

[1] Configurations are therefore also known as walkers, and these terms are used interchangeably.

in configuration space is the same as in the opposite direction,

$$P(\mathbf{R})T(\mathbf{R} \to \mathbf{R}')A(\mathbf{R} \to \mathbf{R}') = P(\mathbf{R}')T(\mathbf{R}' \to \mathbf{R})A(\mathbf{R}' \to \mathbf{R}). \tag{2.6}$$

The transition probability must be ergodic for the Metropolis algorithm to remain valid, that is, any point in the configuration space must be reachable from any other point in a finite number of moves. The configuration then executes a random walk in configuration space and quantities of interest are measured for each configuration and averaged.

One of the drawbacks of the Metropolis algorithm is its inefficiency. The size of the trial move proposed is important in determining how effectively the configuration space is sampled. The move size is generally constrained to be smaller rather than larger as a large proportion of trial moves may be rejected for large moves. However, only a small volume of configuration space will be sampled for short moves and a large number of sequential configurations in the walk will be serially correlated. The issue of efficiency is looked at more closely in Sec. 2.5.

2.2 Variational Monte Carlo

VMC is a simple and elegant method that incorporates the variational principle and Monte Carlo integration scheme to evaluate the energy as the expectation value of the Hamiltonian with an approximate trial wave function[2] Ψ. Ceperley *et al.* [34] first used this method to study fermionic systems.

The variational energy is expressed as

$$\begin{aligned} E_{\mathrm{V}} &= \frac{\int \Psi(\mathbf{R})\hat{H}\Psi(\mathbf{R})\,d\mathbf{R}}{\int \Psi(\mathbf{R})\Psi(\mathbf{R})\,d\mathbf{R}} \\ &= \frac{\int \Psi^2(\mathbf{R})E_{\mathrm{L}}\,d\mathbf{R}}{\int \Psi^2(\mathbf{R})\,d\mathbf{R}}, \end{aligned} \tag{2.7}$$

where $E_{\mathrm{L}} = \Psi^{-1}\hat{H}\Psi$ is the local energy. Using the Metropolis algorithm to sample configuration space, the energy is estimated as

$$E_{\mathrm{VMC}} = \frac{1}{M}\sum_{i=1}^{M} E_{\mathrm{L}}(\mathbf{R}_i), \tag{2.8}$$

where the M configurations $\mathbf{R}_i$ are distributed as

$$P(\mathbf{R}) = \frac{\Psi^2(\mathbf{R})}{\int \Psi^2(\mathbf{R})\,d\mathbf{R}}. \tag{2.9}$$

---

[2]Only real trial wave functions are considered here, as discussed in Sec. 2.4.

The squared error of this estimate is

$$\sigma^2 \approx \frac{1}{M(M-1)} \sum_{i=1}^{M} (E_{\mathrm{L}}(\mathbf{R}_i) - E_{\mathrm{VMC}})^2. \tag{2.10}$$

The importance function $P(\mathbf{R})$ concentrates the sampling of configurations in regions of the Hilbert space that are visited more frequently. Note that the normalization factor $\int \Psi^2(\mathbf{R}) \, d\mathbf{R}$ does not need to be known explicitly as the two occurrences of this factor in Eq. 2.5 cancel out. The expectation values of other operators can be evaluated in an analogous manner.

As the trial wave function approaches an eigenstate, $E_{\mathrm{L}}(\mathbf{R})$ becomes smoother and fewer configurations are needed to obtain a given accuracy. In the limit that the trial wave function is an exact eigenstate of the Hamiltonian ϕ_i with energy eigenvalue E_i, $E_{\mathrm{L}} = \phi_i^{-1} \hat{H} \phi_i = E_i$ is everywhere constant and the variance of the energy becomes zero. The VMC estimate of the energy is then equal to the exact energy for even a finite number of configurations. This zero-variance property motivates one group of optimization methods discussed in Sec. 2.6.

2.3 DIFFUSION MONTE CARLO

The DMC method is fundamentally very different from the VMC method. It improves the wave function by evolving it according to the imaginary-time Schrödinger equation to project out the ground state[3]. In doing so, the DMC method addresses the main shortcoming of the VMC method, namely, its reliance on the functional form of the wave function. Nonetheless, the DMC algorithm still depends on the quality of the wave function and particularly on the quality of the nodal surface. Only the basic DMC algorithm is outlined here; many refinements are made in practice to reduce errors and biases.

2.3.1 IMAGINARY-TIME EVOLUTION

The time-dependent Schrödinger equation (Eq. 1.1) in imaginary time $\tau = it$ is

$$-\frac{\partial \Phi(\mathbf{R}, \tau)}{\partial \tau} = (\hat{H} - E_{\mathrm{T}}) \Phi(\mathbf{R}, \tau), \tag{2.11}$$

where E_{T} is a constant energy offset, and whose solution $\Phi(\mathbf{R}, \tau)$ satisfies

$$\Phi(\mathbf{R}, \tau + d\tau) = e^{-(\hat{H} - E_{\mathrm{T}}) d\tau} \Phi(\mathbf{R}, \tau). \tag{2.12}$$

---

[3]The DMC wave function is not known analytically; only its distribution is known.

Expanding $\Phi(\mathbf{R}, \tau)$ in the eigenfunctions ϕ_i of the Hamiltonian gives

$$\Phi(\mathbf{R}, \tau) = \sum_i c_i \phi_i(\mathbf{R}) e^{-(E_i - E_\mathrm{T})\tau}. \tag{2.13}$$

The oscillating time components of the real-time wave function are transformed into decaying exponentials. Evolving any $\Phi(\mathbf{R}, 0)$ that contains a non-zero component of the ground state $\phi_0(\mathbf{R})$ in imaginary time will project out the ground state which, by definition, has the lowest eigenvalue, E_0. Choosing $E_\mathrm{T} = E_0$ eliminates the time dependence of the wave function and gives the steady-state solution $\Phi(\mathbf{R}, \tau \to \infty) = c_0 \phi_0(\mathbf{R})$.

For a Hamiltonian consisting of kinetic and potential terms $\hat{H} = \hat{T} + \hat{V} = -\frac{1}{2}\nabla^2 + V(\mathbf{R})$, Eq. 2.11 can be written as

$$\frac{\partial \Phi(\mathbf{R}, \tau)}{\partial \tau} = \frac{1}{2}\nabla^2 \Phi(\mathbf{R}, \tau) + (E_\mathrm{T} - V(\mathbf{R}))\Phi(\mathbf{R}, \tau). \tag{2.14}$$

If the potential term is ignored, Eq. 2.14 becomes a diffusion equation with a diffusion constant $D = 1/2$ in the $3N$-dimensional configuration space. If the kinetic term is neglected, Eq. 2.14 becomes a first-order rate equation with a rate constant $E_\mathrm{T} - V(\mathbf{R})$. The diffusion process can be simulated by randomly moving weighted walkers initially distributed as $\Phi(\mathbf{R}, 0)$ at a rate determined by the diffusion constant while the potential term acts as a reweighting of the walkers. The reweighting is more efficiently simulated as a branching process to prevent one walker from dominating over the others [35]. In this process, a branching factor determines if walkers are created or destroyed.

This simple algorithm suffers from two severe limitations. Firstly, the branching factor is proportional to $V(\mathbf{R})$ and thus has divergences when two particles coincide as a result of the Coulomb interactions. The resulting fluctuations in the walker population result in a large variance in the estimated energy. Additionally, as a fermionic wave function necessarily has positive and negative regions, the assumption that Φ can be viewed as a probability distribution is invalid. This is one manifestation of the infamous fermion sign problem, discussed in Sec. 2.3.2. The introduction of importance sampling as discussed in Sec. 2.3.3 addresses both of these issues.

2.3.2 THE FERMION SIGN PROBLEM

The antisymmetry of fermionic wave functions requires that they have both positive and negative regions, preventing the wave function from being directly interpreted as a probability distribution. This fundamental property is at the root of the fermion sign problem which plagues projector QMC methods such as DMC, where walkers are distributed according to the wave function Φ. VMC is unaffected by the fermion sign problem as walkers are distributed according to Φ^2 which is positive everywhere.

The evolution of the wave function in the simple approach described in Sec. 2.3.1

would therefore lead to the nodeless bosonic ground state which has the lowest energy, as shown by the following argument [36].
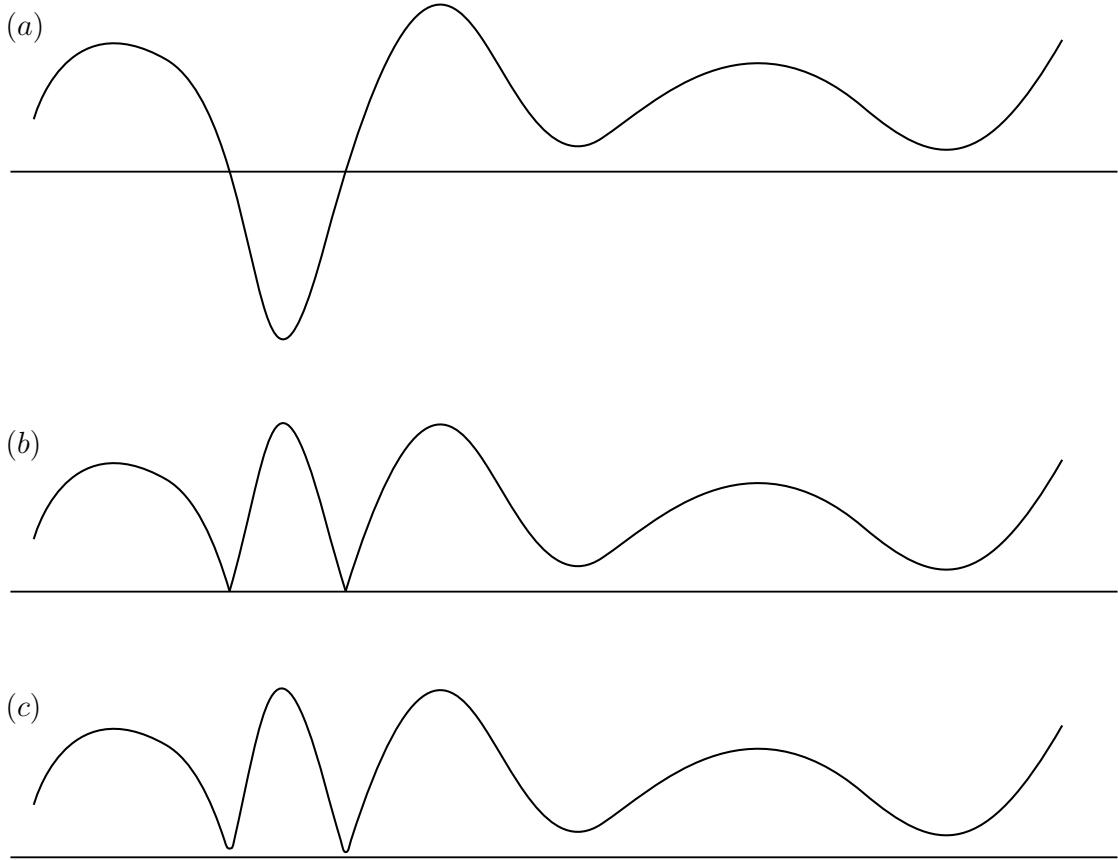
FIGURE 2.1: Evolution of a fermionic wave function into a lower energy bosonic wave function.

Consider fixing all particle positions x_i of the wave function Φ except ξ to obtain the one-dimensional function $\Psi(\xi) = \Phi(x_1, x_2, \ldots, \xi, \ldots)$ depicted in Fig. 2.1(a). Construct a new wave function $\tilde{\Psi}(\xi) = |\Psi(\xi)|$, shown in Fig. 2.1(b). For a real wave function, the energy is

$$E = \frac{\int \Phi \hat{H} \Phi \, d\mathbf{R}}{\int \Phi^2 \, d\mathbf{R}} = \frac{\int [\frac{1}{2} \sum_i (\nabla_i \Phi)^2 + V \Phi^2] \, d\mathbf{R}}{\int \Phi^2 \, d\mathbf{R}}. \tag{2.15}$$

As $\tilde{\Psi}(\xi)$ and $\Psi(\xi)$ have the same values of $(\partial \Phi / \partial \xi)^2$, $(\partial \Phi / \partial x_i)^2$ and Φ^2, they have the same energy. A reduction in this energy can be obtained by smoothing out the kinks in the wave function that occur near the node and decreasing $(\nabla \Phi)^2$ with only a small increase in Φ^2, as shown in Fig. 2.1(c). The same argument can be applied to the other particle coordinates to conclude that the wave function for the bosonic ground state must be nodeless.

To overcome the fermion sign problem, one can construct an antisymmetric function from the difference of two positive symmetric functions, Φ^+ and Φ^-. Release node methods [24] are based on this idea. The two symmetric functions are treated as densities and

two populations of walkers are evolved, one positive and one negative. Although such a method is in principle exact, the signal-to-noise ratio of the population difference decays exponentially in time as both positive and negative populations grow and spread[4].

It has been shown that solving the fermion sign problem is non-deterministic polynomial-time hard (NP-hard) [37]. Nonetheless, DMC calculations are routinely carried out successfully using the fixed-node approximation [38, 39]. The nodes of the DMC wave function are fixed to those of a reference wave function and Eq. 2.14 is solved in each of the nodal pockets with the boundary condition that the wave function goes to zero on the nodal surface. The fixed nodes represent infinite potential barriers that act as sinks for walkers, preventing positive walkers from crossing over to a negative region and vice versa.

The fixed-node approximation works surprisingly well despite being uncontrolled. The fixed-node DMC energy is variational and for an error in the nodal surface of Δ, the error in the energy is $\mathcal{O}(\Delta^2)$ [28]. The nodal surface of the determinant of one-electron orbitals is a good first approximation to the exact nodal surface for many systems. As the wave function is expected to be smooth to minimize kinetic energy, the nodal surface can be assumed to lie in regions of low electron density. Consequently, the energy is expected to be insensitive to small deviations in the nodal surface.

2.3.3 IMPORTANCE SAMPLING

The importance sampling transformation used in DMC [40, 41] replaces the $\Psi^2(\mathbf{R})$ in Eq. 2.9 with $f(\mathbf{R}, \tau) = \Phi(\mathbf{R}, \tau)\Psi(\mathbf{R})$, where $\Psi(\mathbf{R})$ is the input trial wave function and $\Phi(\mathbf{R}, \tau)$ is the DMC wave function. The trial wave function is typically taken from a VMC calculation and acts as a guiding function in the importance sampling. Eq. 2.11 then becomes

$$\frac{\partial f(\mathbf{R}, \tau)}{\partial \tau} = \frac{1}{2}\nabla^2 f(\mathbf{R}, \tau) - \nabla \cdot (\mathbf{v}(\mathbf{R})f(\mathbf{R}, \tau)) - (E_\mathrm{L}(\mathbf{R}) - E_\mathrm{T})f(\mathbf{R}, \tau), \qquad (2.16)$$

where $E_\mathrm{L}(\mathbf{R})$ is the local energy of $\Psi(\mathbf{R})$ and $\mathbf{v}(\mathbf{R}) = \Psi^{-1}(\mathbf{R})\nabla\Psi(\mathbf{R})$ is the drift velocity. The three terms on the right-hand side of Eq. 2.16 represent diffusion, drift and branching respectively. The drift term guides walkers to regions where the magnitude of the trial wave function is larger.

Beyond introducing a drift term, the importance sampling transformation has resolved the two shortcomings of the simple DMC algorithm described in Sec. 2.3.1. Firstly, the potential term V in the branching factor is replaced by E_L, which is much more uniform in configuration space. Secondly, the fixed-node approximation is enforced and the nodes of the DMC wave function Φ are constrained to be those of the trial wave function Ψ. As $f = \Phi\Psi$ must be everywhere positive, Φ and Ψ must be of the same sign everywhere,

---

[4]It is interesting to note that the FCIQMC method described in Sec. 1.2.2 benefits from annihilation of walkers in the discrete space of determinants to overcome the fermion sign problem [4, 9].

change sign together and thus be zero in the same places. The drift term prevents walkers from crossing the nodal surface; the drift velocity grows as a walker approaches a node of Ψ and forces it away. If the time step used to propagate walkers is sufficiently small, walkers never attempt to cross the nodal surface. However, given the finite size of the time step, walkers may diffuse through the nodal surface. The least-biased way of remedying this is to reject such moves[5] [43].

In integral form, Eq. 2.16 becomes

$$f(\mathbf{R}, \tau + d\tau) = \int G(\mathbf{R}' \to \mathbf{R}, d\tau) f(\mathbf{R}', \tau) \, d\mathbf{R}'. \tag{2.17}$$

The time-dependent Green's function[6]

$$G(\mathbf{R}' \to \mathbf{R}, d\tau) = \langle \mathbf{R} | e^{-(\hat{T} + \hat{E}_{\mathrm{L}} - E_{\mathrm{T}})d\tau} | \mathbf{R}' \rangle \tag{2.18}$$

obeys Eq. 2.11 and satisfies the initial condition $G(\mathbf{R}' \to \mathbf{R}, 0) = \delta(\mathbf{R} - \mathbf{R}')$, where $\hat{T} = -\frac{1}{2}\nabla^2 + (\nabla \cdot \mathbf{v}) + (\mathbf{v} \cdot \nabla)$ and $\hat{E}_{\mathrm{L}}$ is the local energy operator. If the Green's function G represents the probability that the particle moves from $\mathbf{R}'$ to $\mathbf{R}$ in imaginary time $d\tau$, a set of walkers initially distributed as $\Psi^2(\mathbf{R})$ can be propagated in imaginary time and their distribution will eventually represent $f = \phi_0 \Psi$, where ϕ_0 is the ground-state wave function. However, the Green's function is not known exactly and must be approximated. The full Green's function is simplified by factoring it into drift-diffusion and branching components,

$$\begin{aligned} G(\mathbf{R}' \to \mathbf{R}, d\tau) &\approx \langle \mathbf{R} | e^{-(\hat{E}_{\mathrm{L}} - E_{\mathrm{T}})d\tau/2} e^{-\hat{T}d\tau} e^{-(\hat{E}_{\mathrm{L}} - E_{\mathrm{T}})d\tau/2} | \mathbf{R}' \rangle \\ &= G_{\mathrm{DD}}(\mathbf{R}' \to \mathbf{R}, d\tau) G_{\mathrm{B}}(\mathbf{R}' \to \mathbf{R}, d\tau). \end{aligned} \tag{2.19}$$

As $\hat{T}$ and $\hat{E}_{\mathrm{L}}$ do not commute, this simplification of the Green's function is only valid for short time steps and becomes exact in the limit $d\tau \to 0$.

2.3.3.1   Drift-diffusion

The drift velocity $\mathbf{v}$ is constant between $\mathbf{R}'$ and $\mathbf{R}$ in the limit of small time steps. The drift-diffusion Green's function is then

$$G_{\mathrm{DD}}(\mathbf{R}' \to \mathbf{R}, d\tau) = \frac{1}{(2\pi d\tau)^{3N/2}} \exp\left(-\frac{(\mathbf{R} - \mathbf{R}' - \mathbf{v}(\mathbf{R}')d\tau)^2}{2d\tau}\right). \tag{2.20}$$

---

[5]Rejecting moves that cross the nodal surface renders the random walk non-ergodic. However, the tiling theorem [42] states that all ground-state nodal pockets are related by permutation symmetry and thus the Schrödinger equation only needs to be solved in one such pocket.

[6]The time-dependent formalism presented here is closely connected to an analogous time-independent Green's function formalism [30].

In the drift-diffusion process, each walker $\mathbf{R}'$ drifts a distance $\mathbf{v}(\mathbf{R}')d\tau$ and diffuses by a random normally-distributed distance χ with variance $d\tau$ such that

$$\mathbf{R} = \mathbf{R}' + \mathbf{v}(\mathbf{R}')d\tau + \chi. \tag{2.21}$$

Under the assumption that the drift velocity is constant between $\mathbf{R}'$ and $\mathbf{R}$, $\hat{T}$ is not Hermitian, G_{DD} is not symmetric and consequently detailed balance is not satisfied. Detailed balance is imposed by a Metropolis accept-reject step [44] in which a move from $\mathbf{R}'$ to $\mathbf{R}$ is accepted with probability

$$A(\mathbf{R}' \to \mathbf{R}) = \min\left(1, \frac{\Psi^2(\mathbf{R})T(\mathbf{R} \to \mathbf{R}')}{\Psi^2(\mathbf{R}')T(\mathbf{R}' \to \mathbf{R})}\right), \tag{2.22}$$

where $T(\mathbf{R} \to \mathbf{R}') = G_{\mathrm{DD}}(\mathbf{R} \to \mathbf{R}')$. The accept-reject step reduces time-step errors as configuration space is sampled correctly regardless of the time step. As rejecting some moves reduces the mean-square distance $\langle r^2 \rangle = 3\tau$ that each electron diffuses in time τ, the effective time step $d\tau_{\mathrm{eff}}$ used in calculating the branching factor is determined for the actual distance diffused [35]:

$$d\tau_{\mathrm{eff}} = d\tau \frac{\langle r^2_{\mathrm{accepted}} \rangle}{\langle r^2_{\mathrm{attempted}} \rangle}. \tag{2.23}$$

2.3.3.2 BRANCHING

The branching Green's function is

$$G_{\mathrm{B}}(\mathbf{R}' \to \mathbf{R}, d\tau) = \exp\left(-\frac{d\tau_{\mathrm{eff}}}{2}(E_{\mathrm{L}}(\mathbf{R}) + E_{\mathrm{L}}(\mathbf{R}') - 2E_{\mathrm{T}})\right). \tag{2.24}$$

The branching process is simulated by allowing a walker to breed additional child walkers or die based on its weight after all electrons have been moved. A walker will produce `int`$(G_{\mathrm{B}}+\eta)$ identical child walkers at its position which then evolve independently, where η is a random number between 0 and 1 and `int` returns the integer component of a real number. Alternatively, a combination of branching and weighting is used. Here, each walker α carries a weight w_i^α that will be multiplied by G_{B}. When this weight becomes large, the walker branches and the non-integer component of the weight, $w_i^\alpha - $`int`$(w_i^\alpha)$, is distributed between the `int`(w_i^α) child walkers. Walkers with weight below a chosen threshold are killed and their weight transferred to another walker.

The reference energy E_{T} is adjusted to limit fluctuations in the size of the population. These changes in E_{T} were not considered in the derivation of the DMC Green's function. As walkers are restrained from multiplying in regions of low local energy and dying out in regions of high local energy, this introduces a positive population control bias in the DMC energy [43]. This bias is inversely proportional to the population size, and is negligible for typical populations. The reference energy also provides an estimate of the ground-state

energy.

If Ψ is equal to the exact ground-state wave function ϕ_0, then $E_L = E_0 = E_T$ and the branching term disappears. Without the branching term, $\Phi \to \Psi$ and $f \to \Psi^2$, and DMC sampling reduces to that of VMC. For an approximate Ψ, the branching term serves to improve the sampling of configuration space by reducing the number of walkers in high energy regions and increasing their number in low energy regions. In effect, branching compensates for the error in Ψ.

There exists some correlation between the walkers as the child walkers resulting from the branching process begin their random walk at the same point in configuration space. This correlation causes the cost of DMC calculations to scale exponentially [45]. Fortunately, the exponential scaling takes over only for large system sizes and/or poor trial wave functions. Most DMC studies in practice concern smaller systems for which this unfavourable scaling is not observed. The VMC algorithm does not rely on branching and thus does not exhibit this exponential scaling.

2.3.4 EXPECTATION VALUES

Once the walk has equilibrated, the statistics of interest are accumulated using as many independent samples as necessary to resolve the estimate to a given accuracy. The mixed distribution $f = \Phi\Psi$ can be used to estimate the expectation value of any Hermitian operator $\hat{A}$ that commutes with the Hamiltonian, giving the mixed estimate

$$A_{\mathrm{DMC}} = \frac{\langle \Phi | \hat{A} | \Psi \rangle}{\langle \Phi | \Psi \rangle} = \frac{\int f(\mathbf{R}) A_{\mathrm{L}}(\mathbf{R}) \, d\mathbf{R}}{\int f(\mathbf{R}) \, d\mathbf{R}} \approx \frac{\sum_i^M W(i) A(i)}{\sum_i^M W(i)}. \tag{2.25}$$

The average value of A_{L} over the $N_C(i)$ configurations present after i time steps is

$$A(i) = \frac{\sum_\alpha^{N_C(i)} w_i^\alpha A_{\mathrm{L}}(\mathbf{R}_i^\alpha)}{\sum_\alpha^{N_C(i)} w_i^\alpha}, \tag{2.26}$$

where $A_{\mathrm{L}}(\mathbf{R}_i^\alpha) = \Psi^{-1}(\mathbf{R}_i^\alpha) \hat{A} \Psi(\mathbf{R}_i^\alpha)$ is the local value of $\hat{A}$ for configuration α. The weight of the ith iteration is

$$W(i) = \sum_\alpha^{N_C(i)} w_i^\alpha. \tag{2.27}$$

The corresponding squared error of the estimate A_{DMC} is

$$\sigma^2 \approx \frac{\sum_i^M W(i)[A(i) - A_{\mathrm{DMC}}]^2}{M \left(\sum_i^M W(i) - \frac{\sum_i^M [W(i)]^2}{\sum_i^M W(i)} \right)}. \tag{2.28}$$

The averages are still evaluated using the trial (VMC) wave function while the branching process weights the sampling of configuration space in DMC. If the equilibration period

is sufficiently long, f can be taken to be f_∞, and the DMC local energy will be exact for the given nodal surface of Ψ. However, some time-step bias remains as the Green's function is only valid for small time steps $d\tau$; this bias must be removed by extrapolating to $d\tau = 0$ or by using a sufficiently small time step.

Fig. 2.2 depicts the average local energy, reference energy and best estimate of energy alongside the walker population for the initial 2×10^5 steps of a DMC calculation for N_2. The first 50000 steps form the equilibration period, where expectation values are not accumulated. The best estimate of the energy is smooth in the accumulation phase and the large numbers of samples are needed only to reduce the error in the energy estimate. It is also clear that the reference energy mirrors the fluctuations in the population such that the population oscillates around the desired population of 2048 walkers.
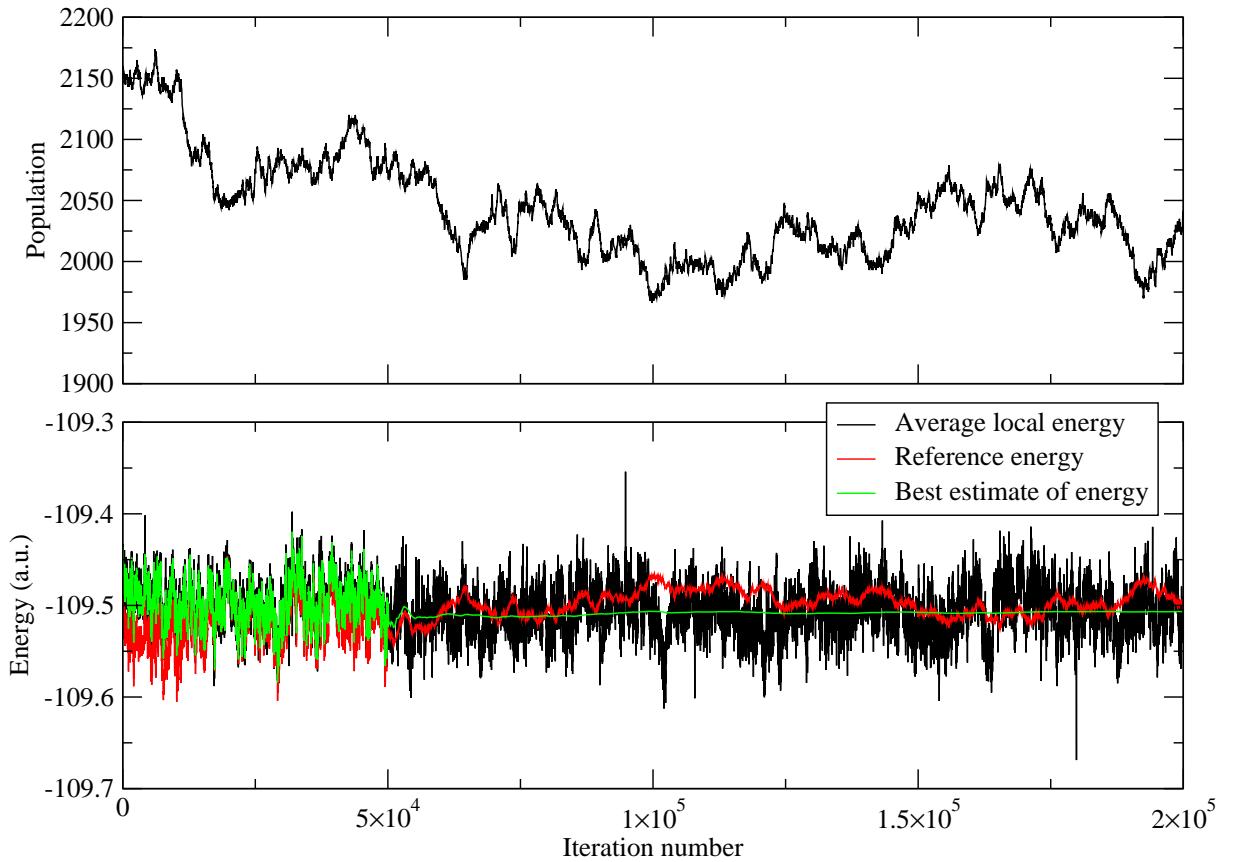
FIGURE 2.2: Energy and population analysis for the first 2×10^5 steps of the DMC calculation for N_2 reported in Chapter 4.

The ground-state wave function is not suitable for estimating expectation values of operators $\hat{A}$ that do not commute with the Hamiltonian as it is not an eigenfunction of $\hat{A}$. The systematic error in both the VMC estimate A_{VMC} and the DMC mixed estimate A_{DMC} are then linear in the error in the wave function Δ, reflecting the fact that the choice of ground-state wave function is not optimal for the importance sampling. This systematic error can be reduced cheaply by combining the VMC and DMC estimates in the extrapolated estimate $A_{ext} = 2A_{DMC} - A_{VMC} + \mathcal{O}(\Delta^2)$, such that the resulting error is

quadratic in the error in the wave function. An unbiased pure estimate can be obtained using the more expensive forward-walking DMC [46, 30] or reptation Monte Carlo [27] methods. Neither of these methods were used in the work presented in this thesis.

2.4 THE TRIAL WAVE FUNCTION

The choice of trial wave function is important in both VMC and DMC. In VMC, the trial wave function is used in the evaluation of expectation values and it directly determines the final accuracy obtained. The accuracy of DMC is determined by the quality of the nodal surface of the wave function. For both methods, the statistical efficiency is dependent on the trial wave function. As VMC and DMC allow any form of trial wave function to be used, it is desired to design a physically-sensible form that builds in sufficient variational freedom compactly.

2.4.1 CHARACTERISTICS OF THE WAVE FUNCTION

The trial wave function Ψ must satisfy several conditions. It should have the correct antisymmetry under particle exchange. Both Ψ and its first derivative must be continuous everywhere except where the potential diverges. The integrals $\int \Psi^2 \, d\mathbf{R}$, $\int \Psi \hat{H} \Psi \, d\mathbf{R}$ and $\int \Psi \hat{H}^2 \Psi \, d\mathbf{R}$ must exist to ensure respectively normalizability, that the energy is finite and that the variance is finite. As the bulk of the computational effort is spent on the evaluation of the wave function and its first and second derivatives, the wave function must have a compact form allowing rapid evaluation. The importance of a compact and physically-meaningful trial wave function is further emphasized by the fact that the error in an estimate may increase with the number of optimizable parameters [47].

Though VMC and DMC allow the use of complex wave functions [48], only real wave functions are used here as the Hamiltonians considered all have time-reversal symmetry and thus are real[7]. Additionally, as the operators of interest such as the Hamiltonian are spin-independent, they commute with both the total spin operator $\hat{S}^2$ and the spin projection operator $\hat{S}_z$. All electron spins are thus assumed to be collinear[8]. The spin-dependence of the wave function can be eliminated and further simplifications can be made as shown below.

---

[7]For real Hamiltonians $\hat{H}$, $\hat{H}\Psi = E_0\Psi$ implies that $\hat{H}\Psi^* = E_0\Psi^*$. Either Ψ and Ψ^* differ by a multiplicative phase factor or they are degenerate. In either case, the linear combination $\Psi + \Psi^*$ is a real eigenfunction, and if Ψ is purely imaginary, $i\Psi$ is a real eigenfunction. If several degenerate solutions exist, additional real eigenfunctions such as $i(\Psi - \Psi^*)$ can be constructed by considering the space of degenerate eigenfunctions orthogonal to these solutions.

[8]QMC is able to simulate non-collinear spins as well [49].

The expectation value of operator $\hat{A}$ is

$$\langle \hat{A} \rangle = \frac{\sum_{\boldsymbol{\sigma}} \int \Psi(\mathbf{X}) \hat{A} \Psi(\mathbf{X}) \, d\mathbf{R}}{\sum_{\boldsymbol{\sigma}} \int \Psi(\mathbf{X}) \Psi(\mathbf{X}) \, d\mathbf{R}}, \tag{2.29}$$

where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)$ is the spin configuration. A fermionic wave function is anti-symmetric with respect to exchange of the spatial and spin coordinates of two electrons. Therefore $\Psi(\mathbf{X})$ can be replaced, subject to a change of sign, by $\Psi(\mathbf{X}')$ which has a spin configuration such that the first $N_\uparrow$ electrons are up-spin electrons and the last $N_\downarrow$ electrons are down-spin electrons:

$$\begin{aligned} \Psi(\mathbf{X}) &= \Psi(\{\mathbf{r}_1, \sigma_1\}, \{\mathbf{r}_2, \sigma_2\}, \ldots, \{\mathbf{r}_N, \sigma_N\}) \\ &= \pm \Psi(\mathbf{X}') = \pm \Psi(\{\mathbf{r}_{i_1}, \uparrow\}, \ldots, \{\mathbf{r}_{i_{N_\uparrow}}, \uparrow\}, \{\mathbf{r}_{i_{N_\uparrow+1}}, \downarrow\}, \ldots, \{\mathbf{r}_{i_N}, \downarrow\}). \end{aligned} \tag{2.30}$$

Relabelling the dummy integration variables $\mathbf{r}_{i_1}, \ldots, \mathbf{r}_{i_N}$ gives

$$\Psi(\mathbf{X}') = \Psi(\{\mathbf{r}_1, \uparrow\}, \ldots, \{\mathbf{r}_{N_\uparrow}, \uparrow\}, \{\mathbf{r}_{N_\uparrow+1}, \downarrow\}, \ldots, \{\mathbf{r}_N, \downarrow\}). \tag{2.31}$$

Substituting $\Psi(\mathbf{X})$ with $\Psi(\mathbf{X}')$ clearly does not affect the expectation value of spin-independent operators. Thus, each spin configuration contributes equally and the sums over the spin configurations cancel. The expectation value of $\hat{A}$ becomes

$$\langle \hat{A} \rangle = \frac{\int \Psi(\mathbf{R}) \hat{A} \Psi(\mathbf{R}) \, d\mathbf{R}}{\int \Psi(\mathbf{R}) \Psi(\mathbf{R}) \, d\mathbf{R}}, \tag{2.32}$$

where $\Psi(\mathbf{R}) = \Psi(\{\mathbf{r}_1, \uparrow\}, \ldots, \{\mathbf{r}_{N_\uparrow}, \uparrow\}, \{\mathbf{r}_{N_\uparrow+1}, \downarrow\}, \ldots, \{\mathbf{r}_N, \downarrow\})$. As $\Psi(\mathbf{R})$ is only antisymmetric with respect to exchange of electrons of the same spin, electrons of different spin are treated as distinguishable particles.

2.4.2   THE SLATER DETERMINANT

The Slater determinant, as introduced in Sec. 1.2.1, is an antisymmetrized product of one-electron orbitals. These orbitals determine the nodal surface and must be of good quality. They are generally obtained from HF or DFT calculations. The Slater determinant as given by Eq. 1.8 is decomposed into a product of up-spin and down-spin components since the wave function only needs to be antisymmetric with respect to exchange of like-spin electrons,

$$\Psi_{\mathrm{D}}(\mathbf{R}) = \begin{vmatrix} \psi_1^\uparrow(\mathbf{r}_1^\uparrow) & \cdots & \psi_1^\uparrow(\mathbf{r}_{N^\uparrow}^\uparrow) \\ \vdots & \ddots & \vdots \\ \psi_{N^\uparrow}^\uparrow(\mathbf{r}_1^\uparrow) & \cdots & \psi_{N^\uparrow}^\uparrow(\mathbf{r}_{N^\uparrow}^\uparrow) \end{vmatrix} \begin{vmatrix} \psi_1^\downarrow(\mathbf{r}_1^\downarrow) & \cdots & \psi_1^\downarrow(\mathbf{r}_{N^\downarrow}^\downarrow) \\ \vdots & \ddots & \vdots \\ \psi_{N^\downarrow}^\downarrow(\mathbf{r}_1^\downarrow) & \cdots & \psi_{N^\downarrow}^\downarrow(\mathbf{r}_{N^\downarrow}^\downarrow) \end{vmatrix}. \tag{2.33}$$

The up-spin orbitals $\psi^\uparrow$ and down-spin orbitals $\psi^\downarrow$ are not necessarily the same, although they may be constructed to be so.

2.4.3 THE JASTROW FACTOR

Standard quantum chemistry methods are obliged to use a large number of Slater determinants to account for both the behaviour of the wave function near particle coalescence and at large inter-particle separations as it difficult to approximate the cusps in the wave function as a truncated sum of smooth functions. These cusps keep the local energy finite. The divergence in the potential energy as charged particles coalesce is cancelled by an opposite divergence in the kinetic energy induced by the cusps. As such, cusps are a feature of the exact wave function. Although advances in explicitly-correlated methods are being made, the difficulty in factorization of the high-dimensional integrals prevents the widespread use of correlating factors in traditional quantum chemistry methods. Methods that employ stochastic integration such as VMC and DMC do not suffer from such problems and are able to use correlating factors.

A Jastrow correlation factor [50] is multiplied to the determinantal part of the wave function to give a Slater-Jastrow (SJ) wave function

$$\Psi(\mathbf{R}) = e^{J(\mathbf{R})}\Psi_{\mathrm{D}}(\mathbf{R}), \tag{2.34}$$

where the Jastrow function $J(\mathbf{R})$ is an explicit function of the inter-particle distances. Thus, the Jastrow factor allows both long- and short-range dynamical correlation effects that depend on the positions of the electrons to be included in a compact and efficient manner. The Jastrow factor is constrained so that the symmetry and boundary properties of $\Psi_{\mathrm{D}}(\mathbf{R})$ are transferred unmodified to $\Psi(\mathbf{R})$.

We have developed a Jastrow factor that goes beyond the standard forms used in the QMC community to better describe electronic correlation. This generic Jastrow factor is described and results obtained are reported in Chapter 4.

2.4.4 BACKFLOW TRANSFORMATIONS

The Jastrow factor is able to account for some correlation and reduce the VMC energy, but it is restricted to be everywhere positive so as to maintain fermionic symmetry. Therefore, the addition of the Jastrow factor cannot alter the nodal surface determined by the determinantal component of the wave function Ψ_{D}. Backflow transformations provide a means to improving the nodal surface, and thus reduce the uncontrolled nodal error in DMC.

A Slater-Jastrow-backflow (SJB) wave function takes the form

$$\Psi(\mathbf{R}) = e^{J(\mathbf{R})}\Psi_{\mathrm{D}}(\mathbf{X}(\mathbf{R})), \tag{2.35}$$

where $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ and

$$\mathbf{x}_i = \mathbf{r}_i + \boldsymbol{\xi}_i(\mathbf{R}) \tag{2.36}$$

are the backflow-transformed quasiparticle coordinates at which the orbitals are evaluated. The correct antisymmetry of the wave function is maintained as antisymmetrization is performed on the backflow-transformed orbitals and the transformation is constrained to not introduce cusps. As the backflow-transformed position of each electron is a function of the position of all the other electrons via the backflow function $\boldsymbol{\xi}_i$, some correlation is built into the wave function.

The motivation for backflow transformations is presented in Chapter 5, along with preliminary work on the development of generalized backflow transformations that allow for greater variational freedom in the wave function.

2.4.5 Multi-determinant expansions

A multi-determinant expansion is necessary to describe the static correlation in systems with near-degeneracy, where excited states have sufficiently low energies to be able to mix with the ground state, or close to the dissociation limit of a molecule. Typically, only a small number of the determinants, those corresponding to the nearly-degenerate excited states, are needed to capture this static correlation. The higher excitations generally describe the dynamical correlation that the Jastrow factor is able to account for more efficiently. The addition of multi-determinants also modifies the nodal surface and can improve DMC energies.

Rather than expanding the wave function in the basis of determinants which are eigenfunctions of $\hat{S}_z$ only, it is preferable to work in the basis of spin-adapted linear combinations of determinants belonging to the same orbital configuration, known as configuration state functions (CSF). The wave function is then an eigenfunction of both $\hat{S}^2$ and $\hat{S}_z$, and has fewer optimizable parameters than if a basis of determinants is used.

For a multi-determinant expansion, the determinantal component of the wave function becomes

$$\Psi_{\mathrm{D}}(\mathbf{R}) = \sum_{j=1}^{N_{\mathrm{CSF}}} c_j \sum_{k=1}^{N_{\mathrm{det}}^j} d_{k,j} D_{k,j}^{\uparrow}(\mathbf{R}_{\uparrow}) D_{k,j}^{\downarrow}(\mathbf{R}_{\downarrow}), \tag{2.37}$$

where N_{CSF} is the number of CSFs, N_{det}^j is the number of determinants D in the jth CSF which has an optimizable weight c_j. The relative weights of the determinants $d_{k,j}$ are fixed to maintain the spin symmetry.

Each excited determinant in a multi-determinant expansion used in QMC typically differs from the ground-state determinant by one or two orbitals, corresponding to single and double excitations. This fact can be used to combine multiple determinants to reduce the size of the expansion and thereby reduce the cost of evaluating the wave function. The determinants can be combined in many different ways, and so constructing the optimal

contraction is a non-trivial task [51]. Such contractions have been found to reduce the number of determinants in the expansion substantially [52].

2.4.6 PAIRING WAVE FUNCTIONS

Rather than constructing a determinantal expansion using single-particle orbitals, pairing determinants can be constructed using orbitals that are a function of the coordinates of two particles. One common example of such a pairing wave function is the Bardeen-Cooper-Schrieffer (BCS) wave function, which is an antisymmetrized product of singlet pairing orbitals. These pairing orbitals are referred to as geminals, and are particularly suited to systems where strong electronic correlation is expected. The explicit pairing of electrons in orbitals introduces correlation in these wave functions, and they are an alternative to multi-determinant expansions.

In the simplest pairing wave function, the determinantal part can be replaced by an antisymmetrized product of geminals (AGP) [53]. For an unpolarized system, the AGP is given by $\Psi_{\text{AGP}}(\mathbf{R}) = \det[\phi(\mathbf{r}_i^\uparrow, \mathbf{r}_j^\downarrow)]$. This definition can easily be extended to spin-polarized systems. The geminals ϕ can be expanded in a basis of single-particle orbitals $\psi_i(\mathbf{r})$. Such wave functions have been used successfully in QMC calculations [54, 55]. Pfaffians are another type of antisymmetrized function based on pairing orbitals that can be used in QMC calculations [56, 57]. Both these antisymmetrized pairing forms can be used in conjunction with a Jastrow factor and backflow transformations.

2.5 ERRORS, STATISTICS AND IMPLEMENTATION

There are several sources of systematic error in VMC and DMC calculations beyond the statistical errors that arise due to the Monte Carlo algorithm used in estimating expectation values.

The only systematic error in a VMC calculation is due to the functional form of the trial wave function. Of the errors and biases resulting from the DMC algorithm, including the time-step and population control biases discussed in Sec. 2.3, the nodal error that arises from the uncontrolled fixed-node approximation is the most severe.

The cost of DMC calculations is found [58] to scale as $Z^{5.5}$ as an increasingly smaller time step must be used to capture the high frequency oscillations of the wave functions in the core region. All-electron calculations are thus only feasible for systems with $Z \lesssim 10$. For heavier atoms, pseudopotentials are used to remove the inert core electrons. The localization [59] or semi-localization [60, 61] of non-local pseudopotentials gives rise to further errors in DMC. Pseudopotentials were not used to obtain any of the results presented in this thesis.

A bulk solid of infinite size cannot be studied and must be approximated by a finite system. Typically, a finite simulation cell is used in conjunction with periodic boundary
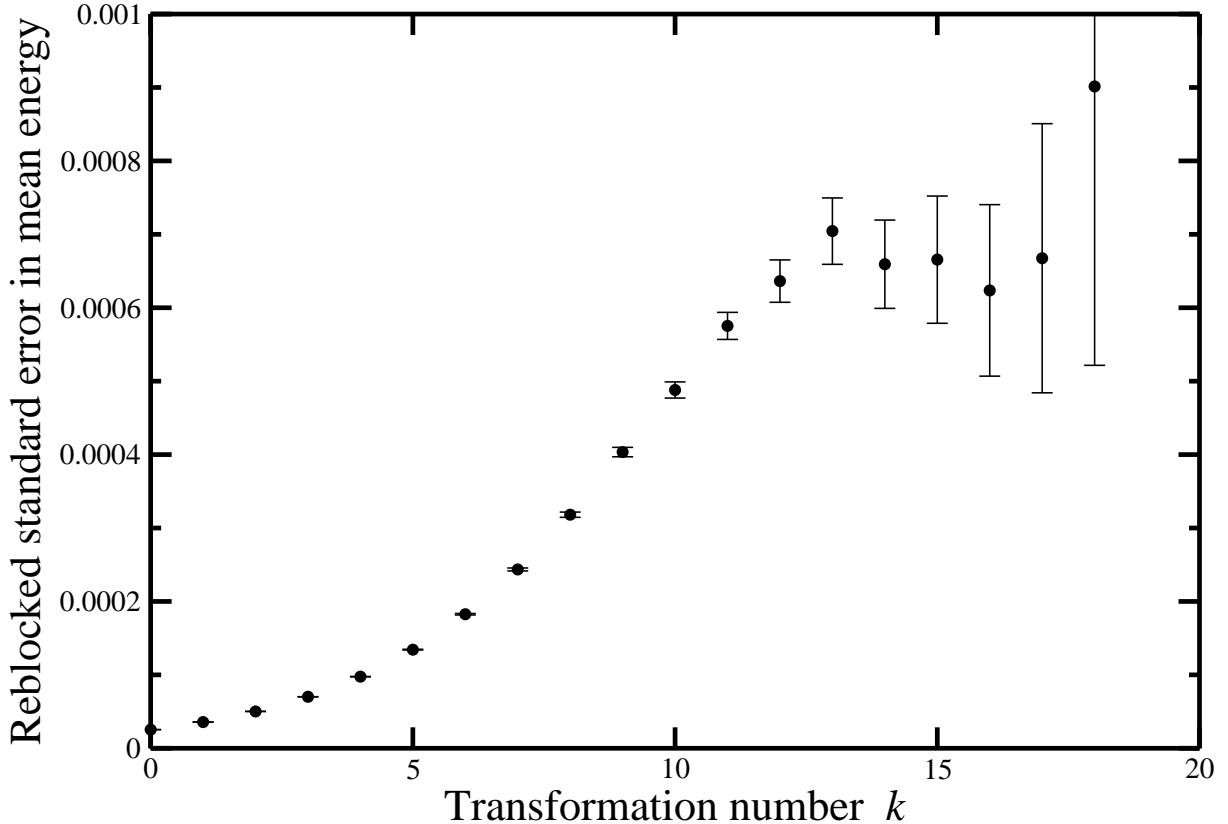
conditions to allow reproduction of the infinite system as well as possible. This approximation introduces finite-size errors into both VMC and DMC results. Various electron-electron potentials and extrapolation schemes have been developed to reduce these errors [62, 63].

Any quantity calculated in the Monte Carlo framework has associated with it a statistical error that determines the confidence intervals of the estimate. Obtaining these error bars is theoretically straightforward. In practice, two problems are encountered: the individual data points are serially correlated and are frequently not Gaussian-distributed. These two issues are expounded in the following sections.

2.5.1 SERIAL CORRELATION

Measurements made on configurations too close sequentially in the random walk will be serially correlated, giving a deceptively low estimate of the variance and thus error. A simplified situation that demonstrates this is a random walk in which a move is only infrequently accepted, say, every T steps. The mean is calculated correctly but the expressions for the squared error, Eqs. 2.10 and 2.28, assume that successive data points are statistically independent and thus underestimate σ^2 by a factor of T. Furthermore, sequential configurations in the walk can be very alike even if moves are accepted frequently and thus the configurations are correlated. In either case, new information regarding the system is gathered only from data points that are separated in the random walk by the correlation time T_{corr}. While both VMC and DMC are affected by serial correlation, it is particularly severe in DMC where the time steps must be small to minimize the time-step bias, and typically over 99% of the steps are accepted.

The correlation time T_{corr} can be estimated in several ways. It can be obtained by approximating the autocorrelation function [30]. The reblocking algorithm [64] can also be used to remove serial correlation and is preferable as it can be performed on-the-fly, reducing memory requirements. Here, the data are divided into blocks containing two adjacent data points which are averaged to give a block average. The variance of these block averages is computed. This is done recursively so that the blocks contain 2^k elements after the kth transformation. Typically the variance increases until $2^k \geq T_{\text{corr}}$. For larger blocks still, the variance does not vary but the error on the variance increases due to the reduced number of samples. The onset of the plateau indicates when serial correlation has been removed and can be used to estimate T_{corr} and the reblocked variance. The reblocked standard error from a DMC calculation on the N_2 molecule is plotted as a function of the transformation number k in Fig. 2.3. The plateau region begins approximately at $k = 12$, giving a standard error of 0.0007.

FIGURE 2.3: Reblocking analysis for the DMC calculation on N_2 reported in Chapter 4.

2.5.2 NON-GAUSSIAN DISTRIBUTIONS OF ESTIMATES

In QMC it is usually assumed that a sufficient number of data points have been accumulated and that they are distributed such that the central limit theorem (CLT) is valid. The choice of $P = \lambda \Psi^2$ where λ is the normalization constant results in VMC estimates that are not Gaussian distributed for many expectation values, including the total energy and its variance [65, 66]. As a result of the singularities in the local energy arising from particle coalescence and the divergence in kinetic energy at the nodal surface, the probability distribution of local energies can be shown to have heavy tails that decay as E^{-4}. The CLT is only weakly valid for the total energy as the energy distribution is Gaussian only for an infinite number of samples; thus outliers occur for finite sampling. The CLT does not apply at all for the variance of the energy, and the random error is not Gaussian even for an infinite number of samples. Outliers are more than a magnitude more likely than expected from Gaussian statistics [66]. In practice, reasonably accurate estimates of the errors in the local energy can be obtained if its distribution is sufficiently narrow. Singularities in other quantities on the nodal surface lead to similar, and occasionally even more severe, non-Gaussian behaviour of estimates.

A sampling distribution without a nodal surface can be chosen to eliminate the algebraic asymptotic tails in the distribution of local energy and its variance [66]. Without heavy non-exponential tails, all moments of the distribution of local energy would exist

and the CLT would be valid. Unbiased estimates that account for finite sampling effects would then be available. The optimal choice of sampling distribution weights the standard Ψ^2 distribution by a factor proportional to the deviation of the local energy from the exact energy [47]. However, this sampling scheme is not practical as the cost of the energy evaluations at every step outweighs the gain in accuracy. Furthermore, the exact energy is not known. A much more efficient nodeless distribution can be constructed by taking a linear combination of the ground-state determinant and an excited state determinant. By sampling the nodal surface of the underlying probability distribution function, this distribution converges much more quickly and thus requires only a fraction of the samples to obtain an estimate to within a given accuracy. Although the benefit of efficient sampling is likely to fall with system size, other sampling strategies can be developed for large systems.

The non-Gaussian behaviour has important consequences for the optimization of wave functions using cost functions based on the variance of the energy or other quantities. In addition to the inaccurate estimates of the parameters, standard sampling suppresses the contribution of the nodal surface. This is undesirable as the methods for improving the nodal surface in VMC such as the inclusion of backflow transformations and multi-determinant expansions rely on sampling in the region close to the nodal surface for successful optimization.

The efficient sampling scheme has been shown [47] to improve the statistical properties of the standard error of VMC estimate as a result of the error in the optimized parameters. Fewer samples are required to obtain an accurate minimum compared to standard sampling. In addition to improving the accuracy of estimates, the efficient sampling scheme gives theoretical justification for averaging converged parameter sets to reduce the error due to optimization [47].

2.5.3 IMPROVING THE EFFICIENCY OF A QMC CALCULATION

The most obvious way in which QMC can be made more efficient is by improving the trial wave function. This is true for both VMC and DMC, in which the quality of the wave function determines the variance of the energy. Using more complex wave functions that take longer to evaluate can often be more efficient overall. The construction of wave functions with sufficient variational freedom and the development of robust minimization schemes are crucial to this end.

Serial correlation between the subsequent configurations in the random walk also reduces the efficiency of a calculation, as discussed in Sec. 2.5.1. In VMC the trial moves are chosen from a Gaussian centred on the walker with variance equal to the VMC time step[9] $d\tau$. To ensure efficient sampling, the time step should be chosen such that the rate

---

[9]The VMC time step is named as such only in analogy to the DMC formalism. As the VMC method

at which the walkers diffuse through configuration space is maximized. If a fraction $\bar{A}$ of M moves is accepted, the root-mean-square distance diffused by an electron is $\sqrt{3M\bar{A}d\tau}$. Too short a time step will result in excessive serial correlation, while a large proportion of trial moves will be rejected if the time step is too large. A rule of thumb that works well in practice is choosing a time step such that 50% of the moves are accepted, i.e., $\bar{A} = 0.5$. Umrigar [67] developed an accelerated Metropolis algorithm uses variable time steps to evolve the system more rapidly and thus sample configuration space more efficiently. By proposing moves for electrons that are proportional to its distance from the nucleus, the shorter length scales of the tightly bound core electrons is distinguished from the longer length scales of the chemically-active valence electrons. Serial correlation is reduced by proposing larger moves in the valence region.

Furthermore, evaluating the contribution of a configuration to expectation values being accumulated can be expensive. Only statistically independent configurations in the random walk should be considered in computing expectation values to ensure that computational effort is not wasted [68].

Rather than moving all electrons at once to a new proposed configuration and then performing the accept-reject step, it is much more efficient to apply an accept-reject step on individual electron moves. A larger time step can be used when electrons are moved individually as the probability of rejecting N one-electron moves is significantly lower than the probability of rejecting a configuration move, which tends to one as $N \to \infty$ for a given time step. The evaluation of a new determinant when using one-electron moves can be evaluated in $\mathcal{O}(N^2)$ operations using the Sherman-Morrison update algorithm [32] instead of the $\mathcal{O}(N^3)$ operations needed to evaluate a determinant from scratch. In the event of a rejected step, only the effort made in moving one electron is wasted. Moving electrons individually is found to be more efficient even with backflow wave functions [69, 68] as the correlation length of configurations is shorter.

Quantities being accumulated are usually calculated after a trial move has been accepted. Instead, by calculating the average of the quantity O as

$$\langle O \rangle = \frac{1}{M} \sum_i \left[A_i O_i(\mathbf{R}) + (1 - A_i) O_i(\mathbf{R}') \right], \tag{2.38}$$

information from both accepted and rejected moves is used and hence the variance of O is reduced [34]. This method is useful in VMC where approximately 50% of the moves are rejected. It is also useful when evaluating expectation values that low probability configurations give large contributions to. One example of such a situation is the the evaluation of the divergent local energy near the nodal surface. It is possible, however, that the benefit of lower variance is outweighed by the high cost of evaluating O_i.

The use of a two-level sampling scheme [70] in VMC further improves efficiency. If

---

is based on the time-independent Schrödinger equation, it is free from the notion of time.

the transition probability is symmetric and ergodic, the acceptance probability of a move from $\mathbf{R}'$ to $\mathbf{R}$ is

$$A(\mathbf{R}' \to \mathbf{R}) = \min\left(1, \frac{\Psi^2(\mathbf{R})}{\Psi^2(\mathbf{R}')}\right). \tag{2.39}$$

For a wave function of the form $\Psi = e^J \Psi_{\mathrm{D}}$, this accept-reject step can be replaced by two sequential accept-reject steps while preserving detailed balance. At the first level, a move is accepted with probability

$$A_{\mathrm{D}}(\mathbf{R}' \to \mathbf{R}) = \min\left(1, \frac{\Psi_{\mathrm{D}}^2(\mathbf{R})}{\Psi_{\mathrm{D}}^2(\mathbf{R}')}\right). \tag{2.40}$$

If the move is rejected at this first level, it is rejected completely and the second level acceptance probability

$$A_{\mathrm{J}}(\mathbf{R}' \to \mathbf{R}) = \min\left(1, \frac{\exp[2J(\mathbf{R})]}{\exp[2J(\mathbf{R}')]}\right) \tag{2.41}$$

need not be computed. To be accepted, the move must be accepted at both the first and second levels. The overall probability of acceptance is then $A(\mathbf{R}' \to \mathbf{R}) = A_{\mathrm{D}}(\mathbf{R}' \to \mathbf{R})A_{\mathrm{J}}(\mathbf{R}' \to \mathbf{R})$. Whether a proposed move is accepted or rejected is typically determined by the ratio of the Slater determinants for the two configurations. As the Jastrow factor does not need to be evaluated for moves rejected at the first level, the computational effort per proposed move is reduced. Note that this sampling scheme cannot be used in conjunction with calculating averages using both accepted and rejected steps as the full acceptance probability is not known.

2.6 OPTIMIZATION

The accuracy of the trial wave function is important in both VMC and DMC calculations. For a small error Δ in the normalized trial wave function, it can be shown that the error in the estimate of the local energy is of $\mathcal{O}(\Delta^2)$. It is thus crucial to improve the analytic form of the trial wave function in VMC by stochastic optimization of the parameters[10].

Optimization in VMC is difficult due to the statistical noise that accompanies an estimate and the presence of non-linear parameters in the wave function. Despite much progress in the development of algorithms for optimizing trial wave functions, the optimization procedure generally consumes the most human and computer time in QMC calculations. This further motivates the development of robust and efficient algorithms that converge quickly to the best parameter sets.

The cost function must correspond to a measure of the quality of the wave function. As the cost function will be evaluated using Monte Carlo integration, it must also have a

---

[10]Parameters have successfully been optimized by hand in DMC [71, 72].

small variance. The various stochastic methods that have been developed for optimizing many-body trial wave functions containing hundreds of parameters broadly fall into two categories: those that minimize a variance-based quantity and those that minimize the energy. Variance minimization was used in most QMC calculations until recently for the reasons discussed below, despite the physical variational principle backing energy minimization.

Before the variance and energy minimization methods are discussed, the correlated sampling scheme which is crucial in robust minimization algorithms is introduced.

2.6.1 CORRELATED SAMPLING

Evaluating derivatives of VMC estimates requires evaluating the difference between quantities with statistical error bars. To resolve the small differences in estimates such as (but not limited to) the variance for slightly different parameter sets, the individual values must be known to a higher accuracy than the difference; otherwise the difference will be swamped by noise. Even as the difference $\Delta = \langle \hat{O} \rangle_A - \langle \hat{O} \rangle_B \to 0$, the error $\sigma_\Delta = \sqrt{\sigma_A^2 + \sigma_B^2} \nrightarrow 0$.

Correlated sampling uses correlation between quantities to reduce the variance of the difference, resulting in more accurate estimates of differences. A set of configurations distributed according to the square of the wave function Ψ_A^2 is generated. The expectation value of $\hat{O}$ for Ψ_B^2 can be evaluated as the average of the local value for Ψ_B, $O_L^B = \Psi_B^{-1} \hat{O} \Psi_B$, weighted by $w_A^B = \Psi_B^2 / \Psi_A^2$ over the configurations distributed as Ψ_A^2,

$$\langle \hat{O} \rangle_B = \frac{\int \Psi_B^2(\mathbf{R}) O_L^B \, d\mathbf{R}}{\int \Psi_B^2(\mathbf{R}) \, d\mathbf{R}} = \frac{\int \Psi_A^2(\mathbf{R}) w_A^B O_L^B \, d\mathbf{R}}{\int \Psi_A^2(\mathbf{R}) w_A^B \, d\mathbf{R}}. \tag{2.42}$$

The correlated sampling estimate of the difference

$$\Delta \approx \frac{1}{M} \sum_i^M O_L^A - \frac{\sum_i^M w_A^B O_L^B}{\sum_i^M w_A^B} \tag{2.43}$$

approaches zero as $\Psi_B \to \Psi_A$, as does the σ_Δ.

The added advantage of correlated sampling is that the same set of configurations can be used for multiple sets of parameters, reducing the cost of generating configurations. For large differences in parameter values, that is, when the weights w_A^B deviate significantly from 1, the correlated sampling estimate is poor.

2.6.2 VARIANCE MINIMIZATION

The zero-variance property of the energy is one important reason why the variance has been preferred as a cost function over the local energy. The known lower bound of zero for the variance of the local energy is attained for an exact trial wave function for any

sampling. The gradient of the variance with respect to wave function parameters is also exactly zero for an exact wave function. More importantly though, variance minimization methods have traditionally been found to be much more numerically stable than their energy-based counterparts [73, 74, 75].

The correlated sampling approach is used to minimize the variance of the local energy [73]. A set of configurations is generated using the wave function with the initial set of parameters $\boldsymbol{A}$. The reweighted variance for the wave function with parameters $\boldsymbol{B}$ is then

$$(\sigma_E^{\boldsymbol{B}})^2 = \frac{\int \Psi_{\boldsymbol{A}}^2(\mathbf{R}) w_{\boldsymbol{A}}^{\boldsymbol{B}} \left[E_{\mathrm{L}}^{\boldsymbol{B}}(\mathbf{R}) - E_{\mathrm{VMC}}^{\boldsymbol{B}}(\mathbf{R}) \right]^2 d\mathbf{R}}{\int \Psi_{\boldsymbol{A}}^2(\mathbf{R}) w_{\boldsymbol{A}}^{\boldsymbol{B}} \, d\mathbf{R}}, \tag{2.44}$$

where $w_{\boldsymbol{A}}^{\boldsymbol{B}} = \Psi_{\boldsymbol{B}}^2 / \Psi_{\boldsymbol{A}}^2$ and

$$E_{\mathrm{VMC}}^{\boldsymbol{B}} = \frac{\int \Psi_{\boldsymbol{A}}^2(\mathbf{R}) w_{\boldsymbol{A}}^{\boldsymbol{B}} E_{\mathrm{L}}^{\boldsymbol{B}}(\mathbf{R}) \, d\mathbf{R}}{\int \Psi_{\boldsymbol{A}}^2(\mathbf{R}) w_{\boldsymbol{A}}^{\boldsymbol{B}} \, d\mathbf{R}}. \tag{2.45}$$

Standard minimization techniques are used to obtain the parameter values that minimize $(\sigma_E^{\boldsymbol{B}})^2$. Formally, this is similar to least-squares fitting of the deviations of the energy estimates from the known lower bound of zero. The existence of a known lower bound contributes to the success of variance minimization.

Since the reweighted variance attains its minimum of zero for an eigenstate of the Hamiltonian regardless of the choice of positive weights w, the weights may be set to unity and the unreweighted variance can be minimized. This is much more robust than reweighted variance minimization which suffers from a numerical instability arising from a few configurations (often only one) acquiring large weights and consequently giving a very small estimate of the variance [74]. The unreweighted variance is not equal to the true variance for a trial wave function even in the limit of perfect sampling, and so the minimization is performed iteratively. A fresh set of configurations generated according to the optimized wave function are used to reoptimize the parameters until they converge to a self-consistent set. Correlated sampling transforms a problem of minimization in the presence of noise into one with a smooth variance landscape for which standard gradient-based minimization methods can be used.

Unreweighted variance minimization can give lower variational energies than reweighted variance minimization and it can be made particularly efficient for optimizing linear parameters in the Jastrow function [75]. However, it is rather poor at optimizing nodal surfaces. The reason for this is that the particle configurations are fixed within an optimization cycle in the correlated sampling scheme, but changing parameters which alter the nodal surface may move the nodal surface through the fixed configurations. The local energy diverges when the nodal surface coincides with a configuration, leading to a poor optimization. Variance-based optimization schemes can be effective in optimizing nodal surfaces if the values of the local energies and/or the weights of configurations near the

nodal surface are limited [75].

The related technique of minimizing the mean absolute deviation MAD of the local energies from the median local energy E_{med} [31],

$$\text{MAD} = \frac{\int \Psi^2(\mathbf{R}) |E_{\text{L}}(\mathbf{R}) - E_{\text{med}}| \, d\mathbf{R}}{\int \Psi^2(\mathbf{R}) \, d\mathbf{R}}, \tag{2.46}$$

is also effective at optimizing the nodal surface as outliers in the energy distribution are not weighted as heavily. The median is also less distorted by outliers than the mean.

2.6.3 ENERGY MINIMIZATION

As the primary quantity of interest in QMC calculation is the ground-state energy for which an upper bound can be estimated in accordance with the variational principle, an efficient energy minimization has been much sought after. The energy and variance minima do not necessarily coincide as the trial wave function generally cannot represent the exact energy eigenstate. Lower total energies can be achieved by minimizing the VMC energy itself. Additionally, wave functions optimized to give the lowest energy have been shown to give better estimates for expectation values other than the energy [76, 77, 78]. A trial wave function that minimizes the energy is also believed to improve the efficiency of DMC [79].

Applying the optimization procedure used for the variance to minimize the energy is unstable for several reasons. In contrast to the variance, an eigenstate of the Hamiltonian does not necessarily give the global minimum of the finite-sampled local energy. This is because the local energy is unbounded from below and hence different samplings of configuration space give different energies. Low estimates of energy can be obtained by selectively sampling configurations with low local energy. Analogously to reweighted variance minimization, reweighted energy minimization experiences instabilities due to large fluctuations in the weights with changes in parameters. However, the weights cannot be modified in the case of the energy. This can be seen by fixing the weights to unity. Unreweighted energy minimization is then equivalent to minimizing the local kinetic energy as the local potential energy is independent of parameters. Altering the weights leads to instabilities as the positions of the energy minima are altered.

Another approach to minimizing the energy would be to write the wave function as a linear sum of basis functions ψ_i and minimize the energy with respect to the coefficients b_i. In matrix notation, this amounts to finding the solution of $\underline{H}\mathbf{b} = E\underline{S}\mathbf{b}$ using diagonalization, where $H_{ij} = \langle \psi_i | \hat{H} | \psi_j \rangle$ and $S_{ij} = \langle \psi_i | \psi_j \rangle$. This method is ubiquitous in quantum chemistry. However the noise in the VMC estimates of $\underline{H}$ and $\underline{S}$ is generally large so that very many configurations must be used. Recasting the standard diagonalization procedure as a least-squares fitting problem [80] similar to the procedure used in variance minimization greatly reduces the number of configurations required.

Assume that the $P+1$ basis functions for P parameters are linearly independent and span an invariant subspace of $\hat{H}$ such that the action of $\hat{H}$ on any ψ_i is described by a linear combination of all ψ_i. One would have then simply have to solve

$$\hat{H}\psi_i(\mathbf{R}) = \sum_{p=0}^{P} \mathcal{E}_{pi}\psi_i(\mathbf{R}) \tag{2.47}$$

by diagonalizing $\mathcal{E}$. For systems of interest, the basis functions do not span an invariant subspace and the solution depends on the choice of configurations. Thus, M configurations, where $M \gg P$, are used to solve the overdetermined problem approximately using a least-squares fit. This method works well in optimizing linear parameters.

To use such a method in VMC to optimize non-linear parameters, the wave function is expanded as

$$\Psi = \Psi_{\boldsymbol{a}_0} + \sum_{p=1}^{P} \delta a_p \left(\frac{\partial \Psi}{\partial a_p}\right)_{\boldsymbol{a}_0} + \mathcal{O}(\delta a^2) \approx \sum_{p=0}^{P} b_p \psi_p, \tag{2.48}$$

where $b_p = \delta a_p$ and $\psi_p = (\partial \Psi/\partial a_p)_{\boldsymbol{a}_0}$. Keeping only the first-order terms in the expansion is often insufficient and the resulting algorithm is unstable. Umrigar and coworkers [81, 82] have developed an energy minimization scheme that is quite robust. In this scheme, the effects of the first-order approximation are reduced by semi-orthogonalization of the basis functions. This method is also extremely effective in optimizing parameters in Ψ that change the nodal surface such as the backflow parameters and CSF coefficients.

While energy minimization is preferred as it gives lower energies than variance minimization, these methods are often used together. Energy minimization is unable to effectively optimize cut-off lengths, which have shallow minima in the energy landscape. The method also converges more quickly when using a good trial wave function, such as one initially optimized with variance minimization.

Chapter 3

STUDIES OF THE FIRST-ROW ATOMS

3.1 INTRODUCTION

FIRST-ROW ATOMS are a natural set of systems to use in learning how to achieve chemical accuracy, which is reached when an error of less than 1 kcal/mol $\simeq$ 1.6 mHa per atom $\simeq$ 43 meV per atom is achieved. Accurate benchmark data are available for light atoms, as are results from many different electronic structure techniques. The cost of all-electron QMC calculations scales with the atomic number [58] Z roughly as $Z^{5.5}$, so that pseudopotentials must be used for heavy atoms, but it is perfectly possible to perform highly-accurate all-electron calculations for atoms up to at least the ten-electron neon atom. We apply VMC and DMC to calculate the ground-state energies and other properties of the atoms Li–Ne and their singly-positively-charged ions. We recover over 98% of the correlation energy for all the atoms and ions studied at the VMC level and over 99% at the DMC level. Chemically-accurate values of the first ionization potentials are obtained. Total energies, scalar relativistic corrections to the energies, mass-polarization terms, and one- and two-electron expectation values are evaluated. We also performed fits to electron and intracule densities.

3.2 TRIAL WAVE FUNCTIONS

Our trial wave functions consist of a multi-determinant expansion which describes static correlation, a Jastrow factor to capture dynamic correlation, and backflow transformations to allow further variations in the nodal surface. These all-electron multi-determinant-Jastrow-backflow wave functions take the form

$$\Psi(\mathbf{R}) = e^{J(\mathbf{R};\mathbf{a})} \sum_{j=1}^{N_{\text{CSF}}} c_j \sum_{k=1}^{N_{\text{det}}^j} d_{k,j} D_{k,j}^{\uparrow}(\mathbf{x}_1,\ldots,\mathbf{x}_{N_\uparrow}) D_{k,j}^{\downarrow}(\mathbf{x}_{N_\uparrow+1},\ldots,\mathbf{x}_N), \qquad (3.1)$$

where $\mathbf{R}$ is the vector of electron positions, $J(\mathbf{R}; \mathbf{a})$ is the Jastrow factor, and $D_{k,j}^{\sigma}(\mathbf{X})$ are the Slater determinants whose orbitals are evaluated at the backflow-transformed coordinates $\mathbf{x}_i = \mathbf{r}_i + \boldsymbol{\xi}_i(\mathbf{R}; \mathbf{b})$. N_{CSF} denotes the total number of CSFs and N_{det}^j is the number of determinants in the jth CSF. The vector $\mathbf{a}$ denotes the parameters in the Jastrow factor, $\mathbf{b}$ those in the backflow transformation, and $\mathbf{c} = (c_1, \ldots, c_{N_{\text{CSF}}})$ the CSF coefficients. As described in Sec. 2.4.5, the coefficients of the determinants $\mathbf{d} = (d_{1,1}, \ldots, d_{N_{\text{det}}, N_{\text{CSF}}})$ are fixed to maintain the proper symmetry of the CSFs.

The Slater determinants and CSFs were generated using the atomic MCHF package ATSP2K [83]. We allowed single and double excitations from the HF ground state up to orbital configurations with principal quantum number $n \leq 7$ and orbital angular momentum quantum number $l \leq 4$. Terms representing excitations from the $1s^2$ core were used for Li, Li$^+$ and Be$^+$ to ensure that double-excitations were included. The CSFs with the largest weights were included in Ψ. Core excitations significantly lowered the MCHF energy of the Be atom, but they did not improve the VMC energy and were therefore not included in the QMC calculations. Excitations from the core become less important for larger Z, and we did not include them for systems with more than three electrons. The high-energy excited-state configurations in the MCHF expansion mostly describe electron-electron cusps, which are captured by the Jastrow factor in QMC calculations. The high-energy MCHF excitations are therefore expected to be much less important in the QMC calculations than in the MCHF ones. Indeed, including very-high-energy excitations serves only to hinder the optimization procedure described in Sec. 3.3 and worsen Ψ. We tested wave functions containing 1, 20 and 50 CSFs for all the atoms, and finally used 50 CSFs for all atoms and ions except O, O$^+$, F and F$^+$, for which we used 100 CSFs. The number of determinants ranged from 171 (Li$^+$) to 4613 (F$^+$).

We used a modified form of the polynomial Jastrow factor proposed by Drummond *et al.* [84] consisting of electron-electron, electron-nucleus and electron-electron-nucleus terms. Each term was written an expansion in powers of $r/(r^\beta + \alpha)$, where r is the interparticle separation and α and β are optimizable parameters, with β constrained to be greater than unity. This basis gave a statistically-significant decrease in the energy for all the atoms compared to an expansion in powers of r when a SJ wave function was used. The improvement was less stark for a SJB wave function. The optimal values of α were found to lie within the range 0.5–17.1 and those of β within the range 1.05–4.67 for the atoms and ions studied. This modification removes the need for cut-offs at large interparticle separations, as the basis functions decay to zero at large r. Based on our tests, we chose expansion orders[1] of 9 for the electron-electron and electron-nucleus parts of the Jastrow factor and an expansion order of 5 for both the electron-electron and electron-nucleus terms in the electron-electron-nucleus Jastrow factor, which gave a total of 118 optimizable parameters.

---

[1]The expansion order is the number of terms in the expansion, i.e., $\sum_{i=0}^{n} a_i r^i$ is of order $n+1$.

The backflow transformation of López Ríos *et al.* [69] was used, with electron-electron and electron-nucleus functions of expansion order 9 and an electron-electron-nucleus function of expansion order 4, resulting in a further 142 optimizable parameters.

In both the Jastrow and backflow functions, the parameter values for anti-parallel spin pairs were allowed to differ from the parameter values for up-spin electron pairs, which were constrained to be equal to the parameter values for down-spin electron pairs. This significantly reduced the number of variable parameters without any noticeable loss in wave function quality. The parallel and anti-parallel-spin cusp conditions were imposed in the Jastrow factor.

3.3 OPTIMIZATION

Various stochastic methods have been developed for optimizing many-body wave functions in QMC calculations as described in Sec. 2.6. We found MAD minimization to be superior to energy minimization for optimizing the cut-off functions, and superior to variance minimization methods for optimizing parameters which alter the nodal surface. We have therefore used MAD minimization in the early stages of the optimizations, but the final optimizations are performed with energy minimization.

We tested optimization of the single-particle orbitals for N, O and F, but found this to have a negligible effect, in agreement with previous atomic studies [85, 86].

We tested several optimization schemes that could potentially reduce the computational effort of wave function optimization. The Jastrow factor and backflow transformation were optimized for a single determinant and then applied to the multi-determinant expansion of a B wave function containing 50 CSFs. Optimizing the CSF coefficients while holding the Jastrow factor and backflow parameters fixed improved the wave function but the final results remained unsatisfactory. This may be expected as the Jastrow and backflow functions capture dynamical correlation overlapping with that described by the addition of CSFs. The CSF coefficients in the B wave function were optimized for one final cycle, as energy minimization of linear coefficients is in general very robust. No improvement was observed, confirming that energy minimization is able to optimize the linear and non-linear parameters simultaneously.

As the optimization process is currently the most costly step in human time and consumes a substantial fraction of the computer time, it is desirable to establish an optimization strategy which is reliable for all of the atoms and ions and may be useful in other systems. Of the several optimization strategies tried, the following consistently gave the best results and was used for all of the final results reported here:

1. Set the CSF coefficients to their MCHF values and the Jastrow parameters **a** to zero. Note that the Jastrow factor is non-zero as the term enforcing the cusp condition is still present.

2. Generate a set of VMC configurations[2] and optimize the Jastrow parameters **a** including the Jastrow basis function parameters α and β, and the CSF coefficients **c** using MAD minimization. We refer to this step as an optimization cycle.

3. Run two more optimization cycles using the parameters obtained in the previous cycle as initial parameters.

4. Optimize the wave function parameters **a** and **c** using VMC energy minimization until converged (usually about 5–8 cycles). The Jastrow basis function parameters α and β are not reoptimized at this stage.

5. Introduce backflow functions with the parameters **b** initially set to zero, and optimize all wave function parameters (**a**, **b**, **c**), including α and β, and the backflow cut-off parameters, using MAD minimization until converged (usually about 3 cycles).

6. Use VMC energy minimization to optimize wave function parameters (**a**, **b**, **c**) until converged (usually about 5–8 cycles). The Jastrow basis function parameters and backflow cut-off parameters are not re-optimized.

We use the fraction of the correlation energy retrieved in a VMC calculation with a given trial wave function Ψ,

$$f_{\mathrm{CE}}[\Psi] = \frac{E_{\mathrm{HF}} - E_{\mathrm{VMC}}[\Psi]}{E_{\mathrm{HF}} - E_{\mathrm{exact}}}, \tag{3.2}$$

as a measure of the quality of Ψ. The improvements in the VMC energies of the atoms at different levels of optimization are shown in Fig. 3.1. The figure clearly shows that VMC energy minimization recovers a significantly larger proportion of the correlation energy than MAD minimization. While this strategy has not been tested for any other systems, we expect it to work well in many cases.

3.4 Results and discussion

3.4.1 Atomic and ionic energies

The VMC optimizations were performed using 5×10^4 statistically-independent particle configurations. One measure of wave function quality is its variance, and Table 3.1 reports the variances of optimized Slater-Jastrow and Slater-Jastrow-backflow wave functions for both single-determinant and multi-determinant Slater forms. The variance is reduced by approximately a factor of 2 or more when a multi-determinant expansion is introduced to

---

[2]The first set of VMC configurations are drawn from the MCHF wave function.
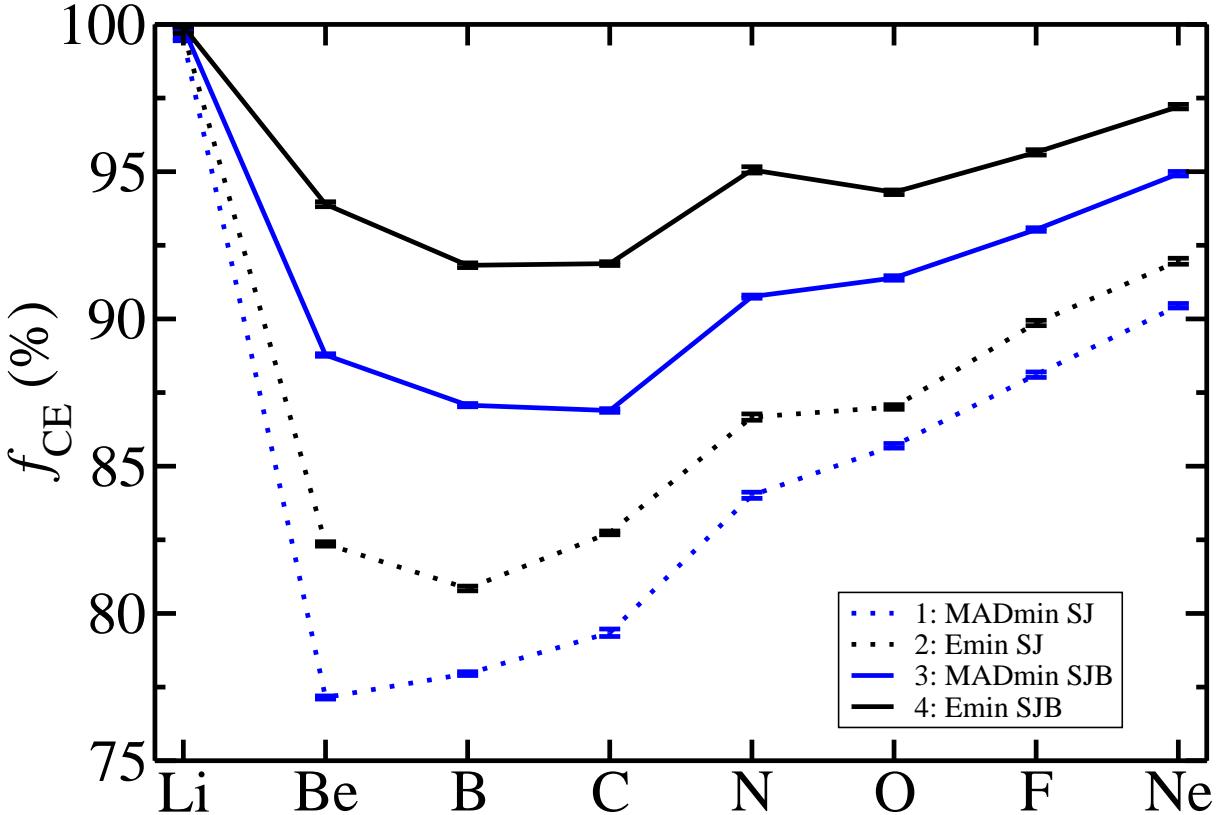
FIGURE 3.1: Percentages of the correlation energy (f_{CE}) retrieved for single-determinant Slater-Jastrow (SJ) and Slater-Jastrow-backflow (SJB) wave functions using mean absolute deviation minimization (MADmin) and VMC energy minimization (Emin).

a Slater-Jastrow-backflow wave function. The improvement is particularly large for Be, B and C, where the variance drops by factors of approximately 10, 8 and 4, respectively. This reflects the strong 2s–2p near-degeneracy effects exhibited in these systems and determinants beyond the HF ground-state configuration must be included to capture the static correlation. With a few exceptions, the variance decreases as the variational energy decreases for a given system. The variance of a wave function after MAD minimization also approaches that after energy minimization as the variational energy decreases.

TABLE 3.1: VMC variances for single-determinant (SD) and multi-determinant (MD) Slater-Jastrow (SJ) and Slater-Jastrow-backflow (SJB) wave functions. All variances are in atomic units.

	SD-SJ	SD-SJB	MD-SJ	MD-SJB
Li	0.00274(3)	0.00130(1)	0.00193(5)	0.00067(2)
Be	0.0443(2)	0.0524(6)	0.01066(6)	0.00526(4)
B	0.0915(2)	0.1434(8)	0.0326(2)	0.01867(8)
C	0.1941(8)	0.1784(7)	0.0819(5)	0.0473(5)
N	0.340(1)	0.263(1)	0.2198(5)	0.1126(5)
O	0.548(1)	0.4763(9)	0.442(1)	0.353(1)
F	0.846(3)	0.619(2)	0.644(3)	0.493(1)
Ne	1.233(3)	0.797(7)	0.623(7)	0.361(2)

The DMC calculations were performed with a target population of 2048 DMC walkers and a minimum of 10^5 steps and a time step corresponding to the smaller of the two used in Ref. [85], ranging from 0.00375 a.u. for Li to 0.00070 a.u. for Ne. These calculations [85] already showed that the errors from these time steps is negligible, and the corresponding errors in the current calculations should be even smaller as the trial wave functions are superior [43].

Table 3.2 gives the VMC and DMC energies and percentages of the correlation energy retrieved for each of the atoms and ions studied. The reference non-relativistic energies, assuming a clamped point nucleus, are taken from Refs. [87, 88]. Percentages of the correlation energy retrieved at the VMC and DMC levels for the neutral atoms in the present work and those of Ref. [85] are compared in Fig. 3.2 and data for singly-charged ions are shown in Fig. 3.3. In both figures, the percentage of the correlation energy required to achieve chemical accuracy is indicated.
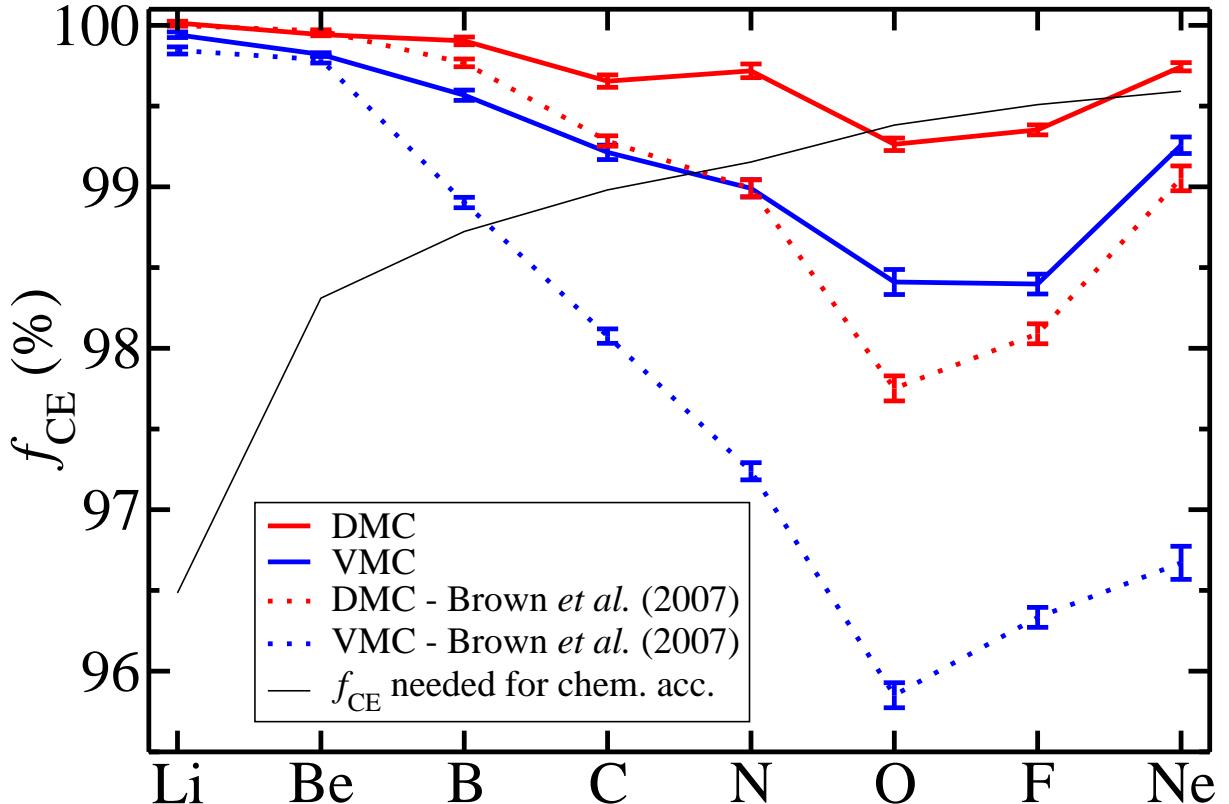
FIGURE 3.2: Percentages of the correlation energy (f_{CE}) retrieved for each atom within VMC and DMC. Chemical accuracy is achieved for Li–N and Ne at the DMC level.

As the critical approximation made in DMC is the fixed-node approximation, the quality of the DMC wave function is expected to decrease with an increasing number of electrons due to the additional complexity of the nodal surface. Using the percentage of the correlation energy retrieved as a measure of the quality of the wave function, we observe this trend for both the atoms and ions with some exceptions. According to Hund's rules, both N and O^+ have a configuration with a half-filled 2p shell that resembles a

TABLE 3.2: VMC and DMC energies of the first-row atoms and ions. Also included are Hartree-Fock energies E_{HF} calculated using ATSP2K [83], reference energies E_{ref} [87, 88], correlation energies $E_{\mathrm{HF}} - E_{\mathrm{ref}}$, and the percentage of the correlation energy recovered at the VMC level, f_{CE} (VMC), and DMC level, f_{CE} (DMC). All energies are in atomic units.

	Li (^{2}S)	Be (^{1}S)	B (^{2}P)	C (^{3}P)	N (^{4}S)	O (^{3}P)	F (^{2}P)	Ne (^{1}S)
VMC	−7.478034(8)	−14.66719(1)	−24.65337(4)	−37.84377(7)	−54.5873(1)	−75.0632(2)	−99.7287(2)	−128.9347(2)
DMC	−7.478067(5)	−14.667306(7)	−24.65379(3)	−37.84446(6)	−54.58867(8)	−75.0654(1)	−99.7318(1)	−128.9366(1)
E_{HF}	−7.432727	−14.573023	−24.529061	−37.688619	−54.400934	−74.809398	−99.409349	−128.547098
E_{ref}	−7.47806032	−14.66736	−24.65391	−37.8450	−54.5892	−75.0673	−99.7339	−128.9376
$E_{\mathrm{HF}} - E_{\mathrm{ref}}$	0.0453333	0.094337	0.124849	0.156381	0.188266	0.257902	0.324551	0.390502
f_{CE} (VMC)	99.94(2)%	99.82(1)%	99.57(3)%	99.21(4)%	98.99(5)%	98.41(8)%	98.40(6)%	99.26(5)%
f_{CE} (DMC)	100.01(1)%	99.943(7)%	99.90(2)%	99.65(4)%	99.72(4)%	99.26(4)%	99.35(3)%	99.74(3)%

	Li$^+$ (^{1}S)	Be$^+$ (^{2}S)	B$^+$ (^{1}S)	C$^+$ (^{2}P)	N$^+$ (^{3}P)	O$^+$ (^{4}S)	F$^+$ (^{3}P)	Ne$^+$ (^{2}P)
VMC	−7.279844(9)	−14.324721(9)	−24.34836(4)	−37.43034(6)	−54.0530(1)	−74.5655(1)	−99.0880(2)	−128.1377(2)
DMC	−7.2799914(3)	−14.324761(3)	−24.34887(2)	−37.43073(4)	−54.05383(7)	−74.56662(7)	−99.0911(2)	−128.1412(2)
E_{HF}	−7.236415	−14.277395	−24.237575	−37.292224	−53.888005	−74.372606	−98.831720	−127.817814
E_{ref}	−7.27991	−14.32476	−24.34892	−37.43103	−54.0546	−74.5668	−99.0928	−128.1431
$E_{\mathrm{HF}} - E_{\mathrm{ref}}$	0.043495	0.047365	0.111345	0.138806	0.166595	0.194194	0.26108	0.325286
f_{CE} (VMC)	99.85(2)%	99.92(2)%	99.50(4)%	99.50(4)%	99.04(6)%	99.33(5)%	98.16(8)%	98.34(6)%
f_{CE} (DMC)	100.009(7)%	100.002(6)%	99.96(2)%	99.78(3)%	99.54(4)%	99.91(4)%	99.35(8)%	99.42(6)%
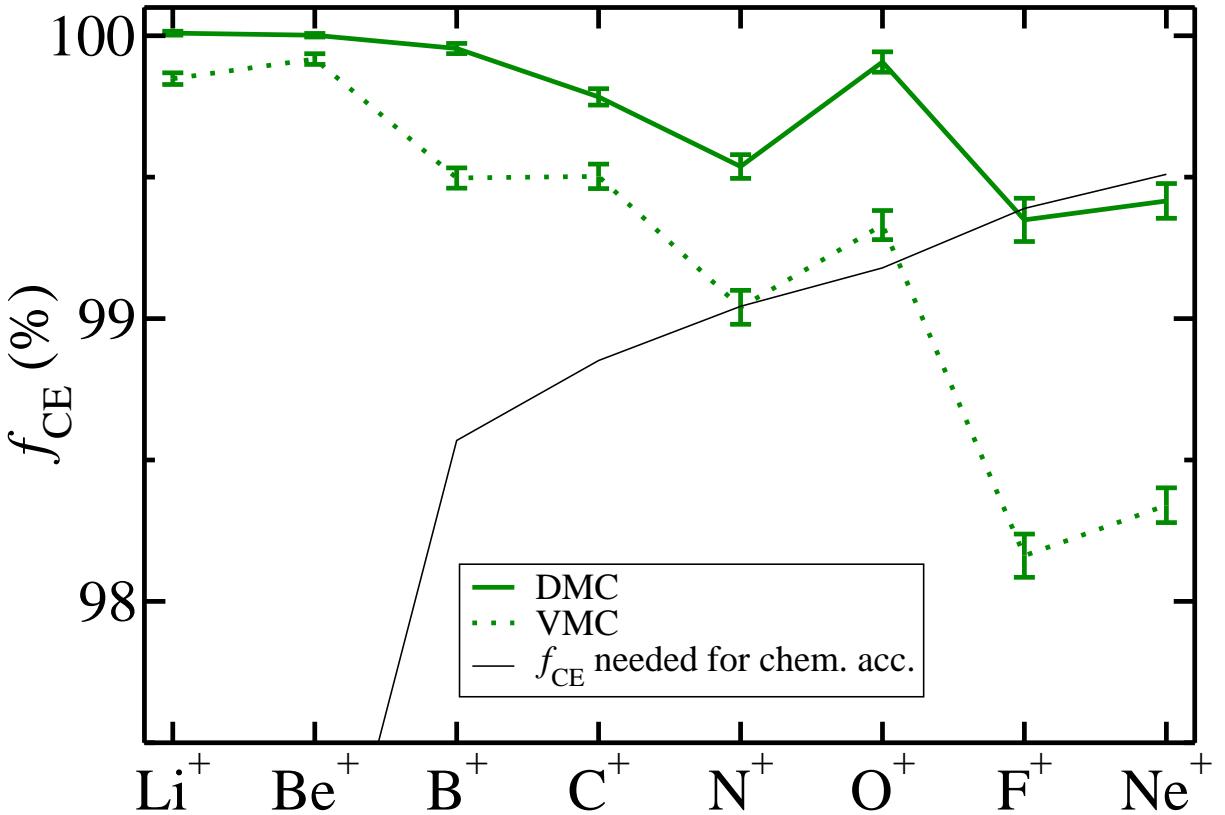
FIGURE 3.3: Percentages of the correlation energy (f_{CE}) retrieved for each ion within VMC and DMC. Chemical accuracy is achieved for Li^+–O^+ at the DMC level. The values for F^+ and Ne^+ are within statistical uncertainty of chemical accuracy.

closed-shell atom. The higher symmetry of such electronic configurations allows a greater number of equivalent determinants to be grouped into a single CSF. For a given number of CSFs, determinants with higher n and l values are included and a larger active space is sampled. This explains the higher quality of the N, O^+ and Ne wave functions. The reverse argument holds for the O and F^+ wave functions, which are of lower quality.

There are several differences between the wave functions used in the present study and those of Ref. [85]. While both calculations relied on the energy minimization scheme of Refs. [81, 82], the implementation used in the current work is more effective and robust[3]. For example, Brown *et al.* [85] were unable to lower the VMC energy of Ne using a multi-determinant expansion, which was easily achieved in the present study. The present optimization strategy is significantly different as we use MAD minimization to first optimize the non-linear parameters at each stage. Brown *et al.* [85] used a Jastrow factor based on an expansion in r, while we have used an expansion in powers of $r/(r^\beta + \alpha)$. We have also employed a larger number of CSFs.

We have obtained more than 99% of the correlation energy at the DMC level for all of the atoms and ions, and at the VMC level for all atoms except O and F and all ions

---

[3]These changes make the matrices involved in VMC energy energy minimization invariant to the scaling of parameters and improve the implementation of the level-shifting [89].

except F^+ and Ne^+. This is a substantially higher accuracy than achieved in the all-electron QMC calculations reported in the literature [85, 90, 54, 91]. For example, the lowest percentage of the correlation energy achieved for a neutral atom in the present study at the VMC level is 98.40(6)% for F, whereas the best previous VMC calculation gave 96.33(6)%, and our lowest percentage in DMC is 99.26(4)% for O compared with the best previous value of 97.83(8)% [85]. We calculated the virial ratios for the atoms and ions in both VMC and DMC, finding them to be within one standard error of the exact value of 2 in each case, with the standard errors lying within the range 0.001–0.02.

We also find that our atomic energies using a single-determinant SJ wave function are significantly better than those obtained within the framework of the transcorrelated method. Umezawa *et al.* [18] and Prasad *et al.* [92] retrieve between 36% (Ne) and 88-89% (Li) of the correlation energy while we retrieved between 80.77(8)% (B) and 99.67(4)% (Li) of the correlation energies for the atoms. The poor quality of their results can be explained in part by the fact that their Jastrow factor comprises only electron-electron terms while ours extends to three-body electron-electron-nucleus terms.

3.4.2 Ionization potentials

Although the total atomic energies can be measured as the sum of the ionization energies, they are not quantities of significant chemical interest. In quantum chemistry one is normally interested in energy differences for which the cancellation of errors between calculations is important. We have therefore calculated the first ionization potentials (IPs) of the atoms Li–Ne as energy differences between the neutral and singly-ionized states. The errors in the calculated IPs from those computed using values from Ref. [88] are shown in Fig. 3.4. Data from the FCIQMC method [6] with an aug-cc-pVQZ basis set for Li, Be and Ne and an aug-cc-pV5Z basis for B–F are shown, together with data from coupled cluster with single and double excitations (CCSD) calculations with a d-aug-cc-pwCV5Z basis and CCSD-F12-HLC data [93]. Each CCSD-F12-HLC energy is the sum of the CCSD energy, an F12 energy which corrects for the finite basis set, and a higher-level correction (HLC) which accounts for the treatment of excitations beyond the doubles in CCSD. It is likely that the CCSD-F12-HLC results [93] are even more accurate than the data of Ref. [88] that we have used as a reference, as they obtain results in closer agreement with experiment when corrections for relativistic effects and the finite nuclear mass are included. However, Klopper *et al.* [93] did not give values for the total energies of the atoms, and therefore we have used the data of Ref. [88] to avoid using different reference data for the total energies and IPs. The differences from using the IP data of Klopper *et al.* [93] are small, as can be seen in Fig. 3.4. Using this data as the reference would not significantly affect the comparisons for Li and Be, but it would slightly worsen the agreement with our results for B, C and Ne and slightly improve it for N, O and F.

The IPs from DMC are within statistical error of chemical accuracy of the reference
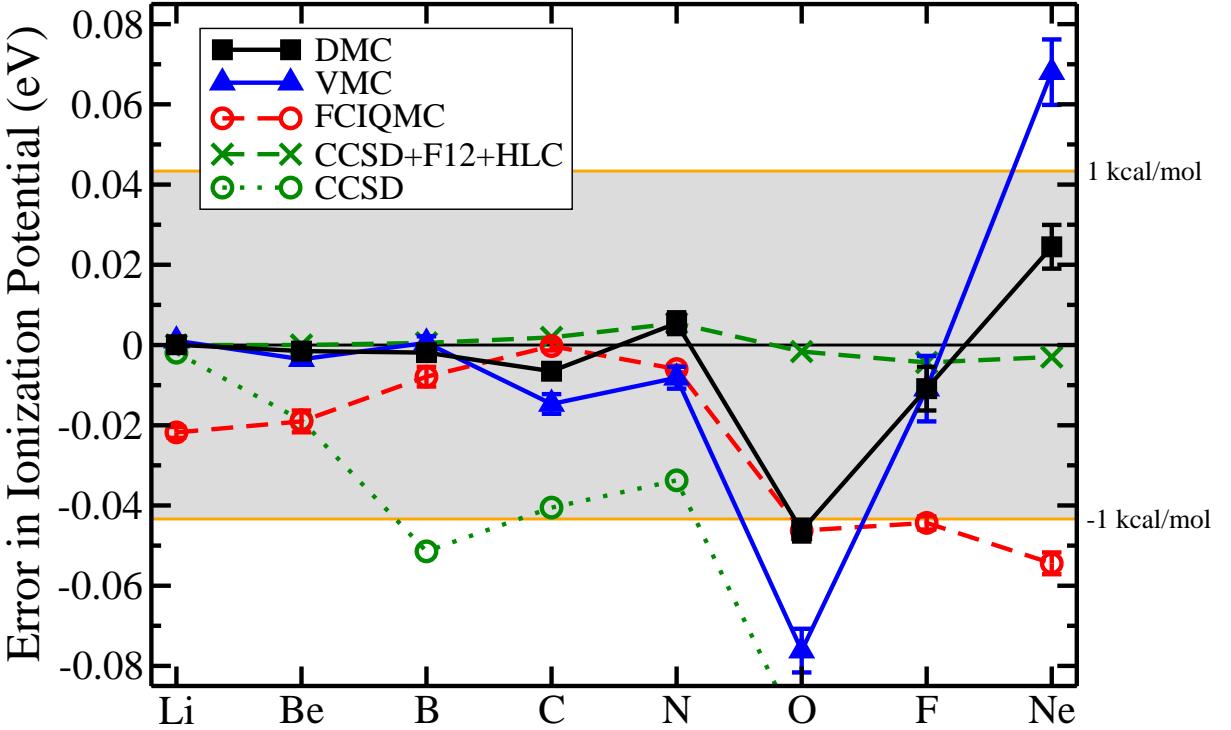
FIGURE 3.4: Errors in the ionization potentials ($\Delta = \text{IP}_{\text{calc}} - \text{IP}_{\text{ref}}$) for the first-row atoms obtained at the VMC and DMC levels compared to those from FCIQMC [6], CCSD [93] and CCSD-F12-HLC [93]. The reference values are taken from Ref. [88]. The shaded region represents chemical accuracy.

data for all atoms. Our errors are smaller than or equal to those in the CCSD results with a d-aug-cc-pwCV5Z basis [93] for all atoms, and smaller than those of the FCIQMC calculations of Booth and Alavi [6] for all atoms except C.

In addition to the data presented in Fig. 3.4, values from DFT calculations using the B3LYP [94], LSDA [95] and PBE-GGA [95] functionals are included in Fig. 3.5. There is significant correlation between the results obtained using different functionals.

The mean deviation, mean absolute deviation and maximum deviation of the IPs from the reference values for these methods and those obtained in DFT using the B3LYP, LSDA and PBE-GGA density functionals are presented in Table 3.3. The mean absolute deviations of the DFT IPs are between 24 and 30 times larger than for our DMC calculations.

The FCIQMC approach [6] is exact up to a basis set convergence error and a small statistical error. They consistently underestimate IPs, perhaps because there are fewer electron-electron and electron-nucleus cusps in an ion than in the corresponding neutral atom. As is clear from Fig. 3.4, we similarly underestimate the IPs in all cases except Ne, but for a different reason, as explained below.

In general, the nodal surface of an ion is easier to describe than that of the corresponding neutral atom. However, for closed-shell atoms such as Ne, the initial restricted HF atomic nodal surface is superior to that of the open-shell ion. The energy of the neutral
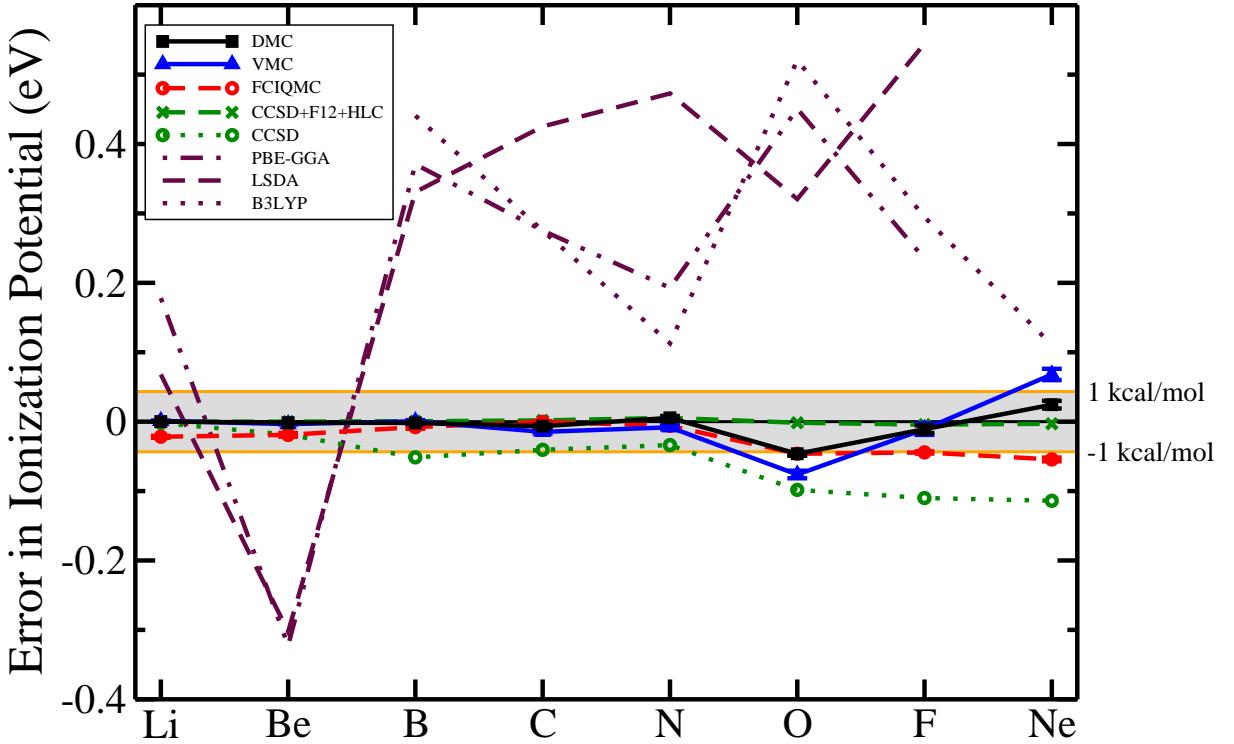
FIGURE 3.5: Errors in the ionization potentials ($\Delta = \mathrm{IP_{calc}} - \mathrm{IP_{ref}}$) for the first-row atoms obtained at the VMC and DMC levels compared to FCIQMC [6], CCSD [93] and CCSD-F12-HLC [93], and those from DFT using the B3LYP [94], LSDA [95] and PBE-GGA [95] functionals. The reference values are taken from Ref. [88]. The shaded region represents chemical accuracy.

TABLE 3.3: Comparison of the mean deviation $\overline{\Delta}$, mean absolute deviation $\overline{|\Delta|}$ and maximum deviation $\Delta_{\max}$ of the ionization potentials obtained from several electronic structure methods. Deviations are from the reference non-relativistic, clamped point nucleus values of Ref. [88]. Averages were taken over Li–Ne, unless otherwise indicated. All values are in electron volts.

	$\overline{\Delta}$	$\overline{|\Delta|}$	$\Delta_{\max}$
VMC	$-0.005(2)$	$0.023(2)$	$0.076(5)$
DMC	$-0.005(1)$	$0.012(1)$	$0.046(2)$
FCIQMC[a]	$-0.0250(7)$	$0.0250(7)$	$0.054(3)$
CCSD[b]	-0.0586	0.0585	0.1140
CCSD-F12-HLC[b]	-0.0001	0.0021	0.0054
B3LYP[c]	0.2925	0.2924	0.5206
LSDA[d]	0.2657	0.3521	0.5447
PBE[d]	0.1971	0.2892	0.4507

[a] Ref. [6].
[b] Ref. [93].
[c] Averages taken over B–Ne values. Ref. [94].
[d] Averages taken over Li–F values. Ref. [95].

atom is more accurate and consequently the IP is overestimated. For reasons given in Sec. 3.4.1, an overestimation of the IP of N and severe underestimation of the IP of O are also observed.

To summarise, the additional complexity that arises for larger atoms manifests itself differently in FCI and DMC. In the former, describing the electron-electron cusps becomes more challenging and requires a larger expansion in determinants, whereas the nodal structure of the larger system is more difficult to describe in DMC.

3.4.3 Electron and electron-pair densities

The electron density[4] is given by

$$\rho(r) = \sum_i \frac{\int \Psi^2(\mathbf{R})\delta(|\mathbf{r}_i| - r)\,d\mathbf{R}}{\int \Psi^2(\mathbf{R})\,d\mathbf{R}}. \tag{3.3}$$

As a function of only one variable, the density is a much simpler quantity than the wave function and is useful in the interpretation of chemical properties. Accurate exchange-correlation functionals for DFT calculations can also be constructed by fitting to electron densities [96, 97].

A fit-based accumulation method was initially tested to fit the density data directly to an analytic linear expansion of basis functions. We tested several different basis functions, including Gaussians, exponentials and Chebyshev polynomials. However, none of these bases could compactly reproduce the short- and intermediate-distance structure of the everywhere-positive density while maintaining the asymptotic large-distance behaviour.

We then simply accumulated the spherically-symmetric electron density in narrow bins and fitted it to an exponential Padé form

$$\rho_{\text{fit}}(r) = \frac{N}{N_{\text{n}}}\left(1 + Ar^{2B}\right)\exp\left(-\frac{k_0 + 2Zr + k_2 r^2 + \ldots + k_{n-1}r^{n-1} + k_n^2 r^n}{1 + l_2^2 r^2 + \ldots + l_{n-2}^2 r^{n-2} + \frac{k_n^2}{2\sqrt{2I}}r^{n-1}}\right), \tag{3.4}$$

where A, k_i and l_i are parameters, N is the number of electrons and N_{n} is a normalization factor. This form satisfies the cusp condition at the origin [98], $\frac{\partial\rho}{\partial r}\big|_{r=0} = -2Z\rho(0)$ and has the correct asymptotic behaviour at large distances [99]: $\lim_{r\to\infty}\rho(r) \sim r^{2B}\exp\left(-2\sqrt{2I}\,r\right)$, where $B = \frac{Z-N+1}{\sqrt{2I}} - 1$. The denominator in the exponential is everywhere positive as all parameters are squared, ensuring the density is well-behaved. The fitted parameter values for the up-spin and down-spin electron densities are given in the supplementary information. We note that a method of obtaining more accurate estimates of densities has been developed [100].

Electron-pair densities are also of interest. They are important in describing atomic

---

[4]The electron density is also referred to as the charge density.

properties such as Hund's rules [101]. There are relationships between experimentally measurable scattering cross sections and electron-pair densities [102, 103]. Quantities such as the Fermi and Coulomb holes which are useful in understanding exchange and correlation can also be written in terms of a pair density [104]. Such quantities can also be used to give better approximations to the DFT functionals. A pair density can be separated into the intracule density which describes the relative motion of a pair of electrons, and the extracule density which describes the motion of the centre of mass of the pair. The moments of the intracule density $\langle r_{ij}^n \rangle$ are also significant and are discussed in Sec. 3.4.4. The spherically-averaged intracule density is defined as

$$h(r) = \frac{1}{2} \sum_{i \neq j} \frac{\int \Psi^2(\mathbf{R}) \delta(|\mathbf{r}_{ij}| - r) \, d\mathbf{R}}{\int \Psi^2(\mathbf{R}) \, d\mathbf{R}}. \tag{3.5}$$

Sarsa *et al.* [105] published a VMC study of the intracule densities of the atoms He–Ne, while more recent atomic QMC studies have explored methods for obtaining more accurate estimates of intracule densities [106].

We fitted the intracule densities separately for parallel spin pairs and anti-parallel spin pairs to

$$h_{\text{fit}}^{\parallel}(r) = \frac{N_{\text{p}}^{\parallel}}{N_{\text{n}}^{\parallel}} \, r^2 \exp\left(-\frac{s_0 - \frac{1}{2}r + s_2 r^2 + \ldots + s_{n-1} r^{n-1} + s_n^2 r^n}{1 + t_2^2 r^2 + \ldots + t_{n-2}^2 r^{n-2} + \frac{s_n^2}{2\sqrt{2I}} r^{n-1}}\right), \tag{3.6}$$

$$h_{\text{fit}}^{\perp}(r) = \frac{N_{\text{p}}^{\perp}}{N_{\text{n}}^{\perp}} \exp\left(-\frac{u_0 - r + u_2 r^2 + \ldots + u_{n-1} r^{n-1} + u_n^2 r^n}{1 + v_2^2 r^2 + \ldots + v_{n-2}^2 r^{n-2} + \frac{u_n^2}{2\sqrt{2I}} r^{n-1}}\right), \tag{3.7}$$

where s_i, t_i, u_i and v_i are parameters, $N_{\text{p}}^{\parallel}$ and $N_{\text{p}}^{\perp}$ are the numbers of parallel and anti-parallel spin pairs, respectively, and $N_{\text{n}}^{\parallel}$ and $N_{\text{n}}^{\perp}$ are normalization factors. Again, the denominator in the exponential is forced to be positive. These forms satisfy the intracule cusp condition for the anti-parallel spin pairs [107] $h'(0) = h(0)$. The higher-order condition [108] $h^{(3)}(0) = \frac{3}{2}h^{(2)}(0)$ must be satisfied for the parallel spin pairs for which $h'(0) = h(0) = 0$. The fitted parameter values for the up-up and down-down parallel spin and up-down anti-parallel spin pair intracule densities are given in supplementary information.

We compare our fit to the electron density obtained using a very accurate Hylleraas wave function [109] for Li^+ in Fig. 3.6. The Hylleraas calculation gave an energy correct to 12 significant figures. Our fit agrees well with the Hylleraas density.

Large statistical fluctuations in the binned data occurs at short inter-particle distances which are not frequently sampled. Similar fluctuations occur at very large inter-particle distances where the density is small and sampling is consequently infrequent. The short-range fluctuations are clear in the fit to the electron density to He given in Fig. 3.7. These fluctuations pose a problem in performing linear-least square fits to the fitting form. Using
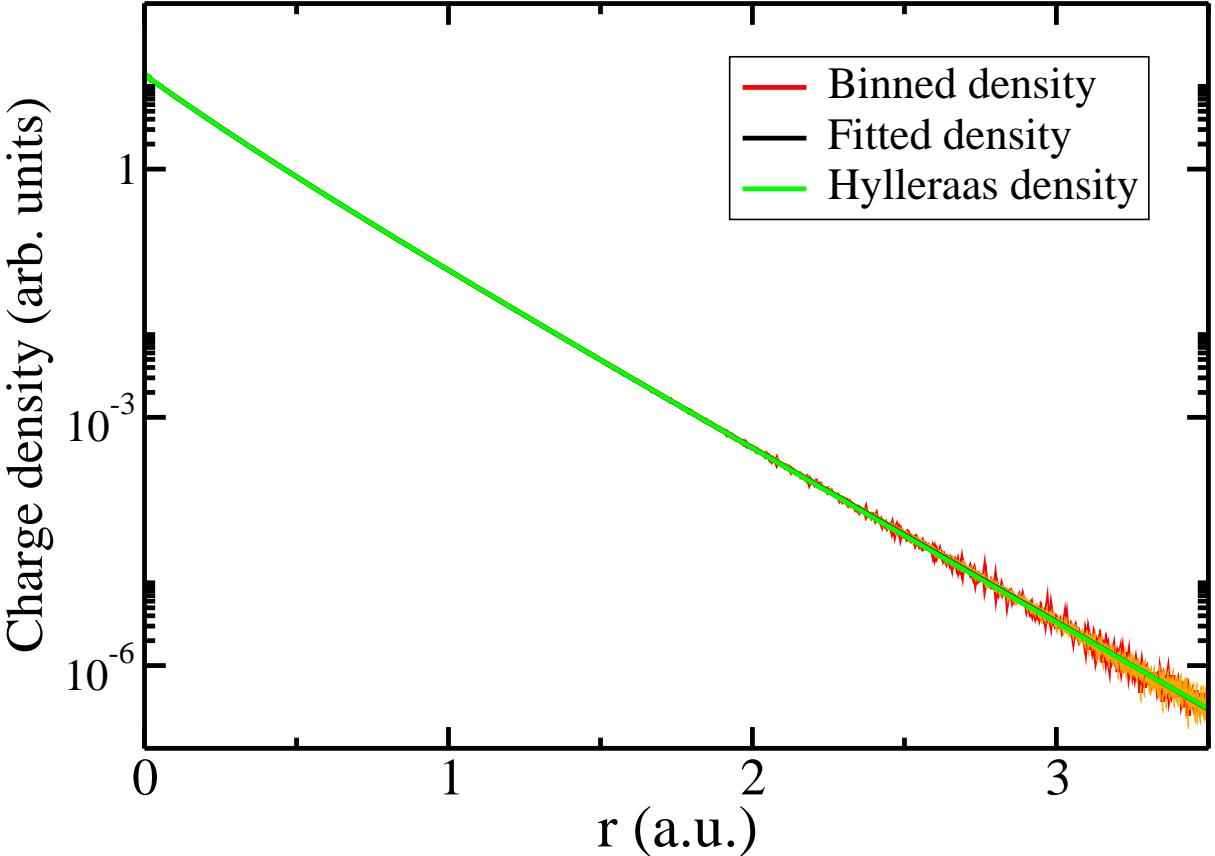
FIGURE 3.6: The total charge density of Li^+ as collected in the binning process and fitted to the form in Eq. 3.4. The error bars in the binned density are shown in yellow. The fit agrees very well with the density of Ref. [109], which was obtained from a Hylleraas calculation.

the zero-variance zero-bias principle of Assaraf and Caffarel [110] to construct improved estimators of quantities such as the electron and intracule densities has been shown to overcome the problem of large variances [100, 106]. Their method does not however give an analytic form of the densities. We believe that Eqs. 3.4, 3.6 and 3.7 are suitable to fitting densities and expect that combining such fitting forms with binned data obtained using zero-variance zero-bias estimators should give highly-accurate densities.

3.4.4 OTHER EXPECTATION VALUES

Scalar relativistic and mass-polarization corrections to the energies of the atoms and ions were computed using first-order perturbation theory within VMC and DMC [111]. For small atoms such as those considered here, first-order perturbation theory gives good approximations to the exact relativistic corrections. The mass-velocity, one-electron Darwin, two-electron Darwin, retardation and mass-polarization terms are given in Table 3.4. We did not calculate spin-dependent relativistic terms such as the spin-orbit and spin-spin terms which require separation of spin and spatial operators. This can in principle be done for a typical QMC wave function; Alexander *et al.* [112] report VMC results for
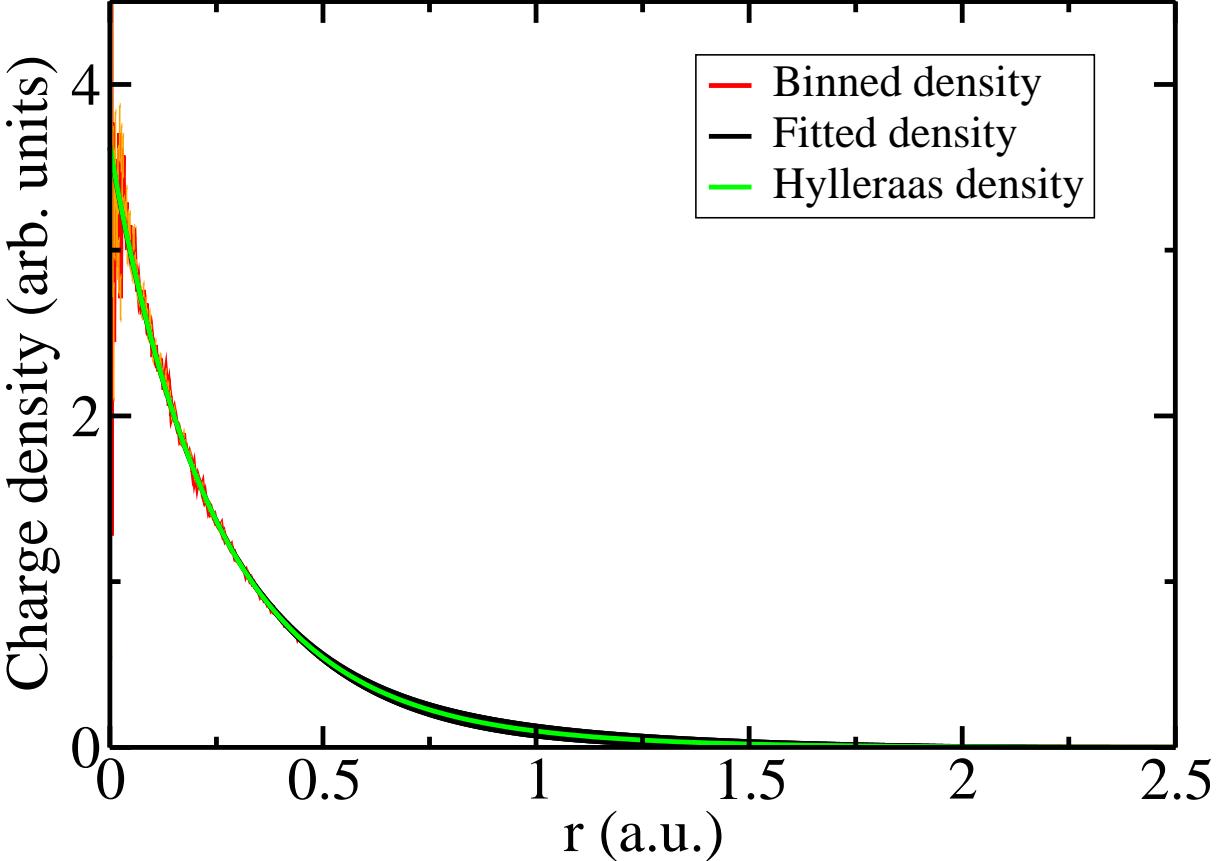
FIGURE 3.7: The total charge density of He as collected in the binning process and fitted to the form of Eq. 3.4. The error bars in the binned density are shown in yellow. The density of Ref. [109], obtained from a Hylleraas calculation, is also plotted.

spin-dependent corrections of electronic states of the He atom. As the VMC and DMC values agree to within one standard error in most cases, we do not give the extrapolated estimates (2×DMC-VMC) in Table 3.4 and instead we quote only the DMC values. Our results for Li, Li$^+$, Be and Be$^+$ are close to those given in Refs. [113], [114] and [115], obtained from Hylleraas calculations, and in Ref. [116], from exponentially-correlated Gaussian calculations.

In Table 3.5 we report some one-electron expectation values for the atoms and ions, while two-electron expectation values are reported in Table 3.6. The moments of the electron and intracule densities have physical significance [117]; for example, the inter-electronic moment $\langle r_{ij}^{-1} \rangle$ represents the average Coulomb repulsion between electrons.

Variational calculations using Hylleraas-type wave functions have given very accurate results for three systems included in our study: the Li atom [118, 119], the Li$^+$ ion [120], and the Be$^+$ ion [121]. Our results for these systems are in good agreement with the Hylleraas data. The data available in the literature for systems with more than three electrons are of much lower accuracy. Our results for $\langle \delta(r_i) \rangle$ and $\langle r_i^{-1} \rangle$ agree well with the Hartree-Fock values of Gálvez and Porras [122] where a comparison can be made, while our values of $\langle r_i^{-2} \rangle$ are larger than theirs. Cohen *et al.* [123] have reported values

TABLE 3.4: Scalar relativistic terms: mass-velocity (MV), electron-nucleus Darwin (D1), two-electron Darwin (D2), spin-spin contact interaction (SSC), retardation (Ret), and mass-polarization (MP) energies calculated at the DMC level. Values from the literature are given for Li, Li$^+$, Be and Be$^+$. All values are in atomic units.

	MV	D1	D2+SSC	Ret	MP
Li	$-0.00417(1)$	$0.00346(1)$	$0.0000914(3)$	$-0.0000232(1)$	$0.0000239(1)$
Be	$-0.01439(2)$	$0.01181(2)$	$0.0002690(5)$	$-0.0000478(1)$	$0.00002815(9)$
B	$-0.0368(1)$	$0.0300(1)$	$0.000598(2)$	$-0.0000585(7)$	$0.0000137(3)$
C	$-0.0790(2)$	$0.0639(2)$	$0.001115(8)$	$-0.000017(2)$	$-0.0000178(5)$
N	$-0.1504(6)$	$0.1207(7)$	$0.00185(4)$	$0.000147(5)$	$-0.000069(2)$
O	$-0.2610(7)$	$0.2086(7)$	$0.00299(3)$	$0.000415(4)$	$-0.0001278(8)$
F	$-0.424(1)$	$0.337(1)$	$0.00451(5)$	$0.000935(7)$	$-0.000195(1)$
Ne	$-0.655(2)$	$0.518(2)$	$0.00646(8)$	$0.00186(1)$	$-0.000303(1)$
Li$^+$	$-0.00411(1)$	$0.00341(1)$	$0.0000895(2)$	$-0.00002291(9)$	$0.00002298(8)$
Be$^+$	$-0.01426(2)$	$0.01171(2)$	$0.0002649(3)$	$-0.0000485(1)$	$0.00002756(7)$
B$^+$	$-0.0377(3)$	$0.0307(3)$	$0.000597(2)$	$-0.0000802(6)$	$0.0000305(3)$
C$^+$	$-0.0796(3)$	$0.0643(3)$	$0.001115(5)$	$-0.000056(1)$	$-0.0000003(4)$
N$^+$	$-0.1506(5)$	$0.1208(5)$	$0.00190(1)$	$0.000071(2)$	$-0.0000494(6)$
O$^+$	$-0.263(2)$	$0.210(2)$	$0.00299(2)$	$0.000388(4)$	$-0.0001143(8)$
F$^+$	$-0.425(2)$	$0.337(2)$	$0.0044(1)$	$0.00091(5)$	$-0.000183(3)$
Ne$^+$	$-0.658(2)$	$0.519(2)$	$0.00636(8)$	$0.001719(9)$	$-0.000287(1)$
Li	-0.00418308[a]	0.00347364[a]	0.0000911359[a]	-0.0000232018[a]	0.0000236819[b]
Li$^+$	-0.00413427[c]	0.00343889[c]	0.000089292[c]	-0.000022791[c]	0.00002259816[c]
Be	-0.01441539[d]	0.011834014[d]	0.0002685577[d]	-0.0000474909[d]	0.0000278121[d]
Be$^+$	-0.0142882124[d]	0.011745724[d]	0.0002644146[d]	-0.00004845370[d]	0.0000273704[d]

[a] Ref. [113].
[b] Ref. [114].
[c] Ref. [115].
[d] Ref. [116].

of $\langle r_i^2 \rangle$ for some atoms, including Be, B and C, calculated within unrestricted HF theory and correlated theories such as FCI. Electron correlation is expected to reduce the size of atoms as measured by $\langle r_i^2 \rangle$. In HF, where correlation is absent, anti-parallel spin electrons can be located close to each other and act to screen electrons from the nucleus. In a correlated system, the electrons position themselves to minimize electron-electron repulsion and maximize nuclear attraction, pulling them closer to the nucleus. The effect of correlation on the electron density decreases with increasing Z as the nuclear attraction dominates the electron-electron repulsion.

Our value of $\langle r_i^2 \rangle$ for Be reported in Table 3.5 is slightly larger than the FCI values of 16.27, while our values for B and C are slightly smaller than the FCI values of 15.54 (B) and 13.84 (C) [123]. In almost all cases the values of the one-electron expectation values (summed over the electrons) are larger for the neutral atoms than for the corresponding ions, as one would expect. However, the expectation values of $\langle \delta(r_i) \rangle$ and $\langle r_i^{-2} \rangle$, which are the most strongly weighted towards the region close to the nucleus, are larger for the

TABLE 3.5: One-electron expectation values: electron moments $\langle r_i^n \rangle$ for $-2 \leq n \leq 3$ and electron density at the coalescence point $\langle \delta(r_i) \rangle$, summed over all electrons i. All values are in atomic units.

	$\langle \delta(r_i) \rangle$	$\langle r_i^{-2} \rangle$	$\langle r_i^{-1} \rangle$	$\langle r_i \rangle$	$\langle r_i^2 \rangle$	$\langle r_i^3 \rangle$
Li	13.79(5)	30.25(4)	5.7193(4)	4.9842(3)	18.300(2)	92.10(1)
Be	35.30(6)	57.59(3)	8.4275(2)	5.9794(1)	16.2986(4)	57.078(2)
B	71.7(3)	93.51(9)	11.3993(7)	6.7446(3)	15.5322(9)	46.011(4)
C	127.2(4)	138.8(1)	14.7065(8)	7.1230(3)	13.7401(7)	33.940(3)
N	206(1)	193.0(1)	18.3491(9)	7.3612(2)	12.1750(5)	25.740(2)
O	312(1)	257.1(2)	22.271(1)	7.6364(2)	11.3283(5)	21.756(2)
F	448(2)	330.8(2)	26.537(1)	7.8166(2)	10.4132(4)	18.003(1)
Ne	619(2)	414.3(3)	31.134(1)	7.9298(2)	9.5220(4)	14.8372(9)
Li$^+$	13.60(4)	29.81(4)	5.3770(4)	1.14539(7)	0.89252(7)	0.8830(1)
Be$^+$	35.01(5)	56.97(2)	7.9760(2)	3.10220(6)	6.5122(2)	18.7046(8)
B$^+$	73.5(8)	93.94(9)	10.9332(7)	4.1791(2)	7.6318(5)	17.736(2)
C$^+$	128.1(6)	138.9(1)	14.1589(8)	4.9235(2)	7.9284(4)	16.072(1)
N$^+$	206.3(9)	193.2(2)	17.727(1)	5.4338(2)	7.7161(4)	13.742(1)
O$^+$	314(3)	257.1(2)	21.618(1)	5.8097(2)	7.3423(3)	11.6156(8)
F$^+$	447(3)	331.3(2)	25.811(1)	6.1624(2)	7.1305(3)	10.3950(7)
Ne$^+$	621(2)	414.9(3)	30.323(1)	6.4236(2)	6.7976(3)	9.0808(5)

ion than for the neutral atom for B/B$^+$, and very similar for C/C$^+$–Ne/Ne$^+$. The larger error bars and lower quality of the wave functions make it more difficult to draw firm conclusions for C/C$^+$–Ne/Ne$^+$.

3.5 CONCLUSIONS

We have calculated energies for the first-row atoms with significantly more accuracy than previous DMC studies. Our DMC energies for the atoms heavier than Li and ions heavier than Be$^+$ are the lowest so far reported from a variational method. Our DMC IPs are also superior to very recent FCIQMC results [6]. Our IPs are, however, substantially less accurate than the CCSD-F12-HLC data of Klopper *et al.* [93]. Our DMC IPs are considerably better than the CCSD values, but the addition of the F12 and HLC corrections leads to errors which are roughly an order of magnitude smaller than those in our DMC calculations. The DMC calculations have the feature that the results are obtained from a single calculation while the cost of calculating the F12 and HLC corrections in the CCSD scheme will increase very rapidly with the number of electrons.

It would be extremely useful if *post hoc* corrections could be developed for QMC methods. One method which has shown some success in VMC is to plot the total energy versus the variance of the local energy using a set of trial wave functions of different qualities [124]. Such a plot normally shows an approximately linear variation so that an extrapolation to zero variance can be performed. The linear variation can be derived by

TABLE 3.6: Two-electron expectation values: inter-electronic moments $\langle r_{ij}^n \rangle$ for $-2 \leq n \leq 3$, the electron-pair density at the coalescence point $\langle \delta(r_{ij}) \rangle$ and the mass-polarization term $-\langle \nabla_i \cdot \nabla_j \rangle$, summed over all electron-pairs ij. All energies are in atomic units.

	$\langle \delta(r_{ij}) \rangle$	$\langle r_{ij}^{-2} \rangle$	$\langle r_{ij}^{-1} \rangle$	$\langle r_{ij} \rangle$	$\langle r_{ij}^2 \rangle$	$\langle r_{ij}^3 \rangle$	$-\langle \nabla_i \cdot \nabla_j \rangle$
Li	0.546(2)	4.386(6)	2.1993(1)	8.6574(4)	36.731(2)	191.00(2)	0.304(2)
Be	1.608(3)	9.532(4)	4.37330(9)	15.2895(2)	52.9958(9)	223.436(6)	0.466(1)
B	3.57(1)	17.45(1)	7.6657(3)	22.4696(6)	67.189(2)	244.63(1)	0.272(6)
C	6.66(5)	29.15(2)	12.5191(4)	29.0654(6)	73.574(2)	225.629(9)	−0.39(1)
N	11.1(2)	45.74(2)	19.2241(5)	35.5092(6)	77.550(2)	204.547(6)	−1.78(5)
O	17.9(2)	68.85(3)	27.9913(6)	42.5224(6)	83.693(2)	200.810(6)	−3.76(2)
F	27.0(3)	99.76(4)	39.2235(7)	49.3275(7)	87.469(2)	189.266(5)	−6.82(4)
Ne	38.6(5)	140.37(5)	53.2639(8)	55.8986(6)	89.587(1)	175.092(4)	−11.23(5)
Li$^+$	0.535(1)	4.088(6)	1.5684(1)	0.86210(7)	0.92684(9)	1.1888(2)	0.293(1)
Be$^+$	1.584(2)	8.901(3)	3.24613(7)	5.26857(9)	13.0819(3)	39.457(1)	0.456(1)
B$^+$	3.57(1)	16.47(1)	5.9706(3)	10.5411(3)	24.698(1)	69.807(4)	0.606(5)
C$^+$	6.67(3)	27.57(2)	10.0797(4)	16.2100(4)	34.253(1)	86.994(4)	−0.007(9)
N$^+$	11.33(7)	43.38(3)	15.9096(5)	21.9744(4)	41.344(1)	93.139(3)	−1.27(2)
O$^+$	17.8(1)	65.03(3)	23.7541(6)	27.8183(5)	46.843(1)	94.164(3)	−3.36(2)
F$^+$	26.3(9)	94.80(4)	33.9176(7)	34.0801(5)	52.794(1)	98.132(3)	−6.39(9)
Ne$^+$	38.0(5)	133.80(5)	46.7451(8)	40.2711(5)	57.231(1)	97.664(2)	−10.62(5)

assuming that the set of wave functions differ by a term of the form $\epsilon\Phi(\mathbf{R})$, where ϵ is a parameter and $\Phi(\mathbf{R})$ is an (unknown) wave function, but there is no guarantee that this assumption is valid. Perhaps a reliable *post hoc* correction scheme can be developed for DMC calculations.

For the most difficult case of the O atom we obtained an error in the energy of 0.74(4)% compared to the error of 2.17(8)% obtained by Brown *et al.* in 2007 [85]. There is every prospect of making substantial further improvements to the VMC and DMC results. Two obvious ways in which this can be achieved is by improving the optimization methods used and by using better wave function forms.

The stochastic optimization techniques used to obtain the optimal values of the wave function parameters have improved greatly in recent years, mainly due to the work of Umrigar and collaborators [81, 82]. The development of VMC sampling strategies which allow more reliable and efficient optimization of wave functions is extremely promising [47].

There have been also major improvements in the available wave function forms [69, 56, 57, 55], and many more such developments can be expected in the coming years. The cost of the QMC calculations reported here increases rapidly with system size because of the use of a multi-determinant expansion. However, we expect that the computational cost could be substantially reduced by using a more efficient representation such as geminal or Pfaffian wave functions. Another idea would be reduce the size of the multi-determinant expansion by more careful selection of the determinants. As the Jastrow factor primarily captures dynamic correlation, it replicates the efforts of a large multi-determinant expansion to some extent. There is no reason to assume that the same set of N determinants are the most important, i.e., have the largest weights, both in the absence and in the presence of a Jastrow factor. It is, however, expected that the weights of higher-excitation determinants decay more rapidly than those of lower energy determinants when a Jastrow factor is introduced. It would be beneficial for QMC calculations if a systematic way to select the determinants could be developed that accounts for static correlation and gives useful freedom in the nodal surface. For example, one could start with a large expansion of 300 CSFs that include up to quadruple excitations[5], and eliminate those with very small coefficients after an initial optimization alongside the Jastrow parameters (and backflow parameters, if included).

---

[5]MCHF calculations show that determinants with triple and quadruple excitations only feature in multi-determinant expansions with over 200 CSFs for an oxygen atom.

Chapter 4

A GENERIC JASTROW CORRELATION FACTOR

4.1 INTRODUCTION

A SLATER DETERMINANT is the simplest fermionic wave function. Its antisymmetric form describes exchange but not correlation. The most fruitful method of going beyond the Slater determinant is to multiply the Slater determinant by a Jastrow factor, which leads to the Slater-Jastrow wave function. The Jastrow factor is normally chosen to depend on the inter-particle separations, which introduces correlation into the wave function. The introduction of a Jastrow factor often leads to the recovery of 80% or more of the correlation energy of electronic systems.

The Jastrow factor is chosen to be everywhere positive and symmetric with respect to the exchange of identical particles in order to maintain the nodal surface defined by the rest of the wave function. It must also have a simple and physically-motivated form and be easy to evaluate. One of the features of the Jastrow factor is that it can conveniently be used to enforce the Kato cusp conditions [125] which determine the behaviour of the wave function when two charged particles approach one another. Enforcing the Kato cusp conditions does not necessarily improve the variational energy, but the reduction in the variance of the local energy is often very important.

DMC can be viewed as VMC with a perfect Jastrow factor, but improving the Jastrow factor can improve DMC calculations in several ways. Improving the trial wave function reduces the time-step errors and (normally very small) population-control errors [43] that the DMC algorithm is subject to. Using highly accurate trial wave functions helps in achieving more accurate results when evaluating expectation values of operators which do not commute with the Hamiltonian in DMC. Similar considerations apply when using non-local pseudopotentials, which involves making some approximations that are ameliorated by improving the trial wave function [59, 61]. As the fundamental limitation on the accuracy of DMC is the quality of the nodal surface, it is desirable to use trial wave

functions with optimizable nodal surfaces as afforded, for example, by multi-determinant wave functions and backflow transformations. A good Jastrow factor can account for the bulk of the dynamical correlation energy, which allows the optimization of parameters which affect the nodal surface to achieve a better nodal surface. The improved form of the wave function is also reflected in the lower variance and hence statistical error bars. The efficiency gain from using a more complex Jastrow factor in DMC is therefore a balance between the increase in cost of evaluating the wave function and the savings from more efficient sampling of configuration space afforded by larger time steps.

Here we introduce a highly flexible form of Jastrow factor which allows for the introduction of a variety of terms involving different numbers of particles. This allows one to quickly implement different functional forms and explore the importance of different correlations in any physical system we study.

Jastrow factors correlating several electrons have been used in earlier calculations, such as those of Refs. [126, 124, 127]. We study the effects of various three-body Jastrow terms and introduce a four-body van der Waals-like term. We also construct anisotropic Jastrow factors that can capture the natural symmetries of a system. We have successfully applied this generic Jastrow factor to a variety of systems, and we report results for the one- and two-dimensional homogeneous electron gases, the Be, B, and O atoms, and the BeH, H_2O, N_2, and H_2 molecules.

We briefly describe the form of the new Jastrow factor in Sec. 4.2. The development of the new Jastrow factor described in Sec. 4.2 and its implementation in CASINO was performed by Pablo López Ríos. I tested the code, reported bugs and produced the results presented in Sec. 4.4 with this new Jastrow factor.

4.2 CONSTRUCTION OF A GENERIC JASTROW FACTOR

QMC methods can be applied to systems which can be generically described as an ensemble of N quantum particles and M sources of external potential. For simplicity, we refer to quantum particles as electrons and to external potentials as nuclei in the rest of this chapter. Our Jastrow factor can be easily extended to other types of quantum particles and external potentials.

Typically $J(\mathbf{R})$ is constructed as a sum of terms, e.g.,

$$J(\mathbf{R}) = J_{e-e}(\mathbf{R}) + J_{e-n}(\mathbf{R}) + J_{e-e-n}(\mathbf{R}) + \ldots, \tag{4.1}$$

where "e–e" stands for "electron-electron," "e–n" for "electron–nucleus," etc. Each of these terms involves different numbers of electrons n and nuclei m. We shall refer to n and m as the electronic and nuclear ranks of a term, respectively, which are constrained to satisfy $n + m \geq 2$, $n \geq 1$, and $m \geq 0$. We have designed a generic Jastrow term

of selectable ranks, $J_{n,m}(\mathbf{R})$, such that the total Jastrow factor is constructed as the exponential of a sum of one or more terms of the desired ranks[1].

The function $J_{n,m}(\mathbf{R})$ is a sum over all sets of n electrons and m nuclei in the system of a parametrized function of the e–e and e–n relative position vectors within each such set. While alternatives exist, a natural way of parametrizing this function for arbitrary values of n and m (implying an arbitrary number of variables in the function) is to expand it in products of functions of the individual e–e and e–n vectors. Thus, we construct our Jastrow factor using these pairwise objects as building blocks.

We name the e–e functions used in the expansion $\Phi_\nu^P(\mathbf{r})$, where $\mathbf{r}$ is the relevant e–e relative position vector, ν is the index of the function within a chosen basis of functions, and P is the e–e dependency index, which allows the use of different optimizable parameters, if present, for parallel- and anti-parallel-spin electron pairs, for example. Similarly, the e–n basis functions are $\Theta_\mu^S(\mathbf{r})$, where $\mathbf{r}$ is the relevant e–n relative position vector, μ is the index of the function within the chosen basis set, and S is the e–n dependency index of the basis set, which allows the use of different parameters for up- and down-spin electrons around a given nucleus, or for different atoms, for example. In the case of non-electronic systems, e–e and e–n dependency indices are used to distinguish between particle types and spins.

We introduce a compact notation for defining $J_{n,m}(\mathbf{R})$. We represent the n electronic indices by the integer vector $\mathbf{i} = (i_1, i_2, \ldots, i_n)$, each of whose components takes a distinct value between 1 and N, and the m nuclear indices by the integer vector $\mathbf{I} = (I_1, I_2, \ldots, I_m)$, each of whose components takes a distinct value between 1 and M. For each term in the Jastrow factor we define the e–e and e–n dependency matrices $\underline{P}$ and $\underline{S}$ of respective sizes $N \times N$ and $N \times M$ containing the dependency indices P_{ij} and S_{iI} for each e–e and e–n pair. The components of $\underline{P}$ and $\underline{S}$ can be made equal depending on the symmetries of the system, including particle distinguishability and geometrical symmetries which make different nuclei equivalent.

It is convenient to use matrices to represent the basis functions involved in the Jastrow factor term. For e–e basis functions, each row and column of the $n \times n$ matrix $\underline{\Phi}$ corresponds to an electron,

$$\underline{\underline{\Phi}}_{\underline{\nu}}^{\underline{P}}(\mathbf{i}) = \begin{pmatrix} 0 & \Phi_{\nu_{i_1 i_2}}^{P_{i_1 i_2}}(\mathbf{r}_{i_1 i_2}) & \cdots & \Phi_{\nu_{i_1 i_n}}^{P_{i_1 i_n}}(\mathbf{r}_{i_1 i_n}) \\ \Phi_{\nu_{i_1 i_2}}^{P_{i_1 i_2}}(\mathbf{r}_{i_2 i_1}) & 0 & \cdots & \Phi_{\nu_{i_2 i_n}}^{P_{i_2 i_n}}(\mathbf{r}_{i_2 i_n}) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{\nu_{i_1 i_n}}^{P_{i_1 i_n}}(\mathbf{r}_{i_n i_1}) & \Phi_{\nu_{i_2 i_n}}^{P_{i_2 i_n}}(\mathbf{r}_{i_n i_2}) & \cdots & 0 \end{pmatrix}. \tag{4.2}$$

Both $\underline{P}(\mathbf{i})$, the $n \times n$ matrix formed by the e–e dependency indices, and the $n \times n$ matrix of e–e expansion indices $\underline{\nu}$ are defined to be symmetric, and this fact has been used in Eq.

---

[1] In this notation, $J_{e-e} \equiv J_{2,0}$, $J_{e-n} \equiv J_{1,1}$, $J_{e-e-n} \equiv J_{2,1}$, etc.

(4.2). Noting that $\mathbf{r}_{ji} = -\mathbf{r}_{ij}$, and restricting the e–e functions to be either symmetric or antisymmetric about the origin, one finds that $\underline{\Phi}$ is symmetric, antisymmetric, or asymmetric depending on whether the functions in the basis set are all symmetric, all antisymmetric, or both types are present, respectively.

For e–n basis functions each row of the $n \times m$ matrix $\underline{\Theta}$ corresponds to an electron and each column to a nucleus,

$$\underline{\Theta}_{\underline{\mu}}^{\mathrm{S}}(\mathbf{i}, \mathbf{I}) = \begin{pmatrix} \Theta_{\mu_{i_1 I_1}}^{S_{i_1 I_1}}(\mathbf{r}_{i_1 I_1}) & \cdots & \Theta_{\mu_{i_1 I_m}}^{S_{i_1 I_m}}(\mathbf{r}_{i_1 I_m}) \\ \vdots & & \vdots \\ \Theta_{\mu_{i_n I_1}}^{S_{i_n I_1}}(\mathbf{r}_{i_n I_1}) & \cdots & \Theta_{\mu_{i_n I_m}}^{S_{i_n I_m}}(\mathbf{r}_{i_n I_m}) \end{pmatrix}. \tag{4.3}$$

We refer to the $n \times m$ matrix of e–n dependency indices as $\underline{S}(\mathbf{i}, \mathbf{I})$, and the $n \times m$ matrix of e–n expansion indices as $\underline{\mu}$.

We write $J_{n,m}$ as a sum of contributions from each group of n electrons and m nuclei in the system,

$$J_{n,m} = \frac{1}{n!m!} \sum_{\mathbf{i}} \sum_{\mathbf{I}} J_{n,m}(\mathbf{i}, \mathbf{I}) = \sum_{\mathbf{i}}^{\mathrm{s.v.}} \sum_{\mathbf{I}}^{\mathrm{s.v.}} J_{n,m}(\mathbf{i}, \mathbf{I}), \tag{4.4}$$

where summations with vector indices represent sums in which every component of the vector is a summation index, and "s.v." (for "sorted vector") indicates that the sum is restricted to vectors whose components are sorted, e.g., $i_1 < i_2 < \ldots < i_n$, which avoids redundant contributions[2]. The contribution from the n-electron and m-nucleus group $\{\mathbf{i}, \mathbf{I}\}$ is

$$J_{n,m}(\mathbf{i}, \mathbf{I}) = \sum_{\underline{\nu}}^{\mathrm{u.t.}} \sum_{\underline{\mu}}^{\mathrm{u.t.}} \lambda_{\underline{\nu},\underline{\mu}}^{\mathrm{P}(\mathbf{i}),\mathrm{S}(\mathbf{i},\mathbf{I})} \prod \underline{\Phi}_{\underline{\nu}}^{\mathrm{P}}(\mathbf{i}) \prod \underline{\Theta}_{\underline{\mu}}^{\mathrm{S}}(\mathbf{i}, \mathbf{I}), \tag{4.5}$$

where λ are the linear parameters, summations with matrix indices represent sums in which every component of the matrix is a summation index, $\prod$ acting on matrices implies the product of all of their components, and "u.t." means that the relevant operation is restricted to the upper-triangular portion of the e–e matrices involved, excluding the diagonal.

4.2.1 INDEXING OF BASIS FUNCTIONS

The components of $\underline{\nu}$ and $\underline{\mu}$ are the e–e and e–n expansion indices. We define the expansion indices so that they can each take any value between 1 and the e–e expansion order p, and between 1 and the e–n expansion order q, respectively. We factorize an optional cut-off

---

[2]For example, $\sum_{i \neq j}^{N} f_{i,j}$ contains $N(N-1)$ terms, but if $f_{i,j} = f_{j,i}$, this operation can be simplified to $\sum_{i<j}^{N} f_{i,j}$, which contains $N(N-1)/2$ terms. In general, the sum of an object which is symmetric with respect to the interchange of any of its n indices, where these indices take non-repeated values, can be rewritten as a sum over sorted sets of indices involving $n!$ times fewer terms than the original sum.

function into Φ_ν^P and Θ_μ^S, so that

$$\Phi_\nu^P(\mathbf{r}) = f^P(\mathbf{r})\phi_\nu^P(\mathbf{r}) \tag{4.6}$$

for $\nu > 0$, and

$$\Theta_\mu^S(\mathbf{r}) = g^S(\mathbf{r})\theta_\mu^S(\mathbf{r}) \tag{4.7}$$

for $\mu > 0$, where f^P and g^S are the e–e and e–n cut-off functions and ϕ_ν^P and θ_μ^S are functions from a suitable basis set. This factorization allows an efficient implementation of localized Jastrow factor terms.

Additionally, we allow expansion indices to take a value of zero with the special meaning that $\Phi_0^P(\mathbf{r}) = \Theta_0^S(\mathbf{r}) = 1$ for all P, S, and $\mathbf{r}$. Note that these zeroth-order functions do not contain cut-off functions. This allows us to construct terms with specialized functional forms, such as those involving dot products of vectorial quantities.

4.2.2   CONSTRAINTS

Constraints on the parameters can be expressed in the form of a system of equations involving the linear parameters and the basis function parameters. We restrict our analysis to linear constraints on the linear parameters, and constraints that can be imposed on the non-linear parameters contained in a basis function independently from the linear parameters and from non-linear parameters in other basis functions.

Linear constraints on the linear parameters can be imposed using Gaussian elimination, as described in Ref. [84]. The matrix of coefficients may depend on the non-linear parameters in the basis functions, if present, and the linear system is usually underdetermined, resulting in a subset of the parameters being determined by the values of the remaining parameters, which can be optimized directly.

4.2.2.1   SYMMETRY CONSTRAINTS

Equation (4.4) imposes the condition that $J_{n,m}(\mathbf{i}, \mathbf{I})$ must not depend on the specific ordering of the electrons and nuclei listed in $\mathbf{i}$ and $\mathbf{I}$. The linear parameters of the Jastrow factor must exhibit a symmetry that implies that a parameter with a given set of superindices $\{\underline{P}(\mathbf{i}), \underline{S}(\mathbf{i}, \mathbf{I})\}$ is determined by another parameter with a permuted set of superindices $\{\underline{P}'(\mathbf{i}), \underline{S}'(\mathbf{i}, \mathbf{I})\}$. This redundancy is removed by considering only one of the possible permutations of $\{\underline{P}(\mathbf{i}), \underline{S}(\mathbf{i}, \mathbf{I})\}$, or the signature. These symmetry constraints amount to equalities between pairs of parameters and must always be imposed, otherwise the trial wave function is unphysical and calculations give erroneous results.

4.2.2.2 CONSTRAINTS AT E–E AND E–N COALESCENCE POINTS

The Coulomb potential energy diverges when the positions of two electrons or an electron and a nucleus coincide. However, the local energy of an eigenstate of the Hamiltonian, including the exact ground-state wave function, is finite and constant throughout configuration space. Divergences in the local energy are therefore not a feature of the exact wave function and can lead to poor statistics in QMC calculations; hence it is important to avoid them. The kinetic energy must diverge to cancel out the potential energy and keep the local energy finite, which is achieved by demanding that the wave function obeys the Kato cusp conditions [125]. For any two charged particles i and j in a two- or three-dimensional system interacting via the Couloumb potential, these are

$$\left(\frac{1}{\Psi} \frac{\partial \hat{\Psi}}{\partial r_{ij}} \right)_{r_{ij} \to 0} = \frac{2 q_i q_j \mu_{ij}}{d \pm 1}, \tag{4.8}$$

where $\hat{\Psi}$ denotes the spherical average of Ψ, q represents charge, $\mu_{ij} = m_i m_j / (m_i + m_j)$ is the reduced mass, m represents mass, d is the dimensionality, and the positive sign in the denominator is for indistinguishable particles and the negative sign is for distinguishable particles. Fixed nuclei are regarded as having an infinite mass.

As typical forms of Ψ_D explicitly depend only on e–n distances, it is common practice to impose the e–n cusp conditions on Ψ_D and the e–e cusp conditions on the Jastrow factor. Our implementation allows the option of applying both types of cusp conditions to the Jastrow factor, which gives flexibility in the choice of Ψ_D and its properties. In particular, we impose the cusp conditions on a single Jastrow factor term, and constrain all other terms in the Jastrow factor so that their contribution to the local kinetic energy is finite at e–e and e–n coalescence points. For non-divergent interaction potentials, such as most pseudopotentials, we simply require that the kinetic energy remains finite at coalescence points. Our implementation is also capable of not applying any constraints at e–e and e–n coalescence points since this is advantageous in some cases [128, 129], as discussed in Sec. 4.4.3.5.

Imposing that the kinetic energy be finite at coalescence points is non-trivial if the Jastrow factor contains anisotropic functions. As two particles coalesce, $\boldsymbol{\nabla} J$ and $\nabla^2 J$ must remain finite, and this gives rise to further constraints that must be satisfied.

4.2.2.3 OTHER CONSTRAINTS

It is possible to construct terms containing dot products by using appropriate constraints. For example, consider the basis functions $\Theta_1(\mathbf{r}) = x$, $\Theta_2(\mathbf{r}) = y$, and $\Theta_3(\mathbf{r}) = z$. In an e–n–n term we can restrict the indices so that $\underline{\mu}$ takes only the values (1 1), (2 2), (3 3), so that the contribution of electron i and nuclei I and J is $\mathbf{r}_{iI} \cdot \mathbf{r}_{iJ}$, provided we also apply a linear constraint that equates the three non-zero linear coefficients.

It is also possible to introduce Boys-Handy-style indexing [13], where the sum of all e–e and e–n indices is restricted to be less than or equal to some fixed integer l.

4.3 BASIS FUNCTIONS AND TERMS

We employ a condensed notation to refer to Jastrow terms that use certain basis functions, cut-off functions and constraints. Each term is represented by a single capital letter, with n and m as subindices. Any other relevant information is given as a superindex. Table 4.1 summarizes the notation for the terms we have introduced.

Possibly the simplest basis set is the natural powers,

$$N_\nu(\mathbf{r}) = r^{\nu-1}, \tag{4.9}$$

as used in the DTN Jastrow factor for the localized u, χ, and f terms [84]. These functions need to be cut off at some radius L, for which purpose the DTN Jastrow factor uses the polynomial cut-off function

$$D(\mathbf{r}) = (r - L)^C \Theta(L - r), \tag{4.10}$$

where L is an optimizable parameter, C is a positive integer, and $\Theta(r)$ is the Heaviside step function. We also use a slightly different version of this cut-off function,

$$P(\mathbf{r}) = (1 - r/L)^C \Theta(L - r), \tag{4.11}$$

which should be numerically superior to $D(\mathbf{r})$.

For simple Jastrow terms we use the natural power basis functions N_ν and the polynomial cut-off functions P or D. We refer to these terms as $N_{n,m}$. $N_{2,0}$, $N_{1,1}$, and $N_{2,1}$ are the equivalent of the DTN u, χ, and f terms, respectively. In the $N_{2,1}$ term, and in any term where more than one electron and one or more nuclei are involved, we choose not to apply e–e cut-off functions, relying instead on the e–n cut-offs to fulfill this role. Additional $N_{n,m}$ terms used here that were not part of the DTN Jastrow factor are $N_{1,2}$, $N_{3,0}$, $N_{1,3}$, $N_{2,2}$, $N_{3,1}$, and $N_{4,0}$. In $N_{n,m}$ we typically use a truncation order in the cut-off function of $C = 3$ to ensures that the local energy is continuous.

A particular variant of $P(\mathbf{r})$ is the anisotropic cut-off function

$$A(\mathbf{r}) = (1 - r/L)^C \Theta(L - r) \sum_i c_i \prod_\beta^d \left[\frac{\mathbf{r} \cdot \hat{\mathbf{u}}_\beta}{r} \right]^{p_\beta^{(i)}}, \tag{4.12}$$

where L is an optimizable parameter, C is a positive integer, d is the dimensionality of the system, $\hat{\mathbf{u}}_\beta$ are unit vectors along d orthogonal directions, c_i are real-valued constants, and

$p_\beta^{(i)}$ are integer exponents, which are constrained so that $\sum_\beta^d p_\beta^{(i)}$ is the same for all values of i. This cut-off function is simply the product of an isotropic cut-off function and a spherical harmonic. For example, with $c_1 = 3$, $c_2 = -1$, $\mathbf{p}^{(1)} = (2\ 1\ 1)$, and $\mathbf{p}^{(2)} = (0\ 3\ 1)$, and the vectors pointing along the Cartesian axes, we obtain

$$A(\mathbf{r}) = (1 - r/L)^C \Theta(L - r) \left[\frac{(3x^2 - y^2)yz}{r^4} \right], \tag{4.13}$$

which is proportional to a real spherical harmonic with $l = 4$. The advantage of describing anisotropy in the cut-off function rather than in the basis functions is that the common spherical harmonic can be factorized out of the sum over expansion indices, which reduces the computational cost. We allow different orientations to be used for different e–e or e–n dependency indices, which is useful to adapt the functional form to, e.g., the geometry of a molecule.

We use $A_{n,m}^{\text{s.h.}}$ to refer to the anisotropic variant of $N_{n,m}$. This term consists of natural power basis functions N_ν and the anisotropic cut-off function A, and "s.h." is a placeholder for the description of the spherical harmonic. For example, for the highly anisotropic N_2 molecule we use terms such as $A_{1,1}^z$, $A_{1,1}^{z^2}$, $A_{2,1}^z$, and $A_{2,1}^{z^2}$.

An alternative to the natural-power basis in finite systems is a basis of powers of fractions which tend to a constant as $r \to \infty$, and therefore do not need to be cut off. We define the basis

$$F_\nu(\mathbf{r}) = \left(\frac{r}{r^b + a} \right)^{\nu - 1}, \tag{4.14}$$

where a and b are real-valued optimizable parameters. Similar basis sets with $b = 1$ have been used in the literature, often in conjunction with Boys-Handy-style indexing [13, 130, 131, 128], and this basis was used in Chapter 3 with an early implementation of the Jastrow factor presented here.

The F_ν basis functions are used in terms $F_{n,m}$, or $F_{n,m}^{b=1}$ when we force $b = 1$ in the basis functions. In some systems it is useful to apply Boys-Handy-style indexing to $F_{n,m}^{b=1}$, and we refer to the resulting term as $B_{n,m}$.

In extended systems it is important to use a basis that is consistent with the geometry of the simulation cell and has the periodicity of the system, such as a cosine basis,

$$C_\nu(\mathbf{r}) = \sum_{\mathbf{G} \in \nu\text{-th star}} \cos\left(\mathbf{G} \cdot \mathbf{r}\right), \tag{4.15}$$

where the $\mathbf{G}$ vectors are arranged in stars defined by the cell geometry.

Terms denoted by $C_{n,m}$ use of the cosine basis functions C_ν. We choose expansion orders so that at least as many $\mathbf{G}$ vectors as electrons in each spin channel are included in the expansion. These terms capture correlation in the corners of the simulation cell beyond the cut-off radius of $N_{n,m}$ terms. They are also important in describing strongly

anisotropic materials. However, the long-range nature of these terms makes them more computationally expensive to evaluate.

$C_{2,0}$ and $C_{1,1}$ correspond to the DTN p and q terms, respectively. While these terms are computationally expensive to evaluate as they are not cut off at any distance, they are also important in describing strongly anisotropic materials.

A suitable basis set for building specialized terms containing dot products is

$$V_\nu(\mathbf{r}) = r^{\texttt{int}[(\nu-1)/d]} \frac{\mathbf{r} \cdot \hat{\mathbf{u}}_{\texttt{mod}(\nu-1,d)+1}}{r}, \tag{4.16}$$

where d is the dimensionality of the system and $\hat{\mathbf{u}}_\beta$ are the d unit vectors parallel to the Cartesian axes. A term constructed using these functions with appropriate constraints would consist of dot products between two vectors multiplied by a natural-power expansion in their moduli.

To test the flexibility of our implementation we have designed an e–e–n–n Jastrow term for describing the correlations associated with van der Waals interactions, which we call $V_{2,2}$. This term is capable of distinguishing between configurations where the electron-nucleus relative position vectors $\mathbf{r}_{iI}$ and $\mathbf{r}_{jJ}$ are parallel from those where they are anti-parallel. Introducing a dot product achieves this effect, and $V_{2,2}$ has the following functional form,

$$\begin{aligned}
V_{2,2} &= \frac{1}{2} \sum_{i\neq j}^N \sum_{I\neq J}^M P(r_{iI}) P(r_{jJ}) \sum_{\nu_{ij}}^p \sum_{\mu_{iI},\mu_{jJ}}^q \lambda_{\nu_{ij}\mu_{iI}\mu_{iJ}} \\
&\quad \times N_{\nu_{ij}}(r_{ij}) N_{\mu_{iI}}(r_{iI}) N_{\mu_{jJ}}(r_{jJ}) \mathbf{r}_{iI} \cdot \mathbf{r}_{jJ}.
\end{aligned} \tag{4.17}$$

We require basis functions to be scalars in our Jastrow factor, so the dot product is separated into its components. Hence, we construct the $V_{2,2}$ term using V_ν for the e–n basis with P as the e–n cut-off functions, and N_ν for the e–e basis. We allow e–n indices to be zero.

TABLE 4.1: Notation for Jastrow terms correlating n electrons and m nuclei using different basis functions.

Name	Basis set	Cut-off function	Special constraints
$N_{n,m}$	Natural powers	Polynomial	None
$F_{n,m}$	Powers of $r/(r^b + a)$	None	None
$B_{n,m}$	Powers of $r/(r + a)$	None	Boys-Handy-style indexing
$A_{n,m}^{\text{s.h.}}$	Natural powers	Anisotropic polynomial	None
$C_{n,m}$	Cosines	None	None
$V_{n,m}$	Natural powers times unit vectors	Polynomial	Dot product

4.4   RESULTS

We have used a variety of methods to optimize our Jastrow factors, namely variance minimization, minimization of the mean absolute deviation of the local energy with respect to the median energy, and linear least-squares VMC energy minimization [81, 82]. All of our final wave functions are energy-minimized except where otherwise stated. Starting with the Hartree-Fock wave function, we progressively introduce Jastrow terms and re-optimize all of the parameters simultaneously. Optimizing the Jastrow factor term-by-term is unnecessary in practical applications, but here it allows us to understand the importance of the different terms. We refer to the total number of optimizable parameters in the wave function as N_p.

To measure the quality of the trial wave function Ψ, we again use the fraction of the correlation energy retrieved in a VMC calculation with a given trial wave function Ψ,

$$f_{\text{CE}}[\Psi] = \frac{E_{\text{HF}} - E_{\text{VMC}}[\Psi]}{E_{\text{HF}} - E_{\text{exact}}}. \tag{4.18}$$

We refer to the difference between the DMC and HF energies as the DMC correlation energy, $E_{\text{HF}} - E_{\text{DMC}}[\Psi]$. The fraction of the DMC correlation energy retrieved in VMC,

$$f_{\text{DCE}}[\Psi] = \frac{E_{\text{HF}} - E_{\text{VMC}}[\Psi]}{E_{\text{HF}} - E_{\text{DMC}}[\Psi]}, \tag{4.19}$$

measures the quality of the Jastrow factor, since a perfect Jastrow factor would make the VMC and DMC energies coincide[3]. We define the fraction of the remaining DMC correlation energy recovered by a wave function Ψ_2 with respect to another Ψ_1 as

$$\frac{E_{\text{VMC}}[\Psi_1] - E_{\text{VMC}}[\Psi_2]}{E_{\text{VMC}}[\Psi_1] - E_{\text{DMC}}[\Psi_2]}. \tag{4.20}$$

As VMC variance tends to its lower bound of zero as Ψ tends to an eigenstate of the Hamiltonian, the variance is also a measure of the overall quality of the trial wave function.

4.4.1   HOMOGENEOUS ELECTRONS GASES

4.4.1.1   ONE-DIMENSIONAL HOMOGENEOUS ELECTRON GAS

We have studied a 1D homogeneous electron gas (HEG) of density parameter $r_s = 5$ a.u. consisting of 19 electrons subject to periodic boundary conditions using a single Slater determinant of plane-wave orbitals. The ground-state energy of an infinitely thin 1D HEG in which electrons interact by the Coulomb potential is independent of the magnetic state,

---

[3]In general, attaining the DMC limit in VMC would require non-analyticities in the Jastrow factor. Nonetheless this theoretical limit is useful for assessing the performance of a Jastrow factor.

and hence we have chosen all the electrons to have the same spin. This system is unusual in that the nodal surface of the trial function is exact, and therefore DMC gives the exact ground-state energy, which we have estimated to be $-0.2040834(3)$ a.u. per electron. Excellent results were reported for this system in Refs. [133, 134] using wave functions with e–e backflow transformations [124, 69] which preserve the (exact) nodal surface of the Slater determinant. Our VMC results for different Jastrow factors are given in Table 4.2.

TABLE 4.2: Energies (E) and VMC variances (V) of the 1D HEG at $r_s = 5$ a.u. using different Jastrow factors. The use of backflow is indicated by "(BF)".

	N_p	E (a.u. per electron)	V (a.u.)	f_{DCE} (%)
HF		-0.191653064		0
$N_{2,0}$	9	$-0.204076(1)$	0.0000654(7)	99.941(8)
$N_{2,0}+C_{2,0}$	18	$-0.2040824(7)$	0.0000168(3)	99.992(6)
$N_{2,0}+C_{2,0}+N_{3,0}$	45	$-0.2040831(2)$	0.00000171(3)	99.998(3)
$N_{2,0}+C_{2,0}+C_{3,0}$	52	$-0.2040832(6)$	0.0000127(3)	99.998(5)
$N_{2,0}+C_{2,0}+N_{3,0}+C_{3,0}$	79	$-0.2040833(2)$	0.00000105(3)	99.999(3)
$N_{2,0}$ (BF)	18	$-0.2040816(5)$	0.00000809(6)	99.986(5)
$N_{2,0}+C_{2,0}$ (BF)	27	$-0.2040833(2)$	0.00000104(3)	99.999(3)
$N_{2,0}+C_{2,0}+N_{3,0}$ (BF)	54	$-0.2040832(1)$	0.00000055(2)	99.998(3)
$N_{2,0}+C_{2,0}+C_{3,0}$ (BF)	61	$-0.20408310(7)$	0.00000020(1)	99.998(3)
$N_{2,0}+C_{2,0}+N_{3,0}+C_{3,0}$ (BF)	88	$-0.20408310(7)$	0.00000020(1)	99.998(3)
DMC		$-0.2040834(3)$		100.000(4)

TABLE 4.3: e–e expansion orders (p) used for the different Jastrow terms in the 1D HEG.

	$N_{2,0}$	$N_{3,0}$	$C_{2,0}$	$C_{3,0}$
p	9	5	10	5

We have investigated the improvement in VMC results when various terms are added to an e–e Jastrow factor $J = N_{2,0}+C_{2,0}$, both with and without backflow transformations. In the absence of backflow, we find that including $N_{3,0}$, $C_{3,0}$, or $N_{3,0} + C_{3,0}$ improves the VMC energy, while the subsequent addition of $C_{4,0}$ yields no further gain. We observe a ten-fold reduction in the variance upon addition of $N_{3,0}$ to $J = N_{2,0} + C_{2,0}$. The $C_{3,0}$ term does not duplicate the $N_{3,0}$ term, and they combine to give a further reduction in variance.

VMC gives an almost exact energy with backflow and $J = N_{2,0} + C_{2,0}$, and therefore no further reduction is possible by including more Jastrow terms. However, the addition of $N_{3,0} + C_{3,0}$ reduces the VMC variance by a factor of five, giving a variance that is an order of magnitude smaller than that reported in Ref. [133] for a similar calculation.

The energy of a SJ wave function with a $J = N_{2,0}+C_{2,0}+N_{3,0}$ containing 45 optimizable parameters is within error bars of the exact (DMC) energy, while a SJB wave function

with only $J = N_{2,0} + C_{2,0}$ and a total of 27 optimizable parameters is required to achieve this. Backflow transformations introduce the variational freedom more compactly than the $N_{3,0}$ term.

4.4.1.2 TWO-DIMENSIONAL HOMOGENEOUS ELECTRON GAS

We have studied a paramagnetic 2D HEG with 42 electrons per simulation cell at $r_s = 35$ a.u., which lies close to the ferromagnetic Wigner crystallization density predicted by Drummond and Needs [72]. Kwon *et al.* [124] found that three-electron correlations are important at low densities, and that the effect of a three-electron Jastrow factor on the VMC energy is comparable to that of backflow. At higher densities, the effects of velocity-dependent backflow transformations become more dominant. This makes low densities appealing for testing higher-rank Jastrow terms. The VMC energy and variance obtained using different Jastrow factors with and without backflow is plotted in Fig. 4.1 and the results are given in Table 4.2.

The addition of an $N_{3,0}$ term to $J = N_{2,0}$ recovers 81% of the remaining DMC correlation energy without backflow and 49% with backflow. The $C_{2,0}$ term further reduces both the VMC energy and variance. The use of a $C_{3,0}$ term recovers 10% of the remaining DMC correlation energy when added to $J = N_{2,0} + C_{2,0}$, but it was not used further since the lack of a cut-off function makes calculations with $C_{3,0}$ too costly for the little benefit it provides.

We have also performed DMC calculations using two different Jastrow factors in the presence of backflow in order to quantify the indirect effect of the quality of the Jastrow factor on the quality of the nodes of the wave function. We obtain a DMC energy of $-0.0277072(1)$ a.u. per electron for $J = N_{2,0}$, and a lower energy of $-0.0277087(1)$ a.u. per electron for $J = N_{2,0} + N_{3,0} + C_{2,0}$. This supports the idea that a better Jastrow factor allows the backflow transformation to shift its focus from the "bulk" of the wave function to its nodes, thus improving the DMC energy.

4.4.2 BE, B AND O ATOMS

While excellent descriptions of these atoms can be obtained within VMC and DMC using multi-determinant wave functions with backflow correlations [132, 85], we have used single-determinant wave functions since we are only interested in the effects of the Jastrow factor. The decrease in quality of the Jastrow factor for heavier atoms can be attributed to the increase in inhomogeneity as Z increases. The higher-order terms are expected to therefore give improve the wave function. We have studied the ground states of the Be, B, and O atoms, corresponding to ^{1}S, ^{2}P, and ^{3}P electronic configurations, respectively. The ATSP2K code [83] was used to generate numerical single-electron HF orbitals tabulated on a radial grid. We have investigated the use of Jastrow factors with up to four-body

TABLE 4.4: Energies (E) and VMC variances (V) of the 2D HEG at $r_s = 35$ a.u. using different Jastrow factors. The use of backflow is indicated by "(BF)".

	N_p	E (a.u. per electron)	V (a.u.)	f_{DCE} (%)	$f_{\text{DCE(BF)}}$ (%)
HF		-0.017121792		0	0
$N_{2,0}$	18	$-0.0275341(8)$	$0.0000620(5)$	98.76(1)	
$N_{2,0}+C_{2,0}$	24	$-0.0275400(8)$	$0.0000593(6)$	98.82(1)	
$N_{2,0}+N_{3,0}$	57	$-0.0276050(6)$	$0.0000378(4)$	99.43(1)	
$N_{2,0}+C_{2,0}+N_{3,0}$	65	$-0.0276112(6)$	$0.0000339(4)$	99.49(1)	
$N_{2,0}+C_{2,0}+C_{3,0}$	132	$-0.0275520(7)$	$0.0000546(6)$	98.93(1)	
$N_{2,0}$ (BF)	35	$-0.0276009(7)$	$0.0000462(4)$		98.982(7)
$N_{2,0}+C_{2,0}$ (BF)	43	$-0.0276046(7)$	$0.0000436(5)$		99.017(7)
$N_{2,0}+N_{3,0}$ (BF)	76	$-0.0276541(5)$	$0.0000269(3)$		99.484(5)
$N_{2,0}+C_{2,0}+N_{3,0}$ (BF)	84	$-0.0276614(5)$	$0.0000232(3)$		99.553(5)
DMC		$-0.0276649(9)$		100.00(1)	99.586(9)
DMC for $N_{2,0}$ (BF)		$-0.0277072(1)$			99.986(1)
DMC for $N_{2,0}+C'_{2,0}+N_{3,0}$ (BF)		$-0.0277087(1)$			100.000(1)
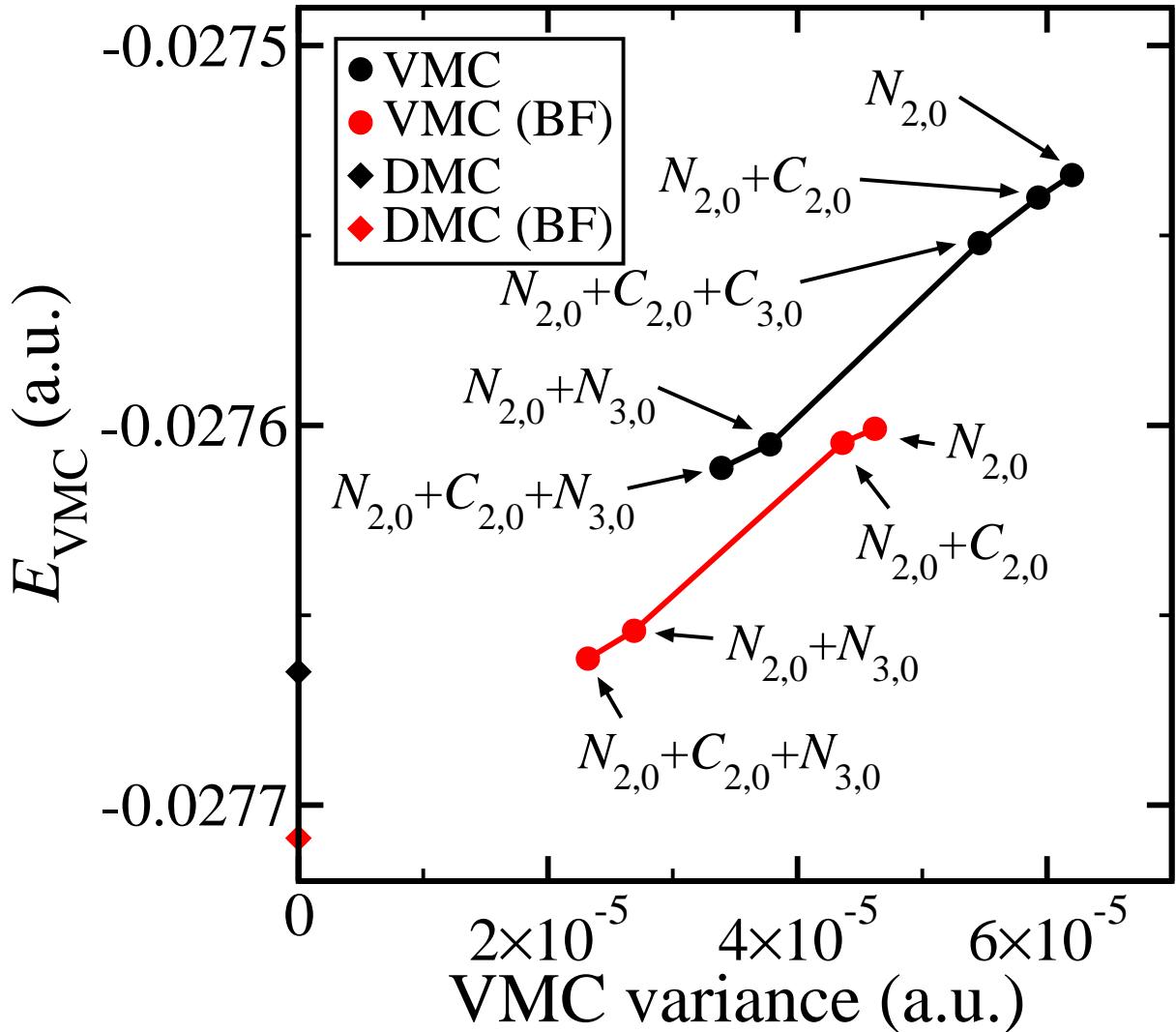
FIGURE 4.1: VMC energies against the VMC variance for the 2D HEG at $r_s = 35$ a.u. using different Jastrow factors, along with the DMC energies for reference. The error bars are smaller than the size of the symbols, and "(BF)" indicates the use of backflow.

TABLE 4.5: e–e expansion orders (p) used for the different Jastrow terms in the 2D HEG.

	$N_{2,0}$	$N_{3,0}$	$C_{2,0}$	$C_{3,0}$
p	9	4	5	3

terms, but we have not used backflow for these systems. The energies of Chakravorty *et al.* [88] have been used as "exact" reference values. Our VMC results for Be, B and O are given in Table 4.6.

TABLE 4.6: Energies (E) and VMC variances (V) for the Be, B and O atoms using different Jastrow factors.

	N_p	E (a.u.)	V (a.u.)	$f_{CE}(\%)$	$f_{DCE}(\%)$
Be atom					
HF		−14.573023		0	0
$F_{2,0}+F_{1,1}+F_{2,1}$	103	−14.65062(7)	0.0445(5)	82.26(7)	92.22(9)
$F_{2,0}+F_{1,1}+F_{2,1}+F_{3,0}$	160	−14.6512(1)	0.0470(3)	82.9(1)	92.9(1)
$F_{2,0}+F_{1,1}+F_{2,1}+F_{3,1}$	170	−14.6522(1)	0.051(1)	83.9(1)	94.0(1)
VMC from Ref. [85]		−14.6311(1)		61.6(1)	69.0(1)
VMC from Ref. [135]		−14.64972(5)		81.30(5)	91.15(7)
DMC		−14.65717(4)		89.20(4)	100.00(7)
Exact from Ref. [88]		−14.66736		100	
B atom					
HF		−24.529061		0	0
$F_{2,0}+F_{1,1}+F_{2,1}$	103	−24.6299(1)	0.093(1)	80.77(8)	90.9(1)
$F_{2,0}+F_{1,1}+F_{2,1}+F_{3,0}$	185	−24.6302(1)	0.0960(5)	81.01(8)	91.1(1)
$F_{2,0}+F_{1,1}+F_{2,1}+F_{3,1}$	195	−24.6309(2)	0.0973(6)	81.6(2)	91.8(2)
VMC from Ref. [85]		−24.6056(2)		61.3(2)	69.0(2)
VMC from Ref. [135]		−24.62936(5)		80.34(4)	90.39(7)
DMC		−24.64002(6)		88.87(5)	100.00(8)
Exact from Ref. [88]		−24.65391		100	
O atom					
HF		−74.809398		0	0
$F_{2,0}+F_{1,1}+F_{2,1}$	103	−75.0341(2)	0.550(2)	87.13(8)	92.97(9)
$F_{2,0}+F_{1,1}+F_{2,1}+F_{3,0}$	185	−75.0368(4)	0.577(2)	88.2(2)	94.1(2)
$F_{2,0}+F_{1,1}+F_{2,1}+F_{3,1}$	195	−75.0381(3)	0.498(2)	88.7(1)	94.6(1)
VMC from Ref. [85]		−75.0233(3)		82.9(1)	88.5(1)
VMC from Ref. [135]		−75.0352(1)		87.55(4)	93.42(6)
DMC		−75.0511(1)		93.72(4)	100.00(6)
Exact from Ref. [88]		−75.0673		100	

TABLE 4.7: e–e and e–n expansion orders (p and q, respectively) used for the different Jastrow factor terms in the Be, B, and O atoms.

	$F_{2,0}$	$F_{1,1}$	$F_{2,1}$	$F_{3,1}$
p	9	–	5	3
q	–	9	5	3

We obtain lower single-determinant VMC energies for the Be, B, and O atoms with $J = F_{2,0}+F_{1,1}+F_{2,1}$ than reported in Refs. [85, 135]. We obtain further small improvements in the VMC energies by including either $F_{3,0}$ or $F_{3,1}$ Jastrow terms, but their combination,

$F_{3,0} + F_{3,1}$, is not found to be advantageous over using the terms individually. This indicates that $F_{3,0}$ and $F_{3,1}$, the latter of which provides a slightly lower VMC energy than the former, have nearly the same effect in these atoms. These three-electron terms should be particularly useful in describing correlations involving electrons in the atomic core region. We expect $F_{3,1}$ to be more useful than $F_{3,0}$ in molecules and solids because it should be able to adapt to the different length scales in these systems, whereas $F_{3,0}$ offers a homogeneous description of three-electron correlations. We have investigated the effect of adding a $F_{4,1}$ term in Be and O, but it does not reduce the VMC energy or variance when added to $J = F_{2,0} + F_{1,1} + F_{2,1} + F_{3,1}$.

Our best VMC energies of $-14.6522(1)$ a.u., $-24.6309(2)$ a.u., and $-75.0381(3)$ a.u. for Be, B and O respectively correspond to fractions of the DMC correlation energy of $94.0(1)\%$, $91.8(1)\%$, and $94.6(1)\%$.

4.4.3   BeH, N$_2$, H$_2$O and H$_2$ molecules

The BeH, N$_2$ and H$_2$O molecules are strongly inhomogeneous and anisotropic systems. We have used basis sets of moderate quality for the single-electron orbitals of BeH and N$_2$ in order to investigate the extent to which the Jastrow factor can compensate for the deficiencies of the basis sets, especially via one-electron terms $N_{1,m}$. For H$_2$O and the H$_2$ triplet we have used very good basis sets. We have also tested anisotropic Jastrow factors in N$_2$, and a van der Waals-like Jastrow factor for H$_2$.

4.4.3.1   BeH molecule

We have studied the all-electron BeH molecule in the $^2\Sigma^+$ ground state configuration at a bond length of 2.535 a.u. [136]. We have used a single-determinant wave function containing Slater-type orbitals generated with the ADF package [137], with which we obtain a reference DMC energy of $-15.24603(4)$ a.u. Our results are given in Table 4.8.

The addition of $N_{1,2}$ to $J = N_{2,0} + N_{1,1} + N_{2,1}$ recovers 11% of the remaining DMC correlation energy. We find no significant gain from adding either an $N_{2,2}$ term or an $N_{3,1}$ term to $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2}$, possibly due to the large number of parameters that needed to be optimized.

4.4.3.2   N$_2$ molecule

We have studied the $^1\Sigma_g^+$ ground state of the N$_2$ molecule at the experimental bond length of 2.074 a.u. [136] HF orbitals were generated in a Slater-type basis using the ADF package [137]. Our VMC results for different Jastrow factors are given in Table 4.10 along with relevant reference energies.

Adding an $N_{1,2}$ term to $J = N_{2,0} + N_{1,1} + N_{2,1}$ recovers 33% of the remaining DMC correlation energy and leads to a significant reduction in the VMC variance. The sub-

TABLE 4.8: Energies (E) and VMC variances (V) for the BeH molecule using different Jastrow factors. We have used a bond length of $r_{\text{BeH}} = 2.535$ a.u. [136].

	N_p	E (a.u.)	V (a.u.)	f_{CE} (%)	f_{DCE} (%)
UHF limit from [138][a]		-15.1536		0	0
UHF (ADF)		-15.1535		-0.1057	-0.1082
$N_{2,0}$	27	$-15.1775(5)$	$0.5681(7)$	$25.3(5)$	$25.9(5)$
$N_{2,0}+N_{1,1}$	63	$-15.2199(2)$	$0.2404(2)$	$70.1(2)$	$71.7(2)$
$N_{2,0}+N_{1,1}+N_{1,2}$	84	$-15.2235(2)$	$0.2343(3)$	$73.9(2)$	$75.6(2)$
$N_{2,0}+N_{1,1}+N_{2,1}$	257	$-15.24054(8)$	$0.0417(3)$	$91.90(8)$	$94.1(1)$
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$	278	$-15.24116(7)$	$0.0392(1)$	$92.56(7)$	$94.73(9)$
VMC from Ref. [139][a]		$-15.212(1)$		$62(1)$	$63(1)$
VMC from Ref. [140]		$-15.228(1)$		$79(1)$	$80(1)$
DMC		$-15.24603(4)$		$97.71(4)$	$100.00(6)$
Exact from Refs. [138][a] and [136]		-15.2482		100	100

[a] Ref. [138] used $r_{\text{BeH}} = 2.537$ a.u. and Ref. [139] used $r_{\text{BeH}} = 2.538$ a.u. We do not expect that these small differences in bond length will affect the comparison between energies significantly.

TABLE 4.9: e–e and e–n expansion orders (p and q, respectively) used for the different Jastrow factor terms in the BeH molecule.

	$N_{2,0}$	$N_{1,1}$	$N_{1,2}$	$N_{2,1}$
p	9	–	–	4
q	–	9	4	4

sequent addition of $N_{2,2}$ provides a reduction in the VMC energy of 13% of the remaining DMC correlation energy. We have tested adding $N_{3,0}$, $N_{3,1}$, and $N_{4,0}$ terms to $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{2,2}$, but neither of these yield any improvements in the VMC energy.

The anisotropy of this system is expected to be captured by terms containing e–n functions that treat the bond as a special direction. We have aligned the z-axis of our reference frame along the N–N bond in our calculations, and $A_{1,1}^z$ is then the simplest explicitly anisotropic term that reflects the geometry of the system. The $A_{1,1}^x$ and $A_{1,1}^y$ terms must be zero by symmetry and we have therefore not used them. There are five spherical harmonics with $l = 2$, which are respectively proportional to xy, xz, yz, $x^2 - y^2$, and $-x^2 - y^2 + 2z^2$. We find that only the last one of these, which we refer to as z^2, has a significant effect on the VMC energy.

The VMC energy with $J = N_{2,0} + N_{1,1} + N_{2,1} + A_{1,1}^z$ is within statistical uncertainty of that with $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2}$, but the former Jastrow factor contains about a third fewer parameters than the latter. The combination of the $N_{1,2}$ and $A_{1,1}^z$ terms into $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2} + A_{1,1}^z$ does not improve the VMC energy compared with the other two Jastrow factors. These results suggest that the terms $N_{1,2}$ and $A_{1,1}^z$ play similar roles in the wave function, which we find reasonable since $N_{1,2}$, although constructed from isotropic basis functions, contains the right variables to capture the symmetry of the molecule in much the same way as $A_{1,1}^z$ does. We have plotted the $A_{1,1}^z$ term for $J = N_{2,0} + N_{1,1} + N_{2,1} + A_{1,1}^z$ and the $N_{1,2}$ term for $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2}$ in Fig. 4.2, where the similarity between the terms can be seen. The value of the $N_{1,2}$ term is roughly the same as that of $A_{1,1}^z$ offset by a positive amount, and this shift is likely to be compensated for by the other Jastrow factor terms. Both terms increase the value of the wave function in the outer region of the molecule with respect to that in the bond region.

We have added different combinations of anisotropic terms to $J = N_{2,0} + N_{1,1} + N_{2,1}$. The e–e–n $A_{2,1}^z$ term retrieves less correlation energy than the e–n $A_{1,1}^z$ term. The $A_{1,2}^z$ term does not improve the $N_{2,0} + N_{1,1} + N_{1,2}$ Jastrow factor and it was not considered further. Combining terms with spherical harmonics of $l = 1$ and $l = 2$ improves the VMC energy significantly with respect to using $l = 1$ only. The anisotropic Jastrow factor $J = N_{2,0} + N_{1,1} + N_{2,1} + A_{1,1}^z + A_{1,1}^{z^2} + A_{2,1}^z + A_{2,1}^{z^2}$, which contains up to e–e–n correlations and has 191 optimizable parameters, recovers 93.3(1)% of the DMC correlation energy.

TABLE 4.10: VMC energies (E) and variances (V) for the N_2 molecule using different Jastrow factors, including explicitly anisotropic terms. We have used a bond length of r_{NN} = 2.074 a.u. [136].

	N_p	E (a.u.)	V (a.u.)	f_{CE} (%)	f_{DCE} (%)
HF limit from Ref. [138]		−108.9929		0	0
HF (ADF)		−108.9917		−0.2185	−0.2339(3)
$N_{2,0}$	18	−109.102(1)	5.275(4)	19.9(2)	21.3(2)
$N_{2,0}+N_{1,1}$	27	−109.3739(6)	3.681(3)	69.4(1)	74.3(2)
$N_{2,0}+N_{1,1}+N_{1,2}$	49	−109.3796(6)	3.595(2)	70.4(1)	75.4(2)
$N_{2,0}+N_{1,1}+N_{2,1}$	80	−109.4441(4)	1.667(2)	82.16(7)	87.9(1)
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$	102	−109.4644(4)	1.149(2)	85.85(7)	91.9(1)
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}+N_{2,2}$	219	−109.4697(4)	1.088(3)	86.82(7)	92.9(1)
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}+N_{2,2}+N_{3,0}$	260	−109.4702(3)	1.083(2)	86.91(5)	93.0(1)
$N_{2,0}+N_{1,1}+A^z_{1,1}$	36	−109.3770(6)	3.670(2)	69.9(1)	74.9(2)
$N_{2,0}+N_{1,1}+N_{2,1}+A^z_{1,1}$	89	−109.4660(3)	1.116(2)	86.14(5)	92.2(1)
$N_{2,0}+N_{1,1}+N_{2,1}+A^z_{1,1}+A^{z^2}_{1,1}$	97	−109.4669(3)	1.073(2)	86.31(5)	92.4(1)
$N_{2,0}+N_{1,1}+N_{2,1}+A^z_{1,1}+A^z_{2,1}$	142	−109.4707(3)	1.072(2)	87.00(5)	93.1(1)
$N_{2,0}+N_{1,1}+N_{2,1}+A^z_{1,1}+A^{z^2}_{1,1}+A^z_{2,1}+A^{z^2}_{2,1}$	191	−109.4714(3)	1.036(4)	87.13(5)	93.3(1)
VMC (SD) from Ref. [135][a]		−109.4520(5)		83.59(9)	89.5(2)
DMC		−109.5060(7)		93.4(1)	100.0(2)
Exact from Ref. [138]		−109.5421		100	100

[a] For r_{NN} = 2.075 a.u. We do not expect that this small difference in bond length will affect the comparison between energies significantly.
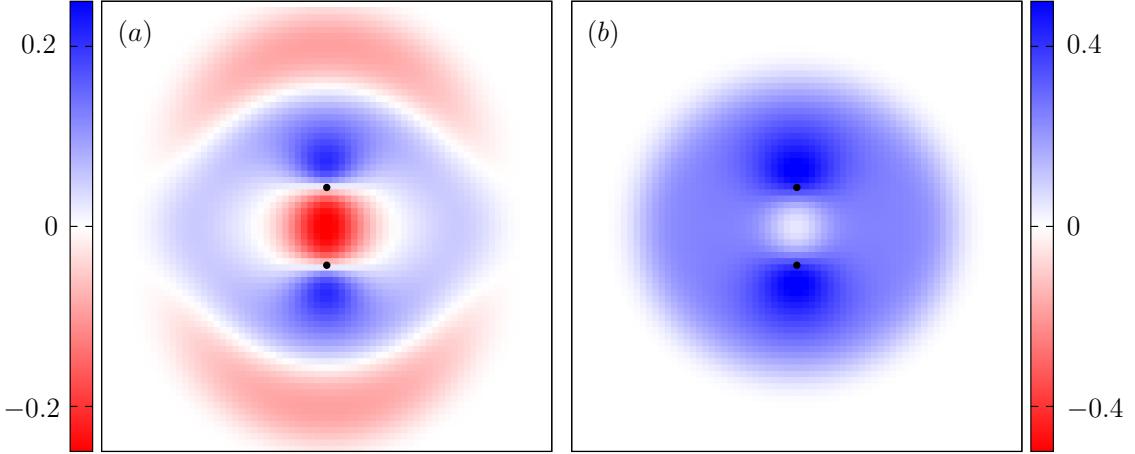
FIGURE 4.2: Plots of the (a) $A_{1,1}^z$ term and (b) $N_{1,2}$ term for N_2 as a function of the position of an electron in a 12 a.u. $\times$ 12 a.u. plane containing the nuclei, indicated by black circles.

This proportion is greater than the 93.0(1)% retrieved by our best isotropic Jastrow factor $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2} + N_{2,2} + N_{3,0}$, which includes more costly e–e–n–n and e–e–e correlations and contains 260 optimizable parameters. We conclude that anisotropic functions are an important tool in the construction of compact Jastrow factors for strongly anisotropic systems.

Toulouse and Umrigar obtained 90% of the DMC correlation energy with a single-determinant wave function [135], and with our best Jastrow factor we retrieve 93% of the DMC correlation energy. We have also optimized a single-determinant backflow wave function with our best Jastrow factor and we obtain a VMC energy of $-109.4820(6)$ a.u. (89% of the correlation energy), which is of similar accuracy to the multi-determinant VMC energy of $-109.4851(3)$ a.u. (89.6% of the correlation energy) obtained by Toulouse and Umrigar.

TABLE 4.11: e–e and e–n expansion orders (p and q, respectively) used for the different Jastrow factor terms in the N_2 molecule.

	$N_{2,0}$	$N_{1,1}$	$N_{1,2}$	$N_{2,1}$	$N_{3,0}$	$N_{2,2}$	$A_{1,1}^{s.h.}$	$A_{2,1}^{s.h.}$
p	9	–	–	4	5	5	–	4
q	–	9	7	4	–	3	9	4

4.4.3.3 H_2O MOLECULE

Single-particle spin-unrestricted HF orbitals for the $^1A_1$ ground state of H_2O were generated using the CRYSTAL Gaussian basis set code [141]. The basis set for O contains 14 s-, 9 p-, and 4 d-functions, and that for H contains 8 s-, 4 p-, and 3 d-functions. Electron-nucleus cusps have been added using the scheme of Ma $et~al.$ [142]. We have simulated a water molecule with a bond length of $r_{OH} = 1.8088$ a.u. and a bond angle

of $\angle_{\mathrm{HOH}} = 104.52°$ [143]. Our VMC results for different Jastrow factors are given in Table 4.12 along with relevant reference energies.

Adding an $N_{1,2}$ term to $J = N_{2,0} + N_{1,1} + N_{2,1}$ gives only a very small improvement for H_2O, compared with the more substantial improvements obtained with this term for BeH and N_2. The $N_{1,2}$ term acts as a correction to the single-electron orbitals, and we believe that it is unimportant in this case because we have used very accurate HF orbitals, whereas the single-electron orbitals used for BeH and N_2 are considerably less accurate. We find additional small improvements to the energy of H_2O from adding $N_{3,0}$ and $N_{3,1}$ terms to $J = N_{2,0} + N_{1,1} + N_{2,1}$.

Clark *et al.* obtained 92% of the DMC correlation energy with a single-determinant wave function in Ref. [144], and with our best Jastrow factor we recover 95.5% of the DMC correlation energy.

TABLE 4.12: Energies (E) and VMC variances (V) for the H_2O molecule using different Jastrow factors. We have used a bond length of $r_{\mathrm{OH}} = 1.8088$ a.u. and a bond angle of $\angle_{\mathrm{HOH}} = 104.52°$ [143].

	N_p	E (a.u.)	V (a.u.)	f_{CE} (%)	f_{DCE} (%)
HF limit from Ref. [138]		-76.0672		0	0
UHF (CRYSTAL)		-76.0667		-0.1348	-0.1407
$N_{2,0}$	18	$-76.1640(6)$	$3.603(7)$	$26.1(2)$	$27.2(2)$
$N_{2,0}+N_{1,1}$	36	$-76.3368(3)$	$3.066(3)$	$72.71(8)$	$75.86(9)$
$N_{2,0}+N_{1,1}+N_{1,2}$	94	$-76.3373(3)$	$3.051(6)$	$72.84(8)$	$76.00(9)$
$N_{2,0}+N_{1,1}+N_{2,1}$	266	$-76.4030(2)$	$0.87(1)$	$90.56(5)$	$94.49(6)$
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$	325	$-76.4035(2)$	$0.812(4)$	$90.70(5)$	$94.63(6)$
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}+N_{3,0}$	410	$-76.4053(2)$	$0.829(5)$	$91.18(5)$	$95.13(6)$
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}+N_{3,1}$	741	$-76.4068(2)$	$0.794(6)$	$91.59(5)$	$95.55(6)$
VMC from Ref. [143]		$-76.3773(2)$		$83.63(5)$	$87.25(6)$
VMC from Ref. [145]		$-76.3803(4)$		$84.4(1)$	$88.1(1)$
VMC from Ref. [144]		$-76.3938(4)$		$88.1(1)$	$91.9(1)$
DMC		$-76.4226(1)$		$95.85(3)$	$100.00(4)$
Exact from Ref. [143]		-76.438		100	

TABLE 4.13: e–e and e–n expansion orders (p and q, respectively) used for the different Jastrow factor terms in the H_2O molecule.

	$N_{2,0}$	$N_{1,1}$	$N_{1,2}$	$N_{2,1}$	$N_{3,0}$	$N_{3,1}$
p	9	–	–	5	5	3
q	–	9	7	5	–	3

4.4.3.4 H_2 SINGLET

We studied the $^1\Sigma_g^+$ singlet spin ground state at the equilibrium bond length $r_{\mathrm{HH}} = 1.4011$ a.u. [136] using cusp-corrected [142] HF orbitals. These orbitals were generated

by optimizing the coefficients and exponents of 13 s-, 6 p- and 4 d-functions using the CRYSTAL Gaussian basis set code [141]. Our results are presented in Table 4.14.
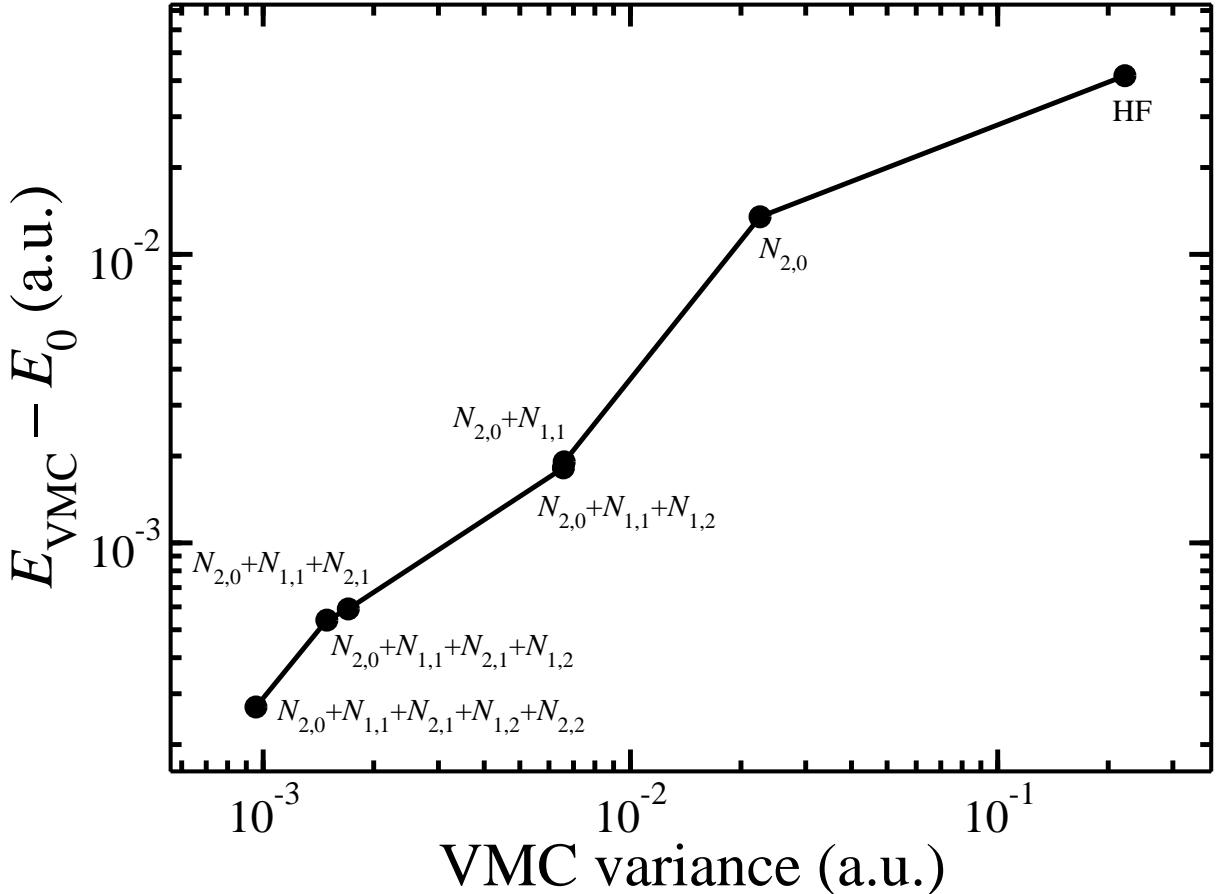
FIGURE 4.3: Difference between the VMC and exact energy against the VMC variance for the H_2 singlet spin ground state using different Jastrow factors.

The $N_{2,0}$ term was able to recover a larger proportion of correlation energy in VMC for the H_2 singlet state than was possible for other systems. When the $N_{1,1}$ term is included, the wave function was able to recover over 95% of the correlation energy. This reflects the simple electronic structure of the system.

Upon addition of the $N_{2,1}$ term to $J = N_{2,0} + N_{1,1}$, the $N_{2,0}$ term e–e cut-off increases from 6.3 a.u. to 7.3 a.u. while the $N_{1,1}$ e–n cut-off increase from 5.1 a.u. to 5.8 a.u. The $N_{2,1}$ e–n cut-off optimizes to 4.0 a.u. These changes in cut-off values demonstrate that the $N_{2,1}$ term is necessary to differentiate length scales even in systems as homogeneous as H_2. The additional variational freedom provided by the $N_{2,1}$ term yields a chemically-accurate ground-state energy for the singlet state of H_2.

The addition of the $N_{1,2}$ term to either $J = N_{2,0} + N_{1,1}$ or $J = N_{2,0} + N_{1,1} + N_{2,1}$ does not lead to a significant decrease in energy but the variance is reduced in the latter case. We conclude that deficiencies in the basis set are largely ameliorated by the homogeneous $N_{1,1}$ term.

The Jastrow factor $N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2} + N_{2,2}$ contains all possible terms that can

TABLE 4.14: Energies (E) and VMC variances (V) for the H_2 molecule in the singlet spin ground state using different Jastrow factors. We have used a bond length of r_{HH} = 1.4011 a.u. [136].

	N_p	E (a.u.)	V (a.u.)	f_{CE} (%)	f_{DCE} (%)
HF limit from Ref. [138]		−1.1336		0	0
HF (CRYSTAL)		−1.1336		0	0
$N_{2,0}$	9	−1.1611(2)	0.02253(5)	67.3(5)	67.4(5)
$N_{2,0}+N_{1,1}$	18	−1.17269(7)	0.00660(1)	95.6(2)	95.8(2)
$N_{2,0}+N_{1,1}+N_{1,2}$	34	−1.17278(8)	0.00657(1)	95.9(2)	96.0(2)
$N_{2,0}+N_{1,1}+N_{2,1}$	76	−1.17401(4)	0.001705(3)	98.9(1)	99.0(1)
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$	92	−1.17406(3)	0.001491(3)	98.98(7)	99.1(1)
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}+N_{2,2}$	160	−1.17433(2)	0.000956(2)	99.64(5)	99.8(1)
VMC from Ref. [128][a]		−1.17447568(21)	0.0000011	99.9994(5)	100.1(1)
DMC		−1.17442(4)		99.9(1)	100.0(1)
Exact from Ref. [146]		−1.174475931399(1)		100.0	

[a] For r_{HH} = 1.4 a.u. We do not expect that this small difference in bond length will affect the comparison between energies significantly. This calculation critically differs from ours in that it does not enforce the cusp conditions.

be constructed for this system and gives a VMC energy within error bars of the DMC energy, which is exact for this system. We conclude that our terms are well-parametrized and account for all the variational freedom needed for a Jastrow factor for the H_2 singlet state.

TABLE 4.15: e–e and e–n expansion orders (p and q, respectively) used for the different Jastrow factor terms in the H_2 singlet.

	$N_{2,0}$	$N_{1,1}$	$N_{1,2}$	$N_{2,1}$	$N_{2,2}$
p	9	–	–	5	4
q	–	9	6	5	3

4.4.3.5 H_2 TRIPLET

The energy of the first triplet spin excited state ($^3\Sigma_u^+$) of H_2 has a very shallow minimum corresponding to a large bond length of nearly 8 a.u. Although the exchange interaction falls exponentially with increasing inter-nuclear separation, Kolos and Wolniewicz found that it contributed significantly to the energy even at the large distance of 10 a.u. [147]. The strong interplay between the attractive dispersion forces and the repulsive exchange interaction requires that both be accounted for to afford an accurate description of the triplet state. This makes the system appealing for studying the construction of four-body Jastrow factor terms to describe van der Waals-like interactions.

We used numerical HF orbitals tabulated on an elliptical grid obtained from the 2DHF package [148] that were kindly generated by John Trail. HF theory predicts no binding for the triplet state at any separation, and therefore any binding that occurs in VMC can be attributed to the Jastrow factor. Unlike the singlet state, the nodal surface of this state is not determined by symmetry and therefore DMC does not give the exact energy.

We have studied the H_2 molecule in the triplet spin state at the inter-nuclear distance of 7.8358 a.u. This separation and the corresponding reference energy of -1.0000208957 a.u. were found by fitting a quadratic function to the data of Staszewska and Wolniewicz [149]. In a preliminary study, we studied the molecule at a variety of other inter-nuclear distances in addition to the equilibrium distance, including the singlet spin state equilibrium distance 1.401 a.u., 2.0 a.u., 4.0 a.u. and 6.0 a.u.

Equilibrium inter-nuclear distance

Previous QMC calculations on H_2 at different inter-atomic distances have used Jastrow factors with up to four-body correlations where the cusp conditions were not enforced [127, 128], instead relying on the variance minimization method to find parameter values that approximately satisfy the cusp conditions. This was found to be advantageous for this system because the additional variational freedom yielded a better description in

VMC than when the cusp conditions were obeyed exactly [129]. The violation of the cusp conditions is potentially catastrophic in DMC calculations, but these studies have restricted the use of such terms to VMC.

We have optimized Jastrow factors consisting of the single e–e–n–n terms $V_{2,2}$, $F_{2,2}^{b=1}$, and $B_{2,2}$ (see Table 4.1) at several expansion orders, where no constraints are enforced at e–e or e–n coalescence points. We have used variance minimization for these Jastrow factors as we found that it produces better results than energy minimization. The results for the single-term Jastrow factors are given in Table 4.16. We have also optimized Jastrow factors consisting of different sums of terms which satisfy the cusp conditions using energy minimization. The results are given in Table 4.17 and are shown graphically in Fig. 4.4.

TABLE 4.16: Energies (E) and VMC variances (V) for H_2 in the triplet spin state at a bond length of r_{HH} = 7.8358 a.u. using different cusp-violating single-term Jastrow factors.

	p	q	N_p	E (a.u.)	V (a.u.)	f_{CE} (%)	f_{DCE} (%)
HF limit				-0.9999828277		0	0
$V_{2,2}$	0	4	11	$-1.0000045(4)$	$0.0000205(1)$	57(1)	57(1)
	3	3	19	$-1.0000100(3)$	$0.0000153(2)$	71.4(8)	71.7(8)
	4	3	25	$-1.0000130(3)$	$0.000013(2)$	79.3(8)	79.6(8)
	0	7	29	$-1.0000090(3)$	$0.0000152(1)$	68.8(8)	69.1(8)
	3	4	31	$-1.0000139(3)$	$0.0000115(1)$	81.6(8)	82.0(8)
	4	4	41	$-1.0000154(2)$	$0.0000083(1)$	85.6(5)	86.0(6)
	3	5	46	$-1.0000157(2)$	$0.00000789(6)$	86.4(5)	86.8(6)
	4	5	61	$-1.0000166(2)$	$0.0000066(1)$	88.7(5)	89.2(6)
	6	5	91	$-1.0000175(2)$	$0.0000060(2)$	91.1(5)	91.5(6)
	5	6	106	$-1.0000178(2)$	$0.000008(1)$	91.9(5)	92.3(6)
$B_{2,2}$	4	4	10	$-1.0000086(3)$	$0.00001317(2)$	67.7(8)	68.0(8)
	5	5	21	$-1.0000179(1)$	$0.00000347(1)$	92.1(3)	92.6(4)
	6	6	43	$-1.00001966(8)$	$0.000001215(8)$	96.8(2)	97.2(3)
	7	7	79	$-1.00002012(6)$	$0.00000067(1)$	98.0(2)	98.4(3)
	8	8	139	$-1.00002028(5)$	$0.000000360(8)$	98.4(1)	98.9(3)
	9	9	229	$-1.00002039(4)$	$0.000000268(7)$	98.7(1)	99.2(3)
	10	10	364	$-1.00002045(3)$	$0.00000021(1)$	98.83(8)	99.3(3)
$F_{2,2}^{b=1}$	2	2	15	$-1.0000176(2)$	$0.00000490(3)$	91.3(5)	91.8(6)
	3	3	82	$-1.00001986(6)$	$0.000000718(7)$	97.3(2)	97.8(3)
	4	4	305	$-1.00002037(3)$	$0.000000269(5)$	98.62(8)	99.1(3)
DMC				$-1.0000207(1)$		99.5(3)	100.0(4)
Exact[a]				-1.0000208957		100	

[a] Exact energy obtained by fitting to the data of Ref. [149].

We have performed the DMC calculations using our best $B_{2,2}$ Jastrow factor and obtain a reference DMC energy of $-1.0000207(1)$ a.u. We have not encountered any statistical problems in the DMC calculations with this cusp-violating wave function. Such issues can occur when the local energy has a negative divergence in a region of configuration space
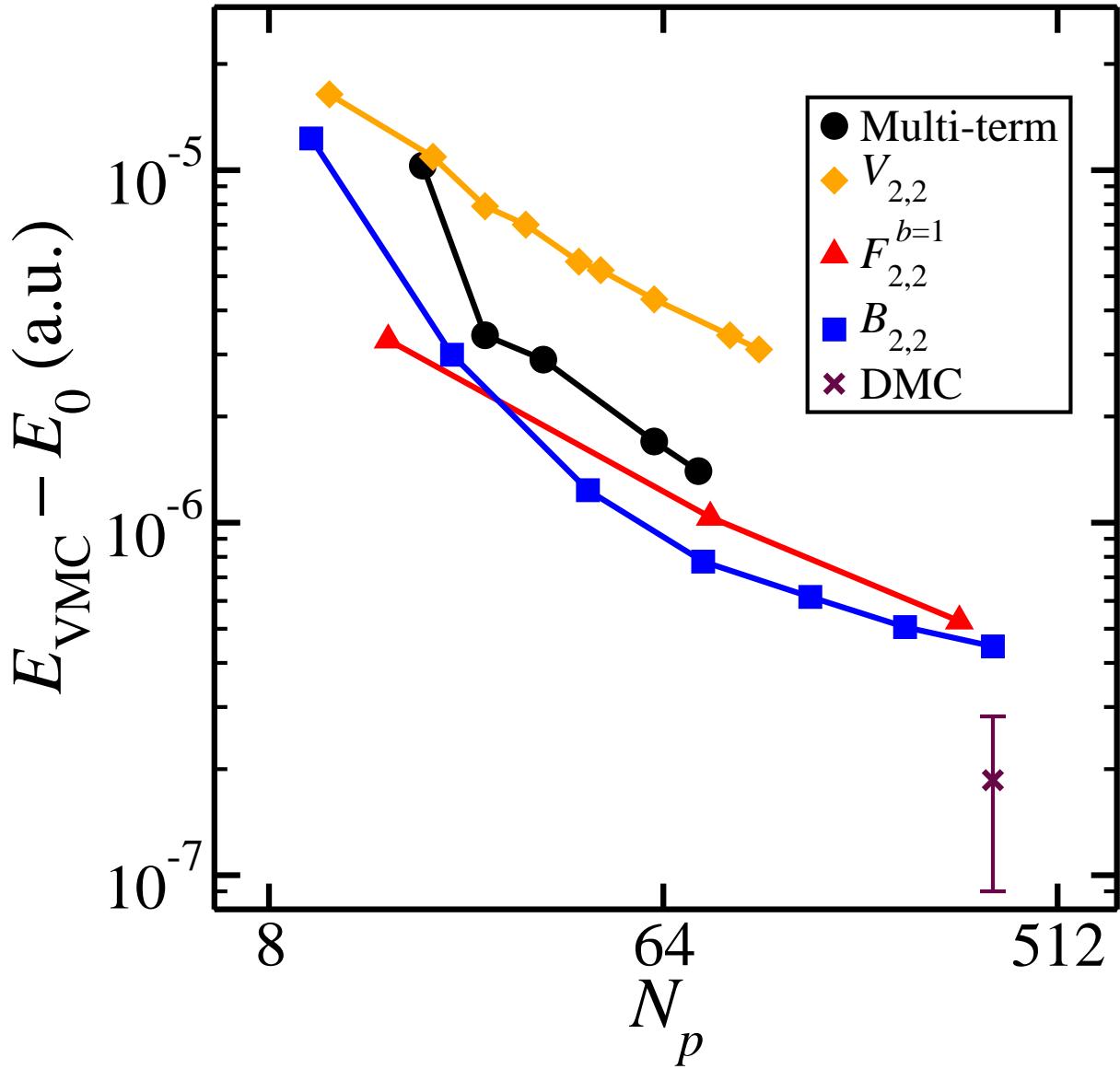
FIGURE 4.4: Difference between the VMC and exact energy against the number of wave function parameters for the H_2 triplet ground state using different Jastrow factors. Only the multi-term Jastrow factor enforces the cusp conditions. The error bars are smaller than the size of the symbol where not shown. All of the wave functions used here predict binding.

TABLE 4.17: Energies (E) and VMC variances (V) for H_2 in the triplet spin state at a bond length of $r_{HH} = 7.8358$ a.u. using different multi-term Jastrow factors.

	N_p	E (a.u.)	V (a.u.)	f_{CE} (%)	f_{DCE} (%)
HF limit		-0.9999828277		0	0
$N_{2,0}$	9	$-0.9999994(4)$	$0.00001932(1)$	44(1)	44(1)
$N_{2,0}+V_{2,2}$	25	$-1.0000175(2)$	$0.000005434(7)$	91.1(5)	91.5(6)
$N_{2,0}+N_{1,1}$	18	$-1.0000106(3)$	$0.00001074(1)$	73.0(8)	73.3(8)
$N_{2,0}+N_{1,1}+V_{2,2}$	34	$-1.0000180(2)$	$0.00000538(1)$	92.4(5)	92.8(6)
$N_{2,0}+N_{1,1}+N_{1,2}$	34	$-1.0000133(2)$	$0.00000969(1)$	80.0(5)	80.4(6)
$N_{2,0}+N_{1,1}+N_{1,2}+V_{2,2}$	50	$-1.0000180(2)$	$0.000005250(7)$	92.4(5)	92.8(6)
$N_{2,0}+N_{1,1}+N_{2,1}$	45	$-1.0000177(2)$	$0.00000476(1)$	91.6(5)	92.1(6)
$N_{2,0}+N_{1,1}+N_{2,1}+V_{2,2}$	61	$-1.0000192(1)$	$0.000003035(9)$	95.5(3)	96.0(4)
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$	61	$-1.0000186(1)$	$0.00000351(1)$	94.0(3)	94.4(4)
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}+V_{2,2}$	77	$-1.0000195(1)$	$0.000002108(6)$	96.3(3)	96.8(4)
DMC		$-1.0000207(1)$		99.5(3)	100.0(4)
Exact[a]		-1.0000208957		100	

[a] Exact energy obtained by fitting to the data of Ref. [149].

with a significant probability of being sampled. We have verified that our wave function causes a negative divergence in the local energy when an electron coalesces with a nucleus, leading us to conclude that the region of influence of this divergence is sufficiently small that statistical problems do not arise in practice.

The $F_{2,2}^{b=1}$ and $B_{2,2}$ terms only differ in that the latter uses Boys-Handy-style indexing, which yields slightly lower VMC energies than standard indexing in most cases for a fixed number of parameters. Our best $F_{2,2}^{b=1}$ and $B_{2,2}$ Jastrow factors retrieve 99% of the DMC correlation energy in VMC.

The $V_{2,2}$ term is designed to describe van der Waals correlations, and contains e–e functions which introduce other correlations. Our best $V_{2,2}$ term recovers 92% of the DMC correlation energy, offering a good description of the system without reaching the accuracy of the more generic $F_{2,2}^{b=1}$ and $B_{2,2}$ terms.

A $V_{2,2}$ term without e–e functions consists of contributions proportional to $\mathbf{r}_{iI} \cdot \mathbf{r}_{jJ}$, where the prefactors depend explicitly on r_{iI} and r_{jJ}, and implicitly on r_{IJ}. This functional form is that of a dipole-dipole interactions. Our best such $V_{2,2}$ term retrieves 69% of the DMC correlation energy, which amounts to $0.0000262(3)$ a.u., and we regard this as a measure of the pure van der Waals correlation energy of this system.

The multi-term Jastrow factors contain the usual $N_{2,0}$, $N_{1,1}$, $N_{1,2}$, and $N_{2,1}$ terms, and for each combination of these we have added a $V_{2,2}$ term without e–e functions obeying the cusp conditions to study its effect. $J = N_{2,0}$ retrieves 44% of the DMC correlation energy, and adding the $V_{2,2}$ term retrieves 85% of the remaining DMC correlation energy. The effectiveness of $V_{2,2}$ progressively drops as more terms are added, and it retrieves 43% of the remaining DMC correlation energy when added to $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2}$.

In all cases, $V_{2,2}$ is found to lower the VMC energy by a larger amount than any of the $N_{n,m}$ terms.

Our best multi-term cusp-enforcing Jastrow factor retrieves 97% of the DMC correlation energy with 77 wave-function parameters, comparable with the 98% retrieved with the cusp-violating $F_{2,2}^{b=1}$ and $B_{2,2}$ terms with a similar number of parameters. For larger systems where van der Waals interactions are important, we expect the violation of cusp conditions to cause statistical problems, and the $V_{2,2}$ term would become an effective way of improving the description of the system in a multi-term Jastrow factor.

TABLE 4.18: e–e and e–n expansion orders (p and q, respectively) used for the different Jastrow factor terms in the multi-term Jastrow factors for the H_2 triplet state at 7.8358 a.u.

	$N_{2,0}$	$N_{1,1}$	$N_{1,2}$	$N_{2,1}$	$V_{2,2}$
p	9	–	–	4	0
q	–	9	6	4	18

Various inter-nuclear distances

In our first study of the H_2 triplet state, we optimized the Jastrow factors term-by-term in two different sequences. In one set of optimizations, we started with the $N_{2,0}$ term and subsequently added the $N_{1,1}$ term, three-body terms and finally the $V_{2,2}$ term. In the second set, we began with the $V_{2,2}$ term and then proceeded to add the $N_{2,0}$, $N_{1,1}$ and three-body terms. We observe small differences in the energy obtained using the two optimization sequences. Starting with the $V_{2,2}$ term gives lower energies for $d = 6$ a.u. and 7.836 a.u. while starting with $N_{2,0}$ is preferable at shorter inter-nuclear distances. We consider this to be reasonable as the $N_{2,0}$ term becomes less important at larger distances while van der Waals contributions described by $V_{2,2}$ term become more important. The percentage of the correlation energy recovered by each Jastrow factor for each distance is given in Table 4.19. The reference energies for all distances are obtained from the data of Ref. [149].

We first considered $V_{2,2}$ terms without an expansion in e–e distances. This resulted in poorer quality wave functions, and all further investigations included a small expansion in e–e distances. As the addition of $N_{2,2}$ to $J = V_{2,2}+N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$ at $d = 7.836$ a.u. did not lead to further improvement, the $N_{2,2}$ was not included in any other calculations.

A Slater determinant comprising Gaussian orbitals was also tested. However, the variances of the energy were a factor of 25–100 times larger than those obtained with numerical orbitals. Furthermore binding was not observed for any combination of terms. This is likely due to the decay of the orbitals at large distance as e^{-r^2} instead of the correct exponential decay. This leads to substantial noise in the large-distance regions resulting in poor optimization of the Jastrow parameters.

TABLE 4.19: Correlation energy retrieved (%) for the H_2 triplet at various inter-nuclear distances d (a.u.) using different Jastrow factors.

$d =$	1.401	2.0	4.0	6.0	7.836
$N_{2,0}$	56.6(1)	53.9(1)	43.2(3)	40.8(6)	43(1)
$N_{2,0}+N_{1,1}$	85.86(7)	76.0(1)	66.3(2)	66.2(5)	75(1)
$N_{2,0}+N_{1,1}+N_{1,2}$	90.55(5)	86.02(7)	76.4(2)	73.7(5)	80.6(8)
$N_{2,0}+N_{1,1}+N_{2,1}$	91.95(5)	84.07(7)	86.6(1)	92.2(3)	94.2(5)
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$	95.56(3)	92.28(5)	90.8(1)	94.2(2)	96.9(3)
$N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}+V_{2,2}$	97.04(2)	95.25(5)	95.58(9)	97.1(2)	96.1(3)
$V_{2,2}$	59.3(1)	59.1(1)	66.7(2)	71.7(5)	73(1)
$V_{2,2}+N_{2,0}$	81.85(7)	82.3(1)	90.7(1)	92.0(3)	91.6(5)
$V_{2,2}+N_{2,0}+N_{1,1}$	91.01(5)	86.74(7)	90.7(1)	92.4(3)	91.3(5)
$V_{2,2}+N_{2,0}+N_{1,1}+N_{1,2}$	95.60(3)	92.33(7)	91.1(1)	93.8(2)	92.4(5)
$V_{2,2}+N_{2,0}+N_{1,1}+N_{2,1}$	95.53(3)	92.96(5)	94.5(1)	97.1(2)	96.1(3)
$V_{2,2}+N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$	97.04(2)	95.20(5)	95.0(1)	97.9(1)	97.1(3)
DMC	98.56(3)	99.29(5)	99.3(1)	99.8(2)	99.9(4)

The cusp conditions are not satisfied by a Jastrow factor consisting of only the $V_{2,2}$ term. The addition of the $N_{2,0}$ term, which satisfies the cusp conditions, reduces the variance by a factor of 3–5 for all inter-atomic distances.

The $N_{1,2}$ term is found to give a non-negligible improvement at all distances, particularly at intermediate distances for $J = N_{2,0}+N_{1,1}+N_{1,2}$. In the presence of the $V_{2,2}$ term, both the $N_{1,1}$ and $N_{1,2}$ terms have a greater impact at $d = 1.401$ a.u. and 2 a.u. than at larger distances. The effect of the $V_{2,2}$ term on the charge density is small enough that the $N_{1,1}$ and $N_{1,2}$ basis-set correction terms together retrieve only 0.4–14% more correlation energy. On the other hand, these terms allow $J = N_{2,0}+N_{1,1}+N_{1,2}$ to recover 33–37% more correlation energy than $J = N_{2,0}$.

We recover over 95% of the correlation energy at VMC level for all distances. Datta et al. [150] are able to recover 99.951(2)% of the correlation energy at VMC level for a separation of $d = 2$ a.u. using a highly-accurate exponential Hylleraas-type form. Our equilibrium distance DMC energy is well within an error bar of the exact energy. It appears that the quality of the HF nodal surface improves with increasing inter-nuclear distance.

For a small system such as the H_2 molecule, visualizing the Jastrow factor gives insight into its evolution as higher-order terms are added. We have plotted the contribution of an electron to the Jastrow function as it is scanned across the plane of the molecule for various Jastrow factors. The nuclei, which we label A and B are separated by the equilibrium bond length of 7.836 a.u. and the second electron is fixed at a distance 1.958 a.u. from nucleus A perpendicular to the bond. The Jastrow factor augments the Slater determinant contribution to the wave function in the blue regions and diminishes it in the red regions. In all cases, the Jastrow functions decay to zero as the electron is moved far from the

molecule and the correct asymptotic behaviour of the one-electron orbitals is retained.

In Fig. 4.5, we plot the Jastrow factor as each term is sequentially added, giving the final $J = N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}+V_{2,2}$ Jastrow factor. The sequential construction of $J = V_{2,2}+N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$ is depicted in Fig. 4.6.
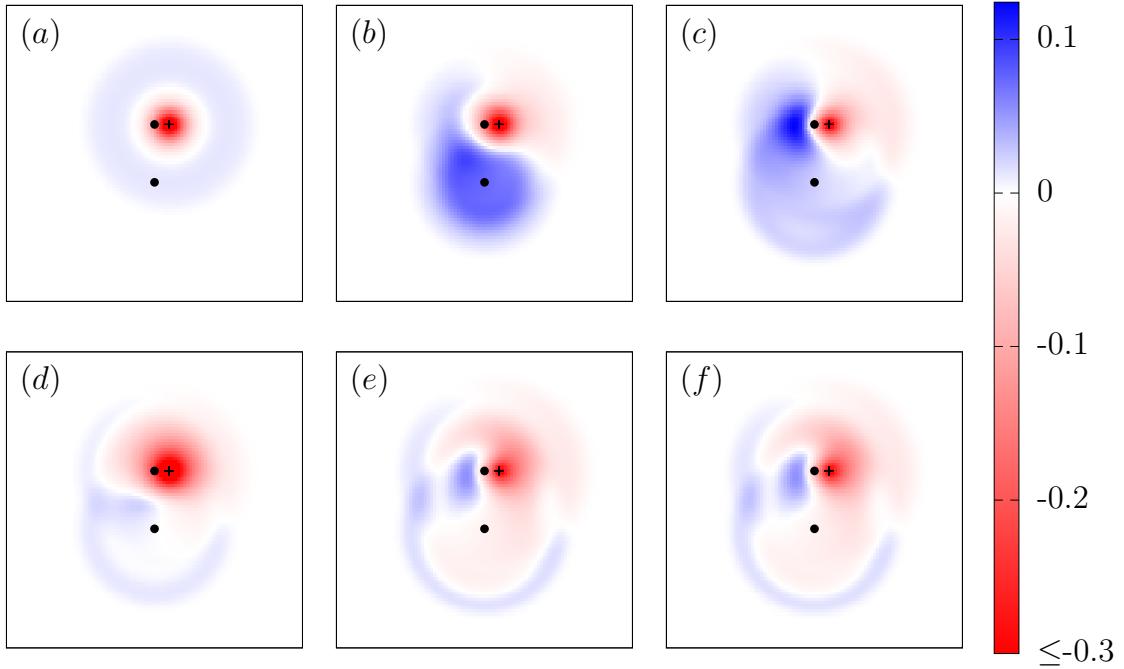
FIGURE 4.5: Plots of the (a) $N_{2,0}$, (b) $N_{2,0}+N_{1,1}$, (c) $N_{2,0}+N_{1,1}+N_{2,1}$, (d) $N_{2,0}+N_{1,1}+N_{1,2}$, (e) $N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$, (f) $N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}+V_{2,2}$ Jastrow functions. The nuclei are indicated by black circles and the fixed electron is indicated by a cross.

The isotropic $N_{2,0}$ term reduces the probability of finding the two electrons close together, as can be seen in Figs. 4.5(a) and 4.6(b). The addition of the $N_{1,1}$ term pulls the charge density away from the region of the fixed electron into the region near nucleus B in Fig. 4.5(b). By increasing the value of the wave function on the opposite side of nucleus A from the fixed electron, the $N_{2,1}$ term makes the system more ionic. This behaviour is observed for inter-nuclear distances of 4 a.u. and 6 a.u. as well as for $J = V_{2,2}+N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$, as seen in Fig. 4.6(d). The $V_{2,2}$ term does not contribute significantly when optimized last (Fig. 4.5(f)), as discussed below.

We also studied the contributions of each term to the final Jastrow factor for each optimization sequence. Plots of the term-wise contributions are given in Fig. 4.7. We see that while the magnitude of the contribution of the $N_{2,0}$, $N_{1,1}$, $N_{2,1}$ and $N_{1,2}$ terms varies, their qualitative shapes remain the same. However, the $V_{2,2}$ term (Fig. 4.7(e)) varies enormously when optimized first compared to when optimized last, resulting in qualitatively very different total Jastrow factors (Fig. 4.7(f)). The $V_{2,2}+N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$ Jastrow factor is dominated by the $V_{2,2}$ term while the $N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}+V_{2,2}$ Jastrow factor is unaffected by its presence. The difference in magnitude and structure of the $V_{2,2}$ term
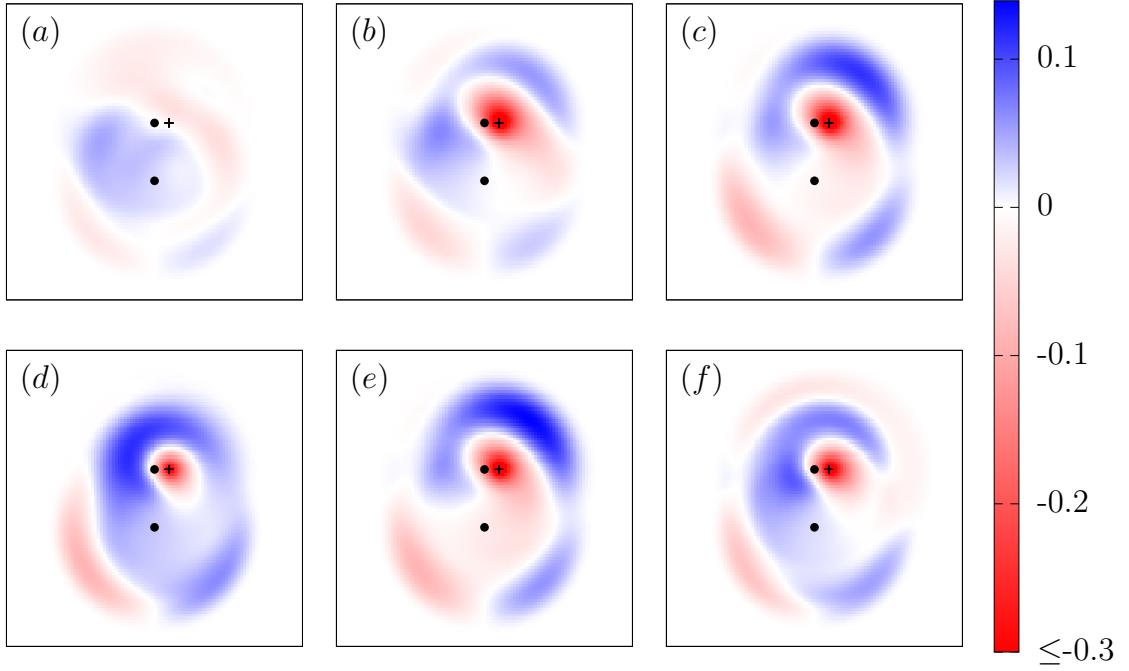
FIGURE 4.6: Plots of the (a) $V_{2,2}$, (b) $V_{2,2}+N_{2,0}$, (c) $V_{2,2}+N_{2,0}+N_{1,1}$, (d) $V_{2,2}+N_{2,0}+N_{1,1}+N_{2,1}$, (e) $V_{2,2}+N_{2,0}+N_{1,1}+N_{1,2}$, (f) $V_{2,2}+N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$ Jastrow functions. The nuclei are indicated by black circles and the fixed electron is indicated by a cross.

is highlighted in Fig. 4.8. We believe that the importance of the sequence in which terms are optimized is a result of the well-known difficulty that VMC energy optimization has in optimizing cut-off lengths[4]. During term-by-term optimization of Jastrow factors, we recommend optimizing important terms first (i.e., those that recover a larger fraction of the correlation energy), and subsequently adding less important terms.

TABLE 4.20: e–e and e–n expansion orders (p and q, respectively) used for the different Jastrow factor terms in the multi-term Jastrow factors for the H_2 molecule at various distances.

	$N_{2,0}$	$N_{1,1}$	$N_{1,2}$	$N_{2,1}$	$V_{2,2}$
p	9	–	–	5	3
q	–	9	6	5	18

4.4.4 DISCUSSION OF MOLECULAR RESULTS

In Fig. 4.9 we have plotted the fraction of the DMC correlation energy retrieved by different Jastrow factor terms for BeH, N_2, H_2O, and the H_2 singlet and triplet states.

---

[4]Shortly after this study was complete, a bug in the energy minimization routine resulting in a less than optimal minimization was found and fixed. It is possible that this dependence on optimization sequence is now weaker or even non-existent.
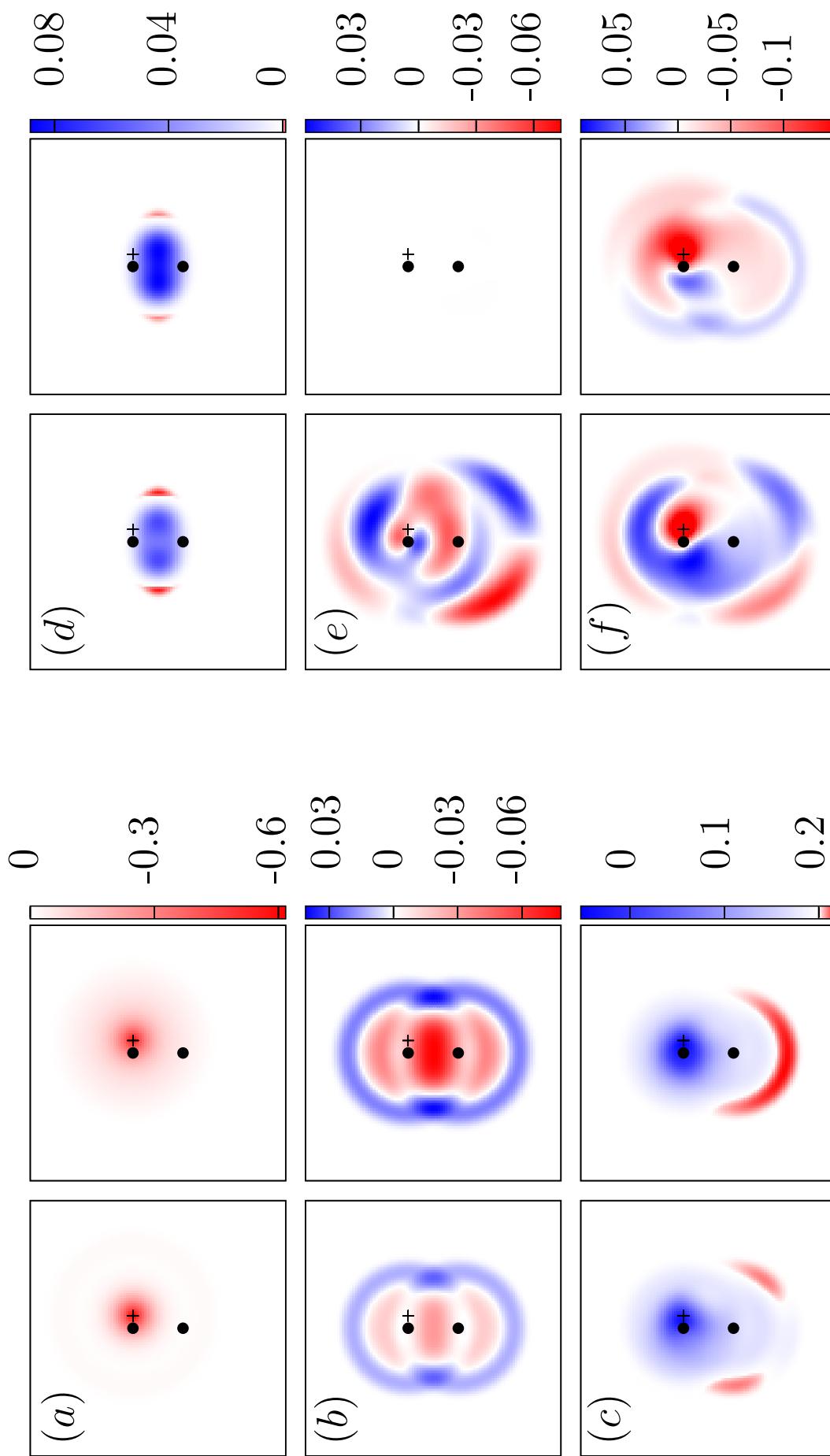
FIGURE 4.7: Plots of the (a) $N_{2,0}$ (b) $N_{1,1}$ (c) $N_{2,1}$ (d) $N_{1,2}$ (e) $N_{2,2}$ terms, and (f) the total Jastrow function. The first column corresponds to an optimization sequence of $V_{2,2}+N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$ and the second column corresponds to an optimization sequence of $N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}+V_{2,2}$. The nuclei are indicated by black circles and the fixed electron is indicated by a cross.
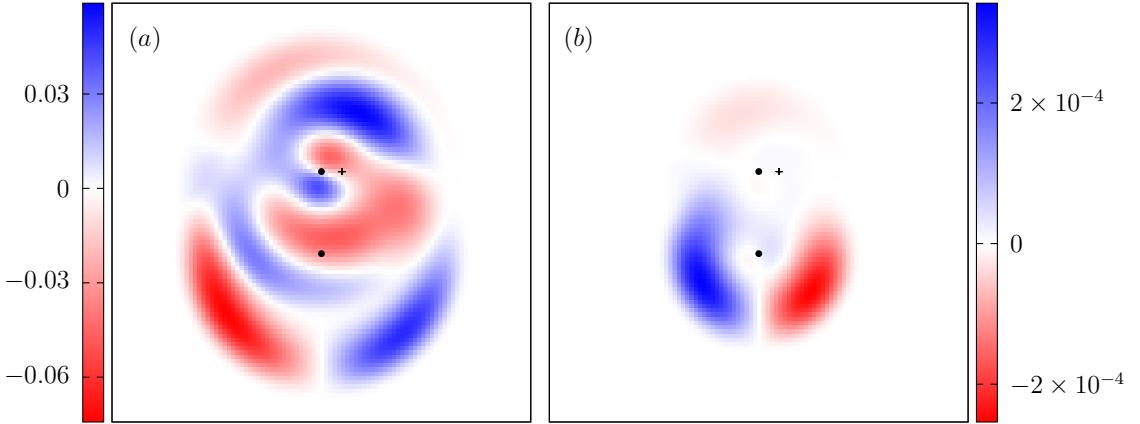
FIGURE 4.8: Plots of the $V_{2,2}$ term when optimized in the (a) $V_{2,2}+N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}$ sequence and the (b) $N_{2,0}+N_{1,1}+N_{2,1}+N_{1,2}+V_{2,2}$ sequence. The nuclei are indicated by black circles and the fixed electron is indicated by a cross.

Our purpose is to visualize the importance of different terms in different systems, and to this end we do not include anisotropic or cusp-violating terms.

The $N_{2,0}$ term represents the simplest description of electronic correlations and typically retrieves 20–25% of the DMC correlation energy. This e–e term distorts the charge density of the HF wave function, and the $N_{1,1}$ term repairs this, typically retrieving an additional 45–50% of the DMC correlation energy. In the case of the more diffuse H_2 molecule the $N_{2,0}$ and $N_{1,1}$ terms have a different relative importance. The $J = N_{2,0}+N_{1,1}$ factor recovers about 95% of the DMC correlation energy for the H_2 singlet and 70–75% of the DMC correlation energy in the other four molecules.

Like $N_{1,1}$, $N_{1,2}$ acts as a correction to the single-electron orbitals. This term provides no significant benefit in H_2O, where we have used high-quality orbitals, but it recovers 7% of the DMC correlation energy for the H_2 triplet. A visual comparison of the $N_{1,1}$ and $N_{1,2}$ terms for the H_2 triplet is given in Fig. 4.10. It is clear that the $N_{1,2}$ term is largely acting in the bond region of the molecule, where there is overlap of the isotropic $N_{1,1}$ terms centred at the two nuclei. Introduction of the $N_{1,2}$ term allows $N_{1,1}$ to be better optimized further away from the bond.

Clearly, the behaviour of the $N_{1,1}$ correction in the bond direction needs to be distinguished. This has been done in two ways in this work. Firstly, introducing a $N_{1,2}$ term recovers 4% more correlation energy for N_2, making it the most important isotropic term beyond the $N_{2,0}$, $N_{1,1}$ and $N_{2,1}$ terms. The $N_{1,2}$ term recovers between 4.5–10% of the H_2 triplet correlation energy at various distances. Secondly, anisotropic cut-offs were used to build in an explicit angular dependence to differentiate the bond direction in the N_2 molecule. Using up to $l = 2$ spherical harmonics, an additional 5.4% of the correlation energy was recovered.

The effect of $N_{1,2}$ in N_2 is noteworthy in that the energy reduction obtained by adding this term to $J = N_{2,0} + N_{1,1}$ is about a factor of four times smaller than when added to
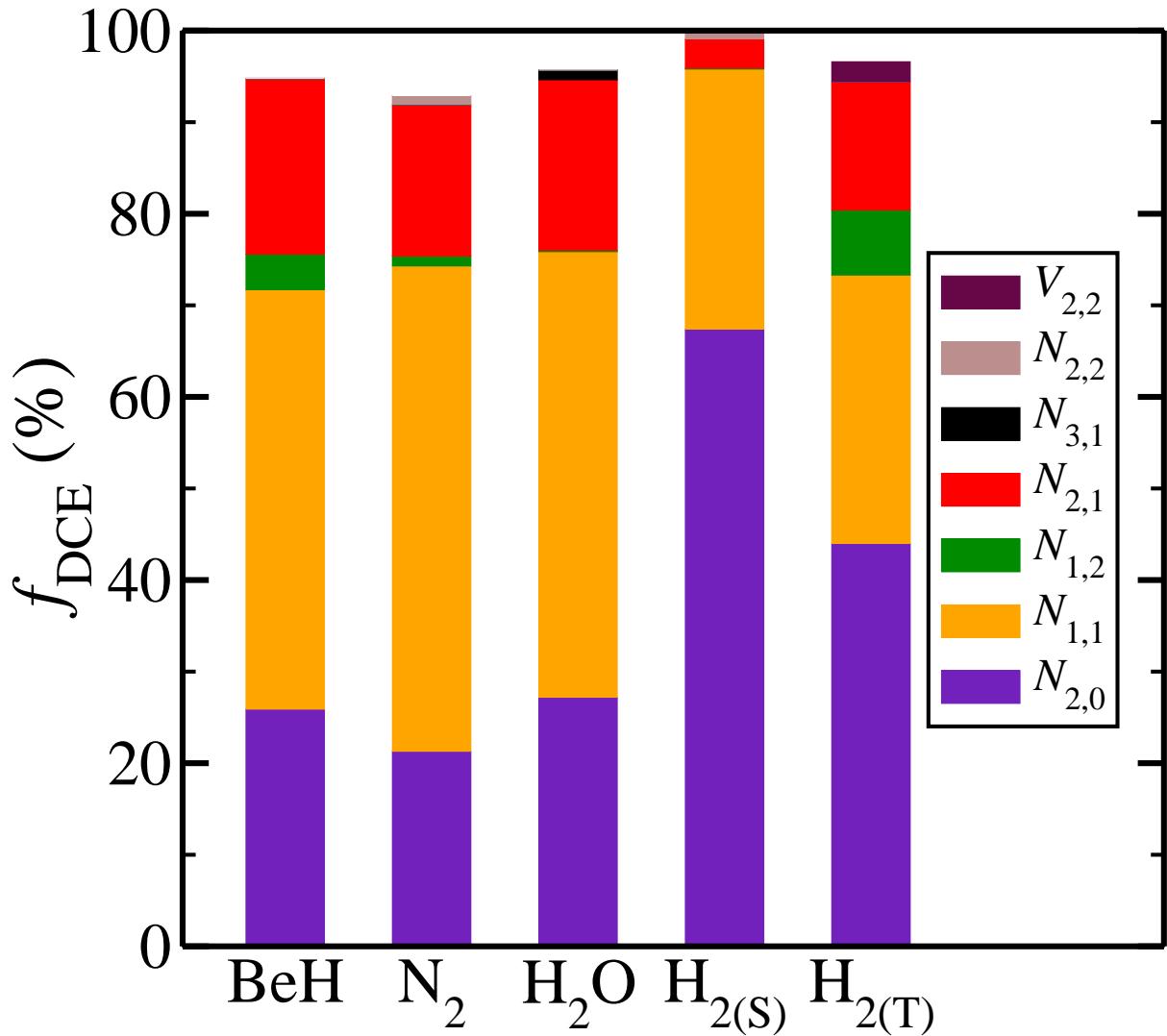
FIGURE 4.9: Fraction of the DMC correlation energy retrieved by different Jastrow factor terms for the BeH, N_2, H_2O, H_2 singlet and H_2 triplet molecules at their equilibrium geometries.
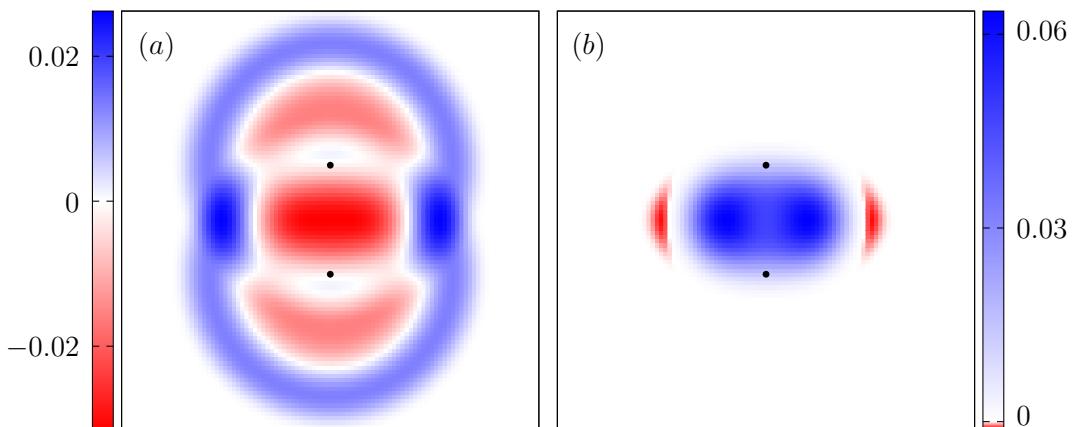
FIGURE 4.10: Plots of the (a) $N_{1,1}$ and (b) $N_{1,2}$ terms. The nuclei are indicated by black circles.

the more accurate $J = N_{2,0} + N_{1,1} + N_{2,1}$. One would expect a term to retrieve more correlation energy when added to a smaller Jastrow factor, and this is the case for $N_{1,2}$ in the other molecules. We think that the distortion in the charge density caused by $N_{2,1}$ in N_2 is such that the single-electron correction effected by $N_{1,2}$ becomes more useful in its presence.

The $N_{2,1}$ term added to $J = N_{2,0} + N_{1,1} + N_{1,2}$ captures an additional 15–20% of the DMC correlation energy for BeH, H_2O and N_2. This demonstrates the importance of the $N_{2,1}$ term in systems with different length scales. The variation of the importance of the $N_{2,1}$ term with distance is made clear by the H_2 triplet. At the short bond length of 1.401 a.u., the 21 term recovers about 6% of the correlation energy and this progressively increases to about 26% at a bond length of 6 a.u.

Higher-order terms added to $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2}$ yield significant gains in relative terms, with e–e–n–n terms retrieving 13% and 43% of the remaining DMC correlation energy remaining for N_2 and the H_2 triplet, respectively, and the e–e–e–n term recovering 17% of the remaining DMC correlation energy for H_2O.

4.4.5 SUMMARY OF RESULTS

Table 4.21 gives a comparison of the best single-determinant non-backflow VMC energies we have found in the literature with those obtained in this work.

4.5 CONCLUSIONS

We have described a generalized Jastrow factor allowing terms that explicitly correlate the motions of n electrons with m static nuclei. These terms can be parametrized using various basis sets, including terms that involve dot products of inter-particle position vectors. We have also introduced anisotropic cut-off functions. The formalism may be applied to systems with particle types and external potentials other than electrons and Coulomb potentials.

Optimization of the wave function is one of the most human- and computer-time consuming tasks in performing QMC calculations. We have performed term-by-term optimizations to understand how different terms in the Jastrow factor contribute to the electronic description of a system, and we hope that our analysis will serve as a guideline for constructing Jastrow factors for other systems.

We have tested these terms on HEGs, atoms, and molecules. The variational freedom from the higher-order terms generally improves the quality of the wave function. It has been argued that higher-order terms can be neglected [131, 126] as the Pauli exclusion principle does not allow for more than two electrons to be close. Huang *et al.* [126] suggest that it would be more economical to improve the wave function by including a

TABLE 4.21: Best single-determinant non-backflow VMC energies (a.u.) found in the literature and those from this work, along with single-determinant DMC and exact energies for reference.

System	This work	Literature	DMC	Exact
1D HEG ($r_s = 5$ a.u., $N=19$)	−0.2040833(2)		−0.2040834(3)	−0.2040834(3)
2D HEG ($r_s = 35$ a.u., $N=42$)	−0.0276112(6)		−0.0277087(1)	
Be	−14.6522(1)	−14.64972(5)[a]	−14.65717(4)	−14.66736
B	−24.6309(2)	−24.62936(5)[a]	−24.64002(6)	−24.65391
O	−75.0381(3)	−75.0352(1)[a]	−75.0511(1)	−75.0673
BeH	−15.2412(3)	−15.228(1)[b]	−15.24603(4)	−15.2482
N$_2$	−109.4714(3)	−109.4520(5)[c]	−109.5060(7)	−109.5421
H$_2$O	−76.4068(2)	−76.3938(4)[d]	−76.4226(1)	−76.438
H$_2$ singlet ($^1\Sigma_g^+$)	−1.17433(2)	−1.17447568(21)[e]	−1.17442(4)	−1.17447593 1399(1)
H$_2$ triplet ($^3\Sigma_u^+$)	−1.00002045(3)		−1.0000207(1)	−1.0000208957

[a] Ref. [135].
[b] Ref. [140].
[c] Ref. [135] (using a slightly different bond length).
[d] Ref. [144].
[e] Ref. [128] (using a slightly different bond length).

multi-determinant wave function than by using higher-order Jastrow terms, specifically a e–e–e–n term. This is often correct but we have concentrated on using a single determinant as our primary goal was to study the Jastrow factor. Of course, our Jastrow factor can be used with other wave function forms. It would be interesting to study whether the hypothesis of Huang *et al.* extends to terms other than the $N_{3,1}$ term.

We have demonstrated the construction and application of an e–e–n–n Jastrow factor term designed to describe van der Waals interactions between atoms. This term retrieves a large fraction of the van der Waals correlation energy in tests on the triplet state of H_2 at the proton separation of minimum total energy.

We have found evidence for the importance of three-electron Jastrow terms in the low-density 1D and 2D HEGs. Improving the Jastrow factor for single-determinant backflow wave functions also leads to improvements in the DMC energy of the 2D HEG. This demonstrates the indirect effect that improving the Jastrow factor can have on improving the nodal surface, as reported in Ref. [124].

We have made efforts to obtain accurate single-determinant VMC energies for most of the systems studied, but for BeH and N_2 we deliberately used inferior one-electron basis sets to see whether we could compensate for this with one-electron Jastrow terms. We find that this goal can be achieved by including an $N_{1,2}$ Jastrow term or anisotropic e–n terms, along with the usual $N_{1,1}$ term.

In strongly inhomogeneous systems, the $N_{1,2}$ term is shown to be important in describing the bond region, allowing the $N_{1,1}$ term to correct the basis set far from the bond. It is conceivable that more compact representations can be constructed by considering bond-centred terms. This idea is motivated by the bond-centred orbitals developed by the quantum chemistry community, and would be an interesting basis for future work.

Chapter 5

ORBITAL-DEPENDENT BACKFLOW TRANSFORMATIONS

5.1 INTRODUCTION

FUNDAMENTALLY, the DMC method is limited by the fixed-node approximation that must be made to overcome the fermion sign problem. Consequently, the DMC energy is limited by the accuracy of the nodal surface of the trial wave function. The Jastrow factor is everywhere positive and cannot modify the nodes. Improving nodes therefore relies on improving the orbital component of the wave function, such as afforded by a multi-determinant expansion or pairing wave functions. Evaluating orbitals at backflow-transformed quasiparticle coordinates can also achieve this. These backflow transformations are the topic of this chapter.

Backflow transformations can be motivated as improvements to the one-electron orbitals used in a wave function consisting of a single Slater determinant. One-electron orbitals do not allow for a description of correlation, but such a wave function is able to describe exchange exactly as a result of the built-in antisymmetry. To account for the anti-parallel spin correlation hole, Wigner and Seitz [151] used a wave function where the up-spin electron orbitals depended parametrically on the positions of the down-spin electrons. The form of the wave function used by Wigner and Seitz is related to the backflow wave function introduced by Feynman [152] and Feynman and Cohen [153]. Feynman and Cohen extended classical backflow, which is related the flow of an incompressible fluid around an impurity, to excitations in pure liquid helium and the ^{4}He system with ^{3}He impurities. For these systems, backflow achieves a flow pattern that conserves the local current and increases the effective impurity mass.

The Slater wave function written as a product of up- and down-spin determinants lacks direct spin coupling as the probability of finding the up-spin electrons in a given configuration is independent of the position of the down-spin electrons and vice versa. The form of backflow transformations used in QMC calculations remedies this deficiency.

The orbitals comprising the Slater determinant

$$\Psi_{\mathrm{D}}(\mathbf{X}(\mathbf{R})) = \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_1(\mathbf{x}_2) & \cdots & \psi_1(\mathbf{x}_N) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & \cdots & \psi_2(\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(\mathbf{x}_1) & \psi_N(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix} \tag{5.1}$$

are evaluated at the backflow-transformed quasiparticle coordinates $\mathbf{x}_i$ which are a function of the positions of all electrons,

$$\mathbf{x}_i = \mathbf{r}_i + \boldsymbol{\xi}_i(\mathbf{R}). \tag{5.2}$$

The contribution to the displacement from anti-parallel spin electrons is found to be larger [69]. Vitiello *et al.* [154] compared the effect of backflow transformations to that of a spin-dependent Jastrow factor. They found that both give similar results, demonstrating a possible equivalence between backflow and spin-dependent correlations.

For homogeneous systems such as electron gases, the backflow function $\boldsymbol{\xi}_i(\mathbf{R})$ is taken to be a function of inter-electron separations. The presence of nuclei introduces inhomogeneity into the system which to some extent is included via electron-nucleus terms. Higher-order electron-electron-nucleus terms are also found to be particularly important for inhomogeneous systems [69]. The inhomogeneous backflow function developed by López Ríos *et al.* [69] is

$$\boldsymbol{\xi}_i(\mathbf{R}) = \sum_{j \neq i}^{N} \eta(r_{ij})\mathbf{r}_{ij} + \sum_{I}^{M} \mu(r_{iI})\mathbf{r}_{iI}$$
$$+ \sum_{j \neq i}^{N} \sum_{I}^{M} \left[\Phi^I(r_{ij}, r_{iI}, r_{jI})\mathbf{r}_{ij} + \Theta^I(r_{ij}, r_{iI}, r_{jI})\mathbf{r}_{iI} \right], \tag{5.3}$$

where η is the e–e term, μ is the e–n term and Φ^I and Θ^I are e–e–n terms.

Backflow transformations are useful in DMC calculations because they can improve the nodal surface. However, this improvement comes at a price. As the backflow-transformed position of each electron is a function of the position of all the other electrons, changing the coordinate of one electron changes the transformed coordinates of all electrons. Each orbital must then be evaluated for each electron configuration. This significantly increases the cost of QMC calculations which then scale as $\mathcal{O}(N^4)$ rather than as $\mathcal{O}(N^3)$. In practice, the added cost of including backflow transformations is lower since only particles within a cut-off distance contribute.

Backflow has also been argued to represent momentum-dependent correlation [155]. Hence it is natural to consider backflow transformations specific to orbitals representing different momentum states. This motivates the development of new backflow transfo-

mations presented here, which are able to accommodate different parametrizations of the backflow function for distinct orbitals. The Slater determinant with these orbital-dependent backflow transformations is then

$$
\Psi_{\mathrm{D}}(\mathbf{X}(\mathbf{R})) = \begin{vmatrix} \psi_1^t(\mathbf{x}_1^t) & \psi_1^t(\mathbf{x}_2^t) & \cdots & \psi_1^t(\mathbf{x}_N^t) \\ \psi_2^u(\mathbf{x}_1^u) & \psi_2^u(\mathbf{x}_2^u) & \cdots & \psi_2^u(\mathbf{x}_N^u) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N^v(\mathbf{x}_1^v) & \psi_N^v(\mathbf{x}_2^v) & \cdots & \psi_N^v(\mathbf{x}_N^v) \end{vmatrix},
\tag{5.4}
$$

where the indices t, u and v represent backflow parameter set indices and

$$
\mathbf{x}_i^n = \mathbf{r}_i + \boldsymbol{\xi}_i^n(\mathbf{R}).
\tag{5.5}
$$

For $t = u = \ldots = v = 1$, we recover orbital-independent backflow. This work is still in progress and the benefits of these orbital-dependent backflow transformations in improving the wave function are yet to be assessed. In the following sections, the required modifications to the algorithms and the rise of additional variational freedom is discussed.

5.2 IMPLEMENTATION

The main changes that need to be made to the existing structure of CASINO to support orbital-dependent backflow transformations relate to the way in which orbitals are indexed, evaluated and used. These changes in orbital management also affect the evaluation of the kinetic energy.

5.2.1 MANAGEMENT OF ORBITALS

Firstly, a list of unique orbitals is constructed. An orbital map is used to map rows of different determinants to the appropriate orbital index. Rather than updating one entire column of a Slater matrix whenever an electron is moved, the orbital index and map structure allow sections of a column to be updated. This is necessary for orbital-dependent backflow as discussed below.

The orbitals must then be classified into groups that have the same backflow transformations. Generally, they can be classified by a number of quantities. For example, plane-wave orbitals can be characterized by their k-vector and band, or by their eigenvalue. Atomic orbitals can be labelled by their principal quantum number n, angular momentum quantum number l and magnetic quantum number m. For each type of basis set, we have constructed a list of quantities that can be used to characterize the orbitals. Each unique orbital is then labelled by its characteristics. The user specifies which of these characteristics is to be used to distinguish orbitals for the purpose of orbital-dependent

backflow transformations. Orbitals sharing a characteristic value form a group and all such orbitals are assigned the same transformation index n, allowing the application of different backflow transformations $\boldsymbol{\xi}_i^n(\mathbf{R})$ to different groups of orbitals.

The orbital map and transformation index data are used to construct an orbital mask, which indicates whether a given orbital belongs to the given transformation group[1]. However, using an orbital mask requires looping over all orbitals to identify those belonging to a group. This operation is computationally inefficient in the cases where only a few orbitals belong to a given group. It can be more efficient to construct orbital ranges for each transformation that store the first and last indices of a sequence of orbitals belonging to the same group.

The construction of the Slater matrix is modified by the existence of orbital-dependent quasiparticle coordinates. An element of the Slater matrix is denoted by $\psi_{lj}^{\kappa,n} = \psi_l^{\kappa,n}(\mathbf{x}_j^n)$, where $\psi_l^{\kappa,n}$ represents a one-electron orbital with transformation index n in the lth row of the κth determinant and the quasiparticle coordinate is evaluated using the nth backflow transformation. Instead of evaluating all orbitals at each of the n sets of quasiparticle coordinates, it is more efficient to loop over the transformation index n and evaluate all the corresponding orbitals ψ^n with transformation index n, regardless of the determinants in which they appear, at the appropriate backflow-transformed electron coordinates $\mathbf{x}^n$. The orbital map is then used to update the Slater matrices appropriately for all determinants with the new orbital values.

These modifications are useful not only in the implementation of orbital-dependent backflow transformations but also for non-backflow calculations. They serve to simplify evaluation routines significantly and unify the underlying structure of wave function evaluation. A speed-up has been observed in the evaluation of certain wave function types, e.g., by avoiding repeated evaluation of orbitals for different determinants. Additionally, they lay the foundations for the integration of other more complex types of wave functions such as geminals and pfaffians into the CASINO code.

5.2.2 KINETIC ENERGY EVALUATION

The total kinetic energy of a system is the sum of the kinetic energy of all the electrons,

$$K = \sum_i^N K_i = \sum_i^N -\frac{1}{2}\Psi^{-1}\nabla_i^2\Psi. \tag{5.6}$$

For several reasons, the kinetic energy of each electron is evaluated as

$$K_i = 2T_i - |\mathbf{F_i}|^2, \tag{5.7}$$

---

[1]Here we are only interested in grouping orbitals based on their transformation index. Orbitals are also grouped by spin, and the grouping can easily be extended to other characteristics.

where

$$T_i = -\frac{1}{4}\nabla_i^2 \ln|\Psi| = -\frac{1}{4}\left[\frac{\nabla_i^2\Psi_{\mathrm{D}}}{\Psi_{\mathrm{D}}} - \left(\frac{\nabla_i\Psi_{\mathrm{D}}}{\Psi_{\mathrm{D}}}\right)^2 + \nabla_i^2 J\right] \tag{5.8}$$

and

$$\mathbf{F}_i = -\frac{1}{\sqrt{2}}\nabla_i \ln|\Psi| = -\frac{1}{\sqrt{2}}\left(\frac{\nabla_i\Psi_{\mathrm{D}}}{\Psi_{\mathrm{D}}} + \nabla_i J\right). \tag{5.9}$$

Firstly, $\langle K_i\rangle = \langle|\mathbf{F}_i|^2\rangle = \langle T_i\rangle$ in VMC. Violation of this condition is indicative of problems in a VMC calculation such as a bug in the code. Secondly, the contribution of the determinantal part of the wave function Ψ_{D} is separated from the contribution of the Jastrow factor, allowing modularization of the code. In what follows, we are only interested in calculating the derivatives of Ψ_{D} as the Jastrow factor is unaffected by backflow transformations.

The basic quantities required for calculating the contribution of the determinantal part of wave function to the local energy are $M_i^\alpha = \nabla_i^\alpha \ln|\Psi_{\mathrm{D}}|$ and $N_i = \nabla_i^2 \ln|\Psi_{\mathrm{D}}|$ which respectively appear in T_i and $\mathbf{F}_i$. In the derivations below, Greek letters α, β and γ represent Cartesian component indices, n and m are transformation indices, l and q are orbital indices and i, j and p are electron indices.

The determinantal component of a multi-determinant-backflow wave function[2] is

$$\Psi_{\mathrm{D}} = \sum_k c_k \prod_\sigma^S D_{k,\sigma}, \tag{5.10}$$

where the kth determinant is written as a product of determinants $D_{k,\sigma}$ of sets of distinguishable particles with index σ. Then,

$$M_i^\alpha = \nabla_i^\alpha \ln|\Psi_{\mathrm{D}}| = \frac{1}{\Psi_{\mathrm{D}}}\sum_k c_k \prod_\sigma D_{k,\sigma} \sum_\tau \frac{\nabla_i^\alpha D_{k,\tau}}{D_{k,\tau}} \tag{5.11}$$

and

$$N_i = \nabla_i^2 \ln|\Psi_{\mathrm{D}}| = -\left(\frac{\nabla_i\Psi_{\mathrm{D}}}{\Psi_{\mathrm{D}}}\right)^2 + \frac{\nabla_i^2\Psi_{\mathrm{D}}}{\Psi_{\mathrm{D}}} = -|\mathbf{M}_i|^2 + \frac{\nabla_i^2\Psi_{\mathrm{D}}}{\Psi_{\mathrm{D}}} \tag{5.12}$$

where

$$\nabla_i^2\Psi_{\mathrm{D}} = \sum_k c_k \prod_\sigma D_{k,\sigma}\left[\sum_\alpha\left(\sum_\tau \frac{\nabla_i^\alpha D_{k,\tau}}{D_{k,\tau}}\right)^2 - \sum_\tau\left(\frac{\nabla_i D_{k,\tau}}{D_{k,\tau}}\right)^2 + \sum_\tau \frac{\nabla_i^2 D_{k,\tau}}{D_{k,\tau}}\right]. \tag{5.13}$$

Once again, care must be taken in the evaluation of $\nabla_i^\alpha D_\kappa$ and $\nabla_i^2 D_\kappa$ as the quasiparticle positions at which the orbitals are evaluated depend on the backflow transformation index of the orbital. Note that the determinant index k and particle group index σ have

---

[2]The corresponding expressions for other types of wave functions such as geminals and pfaffians will be different.

been absorbed into a single index κ below. This index explicitly indicates the dependence of the Slater matrix elements on the determinant. Rather than summing over all orbitals l, we sum over all transformations and restrict the sum over orbitals to those that belong to the given transformation group, i.e., $\sum_l \to \sum_m \sum_{l \in m}$.

In this notation, $\nabla_i^\alpha D_\kappa$ and $\nabla_i^2 D_\kappa$ are given by

$$\nabla_i^\alpha D_\kappa = \frac{\partial D_\kappa}{\partial r_i^\alpha} = \sum_{\substack{m \\ l \in m \\ j}} \frac{\partial D_\kappa}{\partial \psi_{lj}^{\kappa,m}} \sum_\beta \frac{\partial \psi_{lj}^{\kappa,m}}{\partial x_j^{m,\beta}} \frac{\partial x_j^{m,\beta}}{\partial r_i^\alpha} \tag{5.14}$$

and

$$\begin{aligned}
\nabla_i^2 D_\kappa &= \sum_\alpha \frac{\partial}{\partial r_i^\alpha} \left[\sum_{\substack{m \\ l \in m \\ j}} \frac{\partial D_\kappa}{\partial \psi_{lj}^{\kappa,m}} \sum_\beta \frac{\partial \psi_{lj}^{\kappa,m}}{\partial x_j^{m,\beta}} \frac{\partial x_j^{m,\beta}}{\partial r_i^\alpha} \right] \\
&= \sum_{\substack{m \\ l \in m \\ j}} \frac{\partial D_\kappa}{\partial \psi_{lj}^{\kappa,m}} \sum_\beta \frac{\partial \psi_{lj}^{\kappa,m}}{\partial x_j^{m,\beta}} \sum_\alpha \frac{\partial^2 x_j^{m,\beta}}{\partial (r_i^\alpha)^2} \\
&\quad + \sum_{\substack{m \\ l \in m \\ j}} \frac{\partial D_\kappa}{\partial \psi_{lj}^{\kappa,m}} \sum_{\beta,\gamma} \frac{\partial^2 \psi_{lj}^{\kappa,m}}{\partial x_j^{m,\beta} \partial x_j^{m,\gamma}} \sum_\alpha \frac{\partial x_j^{m,\beta}}{\partial r_i^\alpha} \frac{\partial x_j^{m,\gamma}}{\partial r_i^\alpha} \\
&\quad + \sum_{\substack{m \\ l \in m \\ j}} \sum_{\substack{n \\ q \in m \\ p}} \frac{\partial^2 D_\kappa}{\partial \psi_{lj}^{\kappa,m} \partial \psi_{qp}^{\kappa,n}} \sum_{\beta,\gamma} \frac{\partial \psi_{lj}^{\kappa,m}}{\partial x_j^{m,\beta}} \frac{\partial \psi_{qp}^{\kappa,n}}{\partial x_j^{n,\gamma}} \frac{\partial x_j^{m,\beta}}{\partial r_i^\alpha} \frac{\partial x_j^{n,\gamma}}{\partial r_i^\alpha},
\end{aligned} \tag{5.15}$$

where

$$\frac{\partial^2 D_\kappa}{\partial \psi_{lj}^{\kappa,m} \partial \psi_{qp}^{\kappa,n}} = \frac{1}{D_\kappa} \left[\frac{\partial D_\kappa}{\partial \psi_{lj}^{\kappa,m}} \frac{\partial D_\kappa}{\partial \psi_{qp}^{\kappa,n}} - \frac{\partial D_\kappa}{\partial \psi_{lp}^{\kappa,m}} \frac{\partial D_\kappa}{\partial \psi_{qj}^{\kappa,n}} \right]. \tag{5.16}$$

5.3 Variational freedom

In addition to the variational freedom introduced by allowing different backflow parameter sets for different orbitals, we also have further freedom in the choice of orbitals in the Slater matrix when using orbital-dependent backflow transformations. This freedom arises because the Slater determinant is no longer invariant under linear transformation of the ψ orbital basis.

This freedom in the choice of linear combinations of orbitals is not present for wave functions using the traditional orbital-independent backflow transformations. Consider

the Slater determinant

$$D_0 = \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_1(\mathbf{x}_2) & \cdots & \psi_1(\mathbf{x}_N) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & \cdots & \psi_2(\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(\mathbf{x}_1) & \psi_N(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix}. \tag{5.17}$$

We can combine ψ_1 with some proportion c_2 of ψ_2 to give a new determinant

$$D = \begin{vmatrix} \psi_1(\mathbf{x}_1) + c_2\psi_2(\mathbf{x}_1) & \psi_1(\mathbf{x}_2) + c_2\psi_2(\mathbf{x}_2) & \cdots & \psi_1(\mathbf{x}_N) + c_2\psi_2(\mathbf{x}_N) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & \cdots & \psi_2(\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(\mathbf{x}_1) & \psi_N(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix} \tag{5.18}$$

$$= \begin{vmatrix} [\psi_1 + c_2\psi_2](\mathbf{x}_1) & [\psi_1 + c_2\psi_2](\mathbf{x}_2) & \cdots & [\psi_1 + c_2\psi_2](\mathbf{x}_N) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & \cdots & \psi_2(\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(\mathbf{x}_1) & \psi_N(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix}. \tag{5.19}$$

It is easy to show that $D = D_0$ using the properties of determinants.

The determinant is not, however, necessarily invariant when linear combinations of orbitals are used with orbital-dependent backflow. Consider

$$D_0 = \begin{vmatrix} \psi_1^t(\mathbf{x}_1^t) & \psi_1^t(\mathbf{x}_2^t) & \cdots & \psi_1^t(\mathbf{x}_N^t) \\ \psi_2^u(\mathbf{x}_1^u) & \psi_2^u(\mathbf{x}_2^u) & \cdots & \psi_2^u(\mathbf{x}_N^u) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N^v(\mathbf{x}_1^v) & \psi_N^v(\mathbf{x}_2^v) & \cdots & \psi_N^v(\mathbf{x}_N^v) \end{vmatrix}. \tag{5.20}$$

Again, we can combine ψ_1 with some proportion c_2 of ψ_2 to construct a new determinant

$$D = \begin{vmatrix} \psi_1^t(\mathbf{x}_1^t) + c_2\psi_2^u(\mathbf{x}_1^u) & \psi_1^t(\mathbf{x}_2^t) + c_2\psi_2^u(\mathbf{x}_2^u) & \cdots & \psi_1^t(\mathbf{x}_N^t) + c_2\psi_2^u(\mathbf{x}_N^u) \\ \psi_2^u(\mathbf{x}_1^u) & \psi_2^u(\mathbf{x}_2^u) & \cdots & \psi_2^u(\mathbf{x}_N^u) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N^v(\mathbf{x}_1^v) & \psi_N^v(\mathbf{x}_2^v) & \cdots & \psi_N^v(\mathbf{x}_N^v) \end{vmatrix}. \tag{5.21}$$

Unless $t = u$, the linear combination $\psi_1^t(\mathbf{x}_i^t) + c_2\psi_2^u(\mathbf{x}_i^u)$ is a function of two quasiparticle coordinates. We have the freedom to construct a new determinant D' of single-quasiparticle

orbitals by assigning the orbital $\psi_1^t + c_2\psi_2^u$ a new transformation index s:

$$D' = \begin{vmatrix} [\psi_1^t + c_2\psi_2^u]^s(\mathbf{x}_1^s) & [\psi_1^t + c_2\psi_2^u]^s(\mathbf{x}_2^s) & \cdots & [\psi_1^t + c_2\psi_2^u]^s(\mathbf{x}_N^s) \\ \psi_2^u(\mathbf{x}_1^u) & \psi_2^u(\mathbf{x}_2^u) & \cdots & \psi_2^u(\mathbf{x}_N^u) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N^v(\mathbf{x}_1^v) & \psi_N^v(\mathbf{x}_2^v) & \cdots & \psi_N^v(\mathbf{x}_N^v) \end{vmatrix}. \tag{5.22}$$

We can thus optimize the orbitals that comprise the Slater matrix in addition to the backflow functions for each of these orbitals. Generalizing to an arbitrary combination of orbitals, we can write

$$D = \begin{Vmatrix} \begin{bmatrix} 1 & c_{12} & \cdots & c_{1N} \\ c_{21} & 1 & \cdots & c_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N1} & c_{N2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \psi_1^t(\mathbf{x}_1^t) & \psi_1^t(\mathbf{x}_2^t) & \cdots & \psi_1^t(\mathbf{x}_N^t) \\ \psi_2^u(\mathbf{x}_1^u) & \psi_2^u(\mathbf{x}_2^u) & \cdots & \psi_2^u(\mathbf{x}_N^u) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N^v(\mathbf{x}_1^v) & \psi_N^v(\mathbf{x}_2^v) & \cdots & \psi_N^v(\mathbf{x}_N^v) \end{bmatrix} \end{Vmatrix} \tag{5.23}$$

where c_{ij} determines the amplitude of ψ_i in the jth transformed orbital. Each of the c_{ij} can be optimized from its initial value of 0 subject to the constraint that the resulting orbitals are linearly independent. This is equivalent to demanding that the matrix of coefficients be non-singular:

$$\begin{vmatrix} 1 & c_{12} & \cdots & c_{1N} \\ c_{21} & 1 & \cdots & c_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N1} & c_{N2} & \cdots & 1 \end{vmatrix} \neq 0. \tag{5.24}$$

This condition is checked during optimization and parameter sets that do not satisfy it are rejected.

5.4 SYSTEMS OF INTEREST

These modifications are expected to be valuable in studying systems with one-electrons orbitals of very different characters. The 3D HEG is one such system. At high densities, backflow effects are known to become more important [156]. The uniform zero-energy $k = 0$ state differs significantly from the oscillating high-energy states. Given this variation in orbital character, we expect the optimal backflow transformations for each orbital to vary as well.

Systems such as TiO_2 might also benefit from the use of orbital-dependent transformations. The localized d-orbital character of the valence Ti electrons is very different from the more diffuse character of the valence p-electron orbitals in O.

5.5 SUMMARY

Orbital-dependent backflow transformations are expected to be more suitable for obtaining an accurate trial wave function than the system-averaged transformations currently used. Primarily, they will further improve the nodal surface of the wave function and thus bring DMC energies closer to the exact energies. It would be interesting to see if a less complex parametrization of the backflow terms, namely a smaller polynomial expansion, would suffice when using orbital-dependent backflow transformations. This would help limit the cost of including these transformations and allow for better optimization. While the benefits of using other basis functions and higher-order terms are expected to be small, they will be investigated in further work.

Chapter 6

CONCLUSIONS

THE FOCUS OF THIS THESIS is the use of improved wave functions to perform highly-accurate QMC calculations of finite and extended systems.

QMC calculations of the first-row atoms Li–Ne and their singly-positively-charged ions are reported. Multi-determinant-Jastrow-backflow trial wave functions recovered more than 98% of the correlation energy at the VMC level and more than 99% of the correlation energy at the DMC level for both the atoms and ions. We obtained the first ionization potentials to chemical accuracy for all atoms. Scalar relativistic corrections to the energies, mass-polarization terms, and one- and two-electron expectation values are reported. Fits to the electron and intracule densities are also performed.

A flexible framework for constructing Jastrow factors which allows for the introduction of terms involving arbitrary numbers of particles is described. Jastrow factors including various three- and four-body terms, a four-body van der Waals-like term, and anisotropic terms are constructed. They are used in QMC calculations of the one- and two-dimensional homogeneous electron gases, the Be, B, and O atoms, the BeH, H_2O and N_2 molecules, and the singlet and triplet states of the H_2 molecule. Our optimized Jastrow factors retrieve more than 90% of the DMC correlation energy in VMC for each system studied.

Orbital-dependent backflow transformations are motivated. Their implementation in the CASINO QMC code is described. We expect orbital-dependent backflow transformations to play an important role in improving the nodal surface in systems with large variations in orbital character. Finally, some systems are suggested as candidates for testing these transformations.

References

[1] M. Born and R. Oppenheimer, Ann. Phys. **87**, 457 (1927).

[2] A. Szabó and N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory* (Dover Publications, New York, 1996).

[3] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular electronic-structure theory* (Wiley, Chichester, 2000).

[4] G. H. Booth, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009).

[5] D. Cleland, G. H. Booth, and A. Alavi, J. Chem. Phys. **132**, 041103 (2010).

[6] G. H. Booth and A. Alavi, J. Chem. Phys. **132**, 174104 (2010).

[7] D. M. Cleland, G. H. Booth, and A. Alavi, J. Chem. Phys. **134**, 024112 (2011).

[8] J. J. Shepherd, G. Booth, A. Grüneis, and A. Alavi, Phys. Rev. B **85**, 081103 (2012).

[9] J. S. Spencer, N. S. Blunt, and W. M. C. Foulkes, J. Chem. Phys. **136**, 054110 (2012).

[10] C. Hättig, W. Klopper, A. Köhn, and D. P. Tew, Chem. Rev. **112**, 4 (2011).

[11] S. F. Boys and N. C. Handy, Proc. R. Soc. London, Ser. A **309**, 209 (1969).

[12] S. F. Boys and N. C. Handy, Proc. R. Soc. London, Ser. A **310**, 43 (1969).

[13] S. F. Boys and N. C. Handy, Proc. R. Soc. London, Ser. A **310**, 63 (1969).

[14] S. F. Boys and N. C. Handy, Proc. R. Soc. London, Ser. A **311**, 309 (1969).

[15] S. Tsuneyuki, Prog. Theor. Phys. Suppl. **176**, 134 (2008).

[16] S. Ten-no, Chem. Phys. Lett. **330**, 169 (2000).

[17] N. Umezawa and S. Tsuneyuki, J. Chem. Phys. **119**, 10015 (2003).

[18] N. Umezawa, S. Tsuneyuki, T. Ohno, K. Shiraishi, and T. Chikyow, J. Chem. Phys. **122**, 224101 (2005).

[19] R. G. Parr and W. Yang, *Density-functional theory of atoms and molecules* (Oxford University Press, Oxford, 1994).

[20] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[21] M. Levy, Proc. Natl. Acad. Sci. **76**, 6062 (1979).

[22] E. H. Lieb, Int. J. Quantum Chem. **24**, 243 (1983).

[23] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

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

[25] D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).

[26] G. Senatore and N. H. March, Rev. Mod. Phys. **66**, 445 (1994).

[27] S. Baroni and S. Moroni, Phys. Rev. Lett. **82**, 4745 (1999).

[28] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

[29] M. P. Nightingale and C. J. Umrigar, *Quantum Monte Carlo methods in physics and chemistry* (Kluwer Academic Publishers, Dordrecht, 1999).

[30] B. Hammond, W. Lester, and P. Reynolds, *Monte Carlo methods in ab initio quantum chemistry* (World Scientific, Singapore, 1994).

[31] R. J. Needs, M. D. Towler, N. D. Drummond, and P. López Ríos, J. Phys.: Condens. Matter **22**, 023201 (2010).

[32] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes: the art of scientific computing* (Cambridge University Press, Cambridge, 2007).

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

[34] D. Ceperley, G. V. Chester, and M. H. Kalos, Phys. Rev. B **16**, 3081 (1977).

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

[36] R. P. Feynman, *Statistical mechanics: a set of lectures* (Addison Wesley Longman, Reading, 1998).

[37] M. Troyer and U. Wiese, Phys. Rev. Lett. **94**, 170201 (2005).

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

[39] J. B. Anderson, J. Chem. Phys. **65**, 4121 (1976).

[40] R. Grimm and R. Storer, J. Comput. Phys. **7**, 134 (1971).

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

[42] D. M. Ceperley, J. Stat. Phys. **63**, 1237 (1991).

[43] C. J. Umrigar, M. P. Nightingale, and K. J. Runge, J. Chem. Phys. **99**, 2865 (1993).

[44] D. Ceperley, M. H. Kalos, and J. L. Lebowitz, Macromolecules **14**, 1472 (1981).

[45] N. Nemec, Phys. Rev. B **81**, 035119 (2010).

[46] R. Barnett, P. Reynolds, and W. Lester Jr, J. Computat. Phys. **96**, 258 (1991).

[47] J. R. Trail and R. Maezono, J. Chem. Phys. **133**, 174120 (2010).

[48] G. Ortiz, D. M. Ceperley, and R. M. Martin, Phys. Rev. Lett. **71**, 2777 (1993).

[49] N. D. Drummond, PhD thesis, University of Cambridge, Cambridge, 2004.

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

[51] G. Weerasinghe, personal communication.

[52] P. López Ríos, personal communication.

[53] A. C. Hurley, J. Lennard-Jones, and J. A. Pople, Proc. R. Soc. London, Ser. A **220**, 446 (1953).

[54] M. Casula and S. Sorella, J. Chem. Phys. **119**, 6500 (2003).

[55] M. Marchi, S. Azadi, M. Casula, and S. Sorella, J. Chem. Phys. **131**, 154116 (2009).

[56] M. Bajdich, L. Mitas, G. Drobný, L. K. Wagner, and K. E. Schmidt, Phys. Rev. Lett. **96**, 130201 (2006).

[57] M. Bajdich, L. Mitas, L. K. Wagner, and K. E. Schmidt, Phys. Rev. B **77**, 115112 (2008).

[58] A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. E **71**, 066704 (2005).

[59] L. Mitas, E. L. Shirley, and D. M. Ceperley, J. Chem. Phys. **95**, 3467 (1991).

[60] M. Casula, Phys. Rev. B **74**, 161102 (2006).

[61] M. Casula, S. Moroni, S. Sorella, and C. Filippi, J. Chem. Phys. **132**, 154113 (2010).

[62] S. Chiesa, D. M. Ceperley, R. M. Martin, and M. Holzmann, Phys. Rev. Lett. **97**, 076404 (2006).

[63] N. D. Drummond, R. J. Needs, A. Sorouri, and W. M. C. Foulkes, Phys. Rev. B **78**, 125106 (2008).

[64] H. Flyvbjerg and H. G. Petersen, J. Chem. Phys. **91**, 461 (1989).

[65] J. R. Trail, Phys. Rev. E **77**, 016703 (2008).

[66] J. R. Trail, Phys. Rev. E **77**, 016704 (2008).

[67] C. J. Umrigar, Phys. Rev. Lett. **71**, 408 (1993).

[68] R. M. Lee, G. J. Conduit, N. Nemec, P. López Ríos, and N. D. Drummond, Phys. Rev. E **83**, 066706 (2011).

[69] P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. E **74**, 066701 (2006).

[70] M. Dewing, J. Chem. Phys. **113**, 5123 (2000).

[71] N. D. Drummond, Z. Radnai, J. R. Trail, M. D. Towler, and R. J. Needs, Phys. Rev. B **69**, 085116 (2004).

[72] N. D. Drummond and R. J. Needs, Phys. Rev. Lett. **102**, 126402 (2009).

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

[74] P. R. C. Kent, R. J. Needs, and G. Rajagopal, Phys. Rev. B **59**, 12344 (1999).

[75] N. D. Drummond and R. J. Needs, Phys. Rev. B **72**, 085124 (2005).

[76] M. Snajdr and S. M. Rothstein, J. Chem. Phys. **112**, 4935 (2000).

[77] F. J. Gálvez, E. Buendía, and A. Sarsa, J. Chem. Phys. **115**, 1166 (2001).

[78] A. Badinski and R. J. Needs, Phys. Rev. E **76**, 036707 (2007).

[79] D. Ceperley, J. Stat. Phys. **43**, 815 (1986).

[80] M. P. Nightingale and V. Melik-Alaverdian, Phys. Rev. Lett. **87**, 043401 (2001).

[81] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007).

[82] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **126**, 084102 (2007).

[83] C. Froese Fischer, G. Tachiev, G. Gaigalas, and M. R. Godefroid, Comput. Phys. Commun. **176**, 559 (2007).

[84] N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. B **70**, 235119 (2004).

[85] M. D. Brown, J. R. Trail, P. López Ríos, and R. J. Needs, J. Chem. Phys. **126**, 224110 (2007).

[86] N. D. Drummond, P. López Ríos, A. Ma, J. R. Trail, G. G. Spink, M. D. Towler, and R. J. Needs, J. Chem. Phys. **124**, 224104 (2006).

[87] M. Puchalski and K. Pachucki, Phys. Rev. A **78**, 052511 (2008).

[88] S. J. Chakravorty, S. R. Gwaltney, E. R. Davidson, F. A. Parpia, and C. Froese Fischer, Phys. Rev. A **47**, 3649 (1993).

[89] J. R. Trail, personal communication.

[90] E. Buendía, F. J. Gálvez, P. Maldonado, and A. Sarsa, J. Chem. Phys. **131**, 044115 (2009).

[91] K. Hongo, Y. Kawazoe, and H. Yasuhara, Mater. Trans. **47**, 2612 (2006).

[92] R. Prasad, N. Umezawa, D. Domin, R. Salomon-Ferrer, and W. A. Lester, J. Chem. Phys. **126**, 164109 (2007).

[93] W. Klopper, R. A. Bachorz, D. P. Tew, and C. Hättig, Phys. Rev. A **81**, 022503 (2010).

[94] F. De Proft and P. Geerlings, J. Chem. Phys. **106**, 3270 (1997).

[95] M. Ernzerhof and G. E. Scuseria, J. Chem. Phys. **110**, 5029 (1999).

[96] Q. Zhao and R. G. Parr, J. Chem. Phys. **98**, 543 (1993).

[97] Q. Zhao, R. C. Morrison, and R. G. Parr, Phys. Rev. A **50**, 2138 (1994).

[98] E. Steiner, J. Chem. Phys. **39**, 2365 (1963).

[99] M. Levy, J. P. Perdew, and V. Sahni, Phys. Rev. A **30**, 2745 (1984).

[100] R. Assaraf, M. Caffarel, and A. Scemama, Phys. Rev. E **75**, 035701 (2007).

[101] J. Katriel, Phys. Rev. A **5**, 1990 (1972).

[102] R. A. Bonham and M. Fink, *High-energy electron scattering* (Van Nostrand Reinhold, New York, 1974).

[103] R. J. Weiss, *X-ray determination of electron distributions* (Elsevier, Amsterdam, 1966).

[104] R. J. Boyd, C. Sarasola, and J. M. Ugalde, J. Phys. B **21**, 2555 (1988).

[105] A. Sarsa, F. J. Gálvez, and E. Buendía, J. Chem. Phys. **109**, 7075 (1998).

[106] J. Toulouse, R. Assaraf, and C. J. Umrigar, J. Chem. Phys. **126**, 244112 (2007).

[107] A. J. Thakkar and V. H. Smith Jr., Chem. Phys. Lett. **42**, 476 (1976).

[108] A. J. Thakkar, J. Chem. Phys. **84**, 6830 (1986).

[109] Y. Lee and R. J. Needs, unpublished.

[110] R. Assaraf and M. Caffarel, Phys. Rev. Lett. **83**, 4682 (1999).

[111] S. D. Kenny, G. Rajagopal, and R. J. Needs, Phys. Rev. A **51**, 1898 (1995).

[112] S. A. Alexander, S. Datta, and R. L. Coldwell, Phys. Rev. A **81**, 032519 (2010).

[113] F. W. King, D. G. Ballegeer, D. J. Larson, P. J. Pelzl, S. A. Scott, T. J. Prosa, and B. M. Hinaus, Phys. Rev. A **58**, 3597 (1998).

[114] Z. Yan and G. W. F. Drake, Phys. Rev. Lett. **81**, 774 (1998).

[115] C. L. Pekeris, Phys. Rev. **126**, 1470 (1962).

[116] K. Pachucki and J. Komasa, Phys. Rev. Lett. **92**, 213001 (2004).

[117] T. Koga and H. Matsuyama, J. Chem. Phys. **115**, 3984 (2001).

[118] F. W. King, J. Chem. Phys. **102**, 8053 (1995).

[119] Z. Yan and G. W. F. Drake, Phys. Rev. A **52**, 3711 (1995).

[120] A. M. Frolov, J. Chem. Phys. **124**, 224323 (2006).

[121] A. M. Frolov and D. M. Wardlaw, J. Exp. Theor. Phys. **108**, 583 (2009).

[122] F. J. Gálvez and I. Porras, Phys. Rev. A **44**, 144 (1991).

[123] A. J. Cohen, N. C. Handy, and B. O. Roos, Phys. Chem. Chem. Phys. **6**, 2928 (2004).

[124] Y. Kwon, D. M. Ceperley, and R. M. Martin, Phys. Rev. B **48**, 12037 (1993).

[125] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).

[126] C.-J. Huang, C. J. Umrigar, and M. P. Nightingale, J. Chem. Phys. **107**, 3007 (1997).

[127] S. A. Alexander and R. L. Coldwell, J. Chem. Phys. **121**, 11557 (2004).

[128] M. C. Per, S. P. Russo, and I. K. Snook, J. Chem. Phys. **130**, 134103 (2009).

[129] M. C. Per, personal communication.

[130] K. E. Schmidt and J. W. Moskowitz, J. Chem. Phys. **93**, 4172 (1990).

[131] C. Filippi and C. J. Umrigar, J. Chem. Phys. **105**, 213 (1996).

[132] P. Seth, P. López Ríos, and R. J. Needs, J. Chem. Phys. **134**, 084105 (2011).

[133] N. D. Drummond and R. J. Needs, Phys. Rev. Lett. **99**, 166401 (2007).

[134] R. M. Lee and N. D. Drummond, Phys. Rev. B **83**, 245114 (2011).

[135] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **128**, 174101 (2008).

[136] D. Feller, K. A. Peterson, and D. A. Dixon, J. Chem. Phys. **129**, 204105 (2008).

[137] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler, J. Comput. Chem. **22**, 931 (2001).

[138] D. P. ONeill and P. M. W. Gill, Mol. Phys. **103**, 763 (2005).

[139] A. Lüchow and J. B. Anderson, J. Chem. Phys. **105**, 7573 (1996).

[140] N. Nemec, M. D. Towler, and R. J. Needs, J. Chem. Phys. **132**, 034111 (2010).

[141] R. Dovesi, R. Orlando, B. Civalleri, C. Roetti, V. R. Saunders, and C. M. Zicovich-Wilson Z. Kristallogr. **220**, 571 (2005).

[142] A. Ma, M. D. Towler, N. D. Drummond, and R. J. Needs, J. Chem. Phys. **122**, 224322 (2005).

[143] I. G. Gurtubay and R. J. Needs, J. Chem. Phys. **127**, 124306 (2007).

[144] B. K. Clark, M. A. Morales, J. McMinis, J. Kim, and G. E. Scuseria, J. Chem. Phys. **135**, 244105 (2011).

[145] M. Casula, C. Attaccalite, and S. Sorella, J. Chem. Phys. **121**, 7110 (2004).

[146] J. S. Sims and S. A. Hagstrom, J. Chem. Phys. **124**, 094101 (2006).

[147] W. Koos and L. Wolniewicz, J. Chem. Phys. **43**, 2429 (1965).

[148] L. Laaksonen, P. Pyykko, and D. Sundholm, Comput. Phys. Rep. **4**, 313 (1986).

[149] G. Staszewska and L. Wolniewicz, J. Mol. Spectrosc. **198**, 416 (1999).

[150] S. Datta, S. A. Alexander, and R. L. Coldwell, Int. J. Quantum Chem. **111**, 4106 (2011).

[151] E. Wigner and F. Seitz, Phys. Rev. **46**, 509 (1934).

[152] R. P. Feynman, Phys. Rev. **94**, 262 (1954).

[153] R. P. Feynman and M. Cohen, Phys. Rev. **102**, 1189 (1956).

[154] S. A. Vitiello, K. E. Schmidt, and S. Fantoni, Phys. Rev. B **55**, 5647 (1997).

[155] V. R. Pandharipande and N. Itoh, Phys. Rev. A **8**, 2564 (1973).

[156] Y. Kwon, D. M. Ceperley, and R. M. Martin, Phys. Rev. B **58**, 6800 (1998).

Supplementary Information

Electron- and intracule-density fitting parameters

The parameters for the least-linear square fits to the binned electron and intracule densities as described in Sec. 3.4.3 are given here. The number of parameters in each case was chosen to minimize χ^2 while giving a sensible density gradient as $r \to 0$. The errors in the normalization constants for the charge and intracule density fits are of $\mathcal{O}(10^{-3})$ or smaller, except for the intracule densities for B and C, where they are of $\mathcal{O}(10^{-2})$.

TABLE 1: Parameters for fits to the up-spin electron densities for the first-row atoms to the form given by Eq. 3.4.

	Li	Be	B	C	N	O	F	Ne
A	0.34331852	9.1463943×10^{-8}	1.7628238×10^{-7}	4.4433669×10^{-5}	0.12668912	0.051119208	1.4867862×10^{-5}	$-2.2876772 \times 10^{-7}$
a_0	-1.9427389	-2.8719999	-3.5819972	-4.1504508	-4.4901201	-4.9953390	-5.4145965	-5.7370155
a_1	-10.726341	-2.7717923	-1.7318304	-1.9103767	-2.3253100	-1.4286440	-0.92797838	0.038450613
a_2	18.461698	2.4579700	0.55058404	-2.3436076	-9.9236039	-22.829694	-39.871743	-65.981848
a_3	37.955017	5.8307774	0.68213507	0.91893335	-1.7169735	-9.9818289	-24.900394	-37.234060
a_4	21.489245	3.1788571	1.8768540	6.5337572	26.271530	43.939188	60.616172	45.124145
a_5	17.387792	0.50864843	6.8372479	9.5810898	22.593231	43.706729	80.557926	163.44469
a_6	14.230806	0.39390760	7.3872297	7.3726951	7.4677907	15.466883	41.586625	12.148876
a_7	3.0099172	3.0310406	1.8612249	3.5304831	2.6247689	4.3548572	12.828971	—
a_8	3.1083264	5.0409145	2.7129058	1.7353787	1.0609296	3.1057419	5.0510611	—
a_9	3.7714280	2.9646141	1.5440288	1.7707137	1.6953027	1.3894984	3.2851832	—
a_{10}	0.77501797	0.91219253	0.83659939	1.6891232	0.59037946	1.8819974	1.8462949	—
a_{11}	2.6645536	2.4990404	—	1.3253568	—	0.25219046	1.4174319	—
a_{12}	1.5197571	2.4321647	—	1.5937562	—	—	2.0694535	—
a_{13}	0.79252263	1.2617659	—	0.66857505	—	—	0.87670115	—
a_{14}	—	0.79713683	—	—	—	—	—	—
b_2	2.4807918	0.80077965	-0.32459605	0.31562317	0.32191233	0.0019207037	-0.00023781954	0.00041155539
b_3	2.6062509	1.0074396	0.78181314	0.036392385	0.00031730409	-0.00027859323	1.2564952	0.019596266
b_4	-2.7767656	0.47358667	-0.41586349	2.2834406	4.0967284	5.8510075	-7.0913155	9.6146565
b_5	1.5193070	0.82877304	1.4517053	2.0942297	3.3112632	3.3660045	5.1597091	5.3509038
b_6	1.5191360	0.50012316	1.0514682	1.4086337	1.9359045	3.0953395	4.1411743	—
b_7	-1.0894578	-1.0542972	1.6986537	1.2086499	0.87037027	1.7181366	2.5708433	—
b_8	-1.5063730	-0.82266817	1.0393792	1.4265887	0.96909927	0.60535036	1.3786823	—
b_9	0.74284265	0.79568143	0.75651296	1.0281121	0.81835765	1.0203494	1.1711262	—
b_{10}	0.095771470	0.51865936	—	0.81562548	—	0.84309536	1.1448117	—
b_{11}	1.5228313×10^{-5}	0.72207366	—	0.95785229	—	—	1.0175128	—
b_{12}	0.68217537	0.69247333	—	0.82371561	—	—	0.73848311	—
b_{13}	—	0.88296458	—	—	—	—	—	—

TABLE 2: Parameters for fits to the down-spin electron densities for the first-row atoms to the form given by Eq. 3.4.

	Li	Be	B	C	N	O	F	Ne
A	0.0002591391	1.0724580×10^{-6}	-0.070607844	0.043638318	-0.12897792	0.022327365	-0.0046386796	$-7.5168008 \times 10^{-8}$
a_0	-1.9166883	-2.8717112	-3.5698476	-4.1325123	-4.4857100	-5.0251945	-5.3990372	-5.7440271
a_1	-6.1435146	-3.1792968	-1.4902341	-0.60917598	-2.0502598	-1.4004246	-1.9806471	1.8237549
a_2	11.836933	2.0002721	0.53704290	-2.4894699	0.27890212	-8.4141734	-24.477887	-75.944409
a_3	12.258311	7.8378329	2.2731583	-1.4931522	-2.5964052	-10.508224	-25.665812	-16.295836
a_4	4.0785796	8.0049808	1.9300711	7.9475624	11.482135	26.059275	54.154408	51.679106
a_5	-0.40304629	4.0059836	4.6531344	15.363306	29.073909	55.846529	85.490534	129.40521
a_6	2.4619689	-1.1041558	5.1086527	13.276654	22.353448	43.052745	49.553836	11.712816
a_7	0.28697434	-1.5207871	2.3767771	8.6964601	9.2967443	18.951709	22.470956	—
a_8	—	3.8679842	1.7623874	7.3710847	5.8058886	8.7213050	14.391799	—
a_9	—	6.6043443	0.24991113	6.0282823	4.1513132	6.0087073	7.0404810	—
a_{10}	—	2.5843113	—	2.9936105	1.8714768	2.6914039	1.3060337	—
a_{11}	—	0.72111555	—	1.3741130	2.1148610	1.4353010	0.81264509	—
a_{12}	—	2.5233218	—	1.7490020	0.30968492	2.2900839	1.8638744	—
a_{13}	—	1.1885500	—	0.27643302	—	2.0602868	0.80102544	—
a_{14}	—	0.58579413	—	—	—	0.92140705	—	—
b_2	-1.6711567	0.86384140	0.62282715	-0.00019262373	-0.0022995082	-0.0021502060	-0.018392763	-0.0094779950
b_3	1.7518031	1.2909391	0.15374974	0.35646386	0.053928170	-0.0012948228	-0.0019806050	-0.15077294
b_4	0.46122260	1.1325175	0.00097972942	0.87008355	1.2958830	3.3447578	6.7182738	9.4434247
b_5	0.85300888	0.25805011	1.2690095	1.7326799	3.3886270	5.0699758	5.0476887	4.3666283
b_6	0.45616710	-0.51586861	1.1975137	2.1491346	2.7546835	4.3892077	4.4214116	—
b_7	—	-1.0784542	1.1882231	2.2491542	-2.1992376	2.9116329	3.8778238	—
b_8	—	0.83945591	0.88132223	1.9057284	1.4312168	1.8121362	2.2902614	—
b_9	—	0.79020132	—	1.1759509	0.82261517	0.28489201	0.93926377	—
b_{10}	—	0.64401496	—	0.98830434	1.1047262	1.4356842	1.0387582	—
b_{11}	—	0.84839723	—	0.94002906	0.83166248	1.2586701	0.93944725	—
b_{12}	—	0.66561666	—	0.80783307	—	0.90337804	0.69196398	—
b_{13}	—	0.88560412	—	—	—	0.71883093	—	—

TABLE 3: Parameters for fits to the up-spin electron densities for the first-row ions to the form given by Eq. 3.4.

	Li$^+$	Be$^+$	B$^+$	C$^+$	N$^+$	O$^+$	F$^+$	Ne$^+$
A	0.0051328114	0.38300842	$-1.3957019 \times 10^{-6}$	0.049395230	$-7.7911356 \times 10^{-6}$	1.6871922×10^{-6}	0.15276438	$-1.4602176 \times 10^{-6}$
a_0	-1.9116696	-2.8903164	-3.5892466	-4.1594797	-4.6422800	-5.0519294	-5.3904800	-5.7369245
a_1	-1.6709439	-0.87873186	-1.2313807	-0.66296463	-1.4083440	-1.6630645	2.7021433	-2.0510963
a_2	1.3864200	0.58373693	0.31010450	-4.6691894	-12.315791	-26.124430	-58.682608	-68.044415
a_3	3.2486651	2.9777119	0.56674065	0.37731048	1.9757087	-11.131090	-11.068639	-40.366317
a_4	2.4924032	1.6039377	-0.20906535	9.4285728	20.653668	36.955166	76.05914	48.527055
a_5	0.091488300	1.3442881	2.5758995	13.498483	25.427146	48.013974	77.257803	51.648692
a_6	1.4394595	3.5078865	7.1546428	9.2552096	18.637644	27.335278	37.780643	17.857179
a_7	2.6503375	4.4337392	7.4746935	4.1598964	10.922681	11.312453	16.693537	—
a_8	0.96621062	2.5589859	3.2729268	2.7134316	6.0299491	6.1239172	8.7803926	—
a_9	1.3429974	1.1945870	0.70927534	2.3752972	3.3051909	3.5380881	2.3863858	—
a_{10}	1.8473522	2.5627544	2.4699913	1.4182045	2.1999636	1.5630489	-0.27808770	—
a_{11}	—	3.2090348	3.2589167	1.6126473	1.6738315	1.2347535	2.2877226	—
a_{12}	—	1.7199815	1.0225941	1.0514379	1.5694120	1.6079233	2.2261987	—
a_{13}	—	1.4324950	1.3834717	—	1.3755863	1.4248362	—	—
a_{14}	—	0.98077126	0.66417084	—	1.2489497	1.2645058	—	—
a_{15}	—	—	—	—	—	—	—	—
b_2	0.54302214	0.68813389	0.12446733	-0.019240948	-0.44912589	0.33846693	5.0343911×10^{-5}	0.10223657
b_3	-1.1399357	-0.00039714071	0.30989985	-0.0014477705	0.00063720755	0.69457071	6.8778459×10^{-5}	2.1101571
b_4	0.00043800143	0.69258818	0.62289073	1.7263411	3.3250651	-5.8644681	7.7330748	9.5647705
b_5	-0.38524988	1.0814065	1.0961212	-2.5315734	4.5005640	-4.8860274	5.1054505	5.6239195
b_6	0.93378609	0.71934185	1.1048902	2.0477887	3.2570552	-3.2768201	3.4139483	—
b_7	0.65564955	0.55425434	0.98172590	1.5138314	1.4510343	1.8846898	2.2716376	—
b_8	0.610636639	0.83050907	1.2021281	1.1564126	1.3072494	1.7601059	1.6507471	—
b_9	0.62958064	0.87980828	1.4022357	0.89408669	1.5644506	1.1565902	1.4314759	—
b_{10}	—	0.69537014	0.74420922	0.99052342	1.3117581	1.2503444	1.2004102	—
b_{11}	—	0.55345110	0.48900680	0.86252293	1.0234697	1.1038372	0.83713219	—
b_{12}	—	0.78634411	0.99273417	—	0.83188842	0.84194437	—	—
b_{13}	—	0.85019449	0.80499628	—	0.91106835	0.88072166	—	—

TABLE 4: Parameters for fits to the down-spin electron densities for the first-row ions to the form given by Eq. 3.4.

	Li⁺	Be⁺	B⁺	C⁺	N⁺	O⁺	F⁺	Ne⁺
A	$-1.1975211 \times 10^{-7}$	$-2.4277748 \times 10^{-7}$	1.6666289×10^{-5}	$-2.2554092 \times 10^{-5}$	0.0006493159	0.0052644159	0.11701095	0.033045688
a_0	-1.9231800	-2.8368325	-3.5887826	-4.1555416	-4.6397287	-5.0504222	-5.3984164	-5.7277903
a_2	-1.3094033	-239.01631	-1.2956560	-1.5760094	-1.5561329	-1.9027866	1.4522860	0.66342615
a_3	0.55285083	418.62665	0.95117735	0.18094435	1.7143346	4.6910981	-39.448257	-59.158934
a_4	3.0595672	474.24425	-1.2520313	0.35177681	-10.199410	-24.508500	30.436287	-25.599765
a_5	2.1374497	264.14462	-0.037610912	4.1267784	9.1374048	11.306566	40.106674	62.201269
a_6	0.10351703	149.71145	5.6984130	8.9462107	36.579626	75.906067	31.797896	97.773654
a_7	2.0986932	99.885114	8.3411763	8.8806721	16.844302	41.262462	22.477630	16.424269
a_8	2.9758086	73.021390	4.9270617	5.7175137	10.360411	24.047315	20.574033	20.930749
a_9	1.3999542	33.894531	1.3118585	2.9161618	3.2230824	4.6397563	—	—
a_{10}	1.6409543	4.9412125	1.9350441	1.4729397	—	—	—	—
a_{11}	0.10984607	-1.6912689	3.0120610	1.8431733	—	—	—	—
a_{12}	—	1.6218976	1.4868803	0.60955627	—	—	—	—
a_{13}	—	4.4919176	1.3216797	—	—	—	—	—
a_{14}	—	0.48987127	0.87659717	—	—	—	—	—
a_{15}	—	—	—	—	—	—	—	—
b_2	-0.54868428	-9.1552268	-0.0085092824	-0.058707098	0.0045000467	0.00012255747	-0.58599533	-0.0088266927
b_3	-0.89772547	9.4949246	0.39117891	-0.0040471271	-0.010311029	0.00060132461	0.0087297635	-0.076600424
b_4	0.56610545	-7.1463500	0.67477309	1.2039059	1.0174291	0.71455157	4.0868567	8.4935308
b_5	-0.52998023	4.5995273	0.91281505	-1.9360285	3.2404600	4.9838165	9.3250899	8.4992424
b_6	0.87296189	-4.9369987	1.1308559	-1.7131087	3.4840754	4.8765088	2.3616297	5.9950443
b_7	0.72982515	-3.6100980	1.3293243	1.4490485	1.5452275	2.2443343	—	—
b_8	0.66319443	-1.7522292	1.4244468	1.1272928	—	—	—	—
b_9	0.57538290	0.68563932	0.74576349	0.96009092	—	—	—	—
b_{10}	—	0.73317362	0.48069651	0.78836512	—	—	—	—
b_{11}	—	0.87273677	0.94663160	—	—	—	—	—
b_{12}	—	0.58977393	0.87812034	—	—	—	—	—
b_{13}	—	—	—	—	—	—	—	—

TABLE 5: Parameters for fits to the up-spin-up-spin electron-pair intracule densities for the first-row atoms to the form given by Eq. 3.6.

	Li	Be	B	C	N	O	F	Ne
a_0	6.1065917	3.9255932	1.2026733	−0.43893261	−1.7967162	−2.5902780	−3.2627418	−3.7208117
a_2	10.102248	27.397393	15.050222	5.0659883	9.0996460	0.10089390	−2.7666309	−5.1021645
a_3	3.0239738	7.3166748	−14.054645	−3.5608420	−14.198444	−10.786700	−4.7988024	9.0828953
a_4	7.5251002	38.538351	25.438617	8.0270873	5.9046552	−2.6635210	−2.0626111	2.2300920
a_5	1.6881492	20.674499	9.9848824	9.5920353	6.0918842	5.8898622	1.8279229	−0.24947911
a_6	1.9089665	−8.7016632	4.1145611	7.4162083	1.2032377	10.369401	5.8473673	2.5574798
a_7	0.46389805	14.732409	9.2345894	3.2563627	2.9288293	14.638897	7.6246743	0.44260097
a_8	—	15.185904	0.32995064	3.0859328	0.29134275	12.072718	7.7937726	—
a_9	—	0.35445684	1.4336007	4.4045459	—	5.6086891	7.3252424	—
a_{10}	—	3.5198870	0.25010354	1.3396916	—	4.1856940	5.3339935	—
a_{11}	—	2.3584995	—	1.5819677	—	3.9905679	3.5273617	—
a_{12}	—	1.6285084	—	0.19469285	—	1.9827723	2.4590522	—
a_{13}	—	0.26357008	—	—	—	2.1868062	1.8997834	—
a_{14}	—	—	—	—	—	0.33311477	1.9101844	—
a_{15}	—	—	—	—	—	—	0.30450358	—
b_2	−1.1028890	−2.3801132	−2.2700346	1.9806753	1.0691667	2.2070596	−2.3479159	2.1748361
b_3	1.1925429	2.5299741	0.062547305	−0.96456449	−0.0010577033	1.4672477	4.7459899×10^{-5}	0.00097791067
b_4	0.55535023	−1.9531603	1.7875021	2.4320913	1.8912910	2.1783228	2.3621426	0.92347343
b_5	0.83920563	1.8882312	2.6494620	1.4569079	1.0273714	−4.1954299	2.9598140	0.81797319
b_6	—	1.3535868	−0.61004891	1.5243105	1.0593389	2.1357214	2.9075037	—
b_7	—	1.2067544	0.68999833	1.5060977	—	1.9710446	1.6409430	—
b_8	—	1.4602620	0.81273172	0.86735983	—	2.5532010	2.0921500	—
b_9	—	0.91949585	—	0.96169834	—	1.7194802	1.9875127	—
b_{10}	—	0.61338379	—	0.82775707	—	0.98105236	1.4959300	—
b_{11}	—	0.86275779	—	—	—	1.2715184	1.1448211	—
b_{12}	—	—	—	—	—	0.81463977	0.98728371	—
b_{13}	—	—	—	—	—	—	0.77438677	—

TABLE 6: Parameters for fits to the up-spin-down-spin electron-pair intracule densities for the first-row atoms to the form given by Eq. 3.7.

	Li	Be	B	C	N	O	F	Ne
a_0	0.61526412	−0.45401621	−1.2077116	−1.8454732	−2.3867899	−2.8460027	−3.2538261	−3.6160528
a_2	5.9358608	7.8898465	13.289273	13.092863	25.296042	22.858406	27.941130	22.168727
a_3	1.2485362	−6.2570027	−6.7078980	−1.1757649	−34.901550	−24.964177	−56.334606	−65.256373
a_4	1.5296757	15.037601	−5.2824265	−2.5210220	13.136634	−20.039010	−23.071581	−76.656556
a_5	10.337959	13.810730	7.1487641	−1.9239376	5.1813598	10.218548	12.637342	−24.898403
a_6	26.213312	35.620419	−0.80457765	4.4316101	5.7409560	17.229105	19.192329	12.328632
a_7	29.568360	39.412805	0.66585954	12.769011	5.2265771	11.371280	15.640388	31.892458
a_8	18.764349	3.5650191	0.39527579	11.584476	0.68031861	6.9766827	8.0141255	26.584721
a_9	8.7769528	−1.9654578	—	3.6313991	—	2.9813608	1.0597980	11.304558
a_{10}	2.5175545	9.3429902	—	0.76821741	—	0.61280964	1.6550773	5.9075983
a_{11}	−0.073576699	3.1876340	—	3.2608291	—	1.4033784	0.93160090	4.2071639
a_{12}	3.4118648	0.12232887	—	2.9505498	—	1.0070321	—	1.0410535
a_{13}	0.90934168	0.52700746	—	0.52930266	—	—	—	1.5175584
a_{14}	0.42654859	—	—	1.2957111	—	—	—	0.64467668
a_{15}	—	—	—	0.27689057	—	—	—	—
b_2	−0.54865058	1.9416388	0.70513282	1.7461397	1.3128383	2.1329821	2.3621208	3.0225578
b_3	2.2339900	1.2474118	−0.0013060938	−0.00078147196	0.0020754240	0.0062414833	0.0013249887	−0.0069048869
b_4	2.2409313	3.9624968	−0.0022761829	−0.058939582	2.0114822	0.61864298	−4.3265622	−6.0937501
b_5	2.1562168	2.1258147	0.85955088	1.0416219	2.7528404	5.7579290	6.5026344	9.3595662
b_6	3.0819718	−0.019364228	0.62861137	2.7833867	1.5059355	3.1890071	2.4684947	4.2797403
b_7	3.2999674	1.6322528	—	2.2645936	—	1.4413006	1.2067139	3.8637234
b_8	1.3350249	2.9035215	—	2.1289753	—	1.4681711	1.9592937	−4.8464065
b_9	0.00062348440	1.5339398	—	1.9294132	—	1.5641188	0.83492095	2.0143182
b_{10}	$-4.6080849 \times 10^{-5}$	0.56767371	—	0.87940682	—	0.83705574	—	0.36976872
b_{11}	$-2.8685396 \times 10^{-5}$	0.85852615	—	0.64306523	—	—	—	1.6615905
b_{12}	0.96636640	—	—	1.2846726	—	—	—	0.73282070
b_{13}	—	—	—	0.79800424	—	—	—	—

TABLE 7: Parameters for fits to the down-spin-down-spin electron-pair intracule densities for the first-row atoms to the form given by Eq. 3.6.

	Li	Be	B	C	N	O	F	Ne
a_0	—	4.0815651	2.4499188	1.4461641	0.57352755	−1.5269160	−2.8312426	−3.7208612
a_2	—	22.156247	12.357642	3.9114162	4.2828691	−5.0364817	−2.5140508	−1.5604028
a_3	—	40.055019	33.257889	−0.30661256	−1.2558920	−7.6127941	−4.9643529	4.4483608
a_4	—	20.560701	111.55884	2.8236599	0.54403776	34.870034	5.8491221	4.4895372
a_5	—	−7.1137150	13.295365	8.6715290	0.30689219	28.880770	9.1593069	−2.4221168
a_6	—	26.576376	48.290440	6.8303567	—	7.1073175	8.7891068	2.8527301
a_7	—	4.7350942	71.346829	4.2861914	—	−3.2811399	8.8418556	0.42438509
a_8	—	2.4521505	18.371329	8.4722229	—	10.703109	6.5983725	—
a_9	—	2.5193879	3.0863699	4.7671831	—	8.3658091	3.9598092	—
a_{10}	—	1.4957276	6.0651288	−3.4339113	—	−0.62024613	3.5819963	—
a_{11}	—	0.18218378	0.99206348	3.7754244	—	2.0682160	3.0861490	—
a_{12}	—	—	1.7572607	5.9095596	—	0.53805800	1.7683981	—
a_{13}	—	—	0.23918339	4.3561406	—	—	1.8324432	—
a_{14}	—	—	—	2.2325511	—	—	0.24761161	—
a_{15}	—	—	—	0.21641869	—	—	—	—
b_2	—	−2.3862165	−1.1835718	0.85404664	0.51707824	−3.0330298	2.4715861	1.9936198
b_3	—	−3.2533031	5.2646780	0.018108457	−0.00070109191	2.6257468	0.0019281769	−0.0020933309
b_4	—	1.3306505	4.4815591	1.6514310	—	−3.2639446	−2.6030828	0.58170841
b_5	—	0.97795945	2.5380807	−1.7886583	—	0.00076348704	−2.9474069	0.86820782
b_6	—	1.7813673	4.3826141	−1.5625823	—	1.9637202	1.4798700	—
b_7	—	1.3458432	−2.8586607	1.0479958	—	2.0462305	1.3307398	—
b_8	—	0.055544788	−1.7730674	0.61237701	—	0.54569534	2.0882557	—
b_9	—	0.87644108	0.71208722	0.53577068	—	1.1083418	1.2261216	—
b_{10}	—	—	0.77080572	1.3038485	—	0.62575330	0.96767065	—
b_{11}	—	—	0.76343532	1.2279869	—	—	0.98273699	—
b_{12}	—	—	—	1.1117660	—	—	0.76267695	—
b_{13}	—	—	—	0.87097106	—	—	—	—

TABLE 8: Parameters for fits to the up-spin-up-spin electron-pair intracule densities for the first-row ions to the form given by Eq. 3.6.

	Li$^+$	Be$^+$	B$^+$	C$^+$	N$^+$	O$^+$	F$^+$	Ne$^+$
a_0	—	3.7478317	2.2175841	−0.15270330	−1.6060611	−2.7404607	−3.4324839	−3.9980728
a_2	—	5.8362511	2.9195473	5.7206309	2.4092687	−0.63953260	−2.8199671	−1.7304720
a_3	—	3.6414698	−1.1782039	−3.3886223	−3.4692153	−1.2460454	−6.0216468	−5.2100873
a_4	—	8.5159602	1.8777987	14.730761	2.1608578	4.8534518	0.63434288	−0.57462463
a_5	—	12.962116	0.49840186	6.2502236	6.7742753	1.4873129	4.6535082	2.0176292
a_6	—	−3.0085673		0.54366011	9.0763297	1.1209143	6.1793351	2.5109031
a_7	—	2.4740239		0.16658390	7.3282383	2.8772014	5.2688924	3.8466263
a_8	—	15.182606		0.50816931	1.8485765	1.5926772	2.6578247	4.4663581
a_9	—	5.2344283		—	0.069480474	1.8937771	1.9387397	3.5909461
a_{10}	—	−2.6324145		—	2.5342730	0.30993590	2.9124944	3.1088479
a_{11}	—	4.6803444		—	2.6755064	—	2.8419124	3.3553737
a_{12}	—	3.4163240		—	1.0739453	—	1.6991465	2.4659262
a_{13}	—	1.3853975		—	1.7346271	—	1.9510985	1.3313506
a_{14}	—	1.5918068		—	0.36361598	—	0.61820704	2.0562009
a_{15}	—	0.84009893		—	—	—	—	0.50170531
b_2	—	1.0955041	0.00019132364	2.1070312	1.9687728	2.0339525	2.2948842	2.2348082
b_3	—	1.2060727	0.73759678	1.3825309	0.062967704	0.0014859869	−0.0010207953	0.00087648078
b_4	—	1.5322045		1.7273096	2.6327344	1.5077935	2.2850845	1.0143223
b_5	—	1.0239911		0.51119470	1.8407727	1.5061693	2.7731729	−3.5516922
b_6	—	1.1792980		0.26115129	−1.1945176	0.86713802	1.2974783	0.70910498
b_7	—	1.0438200		—	1.2689776	1.1462007	1.6218124	0.77857025
b_8	—	1.2383159		—	1.5960960	0.72318392	1.6881113	2.1557789
b_9	—	1.2864798		—	0.71541339	—	1.0651784	1.9211458
b_{10}	—	0.29003885		—	0.72004903	—	1.3392007	1.0469509
b_{11}	—	0.66495614		—	1.0755849	—	1.0918245	1.0202775
b_{12}	—	0.92312399		—	0.73038441	—	0.70232242	1.0572344
b_{13}	—	0.65988665		—	—	—	—	0.68993938

TABLE 9: Parameters for fits to the up-spin-down-spin electron-pair intracule densities for the first-row ions to the form given by Eq. 3.7.

	Li^+	Be^+	B^+	C^+	N^+	O^+	F^+	Ne^+
a_0	0.62552975	−0.44909445	−1.2545857	−1.8605821	−2.4056123	−2.8562793	−3.2554127	−3.6182335
a_2	7.2373997	9.0685018	11.056703	19.254426	30.712748	37.138856	26.523378	34.758782
a_3	−5.4458166	−6.1049896	−3.2418876	−16.514946	−52.678414	−76.992116	−43.209398	−104.17006
a_4	4.0003507	9.6022044	−1.3597062	−1.3819619	31.668477	49.065662	−34.270903	−24.324272
a_5	6.6837652	7.7361068	8.2264901	9.0486143	0.93955817	−6.1186651	3.2592593	13.600450
a_6	0.85987438	0.83485973	18.045621	2.0283536	1.2101348	10.101845	21.776109	19.630115
a_7	2.8817229	1.7848778	18.523130	1.6399876	2.1774988	2.8177127	25.413800	31.322060
a_8	1.9527465	2.5712309	7.4416874	0.63731812	—	—	17.232586	10.659598
a_9	1.6451738	1.0213159	−0.82976098	—	—	—	5.6718942	3.0066278
a_{10}	0.93036564	1.5986812	1.5392090	—	—	—	0.97376412	2.4129485
a_{11}	—	1.2736447	4.3898207	—	—	—	1.3812832	—
a_{12}	—	0.73855864	2.0428447	—	—	—	1.1339139	—
a_{13}	—	—	0.59014176	—	—	—	1.1309206	—
a_{14}	—	—	0.69501944	—	—	—	—	—
a_{15}	—	—	—	—	—	—	—	—
b_2	0.98059526	1.7280844	1.7598330	0.99197666	0.72177138	1.1740966	2.4377041	2.5463602
b_3	0.39645054	0.0017902607	−0.030295850	0.0024345766	0.00041532634	0.00044548301	0.014616025	0.0010436638
b_4	1.3343393	0.56826552	−0.21872413	0.16154715	2.3825676	3.6044888	−2.2761095	6.7721050
b_5	1.3018818	1.7405761	2.0864065	2.0056585	1.7963695	2.5775632	7.9418104	6.9006922
b_6	0.0006959590327	0.80553042	2.1755355	0.86549684	—	—	3.5128421	−0.0039021165
b_7	0.99609456	0.0031647735	2.2046425	—	—	—	1.4986032	3.5392695
b_8	0.68024380	0.15255942	2.1197151	—	—	—	2.4181383	1.7190358
b_9	—	0.35084652	1.2125677	—	—	—	1.8329593	—
b_{10}	—	1.0248964	0.19052111	—	—	—	0.80120115	—
b_{11}	—	—	0.79244407	—	—	—	1.0422837	—
b_{12}	—	—	1.0174250	—	—	—	—	—
b_{13}	—	—	—	—	—	—	—	—

TABLE 10: Parameters for fits to the down-spin-down-spin electron-pair intracule densities for the first-row ions to the form given by Eq. 3.6.

	Li$^+$	Be$^+$	B$^+$	C$^+$	N$^+$	O$^+$	F$^+$	Ne$^+$
a_0	—	—	2.2235333	1.2061331	0.36576806	−0.26666609	−2.4033157	−3.5529040
a_2	—	—	2.7787711	3.7359748	4.8111471	2.8461879	−5.1120047	−0.46351364
a_3	—	—	−1.0297412	−2.0086771	−2.5445219	3.7931767	−15.063783	−0.12053336
a_4	—	—	1.2006057	1.4319399	1.5007665	3.3894880	39.205563	4.6124179
a_5	—	—	0.48398989	0.45519876	0.60646339	2.6708379	13.175521	3.1893901
a_6	—	—	—	—	—	3.5560091	3.0948920	−0.0039015981
a_7	—	—	—	—	—	0.59045759	15.155183	0.94654716
a_8	—	—	—	—	—	—	6.2130202	4.7954434
a_9	—	—	—	—	—	—	2.1627279	4.8439899
a_{10}	—	—	—	—	—	—	0.10541937	2.4260420
a_{11}	—	—	—	—	—	—	0.48923601	1.2665867
a_{12}	—	—	—	—	—	—	—	2.0008981
a_{13}	—	—	—	—	—	—	—	0.58982983
a_{14}	—	—	—	—	—	—	—	—
a_{15}	—	—	—	—	—	—	—	—
b_2	—	—	0.16179635	0.10747816	0.38034765	−1.6813640	2.8823186	2.1596809
b_3	—	—	0.56794426	0.52719675	0.43148225	0.11759084	1.6120798	−0.0049916896
b_4	—	—	—	—	—	1.1398823	3.1507566	1.1205209
b_5	—	—	—	—	—	0.85039237	2.1130157	1.7065465
b_6	—	—	—	—	—	—	1.3182877	1.5677338
b_7	—	—	—	—	—	—	1.7973564	1.5795226
b_8	—	—	—	—	—	—	0.75034947	1.1894809
b_9	—	—	—	—	—	—	0.099556014	1.0433779
b_{10}	—	—	—	—	—	—	—	0.97702982
b_{11}	—	—	—	—	—	—	—	0.69232902
b_{12}	—	—	—	—	—	—	—	—
b_{13}	—	—	—	—	—	—	—	—