Quantum Monte Carlo: Theory and Application to Atomic, Molecular and Nano Systems

[1,2]Alán Aspuru-Guzik, [1,3]Alexander C. Kollias, [1,3]Romelia Salomón-Ferrer, and [1,3]William A. Lester, Jr.

[1]*Kenneth S. Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley, CA 94720-1460, USA. All authors contributed equally.*

[2]*Gates Millenium Scholarship Holder*

[3] *Chemical Sciences Division, Lawerence Berkeley National Laboratory*

emails: alan@aspuru.com, romelia@berkeley.edu, sasha@okra.cchem.berkeley.edu,walester@lbl.gov

Preface

This review deals with some of the methods known under the designation quantum Monte Carlo (QMC) that have been used for the determination of electronic structure of atomic to many-particle systems. Sources that complement the coverage in this chapter are the reviews [1–10]. and our previous monograph [11]. There are also chapters on QMC contained in selected computational physics texts [12–14].

QMC methods not covered in this review are the auxiliary field [15–18] and path integral QMC approaches [19, 20]. Other areas that will not be discussed are rotational and vibrational states of molecules [21–24] and broad areas of condensed matter [9, 19, 25], and nuclear structure [26, 27] except where relevant to the present focus.

QMC methods have several advantages:

- Computer time scales with system size roughly as N^3, where N is the number of particles of the system. Recent developments have made possible the approach to linear scaling in certain cases.
- Computer memory requirements are small and grow modestly with system size.
- QMC is intrinsically parallel so that associated computer codes are significantly smaller and more readily adapted to parallel computers than basis set molecular quantum mechanics computer programs.

- Basis set truncation and basis set superposition errors are absent in one of the variants of (diffusion MC) of QMC computations.
- Monte Carlo numerical accuracy can be arbitrarily increased. QMC calculations have an accuracy dependence of $\sqrt{T}$, where T is the computer time. This enables one to choose an accuracy range and readily estimate the computer time needed for performing a calculation of an observable with an acceptable error bar.

The purpose of the present paper is to described the commonly used algorithms of QMC for electronic structure and to report some recent developments in the field. We will also review selected recent applications of the approaches.

The chapter is organized as follows. In Section 1, we introduce the topic, as well as enumerate wave function properties that are useful for QMC applications. In Section 2 we describe commonly used QMC algorithms. In Sections 2.7.2 - 2.10 we briefly introduce some special topics that remain fertile research areas. In Sections **??** and 8.3 we present results of calculations of atomization energies, heats of reaction and excited states energies. We comment on work on carbon clusters and transition-metal systems in sections 6.2 and 7. In Section 9.1 we present results of QMC calculations made on silicon nano-clusters. In section 9.2 we describe calculations of two-dimensional quantum dots in vacuum and in the presence of a magnetic field. Finally, in section 9.2.4 we briefly summarize recent progress on excitonic complexes.

Atomic units are used throughout, the charge of the electron e and Planck's normalized constant $\hbar$ are set to unity. In this metric system, the unit distance is the Bohr radius a_0. Energy units are in hartrees (h), kcal/mol, and eV.

1 Introduction

The goal of the QMC method is to solve the Schrödinger equation, which in the time independent form is given by

$$\widehat{H}\Psi_n(\mathbf{R}) = E_n\Psi_n(\mathbf{R}). \tag{1}$$

Here, $\widehat{H}$ is the Hamiltonian operator of the system, with wave function $\Psi_n(\mathbf{R})$ and energy E_n. The index n denotes a particular state. $n = 0, 1, \ldots$. $\mathbf{R}$ is a vector that denotes the $3N$ coordinates of the system of N particles (electrons and nuclei), $\mathbf{R} \equiv \{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$. For molecular systems, in the absence of electric or magnetic fields, the Hamiltonian has the form $\widehat{H} \equiv \widehat{T} + \widehat{V}$, where $\widehat{T}$ is the kinetic energy operator, $\widehat{T} \equiv -\frac{1}{2}\nabla^2_{\mathbf{R}}$, and $\widehat{V}(\mathbf{R})$ is the potential energy operator. For atomic and molecular systems $\widehat{V}$ is the Coulomb potential between particles of charge q_i, i.e., $\widehat{V} \equiv \sum_{ij} \frac{q_{ij}}{\mathbf{r}_{ij}}$.

The first suggestion of a Monte Carlo solution of the Schrödinger equation dates back to Enrico Fermi [28] who indicated that a solution to the stationary state equation

$$-\frac{1}{2}\nabla^2_{\mathbf{R}}\Psi(\mathbf{R}) = E\Psi(\mathbf{R}) - V(\mathbf{R})\Psi(\mathbf{R}) \tag{2}$$

could be obtained by introducing a wave function of the form $\Psi(\mathbf{R},\tau) = \Psi(\mathbf{R})e^{-E\tau}$. This yields the equation

$$\frac{\partial\Psi(\mathbf{R},\tau)}{\partial\tau} = \frac{1}{2}\nabla^2\Psi(\mathbf{R},\tau) - V(\mathbf{R})\Psi(\mathbf{R},\tau). \tag{3}$$

Taking the limit $\tau \to \infty$, in Eq. 3 recovers Eq. 2. If the second term on the right hand side of Eq. 3 is ignored, the equation is isomorphic with a diffusion equation, which can be simulated by a random walk [29, 30], where random walkers diffuse in a $\mathbf{R}$-dimensional space. If the first term is ignored, the equation is a first-order kinetics equation with a position-dependent rate constant, $V(\mathbf{R})$, which can also be interpreted as a stochastic survival probability. A numerical simulation in which random walkers diffuse through $\mathbf{R}$-space, reproduce in regions of low potential, and die in regions of high potential leads to a stationary distribution proportional to $\Psi(\mathbf{R})$, from which expectation values can be obtained.

1.1 Numerical solution of the Schrödinger equation

Most efforts to solve the Schrödinger equation rely on basis sets. These approaches rely almost exclusively on one or a linear combinations of Slater determinants, and include the Hartree-Fock (HF), the density functional theory (DFT), the configuration interaction, (CI) and the multi-configuration self-consistent field (MCSCF) methods. There are perturbation approaches, for example the Møller–Plesset methods (MP(N), N= 2 - 4), and coupled cluster (CC) approaches, which are presently popular computational procedures. An exact basis-set calculation with a given basis set expansion requires $N!$ computer operations, where N is the number of basis functions. A method that competes in accuracy with QMC, such as coupled cluster with singles, doubles, and perturbative treatment of triple excitations, CCSD(T), scales as N^7.[1]

A term that we will use later is correlation energy (CE). It is defined as the difference between the exact nonrelativistic energy and the HF energy in the limit of an infinite basis set [34, 35], *i. e.*,

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}. \tag{4}$$

The CI, MCSCF, MP(N), and CC methods are all directed at generating energies that approach E_{exact}. Other methods that have been developed include dimensional expansions [36], and the contracted Schrödinger equation [37].

Since the pioneering work of the late forties to early sixties [28, 38, 39] the MC and related methods have grown in interest. The QMC methods have an advantage over wave function methods with system size scaling, in the simplicity of algorithms and in trial wave function forms that can be used for importance sampling.

[1] For a more detailed analysis of the scaling of wave-function-based methods see, for example, [31], and [32]. For a general overview of these methods, see ref. [33].

1.2 Properties of the exact wave function

The exact time independent wave function solves Eq. 1. Some analytic properties of this function are very helpful in the construction of trial functions for QMC methods.

For the present discussion, we are interested in the discrete spectrum of the $\widehat{H}$ operator. In most applications the total Schrödinger equation 1 is separated into an electronic Schrödinger equation and a nuclear Schrödinger equation based on the large mass difference between electrons and nuclei. This is the essence of the Born–Oppenheimer (BO) approximation. This separation, in principle, need not be introduced in QMC, but there is the practical benefit that fixing the positions of the nuclei results in the simplest form of the electronic Schrödinger equation.

The wave function also must satisfy the virial, hypervirial, Hellman–Feynman and generalized Hellman–Feynman theorems [40–42].[2] The local energy [43],

$$E_L(\mathbf{R}) \equiv \frac{\widehat{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} \tag{5}$$

is a constant for the exact wave function. This property is crucial for understanding and improving QMC methods.

When charged particles meet, there is a singularity in the Coulomb potential. This singularity must be compensated by a singularity in the kinetic energy, which results in a discontinuity in the first derivative, i.e., a *cusp*, in the wave function when two or more particles meet [44, 45]. For one electron described by an orbital $\phi(\mathbf{r})$ coalescing at a nucleus. Here, $\phi(\mathbf{r}) = \chi(r)Y_l^m(\theta,\phi)$, $\chi(r)$ is a radial function, and $Y_l^m(\theta,\phi)$ is a spherical harmonic with angular and magnetic quantum numbers l and m, the electron–nucleus cusp condition is

$$\left.\frac{1}{\eta(r)}\frac{d\eta(r)}{dr}\right|_{r=0} = -\frac{Z}{l+1}. \tag{6}$$

Here $\eta(r)$ is the radial wave function with the leading r dependence factored out, $\eta(r) \equiv \chi(r)/r^m$, and Z is the atomic number of the nucleus.

Furthermore, $\bar{\rho}(\mathbf{r})$, the spherical average of the electron density, $\rho(\mathbf{r})$,[3] must satisfy another cusp condition, namely,

$$\left.\frac{\partial}{\partial r}\bar{\rho}(r)\right|_{r=0} = -2Z\bar{\rho}(r) \tag{7}$$

at any nucleus. Another condition on $\rho(\mathbf{r})$ is that asymptotically it decays exponentially [46, 47]:

$$\rho(r \to \infty) \approx e^{-2\sqrt{2I_0}r}, \tag{8}$$

[2] The Hellman–Feynman theorem is discussed in Section 2.9.3.

[3] If N is the number of electrons, then $\rho(r)$ is defined by

$$\rho(r) = N\int |\Psi(\mathbf{R})|^2\, d\mathbf{R}.$$

where I_0 is the first ionization potential. This relation can be derived from consideration of a single electron at large distance.

For electron–electron interactions, the cusp condition takes the form

$$\frac{1}{\eta_{ij}(r)}\frac{d\eta_{ij}(r)}{dr}\bigg|_{r_{ij}=0} = \frac{1}{2(l+1)}, \tag{9}$$

where $\eta_{ij}(r)$ is analogous to $\eta(r)$ of Eq. 6[11].

We discuss how to impose properties of the exact wave function on QMC trial functions in Section 2.1.

1.3 Approximate wave functions

James and Coolidge [48] proposed three accuracy tests of a trial wave function, Ψ_T: the root mean square error in Ψ_T

$$\delta_\Psi = \left[\int (\Psi_T - \Psi_0)^2 \, d\mathbf{R}\right]^{\frac{1}{2}}, \tag{10}$$

the energy error

$$\delta_E = E(\Psi_T) - E_0 \tag{11}$$

and the root mean square energy deviation

$$\delta_{E_L} = \left[\int \left|\left(\widehat{H} - E_0\right)\Psi\right|^2 d\mathbf{R}\right]^{\frac{1}{2}} \tag{12}$$

where the local energy is defined by Eq. 5. The calculation of δ_Ψ by QMC requires sampling the exact wave function, a procedure that will be described in Section 2.8.

Several stochastic optimization schemes have been proposed for minimizing expressions (10)–(12). Most researchers have focused on Eq. 12, i.e., minimizing δ_{E_L}; see, for example, ref. [49]. In Section 2.2.5 we turn to stochastic wave function optimization procedures.

2 Algorithms

Here we describe QMC computational methods. All of these methods use MC techniques found widely in other fields, such as operations research, applied statistics, and classical statistical mechanics simulations. Techniques such as importance sampling, correlated sampling and MC optimization are similar in spirit to those described in other MC treatises. The reader is referred to refs. [49–59] for details on the techniques described in this section.

We next present the simple, yet powerful variational Monte Carlo (VMC) method, in which the Metropolis MC [4] method is used to sample a known trial

[4]This algorithm is also known as the M(RT)^2, due to the full list of the authors that contributed to its development, Metropolis, Rosenbluth, Rosenbluth, Teller and Teller, see ref. [60] .

function Ψ_T. We follow with the projector Monte Carlo (PMC) methods that sample the unknown ground state wave function.

2.1 Trial wave functions

In contrast to wave function methods, where the wave function is usually constructed from determinants of orbitals, QMC methods can use arbitrary functional forms for the wave function. Because QMC trial wave functions are not restricted to expansions in one-electron functions, more compact representations are routinely used. In this section, we review the forms commonly used for QMC calculations.

Fermion wave functions must be antisymmetric with respect to exchange of an arbitrary pair of particle coordinates. If they are constructed as the product of N functions of the coordinates, $\phi(r_1, r_2, \ldots, r_N)$, the most general wave function can be constructed enforcing explicit permutation,

$$\Psi(\mathbf{R}, \Sigma) = \frac{1}{\sqrt{(N \cdot M)!}} \sum_{n,m} (-1)^n \widehat{S}_m \widehat{P}_n \phi(r_1, \sigma_1 r_2, \sigma_2, \ldots, r_N, \sigma_N), \qquad (13)$$

where $\widehat{P}_n$ is the nth coordinate permutation operator, $\widehat{P}_n \phi(r_1, r_2, \ldots, r_i, r_j, \ldots, r_N) = \phi(r_1, r_2, \ldots, r_j, r_i, \ldots, r_N)$, and $\widehat{S}_m \phi(\sigma_1, \sigma_2, \sigma_i, \sigma_j, \ldots, \sigma_N) = \phi(\sigma_1, \sigma_2, \sigma_j, \sigma_i, \ldots, \sigma_N)$ is the mth spin coordinate permutation operator. Σ denotes the full spin space, $\Sigma \equiv \{\sigma_1, \ldots, \sigma_N\}$.

If the functions ϕ_i depend only on single-particle coordinates, their antisymmetrized product can be expressed as a Slater determinant

$$D(\mathbf{R}, \Sigma) = \frac{1}{\sqrt{N!}} \det\big|\phi_1, \ldots, \phi_i(r_j, \sigma_j), \ldots, \phi_N\big|. \qquad (14)$$

Trial wave functions constructed from orbitals scale computationally as N^3, where N is the number of particles, compared to $N!$ for the fully antisymmetrized form.[5] The number of evaluations can be reduced by determining which permutations contribute to a particular spin state.

For QMC evaluation of properties that do not depend on spin coordinates Σ for a given spin state, the $M!$ configurations that arise from relabeling electrons, need not be evaluated. The reason is that the Hamiltonian of Eq. 1, contains no magnetic or spin operators and spin degrees of freedom remain unchanged. In this case, and for the remainder of this paper, $\sigma_\uparrow$ electrons do not permute with $\sigma_\downarrow$ electrons, so that the full Slater determinant(s) can be factored into a product of spin-up, $D^\uparrow$, and spin-down, $D^\downarrow$, determinants. The number of allowed permutations is reduced from $(N_\uparrow + N_\downarrow)!$ to $N_\uparrow! N_\downarrow!$ [61, 62].

[5] The evaluation of a determinant of size N requires N^2 computer operations. If the one-electron functions scale with system size as well, the scaling becomes N^3. In contrast, an arbitrary fully antisymmetrized form that cannot be expressed using determinants requires the explicit evaluation of the $N!$ permutations, making the evaluation of this kind of wave function in QMC prohibitive for systems of large N.

The use of various wave function forms in QMC has been explored by Alexander and Coldwell [63], as well as Bertini *et al.* [64]. Some of these antisymmetric descriptions of the wave function are more flexible and require fewer parameters than determinants, but their evaluation is inefficient due to $N!$ scaling.

A good compromise is to use a product wave function of a determinant or linear combination of determinants, e.g., HF, MCSCF, CASSCF, CI, multiplied by a correlation function that is symmetric with respect to particle exchange,

$$\Psi_T = \mathcal{D}\mathcal{F}. \tag{15}$$

Here $\mathcal{D}$ denotes the antisymmetric wave function factor and $\mathcal{F}$ is the symmetric factor. We now describe some of the forms used for $\mathcal{D}$ and then we describe forms for $\mathcal{F}$. Such products are also known as the correlated molecular orbital (CMO) wave functions.

With CMO wave functions, the antisymmetric part of the wave function is constructed as a linear combination of determinants of independent particle functions, ϕ_i (see Eq. 14). The ϕ_i are usually formed as a linear combination of basis functions centered on atomic centers, $\phi_i = \sum_j c_j \chi_j$. The most commonly used basis functions in traditional *ab initio* quantum chemistry are Gaussian functions, which owe their popularity to ease of integration of molecular integrals. Gaussian basis functions take the form

$$\chi_G \equiv x^a y^b z^c e^{-\xi \mathbf{r}^2}. \tag{16}$$

For QMC applications, it is better to use the Slater-type basis functions

$$\chi_S \equiv x^a y^b z^c e^{-\xi \mathbf{r}}, \tag{17}$$

because they rigorously satisfy the electron–nuclear cusp condition of Eq. 6, and the asymptotic property of Eq. 8. Nevertheless, in most studies, Gaussian basis functions have been used, and corrections for enforcing the cusp conditions can be made to improve local behavior close to a nucleus. For example, in one approach [65], the region close to a nucleus is described by a Slater-type function, and a polynomial fit is used to connect the Gaussian region to the exponential. This procedure strongly reduces fluctuations of the kinetic energy of these functions, a desirable property for guided VMC and PMC.

An alternative form for the antisymmetric part of the trial wave-function, $\mathcal{D}$, is the *antisymmetrized geminal power* (AGP) form. It has recently been used in QMC calculations by Casula and Sorella [66]. The AGP function takes the form of

$$\mathcal{D}_{\mathrm{AGP}} \equiv \hat{A}\left[\Phi(\mathbf{r}_1^{\uparrow}, \mathbf{r}_1^{\downarrow})\Phi(\mathbf{r}_2^{\uparrow}, \mathbf{r}_2^{\downarrow}) \cdots \Phi(\mathbf{r}_{N/2}^{\uparrow}, \mathbf{r}_{N/2}^{\downarrow})\right], \tag{18}$$

where the geminal $\Phi(\mathbf{r}_i^{\uparrow}, \mathbf{r}_i^{\downarrow})$ is a two-electron singlet function,

$$\Phi(\mathbf{r}_i^{\uparrow}, \mathbf{r}_i^{\downarrow}) = \phi(\mathbf{r}_i^{\uparrow}, \mathbf{r}_i^{\downarrow}) \frac{1}{\sqrt{(2)}} \left(\alpha(1)\beta(2) - \alpha(2)\beta(1)\right). \tag{19}$$

The pair function $\phi(\mathbf{r}_i^{\uparrow}, \mathbf{r}_i^{\downarrow})$ is expanded in orbitals χ_i, *e. g.*

$$\phi(\mathbf{r}_i^{\uparrow}, \mathbf{r}_i^{\downarrow}) = \sum_i^M \lambda_i \chi_i(\mathbf{r}_i^{\uparrow}) \chi_i^*(\mathbf{r}_i^{\downarrow}), \tag{20}$$

where λ_i are variational parameters. The AGP wave function can be evaluated as the determinant of a matrix, and it can be generalized to include unpaired electrons. Furthermore, the use of more basis functions M than the number of particles N includes some types of static correlation effects, making a single AGP matrix equivalent to several Slater determinants. All the previous properties make the AGP wavefunction a promising ansatz for QMC calculations.

The symmetric part of the wave function is usually built as a product of terms explicitly dependent on inter-particle distance, $\mathbf{r}_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. These functions are usually constructed to reproduce the form of the wave function at electron–electron and electron–nucleus cusps. A now familiar form is that proposed by Bijl [67], Dingle [68], and Jastrow [69] and known as the Jastrow *ansatz*:

$$\mathcal{F} \equiv e^{U(\mathbf{r}_{ij})} \equiv e^{\sum_{i<j} g_{ij}}, \tag{21}$$

where the correlation function g_{ij} is

$$g_{ij} \equiv \frac{a_{ij}\mathbf{r}_{ij}}{1 + b_{ij}\mathbf{r}_{ij}} \tag{22}$$

with constants specified to satisfy the cusp conditions

$$a_{ij} \equiv \begin{cases} \frac{1}{4} & \text{if ij are like spins,} \\ \frac{1}{2} & \text{if ij are unlike spins,} \\ -\mathcal{Z} & \text{if ij are electron/nucleus pairs} \end{cases} \tag{23}$$

Electron correlation for parallel spins is taken into account by the Slater determinant.

This simple Slater–Jastrow *ansatz* has a number of desirable properties. As stated above, scaling with system size for the evaluation of the trial function is N^3, where N is the number of particles in the system, Second, the correct cusp conditions are satisfied at two-body coalescence points and the correlation function g_{ij} correctly approaches a constant at large distance.

The inclusion of 3-body correlation terms has been shown to improve wave function quality. The work of Huang *et al.* [70] shows that if the determinant parameters λ_D are optimized along with the correlation function parameters, λ_C, the nodal structure of the wave function does not significantly improve by including 4-body correlation terms. This finding suggests that increasing the number of determinants is more important than adding fourth- and higher-order correlation terms.

Feynman–Cohen backflow correlation functions [71] have been suggested by Schmidt and Moskowitz [72] for describing three body correlations of atoms

and molecules in U. These functions have been used in trial functions for homogeneous systems such as the electron gas [73, 74]. and liquid helium [75, 76]. The *backflow* concept is based on the conservation of particle current and the variational principle. The procedure involves replacing mean field orbitals by *backflow*-corrected orbitals of the form

$$\phi_n(\mathbf{r}_i) \rightarrow \phi_n\left(\mathbf{r}_i + \sum_{j \neq i} \mathbf{r}_{ij}\nu(\mathbf{r}_{ij})\right), \tag{24}$$

where $\nu(\mathbf{r}_{ij})$ is the backflow function. Pandharipande and Itoh [77] proposed that $\nu(r_{ij})$ should consist of the difference between the $l = 0$ and $l = 1$ states of an effective two-particle Schrödinger equation. Furthermore, Pandharipande *et al.* proposed [78] the inclusion of a $1/r^3$ tail, as originally suggested by Feynman and Cohen,

$$\nu(\mathbf{r}) = \lambda_\nu e^{-\left[\frac{\mathbf{r}_i - \mathbf{r}_j}{\omega_\nu}\right]^2} + \frac{\lambda_{\nu'}}{\mathbf{r}^3}, \tag{25}$$

where, λ_ν, $\lambda_{\nu'}$, and ω_ν are variational parameters. As recently noted by Kwon *et al.* [73] , the incorporation of the full backflow trial function into wave functions involves a power of N increase in computational expense, but yields a better DMC energy for the electron gas.[6]

A common practice is to use orbitals from a mean field calculation with the first term in a series expansion of the backflow contribution in the correlation function $\mathcal{F}$. The advantage of this approach is that orbitals are unperturbed and readily obtainable from mean field computer codes.

The Schmidt and Moskowitz correlation function [72] is a selection of terms from the general form originally proposed in connection with the transcorrelated method [79]:

$$\mathcal{F} = e^{\sum_{I,i<j} U_{Iij}}, \tag{26}$$

where

$$U_{Iij} = \sum_{k}^{N(I)} \Delta(m_{kI} n_{kI}) c_{kI} \left(g_{iI}^{m_{kI}} g_{jI}^{n_{kI}} + g_{jI}^{m_{kI}} g_{iI}^{n_{kI}}\right) g_{ij}^{o_{kI}}. \tag{27}$$

The sum in Eq. 26 goes over I nuclei, ij electron pairs, and the sum in Eq. 27 is over the $N(I)$ terms of the correlation function for each nucleus. The parameters m, n and o are integers. The function $\Delta(m, n)$ takes the value 1 when $m \neq n$, and $\frac{1}{2}$ otherwise. The functions g_{ij} are specified by Eq. 22.

This correlation function defined by Eqs. 26,27 can be shown to have contributions to averaged backflow effects from the presence of electron–electron–nucleus correlations that correspond to values of m, n and o in Eq. 27 of $2, 2, 0$ and $2, 0, 2$. These contributions recover $\approx 25\%$ or more of the total correlation energy of atomic and molecular systems above that from the simple Jastrow term [72].

[6] As discussed in Section 2.3.10, an improved fixed-node energy is a consequence of better nodes of the trial wave function, a critically important characteristic for importance sampling functions in QMC methods.

2.2 Variational Monte Carlo

2.2.1 Formalism

Variational methods involve the calculation of the expectation value of the Hamiltonian operator using a trial wave function Ψ_T. This function is dependent on a set of parameters, Λ, that are varied to minimize the expectation value of the energy, i.e.,

$$\langle \widehat{H} \rangle = \frac{\langle \Psi_T | \widehat{H} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \equiv E[\Lambda] \geq E_0. \tag{28}$$

Equation 28 can be sampled from a probability distribution proportional to Ψ_T^2, and evaluated from the expression

$$\frac{\int d\mathbf{R} \left[\frac{\widehat{H}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \right] \Psi_T^2(\mathbf{R})}{\int d\mathbf{R} \Psi_T^2(\mathbf{R})} \equiv \frac{\int d\mathbf{R} E_L \Psi_T^2(\mathbf{R})}{\int d\mathbf{R} \Psi_T^2(\mathbf{R})} \geq E_0, \tag{29}$$

where E_L is the local energy of Eq. 5. The procedure involves sampling random points in $\mathbf{R}$-space from

$$\mathcal{P}(\mathbf{R}) \equiv \frac{\Psi_T^2(\mathbf{R})}{\int d\mathbf{R} \Psi_T^2(\mathbf{R})}. \tag{30}$$

The advantage of using Eq.30 as the probability density function is that one need not perform the averaging of the numerator and denominator of Eq. 29. The calculation of the ratio of two integrals with the MC method is biased by definition: the average of a quotient is not equal to the quotient of the averages, and Eq. 30 avoids this problem.

In general, sampling is done using the Metropolis method [60], that is well described in Chapter 3 of Ref [53] , and briefly summarized later in this section.

Expectation values can be obtained using the VMC method from the following general expressions [5]:

$$\langle \widehat{O} \rangle \equiv \frac{\int d\mathbf{R} \Psi_T(\mathbf{R})^2 \widehat{O}(\mathbf{R})}{\int d\mathbf{R} \Psi_T(\mathbf{R})^2} \cong \frac{1}{N} \sum_{i=1}^{N} \widehat{O}(\mathbf{R}_i), \tag{31}$$

$$\langle \widehat{O}_d \rangle \equiv \frac{\int d\mathbf{R} \left[\frac{\widehat{O}_d \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \right] \Psi_T(\mathbf{R})^2}{\int d\mathbf{R} \Psi_T(\mathbf{R})^2} \cong \frac{1}{N} \sum_{i=1}^{N} \frac{\widehat{O}_d \Psi_T(\mathbf{R}_i)}{\Psi_T(\mathbf{R}_i)}. \tag{32}$$

Eq. 31 is for a coordinate operator, $\widehat{O}$, and Eq. 32 is preferred for a differential operator, $\widehat{O}_d$.

2.2.2 The generalized Metropolis algorithm

The main idea of the Metropolis algorithm is to sample the electron density, $\Psi_T^2(\mathbf{R})$, using fictitious kinetics that in the limit of large simulation time yields

the density at equilibrium. A coordinate move is proposed, $\mathbf{R} \to \mathbf{R}'$, which has the probability of being accepted given by

$$P(\mathbf{R} \to \mathbf{R}') = \min\left(1, \frac{T(\mathbf{R}' \to \mathbf{R})\Psi_T^2(\mathbf{R}')}{T(\mathbf{R} \to \mathbf{R}')\Psi_T^2(\mathbf{R})}\right), \tag{33}$$

where $T(\mathbf{R} \to \mathbf{R}')$ denotes the transition probability for a coordinate move from $\mathbf{R}$ to $\mathbf{R}'$. Condition 33 is necessary to satisfy the detailed balance condition

$$T(\mathbf{R}' \to \mathbf{R})\Psi_T^2(\mathbf{R}') = T(\mathbf{R} \to \mathbf{R}')\Psi_T^2(\mathbf{R}) \tag{34}$$

which is needed for $\Psi_T^2(\mathbf{R})$ to be the equilibrium distribution of the sampling process.

Several improvements to the Metropolis method have been pursued both in classical and in quantum simulations. These improvements involve new transition probability functions and other sampling procedures. See, for example, refs. [5, 25, 80–85].

A common approach for improving $T(\mathbf{R} \to \mathbf{R}')$ in VMC, is to use the quantum force,

$$\mathbf{F_q} \equiv \nabla \ln\left|\Psi_T(\mathbf{R})^2\right| \tag{35}$$

as a component of the transition probability. The quantum force can be incorporated by expanding $f(\mathbf{R}, \tau) = |\Psi_T(\mathbf{R})^2| = e^{-\ln|\Psi_T^2(\mathbf{R})|}$ in a Taylor series in $\ln|\Psi_T^2(\mathbf{R})|$ and truncating at first order

$$T(\mathbf{R} \to \mathbf{R}') \approx \frac{1}{N} e^{\lambda \mathbf{F_q}(\mathbf{R})\cdot(\mathbf{R}'-\mathbf{R})}, \tag{36}$$

where N is a normalization factor, and λ is a parameter fixed for the simulation or optimized in some fashion; see, for example, ref. [86]. A usual improvement is to introduce a cutoff in $\Delta\mathbf{R} = (\mathbf{R}' - \mathbf{R})$, so that if the proposed displacement is larger than a predetermined measure, the move is rejected.

A good transition probability should also contain random displacements, so that all of phase space can be sampled. The combination of the desired drift arising from the quantum force of Eq. 36 with a Gaussian random move, gives rise to Langevin fictitious dynamics, namely,

$$\mathbf{R}' \to \mathbf{R} + \frac{1}{2}\mathbf{F_q}(\mathbf{R}) + \mathcal{G}_{\delta\tau}, \tag{37}$$

where $\mathcal{G}_{\delta\tau}$ is a number sampled from a Gaussian distribution with standard deviation $\delta\tau$. The propagator or transition probability for Eq. 37 is

$$T_L(\mathbf{R} \to \mathbf{R}') = \frac{1}{\sqrt{4\pi D\delta\tau}^{3N}} e^{-(\mathbf{R}'-\mathbf{R}-\frac{1}{2}\mathbf{F_q}(\mathbf{R})\delta\tau)^2/2\delta\tau} \tag{38}$$

which is a drifting Gaussian, spreading in $\delta\tau$. Using Eq. 37 is equivalent to finding the solution of the Fokker–Planck equation [30]

$$\frac{\partial f(\mathbf{R}, \tau)}{\partial \tau} = \frac{1}{2}\nabla \cdot (\nabla - \mathbf{F_q}) f(\mathbf{R}, \tau). \tag{39}$$

Equation 38 has proved to be a simple and effective choice for a VMC transition probability. More refined choices can be made, usually with the goal of increasing acceptance probabilities in regions of rapid change in $|\Psi_T(\mathbf{R})^2|$, such as close to a nucleus. For a more detailed discussion of this formalism, the reader is directed to Chapter 2 of [11]. More elaborate transition rules can be found in refs. [86–89].

2.2.3 Variational Monte Carlo algorithm

The VMC algorithm is an application of the generalized Metropolis MC method. As in most applications of the method, one needs to insure that the ensemble has achieved equilibrium in the simulation sense. Equilibrium is reached when the ensemble $\mathcal{W}$ is distributed according to Eq. 30. This is usually achieved by performing a Metropolis random walk and monitoring the trace of the observables of interest. When the trace fluctuates around a mean, it is generally safe to start averaging in order to obtain desired properties.

An implementation of the VMC algorithm follows:

1. Equilibration stage
   (a) Generate an initial set of random walker positions, $\mathcal{W}_0$; it can be read in from a previous random walk, or generated at random.
   (b) Perform a loop over N_s steps,
      i. For each r_i of the N_p number of particles,
         A. Propose a move from $\Psi(\mathbf{R}) \equiv \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_{N_p})$ to $\Psi(\mathbf{R}') \equiv \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}'_i, \ldots, \mathbf{r}_{N_p})$. Move from $\mathbf{r}$ to $\mathbf{r}'$ according to

$$\mathbf{r}' \leftarrow \mathbf{r} + \mathcal{G}_{\delta\tau} + \frac{1}{2}\mathbf{F}_\mathbf{q}\delta\tau, \tag{40}$$

            where $\mathcal{G}_{\delta\tau}$ is a Gaussian random number with standard deviation $\delta\tau$,which is a proposed step size, and $\mathbf{F}_\mathbf{q}$ is the quantum force; see Eqs . 36 and Eq. 37..
         (B) Compute the Metropolis acceptance/rejection probability

$$P(\mathbf{R} \to \mathbf{R}') = \min\left(1, \frac{T_L(\mathbf{R}' \to \mathbf{R})\Psi_T^2(\mathbf{R}')}{T_L(\mathbf{R} \to \mathbf{R}')\Psi_T^2(\mathbf{R})}\right), \tag{41}$$

            where T_L is given by Eq. 38.
         (C) Compare $P(\mathbf{R} \to \mathbf{R}')$ with a uniform random number between 0 and 1, $\mathcal{U}_{[0,1]}$. If $P > \mathcal{U}_{[0,1]}$, accept the move, otherwise, reject it.
         (D) Calculate the contribution to the averages $\frac{\widehat{O}_d\Psi_T(\mathbf{R}')}{\Psi_T(\mathbf{R}')}$, and perform blocking statistics as described in Section 2.2.4.
      (ii) Continue the loop until the desired accuracy is achieved.

2.2.4 Statistics

Usually, VMC calculations are performed using an ensemble of $N_{\mathcal{W}}$ random walkers $\mathcal{W} \equiv \{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_\mathbf{N}\}$ that are propagated following $T(\mathbf{R} \rightarrow \mathbf{R}')$ using the probability $P(\mathbf{R} \rightarrow \mathbf{R}')$ to accept or reject proposed moves of ensemble members. Usually, the displacements are small enough that the MC observables evaluated at $\mathbf{R}'$ are statistically correlated with those evaluated at $\mathbf{R}$. This autocorrelation has to be taken into account for estimating the error of the calculation. The variance for an observable, $\widehat{O}$, measured over N_s MC steps of a random walk is

$$\sigma_{\widehat{O}} \equiv \frac{1}{N_S N_{\mathcal{W}}} \big(O_i - \langle O \rangle\big), \tag{42}$$

where $\langle O \rangle$ is the average of the observations, O_i, over the sample. A simple approach to remove auto-correlation between samples is to define a number of blocks, N_B, where each block is an average of N_S steps, with variance

$$\sigma_B \equiv \frac{1}{N_B N_{\mathcal{W}}} \big(O_b - \langle O \rangle\big), \tag{43}$$

where O_b are the N_B observations in block b. If N_B is sufficiently large, σ_B is a good estimator of the variance of the observable over the random walk. The auto-correlation time is a good measure of computational efficiency, and is given by

$$T_{\text{corr}} = \lim_{N_s \rightarrow \infty} N_s \left(\frac{\sigma_B^2}{\sigma_{\widehat{O}}^2} \right). \tag{44}$$

The efficiency of a method depends on the time step [90]. Serial correlation between sample points should vanish for an accurate estimator of the variance. For an observable $\langle O \rangle$, the serial correlation coefficient is defined as

$$\xi_k \equiv \frac{1}{(\langle O^2 \rangle - \langle O \rangle^2)(N-k)} \sum_{i=1}^{N-k} \big(O_i - \langle O \rangle\big)\big(O_{i+k} - \langle O \rangle\big), \tag{45}$$

decays exponentially with k. The correlation length, L, is defined as the number of steps necessary for ξ_k to decay to zero within statistical error. For an accurate variance estimator, blocks should be of at least L steps.

The efficiency of a simulation is inversely proportional to ξ_k. The ξ_k dependence on time step is usually strong [11]; the larger the time step, the fewer steps/block L necessary, and the more points available for calculating the global average $\langle O \rangle$. A rule of thumb is to use an $N_t \approx 10$ times larger than the auto-correlation time to insure statistical independence of block averages, and therefore a reliable variance estimate.

The VMC method shares some of the strengths and weaknesses of traditional variational methods: the energy is an upper bound to the true ground state energy. If reasonable trial functions are used, often reliable estimates of properties can often be obtained.

2.2.5 Trial Wave function optimization

Trial wave functions $\Psi_T(\mathbf{R}, \Lambda)$ for QMC are dependent on variational parameters $\Lambda = \{\lambda_1, \ldots, \lambda_n\}$. Optimization of Λ is a key element for obtaining accurate trial functions. Importance sampling using an optimized trial function increases the efficiency of DMC simulations. There is a direct relationship between trial-function accuracy and the computer time required to calculate accurate expectation values. Some of the parameters λ_i may be fixed by imposing appropriate wave function properties, such as cusp conditions (see Section 1.2).

It is useful to divide Λ into groups distinguished by whether the parameter affects the nodes of the wave function. The Slater determinant parameters, $\lambda_D^{\uparrow\downarrow}$ and the Slater determinant weights, λ_{k_i} change wave function nodal structure [72, 91–96]. The correlation function parameters, λ_C do not change the nodal structure of the overall wave function, and therefore the DMC energy. For some systems, the optimization of λ_C is sufficient for building reliable trial functions for PMC methods, because the correlation function is designed in part to satisfy cusp conditions [44, 45].

Several optimization methods have been proposed previously. Some involve the use of analytical derivatives [62, 97–101], others use of a fixed sample for variance minimization [102], and more recently [91, 103, 104]. Histogram analysis can be useful for the optimization of energy, variance and geometries [105].

The variance functional (VF) [91] is given by

$$VF = \frac{\sum_{i=1}^{N} \left[\frac{\widehat{H}\Psi(\mathbf{R}_i, \Lambda)}{\Psi(\mathbf{R}_i, \Lambda)} - E_T \right]^2 w_i}{\sum_{i=1}^{N} w_i}, \tag{46}$$

where E_T is a trial energy, w_i is a weighting factor defined by

$$w_i(\Lambda) = \frac{\Psi^2(\mathbf{R}_i, \Lambda)}{\Psi^2(\mathbf{R}_i, \Lambda_0)}, \tag{47}$$

and Λ_0 is an initial set of parameters. The sum in Eq. 46 is over configurations initially distributed as $\Psi^2(\mathbf{R}, \Lambda_0)$. Numerical optimization methods are used to find the minimum of VF.

2.2.6 Optimization of the full trial wave function

Recently, several new approaches have appeared for optimizing the determinantal (single or linear combination) part of the wave function simultaneously with a correlation function. Exploration of the full variational freedom of these wave functions shows that significant improvements can be achieved.

The method of Filippi and Fahy [106] combines variance minimization of the correlation function as given by Eq. 46 with the use of a self-consistent field method that takes into account the effect of the correlation function, $\mathcal{F}$. These effects are introduced by means of an energy fluctuation potential (EFP) constructed from discrete R-space configurations. The EFP is represented as

an external potental V_{ext} in the self-consistent-field procedure for the orbitals. Iterative refinement of $\mathcal{F}$ by variance optimization followed by generation of new orbitals using a Hamiltonian containing V_{ext} converges to a set of optimal orbital parameters and determinantal mixing parameters for the given wave function form.

The EFP is obtained from a least squares fit of the local energy to the function,

$$E_R + \sum_{k=1}^{n} \lambda_k \frac{\mathcal{O}_k(\mathcal{D}(\mathbf{R}_i))}{D(\mathbf{R})_i} \tag{48}$$

based on configurations $\mathbf{R}_i$ sampled from $\Psi_T(\mathbf{R})^2$. Here, E_R and λ_k are fitting parameters. The operators $\mathcal{O}_k$ can be either local or non-local, and are coupled to the *ab initio* external potential, e.g.,

$$V_{ext} = V_0 + \sum_{k=1}^{n} \lambda_k \mathcal{O}_k, \tag{49}$$

where V_0 is the external potential of traditional ab initio methods. The wave functions obtained from the EFP method provide excellent guiding functions for DMC.

A similar method was used by Schautz and Fahy [107] to optimize CI coefficients of a truncated multideterminant wave-function. An effective Hamiltonian is constructed by optimizing a least squares fit of the local energy to the function,

$$E_R + \sum_{k=0}^{n} \lambda_k |\Phi_k\rangle\langle\Phi_0| \tag{50}$$

Here, Φ_k are CI determinants, and λ_k and E_R are fitting parameters. The λ_k are used to construct an effective Hamiltonian for the next CI iteration,

$$\hat{H}_{eff} = \hat{H}^0_{eff} + \sum_{k=0}^{n} \lambda_k |\Phi_k\rangle\langle\Phi_0| + \lambda_k^* |\Phi_0\rangle\langle\Phi_k|. \tag{51}$$

Least squares fits of Eqs. 48 and 50 are equivalent to solving a system of stochastic linear equations. For the exact details of the numerical procedure, the reader is referred to the original references [106, 107].

An alternative method for optimizing the full wave function is the use of the stochastic reconfiguration (SR) technique of Sorella [108] and Casula and Sorella[66]. In this approach, the wave function is expanded as Taylor series in the parameters $\Lambda = \{\lambda_1, \ldots, \lambda_n\}$, with $\lambda_k = \lambda_k^0 + \delta\lambda_k$. The series truncated to first order yields,

$$\Psi_T(\Lambda) = \Psi_T(\Lambda^0) + \sum_{k=1}^{n} \delta\lambda_k \frac{\partial \Psi_T(\Lambda^0)}{\partial_{\lambda_k}}. \tag{52}$$

A set of local operators $O^k(\mathbf{R})$ is defined as logarithmic derivatives with respect to the variational parameters,

$$O^k(\mathbf{R}) = \frac{1}{\Psi_T(\mathbf{R}, \Lambda)} \frac{\partial}{\partial \lambda_k} \Psi_T(\mathbf{R}, \Lambda). \tag{53}$$

A renormalized function Ψ_T' is expanded in the subspace of dimension $n+1$, i.e.,

$$|\Psi_T'(\Lambda)\rangle = \sum_{k=0}^{n} \delta\lambda_k O^k |\Psi_T\rangle, \tag{54}$$

where $O^0 = 1$, $\delta\lambda_0 = 1$ and $\delta\lambda_{k>0} \equiv \delta\lambda_k/\delta\lambda_0$.

A projection operator $P_{SR}(E_{SR} - \hat{H})$ is constructed that projects the function on to the space defined by Eq. 54,

$$|\Psi_T'(\Lambda)\rangle = \hat{P}_{SR}(E_{SR} - \hat{H})|\Psi_T\rangle. \tag{55}$$

Here, E_{SR} is a large energy shift used to insure Ψ_T' has a lower energy than Ψ_T. The coefficents $\delta\Lambda \equiv \{\delta\lambda_1, \delta\lambda_2, \ldots, \delta\lambda_n\}$ that correspond to the projected wave function Ψ_T', are found by solving,

$$\sum_l \delta\lambda_l \langle \Psi_T | O^l O^k | \Psi_T \rangle = \langle \Psi_T | O^k (E_{SR} - \hat{H}) | \Psi_T \rangle. \tag{56}$$

With the solutions of Eq. 56, the Λ_k are updated, i.e., $\Lambda_k = \Lambda_k^{(0)} + \delta\Lambda/\delta\Lambda_0$. Iterations are continued until $\delta\lambda_k/\delta\lambda_0 \to 0$. Casula and Sorella [66] note that this system of equation is similar to that derived from Eqs. 48 and 50.

Estimating the overlap of Ψ_T with the ground state wave function, $\langle \Psi_T | \Psi_0 \rangle$, by DMC methods [109] is a very efficient way of assessing wave function quality. There is also a trend that correlates the variational energy of the wave function with the associated variance in a linear relationship [73, 74]. This correlation is expected because the quantities, δ_E, and δ_{E_L}, approach the limits – E_0 and zero, respectively – as wave function quality improves. Observing these quantities during optimization provides a good method of validating the optimization method, as well as assessing wave function quality.

2.3 Projector methods

The QMC approaches of DMC and GFMC are usefully called projector Monte Carlo (PMC) methods.[7] The general idea is to project out a state of the Hamiltonian by iteration of a projection operator $\widehat{P}$. For simplicity, we assume that the desired state is the ground state Ψ_0, but projectors can be constructed for any state,

$$\lim_{i \to \infty} \widehat{P}^i |\Psi_T\rangle \approx |\Psi_0\rangle. \tag{57}$$

[7] This presentation follows refs. [110–112] .

After sufficient iterations i, the contribution of all excited states $|\Psi_{n>0}\rangle$ will be filtered out, and only the ground state is recovered.

If $|\Psi_T\rangle$ is a vector and $\widehat{P}$ is a matrix, then the procedure implied by Eq. 57 is the algebraic power method: If a matrix is applied iteratively to an initial arbitrary vector for a sufficient number of times, only the dominant eigenvector $|\Psi_0\rangle$ will survive. For large i, one has

$$\widehat{P}^i|\Psi_T\rangle = \lambda_0^i \langle\Psi_0|\Psi_T\rangle|\Psi_0\rangle + \mathcal{O}\big(\lambda_1^i\big), \tag{58}$$

where λ_0 is the leading eigenvalue, and λ_1 is the largest sub-leading eigenvalue.

For this approach, it is possible to obtain an estimator of the eigenvalue [51], given by

$$\lambda_0 = \lim_{i\to\infty} \left(\frac{\langle\phi|\widehat{P}^{i+j}|\Psi_T\rangle}{\langle\phi|\widehat{P}^{i}|\Psi_T\rangle} \right)^{\frac{1}{m}}. \tag{59}$$

2.3.1 Markov processes and stochastic projection

For high-dimensional vectors, such as those encountered in molecular electronic structure, the algebraic power method described in the previous section needs to be generalized with stochastic implementation. For this to occur, the projection operator must be symmetric, so that all eigenvalues are real. This is the case for QMC methods because $\widehat{P}$ is a function of the Hamiltonian operator, $\widehat{H}$, which is Hermitian.

A stochastic matrix is a normalized nonnegative matrix. By normalization, we mean that the stochastic matrix columns sum to one, i.e., $\sum_i M_{ij} = 1$. A $\mathbf{R}$-space representation of the normalization condition is a stochastic propagator $M(\mathbf{R},\mathbf{R}')$ that satisfies the condition

$$\int M\big(\mathbf{R},\mathbf{R}'\big)\, d\mathbf{R}' = 1. \tag{60}$$

A Markov chain is a sequence of states obtained from subsequent transitions from state i to state j with a probability related to the stochastic matrix element M_{ij}, or in $\mathbf{R}$-space, $M\big(\mathbf{R},\mathbf{R}'\big)$. Note that any given move only depends on the current state, i. For example, in $\mathbf{R}$-space, this is equivalent to the process

$$\begin{aligned} \pi(\mathbf{R}') &= \int M\big(\mathbf{R}',\mathbf{R}''\big)\pi\big(\mathbf{R}''\big)\, d\mathbf{R}'', \\ \pi(\mathbf{R}) &= \int M\big(\mathbf{R},\mathbf{R}'\big)\pi\big(\mathbf{R}'\big)\, d\mathbf{R}', \\ &\cdots \end{aligned} \tag{61}$$

The sequence of states $S = \{\pi(\mathbf{R}''), \pi(\mathbf{R}'), \pi(\mathbf{R}), \ldots\}$ is the Markov chain.

The propagators of QMC for electronic structure are not generally normalized, therefore they are not stochastic matrices, but one can represent them in terms of the latter by factoring the projection operator as follows,

$$\widehat{P}_{ij} = M_{ij} w_j, \tag{62}$$

where the weights, w_j, are defined by $w_j = \sum_i \widehat{P}_{ij}$. This definition unambiguously defines both the associated stochastic matrix M and the weight vector w.

A MC sampling scheme of $\widehat{P}_{ij}|\Psi_T\rangle$ can be generated by first performing a random walk, and then retaining a weight vector $W(\mathbf{R})$ of the random walkers,

$$\begin{aligned}
\Psi(\mathbf{R}') &\equiv \pi(\mathbf{R}')W(\mathbf{R}') = \int P(\mathbf{R}', \mathbf{R}'')\Psi(\mathbf{R}'')\, d\mathbf{R}'' \\
&= \int M(\mathbf{R}', \mathbf{R}'')B(\mathbf{R}')\Psi(\mathbf{R}'')\, d\mathbf{R}'', \\
\Psi(\mathbf{R}) &\equiv \pi(\mathbf{R})W(\mathbf{R}) = \int P(\mathbf{R}, \mathbf{R}')\Psi(\mathbf{R}')\, d\mathbf{R}' \\
&= \int M(\mathbf{R}, \mathbf{R}')B(\mathbf{R})\Psi(\mathbf{R}')\, d\mathbf{R}', \qquad (63)
\end{aligned}$$

Here, $B(\mathbf{R})$ is the function that determines the weight of the configurations at each state of the random chain. This leads to a generalized stochastic projection algorithm for unnormalized transition probabilities that forms the basis for population Monte Carlo (PopMC) algorithms, which are are also used for statistical information processing and robotic vision [113]. A generalized PopMC stochastic projection algorithm, represented in $\mathbf{R}$-space, follows:

1. *Initialize*
   Generate a set of $\mathbf{n}$ random walkers, located at different spatial positions, $\mathcal{W} \equiv \{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R_n}\}$, where $\mathbf{R}_i$ denotes a Dirac delta function at that point in space, $\delta(\mathbf{R} - \mathbf{R}_i)$. These points are intended to sample a probability density function $\Phi(\mathbf{R})$.

2. *Move*

   (a) Each walker j is moved independently from $\mathbf{R}$ to a new position $\mathbf{R}'$, according to the transition probability

   $$T(\mathbf{R} \to \mathbf{R}') \equiv M(\mathbf{R}, \mathbf{R}'). \qquad (64)$$

   (b) Ensure detailed balance if $T(\mathbf{R} \to \mathbf{R}') \neq T(\mathbf{R}' \to \mathbf{R})$ by using a Metropolis acceptance/rejection step as in Eq. 41.

3. *Weight*

   (a) Calculate a weight vector using a weighting function $B(\mathbf{R}_i)$,

   $$w_i^* = B(\mathbf{R}_i). \qquad (65)$$

   The ideal weight function preserves normalization of $\widehat{P}(\mathbf{R}, \mathbf{R}')$ and maintains individual weights w_i close to unity.

   (b) Update the weight of the walker, multiplying the weight of the previous iteration by the weight of the new iteration,

   $$w_i' = w_i^* w_i. \qquad (66)$$

4. *Reconfiguration*

   (a) Split walkers with large weights into multiple walkers with weights that add up to the original weight.

   (b) Remove walkers with small weight.

Step 4 is necessary to avoid statistical fluctuations in the weights. It is a form of importance sampling that makes the calculation stable over time. Some algorithms omit this step; see, for example, the work of Caffarel and Claverie [61], but it has been proved that such calculations eventually diverge [114]. There is a slight bias associated with the introduction of step 4 together with population control methods, that will be discussed in Section 2.3.7. When step 4 is used, $B(\mathbf{R})$ is also referred in the literature as a *branching factor*.

It is important to recall that PopMC algorithms are not canonical Markov Chain Monte Carlo (MCMC) algorithms [58, 115], in the sense that the propagator used is not normalized, and therefore factoring the propagator into a normalized transition probability and a weighting function is required.

2.3.2 Projection operators or Green's functions

Different projection operators lead to different QMC methods. If the resolvent operator,

$$\widehat{P}(\widehat{H}) \equiv \frac{1}{1+\delta\tau(\widehat{H}-E_R)}, \tag{67}$$

is used, one obtains Green's function Monte Carlo (GFMC) [39, 116]. This algorithm will be described in Section 2.3.9. If the imaginary time evolution operator is used, i.e.,

$$\widehat{P}(\widehat{H}) \equiv e^{-\delta\tau(\widehat{H}-E_R)}, \tag{68}$$

one has the DMC method [117, 118], which is discussed in the following Section 2.3.5.

For finite $\delta\tau$, and for molecular systems, the exact projector is not known analytically. In GFMC, the resolvent of Eq. 67 is sampled by iteration of a simpler resolvent, whereas for DMC, the resolvent is known exactly as $\tau \to 0$, so an extrapolation to $\delta\tau \to 0$ is done.

Note that any decreasing function of $\widehat{H}$ can serve as a projector. Therefore the possibility of new QMC methods remains.

2.3.3 Imaginary time propagator

If one transforms the time-dependent Schrödinger equation (Eq. 2) to imaginary time τ, i.e.,

$$it \to \tau, \tag{69}$$

then one obtains, after introducing an energy offset E_R as reference energy,

$$\frac{\partial}{\partial\tau}\Psi(\mathbf{R},\tau) = (\widehat{H}-E_R)\Psi(\mathbf{R},\tau). \tag{70}$$

For real $\Psi(\mathbf{R},\tau)$, Eq. 70 has the advantage of being in $\mathcal{R}^N$, whereas Eq. 2 has, in general, complex solutions.

Equation 70 can be cast into integral form,

$$\Psi(\mathbf{R},\tau+\delta\tau) = \lambda_\tau \int G\big(\mathbf{R},\mathbf{R}',\delta\tau\big)\Psi\big(\mathbf{R}',\tau\big)\,d\mathbf{R}'. \tag{71}$$

The Green's function, $G(\mathbf{R}',\mathbf{R},\delta\tau)$, satisfies the same boundary conditions as Eq. 70:

$$\frac{\partial}{\partial\tau}G\big(\mathbf{R},\mathbf{R}',\delta\tau\big) = \big(\widehat{H}-E_T\big)G\big(\mathbf{R},\mathbf{R}',\delta\tau\big) \tag{72}$$

with the initial conditions associated with the propagation of a Dirac delta function, namely,

$$G\big(\mathbf{R},\mathbf{R}',0\big) = \delta\big(\mathbf{R}-\mathbf{R}'\big). \tag{73}$$

The form of the Green's function that satisfies Eq. 72, subject to Eq. 73, is

$$G\big(\mathbf{R},\mathbf{R}',\delta\tau\big) = \langle\mathbf{R}|e^{-\tau(\widehat{H}-E_R)}|\mathbf{R}'\rangle \tag{74}$$

which can be expanded in eigenfunctions, Ψ_α, and eigenvalues E_α of the system, i.e.,

$$G\big(\mathbf{R},\mathbf{R}',\delta\tau\big) = \sum_\alpha e^{-\tau(E_\alpha-E_R)}\Psi^*_\alpha\big(\mathbf{R}'\big)\Psi_\alpha(\mathbf{R}). \tag{75}$$

For an arbitrary initial trial function, $\Psi(\mathbf{R})$, as $\tau\to\infty$, one has

$$\begin{aligned}\lim_{\tau\to\infty} e^{-\tau(\widehat{H}-E_T)}\Psi &= \lim_{\tau\to\infty}\int G\big(\mathbf{R}',\mathbf{R},\tau\big)\Psi\big(\mathbf{R}'\big)\,d\mathbf{R}' \\ &= \lim_{\tau\to\infty}\langle\Psi|\Psi_0\rangle e^{-\tau(E_0-E_R)}\Psi_0,\end{aligned} \tag{76}$$

and only the ground state wave function Ψ_0 is obtained from any initial wave function. Therefore, the imaginary time evolution operator can be used as a projection operator as stated at the beginning of this section.

2.3.4 Diffusion Monte Carlo stochastic projection

Due to the high dimensionality of molecular systems, a MC projection procedure is used to obtain expectation values. In this approach, the wave function is represented as an ensemble of delta functions, also known as configurations, walkers, or *psips* (psi-particles),

$$\Psi(\mathbf{R}) \longleftrightarrow \sum_k \delta(\mathbf{R}-\mathbf{R_k}). \tag{77}$$

The wave function is propagated in imaginary time using a Green's function. In the continuous case, one can construct a Neumann series

$$\begin{aligned}\Psi^{(2)}(\mathbf{R},\tau) &= \lambda_1\int G\big(\mathbf{R}',\mathbf{R},\tau_2-\tau_1\big)\Psi^{(1)}\big(\mathbf{R}'\big)\,d\mathbf{R}, \\ \Psi^{(3)}(\mathbf{R},\tau) &= \lambda_2\int G\big(\mathbf{R}',\mathbf{R},\tau_3-\tau_2\big)\Psi^{(2)}\big(\mathbf{R}'\big)\,d\mathbf{R}, \\ &\cdots\end{aligned} \tag{78}$$

Equation 78 is a specific case of the PopMC propagation of Section 2.3.1. A discrete Neumann series can be constructed in a similar way,

$$\Psi^{(n+1)}(\mathbf{R}, \tau + \delta\tau) \longleftrightarrow \lambda_k \sum_k G^{(n)}\big(\mathbf{R}, \mathbf{R}', \delta\tau\big). \tag{79}$$

A stochastic vector of configurations $\mathcal{W} \equiv \{\mathbf{R}_1, \ldots, \mathbf{R}_n\}$ is used to represent $\Psi(\mathbf{R})$ and is iterated using $G^{(n)}(\mathbf{R}, \mathbf{R}', \delta\tau)$.

2.3.5 The form of the propagator

Sampling Eq. 74 can not be done exactly, because the argument of the exponential is an operator composed of two non-commuting terms. In practice, approximate forms of the propagator are used.

In the short-time approximation (STA), the propagator $G(\mathbf{R}, \mathbf{R_k}, d\tau)$ is approximated as if the kinetic and potential energy operators commuted with each other, i.e.,

$$e^{(T+V)\delta\tau} \approx e^{T\delta\tau} \cdot e^{V\delta\tau} + \mathcal{O}\big((\delta\tau)^2\big) \equiv G_{ST} \equiv G_D \cdot G_B. \tag{80}$$

The Green's function becomes the product of a diffusion factor G_D and a branching factor G_B. Both propagators are known:

$$G_D = (2\pi\tau)^{-3N/2} e^{-\frac{(\mathbf{R}-\mathbf{R}')^2}{2\tau}} \tag{81}$$

and

$$G_B = e^{-\delta\tau(V(\mathbf{R})-2E_T)}. \tag{82}$$

G_D is a fundamental solution of a Fourier equation that describes a diffusion process in wave function space, and G_B is the fundamental solution of a first-order kinetic birth–death process.

The Campbell–Baker–Hausdorff (CBH) formula,

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]+\frac{1}{12}[(A-B),[A,B]]+\cdots} \tag{83}$$

makes possible more accurate decompositions, such as an expansion with a cubic error $\mathcal{O}((\delta\tau)^3)$,

$$e^{\delta\tau(T+V)} = e^{\delta\tau(V/2)} e^{\delta\tau T} e^{\delta\tau(V/2)} + \mathcal{O}\big((\delta\tau)^3\big). \tag{84}$$

There are also more sophisticated second-order [111] and fourth-order [119, 120] expansions that reduce the error considerably and generate higher accuracy DMC algorithms at the expense of a more complex propagator.

The most common implementation uses G_D as a stochastic transition probability $T(\mathbf{R} \to \mathbf{R}')$, and G_B as a weighting or branching factor, $B(\mathbf{R})$. Sampling Eq. 81 can be achieved by obtaining random variates from a Gaussian distribution of standard deviation δ_τ.

2.3.6 Importance sampling

Direct application of the algorithm of the previous section to systems governed by the Coulomb potential leads to large population fluctuations. These arise because the potential $\widehat{V}(\mathbf{R})$ can become unbounded and induce large fluctuations in the random walker population. A remedy, importance sampling, was first used for GFMC by Kalos [39] and extended to the DMC method by Ceperley and Alder [25].

In importance sampling Monte Carlo, the goal is to reduce fluctuations, by multiplying the probability distribution by a known trial function, $\Psi_T(\mathbf{R})$, that is expected to be a good approximation for the wave function of the system. Rather than $\Psi(\mathbf{R},\tau)$, one samples the product

$$f(\mathbf{R},\tau) = \Psi_T(\mathbf{R})\Psi(\mathbf{R},\tau). \tag{85}$$

Multiplying Eq. 71 by $\Psi_T(\mathbf{R})$ yields

$$f(\mathbf{R},\tau+d\tau) = \int K\left(\mathbf{R}',\mathbf{R},\delta\tau\right) f\left(\mathbf{R}',\tau\right) d\mathbf{R}', \tag{86}$$

where $K(\mathbf{R},\mathbf{R}',\delta\tau) \equiv e^{-\tau(\widehat{H}-E_T)}\frac{\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R}')}$. Expanding K in a Taylor series gives

$$K = Ne^{-(\mathbf{R}_2-\mathbf{R}_1+\frac{1}{2}\nabla \ln \Psi_T(\mathbf{R}_1)\delta\tau)^2/(2\delta\tau)} \times e^{-(\frac{\widehat{H}\Psi_T(\mathbf{R}_1)}{\Psi_T(\mathbf{R}_1)}-E_T)\delta\tau} \equiv K_D \times K_B. \tag{87}$$

Equation 87 is closely associated with the product of the kernel of the Smoluchowski equation, which describes a diffusion process with drift, multiplied by a first-order rate process. Here the rate process is dominated by the local energy, instead of the potential. The random walk is modified by appearance of a drift term that moves configurations to regions of large values of the wave function. This drift is the quantum force of Eq. 35.

The excess local energy $(E_T - E_L(\mathbf{R}))$ replaces the excess potential energy in the branching term exponent, see Eq. 82. The local energy has kinetic and potential energy contributions that tend to cancel each other, giving a smoother function. If $\Psi_T(\mathbf{R})$ is a reasonable function, the excess local energy will be nearly a constant. The regions where charged particles meet remain finite by enforcing the cusp conditions on $\Psi_T(\mathbf{R})$ (see Section 1.2).

The local energy is an estimator of the energy with a low statistical variance, so it is preferred over other possible choices for an estimator. A simple average of the local energy yields the estimator of the energy of the quantum system,[8]

$$\begin{aligned}
\langle E_L\rangle &= \int f(\mathbf{R},\tau\to\infty)E_L(\mathbf{R})\,d\mathbf{R}\Big/\int f(\mathbf{R})\,d\mathbf{R} \\
&= \int \Psi(\mathbf{R})\Psi_T(\mathbf{R})\left[\frac{\widehat{H}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}\right] d\mathbf{R}\Big/\int \Psi(\mathbf{R})\Psi_T(\mathbf{R}) \\
&= \int \Psi(\mathbf{R})\widehat{H}\Psi(\mathbf{R})\,d\mathbf{R}\Big/\int \Psi(\mathbf{R})\Psi_T(\mathbf{R})\,d\mathbf{R} \\
&= E_0.
\end{aligned} \tag{88}$$

[8] For other energy estimators, see the discussion in refs. [11, 116]

Simple averaging of the local energy yields the DMC energy estimator,

$$\langle E_L \rangle = \lim_{N_s \to \infty} \frac{1}{N_s} \sum_i^{N_w} E_L(\mathbf{R}_i). \tag{89}$$

Because the importance sampled propagator, $K(\mathbf{R}, \mathbf{R}', \delta\tau)$, is only exact to a certain order, for an exact estimator, it is necessary to extrapolate to $\delta\tau = 0$ for several values of $\langle E_L \rangle$.

Importance sampling with appropriate trial functions, such as those used for accurate VMC calculations, can increase the efficiency of the random walk by several orders of magnitude. In the limit of the exact trial function as the importance sampling function, only a single evaluation of the local energy yields the exact answer. Importance sampling has made atomic, molecular and nanostructure calculations feasible. Note that the quantum force of Eq. 87 also moves random walkers away from nodal regions to regions of large values of the trial wave function, reducing the number of attempted node crossings by typically several orders of magnitude.

2.3.7 Population control

If left uncontrolled, the population of random walkers will eventually vanish or fill all computer memory. Therefore, some form of population control is needed to stabilize the number of random walkers. Control is usually achieved by slowly changing E_T as the simulation progresses. As more walkers are produced in the procedure, one needs to lower the trial energy E_T or, if the population starts to decrease, then one needs to raise E_T. This can be achieved by periodically changing the trial energy in the appropriate direction. One version of the adjustment is to use

$$E_T = \langle E_0 \rangle + \alpha \ln \frac{N_w^0}{N_w}, \tag{90}$$

where $\langle E_0 \rangle$ is the best approximation to the eigenvalue of the problem to this point, α is a parameter that should be as small as possible while still having a population control effect, N_w^0 is the number of desired random walkers, and N_w is the current number of random walkers.

This simple population control procedure has a slight bias if the population control parameter α is large, or if the population is small. The bias observed goes as $1/N_w$, and, formally a $N_w \to \infty$ extrapolation is required. Bias is absent in the limit of an infinite population.

A recently resurrected population control strategy, stochastic reconfiguration [114, 121–123] follows the work of Hetherington [110]. In this algorithm, walkers carry a weight, but the weight is recomputed at each step to keep the population constant. The idea behind this method is to control the global weight $\overline{w}$ of the population,

$$\overline{w} = \frac{1}{N_w} \sum_{i=1}^{N_w} w_i \tag{91}$$

by introducing a renormalized individual walker weight, ω_i, defined by

$$\omega_i \equiv \frac{w_i}{\overline{w}}. \tag{92}$$

Another stochastic reconfiguration scheme proposes setting the number of copies of walker i for the next step to be proportional to the renormalized walker weight ω_i. This algorithm has shown to have less bias than the approach following Eq. 90. It also has the advantage of having the same number of walkers at each step and simplifies implementation of the algorithm for parallel computers.

2.3.8 Diffusion Monte Carlo algorithm

There are several versions of the DMC algorithm. The approach presented here focuses on simplicity; see refs. [118, 124, 125].

1. Initialize an ensemble $\mathcal{W}$ of $N_{\mathcal{W}}$ configurations, distributed according to $P(\mathbf{R})$ for $\Psi_T(\mathbf{R})$; for example, use the random walkers obtained from a previous VMC run.

2. For every configuration in $\mathcal{W}$:

   (a) Propose an electron move from $\mathbf{R} \equiv \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_{N_p})$ to $\mathbf{R}' \equiv \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}'_i, \ldots, \mathbf{r}_{N_p})$. The short-time approximation propagator, $K(\mathbf{R}, \mathbf{R}'; \delta\tau)$, has an associated stochastic move

   $$\mathbf{R}' \rightarrow \mathbf{R} + \mathbf{F_q}(\mathbf{R})\delta\tau + \mathcal{G}_{\delta\tau}. \tag{93}$$

   (b) Enforce the fixed node constraint: if a random walker crosses a node, i.e., $\text{sign}(\Psi_T(\mathbf{R})) \neq \text{sign}(\Psi_T(\mathbf{R}'))$, then reject the move of the current electron and proceed to treat the next electron.

   (c) Compute the Metropolis acceptance/rejection probability

   $$P(\mathbf{R} \rightarrow \mathbf{R}') = \min\left(1, \frac{K_D(\mathbf{R}, \mathbf{R}'; \delta\tau)\Psi_T^2(\mathbf{R}')}{K_D(\mathbf{R}', \mathbf{R}; \delta\tau)\Psi_T^2(\mathbf{R})}\right), \tag{94}$$

   where K_D is the diffusion and drift transition probability given by Eq. 87.

   (d) Compare $P(\mathbf{R} \rightarrow \mathbf{R}')$ with an uniform random number between 0 and 1, $\mathcal{U}_{[0,1]}$. If $P > \mathcal{U}_{[0,1]}$, accept the move, otherwise, reject it.

(3) Calculate the branching factor G_B for the current configuration

$$B(\mathbf{R}, \mathbf{R}') = e^{(E_R - \frac{1}{2}(\frac{\widehat{H}\Psi_T(\mathbf{R}')}{\Psi_T(\mathbf{R}')} + \frac{\widehat{H}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}))\delta\tau}. \tag{95}$$

(4) Accumulate all observables including the energy. The contributions, O_i, are weighted by the branching factor, i.e.,

$$O_T^{(n+1)} = O_T^{(n)} + B(\mathbf{R}, \mathbf{R}')O_i(\mathbf{R}), \tag{96}$$

where $O_T^{(n)}$ is the cumulative sum of the observable at step n.

(5) Generate a new generation of random walkers, reproducing the existing population, creating an average $B(\mathbf{R}, \mathbf{R}')$ new walkers out of a walker at $\mathbf{R}$. The simplest procedure for achieving this goal is to generate n new copies of $\mathbf{R}$ where $n = \text{int}(B(\mathbf{R}, \mathbf{R}') + \mathcal{U}_{[0,1]})$.

(6) Perform blocking statistics (see Section 2.2.4), and apply population control (see Section 2.3.7)

   (a) One choice is to update the reference energy E_R at the end of each accumulation block,

$$E_R \leftarrow E_R + E_R^{\omega} * \langle E_B \rangle, \tag{97}$$

   where E_R^{ω} is a re-weighting parameter, usually chosen to be ≈ 0.5, and $\langle E_B \rangle$ is the average energy for a block B.

   (b) Discard a relaxation time of steps, N_{Rel}, which is of the order of a tenth of a block, because moving the reference energy induces the most bias in unit relaxation time.

3. Continue the loop until the desired accuracy is achieved.

Umrigar *et al.* [125] proposed several modifications to this algorithm to reduce time-step error. These modifications concentrate on improving the propagator in regions where the short-time approximation performs poorly; namely, near wave function nodes and Coulomb singularities. These propagator errors are expected, because the short-time approximation propagator assumes a constant potential over the move interval, which is a poor approximation in regions where the Coulomb interaction diverges.

2.3.9 Green's function Monte Carlo

The GFMC method is a QMC approach that has the advantage of having no time-step error. It has been shown to require more computer time than DMC, and therefore, has been applied to atomic and molecular systems less frequently than the latter. Good descriptions of the method can be found in refs. [39, 53, 126–128]. The GFMC approach is a PopMC method for which the projector for obtaining the ground state Green's function is the standard resolvent of the Schrödinger equation (see Eq. 67). The integral equation for this case, takes the form

$$\Psi^{(n+1)} = \left[\frac{E_T + E_C}{\widehat{H} + E_C} \right] \Psi^{(n)} \tag{98}$$

where the constant E_C is positive and fulfills the condition that $|E_C| > |E_0|$, and E_T is a trial energy. The resolvent of Eq. 98 is related to the DMC propagator by the one-sided Laplace transform

$$\frac{1}{\widehat{H} + E_C} = \int_0^{\infty} e^{-(\widehat{H} + E_C)\tau} \, d\tau. \tag{99}$$

The integral 99 is evaluated by MC. After equilibration, the sampled times have a Poisson distribution with a mean of $\frac{N_s}{E_0+E_C}$ after N_s steps. The parameter E_C controls the average time step.

The Green's function is not known in closed form, but it can be sampled by MC. This may be done by rewriting the resolvent in the form

$$\frac{1}{\widehat{H}+E_C} = \frac{1}{\widehat{H}_U+E_C} + \frac{1}{\widehat{H}_U+E_C}\left(\widehat{H}_U-\widehat{H}\right)\frac{1}{\widehat{H}+E_C}. \tag{100}$$

The Hamiltonian $\widehat{H}_U$ represents a family of solvable Hamiltonians. To sample the Green's function, one samples the sum of terms on the right-hand side of Eq. 100. The Green's functions associated with $\widehat{H}$ and $\widehat{H}_U$ satisfy the relations

$$\left(\widehat{H}+E_C\right)G\left(\mathbf{R},\mathbf{R}'\right) = \delta\left(\mathbf{R}-\mathbf{R}'\right), \tag{101}$$

$$\left(\widehat{H}_U+E_C\right)G_U\left(\mathbf{R},\mathbf{R}'\right) = \delta\left(\mathbf{R}-\mathbf{R}'\right). \tag{102}$$

The most commonly used form of $\widehat{H}_U$ is

$$\widehat{H}_U = \frac{1}{2}\nabla_R^2 + U, \tag{103}$$

where U is a potential that is independent of $\mathbf{R}$. It is convenient to have $G_U(\mathbf{R},\mathbf{R}')$ vanish at the domain boundary. $\widehat{H}_U$ should be a good approximation to $\widehat{H}$ in the domain to achieve good convergence. The $\mathbf{R}$-space representation of Eq. 100 is

$$\begin{aligned} G\left(\mathbf{R},\mathbf{R}'\right) &= G_U\left(\mathbf{R},\mathbf{R}'\right) - \int_S d\mathbf{R}'' G\left(\mathbf{R},\mathbf{R}''\right)\left[-\hat{n}\cdot\nabla G_U\left(\mathbf{R}'',\mathbf{R}\right)\right] \\ &\quad + \int_V d\mathbf{R}'' G\left(\mathbf{R},\mathbf{R}''\right)\left[U-V\left(\mathbf{R}''\right)\right]G_U\left(\mathbf{R}'',\mathbf{R}'\right). \end{aligned} \tag{104}$$

2.3.10 Fixed-node approximation

In this section, we discuss the implications of the fermion character of $\Psi(\mathbf{R})$. It is an excited state in a manifold containing all the fermionic and bosonic states. A fermion wave function has positive and negative regions that are difficult to sample with the DMC algorithm described in Section 2.3.8. For real wave functions, $\Psi(\mathbf{R})$ contains positive and negative regions, $\Psi^+(\mathbf{R})$, and $\Psi^-(\mathbf{R})$ that, in principle, could be represented as probabilities. The sign of the wave function could be used as an extra weight for the random walk. In practice, doing so leads to a slowly convergent method.

Returning to the importance sampling algorithm, recall that the initial distribution, $|\Psi(\mathbf{R})|^2$, is positive. Nevertheless, the Green's function, $K(\mathbf{R},\mathbf{R}')$, can become negative if a random walker crosses a node of the trial wave function. Again, the sign of $K(\mathbf{R},\mathbf{R}')$ could be used as a weight for sampling $|K(\mathbf{R},\mathbf{R}')|$. The problem is that the statistics of this process lead to exponential growth of the variance of the observable.

The simplest approach to avoid exponential growth is to forbid moves in which the distribution f changes sign, see Eq. 85. This boundary condition on permitted moves is the defining characteristic of the fixed-node approximation (FNA). The nodes of the sampled wave function are *fixed* to be the nodes of the trial wave function. The FNA is an inherent feature of the DMC method, which is, by far, the most commonly used method for atomic and molecular MC applications [117], [118].

The fixed-node energy is an upper bound to the exact energy of the system. In fact, the fixed-node wave function is the best solution for that fixed set of nodes. The DMC method has much higher accuracy than the VMC method. For atomic and molecular systems, it is common to recover 95–100% of the CE, cf. Section 1.1, whereas the CE recovered with the VMC approach is typically less than 80% of the total.

The FNA is the most commonly imposed boundary condition. It satisfies the variational principle, i.e., FNA solutions approach the exact energy from above. This is an useful property, but one that does not facilitate the search for exact results, because there is no known general way to parametrize the nodal surface and to vary it to obtain the exact solution. In the following section, we describe methods that impose no approximate boundary conditions on the wave function.

2.4 Exact methods

Probably the most important algorithmic challenge that remains to be explored is the "node problem". Although progress has been made on systems that contain up to a dozen electrons [129–133], a stable algorithm that can sample the exact wave function without resorting to the FNA remains to be determined. In this section, we discuss a family of methods that avoid the FNA. These approaches yield exact answers at the expense of a considerable increase in computational time.

The Pauli antisymmetry principle imposes a boundary condition on the wave function. It is the requirement that the exchange of like-spin electrons changes the sign of the wave function. This condition is global and holds even for an algorithm that is fundamentally local such as those of QMC methods.

2.4.1 Release node method

The evolution operator, $e^{-\tau(\widehat{H}-E_T)}$, is symmetric and has the same form for both fermions and bosons. Its straightforward application to an arbitrary initial wave function $|\Psi_0\rangle$ leads to collapse to the overall ground state (bosonic) wave function, as can be seen from Eq. 76.

An arbitrary fermion wave function, $\Psi(\mathbf{R})$, can be separated into two functions $\Psi^+(\mathbf{R})$ and $\Psi^-(\mathbf{R})$ as follows,

$$\Psi^{\pm}(\mathbf{R},\tau) \equiv \frac{1}{2}\big[\,|\Psi(\mathbf{R},\tau)| \pm \Psi(\mathbf{R},\tau)\big]. \tag{105}$$

Note that the original trial wave function is recovered as

$$\Psi(\mathbf{R},\tau) = \Psi^{+}(\mathbf{R},\tau) - \Psi^{-}(\mathbf{R},\tau). \tag{106}$$

The released node (RN) algorithm involves two independent DMC calculations, using Ψ^{+} and Ψ^{-} as the wave functions to evolve

$$\begin{aligned}
\Psi(\mathbf{R},\tau) &= \int G\big(\mathbf{R},\mathbf{R}',\delta\tau\big)\Psi\big(\mathbf{R}',0\big)\,d\mathbf{R}' \\
&= \int G\big(\mathbf{R},\mathbf{R}',\tau\big)\Psi^{+}\big(\mathbf{R}',0\big)\,d\mathbf{R}' - \int G\big(\mathbf{R},\mathbf{R}',\tau\big)\Psi^{-}\big(\mathbf{R}',0\big)\,d\mathbf{R}' \\
&= \Psi^{+}(\mathbf{R},\tau) - \Psi^{-}(\mathbf{R},\tau).
\end{aligned} \tag{107}$$

The time evolution of the system can be followed from the difference of separate simulations for $\Psi^{\pm}(\mathbf{R})$. Note that both distributions are always positive during the simulation, and that they decay to the ground state bosonic wave function. This decay is problematic because the "signal-to-noise" ratio in this method depends on the difference between these two distributions. The decay of the difference $\Psi^{+}(\mathbf{R},\tau) - \Psi^{-}(\mathbf{R},\tau)$ goes roughly as $e^{-\tau(E_F - E_B)}$, where E_F is the lowest fermion state energy and E_B is the bosonic ground state energy.

For this method to be practical, one needs to start with the distribution of a good fermion trial wave function. The distribution will evolve from this starting point to the bosonic ground state at large imaginary time τ. In an intermediate "transient" regime one can collect information on the exact fermion wave function.

The energy can be estimated from the expression

$$\begin{aligned}
E_{\mathrm{RN}}(\tau) &= \frac{\int \Psi(\mathbf{R},\tau)\widehat{H}\Psi_T(\mathbf{R})\,d\mathbf{R}}{\int \Psi(\mathbf{R},\tau)\Psi_T\,d\mathbf{R}} \\
&= \frac{\int \Psi^{+}(\mathbf{R},\tau)\widehat{H}\Psi_T(\mathbf{R})\,d\mathbf{R}}{\int[\Psi^{+}(\mathbf{R},\tau) - \Psi^{-}(\mathbf{R},\tau)]\Psi_T(\mathbf{R})\,d\mathbf{R}} \\
&\quad - \frac{\int \Psi^{-}(\mathbf{R},\tau)\widehat{H}\Psi_T(\mathbf{R})\,d\mathbf{R}}{\int[\Psi^{+}(\mathbf{R},\tau) - \Psi^{-}(\mathbf{R},\tau)]\Psi_T(\mathbf{R})\,d\mathbf{R}} \\
&= E_F.
\end{aligned} \tag{108}$$

In the release-node method [130], a fixed-node distribution is propagated as usual, but now two sets of random walkers are retained: $\mathcal{W}_{\mathrm{FN}}$, the fixed node ensemble, and $\mathcal{W}_{\mathrm{RN}}$, the release-node ensemble. Walkers are allowed to cross nodes, and when they do, they are transferred from $\mathcal{W}_{\mathrm{FN}}$ to $\mathcal{W}_{\mathrm{RN}}$. Also a count is made of the number of iterations that a walker has survived, $\mathcal{S}_{\mathrm{RN}} = \{s_1,\ldots,s_{N_w}\}$. This index is used to bin walkers by age. Each time a walker crosses a node, a summation weight associated with it, $\Omega_{\mathrm{RN}} = \{\omega_1,\ldots,\omega_{N_w}\}$ changes sign. These weights determine the sign of the walker contribution to global averages.

The released node energy can be calculated using the estimator,

$$E_{\rm RN} = \frac{\sum_{i=1}^{Nw} \omega_i \frac{\Psi_T(\mathbf{R}_i)}{\Psi(\mathbf{R}_i)} E_L(\mathbf{R}_i)}{\sum_{i=1}^{Nw} \omega_i \frac{\Psi_T(\mathbf{R}_i)}{\Psi(\mathbf{R}_i)}}. \tag{109}$$

2.4.2 Fermion Monte Carlo

From the previous section one can infer that if a method in which the distribution does not go to the bosonic ground state, but stays in an intermediate regime, then it will not have the deficiency of exponential growth of "signal to noise". This leads to the fermion Monte Carlo (FMC) method. The approach [134–139] involves correlated random walks that achieve a constant signal to noise.

The expectation value of Eq. 109 for an arbitrary distribution of signed walkers can be rewritten as

$$\langle E_{\rm FMC} \rangle = \frac{\sum_{i=1}^{Nw} \left[\frac{\widehat{H}\psi_T(\mathbf{R}_i^+)}{\Psi_G^+(R_i^+)} - \frac{\widehat{H}\psi_T(\mathbf{R}_i^-)}{\Psi_G^-(R_i^-)} \right]}{\sum_{i=1}^{Nw} \left[\frac{\psi_T(R_i^+)}{\Psi_G^+(R_i^+)} - \frac{\psi_T(R_i^-)}{\Psi_G^-(R_i^-)} \right]}, \tag{110}$$

where $\Psi_G^{\pm}(\mathbf{R}^{\pm})$ are the guiding functions for a pair of random walkers $P_i = \{\mathbf{R}_i^+, \mathbf{R}_i^-\}$. Note that the variance of the energy estimator of Eq. 110 goes to infinity as the difference between the two populations goes to zero, i.e., the denominator,

$$\mathcal{D} \equiv \sum_{i=1}^{Nw} \left[\frac{\psi_T(R_i^+)}{\Psi_G^+(R_i^+)} - \frac{\psi_T(R_i^-)}{\Psi_G^-(R_i^-)} \right], \tag{111}$$

goes to zero as the simulation approaches the bosonic ground state. A procedure that does not change $\langle E_{\rm FMC} \rangle$, Eq. 110, would be to cancel positive and negative random walkers whenever they meet [129]. Although random walks are guaranteed to meet in one dimension, they need not meet in several dimensions, due to the exponentially decaying walker density in $\mathbf{R}$-space. Besides, cancellation has to be combined with other procedures to insure a stable algorithm.

Cancellation can be increased by introducing correlation between the random walkers. Recall the diffusion step in DMC, in which walkers diffuse from $\mathbf{R}$ to $\mathbf{R}'$ following G_D of Eq. 81. In the DMC algorithm, this is implemented stochastically by updating the coordinates of the random walkers with a random displacement taken from a Gaussian distribution with a variance of $\delta\tau$,

$$\mathbf{R}'^{+} \rightarrow \mathbf{R}^+ + \mathcal{G}_{\delta\tau}^+ \quad \text{and} \quad \mathbf{R}^- \rightarrow \mathbf{R}^- + \mathcal{G}_{\delta\tau}^-. \tag{112}$$

If one introduces correlation between the Gaussian vectors, $\mathcal{G}_{\delta\tau}^+$ and $\mathcal{G}_{\delta\tau}^-$, the expectation value of Eq. 110 is unaffected, because it is linear in the density of random walkers.

An efficient cancellation scheme can be achieved if the Gaussian vectors are correlated as follows:

$$\mathcal{G}_{\delta\tau}^- = \mathcal{G}_{\delta\tau}^+ U^- - 2\left(\mathcal{G}_{\delta\tau}^+ \cdot \frac{(\mathbf{R}^+ - \mathbf{R}^-)}{|\mathbf{R}^+ - \mathbf{R}^-|^2} \right) \cdot (\mathbf{R}^+ - \mathbf{R}^-). \tag{113}$$

Equation 113 accounts for reflection along the perpendicular bisector of the vector that connects the pair, $\mathbf{R}^+ - \mathbf{R}^-$. This cancellation scheme generates a correlated random walk in one dimension along the vector $\mathbf{R}^+ - \mathbf{R}^-$. This one-dimensional random walk is independent of the number of dimensions of the physical system, and therefore overcomes the cancellation difficulties mentioned above. Walkers are guaranteed to meet under these conditions.

The modifications to the DMC algorithm mentioned to this point are necessary, but not sufficient for achieving a stable algorithm. If one were to interchange the random walker populations, $\{\mathbf{R}_1^+, \ldots, \mathbf{R}_{N_w}^+\} \leftrightarrow \{\mathbf{R}_1^-, \ldots, \mathbf{R}_{N_w}^-\}$, the fictitious dynamics would not be able to distinguish between the two populations, leading to a random walk with two degenerate ground states. Specifically, a ground state in which all the positive walkers $\mathbf{R}^+$ are marginally on the positive region of the wave function, and vice versa, $\{\Psi^+(\mathbf{R}^+), \Psi^-(\mathbf{R}^-)\}$ and $\{\Psi^+\{\mathbf{R}^-\}, \Psi^-(\mathbf{R}^+)\}$. This *plus–minus* symmetry can be broken by using two distinct guiding functions. An example is the guiding function

$$\Psi_G^{\pm} = \sqrt{\Psi_S^2(\mathbf{R}) + c^2\Psi_A^2(\mathbf{R})} \pm c\Psi_A(\mathbf{R}), \tag{114}$$

where $\Psi_S(\mathbf{R})$ is a symmetric function under permutation of electron labels; $\Psi_A(\mathbf{R})$ is an antisymmetric function, and c is a small adjustable parameter. The guiding functions of Eq. 114 are almost equal, which provides nearly identical branching factors for the walker pair. They are positive everywhere, a requirement for the DMC algorithm, and are symmetric under permutation of the coordinates, $\Psi_G^+(\widehat{\mathbf{P}}\mathbf{R}) = \Psi_G^-(\mathbf{R})$.

The use of different guiding functions in the respective regions is the last required ingredient for a stable algorithm. It breaks the *plus–minus* symmetry effectively, because the drift dynamics is different because the quantum force of Eq. 35 is distinct for each population. For a complete description of the FMC algorithm, the reader is referred to Ref. [137].

The denominator of Eq. 111 is an indicator of stability of the algorithm. It is a measure of the antisymmetric component of the wave function. FMC calculations have mantained stable denominators for thousands of relaxation times, indicating the stability of the algorithm.

Early versions of the method [129] did not scale well with system size due to the use of uncorrelated cancellation schemes. Nevertheless, researchers have successfully applied the method to several small molecular systems and obtained solutions to the Schrödinger equation with no systematic error [131, 140–142]. An early version of the FMC algorithm, with GFMC propagation and without correlated dynamics, is known as exact quantum Monte Carlo (EQMC).

2.5 Zero variance principle

Besides improving the trial function Ψ_T, one can improve the procedure for the evaluation of the operator $\widehat{O}$. One way of proceeding is to renormalize the operator in such a way that it has the same expectation value, but lower variance. Recently Assaraf and Caffarel [143, 144] have shown how to construct

such operators for the energy and energy derivatives with respect to nuclear position. One defines a trial operator $\widehat{H}_V$ and auxiliary trial function Ψ_V such that the evaluation of a renormalized observable $\overline{O}$ will have a variance that is smaller than that of the original observable $\widehat{O}$, and in the exact limit can be made zero.

To develop this concept, let us construct a trial operator $\widehat{H}_V$ such that,

$$\int \widehat{H}_V(\mathbf{R},\mathbf{R}')\sqrt{\pi(\mathbf{R}')}\,d\mathbf{R}' = 0, \tag{115}$$

where $\pi(\mathbf{R}')$ is the MC distribution. For example, in VMC the MC distribution is the wave function squared, $\Psi_T(\mathbf{R})^2$, and in DMC it is the mixed distribution of Eq. 85. Next, define a renormalized observable $\overline{O}(\mathbf{R})$ related to the observable $\widehat{O}(\mathbf{R})$ given by

$$\overline{O}(\mathbf{R}) = \widehat{O}(\mathbf{R}) + \frac{\int \widehat{H}_V(\mathbf{R},\mathbf{R}')\psi_V(\mathbf{R}')\,d\mathbf{R}'}{\sqrt{\pi(\mathbf{R}')}}. \tag{116}$$

The mean of the rescaled operator is formally

$$\langle\overline{O}\rangle = \frac{\int \widehat{O}(\mathbf{R})\pi(\mathbf{R})\,d\mathbf{R} + \frac{\int\int \pi(\mathbf{R})\widehat{H}_V(\mathbf{R},\mathbf{R}')\Psi_V(\mathbf{R}')\,d\mathbf{R}\,d\mathbf{R}'}{\sqrt{\pi(\mathbf{R})}}}{\int \pi(\mathbf{R})\,d\mathbf{R}} \tag{117}$$

which by property 115 is the same as the mean for the unnormalized operator:

$$\langle\overline{O}\rangle = \langle\widehat{O}\rangle. \tag{118}$$

The operator $\overline{O}$ can be used as an unbiased estimator, even though statistical errors for $\overline{O}$ and $\widehat{O}$ can be quite different. The goal of this kind of importance sampling is to reduce fluctuations by construction of such an operator.

The implementation of the procedure requires optimization of a set of parameters of the auxiliary trial wave function, $\Psi_V(\mathbf{R},\Lambda_V)$, using the minimization functional

$$\int \widehat{H}_V(\mathbf{R},\mathbf{R}')\Psi_V(\mathbf{R},\Lambda_V)\,d\mathbf{R}' = -\left[\overline{O}(x) - \langle\overline{O}\rangle\right]\sqrt{\pi(\mathbf{R})}. \tag{119}$$

After the parameters Λ_V are optimized, one can run a simulation to average $\overline{O}$, instead of $\widehat{O}$. A possible choice of auxiliary Hamiltonian [143] is

$$H_V = -\frac{1}{2}\nabla^2_{\mathbf{R}} + \frac{1}{2\sqrt{\pi(\mathbf{R})}}\nabla^2_{\mathbf{R}}\sqrt{\pi(\mathbf{R})}. \tag{120}$$

Note that when Eq. 120 is applied to $\sqrt{\pi(\mathbf{R})}$, the $\mathbf{R}'$ integration vanishes by construction. The choice of auxiliary wave function is open, and an interesting observation is that for any choice of auxiliary trial wave function, minimization

of the normalization factor of $\Psi_V(\mathbf{R})$, will reduce fluctuations in the auxiliary observable

$$\sigma(\overline{O})^2 = \sigma(\widehat{O})^2 - \frac{\left\langle \frac{\widehat{O}(\mathbf{R}) \int \widehat{H}(\mathbf{R},\mathbf{R}')\Psi_V(\mathbf{R}')\, d\mathbf{R}'}{\sqrt{\pi(\mathbf{R})}} \right\rangle^2}{\left\langle \left[\frac{\int \widehat{H}(\mathbf{R},\mathbf{R}')\Psi_V(\mathbf{R}')\, d\mathbf{R}'}{\sqrt{\pi(\mathbf{R})}} \right]^2 \right\rangle}, \tag{121}$$

because the second term on the right hand side of Eq. 121 is always negative

This variance reduction technique, applied to VMC and GFMC simulations, has achieved an order of magnitude reduction in computational effort [143]. It can also be used to calculate energy derivatives [144].

2.6 Fermion nodes

For the ground state of a bosonic system, i.e., one for which the wave function has the same sign everywhere, QMC provides an exact solution in a polynomial amount of computer time. It is a solved problem. Research on fermion systems seeks an algorithm that has the same favorable scaling, but for wave functions that have both positive and negative regions.

Investigation of nodes has been pursued [145–149] to understand the properties of the nodes of fermion wave functions.

The full nodal hyper-surfaces of a wave function, $\Psi(\mathbf{R})$, where $\mathbf{R}$ is a $3N$-dimensional vector and N is the number of fermions in the system is a $3(N-1)$-dimensional function we label $\eta(\mathbf{R})$. For this function, symmetry requirements determine a $(3N-3)$-dimensional surface, the *symmetry* sub-surface, $\sigma(\mathbf{R})$. This is unfortunate, because even though that $\sigma(\mathbf{R}) \subset \eta(\mathbf{R})$, the remainder of the nodal surface, the *peculiar* nodal surface, $\varpi\mathbf{R})$ which is a function of the specific form of the nuclear and inter-electronic potential, is difficult to specify *a priori* for an arbitrary system. Note that $\sigma(\mathbf{R}) \cup \varpi(\mathbf{R}) = \eta(\mathbf{R})$.

Understanding nodal properties is important for further development of QMC methods.Ceperley discusses general properties of wave function nodes[146]. A listing follows.

1. The *coincidence planes* $\pi(\mathbf{r}_i = \mathbf{r}_j)$, are located at nodes when two electrons have the same spin, i.e., $\delta_{\sigma_{ij}} = 1$. In more than 1 dimension, $\pi(\mathbf{R})$ is a subspace through which the complex nodal surface passes through. Note that $\pi(\mathbf{R}) \subset \sigma(\mathbf{R})$.

2. The nodes possess all the symmetries of the ground state wave function.

3. The nodes of the many-body wave function are distinct from the orbital nodes $\phi_i(\mathbf{r})$ of Section 2.1.

4. For degenerate wave functions, node positions are arbitrary. For a p-fold degenerate energy level, one can pick $p-1$ points in $\mathbf{R}$ and find a linear transformation for which the transformed wave functions vanish at all but one of these points.

5. A *nodal cell* $\Omega(\mathbf{R})$ around a point $\mathbf{R}$ is defined as the set of points that can be reached from $\mathbf{R}$ without crossing a node. For potentials of present interest, the ground state nodal cells have the *tiling-property*: any point $\mathbf{R}'$ that is not on a node is related by symmetry to a point in $\Omega(\mathbf{R})$. This implies that there is only one type of nodal cell: all other cells are copies that can be accessed by relabeling the particles. This property is the generalization to fermions of the theorem that the wave function of the bosonic ground state is a nodeless wave function.

Ceperley suggests that DMC simulations benefit from the tiling property [146]. One only needs to sample one nodal cell, because all cells are equivalent. Any trial function resulting from a strictly mean field theory, such as the LDA approximation, will satisfy the tiling property.

Glauser *et al.* showed that simple HF wave functions of first-row atoms have four nodal regions (two intersecting nodal surfaces) instead of two regions [147]. This structure is attributed to factorizing the wave function into two distinct Slater determinants, $D^{\uparrow}$ and $D^{\downarrow}$, each composed of two surfaces, one for the $\uparrow$ and one for the $\downarrow$ electron, as discussed in Section 2.1.

Recently, after analysis of wave functions for He, Li and Be, it was conjectured by Bressanini *et al.*, that any atomic wave function can be factored as follows [149],

$$\Psi(\mathbf{R}) = N(\mathbf{R})e^{f(\mathbf{R})}, \tag{122}$$

where $N(\mathbf{R})$ is antisymmetric polynomial of finite order, and $f(\mathbf{R})$ is a positive definite function. A weaker conjecture is that N may not be a polynomial, but can be closely approximated by a lower-order antisymmetric polynomial. The variables in which N should be expanded are inter-particle coordinates. For example, for all ^3S states of two-electron atoms, the nodal factor $N(\mathbf{R})$ in Eq. 122 is

$$N(\mathbf{r}_1, \mathbf{r}_2) = \mathbf{r}_1 - \mathbf{r}_2, \tag{123}$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the coordinates of the two electrons.

2.7 Treatment of heavy elements

There is a steep computational dependence of QMC methods with atomic number Z. The computational cost of QMC methods has been estimated to scale as $Z^{5.5-6.5}$ [150, 151]. This has motivated the replacement of the core electrons by effective core potentials (ECPs). With this modification, scaling with atomic number is improved to $Z^{3.4}$ [151]. Other approaches involve the use of core-valence separation schemes [152] model potentials [153], and effective Hamiltonians [154, 155].

2.7.1 Effective core potentials

In the ECP method [156–160], core electrons are simulated by an effective potential acting on the valence electrons. The effective Hamiltonian for these

electrons is

$$\mathcal{H}_{\rm val} = \sum_i \frac{-Z_{\rm eff}}{r_i} + \sum_{i<j} \frac{1}{r_{ij}} + \sum_i \mathcal{W}_i(\mathbf{r}), \tag{124}$$

where i and j designate the valence electrons, $Z_{\rm eff}$ is the effective nuclear charge in the absence of core electrons, and $\mathcal{W}$ is the pseudopotential operator. The latter can be written,

$$\mathcal{W}(\mathbf{r}) = \sum_{l=0}^{\infty} \mathcal{W}_l(\mathbf{r}) \sum_m |lm\rangle\langle lm|, \tag{125}$$

where l and m are the angular momentum and magnetic quantum numbers. The projection operator $\sum_m |lm\rangle\langle lm|$, connects the pseudopotential with the one-electron valence functions. A common approximation to Eq. 125 is to assume that the angular momentum components of the pseudopotential, $w_l(\mathbf{r})$ do not depend on l for $l > L$, the angular momentum of the core. This approximation leads to the expression

$$\mathcal{W}(\mathbf{r}) = \mathcal{W}_{L+1}(\mathbf{r}) + \sum_{l=0}^{L} \big(\mathcal{W}_l(\mathbf{r}) - \mathcal{W}_{L+1}(\mathbf{r})\big) \sum_m |lm\rangle\langle lm|. \tag{126}$$

The operator 126 can be applied to a valence orbital, i.e., pseudo-orbital, $\phi_l(\mathbf{r})$. This function is usually represented by a polynomial expansion for distances less than a cutoff radius, $r < r_c$, and by a fit to the all-electron orbital for $r > r_c$.

Rapid fluctuations in potential terms can cause the first-order propagator of Eq. 84 to fail to describe the wave function accurately and leads to seeking a slowly varying ECP for QMC simulations. Greeff and Lester [161, 162] proposed the use of norm-conserving soft ECPs for QMC. Soft ECPs derive their name from the property of being finite at the nucleus and leads to a pseudo-orbital with no singularities at the origin in the kinetic energy. The associated effective potential has no discontinuities or divergences.

2.7.2 Embedding methods

For the treatment of large systems, a commonly used approach with wave function-based methods is to use embedding schemes, in which a region of high interest of a large system is treated by an accurate procedure for the remainder of the system is described by a less accurate method. Recent work by Flad *et al.* [163] has extended the methodology to QMC methods. In the latter approach, a mean field calculation is performed for the whole system. An electron localization procedure is performed, the orbitals to be correlated are chosen and separated from the remaining orbitals. An effective Coulomb and exchange potential $\widehat{V}_E$ is constructed $\widehat{V}_E$ which is added to the standard Hamiltonian of Eq. 1 to construct an effective Hamiltonian ,$\widehat{H}_E$, for QMC calculations. Localization procedures, similar to those required for ECPs, are needed for representing the effect of nonlocal terms.

The effective Hamiltonian, $\widehat{H}_E$, takes the form

$$\widehat{H}_E = \widehat{H}_{\text{int}} + \widehat{V}_{\text{ext}} + \widehat{J}_{\text{ext}} + \widehat{K}_{\text{ext}} + \widehat{S}_{\text{ext}} \tag{127}$$

where $\widehat{H}_{\text{int}}$ is the Hamiltonian for the QMC active region, $\widehat{V}_{\text{ext}}$ is the Coulomb potential exerted by external nuclei, and $\widehat{J}_{\text{ext}}$ represents Coulomb repulsions, The term $\widehat{K}_{\text{ext}}$ denotes exchange interactions, and $\widehat{S}_{\text{ext}}$ is a shift operator that prevents the wave function expanded in core orbitals ϕ_c by raising their single-particle energies to infinity, and is given by

$$\widehat{S}_{\text{ext}} = \lim_{\lambda\to\infty} \lambda \sum_{\alpha}^{\text{int}} \sum_{\beta}^{\text{ext}} |\phi_p(\mathbf{r})\rangle\langle\phi_p(\mathbf{r})| \, d\mathbf{r}. \tag{128}$$

Here λ is an effective orbital coupling constant that is derived from considering single and double excitations into core and virtual orbitals of the system. The Coulomb term, $\widehat{J}_{\text{ext}}$, and the external Coulomb potential, $\widehat{V}_{\text{ext}}$, are local potentials, and can be evaluated within QMC without further approximation. The remaining terms require localization approximations that have been discussed in detail [163].

2.8 The exact wave function and quantities that do not commute with the Hamiltonian

Given a trial wave function $\Psi_T(\mathbf{R})$, atomic and molecular properties are readily computed in VMC calculations [164]. Expectation values are calculated directly from Eqs. 31 and 32. The accuracy of the results obtained with VMC depend on the quality of $\Psi_T(\mathbf{R})$.

To obtain expectation values of operators that do not commute with the Hamiltonian in an importance sampled PMC calculation, one needs to extract the exact distribution $\Psi^2(\mathbf{R})$ from the mixed distribution $f(\mathbf{R}) = \Psi(\mathbf{R})\Psi_T(\mathbf{R})$. The expectation values for an operator $\widehat{O}$, $\langle\Psi(\mathbf{R})|\widehat{O}|\Psi_T(\mathbf{R})\rangle$ and $\langle\Psi_T(\mathbf{R})|\widehat{O}|\Psi(\mathbf{R})\rangle$, are not the same. MC sampling requires knowledge of the exact ground state distribution: a mixed distribution is not sufficient to obtain the exact result.

If the operator $\widehat{O}$ is a multiplicative operator, then the algorithms described in this section will be pertinent. Treatment of nonmultiplicative operators, are described in Section 2.9.

2.8.1 Extrapolation method

An approximate procedure for estimating the ground state distribution can be obtained by extrapolation from the mixed and VMC distributions. This procedure is valuable because no modifications of the canonical VMC and PMC algorithms are needed. Extrapolation can fail even in very simple cases [20], but it also has provided very accurate results in more favorable cases [165]. The mixed estimator of a coordinate operator $\widehat{O}$ is

$$\langle\widehat{O}\rangle_m = \frac{\int \Psi(\mathbf{R})\widehat{O}\Psi_T(\mathbf{R})\, d\mathbf{R}}{\int \Psi(\mathbf{R})\Psi_T(\mathbf{R})\, d\mathbf{R}}, \tag{129}$$

which is to be distinguished from the pure estimator

$$\langle \widehat{O} \rangle_p = \frac{\int \Psi(\mathbf{R}) \widehat{O} \Psi(\mathbf{R})\, d\mathbf{R}}{\int \Psi(\mathbf{R}) \Psi(\mathbf{R})\, d\mathbf{R}}. \tag{130}$$

Equation 129 can be rewritten in a Taylor series in the difference between the exact and approximate wave functions, $\delta\Psi \equiv \Psi(\mathbf{R}) - \Psi_T(\mathbf{R})$,

$$\langle \widehat{O} \rangle_m = \langle \widehat{O} \rangle_p + \int \Psi \big(\langle \widehat{O} \rangle_p - \widehat{O}(\mathbf{R}) \big) \delta\Psi \, d\mathbf{R} + \mathcal{O}\big((\delta\Psi)^2\big). \tag{131}$$

A similar expansion can be constructed for the variational expectation, $\langle \widehat{O} \rangle_v$,

$$\langle \widehat{O} \rangle_v = \langle \widehat{O} \rangle_p + 2 \int \Psi \big(\langle \widehat{O} \rangle_p - \widehat{O}(\mathbf{R}) \big) \delta\Psi \, d\mathbf{R} + \mathcal{O}\big((\delta\Psi)^2\big). \tag{132}$$

Combining Eqs. 131 and 132, we can arrive at an expression with a second-order error,

$$\langle \widehat{O} \rangle_e = 2\langle \widehat{O} \rangle_m - \langle \widehat{O} \rangle_v = \langle \widehat{O} \rangle_p + \mathcal{O}\big((\delta\Psi)^2\big), \tag{133}$$

where $\langle \widehat{O} \rangle_e$ is an extrapolation estimate readily available from VMC and PMC calculations.

2.8.2 Future walking

The future walking method can be combined with any importance sampled PMC method that leads to a mixed distribution. If one multiplies both sides of Eq. 129 by the ratio $\Psi(\mathbf{R})/\Psi_T(\mathbf{R})$, one recovers Eq. 130. The ratio is obtained from the asymptotic population of descendants of a single walker [166] .

A walker in $\mathbf{R}$-space can be represented as a sum of eigenfunctions of $\widehat{H}$:

$$\delta\big(\mathbf{R}' - \mathbf{R}\big) = \Psi\big(\mathbf{R}'\big) \sum_{n=0}^{\infty} c_i(\mathbf{R}) \Psi_n(\mathbf{R}). \tag{134}$$

The coefficients $c_i(\mathbf{R})$ can be obtained by multiplying Eq. 134 by $\Psi(\mathbf{R}')/\Psi_T(\mathbf{R}')$ and integrating over $\mathbf{R}'$:

$$c_n(\mathbf{R}) = \int \delta\big(\mathbf{R}' - \mathbf{R}\big) \frac{\Psi(\mathbf{R}')}{\Psi_T(\mathbf{R}')}\, d\mathbf{R}' = \frac{\Psi(\mathbf{R})}{\Psi_T(\mathbf{R})}. \tag{135}$$

One wants to know the contribution to the ground state wave function, $c_0(\mathbf{R})$ of the walker at $\mathbf{R}$. If propagated for sufficiently long time, all coefficients $c_i(\mathbf{R}) \neq c_0(\mathbf{R})$ for the random walker will vanish. This can be seen from the decay in τ of Eqs. 75 and 76.

If one defines $P_\infty(\mathbf{R})$ to be the asymptotic population of walkers descended from a random walker at $\mathbf{R}$, one finds,

$$\begin{aligned} P_\infty(\mathbf{R}) &= \int c_0(\mathbf{R}) e^{-(E_0 - E_T)\tau} \Psi\big(\mathbf{R}'\big) \Psi_T\big(\mathbf{R}'\big)\, d\mathbf{R}' \\ &= \frac{\Psi(\mathbf{R})}{\Psi_T(\mathbf{R})} e^{-(E_0 - E_T)\tau} \big\langle \Psi(\mathbf{R}) \big| \Psi_T(\mathbf{R}) \big\rangle. \end{aligned} \tag{136}$$

To obtain $P_\infty(\mathbf{R})$ in a PMC algorithm, one keeps a list of all descendants of each walker $\mathbf{R}_i$ at each time step τ_j. The number of steps needed to keep track of descendants N_d is a critical parameter. The statistical error of the asymptotic walker population grows in the limit $N_d \to \infty$, and if only few steps are used, a bias is encountered by nonvanishing contributions from excited states $c_i(\mathbf{R}) \neq c_0(\mathbf{R})$. Efficient algorithms for keeping track of the number of descendants can be found in the literature [11, 61, 165–169]. Recently, a method that does not require keeping explict track of the number of descendants has been proposed [170, 171].

Trial wave function overlap with the exact ground state can also be obtained with these methods, as shown by Hornik *et al.* [109]. These methods have been applied for obtaining dipole moments [172], transition dipole moments [173] and oscillator strengths [174]. Other methods for obtaining the exact distribution that are not discussed here are bilinear methods [175], and time correlation methods [176].

2.9 Force calculations

Most QMC applications assume the Born–Oppenheimer (BO) [177] approximation. In this approximation, the nuclear coordinates $\mathcal{R}$ are fixed at a certain position during the calculation.[9] The wave function and energy depend parametrically on the nuclear coordinates $\mathbf{R}$, i.e., $E(\mathcal{R})$ and $\Psi(\mathbf{R}, \mathcal{R})$. We will omit this parametric dependence for the remainder of the discussion, and simplify the arguments to E and $\Psi(\mathbf{R})$, where appropriate.

Forces are derivatives of the energy with respect to nuclear displacement:

$$F(\mathcal{R}) = -\nabla_{\mathcal{R}} E(\mathcal{R}). \tag{137}$$

Because of the stochastic nature of QMC, obtaining forces in the approach is a difficult task. Generally, QMC calculations are carried out using geometries obtained with a different quantum chemical method such as density functional theory (DFT) [179], or wave function methods. Whereas DFT and HF wave function methods can use the Hellman–Feynman theorem for the calculation of forces, a straightforward application of the theorem in QMC leads to estimators with very large variance.

2.9.1 Analytic derivative methods

The calculation of analytic derivative estimators is a costly process both for wave function-based methods, and for QMC methods. Fortunately, in QMC one does not have to evaluate derivatives at each step, but rather sample points intermittently, both to reduce computer time and serial correlation.

The local energy estimator for a DMC mixed distribution is

$$E_0 = \langle E_L \rangle = \frac{\int \Psi_0(\mathbf{R}) E_L(\mathbf{R}) \Psi_T(\mathbf{R})\, d\mathbf{R}}{\int \Psi_0(\mathbf{R}) \Psi_T(\mathbf{R})\, d\mathbf{R}}. \tag{138}$$

[9] The QMC method can be used for calculations without the BO approximation, but applications to date have been to nodeless systems, such as H_2 [178].

The gradient of Eq. 138 involves derivatives of the unknown exact wave function $\Psi_0(\mathbf{R})$, and the trial wave function $\Psi_T(\mathbf{R})$. Derivatives of $\Psi_0(\mathbf{R})$ have to be obtained with a method devised for sampling operators that do not commute with the Hamiltonian, that are described in Section 2.8. This approach leads to an exact estimator for the derivative, but with the added computational complexity inherent in those methods. A simple approximation can be used of replacing the derivatives of $\Psi_0(\mathbf{R})$ with those of $\Psi_T(\mathbf{R})$ to obtain

$$\nabla_{\mathcal{R}} E_0 \approx \left\langle \nabla_{\mathcal{R}} E_L(\mathbf{R}) \right\rangle + 2\left\langle E_L \frac{\nabla_{\mathcal{R}} \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \right\rangle - 2E_0 \left\langle \frac{\nabla_{\mathcal{R}} \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \right\rangle. \tag{139}$$

The derivatives of $\Psi_T(\mathbf{R})$ are readily obtainable from the known analytic expression of $\Psi_T(\mathbf{R})$.

The exact derivative involves the cumulative weight of configuration $\mathbf{R}_i$ at time step s, $\mathbf{R}_i^s$,

$$\overline{B}_i = \prod_{s_0}^{s} B\left(\mathbf{R}_i^s\right), \tag{140}$$

where the range $s_0 - s$ is the number of generations for the accumulation of the cumulative weight, and $B(\mathbf{R}_i^s)$ is the PMC branching factor of Eqs. 65 and 95. This is an application of the future walking method of Sec. 2.8.2. The energy expression using cumulative weights is

$$E_0 = \frac{\int \Psi_T^2(\mathbf{R}) \overline{B}(\mathbf{R}) E_L(\mathbf{R})\, d\mathbf{R}}{\int \Psi_T^2(\mathbf{R}) \overline{B}(\mathbf{R})\, d\mathbf{R}}. \tag{141}$$

and the derivative of 141 is,

$$\begin{aligned} \nabla_{\mathcal{R}} E_0 \quad = \quad & \left\langle \nabla_{\mathcal{R}} E_L(\mathbf{R}) \right\rangle + 2\left\langle E_L(\mathbf{R}) \frac{\nabla_{\mathcal{R}} \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \right\rangle - 2E_0 \left\langle \frac{\nabla_{\mathcal{R}} \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \right\rangle \\ & + \left\langle E_L(\mathbf{R}) \frac{\nabla_{\mathcal{R}} \overline{B}(\mathbf{R})}{\overline{B}(\mathbf{R})} \right\rangle - E_0 \left\langle \frac{\nabla_{\mathcal{R}} \overline{B}(\mathbf{R})}{\overline{B}(\mathbf{R})} \right\rangle. \end{aligned} \tag{142}$$

Analytic energy derivatives have been applied to H_2 [167], LiH and CuH [180]. Higher order derivatives can be obtained as well. Details on the latter can be found in Refs [181, 182].

2.9.2 Correlated sampling

An efficient approach to force calculation is correlated sampling, which is a MC method that uses correlation between similar observations to reduce the statistical error of the sampling. If one represents Eq. 137 in a finite difference scheme for evaluating a derivative along the difference vector,$\mathbf{r}_d$,

$$\frac{\partial E}{\partial \mathbf{r}_d} \approx \frac{E(\mathcal{R} + \mathbf{r}_d) - E(\mathcal{R})}{\mathbf{r}_d}, \tag{143}$$

then one obtains an approximate energy derivative along the $\mathbf{r}_d$. If two separate calculations are carried out, with a statistical error of the energies of σ_E, the statistical error for the difference σ_d is approximately

$$\sigma_d \approx \frac{\sigma_E}{\mathbf{r}_d}. \tag{144}$$

One can see that because $\mathbf{r}_d$ is a small perturbation, of ≈ 0.01 a.u., the statistical error of the difference will be several times larger than the statistical error of the energies. If $\mathbf{r}_d$ is sufficiently small, a single random walk can be performed, while evaluating the energy at the original and perturbed geometries, $E[\Psi(\mathcal{R})]$ and $E[\Psi(\mathcal{R}+\mathbf{r}_d)]$. In this case, both the primary ($\mathcal{R}$) and secondary ($\mathcal{R}$) walks will be correlated, and therefore have a lower variance than uncorrelated random walks.

To implement correlated sampling for forces in a PMC algorithm, Filippi and Umrigar [183] used expressions including branching factors $B(\mathbf{R})$, re-optimized the parameters of the wave function Λ for each perturbed geometry, and performed additional coordinate transformations. Implementation details of the correlated sampling method for derivatives are described in detail in refs. [183, 184].

2.9.3 Hellman–Feynman derivatives and the zero variance theorem

The Hellman–Feynman theorem states that the force can be obtained from the gradient of the potential

$$\langle \nabla_{\mathcal{R}} E_0 \rangle = -\frac{\int \Psi^2(\mathbf{R}) \nabla_{\mathcal{R}} V(\mathbf{R})\, d\mathbf{R}}{\int \Psi^2(\mathbf{R})\, d\mathbf{R}}, \tag{145}$$

where $V(\mathbf{R})$ is the Coulomb potential for the system. A QMC estimator of the Hellman–Feynman force, $\mathbf{F}_{\mathrm{HF}} \equiv -\nabla_{\mathcal{R}} V(\mathbf{R})$, can be constructed, but it has infinite variance. This property is a consquence of the dependence of the force at short electron–nucleus distances, $\mathbf{r}_{iN}$. Because the variance of $\mathbf{F}_{\mathrm{HF}}$ depends on $\langle \mathbf{F}^2_{\mathrm{HF}} \rangle$, one has a singularity. Furthermore, the Hellman–Feynman theorem only holds for exact wave functions and in the Hartree-Fock limit, in general basis set errors need to be accounted for [185]. Also, the fixed-node approximation introduces an extra requirement on the nodal surface. The former has to be independent of the position of the nuclei, or it has to be the exact one. An elaborate discussion of this issue can be found in refs. [186, 187]. Lu [188, 189] has been succesful in avoiding these difficulties with the use of floating spherical gaussian basis functions and gaussian geminals.

One solution to the infinite variance problem is to evaluate the forces at a cutoff distance close to the nuclear position, and then extrapolate to the nuclear position [181]. This approach has the problem that the extrapolation procedure is difficult, because the variance increases as the cutoff value is decreased.

As discussed in Section 2.5, renormalized operators can be obtained in such a way that they have the same expectation value, but lower variance. Recently,

Assaraf and Caffarel [144] introduced a renormalized Hellman–Feynman operator

$$\overline{\mathbf{F}}_{\mathrm{HF}} = \mathbf{F}_{\mathrm{HF}} + \left[\frac{\widehat{H}_V \Psi_V(\mathbf{R})}{\Psi_V(\mathbf{R})} - \frac{\widehat{H}_V \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}\right] \frac{\Psi_V(\mathbf{R})}{\Psi_T(\mathbf{R})}. \tag{146}$$

Here $\Psi_V(\mathbf{R})$ is an auxiliary wave function and $\widehat{H}_V$ is an auxiliary Hamiltonian. The variance of operator (146) can be shown to be finite, and therefore smaller than the variance of $\mathbf{F}_{\mathrm{HF}}$. The form of $\Psi_V(\mathbf{R})$ proposed in [144] is a simple form that cancels the singularities of the force in the case of a diatomic molecule. General forms of the auxiliary wave function can be constructed.

2.9.4 Geometry Optimization

The solution mapping (SM) methodology has evolved from the needs of reactive flow simulations of combustion processes, where the large size of reaction models has driven the development of economic numerical strategies [190, 191]. In the SM approach an approximation is sought not to the mathematical equations that define the model, but to the solution of these equations [192]. The approximation is developed through the statistical techniques of response surface design by performing a relatively small number of computer experiments with the original model and fitting the numerical results with simple functions such a polynomials [193, 194].

The computer experiments are preformed at preselected combinations of the factorial variables, and the set of these computer experiments is called a factorial design. Factorial designs originate from rigorous analysis of variance with the objective of minimizing the number of computer experiments that must be performed in order to gain the desired information.

As a test of the SM method for geometry optimization, Schuetz *et al* [195] applied it to formaldehyde. A total of 15 single-point energy evaluations were carried out with single determinant ECP VMC and DMC methods. The QMC energies were fit to a second order polynomial. The resultant function yielded optimized bond lengths and bond angle in the DMC method that lie within experimental error. The VMC optimized geometry is within experimental error except for the CH bond length which is slightly underestimated. The resulting quadratic represenatation of the potential energy surface in the region of the minimum was used to calculate three force constants and harmonic frequencies. The accuracy of the force constants and vibrational frequencies were relatively low. The authors indicate that this is in part due to the quality of the fitting function to the VMC and DMC singlet point energies.

2.9.5 Variational Monte Carlo dynamics

Correlated sampling can be combined with a fictitious Lagrangian technique, similar to that developed by [196] in a way first proposed by [197] for geometry optimization. In this approach, the expectation value of the Hamiltonian is

treated as a functional of the nuclear positions and the correlation parameters:

$$\langle \widehat{H} \rangle = \frac{\langle \Psi | \widehat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = E\big[\{\Lambda\}, \{\mathcal{R}\}\big]. \tag{147}$$

With Eq. 147, a fictitious Lagrangian can be constructed of the form

$$L = \sum_{\alpha} \frac{1}{2} \mu_\alpha \lambda_\alpha'^2 + \sum_{I} \frac{1}{2} M_I \mathcal{R}_I'^2 - E\big[\{\Lambda\}, \{\mathcal{R}\}\big], \tag{148}$$

where M_I are the nuclear masses and μ_a are the fictitious masses for the variational parameters, λ_α. The modified Euler–Lagrange equations can be used for generating dynamics for the sets of parameters, $\{\mathcal{R}\}$ and $\{\Lambda\}$,

$$M_I \mathcal{R}_I'' = \nabla_{R_I} E, \tag{149}$$

$$\mu_\alpha \lambda_\alpha'' = \frac{\partial E}{\partial \lambda_\alpha}. \tag{150}$$

A dissipative transformation of Eqs. 149 and 150, where the masses M_I and μ_α are replaced by damped masses $\widetilde{M_I}$ and $\widetilde{\mu_\alpha}$ can be used for geometry optimization. A more elaborate approach that attempts to include quantum effects in the dynamics is described in [198] .

2.10 Linear Scaling QMC

In insulator systems, interactions between electrons diminish rapidly with distance between them. Several computational methods have been devised to take advantage of this property. Quantum simulations of complex chemical systems have recently become accessible in QMC with the development of linear-scaling methods. One such method is the local diffusion Monte Carlo method (LDMC); see refs. [199, 200]. The reader is referred to these references for further details on the implementation and efficiency of the method. A brief description follows.

The time-determining step of most QMC codes for chemical applications is the calculation of local energy averages of Eq. 89. Local energy evaluation requires most (usually $\approx 95 - 99\%$) of the total computational effort of a calculation.

We now analyze the computational effort in sufficient detail to indicate the strategy leading to linear scaling. The computer time for a given QMC calculation is proportional to the number of evaluations of the local energy function

$$t_{CPU} = N \times t_{E_L} \tag{151}$$

where t_{E_L} is the time for a single evaluation of the local energy, and N is the number of evaluations during the MC run.

To achieve a given accuracy requires a certain number of evaluations N. The scaling of N with system size depends on the type of system studied and on trial wave function quality. If a given accuracy is sought, regardless of system size, the number of evaluations scales roughly linearly with system size.

We next describe methods for improving the scaling of t_{E_L}.

2.10.1 Linear scaling of the local energy

For evaluating correlated molecular orbital (CMO) wave functions described in Eq. 85, the following steps are performed

1. Evaluation of basis functions $\chi_k(\mathbf{r})$.
2. Evaluation of MOs and their first and second derivatives with respect to electronic positions $\phi_k(\mathbf{r})$, $\nabla\phi_k(\mathbf{r})$, and $\nabla^2\phi_k(\mathbf{r})$.
3. Calculation of the Slater determinant(s) and their first and second derivatives with respect to the coordinates of the configuration, i.e., $\mathcal{D}(\mathbf{R})$, $\nabla_{\mathbf{R}}\mathcal{D}(\mathbf{R})$, and $\nabla^2_{\mathbf{R}}\mathcal{D}(\mathbf{R})$, where $\mathbf{R} \equiv \{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\}$.
4. Calculation of the correlation function $\mathcal{F}$, see Eqs. 26 and 27.
5. Calculation of the potential energy, $V(\mathbf{R}) \equiv V_{ee}(\mathbf{R}) + V_{eN}(\mathbf{R})$.

In the succeding we describe the procedure for Gaussian or Slater expansions suggested by Manten and Lüchow [200]. For plane-wave calculations, analogous steps are described in ref. [199].

(1) Basis Functions

Molecular orbitals (MOs) are given as linear combination of M basis functions,

$$\phi_k(\mathbf{r}_i) = \sum_{n=0}^{M} c_{kn}\chi_{kn}(\mathbf{r}_i). \tag{152}$$

The form of the basis functions χ_{kn} is usually employed is that of Eqs. 16 or 17. Linear scaling can be achieved by eliminating the contributions of basis functions that lie beyond a predetermined cut-off radius, i.e. $|\mathbf{r}_i - \mathbf{r}_{ck}| > r_{cut}$, where $\mathbf{r}_{ck}$ is the center of the k-th basis function, and r_{cut} is the cut-off radius. The cut-off criterium for the radius used by Manten and Lüchow [200] is the absolute value of the laplacian of the basis function, i.e., $|\nabla^2\chi_{kn}|$.

(2) Molecular Orbitals

Evaluation of the MOs is a cubic (N^3) procedure that is achieved by a sum over the particle index i in Eq. 152. Procedures, such as those described in refs. [201–204] can be used to construct localized MOs (LMOs). By using LMOs, a constant number of basis functions is required for each electronic evaluation. In practice, the methods described above require that LMOs remain orthogonal. This constraint is evidenced in small contributions to the LMOs at centers that lie far from the main contribution to the orbital. To achieve linear scaling, cut-offs are introduced. One discards the remote small contributions and uses the modified LMOs to construct the MC guiding function.

(3) Slater Determinant(s)

The linear algebra operations for inverting a slater determinant require N^3 operations. Usually, these operations require modest time compared to the needed for MO evaluation. If LMOs are used, the Slater determinant is sparse

and the inverse can be obtained using sparse LU decomposition [205]. The later scales linearly with system size.

(4) Correlation Function

The g_{ij} terms in correlation function of Eq. 27 can be made to decay exponentially with respect to interparticle coordinates, by using an exponentially scaled reduce distance,

$$g_{ij} = 1 - exp^{-\lambda_c r_{ij}}. \tag{153}$$

The exponent λ_c can be made sufficently small without notably decreasing the quality of the wave function. Analogous to the case of basis functions, the g_{ij} terms can be ignored after a cut-off radius, thus achieving linear scaling.

(5) Potential Energy

This is usually the part of the local energy computation that requires the least computer time. The scaling is quadratic with respect to system size, due to the double sum over electrons and nuclei. If necessary, fast multipole methods can be employed ([206]. Section 9 reviews several applications to large systems that would otherwise be infeasible without the use of LDMC.

3 Applications - general comments

A range of chemical reactions, atoms, molecules, and solids has been studied using the QMC method. Properties, including atomization energies, heats of formation, optical transitions, and relative stabilities, have been computed to a high degree of accuracy with this method. The DMC approach typically recovers $\approx$95% of the correlation energy and does not suffer from the strong basis set dependence of other *ab initio* methods such as HF, DFT, and coupled-cluster (CC) methods. There are however, systematic errors that can be encountered in QMC calculations: fixed-node error in DMC and localization error with ECPs, dicussed in section. The fixed-node error can be as much as $\approx$5% of the total correlation energy. The localization error is typically not of significance and is often masked by the fixed-node error.

4 Atoms

The literature on QMC calculations and methods for atoms is growing rapidly. Various correlation functions, including variants of the Jastrow [207] and Boys-Handy functions [79], as well other compact forms, have been used to calculate the total energy of first-row atoms. Many procedures have been proposed to optimize correlation function parameters in conjunction with MO and linear coefficients in multi-determinant trial functions; see, for example Refs. [62, 95, 101, 103, 208–210]. In the next section, we comment on neutral and charged species with particular attention to the work of Schmidt and Moskowitz [72] and Galvez *et al.*[211].

In Table 1, VMC energies for first-row atoms are compared to estimated exact non-relativistic total energies (E_s) [212]. The VMC energies were calculated

using correlation functions containing 7 (E_7), 9 (E_9) and 17 (E_{17}) parameters. The 17-parameter function recovers 100% of the He correlation energy and 97% of the Li correlation energy. For Be, B, and C the correlation energy recovered is less than that for other first-row atoms due to known 2s-2p degeneracy effects.

Galvez *et al.* [211] carried out a VMC study using the 17- and 42-parameter variants of the Schmidt-Moskowitz Boys-Handy (SMBH) correlation function. Non-dynamic correlation due to the 2s-2p near degeneracy was accounted for using a 2-CSF wave function of the form

$$\psi = \psi_1 + \lambda\psi_2$$

where λ is a variational parameter, ψ_1 is the $1s^2 2s^2 2p^k$ CSF, and ψ_2 is the $1s^2 2p^{k+2}$ configuration state function (CSF). A DMC CSF is typically written

$$\psi_I = \sum_i d_i D^{\alpha}_{iI} D^{\beta}_{iI} \tag{154}$$

where d_i is a linear coefficient and D^{α} and D^{β} are determinants of MOs of α and β spin electrons. For four-, five-, and six-electron systems, k = 0, 1, and 2, respectively. Multi-configuration HF (MCHF) wave functions of these descriptions were used as VMC trial functions

The MCHF trial functions with a 17-parameter SMBH correlation function recovered 80-100% of the correlation energy; see Table 1. In particular, for Be 97.2%, for B 87%, and C for 83% of the correlation energy was recovered compared to 68%, 69%, and 72%, respectively, using a HF trial function. Increasing the number of correlation parameters to 42 from 17 improves the correlation energy recovered by $\sim 2\%$ for all first-row atoms.

Galvez *et al.* also used MCHF trial functions to compute VMC energies for first-row positive and negative ions [211]. These results are compared with VMC results of Schmidt-Moskowitz and estimated exact total non-relativistic energies in Table 2; note bound states do not exist for $Be^-$and N^-. For the bound systems, the MCHF trial function multiplied by the 17-parameter SMBH function of Galvez *et al.* recovered 82-100% of the correlation energy while a 9-parameter SMBH-HF trial function of Schmidt and Moskowitz at the VMC level recovered 60-88% of the correlation energy; see Table 2. For B^-, C^-, O^- and F^- the MCHF wave function also with the 17 correlation parameters recovered 81.4-96.5% of the correlation energy compared to 73-85% obtained with the HF with tbe 17 parameter correlation function. Increasing the number of parameters in the correlation function to 42 coupled with the MCHF trial function did not yield significant improvement in total energy relative to the MCHF 17-parameter wave functions.

In Table 3. we compare computed ionization potentials (IP) and electron affinities (EA) of first-row atoms to experiment. The IP and EA results obtained with HF and MCHF trial functions are in good agreement with experiment, except for the B and F EAs. For B the EA is underestimated because the trial functions does not properly contain the more diffuse p orbitals in B^- as

addressed by Botch and Dunning [213]. For F the VMC EA is greatly improved using a 2-CSF trial function

Greeff *et al.* computed several properties of Al and Al_2, including the IPs and binding energies [214]. In the DMC calculations, the 1s, 2s, and 2p electrons were replaced with the Stevens Basch and Krauss (SBK) ECP, resulting in effective 3- and 6-electron systems for Al and Al_2, respectively. To determine the ECP localization error in the DMC calculations, two trial functions were calculated; the first consisted of an 8-parameter SMBH correlation function and the second a simple Jastrow function. The VMC energy with a simple Jastrow recovered 33% of the correlation energy defined as the difference between the SCF energy and the DMC energy. The DMC energy obtained with the ECP and simple Jastrow function is 1.93320(57) h compared to the 8-parameter SMBH function that yielded an energy of -1.93339(43) h. The DMC energies agree to within statistical uncertainty, indicating that the ECP localization error is small for Al. However, this cannot be assumed to hold for all atoms. In transition metals using argon-core ECPs, the localization error has been found to be large [215].

DMC calculation for ground state Al_2 were carried out at the experimental bond length of 5.140 bohr, while for the Al^+ and Al^- states, Jones bond lengths were used [216]. The DMC IP for $Al(^2P)$ is smaller than experiment by 0.023(13) eV. The DMC calculation is expected to underestimate the IP of Al because there is no fixed-node error for $Al^+(^1S)$, which is a nodeless two-electron system with the ECP, while the energy of $Al(^2P)$ is governed by the fixed-node approximation. These results are in better accord with experiment than the previous CCD+ST calculations of Sunil and Jordan [217] and the MRCI result of Bauschlicher *et al.* [218]. The EA of Al is in agreement with the experimental value of 0.44(1) eV and is comparable in accuracy to large-scale CI calculations. The DMC binding energies for the state is in agreement with contracted MRCI (CMR-CI) and MRCI calculations, although these values lie just within experimental uncertainty and 0.1-0.05 eV above previous *ab initio* calculations.

Removing an electron from the p orbital of the ground state of Al_2 yields the ground state of Al_2^+. The adiabatic and verticals IPs for Al_2^+ were carried out using the experimental Al_2 bond length. The DMC vertical IP is significantly larger than that of Sunil and Jordan [217]. This is likely due to different bond lengths used in the calculations, however, this result is smaller than the experimental value of Harrington and Weishaar [219]. The binding energies for $Al_2^+$and Al_2^- are larger than previous calculations, while the electron affinity of Al_2 and the binding energy of Al_2^- are in agreement with the results of Sunil and Jordan [217]. The fact that Greeff *et al.* obtain an IP and EA for Al atom in somewhat better agreement with experiment leads to the belief that these fixed-node calculations obtain more of the valence correlation energy than the CC calculations.

5 Diatomics

5.1 Diatomics - homonuclear

Filippi and Umrigar [93] calculated the VMC and DMC total energies for the first-row homonuclear diatomics, Li_2, Be_2, B_2,C_2, N_2, and O_2 at the experimental bond lengths. The exponents and MO coefficients of the Slater basis functions were obtained by variance optimization at the VMC level of theory. The Slater determinant was multiplied by a generalized Jastrow, siimilar to an earlier study [91]. VMC and DMC calculations were carried out using an accelerated Metropolis algorithm and a small-time-step-error algorithm, respectively [87, 125].

To determine the percentage of correlation energy recovered the QMC total energies were compared to Cade and Huo's [220] HF-limit energies. In order to judge the quality of the trial functions, VMC and DMC energies, (E_{VMC} and E_{DMC}) and the standard deviations (σ_{vmc}) of the VMC local energies were compared. For systems that have multi-configurational character, inclusion of higher order CSFs to HF configuration leads to significant improvement in VMC and DMC energies. The correlation energy recovered with a single CSF, plus correlation function with VMC ranged from 67.82(9)% for C_2 to 82.26(5)% for Li_2, while the DMC correlation energy recovered varied from 88.1(2)% for C_2 to 96.5(1)% for Li_2. The correlation energy recovered with multiple CSFs with VMC ranged from 83.10(7)% for B_2 to 87.56(6)% for Be_2, while the DMC correlation energy recovered varied from 93.1% for N_2 to 98.7% for Li_2.

For both VMC and DMC the correlation energy recovered with multi-determinant wave functions increased with Z, for $Z \geq 6$. The auto-correlation times of multi-determinant trial wave functions are reduced compared to the single-determinant wave functions. In addition, local energy fluctuations for single and multi-determinant trial functions increase almost linearly with Z. The multi-determinant wave functions do not exhibit significant variance reduction relative to the single-determinant trial functions.

5.2 Diatomics - heteronuclear diatomics

The dissociation energy of first-row heteronuclear diatomics, LiH, BeH, BH, CH, NH, OH, and FH were examined by Luchow and Anderson with the VMC and DMC method [221]. Trial functions for these calculations consisted of a HF function and the SMBH correlation function at the experimental bond distance in all hydrides except BeH for which a MCSCF trial wave function consisting of 6 CSFs was also used.

The of correlation energy recovered by DMC ranged from 99.9(2)% (LiH) to 92.1(3)% (CH). For LiH, CH, OH, and FH the DMC dissociation energies are within 0.2 kcal/mol of experiment, taking account of the statistical uncertainty of the DMC calculations. The VMC results underestimate experiment by ~2-20 kcal/mol. The authors suggest that this is due to the inadequacy of the optimization procedure. For BeH, the single reference trial function overestimated

the dissociation energy by ~2 kcal/mol. Using a trial function that includes the $1s^2 2p^2$ configuration the DMC calculations yielded a dissociation energy in agreement with experimental value of 49.8(3) kcal/mol.

6 Carbon and silicon compounds

6.1 Small hydrocarbons

Barnett *et al.* examined multi-determinant expansions for acetylene and its dissociation fragments C_2, CH, and C_2H [96]. Multi-determinant wave functions for C, C_2, CH, C_2H, and C_2H_2, were obtained from complete active space SCF (CASSCF) calculations. The CSFs were obtained from CASSCF calculations in which the CSFs with a weight above a predetermined threshold were retained.

The trial wave functions included a SMBH correlation function. The correlation function parameters were minimized in a fixed-sample variance minimization procedure in which the walkers are held fixed and the energy is minimized according to the variance functional.

For C, CH, and C_2, fewer than 20 CASSCF determinants for each system generated greater than 96% of the correlation energy. The C_2H fragment is less well described with correlatin energy recovered in the range of 88.6(6) - 90.5(2)%, with 25 determinants needed for the most favorable case. For C_2H_2, 66 determinants recovered 97.9% of the correlation energy.

Larger determinant expansions of trial functions and satisfaction of electron-nucleus cusp conditions would likly result in better agreement with experiment. Note that the energy converges slowly with respect to the number of determinants in the trial function. A possible contribution to this trend is that only low angular momentum s and p basis functions were used to construct the MOs. To improve convergence, higher angular momentum basis functions such as d, f, and g are likely needed [96].

6.2 Carbon clusters

Carbon clusters are found in interstellar space in addition to various regions of the earth. In the laboratory, carbon clusters, including fulleres, are typically formed by a combination of laser vaporization and annealing processes of graphite. Mass spectra indicate that these experimental techniques yield a variety of cluster sizes. The number of C atoms in these clusters (C_n) range from $10 \leq n \leq 18$ and $32 \leq n \leq 60$. See, for example, the mass spectra described in refs. [222] and [223]. These clusters display a variety of physical and chemical properties that are due to bonding, electrostatic interactions and other factors. Clusters at the lower end of the mass range are believed to be mostly monocyclic rings containing both even and odd numbers of C atoms. Other possible structures include chains and polycyclic rings, although monocyclic structures are believed to be the lowest energy isomers. The clusters at the higher end of the mass scale include an abundance of fullerene structures as well as chain,

mono and polycylic rings. The abundance of the types of clusters generated experimentally depends heavily on experimental conditions. For reviews of experimental and theoretical studies of small C clusters, see, for example, Van Orden and Saykally [224], Weltner and Van Zee [225], and Martin *et al.* [226].

Factors that influence cluster stability include vibrational energies, energetic barriers and thermodynamic stability to rearrangement. Clusters are generally formed at temperatures of 1000 K or greater. At these temperatures, vibrational contributions to the free energy can be significant. Clusters with fullerene-type structures, are more rigid than ring systems. This rigidity causes the fullerenes to have higher vibrational free energies than ring systems, which have many low-lying vibrational energies. As a result, vibrational effects tend to favor ring isomers at high temperatures. High energy barriers to formation explain the absence experimentally of the sheet/bowl structures that have been predicted to be stable at 0K. Thermodynamic stability to rearrangements of different cluster sizes determines which isomers are observed but does not however, determine which cluster sizes are observed. Clusters in the intermediate range have smaller energy differences among competing structures than large clusters, for which energy differences are much greater.

There are several stable isomers of C_{20}; these include such as chains, rings, bowls, and cages (fullerenes). The bowl is essentially a curved pentagonal piece of graphite, while the cage structure is the smallest example of a fullerene. It is not known whether the C_{20} cage structure is stable and several studies have shown conflicting results. For example, Raman spectra indicate the existence of a chain isomer, but not bowl and cage structures [227].

Of the small fullerenes, the most abundant is C_{32}. This system is thought to form through a mechanism in which larger fullerenes are formed by the addition of carbon atoms to smaller fullerenes. It is believed that the first fullerenes are formed for $n \leq 32$; the smallest consists of 20 carbon atoms, although it is not the most stable isomer of C_{20}. Fullerene stability is achieved in the range of $20 < n \leq 30$. Studies examining the stability of fullerenes are discussed below. Factors such as electronic correlation and electrostatic interactions are important in determining relative stability and binding energies of carbon clusters. In certain cases, energy differences between isomers are small due to broad minima on the potential energy surfaces, as in the case of C_8 and C_{10}. Described below are DMC calculations applied to a variety of carbon clusters, including the $C_8 - C_{20}$ isomers which are the work of Sokolova *et al.*, Grossman *et al.*, and Torelli *et al.* [228–230].

6.2.1 Carbon rings and chains

The monocyclic C_n rings, with $k = 4n+2$ ($n = 1-4$) are examples of precursors to the formation of fullerenes and nanotubes. In annealing processes, a non-fullerene C_{60}^+ ion can be converted into a fullerene and a large monocyclic ring isomer. Second-order Jahn-Teller distortion, aromaticity, and Peierls instability cause these ring systems to be stabilized by different mechanisms. Torelli *et al.* [229] and Shlyakhter *et al.* [230] examined specific ring and chain systems to

determine which isomers are most stable.

There are eight possible isomers of C_8 and C_{10}. These are comprised of the singlet and triplet states of the linear and cyclic structures with either acetylenic or cumulenic bonding for both C_8 and C_{10} isomers. For the eight structures, the lowest energy structures from ab initio calculations are the linear cumulenic and cyclic acetylenic structures for C_8, and the cyclic cumulenic, the cyclic acetylenic and the cyclic cumulenic. The geometries of these structures are found in regions of broad minima on the potential energy surfaces; the geometries are insensitive to small variations in nuclear positions.

The ring systems examined by Torelli *et al.* [229] are a cumulenic ring with full D_{nh} symmetry, with equal bond angles and bond lengths, and two distorted ring structures with either alternating bond lengths or alternating angles. The trial functions used in this series of calculations were obtained from MCSCF calculations. Including single excitations in the trial function resulted in significant lowering of the DMC energies and indicate that C_{4n+2} ground state structures have alternated carbon-carbon bond at all sizes, while the cumulenic isomer is a transition state.

DMC calculations were carried out for rings isomers from C_6 to C_{18} [229]. For carbon ring labeling see ref. [229] The B structure is the most stable C_6 isomer because of overlap of the in-plane π orbitals. Isomer A is a transition state for the angle alternating structure, while isomer C is unstable. For the C_{10} ring isomer, the B isomer is the most stable, lying $\sim$1 kcal/mol below the A transition structure, in agreement with CCSD(T) calculations of Martin and Taylor [231]. Isomer B is stabilized by second-order Jahn-Teller distortion, while A is stabilized by aromaticity. The acetylenic isomer appears to be unstable in DMC. In two previous studies, DMC results for this structure indicated that it was stable - a result that is likely due to the use of a HF trial function [232].

The B isomer is unstable for the C_{14} and C_{18} isomers because in-plane π orbital overlap is reduced due to the increased ring radius. The dimer of isomer C is favored at the MCSCF and DMC levels of theory, while LDA favors the aromatic structure due to overestimation of the correlation energy.

Shlyakhter *et al.* who also examined these carbon ring systems, constructed trial wave functions from single determinant SCF calculations multiplied by a 10-parameter SMBH correlation function [230]. To satisfy the electron-nuclear cusp condition, each contracted 1s Gaussian orbital was replaced by a Slater-type orbital.

In this study the DMC total energies for the C_8 and C_{10} structures differ by $\sim$0.5 hartrees. Compared to the CCSD(T) and MP4 methods, the DMC energy for C_8 is lower by 0.4 and 0.8 hartrees, respectively. For C_{10}, the DMC total energy is 0.8 hartrees lower than CCSD(T) and MP2 energies. The DMC calculations predict the cyclic C_{10} structures to be 55-65 kcal/mol lower in energy than the linear C_{10} structure, in agreement with CISD results of Liang and Schaefer [233]and Parasuk and Almlof [234]. The cyclic structure is slightly higher in energy than the cyclic C_{10} structure.

6.2.2 C_{20} clusters

Of the many C_{20} isomers, the ring, bowl, and cage are of lowest energy. Calculations using HF, MP2, CCSD and CCSD(T), LDA, and B3LYP methods give a variety of results dependent usually on size and quality of the basis set. In addition, the number of electrons in these systems limits the size of the basis set used. For example, the CCSD(T) method is limited to basis set of cc-pVDZ quality. The HF method predicts the ring to be most stable; the ring/cage separation is ~4.5 eV. MP2 calculations using s small basis set, *e.g.* 6-31G(d,p), predict the cage as most stable while similar calculations with a large basis set e.g. cc-pVTZ, favors the bowl. The CCSD method favors the bowl while CCSD(T) favors the cage. DFT calculations using LDA predict the reverse of the HF ordering with a ring/cage energy difference of 4.0 eV. GGA results favor the HF ordering [8].

DMC calculations by Grossman *et al.* [228] using single determinant trial wave functions and non-local pseudopotentials, indicate that the bowl is the most energetically stable isomer. This finding contradicts DFT results, subsequent all-electron DMC calculations confirm te ECP-DMC results[235].

The ring structure contains large Coulomb and exchange contributions. The correlation energy in the ring is 3.4(2) eV less than the cage structure. The cage contains the largest amount of correlation energy due to the number of sp^2 hybrid bonds. The bowl structure has less correlation energy than the cage due to electrostatic interactions and, as a result, has the lowest total energy.

Comparisons of the HF and LDA energies reveal that the cage and bowl energies vary by less than 0.5 eV for different geometries. The various ring geometries indicate that this structure is energetically more than twice as sensitive as the cage. The effect of different optimization schemes on relative energies is small; see, for example, the study by Taylor *et al.* [236].

6.2.3 Larger fullerenes $C_{24} - C_{32}$

As stated earlier, stable fullerenes occur in the range of 20-30 atoms. At the lower end of this range is the C_{20} cage isomer, the first example of an unstable fullerene, while at the upper end of the range C_{32} fullerenes are found in abundance. To identify the first stable fullerene, Kent *et al.*[237] examined a variety of isomers ranging from C_{24} to C_{32}. In this range, three classes of isomers are energetically competitive: fullerenes, planar sheets, bowls, and monocyclic rings.

The isomers that were examined in this study included five isomers of C_{24}, three isomers of C_{26} and three isomers of C_{28}. Geometries for the isomers were obtained at the B3YLP/cc-pVDZ level of theory. The sensitivity of the energy to geometry was ascertained by comparing the B3LYP and BLYP geometries of the C_{26} ring and D_6 fullerene. Differences in geometries were 0.03 A for bond lengths and 0.04° for bond angles. Differences in energy of 0.14 eV were obtained for the BLYP and B3LYP relaxed geometries [237].

For the C_{24} isomers, DMC calculations predict the graphitic sheet to be more

stable than the ring and cage structures. The stability of the sheet is primarily due to the low strain of the hexagonal rings. Of the C_{26}, C_{28}, and C_{32} isomers, the fullerene C_{26} was determined to be the most stable structure, indicating that it is the smallest stable fullerene. The results for C_{32} are consistent with cluster experiments that indicate a large abundance of C_{32} fullerenes.

Reduction of ring strain causes the binding energies per atom of the ring and fullerene structures to gradually increase with cluster size. However, the presence of ring strain in the fullerene structures cause the binding energy to increase more steeply for these structures than for rings. The large number of high-energy pentagonal rings and the compactness of the sheet/bowl structures causes the binding energies to vary less smoothly with system size than for the other structures. Energies of the fullerene and ring structures cross at C_{24}, indicating that below this size rings are more stable, and above it, fullerenes are more stable.

Results for C_{24} isomers indicate that the graphitic sheet is most stable, in comparison with C_{20} calculations by Grossman *et al.* [228] and Sokolova *et al.* [235] that indicate that a bowl structure is the most stable. This means that a range exists in which the sheet/bowl structures are energetically more stable. Neither the sheet or bowl structure has been identified in experimental studies. The DMC calculations of the C_{26} and C_{28} fullerenes are several eV per cluster lower in energy than the other isomers.

6.3 Hydrogenated silicon clusters

The absorption and emission of light in Si quantum dots and porous Si can be modeled using hydrogenated Si clusters. Optical properties of Si clusters are influenced by electron-hole interactions. Theoretical studies have been carried out using empirical tight-biding DFT, and perturbation theory techniques. Greeff and Lester computed DMC bond and atomization energies of SiH_n ($n = 1-4$), Si_2H_6, and Si_3H_8, SiH_3^- [238]. The calculated atomization energies for SiH_4, Si_2H_6, and Si_3H_8 showed that no subsequent corrections were needed that assumed that the final state of Si is amorphous. Grossman *et al.* have reported IPs and EAs,and first excited singlet and triplet energies of SiH_4 [239]. Needs *et al.* and Porter *et al.* determined the excitation energies of the lowest triplet and bright, dipole allowed, singlet transitions of SiH_4, Si_2H_6, Si_5H_{12}, and $Si_{10}H_{16}$ clusters at the CIS and DMC levels of theory [240–242]. A more in-depth discussion of the optical properties of quantum dots and QMC calculations of these systems are discussed below.

In the DMC study by Greeff and Lester the geometries for SiH, SiH_2 and SiH_4 were obtained from Allen and Schaefer [243]; the geometry for $Si_2H_6$was obtained from that of Leszczynski *et al.* [244] and the SiH_3 and Si_3H_8 geometries were calculated at the HF level. The wave functions used in this study consisted of single HF Slater determinants multiplied by a 9-parameter SMBH correlation function; the core electrons were described by SBK ECPs. The calculated DMC atomization energies for SiH_n ($n = 1-4$) were in agreement with large basis set CCSD(T) calculations of Grev *et al.* [245]. The DMC atom-

ization energy for SiH_2 was found to be in accord with experiment. For SiH, SiH_3, and SiH_4 the atomization energies were found to be ~1.5 kcal/mol above the experimental values; for the Si_2H_6 and Si_3H_8 hydrides the DMC results are within 2.5 kcal/mol of experiment. The VMC and DMC results for the Si-H bond energies for SiH_4, SiH_3, SiH_2, SiH are within 2 kcal/mol of experiment while the Si-Si bond energies of the Si_2H_6 and Si_3H_8 hydrides are within 3-4 kcal/mol of experiment.

Porter *et al.* performed DMC calculations for SiH_4 using CIS and time-dependent (TD-LDA) wave functions[241]. In this study, the ground and excited state energies and the vertical ionization potentials were calculated. The Si core electrons were replaced with a HF ECP.

For the ground state of SiH_4 the LDA optimized structure was used. The DMC atomization energy of 13.38(2) eV is close to the value of 13.82(2) eV obtained by Greeff and Lester [238]. These results are in close agreement with CCSD(T) CBS results of Dixon and Feller [246]. The DMC vertical IP of 12.6(1) eV is in close agreement with the experimental values of 12.36 and 12.85 eV [241].

The first few dipole-allowed singlet excited states of silane were examined by Porter *et al.*; these include the $2t^2 \rightarrow 4s$, $2t^2 \rightarrow 4p$, $2t^2 \rightarrow 5s$, and $2t^2 \rightarrow 4d$ excited states. Note no experimental data are available for the $2t^2 \rightarrow 4p$ and $2t^2 \rightarrow 4d$ transitions [240, 242]. DMC calculations were performed on the excited states using CIS trial functions and compared to CIS, MRCI and experimental results. For the $2t^2 \rightarrow 4s$ and $2t^2 \rightarrow 5s$ transitions, the CIS results are in close agreement with the experimental value of 9.7 eV while the DMC calculations using a CIS trial function and MRCI results are found to be ~0.2 and ~-0.3 eV below experiment, respectively.

To gauge the quality of the CIS trial function DMC calculations were carried out on the ground and lowest singlet and triplet excited states using HF, TD-LDA and CASSCF trial functions. The HF trial function was obtained by promoting an electron from the highest occupied molecular orbital (HOMO) to the lowest-unoccupied molecular orbital (LUMO). The DMC results for the singlet transition with CIS, TD-LDA, and CASSCF trial wave functions are the same, ~ 0.3 eV below the experimental value of 9.7 eV, while the HF trial function overestimates the energy splitting by ~0.9 eV. For the triplet energy splitting the HF trial function overestimates by 0.5 eV while the CIS and CASSCF trial wave functions overestimate the triplet transition by ~0.1 eV. The LDA trial functions yields results in agreement with the experimental results of 8.7 eV.

Porter *et al.* also examined the lowest-triplet, lowest-bright (dipole allowed)-singlet and lowest-dark (dipole-forbidden singlet transitions of SiH_4, Si_2H_6, Si_5H_{12}, and $S_{10}H_{16}$ clusters[242]. DMC calculations were carried out using CIS and LDA trial functions. The DMC results were compared to CIS and TD-LDA results as well as experiment where available.

For the lowest singlet and lowest triplet state, excitation energies decrease with increasing cluster size. CIS overestimates singlet excitation energies due to a poorer description of the correlation energy in the excited state than in the

ground state energies. The DMC results are lower in energy than the CIS results and underestimate the experimental results. The discrepancy with experiment is due to the role of vibrational and Jahn-Teller effects in these systems, which tend to lower the excitation energies and were not included in the DMC calculations. For the lowest singlet and triplet states, the TD-LDA results are lower than experiment, while for the dark singlet excited state, the TD-LDA results are lower than the DMC results.

7 Transition metal systems

The pursuit of accurate electronic structure calculations of transition metals is complicated by several factors including relativistic effects, near degeneracy of 3d, 4s, and 4p levels, compactness and double occupancy of d shells, and semi-core character of 3s and 3p shells. The latter two effects are significant and are responsible for atypically large correlation energies of 20 eV or more. The core electrons of transition metals are very tightly bound due to large Z and relativistic effects. For these reasons, it is not possible to carry out all-electron calculations for transition metal systems and achieve desireable variance.

Wave functions from independent particle theories such as HF used in the DMC calculations can provide excellent estimation of transition energies, however, large discrepancies with experiment are observed for atomization energies. This suggests that electron-electron correlation has an important effect on the nodes of the trial wave functions.

The largest part of 3d, 3p, and 3d correlation energy arises from the repulsive electron-electron cusp the wave function. The electron-electron cusp is not well described by most correlated methods, which exhibit slow convergence in recovering the cusp. The QMC method explicitly accounts for the cusp and offers an attractive alternative for treating molecular systems containing transition metals.

To date, there are several QMC studies of transition metals systems. For atoms and dimers, recent QMC applications include the Fe atom by Mitas who calculated IPs and EAs and the $5D \rightarrow 3F$ and $5D \rightarrow 4F$ transitions [247]; the Sc and Y ionization potentials computed by Christiansen [248]; the 5F-3F splitting in Ti and the singlet-triplet splitting of TiC by Sokolova and Luchow [249]; the $d^5s^2 \rightarrow d^6s$ transition in Mn and the binding energies of TiO and MnO by Wagner and Mitas [250]; the ground state properties of Hg_2 by Flad and Dolg [251]; and the group 12 dimers of Zn, Cd, and Hg by Schautz and Dolg [252, 253]. The number of QMC studies of TM clusters is small. Dolg and Flad explored size dependent properties of Hg clusters [254], Ovcharenko *et al.* examined small Cu-doped silicon clusters [255] and El Akramine *et al.* studied the interaction of CO with a model Cr (110) surface [256].

7.1 Fe atom

Mitas studied Fe-atom transitions using Ne-core ECPs [247]. The orbitals for Slater determinants were obtained by the numerical HF method. The trial function included excitations to account for near-degeneracy effects.

The VMC method yielded a small variational bias for the IP and the $5D \rightarrow 5F$ excitation energy, while a large bias was found for the $5D \rightarrow 3F$ excitation energy and the EA. This discrepancy is believed due to a less accurate trial function for the negative ion and the 3F state. Errors arsing from using a ECP were determined to be $\sim$0.1 eV. Other possible sources of error include non-rigorous treatment of relativistic effects and the fixed node error. Further DMC calculations determined that ECP localization error was small.

7.2 TiC

There have been many theoretical and experimental studies of the titanium carbon cluster, Ti_8C_{12} discovered by Guo *et al.* [257]. An entire family of titanium-carbon clusters has been created, called metallo-carbohedrenes, which have unusual stability. There are two possible structures for the $T_{18}C_{12}$ cluster; a cage-like pentagonal dodecahedron (T_h symmetry) and a tetra-capped tetrahedral cage structure (T_d symmetry).

Diatomic TiC has been investigated at the CCSD(T) and MRCI levels of theory. To date, no experimental data exists for TiC, and comparisons are restricted to other ab initio calculations. For Ti, Ne core ECP constructed by Hay and Wadt [258, 259] was used, which leaves the 3s and 3p as well as 3d and 4s electrons in the valence space. For C, the SBK ECP was used. The 5F-3F splitting of Ti and the energy splitting of TiC were computed [260]. The DMC results for the Ti 5F-3F splittings are in good agreement with non-relativistic limit estimate. However, for the DFT methods, only the hybrid B3LYP functional gives the correct order of states. The MRCI and CCSD(T) methods over-estimate the energy splitting of the states [261].

For the lowest triplet state of TiC, a ROHF trial function was used for the DMC calculations with an optimized geometry obtained at the MP2 level of theory. The singlet excited state is low-lying and has significant dynamic and non-dynamic contributions to the correlation energy. For this state, the geometry was obtained using s 55-term MR-SDCI calculation; the DMC trial function was constructed by retaining only the 4 dominant CSFs of the MR-SDCI calculation.

The calculated energy splitting of TiC varies with method. Both the MR-CI and CCI calculations give a single-triplet splitting $\sim$0.4 eV smaller than the DMC result. The DMC values of the binding energy are 0.5 eV larger than the MR-CI and CCI values.

7.3 Transition metal oxides

Wagner and Mitas calculated Mn atom $s \rightarrow d$ promotion energies as well as the TiO and MnO binding energies. The trial function consisted of orbitals obtained from HF, MCSCF and DFT calculations [250]. Core electrons were replaced by ECPs, and for Mn and Ti, the ECPs were generated from Dirac-Fock calculations with a Ne-core, i.e., 3s 3p electrons were in the valence space and a soft ECP was used for O.

In the MCSCF calculations, the transition metal 3s and 3p states were frozen and single and double excitations from higher orbitals carried out to obtain an improved trial function. The natural orbitals (NOs) of the MCSCF calculation were used in a subsequent CI calculation. For the DMC calculations, determinants with a weight above 0.05 for MnO and 0.01 for TiO were retained. These cutoffs in the CI weights reduced the number of trial function determinants to approximately 20. For MnO, calculations involving active spaces of 10, 12, and 17 orbitals were performed in both the MCSCF and DMC methods. The larger active space improved the MCSCF energy, but had little effect on the DMC energy.

The $s \rightarrow d$ excitation was examined with HF, DFT, and DMC methods. The HF method overestimates the excitation energies by 1.3 eV while LDA underestimates it by 0.62 eV. The GGA functionals do not show consistent improvement over LDA energies. DMC energies obtained using either HF or B3LYP trial functions yield excitations energies in accord with experiment.

Using HF trial functions, DMC calculations the TiO binding energy is underestimated by 0.6 eV and that MnO by 0.8 eV relative to experiment. It is anticipated that these discrepancies can be reduced with multi-determinant MCSCF trial functions. For TiO, the DMC binding energy obtained using a B3LYP trial function is improved over the value obtained with the MCSCF trial function. For MnO, no improvement in binding energy is achieved; both the B3LYP and MCSCF trial function yields results that are in agreement with experiment.

8 Benchmark studies

8.1 Gaussian-1 set

The Gaussian-1 (G1) set initally was comprised of 31 neutral molecules that served as the test set for the G1 method [262]. Later, a set of 24 molecules was added to represent a wider variety of chemical species [263]. The G1 method has a mean absolute deviation (ε_{mad}) from experiment of 1.6 kcal/mol and a maximum deviation of 7.4 kcal/mol on the G1 set. The experimental uncertainty for much of the experimental data such as atomization energies and heats of formation is less than 0.5 kcal/mol.

Recently, in order to gauge the accuracy and predictive capability of QMC, Grossman calculated the atomization energies of the G1 set using single determinant DMC [264]. The core electrons were replaced by the SBK ECPs and a basis

set of 6-311++G(2d,2p) quality was used for the valence electrons [157]. The average absolute deviation of the DMC results is 2.9 kcal/mol with a maximum deviation from experiment of 14 kcal/mol. Systems found to have the greatest deviation from experiment in descending order are: SO_2, NO_2, ClO, CN, LiF, and NO. The 27 poorest results underbind, on average, by 2.7 kcal/mol while the 27 best cases overbind by just 0.2 kcal/mol. Possible sources of discrepancy between theory and experiment are: fixed-node errors, experimental zero point energies, geometry specifications, and experimental atomization energies. Grossman's G1 set results are also presented in Table 4.

Comparing to other theoretical methods, the ε_{mad} for LDA heat of formation is $\approx$40 kcal/mol and 2.5 kcal/mol for DFT using either the B3LYP or B3PW91 generalized gradient approximation (GGAs). The ε_{mad} for CCSD(T)/aug-cc-pVQZ is 2.8 kcal/mol which is reduced to 1.3 kcal/mol by extrapolating to the complete basis set (CBS) limit. Feller and Peterson's CCSD(T) calculations at the CBS limit yielded ε_{mad} of 0.7 - 0.8 kcal./mol for 73 molecules chosen from the G2 set of 155 molecules [265, 266]. The largest deviations were 2-3 kcal/mol above experiment. The largest available correlation consistent basis sets were used in this study (up to aug-cc-pV6Z for some systems). These CBS results for the G1 set are presented in Table 1.

In several recent studies, Lu determined the atomization energy for a number of first-row polyatomics using the OUDQMC (Ornstein-Uhlenbeck diffusion Monte Carlo) method [189, 267, 268]. The results of this study can be compared to experimental atomization energies and with Feller and Peterson's CBS results and Grossman's results by correcting for zero point energy [265]. The OUDQMC results were obtained with ZPEs estimated following the procedure of Grev *et al.*[269].

Because only 14 systems of the G1 set are found in the OUDQMC study, it is difficult to draw generalizations. Agreement is found among these approaches for H_2O, NH_3, HF, and H_2CO to within statistical uncertainty. The OUDQMC atomization energies for F_2, LiF, N_2, CO_2 and CH_4 are in better agreement with experiment than Grossman's values and reflect improved nodes of the variationally more flexible trial functions of the OUDQMC approach. Atomization energies obtained by the various methods are listed in Table 1.

8.2 Reactions

Manten and Luchow performed all-electron DMC calculations on a set of 20 molecules and 17 reactions selected by Klopper *et al.* as tests of the accuracy of the CCSD(T) method [65, 270]. The DMC calculations were carried out with HF trial wave functions. The *1s* and *2s* basis set functions of the cc-pVTZ basis set were adjusted to satisfy the electron-nucleus cusp condition using a fitting function:

$$f(r) = ae^{-\alpha r} + br + c \tag{155}$$

where the parameters a, α, b, and c were obtained by numerically fitting

Eq. 155 to the basis set. In the region near the basis function origin, the fitted function was smoothly connected to the basis set by enforcing continuity of first and second derivatives. The variance of the DMC calculations is reduced because the Laplacian near the nucleus is smooth, which improves the local energy during the random walk. In addition, the time step error of the calculation is drastically reduced [65].

Geometries for the DMC calculations were obtained at the MP2/cc-pVTZ level of theory. The DMC total energies for the 20 molecules are lower than the CCSD(T)/cc-pVXZ (X=D,T) and CBS energies in part owing to the frozen core approximation of the CCSD(T) calculations. The DMC and CCSD(T)/cc-pVXZ (X=D,T) total energies for the molecules studied are presented in Table 5. The DMC total energies are on average 16 kcal/mol above the CBS energies obtained with the quadratic CI atomic pair natural orbital (QCI/APNO) method; the largest difference is for O_3 and the smallest is for H_2.

For the 17 reactions, the all-electron DMC energies differ from experiment less than the CCSD(T)/cc-pVDZ results and more than those of the CCSD(T)/cc-pVTZ method. The CC methods are strongly basis set dependent as shown by the slow convergence of correlation energy with basis set [271]. Only for the largest basis sets, such as cc-pVQZ and cc-pV5Z, do errors of the CC methods appear [271].

Compared to experiment, the DMC method underestimates the heat of reaction for all but one reaction, while CCSD(T) overestimates the majority of the reactions using the cc-pVDZ and cc-pVTZ basis sets; see Table 6. For the three levels of theory, O_3 reaction differs the most from experiment. To accurately describe this reaction a multi-configuration approach is likely needed; both the CCSD(T) and DMC calculations are single reference.

The DMC ε_{mad} is 5.9 kcal/mol which is to be compared with 8.2 kcal/mol for CCSD(T)/cc-pVDZ, 2.4 kcal/mol for CCSD(T)/cc-pVTZ, and 0.6 kcal/mol at the CCSD(T)/CBS limit. The ε_{mad} for QCI/APNO CBS is 0.8 kcal/mol [270]. For the G2 and B3LYP methods, ε_{mad} is 2.6 kcal/mol [266].

Omitting the reaction involving ozone and formation of HOF+HF, the ε_{mad} for DMC and for CCSD(T)/cc-pVTZ decreases by $\sim$1 kcal/mol while for the G2 and B3LYP, the improvement is $\sim$0.5 kcal/mol. The maximum absolute error (MAE) for CCSD(T)/cc-pvTZ is 0.7 kcal/mol and 13.1 kcal/mol for DMC. The DFT, G2, and CBS results have a smaller MAE range of 7.9 - 1.4 kcal/mol. The HOF+HF reaction is significantly underestimated by DMC. In part, this is due to F_2 for which the DMC atomization energy has been shown to be significantly understimated in calculations[264].

8.3 Excited states

Excitation energies are labeled as "1/N" effects because the change in energy is inversely proportional to the number of electrons in the system, i.e., as the number of electrons increases, the spacing of the energy levels decreases as 1/N. Therefore, precise calculations are required to determine the change in energy from the ground state within the statistical uncertainty encountered in QMC

calculations. Single determinant trial HF wave functions can be used to describe open-shell triplet states. Other excited states such as open-shell singlet states, require multi-configuration trial functions. For example, DMC calculations by Grimes *et al.* [167] examined the excited states of H_2 using a simple MCSCF trial function and obtained results for the B $^1\Sigma_u^+$and E $^1\Sigma_g^+$states that are in excellent agreement with the exact energies of Kolos and Wolniewicz [272]. There are several other examples of excited-state QMC calculations in the literature, including the excited triplet states of silane, and hydrogenated silicon clusters, discussed in Sec. 6.3. More recently, calculation of excited states of free-base Porphyrin agree with experiment within statistical error [**?**].

9 Nano-systems - introduction

As atomic and molecular systems become smaller in size, interesting changes in physical properties occur. The volume to surface area ratio decreases which makes surface properties have an increasingly important role in the behavior of the system. Further, quantum effects start to appear owing to the diminished space and one has quantum confinement. Nano materials constitute physically interesting systems for study as their size lies between the macroscopic and the microscopic world making them ideal objects for study of this transition region. Nano materials are also being widely studied because of their potential use in different fields as diverse as biology, laser technology and optoelectronics.

Because of their size, typical lack of symmetry and unusually large electronic correlation interactions, nano materials present difficulties for most quantum chemical methods. A wide range of electronic structure approaches have been used to study nano systems and include the familiar ab initio methods of HF, DFT and time-dependent DFT (TD-DFT) as well as semiempirical approaches. While these methods have provided considerable insight, the accuracy has not always been adequate for the purpose. DFT methods are the most widely applied to these systems, however, there is presently no way of systematically improving the results obtained, nor is there a satisfactory approach for describing exited states, which limits their use. As noted above, QMC methods are a proven approach for recovering correlation energy and furthermore, these methods scale, in general, as $O(N^3)$ with system size.

9.1 Silicon nano-clusters

Nano semiconductors have been widely studied in the past few years. As electronic devices have become smaller, the need for understanding the behavior of these materials has become crucial. As semiconductors are reduced in size, the optical band gap increases. Semiconductor nano-clusters, commonly known as quantum dots, usually emit light in the visible region of the spectrum. Since the surface to volume ratio is large for these systems, surface reconstruction has important implications for their properties. Nano-clusters formed of atoms from the II and IV groups of the periodic table such as CdSe, have been widely stud-

ied experimentally [273–275]. These materials are easily synthesized with small size dispersion and with well passivated surfaces. For quantum dots, the main property affecting the behavior of the system is thought to be quantum confinement, which arises from the smallness of the spatial region in which electrons are restricted. As the system size decreases, the optical gap increases and the emission spectrum narrows relative to bulk values. Surface reconstruction has not been studied in detail in these systems, but it is thought to have a much smaller impact than quantum confinement [276]. On the other hand, group IV semiconductors such as Si clusters are more difficult to produce with a specific size, and passivator effects are crucial in controling their optical properties. The potential uses of Si nano-clusters are considerable, especially in silicon-based electronics and biological imaging. For Si nano-clusters, the specific surface chemistry, in addition to quantum confinement, must be taken into account in order to be able to interpret and to predict phenomena correctly.

9.1.1 Optical gaps and size dependence

Silicon quantum dots can be manufactured in different sizes ranging from a few atoms to large clusters. Small clusters have been frequently studied because of their large luminescence. It has been observed that different clusters emit light at different wavelengths, opening the possibility of several optical applications. This behavior is also interesting theoretically because of the transition from atomic-like properties to those of the bulk. Silicon-based technological applications are varied is size requirement , but the need for smaller materials has increased with the technological demand for size reduction of various types of components. As a result, understanding the behavior of such materials and the connection of that behavior with the electronic structure as clusters grow is crucial to further understanding and prediction.

To study optical gaps requires the use of methods that make possible calculations with consistency over a range of system sizes for the particular transition of interest. A typical model for the study of bulk Si is that of an Si cluster with dangling bonds capped with H atoms.

Optical gaps have usually been calculated in QMC as the energy difference between the first singlet exited state and the ground state energy.

$$E_{opt} = E^* - E_{GS} \tag{156}$$

This is done by promoting an electron from the HOMO to the LUMO orbital to calculate the dot in its first exited singlet state.

Calculations of the optical gap of silicon nano-clusters, Eq. 156, with sizes ranging from 0 to 3.0nm capped with H and with no reconstructed surfaces by Williamson *et al.* [277], show that the optical gap decreases as the size of the cluster increases and quantum confinement is reduced; see Fig. 1 of ref. [277]. All levels of theory show the same trend. In general, semi-empirical and LDA approaches, even in the time dependent TD-LDA formulation, yield gaps that are too low, with an error of 1-2 eV compared to DMC. The GW approximation

does significantly better than the other methods; however, it scales poorly with system size, which limits its applicability.

Optical gaps obtained with DFT using B3LYP demonstrate that this method has smaller errors than LDA but, at low electronic densities, it overestimates optical gaps while for big clusters it underestimates this property, as shown in Fig. 2 of ref. [277]. It is probably not surprising that B3LYP gives better results than LDA for these systems because silane (SiH_4) was one of the molecules used to parametrize the functional.

Calculations by Mitas *et al.* [278] show that silicon dots of different sizes have different gap sizes and can emit at different wavelengths. For example, dots of 1.0 (Si_{29}), 1.67 (Si_{123}), 2.15 and 2.9 nm size can emit light in the UV/blue, green, yellow and red respectively. This can have great applicability in computer monitor displays, biomedical tagging and flash memory.

9.1.2 Optical gaps and surface reconstruction dependence

Surface reconstruction can dramatically change the absorption spectrum of Si dots and it has been proposed as a source of optical activity in ultra-small particles. A reconstructed surface H-Si-Si-H dimer is generated each time two H atoms are eliminated from the surface. Based on experiments and theoretical calculations of reconstructed Si surfaces, there exist two possible geometries for the H-Si-Si-H dimer, buckled and symmetric [279]. Despite the many studies carried out on this system, controversy remains over which of the structures constitutes the ground state. A DMC calculation has been carried out on a small model of the system to represent the relevant part of the surface. In this calculation a spin DFT (SDFT) [280] wave function was used as the trial wave function and gives the buckled dimer as the ground state.

Mitas *et al.* [273] have reported calculations of a silicon nano-cluster with reconstructed surfaces. Calculations for the absorption spectrum, absorption edge and some other properties for $Si_{29}H_x$ nano-clusters prototypes, where x corresponds to the number of H atoms used to terminate the surface on each structure are presented in [273].

The absorption spectra were calculated using DFT, CIS and DMC. The latter approach was used to calculate the absorption edge which is used as a shift correction to CIS generated spectra. The CIS spectra are known to correctly reproduce the number and relative positions of bands, but not the exact values of the positions due to incomplete treatment of electron correlation. The spectrum obtained was broadened with a Gaussian function to account for temperature effects.

The spectra shown in ref. [273] agree well with experimental data. The overall spectra and the band strengths are in good agreement. However, there are some differences in structural details. As the number of reconstruction sites increases, the absorption edge of the dot is shifted to lower energies. For example, $Si_{29}H_{36}$, which has all superficial Si bonds terminated in H, the absorption edge is found at 4.8eV, while for $Si_{29}H_{24}$, which contains six reconstructed sites, and $Si_{29}H_{12}$, which contains twelve reconstructed sites, the absorption edges are

found at 3.5 and 0.1 eV, respectively. With the loss of more than 24 H atoms, the gap for this system drops sharply due to a change from a sp^3 to a sp^2 network. Absorption spectra for each of these clusters show highly molecular (discrete) band patterns, with $Si_{29}H_{12}$ being the most complicated.

Steps are another way of reducing H coverage of the surface. A step is created when a Si-Si bond is formed from neighboring Si atoms - one Si bonds to two H atoms and the other bonds to a H atom and another Si. The effect of surface dimers appears to be important for small clusters. Studies of larger clusters, ranging in size from 53 to 331 silicon atoms, show that the presence of steps in the surface is more important than dimers for optical properties such as the optical gap. The presence of dimers in the surface helps to lower the optical gap and steps in the surface enhance this effect. These structures cannot be formed in very small dots. Therefore, the main defect determining the size of the optical gap in these H terminated silicon clusters is the number of surface dimers.

9.1.3 Optical gaps and passivant effect

Passivants are substances that are added during cluster fabrication to terminate their growth. Computational results by Puzder *et al.* [276, 281–283] and Belomoin *et al.* [284] indicate a close connection between the type of passivant used to terminate Si bonds in a quantum dot and the optical properties observed for the system. Table 7 shows results for the optical gap obtained for $Si_{35}H_{36}$ and similar dots when one H atom is replaced by another element or molecule, or substituting two H atoms if the passivant generates a double bond when it links to the surface.

The results in Table 7 indicate that singly bonded passivants do not appreciatively contribute to optical transitions even when they are highly electronegative. Doubly-bonded passivants, however, do affect the optical gap. The reason is attributed to the distortion in the sp^3 network at the surface.

Contaminants in the synthesis of these materials can lead to dots with very different behaviors than those anticipated. Such behavior opens the possibility of constructing Si-based dots with desired properties.

Results by Puzder *et al.* [276, 281–283] show that increasing the dot size with oxygen as a contaminant also reduces the optical gap, due to the reduction of quantum confinement. The rate of decrease of the optical gap with oxygen is, however, much slower than that observed for dots containing only H termination. This can be appreciated in Fig. 1 of ref. [276].

Increasing the amount of oxygen in a dot decreases the gap even further. Oxygen can bond with Si atoms in a dot in different ways. It can form a double bond with a single Si atom or bond with adjacent Si atoms to form a bridge or it can occupy an interstitial site. A gap reduction is found again related to the extent to which these bonds distort the sp^3 network. The biggest effect is observed with oxygen doubled bonded to a Si atom. The bridged bond also reduces the gap to a large extent, while the interstitial oxygen does not seem to affect the gap significantly. However, if no SiH_3 termination exists, as in

the interstitial case, the lattice distortion might be larger than that discussed, which would produce a bigger effect in the gap.

All the calculations referred to in this section were performed at geometries that are modifications of a bulk Si diamond-like cluster. The changes arise from reconstruction of the surface owing to passivant termination. Depending on the passivant, the surface follows either a tetrahedral sp^3 network or an sp^2 network. The distances used for the passivant-Si bonds were taken from smaller molecules; for example, the Si-H distance was taken to be that in SiH_4. Structures were then relaxed to zero temperature using molecular dynamics simulations [196].

9.1.4 Synthesis dynamics effects

The previous two sections were concerned with properties observed in Si nano-clusters having all Si atoms situated in a diamond-like structure throughout the dot. All effects discussed above are related to changes arising from modifications in the structure of the surface. Although these studies are of great value for understanding the importance surface chemistry can have in dictating the behavior of materials in this size range, they do not always fully explain experimental results.

Synthetic procedures can have a large effect on the properties of Si dots. Small Si nano-clusters produced at high temperatures by methods such as vapor deposition tend to develop a non-crystalline core structure. Crystalline structures of these clusters do not seem to be dynamically favored unless the process involves starting from a crystalline structure, such as etching or sonification, done at low temperatures or with rapid passivation.

For nano-particles larger than 3nm, describing the core as a tetrahedral network is a good description; for smaller dots this is not always the case [285–289]. There is a large difference in the optical gap observed for a Si dot depending on the way it was fabricated. The Si dots obtained by chemical vapor deposition have optical gaps that are several eV lower than fabricated using inverse micelles [290, 291]. The significant point is that alternative fabrication techniques lead to different structures in the dot core and therefore different physical properties.

In Fig. 1, Draeger *et al.* [290, 292] show that different molecular structures, generated for an amorphous cluster of 29-30 Si atoms using a first-principles molecular dynamics approach (FPMD) [196], have different optical gaps. Structures presented in Fig. 1 were obtained starting from an amorphous Si cluster generated by running FPMD calculation of bare Si a cluster at high temperature, and later adding H atoms to passivate the surface. As shown in Fig. 1, structures (a) and (c) correspond to atomic arrangements that would be obtained at low temperatures, and therefore have a crystal-like structure. Two different structures are shown; (a) $Si_{29}H_{24}$ having a single Si as a core, or (c) $Si_{30}H_{22}$ with two tetrahedral Si atoms, double-core, structure. Structures (b) and (d) correspond to dots of 29 and 30 Si atoms, respectively, obtained by FPMD as explained above. While the Si_{30} structure does produce a double-core structure as the ideal case, the Si_{29} structure also forms a double-core unlike the ideal case. At high temperature, structures with larger numbers of core atoms are

more favorable (double-core), and therefore, in this case, a core of 30 atoms is more favorable, by symmetry arguments, than the regular 29-atom core, see Fig. 1 (b) and (d). Structures with a double core formed easily in all FPMD simulations [290]. During the formation of the cluster at high temperature, bare Si dots become amorphous. In an attempt to saturate as many dangling surface bonds as possible, a double-core structure is formed. Transitions between double- and single-core structures are highly unfavorable, even when enough H atoms exist to generate a single-core dot, since the global surface reconstruction would involve the breaking of several bonds with resultant high activation energy. The energy of the cluster formed when the geometry is relaxed at high temperature is considerably different from that of a cluster with an ordered core.

DMC and LDA calculations by Draeger *et al.* [290] on Si clusters are presented in Table 7. The results for two geometries, the crystal-like structure and one obtained by FPMD, using for the latter the geometry corresponding to the cluster that gave the largest gap in relaxation, show that optical gaps obtained in double-core structures are smaller than those from single-core dots and agree well with experiment [293, 294].

Energies of relaxed double-core clusters show good agreement with experimental data for small dots. Small dots favor the formation of amorphous cores and usually resulting in a double tetrahedral core. Calculations of crystalline structures show a significantly larger gap. For large clusters, a single, ordered core obtained by FPMD (at low temperature) or by taking the diamond-like structure of bulk Si show good agreement with experiment. Large clusters tend to favor the crystal tetrahedral unique core structure.

9.2 Two-dimensional quantum dots - introduction

Two-dimensional quantum dot can be constructed by bringing together two thin semiconductor layers, generally GaAs, and generating a two dimensional electron gas (2DEG) between those layers. Electrons are restricted to move in only two dimensions, which is the reason these dots are referred to as two-dimensional quantum dots (2DQD). The electrons are free to move in the plane and can be modeled as free particles. These electrons can be further confined by applying voltages, gates, and similar constraints. Depending on the type of confinement, the dot dimensionality can be reduced to one or zero dimensions, and the geometry of the system can be defined. For example, a one-dimensional electron gas (1DEG) also known as a quantum wire can be generated. The most common type of confinement studied is a harmonic potential. This confinement reduces the dimensionality to zero (0DEG). In this circumstance, electrons are confined to a well, similar to the situation in an atom. For this reason, such dots are usually referred to as artificial atoms. Properties such as addition energy, shell structure and spin configuration can be determined experimentally. For quantum compting applications, these devices must be small and be kept at low temperatures to avoid vibrations and possible destruction of quantum coherence.

Both 2DQD and 0DQD have recently grown in interest because of their potential use in the construction of quantum computers and nano electronic

devices. Understanding the electronic properties of these systems as well as their interaction with electric currents and magnetic fields is of great importance to the development of these applications. Owing to its size, a QD requires quantum mechanics to be understood, however, the electronic structure can only be solved exactly for dots containing a small number of electrons. Theoretical methods able to describe electron correlation accurately are needed to describe and to predict experimental results. As with many other fields, the methods of choice for studying these systems are HF, DFT in the LDA approximation, SDFT [280], current SDFT (CSDFT) [295], exact diagonalization methods where possible, and QMC.

Although LDA offers a fast and efficient method of calculation, the approach contains uncontrolled approximations and the results obtained are not eigenfunctions of the square of the total spin operator S^2. HF theory has the HF exchange energy, but by definition no correlation energy. Correlation effects are essential for describing these systems correctly. LDA includes an approximation to both exchange and correlation. Exact diagonalization methods offer a way of obtaining exact results, but these procedures are available only for highly symmetric systems with a small number of electrons ($n <\sim 7$). The QMC method is a good alternative for treating these problems because correlation can be readily included and large systems can be examined with high accuracy within reasonable computational resources. In addition, the computational cost of VMC is comparable to the cost of CDFT or other mean field methods.

The total energy of these systems can be written as a sum of the contributions arising from the kinetic energy, the Coulomb interaction with the external potential (usually considered to be parabolic) and the interaction with the magnetic field, B, if existent, which is also usually given as a parabolic function.

Quantum dots are about 100 times larger than atoms; therefore the interaction with a magnetic field is strong, which gives rise to large gaps between high- and low-spin configurations. It is observed that at low magnetic field, Hunds rule applies in the orbital filling, giving rise to high spin states. For high magnetic fields, however, all spins align obtaining what is called the maximum spin density droplet state. At even higher magnetic field values, lower states are emptied to populate higher ones. As the magnetic field increases, transitions to higher spin states are encountered.

9.2.1 Quantum dots in vacuum

Harju, *et al.* [296] have reported a series of studies that lend validity to the applicability of VMC to describe accurately QD systems. In this work they test the applicability of VMC with calculations on a very simplified version of a 2DQD, namely, two interacting electrons in the confining potential created by the quantum dot. This system has been studied extensively by other methods [296–303]. The Hamiltonian of the system is

$$H = \sum_{i=1}^{2}(-\frac{\hbar^2}{2m^*}\nabla_i^2 + \frac{\omega_c}{}L_{z,i} + \frac{m^*\Omega_0^2}{2}r_i^2) \tag{157}$$

where m^* and e are the effective electron mass and charge, respectively. In a center-of-mass representation Eq. 157 is separated into an harmonic-like equation, dependent on the center-of-mass coordinates and an equation dependent on the relative coordinates of the electrons.

$$\Psi_{CM} = exp(-\frac{R^2}{l^2}) \tag{158}$$

Equation 158 conserves the angular momentum of the electron pair so that the solution can be written as a product of radial and angular factors. The radial factor can be interpreted as the two-body Jastrow correlation factor of QMC trial functions [296].

Harju *et al.* show that VMC results obtained in this way are in excellent agreement with exact numerical results; see Fig. 2 of ref. [296]. This finding calls attention to VMC as an effective alternative method for studying large quantum dots. The authors followed with a study of a 2DQD interacting with weak magnetic fields [304]. The latter case is discussed in the following section 9.2.2.

As mentioned, a confining potential generates atom-like energy levels. Harju *et al.* performed calculations in the weak-confinement limit to show that as this limit is approached, quantum effects disappear and the system description becomes the classical solution of point charges in an external magnetic field. This state is referred to as the Wigner molecule. In this work they compare the VMC results along this transition with the most accurate energies available. As the confinement gets weaker the system spin becomes polarized before electrons attain classical positions; see Fig. 1 of ref. [305]. The transition to the Wigner molecule occurs roughly at the same density as in the 2DEG, however, it does not go through a polarization before the transition.

Calculations for systems containing more than two electrons have been carried out demonstrating good agreement between VMC and exact methods [305, 306]. Pederiva *et al.* [307, 308] also present comparisons between VMC and DMC calculations for systems containing various numbers of electrons using a model similar to that of Harju. Reference contains corrected results obtained using a trial wave function with an improved Jastrow factor. The results in each of these studies show good agreement between both methods except for the three-electron 2DQD calculation in the first paper by Pederiva *et al.* Harju *et al.* also show VMC results for the same systems treated by Pederiva *et al.* using an optimized independent-particle function a trial function with a two-body jastrow correlation function and obtaining good agreement with DMC. The total energy of the three-electron dot is presented in Table 8. Thee results confirms the importance of the trial function for accurate VMC results.

The results shown in Table 8 were obtained using an optimized wave function and agree almost completely with the ones obtained using DMC. The errors presented in VMC calculations are usually much smaller than the SDFT error [305, 307, 308].

Harju *et al.* also present VMC energies for a six-electron QD and compare results with those obtained using an exact diagonalization method; see Table 9.

As in the previous case, VMC and exact diagonalization are in good agreement. The VMC energies are lower than those of the exact diagonalization method. This difference is ascribed to the inability of the exact diagonalization approach to describe the many body wave function with a limited basis set [305]. Calculations of the radial pair distribution show that as the confinement potential weakens, the peaks that represent the positions of the two electrons narrow and grow higher, showing that the electrons become localized as confinement is reduced and approach the classical description.

With different confining procedures, it is possible to obtain quantum dots with different geometries. By applying an infinite hard-wall potential, for example, one can generate a good model for real rectangular dots. The latter case has been studied by several authors [309–313]. The simplest model for this case corresponds to the particle in a box. In this case, it is a rectangular box of dimension L by β L, where β is a deformation parameter that gives the ratio of the lengths of the two-dimensional box. In this model, electrons are considered to be noninteracting. The solution of this problem takes the form of an energy dependent on the deformation parameter and the two quantum numbers of a particle in a 2D box.

$$E_{n_x,n_y} = \frac{1}{2}(\frac{n_x^2}{\beta} + \beta n_y^2) \tag{159}$$

For a square quantum dot, the deformation parameter is unity and the energy equation described before Eq.159 predicts degenerancies for each energy state. This degenerancy leads to the introduction of magic electron numbers corresponding to occupations of closed shell configurations, which for a square dot are N = 2, 6, 8, 12, etc. As one deforms the geometry by making one side shorter than the other, the degenerancy is lifted in a manner similar to that encountered for an anisotropic harmonic potential (elliptical potential), but without the regularly spaced junctions of the eigenstates. As the deformation is increased, the system geometry approaches that of a one-dimensional system. This makes quantization in the longer direction the determining factor for energy level occupation.

In Fig. 4 of ref. [314], Rasanen *et al.* show spectra for the addition energy obtained for a rectangular dot that contains up to 16 electrons with various deformation coefficients. The addition energy is defined as the energy involved in adding an electron to an N electron dot. The agreement among SDFT and VMC is generally quite good for these calculations. For the square quantum dot the magic numbers can be easily observed as peaks in the plot. The large energy that is occasionally observed for N=4, even though it is not a magic number, is a consequence of being a half-filled shell. It is also clear that the shape of the spectrum varies strongly with deformation. At high deformation, $\beta > 2$, the trend corresponds to an effective one-dimensional system.

Discussion to this point has been on systems composed of non-interacting particles. We now turn to systems of interacting particles. Rasanen et.al. [314] have compared the energy of a system containing N non-interacting particles

with the ground state of the corresponding system of interacting electrons; see Fig. 5 of ref. [314]. For $\beta = 1$, the system has a square shape and the S = 1 ground state, which corresponds to a partially spin polarized configuration, is found at N = 4, 10, and 14, corresponding to half-filled shells. As the dot is deformed, the system changes to an S = 0 state. In general, polarization occurs at larger values of the deformation parameter for the interacting case than for the non-interacting case. This behavior suggests that for these systems, the effective deformation of the rectangle is smaller than the external potential, which can be explained by the fact that in a rectangular box, the electron density is increased toward the shorter side. In the case of elliptical dots, the observed behavior is the opposite since the deformation of the system causes the electronic density to concentrate at the center of the ellipse.

Comparisons of the energetic results of this type of systems with experiment is not always straightforward. There may be irregularities in the experimental dot that lead to unexpected behavior. Nevertheless, the calculated spectra obtained in this way are quite similar to experimental spectra.

9.2.2 Quantum dots in magnetic fields

The interaction of QD's with magnetic fields is important for ascertaining the possible uses of these systems in nano-electronics. Many papers have been published on quantum dots in magntic fields. The most desirable method is one that gives exact results, which makes exact diagonalization compelling. This method is limited, however, by its scaling with system size and finite-basis errors. QMC has proved to offer advantages for this problem because its high accuracy has been found comparable to the exact method and it scales better with system size.

The VMC method can be used to address these systems for a wide range of magnetic strengths. Simple wave functions can be constructed that not only give accuracies comparable to exact methods, but are also easier to interpret.

As mentioned above, the interaction of a QD with a magnetic field is usually strong. The effect of an external magnetic field and the electron-electron interaction energy are usually of comparable strength in a typical QD [315]. As the field strength increases, the system typically undergoes several transitions to higher spin states. The state where all electrons that compose the dot align with the magnetic field is called a maximum-density droplet; it is a state of full spin polarization and occupation of the lowest Landau levels (LLL).

In Fig. 5 of ref. [315], Saarikoski *et al.* compare the energy of a six-electron dot using VMC, SDFT and CSDFT as a function of magnetic field strength. The transitions between different spin states are evident. However, it can be seen that the DFT methods and QMC are not always in agreement. In SDFT and CSDFT, the transition from the maximum-density droplet state to the $L_z = 21$, $S_z = 3$ state occurs at a magnetic field of 10T, while for VMC the transition occurs at 9T. It can also be seen that VMC and SDFT give a transition at roughly 4T to a state with $L_z = 7$, $S_z = 1$ that is absent in the CSDFT calculation. These differences reflect the difficulty in constructing an

exchange-correlation functional that remains valid over the full range of field strength. According to Saarikoski *et al.* [315], beyond the maximum-density droplet state, it is hard to converge the CSDFT method.

For an arbitrary system, the confining potential landscape can play an important role in determining the properties of a quantum dot. The presence of impurities can alter the behavior of the given system. Analysis done to ascertain the effect of random disorder, chaotic dots and impurities can be found in refs. [316–326].

In Fig. 2 of ref. [327], Guclu *et al.* plot the ground state energy dependence with respect to magnetic field strength for ordered (pure) and disordered dots containing 4, 6, 9 and 10 electrons. For these calculations, the dot was simulated as a standard QD in a symmetric parabolic potential. The disordered dots were simulated by introducing impurities represented by randomly positioned Gaussian-like potentials using unrestricted HF (UHF) and DMC. The latter method gives a much better result than UHF. In addition, calculations for N=4 (not shown) were performed by exact diagonalization [327] show excellent agreement with DMC results. The vertical lines shown on the graphs represent the transition to different spin states obtained by DMC in the standard dot. The total spin in the z direction is indicated in each graph by an integer placed between the spin transition lines corresponding to the particular state. For large magnetic field, close to the maximum density droplet region, UHF results become much better than at weak fields where they do not give accurate energies or correct spin states. This can be understood by the fact that the maximum density droplet state has very small correlation energy and can be described reasonably well by a single Slater determinant. The effect of disorder in these systems is evident in all plots. The energy trends shown by the disordered dots are not as clear as those for pure dots, i.e., the transitions between different spin states nearly disappear. This effect can make the observation of these transitions hard to observe experimentally. The value of the magnetic field at which a transition takes place is also changed and in some cases even new transitions that did not take place in the pure dot appear.

In Fig. 4 of ref. [327] the dependence of the addition energy on the number of electrons is calculated for different disordered states and compared to the corresponding pure quantum dot shown in the bottom line. It is evident from these results that, for closed shells, *i.e.* for magic n umbers N = 2, 6, 12, Hunds rule applies even for highly disordered dots. This behavior is indicated in the figure mentioned as a persistent peak at those positions that have S = 0, which indicates the existence of the closed shell. For open shells, the rule is not always obeyed. For systems with a small number of electrons, *i.e.*, for systems that have only low shells populated, they seem to be less affected by disorder than systems with higher shells populated. These results are significant because they provide important information for experimentalists because these effects have implications for interpreting spectroscopy data, since electronic transitions are less clear and the magnetic field needed for promoting those transitions does not correspond to the one predicted in non-disordered dots.

9.2.3 Interacting quantum dots

The system of two quantum dots containing one electron each, similar to the H_2 molecule, has been proposed as a basic elementary gate in quantum computers. This idea can be generalized to different arrangements of QD, *e.g.* linear chains, etc. [328–330] There are a variety of materials and tunable parameters including the confining potential that can be used to generate different behaviors. These systems are referred to as quantum dot molecules.

The most studied geometry is one with two QDs laterally coupled at the interface of GaAs/AlGaAs plates, each QD containing one electron. The dots must be sufficiently close that electron tunneling between these dots can occur.

Figure. 7 of ref. [315] shows the dependence of the energy of the coupled dot system with S = 0 and S = 1 with respect to dot separation. One sees that as the distance increases, the energy is reduced. The DFT values are in good quantitative agreement with VMC results at small inter-dot distances for the singlet-triplet gap. At large distances, the agreement is not as good, SDFT leads to a symmetry breaking. There is only a qualitative agreement between SDFT and VMC in this regime.

Because of their potential use in nano electronics, the effect of a magnetic field on interacting QDs holds considerable interest. Results for the dependence of the energy difference of the triplet and singlet states of the dot with respect to magnetic field strength calculated with CSDFT are presented in Fig. 8 of ref. [315] for two dots separated by 2.73 bohr. As shown, CSDFT predicts a singlet to triplet transition at 1.5T. At low magnetic fields, the CSDFT results are in accord with QMC, but as the magnetic field is increased, the CSDFT and DMC results begin to deviate from each other. The singlet to triplet transition holds high interest because it opens the possibility of using QDs as qubits in the construction of quantum computers. The CSDFT and QMC results are in reasonable agreement at the value of the magnetic field needed for such a singlet-triplet transition. In general, CSDFT is in good agreement with QMC in regimes of finite magnetic fields; however, the results depend on the choice of exchange-correlation functional.

9.2.4 Quantum dot excitonic complexes

Another aspect of QDs that has generated high interest recently, both experimentally and theoretically, is excitonic complexes. These systems have considerable importance in semiconductor physics and spectroscopy. An exciton corresponds to the bound state formed by an electron with an effective mass m_e and a positively charged hole of mass m_h left behind when the electron is excited to a higher level. The ratio of m_e to m_h characterizes the system. The limit where this ratio goes to zero is designated the hydrogenic limit, while the limit where the ratio goes to one is labeled the positronic limit. The ratio for a given exciton depends on the material and other factors. Also, bound states can be formed consisting of n_e electrons and n_h holes, similar to the case of an electron-proton system in which one has bound systems such as H, H_2, H_2^+,

H^-, H_3^+. An important feature of excitonic complexes is that the emission in QD lasers arises from exciton recombination. Single-dot spectroscopy has been used to resolve the energy of excitons. One of the most intriguing properties is binding energy enhancement that arises from excitons in a two-dimensional QD compared to a three-dimensional system, are much larger than bulk values. This behavior is due to confinement of the hole and the electron in the same quantum well.

There are a number of different energy values that are important to calculate for an excitonic complex. For example, one value of interest for a single exciton system is the exciton transition energy, *i.e.*,

$$E_X = E_{1,1}(e_0^1, h_0^1) - E_{0,0} \tag{160}$$

Here the notation for energy levels is E_{e_0,h_0}. This is defined as the difference between the energy of a dot having an electron in level e_0 and a hole in level h_0, and a hole in its ground state. Typical values for this quantity for group III-V or II-IV dots are 1-3 eV.

Another important property is the exciton binding energy ΔX, which is defined as the difference between the total energy of a system composed of two dots infinitely separated, one having an electron in level e_0 and the other having a hole in level h_0, and the energy of a quantum dot containing an exciton,

$$\Delta X = E_{1,0} + E_{0,1} - E_{1,1} - E_{0,0} \tag{161}$$

Typical energy values are in the range 10-200 meV. Similar properties can be defined for systems containing more than one exciton beginning with biexcitons. For example, the biexciton binding energy ΔXX is defined as the difference between twice the exciton energy and the biexciton energy $E_{2,2}$, *i.e.*, the difference in energy of having two excitons in a single dot compared to the energy of having two dots each with a single exciton. This energy is typically of the order of 1-6meV. The presence of biexcitons has been observed experimentally in single dot spectroscopy as a red shift on exciton luminescence. For multiple excitons there are two other quantities of interest. One is the charging energy, which is defined as the minimum energy required to add an extra electron-hole pair to an N-1 exciton system. The other is the addition energy, which in this context is defined as the difference in energy between successive multi-electron charging energies. QMC has been proved to be useful in calculating accurate cuantities, such as the ones mentioned before, for these systems.

In Table 1 of ref. [331], results are presented for a simplified model of a spherical dot. The column labelled as magnitude corresponds to the QMC value. As can be seen QMC is able to give excellent results for a variety of properties involving excitons.

The dependence of these quantities on dot size has also been investigated. Figure 2 of ref. [331] shows the behavior of the energy of an electron, a hole and the exciton and biexciton binding energies as a function of the size of the dot. The geometry used for the system is the one shown in the panel on the binding energy plot. It can see that as the radius of the dot increases, the

material has a drastic change in its behavior around 2nm; see the first two plots from the figure mentioned previously. For these systems, Table 2 and the last two plots of Fig. 2 of ref. [331] show that the correlation energy is of order 1% of the total energy. There are other quantities, however, including like exciton addition energies which have large correlation energy contributions. While for exciton binding energies CI recovers nearly 98% of the QMC energy, for biexcitonic binding energies CI calculations are relatively poor compared to QMC, recovering in some cases only 65% of the energy.

Charged excitons are also an important field of study. The description of these systems is important because luminescence lines due to charged excitons have sometimes been wrongly identified as biexcitons. For more information on these systems, please refer to the following references [332–334].

List of symbols

We denote:

Z_j Atomic nuclear charge,

$\mathcal{R}$ Set of coordinates of clamped particles (under the Born–Oppenheimer approximation): $\mathcal{R} \equiv \{\mathcal{R}_1, \ldots, \mathcal{R}_m\}$,

r_i Electronic coordinate in a Cartesian frame $r_i = \{x_i \ldots x_d\}$, where the number of dimensions, d is 3 for most chemical applications,

$\mathbf{R}$ Set of coordinates of all n particles treated by QMC, $\mathbf{R} \equiv \{r_1, \ldots, r_n\}$,

σ_i Spin coordinate for an electron, $\sigma_\uparrow$ is spin-up and $\sigma_\downarrow$ for spin-down particles,

Σ Set of spin coordinates for particles, $\Sigma \equiv \{\sigma_1, \ldots, \sigma_N\}$,

Ψ_0 Exact ground state wave function,

Ψ_i ith exact wave function,

$\Psi_{T,i}$ Approximate trial wave function for state i,

ϕ_k Single particle molecular orbital (MO),

$D(\phi_k)$ Slater determinant of k MOs: $D = \frac{1}{\sqrt{N!}} \det |\phi_1, \ldots, \phi_N|$,

$D^\uparrow(\phi_k^\uparrow), D^\downarrow(\phi_k^\downarrow)$, Spin factored Slater determinants for spin up ($\uparrow$) and spin down ($\downarrow$) electrons,

$\widehat{H}$ Hamiltonian operator: $\widehat{H} \equiv \widehat{T} + \widehat{V}$,

$\widehat{T}$ Kinetic energy operator $-\frac{1}{2}\nabla_R^2 \equiv -\frac{1}{2}\sum_i \nabla_i^2$,

$\widehat{V}$ Potential energy operator, for atomic and molecular systems: $\widehat{V} = -\sum_{ij} \frac{Z_j}{r_{ij}} + \sum_{i<j} \frac{1}{r_{ij}} + \sum_{i<j} \frac{Z_i Z_j}{r_{ij}}$,

τ Imaginary time, $\tau = it$,

$\rho(\mathbf{r})$ Electronic density,

E_L Local Energy, $\widehat{H}\Psi_T(\mathbf{R})/\Psi_T(\mathbf{R})$,

$\mathcal{U}_{[a,b]}$ Uniform random variate in the interval $[a, b]$,

$\mathcal{G}_\sigma$ Gaussian random variate of variance σ,

$\sigma_{\widehat{O}}$ Monte Carlo variance for observable $\widehat{O}$,

$\mathcal{W}$ Ensemble of random walkers, $\mathcal{W} \equiv \{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R_n}\}$,

$\mathcal{P}(\mathbf{R})$ Monte Carlo probability density function,

$P(\mathbf{R} \rightarrow \mathbf{R}')$ Monte Carlo transition probability,

$B(\mathbf{R}, \mathbf{R}')$ Branching factor for Population Monte Carlo algorithms,

$G(\mathbf{R}, \mathbf{R}'; \tau - \tau')$ Time dependent Green's function,

$G_{\mathrm{ST}}(\mathbf{R}, \mathbf{R}'; \delta\tau)$ Time dependent short time Green's function for the Schrödinger equation,

$\mathbf{F_q}$ Quantum force, $\mathbf{F_q} \equiv \nabla \ln |\Psi_T(\mathbf{R})^2|$,

$\mathcal{D}$ Denominator in fermion Monte Carlo, $\mathcal{D} \equiv \sum_{i=1}^{Nw} \left[\frac{\psi_T(R_i^+)}{\Psi_G^+(R_i^+)} \frac{\psi_T(R_i^-)}{\Psi_G^-(R_i^-)}\right]$,

M_I Ionic mass,

μ_α Fictitious parameter mass.

Acknowledgments

A.A.-G. was a holder of a Gates Millenium Fellowship during the preparation of this chapter. W.A.L. was supported by the Director, Office of Science, Office of Basic Energy Sciences, Chemical Sciences Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098, and by the National Science Foundation through the CREST Program (HRD-9805465).

References

[1] K. E. SCHMIDT, Variational and Green's Function Monte Carlo Calculations of Few Body Systems, in *Conference on Models and Methods in Few Body Physics*, Lisbon, 1986.

[2] W. A. LESTER, JR. and B. L. HAMMOND, *Annu. Rev. Phys. Chem.* **41**, 283 (1990).

[3] D. M. CEPERLEY and L. MITAS, Quantum Monte Carlo Methods in Chemistry, in *New Methods in Computational Quantum Mechanics*, edited by I. PRIGOGINE and S. A. RICE, volume XCIII of *Adv. Chem. Phys.*, John Wiley and Sons, 1996.

[4] P. H. ACIOLI, *J. of Mol. Struct. (Theochem)* **394**, 75 (1997).

[5] D. BRESSANINI and P. J. REYNOLDS, *Adv. Chem. Phys.* **105**, 37 (1998).

[6] L. MITAS, Diffusion Monte Carlo, in *Quantum Monte Carlo Methods in Physics and Chemistry*, edited by M. P. NIGHTINGALE and C. J. UMRIGAR, volume 525, p. 247, Kluwer Academic Publishers, 1998.

[7] J. B. ANDERSON, Quantum Monte Carlo: Atoms, Molecules, Clusters, Liquids and Solids, in *Reviews in Computational Chemistry*, edited by K. B. LIPKOWITZ and D. B. BOYD, volume 13, p. 133, John Wiley and Sons, New York, 1999.

[8] A. LÜCHOW and J. B. ANDERSON, *Ann. Rev. Phys. Chem.* **51**, 501 (2000).

[9] M. FOULKES, L. MITAS, R. NEEDS, and G. RAJAGOPAL, *Rev. Mod. Phys.* **73**, 33 (2001).

[10] A. ASPURU-GUZIK and W. A. LESTER, JR., Quantum Monte Carlo methods for the Solution of the Schrödinger equation for molecular systems, in *Computational Chemistry*, edited by C. LE BRIS, volume X of *Handbook of Numerical Analysis*, p. 485, Elsevier, Amsterdam, The Netherlands, first edition, 2003.

[11] B. L. HAMMOND, W. A. LESTER, JR., and P. J. REYNOLDS, *Monte Carlo Methods in Ab Initio Quantum Chemistry*, World Scientific, Singapore, 1994.

[12] S. E. KOONIN and D. C. MEREDITH, *Computational physics, FORTRAN version*, Addison Wesley, third edition, 1995.

[13] H. GOULD and J. TOBOCHNIK, *An Introduction to Computer simulation Methods: Applications to Physical Systems*, Addison Wesley, second edition, 1996.

[14] J. M. THIJSSEN, *Computational Physics*, Cambridge Univ. Press, first edition, 1999.

[15] D. M. CHARUTZ and D. NEUHAUSER, *J. Chem. Phys.* **102**, 4495 (1995).

[16] N. ROM., D. M. CHARUTZ, and D. NEUHAUSER, *Chem. Phys. Lett.* **270**, 382 (1997).

[17] R. BAER, M. HEAD-GORDON, and D. NEUHAUSER, *J. Chem. Phys.* **109**, 6219 (1998).

[18] R. BAER and D. NEUHAUSER, *J. Chem. Phys.* **112**, 1679 (2000).

[19] D. M. CEPERLEY, *Rev.Mod. Phys.* **67**, 279 (1995).

[20] A. SARSA, K. E. SCHMIDT, and W. R. MAGRO, *J. Chem. Phys.* **113**, 1366 (2000).

[21] D. M. BENOIT and D. C. CLARY, *J. Chem. Phys.* **113**, 5193 (2000).

[22] Y. KWON, P. HUANG, M. V. PATEL, D. BLUME, and K. B. WHALEY, *J. Chem. Phys.* **113**, 6469 (2000).

[23] D. C. CLARY, *J. Chem. Phys.,* **114**, 9725 (2001).

[24] A. VIEL and K. B. WHALEY, *J. Chem. Phys.* **115**, 10186 (2001).

[25] D. M. CEPERLEY and B. J. ALDER, *Phys. Rev. Lett.* **45**, 566 (1980).

[26] R. B. WIRINGA, S. C. PIEPER, J. CARLSON, and V. R. PANDHARIPANDE, *Phys. Rev. C* **62**, 14001 (2000).

[27] S. C. PIEPER and R. B. WIRINGA, *Annu. Rev. Nucl. Part. Sci* **51**, 53 (2001).

[28] N. METROPOLIS and S. ULAM, *J. Amer. Statist. Assoc.* **44**, 335 (1949).

[29] ALBERT EINSTEIN, *Investigations in the theory f Brownian Motion*, Metheun & Co. Ltd, 1926, English translation of Einstein's original paper.

[30] R. COURANT, K. O. FRIEDRICHS, and H. LEWY, *Math. Ann.* **100**, 32 (1928).

[31] M. HEAD-GORDON, *J. Phys. Chem.* **100**, 13213 (1996).

[32] K. RAGHAVACHARI and J. B. ANDERSON, *J. Phys. Chem.* **100**, 12960 (1996).

[33] E. CANCÈS, M. DEFRANCESCHI, W. KUTZELNIGG, C. LE BRIS, and Y. MADAY, The models of quantum chemistry and their discretizations, in *Computational Chemistry*, edited by C. LE BRIS, volume X of *Handbook of Numerical Analysis*, chapter 1, Elsevier, Amsterdam, The Netherlands, first. edition, 2003.

[34] P.-O. LÖWDIN, *Adv. Chem. Phys.* **2**, 207 (1959).

[35] G. SENATORE and N. H. MARCH, *Rev. Mod. Phys.* **66**, 445 (1994).

[36] D. WATSON, M. DUNN, T. C. GERMANN, D. R. HERSCHBACH, and D. Z. GOODSON, Dimensional Expansions for Atomic Systems, in *New Methods in Quantum Theory*, edited by C. A. TSIPIS, V. S. POPOV, D. R. HERSCHBACH, and J. AVERY, p. 83, Kluwer Academic, Dortecht, 1996.

[37] D. A. MAZZIOTTI, *Phys. Rev. A* **57**, 4219 (1998).

[38] M. D. DONSKER AND M. KAC, *J. Res. Nat. Bur. Standards* **44**, 551 (1950).

[39] M. H. KALOS, *Phys. Rev.* **68** (1962).

[40] H. HELLMAN, *Einführung in die Quanten Theorie*, Deuticke, 1937.

[41] R. P. FEYNMAN, *Phys. Rev.* **56**, 340 (1939).

[42] M. WEISSBLUTH, *Atoms and molecules*, Academic Press, New York, 1978, See pages 570-572.

[43] A. A. FROST, R. E. KELLOG, and E. A. CURTIS, *Rev. Mod. Phys.* **32**, 313 (1960).

[44] T. KATO, *Comm. in Pure. and Appl. Math.* **10**, 151 (1957).

[45] C. R. MEYERS, C. J. UMRIGAR, J. P. SETHNA, and J. D. MORGAN, *Phys. Rev. A* **44**, 5537 (1991).

[46] M. M. MORELL, R. G. PARR, and M. LEVY, *J. Chem. Phys.* **62**, 549 (1975).

[47] E. R. DAVIDSON, *Reduced Density Matrices in Quantum Chemistry*, Acad. Press, New York, 1976.

[48] H. H. JAMES and A. S. COOLIDGE, *Phys. Rev.* **51**, 860 (1937).

[49] K. MCDOWELL, *Int. J. Quant. Chem: Quant. Chem. Symp.* **15**, 177 (1981).

[50] W. F. BAUER, *J. Soc. Indst. Appl. Math.* **6**, 438 (1958), A very well described introduction to MC methods by an applied mathematician. Cites the Courant and Levy 1928 paper.

[51] J. M. HAMMERSLEY and D. C. HANDSCOMB, *Monte Carlo Methods*, Methuen, London, 1964.

[52] JOHN H. HALTON, *SIAM Review* **12**, 1 (1970).

[53] M. H. KALOS and P. A. WHITLOCK, *Monte Carlo Methods Volume 1: Basics*, Wiley, N.Y., 1986.

[54] W. W. Wood and J. J. Erpenbeck, *Ann. Rev. Phys. Chem.* **27**, 319 (1976).

[55] G. S. Fishman, *Monte Carlo: Concepts, Algorithms and Applications*, Springer-Verlag, 1a ed. edition, 1996.

[56] I. M. Sobol, *A Primer for the Monte Carlo Method*, CRC Press, 1994.

[57] I. Manno, *Introduction to the Monte-Carlo Method*, Akademiai Kiado, Budapest, 1999.

[58] A. Doucet, N. de Freitas, N. Gordon, and A. Smith, editors, *Sequential Monte Carlo methods in practice*, Springer Verlag, New York, 2001.

[59] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer Verlag, New York, 2001.

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

[61] M. Caffarel and P. Claverie, *J. Chem. Phys.* **88**, 1088 (1988).

[62] S.-Y. Huang, Z. Sun, and W. A. Lester, Jr., *J. Chem. Phys.* **92**, 597 (1990).

[63] S. A. Alexander and R. L. Coldwell, *Int. J. Quant. Chem.* **63**, 1001 (1997).

[64] L. Bertini, D. Bressanini, M. Mella, and G. Morosi, *Int. J. of Quant. Chem.* **74**, 23 (1999).

[65] S. Manten and A. Lüchow, *J. Chem. Phys.* **115**, 5362 (2001).

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

[67] A. Bijl, *Physica* **7**, 869 (1940).

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

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

[70] C. Huang, C. J. Umrigar, and M. P. Nightingale, *J. Chem. Phys.* **107**, 3007 (1997).

[71] R. P. Feynman and M. Cohen, *Phys. Rev.* **102**, 1189 (1956).

[72] K. E. Schmidt and J. W. Moskowitz, *J. Chem. Phys.* **93** (1990).

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

[74] Y. Kwon, D. M. Ceperley, and R. M. Martin, *Phys. Rev. B.* **58**, 6800 (1998).

[75] K. Schmidt, M. H. Kalos, M.l A. Lee, and G. V. Chester, *Phys. Rev. Lett.* **45**, 573 (1980).

[76] J. Casulleras and J. Boronat, *Phys. Rev. Lett.* **84**, 3121 (2000).

[77] V. R. Pandharipande and N. Itoh, *Phys. Rev. A* **8**, 2564 (1973).

[78] V. R. Pandharipande, S. C. Pieper, and R. B. Wiringa, *Phys. Rev. B* **34**, 4571 (1986).

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

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

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

[82] M. Rao and B. J. Berne, *J. Chem. Phys.* **71**, 129 (1979).

[83] C. Pangali, M. Rao, and B. J. Berne, *Mol. Phys.* **37**, 1773 (1979).

[84] J. B. Anderson, *J. Chem. Phys.* **73**, 3897 (1980).

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

[86] M. L. Stedman, W. M. C. Fowlkes, and M. Nekovee, *J. Chem. Phys.* **109**, 2630 (1998).

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

[88] Z. Sun, M. M. Soto, and W. A. Lester, Jr., *J. Chem. Phys.* **100**, 1278 (1994).

[89] D. Bressanini and P. J. Reynolds, *J. Chem. Phys.* **111**, 6180 (1999).

[90] S. M. Rothstein and J. Vbrik, *J. Comp. Phys.* **74**, 127 (1988).

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

[92] Z. Sun, R. N. Barnett, and W. A. Lester, Jr., *Chem. Phys. Lett.* **195**, 365 (1992).

[93] C. Filippi and C. J. Umrigar, *J. Chem. Phys.* **105**, 213 (1996).

[94] H.-J. Flad, M. Caffarel, and A. Savin, Quantum Monte Carlo Calculations with Multi-Reference Trial Wave Functions, in *Recent Advances in Quantum Monte Carlo Methods*, edited by W. A. Lester, Jr., chapter 5, pp. 73–98, Wold Scientific, 1997.

[95] R. N. Barnett, Z. Sun, and W. A. Lester, Jr., *Chem. Phys. Lett.* **273**, 321 (1997).

[96] R. N. Barnett, Z. Sun, and W. A. Lester, Jr., *J. Chem. Phys.* **114**, 2013 (2001), J. Chem. Phys.

[97] H. Bueckert, S. M. Rothstein, and J. Vrbik, *Can. J. Chem.* **70**, 366 (1992).

[98] A. Lüchow and J. B. Anderson, *J. Chem. Phys.* **105**, 7573 (1996).

[99] H. Huang and Z. Cao, *J. Chem. Phys.* **104**, 1 (1996).

[100] H. Huang, Q. Xie, Z. Cao, Z. Li, Z. Yue, and L. Ming, *J. Chem. Phys.* **110**, 3703 (1999).

[101] X. Lin, H. Zhang, and A. M. Rappe, *J. Chem. Phy.s* **112**, 2650 (2000).

[102] H. Conroy, *J. Chem. Phys.* **41**, 1331 (1964).

[103] Z. Sun, S.-Y. Huang, R. N. Barnett, and W. A. Lester, Jr., *J. Chem. Phys.* **93**, 5 (1990).

[104] M. P. Nightingale and C. J. Umrigar, Monte Carlo Optimization of Trial Wave Functions in Quantum Mechanics and Statistical Mechanics, in *Recent Advances in Quantum Monte Carlo Methods*, edited by W. A. Lester, Jr., chapter 12, pp. 201–227, World Scientific, 1997.

[105] M. Snajdr, J. R. Dwyer, and S. M. Rothstein, *J. Chem. Phys.* **111**, 9971 (1999).

[106] C. Filippi and S. Fahy, *J. Chem. Phys.* **112**, 3532 (2000).

[107] F. Schautz and S. Fahy, *J. Chem. Phys.* **116**, 3533 (2002).

[108] S. Sorella, *Phys. Rev. B* **6402** (2001).

[109] M. Honik, M. Snajdr, and S. M. Rothstein, *J. Chem. Phys.* **113**, 3496 (2000).

[110] J. H. Hetherington, *Phys. Rev. A.* **30**, 2713 (1984).

[111] S. A. Chin, *Phys. Rev. A* **42**, 6991 (1990).

[112] N. J. Cerf and O. C. Martin, *Int. J. of Mod. Phys.* **6**, 693 (1995).

[113] Y. Iba, *Trans. Japanese. Soc. of Art. Int.* , 279 (2000).

[114] R. Assaraf, M. Caffarel, and A. Khelif, *Phys. Rev. E* **61**, 4566 (2000).

[115] F. B. Jr. and R. Fuller, *Mathematics of Clasical and Quantum Physics*, Dover, 1992.

[116] D. M. Ceperley and M. H. Kalos, Quantum Many-Body problems, in *Monte Carlo Methods in Statistical Physics*, edited by K. Binder, Springer-Verlag, New York, 2nd. edition, 1986.

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

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

[119] S. A. Chin., *Phys. Lett. A.* **226**, 344 (1997).

[120] H. A. Forbert and S. A. Chin, *Int. J. Mod. Phys. B* **15**, 1752 (2001).

[121] S. Sorella, *Phys. Rev. Lett.* **80**, 4558 (1998).

[122] M. C. Buonaura and S. Sorella, *Phys. Rev. B* **61**, 2559 (1998).

[123] S. Sorella and L. Capriotti, *Phys. Rev. B* **61**, 2599 (2000).

[124] M. F. DePasquale, S. M. Rothstein, and J. Vrbik, *J. Chem. Phys.* **89**, 3629 (1988).

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

[126] D. M. Ceperley, *J. Comp. Phys.* **51**, 404 (1983).

[127] D. W. Skinner, J. W. Moskowitz, M. A. Lee, K. E. Schmidt, and P. A. Whitlock, *J. Chem. Phys.* **83**, 4668 (1985).

[128] K. E. Schmidt and J. W. Moskowitz, *J. Stat. Phys.* **43**, 1027 (1986).

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

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

[131] S. Zhang and M. H. Kalos, *Phys. Rev. Lett.* **67**, 3074 (1991).

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

[133] Z. Liu, S. Zhang, and M. H. Kalos, *Phys. Rev. E.* **50**, 3220 (1994).

[134] M. H. Kalos and K. E. Schmidt, *J. Stat. Phys.* **89**, 425 (1997).

[135] M. H. Kalos and F. Pederiva, Fermion Monte Carlo, in *Quantum Monte Carlo Methods in Physics and Chemistry*, edited by M. P. Nightingale and C. J. Umrigar, volume 525, pp. 263–286, Kluwer Academic Publishers, 1998.

[136] F. Pederiva and M. H. Kalos, *Comp. Phys. Comm.* **122**, 445 (1999).

[137] M. H. Kalos and F. Pederiva, *Phys. Rev. Lett.* **85**, 3547 (2000).

[138] M. H. Kalos and F. Pederiva, *Physica A* **279**, 236 (2000).

[139] A. Aspuru-Guzik, M. Kalos, and W. A. Lester, Jr., Fermion Monte Carlo Study of the Beryllium Atom, in *The Monte Carlo Method in the Physical Sciences*, edited by J. Gubernatis, volume 690 of *AIP Conf. Proc.*, p. 371, 2003.

[140] A. Bhattacharya and J. B. Anderson, *Phys. Rev. A* **49**, 2441 (1994).

[141] A. Bhattacharya and J. B. Anderson, *J. Chem. Phys.* **100**, 8999 (1994).

[142] J. B. Anderson, *J. Chem. Phys.* **115**, 4546 (2001).

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

[144] R. Assaraf and M. Caffarel, *J. Chem. Phys.* **113**, 4028 (2000).

[145] M. Caffarel, P. Claverie, C. Mijoule, J. Andzelm, and D. R. Salahub, *J. Chem. Phys.* **90**, 990 (1989).

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

[147] W. A. Glauser, W. R. Brown, W. A. Lester, Jr., D. Bressanini, B. L. Hammond, and M. L. Kosykowski, *J. Chem. Phys.* **97**, 9200 (1992).

[148] M. Caffarel, X. Krokidis, and C. Mijoule, *Europhys. Lett.* **7**, 581 (1992).

[149] D. Bressanini, D. M. Ceperley, and P. J. Reynolds, What do we know about wave function nodes?, in *Recent Advances in Quantum Monte Carlo Methods, Part II*, edited by S. M. Rothstein, W. A. Lester, Jr., and S. Tanaka, Singapore, 2002, World Scientific.

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

[151] B. L. Hammond, P. J. Reynolds, and W. A. Lester, Jr., *J. Chem. Phys.* **87**, 1130 (1987).

[152] V. N. Staroverov, P. Langfelder, and S.M. Rothstein, *J. Chem. Phys.* **108**, 2873 (1998).

[153] T. Yoshida and K. Iguchi, *J. Chem. Phys.* **88**, 1032 (1988).

[154] J. Carlson, J. Moskowitz, and K. Schmidt, *J. Chem. Phys.* **90**, 1003 (1989).

[155] G. Bachelet, D. Ceperley, and M. Chiocchetti, *Phys. Rev. Lett.* **62**, 2088 (1989).

[156] L. Szasz, *Pseudopotential Theory of Atoms and Molecules*, Wiley, New York, 1985.

[157] M. Krauss and W. J. Stevens., *Annu. Rev. Phys. Chem.* , 357 (1984).

[158] P. A. Christiansen, W. C. Ermler, and K. S. Pitzer, *Annu. Rev. Phys. Chem.* **35**, 357 (1984).

[159] K. Balasubramanian and K.S. Pitzer, *Adv. Chem. Phys.* **77**, 287 (1987).

[160] M. Dolg, Effective Core Potentials, in *Modern Methods and Algorithms of Quantum Chemistry*, edited by J. Grotendorst, volume 1, pp. 479–508, John von Neumann Institute for Computing, 2000.

[161] C. W. Greeff and W. A. Lester, Jr., *J. Chem. Phys.* **109**, 1607 (1998).

[162] I. Ovcharenko, A. Aspuru-Guzik, and W. A. Lester, Jr., *J. Chem. Phys.* **114**, 7790 (2001).

[163] H.-J. Flad, F. Schautz, Y. Wang, and M. Dolg, Quantum Monte Carlo Study of Mercury Clusters, in *Recent Advances in Quantum Monte Carlo Methods, Part II*, edited by S. M. Rothstein, W. A. Lester, Jr., and S. Tanaka, Singapore, 2002, World Scientific.

[164] S. A. Alexander, R. L. Coldwell, G. Aissing, and A. J. Thakkar, *Int. J. Q. Chem.* **26**, 213 (1992).

[165] K. S. Liu, M. H. Kalos, and G. V. Chester, *Phsy. Rev. A* **10**, 303 (1974).

[166] R. N. Barnett, P. J. Reynolds, and W. A. Lester, Jr., *J. Comp. Phys.* **96**, 258 (1991).

[167] P. J. Reynolds, R. N. Barnett, B. L. Hammond, R. M. Grimes, and W. A. Lester, Jr., *Int. J. of Quant. Chem.* **29**, 589 (1986).

[168] A. L. L. East, S. M. Rothstein, and J. Vrbik, *J. Chem. Phys.* **89**, 4880 (1988).

[169] P. Langfelder, S.M. Rothstein, and J. Vrbik, *J. Chem. Phys.* **107**, 8525 (97).

[170] J. Boronat and J. Casulleras, *Comp. Phys. Comm.* **122**, 466 (1999).

[171] A. Sarsa, J. Boronat, and J. Casulleras, *J. Chem. Phys.* **116**, 5956 (2002).

[172] F. Schautz and H.-J. Flad, *J. Chem. Phys.* **110**, 11700 (1999).

[173] R. N. Barnett, P. J. Reynolds, and W. A. Lester, Jr., *J. Chem. Phys.* **96**, 2141 (1992).

[174] R. N. Barnett, P. J. Reynolds, and W. A. Lester, Jr., *Int. J. of Quant. Chem.* **42**, 837 (1992).

[175] S. W. Zhang and M. H. Kalos, *J. Stat. Phys.* **70**, 515 (1993).

[176] D. M. Ceperley and B. Bernu, *J. Chem. Phys.* **89**, 6316 (1988).

[177] M. Born and J. R. Oppenheimer, *Ann. Phys.* **84**, 457 (1927).

[178] C. A. Traynor, J. B. Anderson, and B. M. Boghosian, *J. Chem. Phys.* **94**, 3657 (1994).

[179] R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules*, Oxford University Press, 1989.

[180] J. Vrbik, D. A. Lagare, and S. M. Rothstein, *J. Chem. Phys.* **92**, 1221 (1990).

[181] J. Vrbik and S. M. Rothstein, *J. Chem. Phys.* **96**, 2071 (1992).

[182] P. Belohorec, S.M. Rothstein, and J. Vrbik, *J. Chem. Phys.* **98**, 6401 (1993).

[183] C. Filippi and C. J. Umrigar, Interatomic Forces and Correlated Sampling in Quantum Monte Carlo, in *Recent Advances in Quantum Monte Carlo Methods, Part II*, edited by S. M. Rothstein, W. A. Lester, Jr., and S. Tanaka, Singapore, 2002, World Scientific.

[184] Z. Sun, W. A. Lester, Jr., and B. L. Hammond, *J. Chem. Phys.* **97**, 7585 (1992).

[185] C. J. Umrigar, *Int. J. of Quant. Chem.* **23**, 217 (1989).

[186] K. C. Huang, R. J. Needs, and G. Rajagopal, *J. Chem. Phys.* **112**, 4419 (2000).

[187] F. Schautz and H.-J. Flad, *J. Chem. Phys.* **112**, 4421 (2000).

[188] S. Lu, *J. Chem. Phys.* **114**, 3898 (2001).

[189] S. Lu, *J. Chem. Phys.* **118**, 6152 (2003).

[190] M. Frenklach, Solution mapping approach to modeling combustion, in *Computational Fluid and Solid Mechanics*, edited by K. J. Bathe, pp. 1177–1179, New York, Elsevier, 2001.

[191] M. Frenklach, H. Wang, and M. J. Rabinowitz, *Prog. Energy Combust. Sci.* **18**, 47 (1992).

[192] M. Frenklach, Modeling, in *Combustion Chemistry*, edited by J. Gardiner, W. C., pp. 423–453, Springer-Verlag, New York, 1984.

[193] R. Myers and D. Montgomery, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, Wiley, New York, 2002.

[194] G. E. P. Box and N. Draper, *Empirical Model-Building and Response Surfaces*, Wiley, New York, 1987.

[195] C. A. Schuetz, A. C. Kollias, M. Frenklach, and W. A. Lester, Jr., *J. Chem. Phys.* **119**, 9386 (2003).

[196] R. Car and M. Parinello, *Phys. Rev. Lett.* **55**, 2471 (1985).

[197] S. Tanaka, *J. Chem. Phys.* **100**, 7416 (1994).

[198] S. Tanaka, Ab Initio apporach to vibrational properties and quantum dynamics of molecules, in *Recent Advances in Quantum Monte Carlo Methods, Part II*, edited by S. M. Rothstein, W. A. Lester, Jr., and S. Tanaka, Singapore, 2002, World Scientific.

[199] A. Williamson, R. Hood, and J. Grossman, *Phys. Rev. Lett.* **8724** (2001).

[200] A. Luchow and S. Manten, *J. Chem. Phys.* **119**, 1307 (2003).

[201] S. Boys, *Rev. Mod. Phys.* **32**, 296 (1960).

[202] C. Edmiston and K. Ruedenberg, *Rev. Mod. Phys.* **35**, 457 (1963).

[203] J. Pipek and P. Mezey, *J. Chem. Phys.* **90**, 4916 (1989).

[204] S. Liu, J. Perez-jorda, and W. Yang, *J. Chem. Phys.* **112**, 1634 (2000).

[205] J. Demmel and X. Li, *ACM Transactions on Mathematical Software* **29**, 110 (2003).

[206] L. Greengard and V. Rokhlin, *J. Comp. Phys.* **73**, 325 (1987).

[207] R. Jastrow, *Phys. Rev.* **98**.

[208] S. A. Alexander, R. L. Coldwell, H. J. Monkhorst, and J. D. Morgan III, *J. Chem. Phys.* , 6611 (1991).

[209] S. Alexander, R. Coldwell, and J. Morgan, *J. Chem. Phys.* **97**, 8407 (1992).

[210] Z. Sun, R. Barnett, and W. Lester, *J. Chem. Phys.* **96**, 2422 (1992).

[211] F. Galvez, E. Buendia, and A. Sarsa, *Int. J. Quantum. Chem.* **87**, 270 (2002).

[212] A. Veillard and E. Clementi, *J. Chem. Phys.* **49**, 2415 (1968).

[213] B. Botch and T. Dunning, *J. Chem. Phys.* **76**, 6046 (1993).

[214] C. W. Greef and W. A. Lester, Jr., *J. Chem. Phys.* **104**, 1973 (1996).

[215] L. Mitas, in *Computer Simulation Studies in Condensed Matter Physics*, p. 94, Springer, Berlin, v edition, 1993.

[216] R. Jones, *Chem. Phys. Lett.* **99**, 1194 (1993).

[217] K. Sunil and K. Jordan, *J. Phys. Chem.* **92**, 2774 (1988).

[218] C. Bauschlicher, L. Barnes, and P. Taylor, *J. Phys. Chem.* **93**, 2932 (1989).

[219] J. Harrington and J. Weisshaar, *J. Chem. Phys.* **93**, 854 (1990).

[220] P. Cade and W. Huo, *J. Chem. Phys.* **47**, 614 (1967).

[221] A. Luchow and J. Anderson, *J. Chem. Phys.* **105**, 7573 (1996).

[222] H. Kietzmann, R. Rochow, G. Gantefor, W. Eberhardt, K. Vietze, G. Seifert, and P. Fowler, *Phys. Rev. Lett.* **81**, 5378 (1998).

[223] H. Handschuh, G. Gantefor, B. Kessler, P. Bechthold, and W. Eberhardt, *Phys. Rev. Lett.* **74**, 1095 (1995).

[224] A. Van Orden and R. Saykally, *Chem. Rev.* **98**, 2313 (1998).

[225] W. Weltner and R. Vanzee, *Chem. Rev.* **89**, 1713 (1989).

[226] J. Martin, J. Francois, and R. Gijbels, *J. Mol. Struct.* **294**, 21 (1993).

[227] A. Ott, G. Rechtsteiner, C. Felix, O. Hampe, M. Jarrold, R. Van Duyne, and K. Raghavachari, *J. Chem. Phys.* **109**, 9652 (1998).

[228] J. Grossman, L. Mitas, and K. Raghavachari, *Phys. Rev. Lett.* **75**, 3870 (1995).

[229] T. Torelli and L. Mitas, *Phys. Rev. Lett.* **85**, 1702 (2000).

[230] Y. Shlyakhter, S. Sokolova, A. Luchow, and J. Anderson, *J. Chem. Phys.* **110**, 10725 (1999).

[231] J. Martin and P. Taylor, *J. Phys. Chem.* **100**, 6047 (1996).

[232] J. C. Grossman and L. Mitas, in *Recent Advances in Quantum Monte Carlo Methods, Part II*, edited by S. M. Rothstein, W. A. Lester, Jr., and S. Tanaka, Singapore, 2002, World Scientific.

[233] C. X. Linag and H. F. Schaefer, *Chem. Phys. Lett.* **169**, 150 (1990).

[234] V. Linag and J. Almlof, *Theor. Chim. ACTA* **83**, 227 (1992).

[235] S. Sokolova, A. Luchow, and J. Anderson, *Chem. Phys. Lett.* **323**, 229 (2000).

[236] P. Taylor, E. Bylaska, J. Weare, and R. Kawai, *Chem. Phys. Lett.* **235**, 558 (1995).

[237] P. Kent, M. Towler, R. Needs, and G. Rajagopal, *Phys. Rev. B* **62**, 15394 (2000).

[238] C. W. Greeff and W. A. Lester, Jr., *J. Chem. Phys.* **106**, 6412 (1997).

[239] J. Grossman, M. Rohlfing, L. Mitas, S. Louie, and M. Cohen, *Phys. Rev. Lett.* **86**, 472 (2001).

[240] R. Needs, P. Kent, A. Porter, M. Towler, and G. Rajagopal, *Int. J. Quantum Chem.* **86**, 218 (2002).

[241] A. Porter, O. Al-Mushadani, M. Towler, and R. Needs, *J. Chem. Phys.* **114**, 7795 (2001).

[242] A. Porter, M. Towler, and R. Needs, *Phys. Rev. B* **6403** (2001).

[243] W. Allen and H. Schaefer, *Chem. Phys.* **108**, 243 (1986).

[244] J. Leszczynski, J. Huang, P. Schreiner, G. Vacek, J. Kapp, P. Schleyer, and H. Schaefer, *Chem. Phys. Lett.* **244**, 252 (1995).

[245] R. Grev and H. Schaefer, *J. Chem. Phys.* **97**, 8389 (1992).

[246] D. Feller and D. Dixon, *J. Phys. Chem. A* **103**, 6413 (1999).

[247] L. Mitas, *Phys. Rev. A* **49**, 4411 (1994).

[248] P. A.Christiansen, *J. Chem. Phys.* **95**, 361 (1991).

[249] S. Sokolova and A. Luchow, *Chem. Phys. Lett.* **320**, 421 (2000).

[250] L. Wagner and L. Mitas, *Chem. Phys. Lett.* **370**, 412 (2003).

[251] H. Flad and M. Dolg, *J. Phys. Chem.* **100**, 6152 (1996).

[252] H. Flad, F. Schautz, Y. Wang, M. Dolg, and A. Savin, *European Physical Journal D* **6**, 243 (1999).

[253] F. Schautz, H. Flad, and M. Dolg, *Theor. Chem. Acc.* **99**, 231 (1998).

[254] M. Dolg and H. Flad, *Mol. Phys.* **91**, 815 (1997).

[255] I. V. Ovcharenko, W. A. Lester, Jr., C. Xiao, and F. Hagelberg, *J. Chem. Phys.* **114**, 9028 (2001).

[256] O. El Akramine, W. A. Lester, Jr., X. Krokidis, C. Taft, T. Guimaraes, A. Pavao, and R. Zhu, *Mol. Phys.* **101**, 277 (2003).

[257] B. Guo, K. Kerns, and A. Castleman, *Science* **255**, 1411 (1992).

[258] P. Hay and W. Wadt, *J. Chem. Phys.* **82**, 270 (1985).

[259] P. Hay and W. Wadt, *J. Chem. Phys.* **82**, 299 (1985).

[260] A. S. Sokolova, *Chem. Phys. Lett.* **320**, 421 (2000).

[261] M. Hack, R. Maclagan, G. Scuseria, and M. Gordon, *J. Chem. Phys.* **104**, 6628 (1996).

[262] J. Pople, M. Headgordon, D. Fox, K. Raghavachari, and L. Curtiss, *J. Chem. Phys.* **90**, 5622 (1989).

[263] L. Curtiss, C. Jones, G. Trucks, K. Raghavachari, and J. Pople, *J. Chem. Phys.* **93**, 2537 (1990).

[264] J. Grossman, *J. Chem. Phys.* **117**, 1434 (2002).

[265] D. Feller and K. Peterson, *J. Chem. Phys.* **110**, 8384 (1999).

[266] L. Curtiss, K. Raghavachari, P. Redfern, and J. Pople, *J. Chem. Phys.* **106**, 1063 (1997).

[267] S. Lu, *J. Chem. Phys.* **118**, 9528 (2003).

[268] S. Lu, *J. Chem. Phys.* **287**, 297 (2003).

[269] R. Grev, C. Janssen, and H. Schaefer, *J. Chem. Phys.* **95**, 5128 (1991).

[270] W. Klopper, K. Bak, P. Jorgensen, J. Olsen, and T. Helgaker, *J. Phys. B* **32**, R103 (1999).

[271] T. Helgaker, J. P., and O. J., *Molecular electronic-structure theory*, Wiley, New York, 2000.

[272] W. Kolos and L. Wolniewicz, *J. Chem. Phys.* **49**, 404 (1968).

[273] L. Mitas, J. Therrien, R. Twesten, G. Belomoin, and M. H. Nayfeh, *Appl. Phys. Lett.* **78**, 1918 (2001).

[274] V. I. Klimov, A. A. Mikhailovsky, S. Xu, A. Malko, J. A. Hollingsworth, C. A. Leatherdale, H. J. Eisler, and M. G. Bawendi, *Science* **290**, 314 (2000).

[275] F. A. Reboredo, A. Franceschetti, and A. Zunger, *Appl. Phys. Lett.* **75**, 2972 (1999).

[276] A. Puzder, A. J. Williamson, J. C. Grossman, and G. Galli, *Materials Science and Engineering B-Solid State Materials for Advanced* **96**, 80 (2002).

[277] A. J. Williamson, J. C. Grossman, R. Q. Hood, A. Puzder, and G. Galli, *Phys. Rev. Lett.* **89** (2002).

[278] G. Belomoin, J. Therrien, A. Smith, S. Rao, R. Twesten, S. Chaieb, M. H. Nayfeh, L. Wagner, and L. Mitas, *Appl. Phys. Lett.* **80**, 841 (2002).

[279] P. Bokes, I. Stich, and L. Mitas, *Chem. Phys. Lett.* **362**, 559 (2002).

[280] O. Gunnarsson and B. I. Lundqvist, *Phys. Rev. B* **13**, 4274 (1976).

[281] A. Puzder, A. J. Williamson, J. C. Grossman, and G. Galli, *Phys. Rev. Lett.* **88** (2002).

[282] A. Puzder, A. J. Williamson, J. C. Grossman, and G. Galli, *Physica Status Solidi B-Basic Research* **233**, A1 (2002).

[283] A. Puzder, A. J. Williamson, J. C. Grossman, and G. Galli, *J. Chem. Phys.* **117**, 6721 (2002).

[284] G. Belomoin, E. Rogozhina, J. Therrien, P. V. Braun, L. Abuhassan, M. H. Nayfeh, L. Wagner, and L. Mitas, *Phys. Rev. B* **65** (2002).

[285] W. C. W. Chan and S. M. Nie, *Science* **281**, 2016 (1998).

[286] G. Allan, C. Delerue, and M. Lannoo, *Phys. Rev. Lett.* **78**, 3161 (1997).

[287] C. C. Chen, A. B. Herhold, C. S. Johnson, and A. P. Alivisatos, *Science* **276**, 398 (1997).

[288] D. J. Norris and M. G. Bawendi, *J. Chem. Phys.* **103**, 5260 (1995).

[289] M. G. Bawendi, W. L. Wilson, L. Rothberg, P. J. Carrol, T. M. Jedju, M. L. Steigerwald, and L. E. Brus, *Phys. Rev. Lett.* **65**, 1623 (1990).

[290] E. W. Draeger, J. C. Grossman, A. J. Williamson, and G. Galli, *Physica Status Solidi B-Basic Research* **239**, 11 (2003).

[291] J. P. Wilcoxon, G. A. Samara, and P. N. Provencio, *Phys. Rev. B* **60**, 2704 (1999).

[292] E. W. Draeger, J. C. Grossman, A. J. Williamson, and G. Galli, *Phys. Rev. Lett.* **90** (2003).

[293] J. D. Holmes, K. Ziegler, R. C. Doty, L. E. Pell, K. P. Johnston, and B. A. Korgel, *J. Am. Chem. Soc.* **123**, 3743 (2001).

[294] G. Belomoin, J. Therrien, A. Smith, S. Rao, R. Twesten, S. Chaieb, M. H. Nayfeh, L. Wagner, and L. Mitas, *Appl. Phys. Lett.* **80**, 841 (2002).

[295] G. Vignale and M. Rasolt, *Phys. Rev. B* **37**, 10685 (1988).

[296] A. Harju, V. A. Sverdlov, B. Barbiellini, and R. M. Nieminen, *Physica B* **255**, 145 (1998).

[297] R. B. Laughlin, *Phys. Rev. B* **27**, 3383 (1983).

[298] D. Pfannkuche, V. Gudmundsson, and P. A. Maksym, *Phys. Rev. B* **47**, 2244 (1993).

[299] M. Taut, *J. Phys. A* **27**, 1045 (1994).

[300] Y. M. Blanter, N. E. Kaputkina, and Y. E. Lozovik, *Physica Scripta* **54**, 539 (1996).

[301] J. L. Zhu, J. Z. Yu, Z. Q. Li, and Y. Kawazoe, *Journal of Physics-Condensed Matter* **8**, 7857 (1996).

[302] F. Bolton, *Phys. Rev. B* **54**, 4780 (1996).

[303] M. Dineykhan and R. G. Nazmitdinov, *Phys. Rev. B* **55**, 13707 (1997).

[304] A. Harju, V. A. Sverdlov, R. M. Nieminen, and V. Halonen, *Phys. Rev. B* **59**, 5622 (1999).

[305] A. Harju, S. Siljamaki, and R. M. Nieminen, *Phys. Rev. B* **65** (2002).

[306] S. M. Reimann, M. Koskinen, and M. Manninen, *Phys. Rev. B* **62**, 8108 (2000).

[307] F. Pederiva, C. J. Umrigar, and E. Lipparini, *Phys. Rev. B* **62**, 8120 (2000).

[308] F. Pederiva, C. J. Umrigar, and E. Lipparini, *Phys. Rev. B* **68**, 089901 (2003).

[309] G. W. Bryant, *Phys. Rev. Lett.* **59**, 1140 (1987).

[310] C. E. Creffield, W. Hausler, J. H. Jefferson, and S. Sarkar, *Phys. Rev. B* **59**, 10719 (1999).

[311] E. Wigner, *Phys. Rev.* **46**, 1002 (1934).

[312] D. G. Austing, S. Sasaki, S. Tarucha, S. M. Reimann, M. Koskinen, and M. Manninen, *Phys. Rev. B* **60**, 11514 (1999).

[313] I. H. Lee, Y. H. Kim, and K. H. Ahn, *Journal of Physics: Condensed Matter* **13**, 1987 (2001).

[314] E. Rasanen, H. Saarikoski, V. N. Stavrou, A. Harju, M. J. Puska, and R. M. Nieminen, *Phys. Rev. B* **67** (2003).

[315] H. Saarikoski, E. Rasanen, S. Siljamaki, A. Harju, I. J. Puska, and R. M. Nieminen, *European Physical Journal B* **26**, 241 (2002).

[316] R. Berkovits, *Phys. Rev. Lett.* **81**, 2128 (1998).

[317] C. M. Canali, *Phys. Rev. Lett.* **84**, 3934 (2000).

[318] H. Tamura and M. Ueda, *Phys. Rev. Lett.* **79**, 1345 (1997).

[319] J. H. Ahn, K. Richter, and I. H. Lee, *Phys. Rev. Lett.* **83**, 4144 (1999).

[320] A. Cohen, K. Richter, and R. Berkovits, *Phys. Rev. B* **60**, 2536 (1999).

[321] S. Levit and D. Orgad, *Phys. Rev. B* **60**, 5549 (1999).

[322] D. Ullmo and H. U. Baranger, *Phys. Rev. B* **64**, 245324 (2001).

[323] K. Hirose, F. Zhou, and N. S. Wingreens, *Phys. Rev. B* **63**, 075301 (2001).

[324] K. Hirose and N. S. Wingreens, *Phys. Rev. B* **65**, 193305 (2002).

[325] Y. Wan, G. Ortiz, and P. Phillips, *Phys. Rev. B* **55**, 5313 (1997).

[326] E. Lee, A. Puzder, M. Y. Chou, T. Uzer, and D. Farrelly, *Phys. Rev. B* **57**, 12281 (1998).

[327] A. D. Guclu, J. S. Wang, and H. Guo, *Phys. Rev. B* **68** (2003).

[328] K. Kang, M. C. Cha, and S. R. E. Yang, *Phys. Rev. B* **56**, R4344 (1997).

[329] A. Harju, S. Siljamaki, and R. M. Nieminen, *Phys. Rev. Lett.* **88** (2002).

[330] H. Saarikoski, M. J. Puska, and R. M. Nieminen, *Int. J. Quantum Chem.* **91**, 490 (2003).

[331] J. Shumway, A. Franceschetti, and A. Zunger, *Phys. Rev. B* **6315** (2001).

[332] T. Tsuchiya, *J. Lumin.* **87-9**, 509 (2000).

[333] T. Tsuchiya and S. Katayama, *Solid-State Electron.* **42**, 1523 (1998).

[334] A. J. Williamson, A. Franceschetti, and A. Zunger, *Europhysics Letters* **53**, 59 (2001).

[335] J. Moskowitz and K. Schmidt, *J. Chem. Phys.* **97**, 3382 (1992).

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

[337] T. Andersen, H. Haugen, and H. Hotop, *J. of Phys. Chem. Ref. Data* **28**, 1511 (1999).

[338] J. Chase, M.W., C. Davies, J. Downey, J.R., D. Frurip, R. McDonald, and A. Syverud, *JANAF Thermochemical Tables*, American Chemical Society, American institute of Physics, Midland, Michigan, 1986.

[339] S. Lu, *J. Chem. Phys.* **118**, 9528 (2003).

[340] S. Lu, *J. Chem. Phys.* **118**, 6152 (2003).

[341] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, and J. A. Pople, Gaussian 98 (Revision A7), Gaussian, Inc. Pittsburg, Pa., 1998.

Table 1 VMC energies (hartrees) for first-row atoms and percentages of correlation energy recovered ($\%E_{corr}$). for SMBH correlation functions of 7,9,17, and 42 terms. VMC uncertainties given in parenthesis.

	E_7^a		E_9^a		E_{17}^a		E_{17}^b		E_{42}^b
Atoms	Energy	$\%E_{corr}$	Energy	$\%E_{corr}$	Energy	$\%E_{corr}$	Energy	$\%E_{corr}$	Energy
He	-2.8997(2)	90	-2.9029(1)	98	-2.9036(1)	100	-2.903660(5)	99	-2.903717(8)
Li	-7.4746(6)	92	-7.4731(6)	89	-7.4768(3)	97	-7.4770(1)	97	-7.47722(4)
Be	-14.6259(7)	56	-14.6332(8)	64	-14.6370(6)	68	-14.6647(1)	97.2	-14.6659(2)
B	-24.5946(7)	52	-24.6113(8)	66	-14.6165(6)	69	-24.6383(6)	87	-24.6403(5)
C	-37.7721(7)	53	-37.7956(7)	68	-37.8017(6)	72	-37.8192(4)	83	-37.8225(4)
N	-54.5019(7)	54	-54.5390(6)	78	-54.5456(6)	77	-54.5525(7)	80.4	-54.5563(2)
O	-74.9469(6)	53	-75.0109(4)	78	-75.0146(7)	80	-75.0226(3)	82	-75.0270(1)
F	-99.5746(9)	51	-99.6685(5)	80	-99.6736(7)	82	-99.6881(7)	86	-99.6912(2)
Ne	-128.7689(8)	57	-128.8771(5)	85	-128.8796(6)	85	-128.89307(7)	88.7	-128.8910(2)

[a] HF trial function multiplied by 7- (E_7),9- (E_9), or 17- (E_{17})term SMBH correlation function, from ref. [72]

[b] Two determinant MCHF trial function multiplied by 17- (E_{17})or 42- (E_{42}) term SMBH correlation function, from ref. [211]

[c] Estimated exact total non-relativistic energies, from ref. [212]

Table 2 VMC energies (hartrees) for first-row positive and negative ions and percentage of correlation energy recovered ($\%E_{corr}$) for SMBH correlation functions of 9, 17, and 42 terms. VMC uncertainties given in parenthesis.

	E_9^a		E_{17}^a		E_{17}^b		E_{42}^b		
Atoms	Energy	$\%E_{corr}$	Energy	$\%E_{corr}$	Energy	$\%E_{corr}$	Energy	$\%E_{corr}$	E_s
Li^+	-7.2795(1)	88	-	-	-7.279889(7)	100	-7.279911(5)	100	-7.279913
Be^+	-14.3191(2)	88	-	-	-14.3238(2)	97.9	-14.3241(1)	98.5	-14.3248
B^+	-24.3040(7)	60	-	-	-24.3455(3)	97	-24.3468(4)	98.1	-24.34983
C^+	-37.3807(7)	64	-	-	-37.4132(3)	87	-37.4144(8)	88.0	-37.43103
N^+	-54.0071(2)	71	-	-	-54.0258(8)	83	-54.0286(5)	84.4	-54.0546
O^+	-74.5211(4)	77	-	-	-74.5304(10)	81.3	-74.5332(12)	82.7	-74.5668
F^+	-99.0292(9)	76	-	-	-99.047(3)	82.5	-99.054(1)	85.1	-99.0928
Ne^+	-128.0794(3)	81	-	-	-128.097(2)	85.8	-1298.100(1)	86.7	-128.1431
Li^-	-	-	-	-	-7.4982(1)	96.5	-7.4989(3)	97.4	-7.5008
B^-	-24.6148(5)	66	-24.6243(4)	73	-24.6374(5)	82	-24.6411(8)	84.1	-24.6642
C^-	-37.8375(7)	70	-37.8433(6)	76	-54.8575(9)	81.4	-37.8617(9)	83.7	-37.89138
O^-	-75.0479(4)	78	-75.0612(4)	82	-75.072(3)	85.2	-75.079(1)	87.3	-75.121
F^-	-99.7840(6)	81	-99.8006(3)	85	-99.8136(10)	88.6	-99.8210(10)	90.5	-99.85889

[a] HF trial function multiplied by 7- (E_7),9- (E_9), or 17- (E_{17})term SMBH correlation function, from ref. [335].

[b] Two determinant MCHF trial function multiplied by 17- (E_{17})or 42- (E_{42}) term SMBH correlation function, from ref. [211].

Table 3 VMC ionization potentials (IP) and electron affinities (EA) of first-row atoms for SMBH correlation functions of 9 and 42 terms (eV).

	IP			EA		
Atom	E_9^a	E_{42}^b	Exp[c]	E_{17}^a	E_{42}^b	Exp[d]
Li	5.27(2)	5.38(1)	5.3896	-	0.60(1)	0.62
B	8.55(2)	7.99(3)	8.3	0.24(2)	0.02(2)	0.28
C	11.29(3)	11.10(3)	11.26	1.27(2)	1.07(4)	1.26
N	14.7(2)	14.35(2)	14.55	-	-	-
O	13.33(2)	13.44(4)	13.62	1.30(2)	1.42(3)	1.46
F	17.40(2)	17.34(3)	17.45	3.06(4)	3.53(3)	3.3
Ne	21.71(2)	21.52(3)	21.62	-	-	-

[a] HF trial function multiplied by a 9- (E_9) or 17- (E_{17})term SMBH correlation function, from ref. [335]
[b] MCHF trial function multiples by 42-term (E_{42}) SMBH correlation function, from ref [211]
[c] Experimental ionization potentials, from ref. [336]

[d] Experimental electron affinities, from ref. [337]

Table 4 Atomization energies of the G1 set of molecules (kcal/mol)

Molecule	CCSD(T)[a]	DMC[b]	Exp.[c]
LiH	56.0	55.3(2)	56.00(1)
		56.72(3)[d]	
BeH	47.6	43.0(2)	46.90(1)
		54.5(4.0)[d]	
CH	80.0	79.5(2)	79.90(2)
		79.3(3)[e]	
CH_2 ($^3B_1$)	179.5	181.9(4)	179.6(4)
CH_2 ($^1A_1$)	170.4	169.7(4)	170.6(4)
CH_3	289.0	290.0(2)	289.3(2)
CH_4	392.2	395.0(2)	392.5(1)
		392.67(4)[d]	
NH	78.2	78.2(4)	79.0(4)
NH_2	169.4	169.2(4)	170.0(3)
NH_3	276.2	276.5(2)	276.7(1)
		276.87(7)[d]	
OH	101.7	101.2(3)	101.4(3)
H_2O	218.9	219.4(2)	219.35(1)
		219.49(5)	
HF	135.1	135.9(2)	135.2(2)
		135.3(9)[d]	
		135.62(6)[d]	
$SiH_2(^1A_1)$	145.9	145.5(2)	144.4(2)
SiH_2 ($^3B_1$)	124.6	125.8(2)	123.4(2)
SiH_3	214.6	215.1(2)	214(1)
SiH_4	304.6	305.8(2)	302.6(5)
PH_2	145.9	143.7(2)	144.7(6)
PH_3	227.0	224.8(2)	228.6(4)
H_2S	173.8	172.1(4)	173.1(2)
HCL	102.3	103.4(4)	102.2(5)
Li_2	23.9	23.5(2)	23.9(7)
		24.16(4)[d]	
LiF	137.1	145.1(4)	138(2)
		137.02(17)[d]	
C_2H_2	388.9	390.0(4)	386.9(2)
		388.93(33)[d]	
		380.7(5)[e]	
C_2H_4	531.6	533.5(4)	531.9(1)
C_2H_6	665.9	669.3(4)	666.3
CN	175.9	170.5(4)	178(2)
HCN	303.0	302.0(8)	301(2)
		303.12(55)[d]	
CO	255.7	253.2(3)	256.2(2)

[a]CBS CCSD(T) results obtained from ref. [265].

[b]DMC results obtained from ref. [264] unless noted otherwise.

[c]Experimental atomization results obtained from ref. [338].

[d]Obtained from Shih-Lu [268, 339, 340] and corrected form zero point energies as given by Grev *et al.*[269]

[e]Obtained from Barnett *et al.*[96]

Table 5 Total DMC and CCSD(T) energies for 20 molecules (hartrees)

		CCSD(T)[b]			
Molecules	QCI/APNO[a]	cc-pVDZ	cc-pVTZ	CBS	DMC[c]
H_2	1.1745	1.1634	1.1723	1.175	1.1739(1)
CH_2	39.0774	39.022	39.0614	39.1304	39.1165(3)
CH_4	40.4574	40.3868	40.4381	40.5125	40.5005(3)
NH_3	56.5039	56.402	56.4732	56.5598	56.5485(4)
H_2O	76.3757	76.241	76.3322	76.432	76.4207(2)
C_2H_2	77.2211	77.1092	77.1876	77.3287	77.3110(4)
C_2H_4	78.4744	78.3544	78.4388	78.5823	78.5644(3)
HNC	93.3159	93.1632	93.2513	93.4253	93.3737(3)
HCN	93.2921	93.1884	93.2751	93.4017	93.3987(4)
HF	100.394	100.2275	100.3379	100.451	100.4466(3)
N_2	109.422	109.2753	109.3739	109.533	109.5046(3)
N_2H_2	110.531	110.367	110.478	110.642	110.6054(5)
CO	113.207	113.0544	113.1555	113.317	113.2877(4)
H_2CO	114.389	114.2183	114.3338	114.5	114.4739(4)
HNO	130.362	130.171	130.2984	130.474	130.4371(4)
H_2O_2	151.44	151.1937	151.3586	151.551	151.5213(3)
HOF	175.426	175.1519	175.3343	175.54	175.5120(4)
CO_2	188.418	188.1475	188.3271	188.586	188.5429(4)
F_2	199.398	199.0975	199.2961	199.515	199.4841(4)
O_3	225.245	224.9091	225.1326	225.418	225.3410(4)

[a] CBS-quadratic-CI atomic-pair natural orbital (QCI/APNO) energies, from ref. [270]

[b] CCSD(T) valence energies from ref. [270].

[c] DMC energies from ref. [65]

Table 6 Deviation from experiment of calculated heats for reaction for 17 reactions (kcal/mol)

	Other			CCSD(T)				
Reaction	DFT[a]	G2[b]	CBS[c]	cc-pVDZ	cc-pVTZ	∞[d]	DMC[e]	Exp.[f]
$CH_2 + H_2 \rightarrow CH4$	0.24	2.39	0.00	2.87	0.96	0.24	-1.91	-130.02(48)
$C_2H_2 + H_2 \rightarrow C_2H_4$	-1.20	0.24	-0.96	-2.39	-0.72	-0.72	-1.43	-48.52(48)
$C_2H_2 + 3H_2 \rightarrow 2CH_4$	-0.96	1.43	-0.96	-2.63	-1.20	-0.24	0.96	-106.60(48)
$N_2H_2 \rightarrow N_2 + H_2$	2.87	-0.72	0.00	-3.59	-1.20	0.24	-4.30	-41.59
$CO + H_2 \rightarrow H_2CO$	-3.11	0.96	-0.24	4.30	0.96	-0.48	-2.63	-5.02(24)
$N_2 + 3H_2 \rightarrow 2NH_2$	-0.48	4.06	0.72	14.58	3.82	-0.24	-5.26	-39.20(24)
$F_2 + H_2 \rightarrow 2HF$	5.50	-0.24	1.20	12.43	4.06	-0.48	-13.14	-134.56(24)
$O_3 + H_2 \rightarrow 3H_20$	5.74	5.02	1.20	18.88	4.06	-3.11	-27.96	-222.9(48)
$H_2CO + 2H_2 \rightarrow H_2O + CH_4$	4.06	3.82	0.72	7.89	2.15	0.24	-2.63	-59.99(24)
$H_2O_2 + H_2 \rightarrow 2H_2O$	4.54	1.20	0.72	8.60	3.11	0.72	-4.54	-87.24(48)
$CO + 3H_2 \rightarrow CH_4 + NH_3$	0.96	5.02	0.48	12.19	3.11	-0.24	-5.26	-65.01(24)
$HNC + 3H_2 \rightarrow CH_4 + NH_3$	0.00	3.59	-0.24	7.17	1.43	-0.24	-4.30	-76.48(72)
$HNO + 2H_2 \rightarrow H_2O + NH_3$	3.59	4.30	0.96	14.34	3.59	-0.48	-9.80	-106.12(24)
$HNC \rightarrow HCN$	1.43	0.24	0.48	-0.72	0.24	0.24	-0.48	-15.30
$H_2O + F_2 \rightarrow HOF + HF$	2.39	1.43	2.87	5.26	3.35	2.63	-2.87	-30.83(96)
$CO_2 + 4H_2 \rightarrow 2H_2O + CH_2$	7.89	6.69	1.43	15.77	4.06	0.00	-6.69	-58.32(24)
$2CH_2 \rightarrow C_2H_4$	-0.24	3.59	0.00	5.74	2.15	-0.24	-6.45	-201.72(72)
Mean error	1.95	2.53	0.49	7.10	2.00	-0.13	-5.81	
Absolute error	2.66	2.64	0.77	8.20	2.36	0.63	5.92	
Root-mean-square-error	3.47	3.28	1.03	9.78	2.70	1.05	6.34	
Maximum absolute error	7.89	6.69	1.43	15.77	3.82	0.72	13.14	

[a] Resuls obtained using the B3LYP/6-311++G(2df,p) approach
[b] Gaussian-2 model as implemented in the Gaussian program[341]; Core correlation is not inlcuded.
[c] CBS-QCI/APNO method as implemented in the Gaussian program[341]; Core correlation in not included.
[d] Extrapolation to the limit of a complete basis.
[e] DMC results obtained from ref.[65]
[f] Experimental heats of reaction obtained from ref. [270]

Table 7 Effect of pasivants on the optical gap of Si nano-clusters

Passivant	Bond type	Property	Optical gap (eV)	Effect
H[a]	single	low electronegativity	3.4	no effect
O[a]	double	high electronegativity	2.2	reduction
OH[a]	single	lone pair	3.3	small effect
F[a]	single	high electronegativity	3.2	small effect
Cl[a]	single	high electronegativity	3.3	small effect
CH_2[a]	double	low electronegativity	2.5	reduction
S[a]	double	medium electronegativity	1.8	reduction
Butylamine(N)[b]	single	Lewis base	3.25	small effect
Pentane(C)[b]	single	not a Lewis base	3.55	small effect

[a] Ref. [276, 281–283] [b] Ref. [284]

Table 8 Energy gaps (in eV) for silicon nano-clusters using both LDA and QMC. For strucutres created with FPMD, a range of LDA gap energies is given, with QMC gaps claculations only for the structure with the largest LDA gap.

Cluster	E_{gap}^{LDA}	E_{gap}^{QMC}
$Si_{29}H_{22}^{FPMD}$	2.2-2.3	3.3(1)
$Si_{29}H_{24}^{nano2}$	2.9	4.1(1)
$Si_{30}H_{22}^{ideal}$	2.2	3.2(1)
$Si_{30}H_{22}^{FPMD}$	2.2-2.4	3.4(1)

Table 9 Comprison of the results obtained for a three electron quantum dot (meV).

Author	VMC	DMC
Pederiva *et al.*[a]	29.669(3)	26.488(3)
Harju *et al.*[b]	26.563(1)	NA
Harju *et al.*[c]	26.5406(8)	NA

[a] Ref. [307] Results obtained using a wave function without optimizing the exponential parameters.

[b] Ref. [304] Results obtained using a wave function optimizing the exponential parameters.

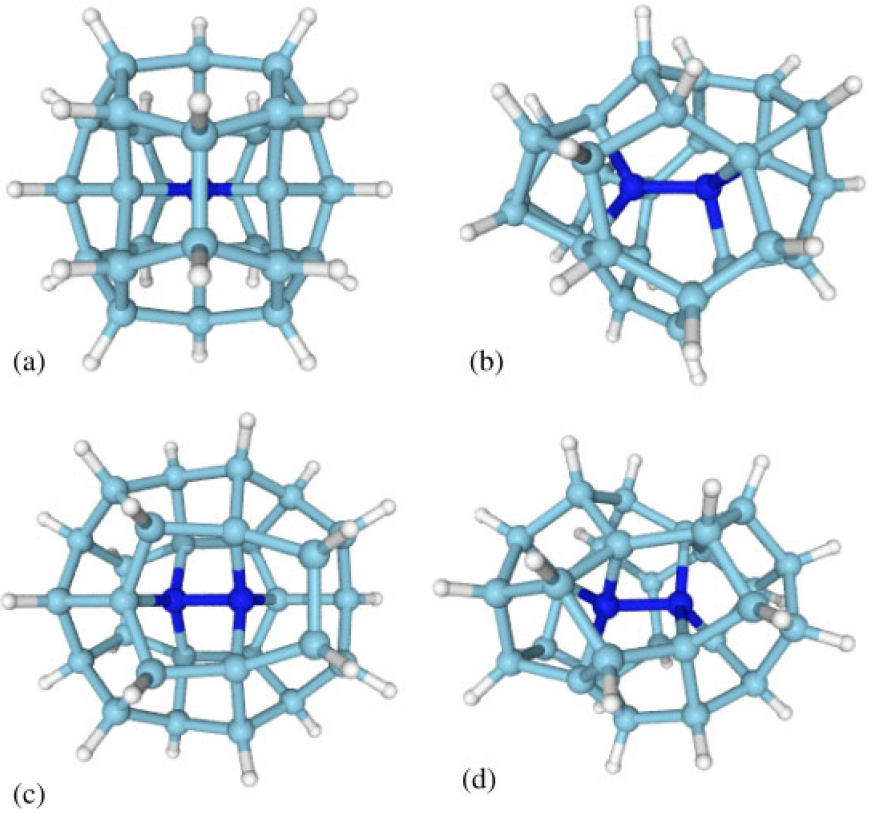

Figure 1: Results of FPMD synthesis for Si29 and Si30, compared with low-energy ideal structures.