10   SCIENTIFIC HIGHLIGHT OF THE MONTH

Quantum Monte Carlo and the CASINO program : highly accurate total energy calculations for finite and periodic systems

M. D. Towler

TCM Group, Cavendish Laboratory, University of Cambridge,
Madingley Road, Cambridge CB3 0HE, UK

Abstract

For practical computational electronic structure calculations an important and complementary alternative to density functional theory in situations where high accuracy is required is the *quantum Monte Carlo* method. This has many attractive features for probing the electronic structure of real systems. In particular, it is an explicitly many-particle technique which takes electron correlation into account from the outset, and gives consistent, highly accurate results while at the same time exhibiting favourable (cubic or better) scaling of computational cost with system size. In this article I briefly review the present state of the art, and outline the implementation of the method in the Cambridge quantum Monte Carlo code 'CASINO' [1, 2].

1   Introduction

The continuum Quantum Monte Carlo (QMC) method has been developed to calculate the properties of assemblies of interacting quantum particles. It is generally capable of doing so with great accuracy. The various different techniques which lie within its scope have in common the use of *random sampling*, and this is used because it represents by far the most efficient way to integrate numerically (wave) functions in many dimensions. In this article I shall give a brief introduction to the two most common types of QMC, namely variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC) [3, 4]. As we shall see, the first of these techniques is simple in concept and is designed just to sample a given trial wave function and calculate the expectation value of the Hamiltonian using Monte Carlo numerical integration. DMC is one of a class of so-called 'projector' methods which attempt the much more difficult job of simultaneously creating and sampling the unknown exact ground state wave function. Other variants, including those aimed at expanding the scope of the method to finite temperature such as path integral Monte Carlo (PIMC) [5, 6], or those designed to find the exact non-relativistic energy overcoming the small fixed-node approximation made in DMC (such as fermion Monte Carlo (FMC) [7, 8, 9]) will not be discussed in any

detail here. The interested reader is invited to consult the literature for more detailed discussions (the extensive bibliography in Ref. [4] is a good place to start).

QMC is perhaps best known for its early application to the homogeneous electron gas by Ceperley and Alder [10], the results of which were used to develop accurate parametrizations of the local density approximation to density functional theory (DFT) in the early 1980s. However, it is of course perfectly possible to apply the method to real systems with atoms, and for small molecules containing helium and hydrogen QMC gives total energies with an extraordinary accuracy greater than 0.01 kcal/mole ($\approx 1.5 \times 10^{-5}$ Ha or 4×10^{-4} eV). In one well-known QMC study of the H + H_2 $\longrightarrow H_2$+H potential energy surface tens of thousands of points with accuracies close to this value were computed [11]. Despite such capabilities the technology of QMC is neither mature nor particularly widely used; its routine application to arbitrary finite and periodic systems, particularly those containing heavier atoms, has long been just out of reach and there are still many open methodological and algorithmic problems to interest the computational electronic structure theorist. The situation is clearly changing however, and it ought now to be a matter of routine for people to perform accurate *a posteriori* QMC evaluations of the correlation energy starting from the results of molecular orbital or band theory calculations. Systems and problems for which accurate calculations of the total energy actually matter, and for which DFT (for example) is not sufficiently accurate, are likely more numerous than is generally believed. To this end, our group in the Cavendish Laboratory of Cambridge University has over many years developed a general-purpose QMC computer program - CASINO [1, 2]. This is capable of performing both variational and diffusion Monte Carlo calculations on a wide variety of systems, which may be finite (atoms or molecules) or have periodic boundary conditions in one, two or three dimensions, i.e., polymers, slabs (surfaces) or crystalline solids. The code may also be used to study situations without an external potential such as the homogeneous electron gas, Wigner crystals, and various electron-hole phases. We will describe CASINO in more detail presently.

One of the main reasons that QMC is currently attracting so much interest is the scaling behaviour of the necessary computational effort with system size. This is favourable enough that we can continue to apply the method to systems as large as are treated in conventional DFT (albeit with a considerably bigger pre-factor, and thus probably not on the same computers). In fact, QMC represents the most accurate method available for medium-sized and large systems. Other quantum chemistry correlated wave function methods based on the 'standard model' of multideterminant expansions - such as configuration interaction or high-order coupled cluster theory - are capable of similar accuracy for systems containing a few electrons, but as the size of the molecule is increased they quickly become too expensive. Standard Quantum Monte Carlo calculations scale as the third power of the system size (the same as DFT), and are capable of treating solid crystalline phases as well as molecules. The largest calculations done to date on the more expensive periodic systems using the regular algorithm include almost 2000 electrons per cell in the three-dimensional electron gas [12], 1732 electrons (432 atoms) per cell in crystalline silicon [13], and 1024 electrons (128 atoms) per cell in antiferromagnetic nickel oxide [14]. Furthermore the natural observation has been made that provided localized molecular or crystalline orbitals are used in constructing the QMC trial wave function, and provided these orbitals are expanded in a localized basis set, then the scaling of the basic algorithm can be substantially improved, even as far as linear scaling in some cases [15, 16]. This capability, to be discussed in Section 4.2, has recently been introduced into CASINO and this should considerably extend the range of problems that may be studied.

Before we go further, it will be useful to list some other favourable properties of the method :

- For most practical purposes the 'basis set problem' is essentially absent in DMC; errors due to the use of a finite basis set are very small since the many-electron wave function is not represented directly in terms of a basis set, but rather by the distribution of an ensemble of particles evolving in (imaginary) time. The basis set that is employed in DMC is just used to expand a guiding function required for importance sampling. The final DMC energy depends only weakly on the nodal surface of this function

(i.e., the set of points in configuration space on which the function is zero).

- The QMC algorithm is intrinsically parallel and Monte Carlo codes are thus easily adapted to parallel computers and scale linearly with the number of processors. There are no memory or disk bottlenecks even for relatively large systems.

- We can use many-electron wave functions with explicit dependence on interparticle distances and no need for analytic integrability.

- We can calculate ground states, some excited states, chemical reaction barriers and other properties within a single unified framework. The method is size-consistent and variational.

One may ask why one should formulate a method based on the many-electron wave function when so much stress is normally placed on reducing the number of variables in the quantum problem (by using, e.g., density, Green's functions, density matrices or other quantities which depend on fewer independent variables). The main point is that the many-electron wave function satisfies a rather well-known fundamental equation [17]:

$$\hat{H}\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = E\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) \tag{1}$$

The price we pay for reformulating the problem in terms of the density is that we no longer know the exact equation satisfied by the density. In DFT, the complicated many-body problem is effectively relocated into the definition of the exchange-correlation functional, whose mathematical expression is not currently known and unlikely ever to be known exactly. The inevitable approximations to this quantity substantially reduce the attainable accuracy.

Widely-used standard solid state texts often refute the possibility of solving the many-electron Schrödinger equation directly in any meaningful way for large crystalline systems. To take a particular example, Ashcroft and Mermin [18] state that, 'one has no hope of solving an equation such as Eq. 1' and one must reformulate the problem in such a way as 'to make the one-electron equations least unreasonable'. The key simplifying physical idea to allow one to treat crystalline solids however, is not the use of one-electron orbitals but simply the imposition of periodic boundary conditions. One can then have an explicitly correlated many-body wave function (i.e., with explicit dependence on the interparticle separations), in a box, embedded in an infinite number of copies of itself. The 'particles' in the many-body wave function are no longer individual electrons, but electrons in the cell we are focussing on moving in tandem with all of their periodic images. Clearly in order for this to have any chance of being an accurate approximation, the range of the electron-electron pair correlation function must be substantially shorter than the repeat distance and the box must be large enough so that the forces on the particles within it are very close to those in the bulk. If not, then we may get substantial 'finite-size errors'.

This problem is analogous to, but not quite the same as, the problem of representing an infinite system in DFT calculations. In that case Bloch's theorem is used in the extrapolation to infinite system size so that the problem of calculating an infinite number of one-electron states reduces to the calculation of a finite number (equal to the number of electrons in the primitive cell) of states at an infinite number of $\mathbf{k}$ points. Because the energy of states at $\mathbf{k}$ points which are close together are very similar, the $\mathbf{k}$ space may be 'sampled' and if this is done efficiently the calculated energy per cell approaches that in the infinite system. The situation in QMC is a little different since the explicit correlation between electrons means that the problem cannot be reduced to the primitive cell; a one-electron wave function on a $2 \times 2 \times 2$ $\mathbf{k}$ point grid corresponds to a many-electron wave function for a $2 \times 2 \times 2$ supercell in real space. There is a 'many-body Bloch theorem' expressing the invariance of the Hamiltonian under translations of all electrons by a primitive lattice vector or of a single electron by a supercell lattice vector [19], and thus there are two $\mathbf{k}$ vectors associated with the periodic many-body wave function. The error analagous to inadequate Brillouin zone sampling can be made smaller either by increasing the size of the simulation cell

or by choosing the **k** values using 'special **k** point' techniques [20]. An additional type of finite-size error arises in periodic QMC calculations (though not in DFT) when calculating interactions between particles with long-range Coulomb sums. The difference is that in QMC we deal with instantaneous positions of electron configurations, rather than with the interaction of averaged densities. When using the standard Ewald formulation [21, 22], the choice of boundary conditions leads to an effective depolarization field which is supposed to cancel the field due to notional surface charges. As all periodic copies of the simulation cell contain, for example, the same net dipole due to the random arrangement of electrons with respect to nuclei the interaction of the dipoles and the field gives rise to 'Coulomb finite size errors'. These can be substantially reduced by using special techniques [23].

A few years ago in his Nobel prize-winning address Walter Kohn suggested that the many-electron wave function is not a legitimate scientific concept for more than about a thousand particles [24]. Does this mean we're in trouble if we use them in QMC? The main idea behind this statement is that the overlap of an approximate wave function with the exact one will tend exponentially to zero as the number of particles increases, unless one uses a wave function in which the number of parameters increases exponentially. Such an object would not be computable for large systems. This is indeed true, as is easy to verify by calculating the overlap integral directly using VMC [25]. Note that one does not need the exact wave function itself to perform this calculation. Rather one can evaluate the overlap between a single-determinant wave function on its own and multiplied by a Jastrow correlation function. The fact that these share the same nodal surface does not matter since Kohn's argument is based solely on the high-dimensionality of the overlap integrals rather than, say, the explicit cancellation of positive and negative regions. However, his objection is almost certainly not relevant to the sort of QMC calculations discussed here. Clearly the successful DMC calculations of systems containing up to 2000 electrons mentioned earlier provide some evidence in this direction. Kohn's arguments were used to motivate density functional theory (DFT), but it is possible to argue that, within the standard Kohn-Sham formulation, DFT suffers from exactly the same overlap 'catastrophe'. For a large system the overlap of the determinant of Kohn-Sham orbitals with the exact one will go to zero because of the inevitable numerical inaccuracies and the approximations to the exchange-energy functional. Fortunately the overlap 'catastrophe' is irrelevant to calculating the quantities of interest. As Kohn himself points out, we are interested in quantities such as the total energy, which can be accurate even when the overlap with the exact wave function goes to zero. To get the energy right it is required only that relatively low-order correlation functions (such as the pair-correlation function) are well-described and QMC seems to manage this very well.

To understand how accurate the total energies must be we note that the main goal is to calculate the energy difference between two arrangements of a set of atoms. The desired result might be the energy required to form a defect, or the energy barrier to some process, or whatever. All electronic structure methods for large systems rely on a cancellation of errors in energy differences. For such error cancellations to occur we require that the error in the energy per atom is proportional to the number of atoms. If this condition was not satisfied then, for example, the cohesive energy would not have a well defined limit for large systems. Many VMC (and DMC) calculations have demonstrated that the commonly-used Slater-Jastrow form of many-body wave function leads to errors which are proportional to the number of atoms, and typically gives between 70 and 80% of the correlation energy independent of system size. Additional requirements on QMC algorithms are that the number of parameters in the trial wave function must not increase too rapidly with system size and that the wave function be easily computable. Fortunately the number of parameters in a Slater-Jastrow wave function increases only linearly with system size or at worst as the square of the system size, and it can be evaluated in a time which rises as a low power of the system size.

2 QMC algorithms

In this section, we will look at the basic ideas and algorithms underlying VMC and DMC.

2.1 Variational Monte Carlo

2.1.1 Basics

With variational methods we must 'guess' an appropriate many-electron wave function which is then used to calculate the expectation value of the Hamiltonian operator. In general this wave function will depend on a set of parameters $\{\alpha\}$ which can be varied to optimize the function and minimize either the energy or the statistical variance. The energy thus obtained is an upper bound to the true ground state energy,

$$\frac{\langle \Psi_T(\{\alpha\})|\hat{H}|\Psi_T(\{\alpha\})\rangle}{\langle \Psi_T(\{\alpha\})|\Psi_T(\{\alpha\})\rangle} = E(\{\alpha\}) \geq E_0 \; . \tag{2}$$

The expectation value of the Hamiltonian $\hat{H}$ with respect to the trial wave function Ψ_T can be written as

$$\langle \hat{H} \rangle = \frac{\int E_L(\mathbf{R})\Psi_T^2(\mathbf{R})\, d\mathbf{R}}{\int \Psi_T^2(\mathbf{R})\, d\mathbf{R}}, \tag{3}$$

where $\mathbf{R}$ is a $3N$ dimensional vector giving the coordinates $(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ of the N particles in the system, and $E_L(\mathbf{R}) = \frac{\hat{H}(\mathbf{R})\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}$ is known as the *local energy*.

We can evaluate this expectation value by using the Metropolis algorithm [26] to generate a sequence of configurations $\mathbf{R}$ distributed according to $\Psi_T^2(\mathbf{R})$ and averaging the corresponding local energies,

$$\langle \hat{H} \rangle = \frac{1}{M}\sum_{i=1}^{M} E_L(\mathbf{R}_i) = \frac{1}{M}\sum_{i=1}^{M} \frac{\hat{H}\Psi_T(\mathbf{R}_i)}{\Psi_T(\mathbf{R}_i)} \; . \tag{4}$$

The question of whether or not we get the right answer with this approach is just one of *complexity* ; can we create a wave function with enough variational freedom so that the energy approaches the exact (non-relativistic) ground state energy? The answer in general is no. There is no systematic way in which one can improve the wave function until the correct answer is reached, and in general, we shouldn't normally expect to recover much more than 80% of the correlation energy in this way. As we shall see, the the final 20% or so can be calculated by feeding the VMC wave function into a projector method such as DMC. This, to my mind, is the main use of VMC and in our laboratory we rarely use it as a method in its own right when performing calculations. With this attitude, it is not generally necessary to kill oneself optimizing wave functions in order to recover an extra 1% of the correlation energy with VMC - it is better to use DMC and let the computer do the work for you. Although the efficiency of the DMC calculations is increased with more accurate trial functions, the final DMC energy does not in principle depend on that part of the wave function that we generally optimize.

2.1.2 The form of the trial wave function

Clearly, however, for VMC the choice of the trial function is particularly important as it directly determines the accuracy of the calculation ; the answer will approach the true energy from above as we use better and better wave functions. An additional consideration is the 'zero variance principle'. As the trial function approaches an exact eigenstate the local energy $\frac{\hat{H}\Psi}{\Psi}$ approaches a constant, E, everywhere in configuration space (check the Schrödinger equation again!) and hence the variance approaches zero.

Through its direct influence on the variance of the energy the accuracy of the trial wave function thus determines the amount of computation required to achieve a specified accuracy. When optimizing wave functions, one can therefore choose to use energy or variance as the objective function to be minimized.

The fact that arbitrary wave function forms can be used is one of the defining characteristics of QMC. We do not need to be able to integrate the wave function analytically, as is done for example in quantum chemistry methods with Gaussian basis functions. We just need to be able to *evaluate* it at a point in the configuration space. If the electrons and nuclei have certain fixed positions in space, what is the value of the wave function? This being the case, we can use correlated wave functions which depend explicitly on the distances between particles. The most commonly-used functional form is known as the Slater-Jastrow wave function [27]. This consists of a single Slater determinant (or sometimes a linear combination of a small number of them) multiplied by a positive-definite Jastrow correlation function which is symmetric in the electron coordinates and depends on the inter-particle distances. The Jastrow function allows efficient inclusion of both long and short range correlation effects. As we shall see however, the final DMC answer depends only on the nodal surface of the wave function and this cannot be affected by the nodeless Jastrow. In DMC it serves mainly to decrease the amount of computer time required to achieve a given statistical error bar and improve the stability of the algorithm.

The basic functional form of the Slater-Jastrow function is

$$\Psi(\mathbf{X}) = e^{J(\mathbf{X})} \sum_n c_n D_n(\mathbf{X}) \ , \tag{5}$$

where $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ and $\mathbf{x}_i = \{\mathbf{r}_i, \sigma_i\}$ denotes the space-spin coordinates of electron i, $e^{J(\mathbf{X})}$ is the Jastrow factor, the c_n are coefficients, and the $D_n(\mathbf{X})$ are Slater determinants of single-particle orbitals,

$$D(\mathbf{X}) = \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_1(\mathbf{x}_2) & \ldots & \psi_1(\mathbf{x}_N) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & \ldots & \psi_2(\mathbf{x}_N) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_N(\mathbf{x}_1) & \psi_N(\mathbf{x}_2) & \ldots & \psi_N(\mathbf{x}_N) \end{vmatrix} . \tag{6}$$

The orbitals in the determinants are often obtained from self-consistent DFT or Hartree-Fock calculations and are assumed to be products of spatial and spin factors,

$$\psi_\alpha(\mathbf{x}) = \psi_\alpha(\mathbf{r})\delta_{\sigma,\sigma_\alpha} \ , \tag{7}$$

Here $\delta_{\sigma,\sigma_\alpha} = 1$ if $\sigma = \sigma_\alpha$ and zero otherwise. If the determinant contains $N_\uparrow$ orbitals with $\sigma_\alpha = \uparrow$ and $N_\downarrow = N - N_\uparrow$ with $\sigma_\alpha = \downarrow$, it is an eigenfunction of $\hat{S}_z$ with eigenvalue $(N_\uparrow - N_\downarrow)/2$. To avoid having to sum over spin variables in QMC calculations, one generally replaces the determinants D_n by products of separate up- and down-spin determinants,

$$\Psi(\mathbf{R}) = e^{J(\mathbf{R})} \sum_n c_n D_n^\uparrow(\mathbf{r}_1, \ldots, \mathbf{r}_{N_\uparrow}) D_n^\downarrow(\mathbf{r}_{N_\uparrow+1}, \ldots, \mathbf{r}_N) \ , \tag{8}$$

where $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ denotes the spatial coordinates of all the electrons. This function is not antisymmetric under exchange of electrons with opposite spins but it can be shown that it gives the same expectation value as $\Psi(\mathbf{X})$ for any spin-independent operator. Note that the use of wave function forms in QMC which allow one to treat non-collinear spin arrangements and the resultant vector magnetization density is an interesting open problem, and we are currently working on developing such an algorithm [28].

The full Jastrow functions that we typically use in CASINO contain one- and two-electron terms and may be inhomogeneous, i.e., depend on the distances of the electrons from the ions. Their functional

form is quite complicated and involves sums in Chebyshev polynomials. However just to give you a feel for it, a simple Jastrow function for a homogeneous system such as the electron gas might look like :

$$e^{J(\mathbf{R})} \quad \text{with} \quad J(\mathbf{R}) = -\sum_{i>j} u_{\sigma_i,\sigma_j}(r_{ij}) \quad \text{and} \quad u_{\sigma_i,\sigma_j}(r_{ij}) = \frac{A}{r_{ij}}\left(1 - e^{-r_{ij}/F_{\sigma_i,\sigma_j}}\right) \ . \tag{9}$$

where r_{ij} is the distance between electrons i and j, and F is chosen so that the electron-electron cusp conditions are obeyed i.e. $F_{\uparrow\uparrow} = \sqrt{2A}$ and $F_{\uparrow\downarrow} = \sqrt{A}$. The value of A, and the coefficients in the Chebyshev expansion in the full version of the Jastrow function, will need to be optimized using, for example, variance minimization.

2.1.3 Optimization of trial wave functions

In our discussion so far we have said nothing about how we actually perform the wave function optimization, and this is clearly a critical step. Our code CASINO, for example, allows optimization of the parameters in the Jastrow factor, of the coefficients of the determinants of a multi-determinant wave function, and of various parameters in specialized wave functions such as pairing parameters in electron-hole gases and Padé coefficients in orbitals in Wigner crystals. Optimizing the orbitals themselves in the presence of the Jastrow factor would be a good thing to be able to do, since it optimizes the nodal surface in some sense and allows improvement of the DMC answer. This remains an active problem, although some progress has been made [29, 30].

There are many approaches to wave function optimization, but as far as our code is concerned, optimization of the wave function is achieved by minimizing the variance of the energy,

$$\sigma_E^2(\alpha) = \frac{\int \Psi^2(\alpha)\,[E_L(\alpha) - E_V(\alpha)]^2\,d\mathbf{R}}{\int \Psi^2(\alpha)\,d\mathbf{R}} \ , \tag{10}$$

where α is the set of parameters, and E_V is the variational energy. There are a number of reasons for preferring variance minimization over energy minimization (including the fact that the variance has a known lower bound of zero) but the most important is that it appears to show better numerical stability, particularly in large systems [31].

Minimization of σ_E^2 is carried out via a correlated-sampling approach in which a set of configurations distributed according to $\Psi^2(\alpha_0)$ is generated, where α_0 is an initial set of parameter values. The variance $\sigma_E^2(\alpha)$ is then evaluated as

$$\sigma_E^2(\alpha) = \frac{\int \Psi^2(\alpha_0)\,w(\alpha)\,[E_L(\alpha) - E_V(\alpha)]^2\,d\mathbf{R}}{\int \Psi^2(\alpha_0)\,w(\alpha)\,d\mathbf{R}} \ , \tag{11}$$

where

$$E_V(\alpha) = \frac{\int \Psi^2(\alpha_0)\,w(\alpha)\,E_L(\alpha)\,d\mathbf{R}}{\int \Psi^2(\alpha_0)\,w(\alpha)\,d\mathbf{R}} \ , \tag{12}$$

and the integrals contain a weighting factor, $w(\alpha)$, given by

$$w(\alpha) = \frac{\Psi^2(\alpha)}{\Psi^2(\alpha_0)} \ . \tag{13}$$

The parameters α are adjusted until $\sigma_E^2(\alpha)$ is minimized. The advantage of the correlated-sampling approach is that one does not have to generate a new set of configurations every time the parameter values are changed.

In order to create a set of configurations distributed according to $\Psi^2(\alpha_0)$, a VMC 'configuration generation' run must first be carried out. This is just an ordinary VMC run where 'snapshots' of the system are

created by writing the current electron positions and associated interaction energies to a file. To perform the actual minimization, we then use an algorithm which performs an unconstrained minimization (without requiring derivatives) of a sum of m squares of functions which contain n variables, where $m \geq n$. Having generated a new set of parameters with this algorithm, we then carry out a second configuration generation run with these parameters followed by a second variance minimization run, and so on. Generally very few such 'iterations' are required before the minimum is approached. The procedure can be, and in CASINO is, thoroughly automated and providing a systematic approach is adopted, optimizing VMC wave functions is not the complicated time-consuming business it once was. This is particularly the case if one requires only the optimized wave function for input into a DMC calculation, rather than being concerned with lowering the VMC energy as much as possible.

2.1.4 VMC conclusions

Although VMC can be quite powerful when applied to the right problem, the necessity of guessing the functional form of the trial function limits its accuracy and there is no known way to *systematically* improve it all the way to the exact non-relativistic limit. In practice therefore, the main use of VMC is in providing the optimized trial wave function required as an importance sampling function by the much more powerful DMC technique, which we now describe.

2.2 Diffusion Monte Carlo

Let us imagine that we are ignorant, or have simply not been paying attention in our quantum mechanics lectures, and that we believe that the wave function of the hydrogen atom looks like a square box centred on the nucleus. If we tried to calculate the expectation value of the Hamiltonian using VMC we would obtain an energy which was substantially in error. What DMC can do, in essence, is to correct the functional form of the guessed square box wave function so that it looks like the correct exponentially-decaying one before calculating the expectation value. This is a nice trick if you can do it, particularly in cases where we have very little idea of what the exact ground state wave function looks like (that is, almost all of them). Clearly the necessary algorithm is rather more involved than the VMC.

Essentially then, the DMC method is a stochastic projector method for evolving the imaginary-time Schrödinger equation,

$$-\frac{\partial \Psi(\mathbf{R}, t)}{\partial t} = -\frac{1}{2}\nabla^2\Psi(\mathbf{R}, t) + (V(\mathbf{R}) - E_T)\Psi(\mathbf{R}, t) \ . \tag{14}$$

Here the real variable t measures the progress in imaginary time and $\mathbf{R}$ is a $3N$-dimensional vector of the positions of the N electrons. $V(\mathbf{R})$ is the potential energy operator, E_T is an energy offset which only affects the normalization of the wave function Ψ, and $\nabla = (\nabla_1, \nabla_2, \ldots, \nabla_N)$ is the $3N$-dimensional gradient operator.

This equation has the property that an initial starting state $\Psi(\mathbf{R}, t = 0)$ decays towards the ground state wave function. In DMC the time evolution of Eq. 14 may be followed using a stochastic technique in which $\Psi(\mathbf{R}, t)$ is represented by an ensemble of $3N$-dimensional electron configurations ('walkers'), $\{\mathbf{R}_i\}$. The time evolution of these configurations is governed by the Green's function of Eq. 14. Within the short time approximation the Green's function separates into two processes: random diffusive jumps of the configurations arising from the kinetic term and creation/destruction of configurations arising from the potential energy term.

Unfortunately this simple algorithm suffers from two very serious drawbacks. The first is that we have implicitly assumed that Ψ is a probability distribution, even though its fermionic nature means that it must have positive and negative parts. The second problem is less fundamental but in practice very severe.

The required rate of removing or adding configurations diverges when the potential energy diverges, which occurs whenever two electrons or an electron and a nucleus are coincident. This leads to extremely poor statistical behaviour.

These problems are dealt with at a single stroke by introducing an importance sampling transformation. If we consider the *mixed distribution* $f = \Psi_T \Psi$, where Ψ_T is known as the trial or guiding wave function, and substitute into Eq. 14 we obtain

$$-\frac{\partial f(\mathbf{R}, t)}{\partial t} = -\frac{1}{2}\nabla^2 f(\mathbf{R}, t) + \nabla \cdot [\mathbf{v}_D(\mathbf{R}) f(\mathbf{R}, t)] + (E_L(\mathbf{R}) - E_T) f(\mathbf{R}, t) , \qquad (15)$$

where $\mathbf{v}_D(\mathbf{R})$ is the $3N$-dimensional *drift velocity* defined by

$$\mathbf{v}_D(\mathbf{R}) = \nabla \ln |\Psi_T(\mathbf{R})| = \frac{\nabla \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} , \qquad (16)$$

and

$$E_L(\mathbf{R}) = \Psi_T^{-1} \left(-\frac{1}{2}\nabla^2 + V(\mathbf{R}) \right) \Psi_T , \qquad (17)$$

is the local energy. This formulation imposes the *fixed-node approximation* [32]. The nodal surface of a wave function is the surface on which it is zero and across which it changes sign. The nodal surface of Ψ is constrained to be the same as that of Ψ_T and therefore f can be interpreted as a probability distribution. The time evolution generates the distribution $f = \Psi_T \Psi$, where Ψ is the best (lowest energy) wave function with the same nodes as Φ_T. The problem of the poor statistical behaviour due to the divergences in the potential energy is also solved because the term $(V(\mathbf{R}) - E_S)$ in Eq. 14 has been replaced by $(E_L(\mathbf{R}) - E_T)$ which is much smoother. Indeed, if Ψ_T was an exact eigenstate then $(E_L(\mathbf{R}) - E_T)$ would be independent of position in configuration space. Although we cannot in practice find the exact Ψ_T it is possible to eliminate the divergences in the local energy by choosing a Ψ_T which has the correct cusp-like behaviour whenever two electrons or an electron and a nucleus are coincident [33]. The fixed-node approximation implies that we solve independently in different nodal pockets, and at first sight it appears that we have to solve the Schrödinger equation in every nodal pocket, which would be an impossible task in large systems. However, the *tiling theorem* for exact fermion ground states [34, 35] asserts that all nodal pockets are in fact equivalent and therefore one only need solve the Schrödinger equation in one of them. This theorem is intimately connected with the existence of a variational principle for the DMC ground state energy [35].

A DMC simulation proceeds as follows. First we pick an ensemble of a few hundred configurations chosen from the distribution $|\Psi_T|^2$ using VMC and the standard Metropolis algorithm. This ensemble is evolved according to the short-time approximation to the Green function of the importance-sampled imaginary-time Schrödinger equation (Eq. 15), which involves biased diffusion and addition/subtraction steps. The bias in the diffusion is caused by the importance sampling which directs the sampling towards parts of configuration space where $|\Psi_T|$ is large. After a period of equilibration the excited state contributions will have largely died out and the configurations start to trace out the probability distribution $f(\mathbf{R})/ \int f(\mathbf{R}) \, d\mathbf{R}$. We can then start to accumulate averages, in particular the DMC energy, E_D, which is given by

$$E_D = \frac{\int f(\mathbf{R}) E_L(\mathbf{R}) \, d\mathbf{R}}{\int f(\mathbf{R}) \, d\mathbf{R}} \approx \sum_i E_L(\mathbf{R}_i) . \qquad (18)$$

This energy expression would be exact if the nodal surface of Ψ_T was exact, and the fixed-node error is second order in the error in the nodal surface of Ψ_T (when a variational theorem exists [35]). The accuracy of the fixed node approximation can be tested on small systems and normally leads to very satisfactory results. The trial wave function limits the final accuracy that can be obtained because of the fixed-node approximation and it also controls the statistical efficiency of the algorithm. Like VMC, the

DMC algorithm satisfies a zero-variance principle, i.e., the variance of the energy goes to zero as the trial wave function goes to an exact eigenstate.
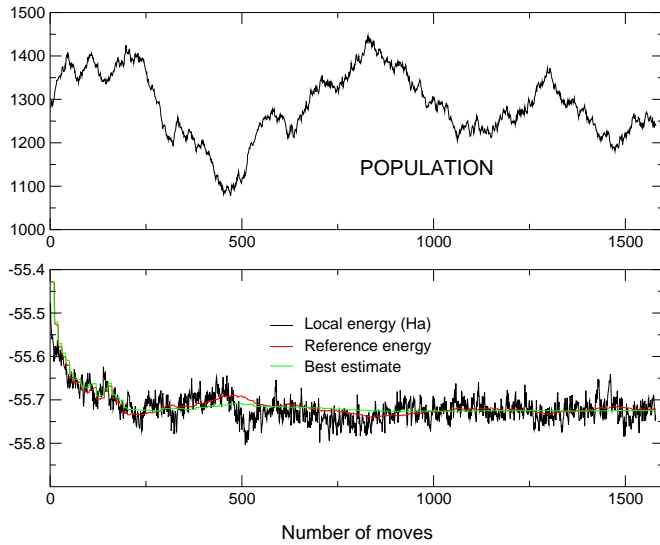

Figure 1: *DMC simulation of solid antiferromagnetic NiO. In the lower panel, the noisy black line is the local energy after each move, the green (or possibly grey) line is the current best estimate of the DMC energy, and the red (slightly different grey) line is E_T in Eqn. 15 which is varied to control the population of configurations through a feedback mechanism. As the simulation equilibrates the best estimate of the energy, initially equal to the VMC energy, decreases significantly then approaches a constant - the final DMC energy. The upper panel shows the variation in the population of the ensemble during the simulation as walkers are created or destroyed.*

3   Miscellaneous issues

In this section I will discuss some practical issues related to VMC and DMC.

3.1   More about trial wave functions

Single-determinant Slater-Jastrow wave functions often work very well in QMC calculations. For example, in the ground state of the carbon pseudo-atom, single determinant HF theory retrieves about 98.2% of the total energy. The residual 1.8% is the correlation energy, amounting to 2.7 eV in this case, which is very important for an accurate description of chemical bonding. A determinant of HF orbitals gives the lowest energy of all single-determinant wave functions and DFT orbitals are often very similar to them. These orbitals are not optimal when a Jastrow factor is included, but it turns out that the Jastrow factor does not change the detailed structure of the optimal orbitals very much, and the changes are well described by a fairly smooth change to the orbitals, which is conveniently included in the Jastrow factor.

How might we improve on the HF or DFT orbitals in the presence of the Jastrow factor? Direct optimization of the whole Slater-Jastrow wave function including both the orbitals and Jastrow factor has not been performed for large systems due to the computational cost. A promising technique [29, 30] is to optimize the potential that generates the orbitals rather than the orbitals themselves. Grossman and Mitas [36] have used a determinant of the *natural orbitals* which diagonalize the one-electron density matrix. It is not immediately clear why this should be expected to work in QMC, but the motivation appears to be that the convergence of configuration interaction expansions is improved by using natu-

ral orbitals instead of HF orbitals. However, the calculation of reasonably accurate natural orbitals is computationally demanding, and this approach is not attractive for large systems.

Another approach is to introduce 'backflow' correlations which were originally derived from a current conservation argument by Feynman and Cohen [37] to provide a picture of the excitations in liquid ^{4}He, although they are also helpful in fermionic systems. In the backflow trial function the electron coordinates $\mathbf{r}_i$ appearing in the Slater determinants of Eq. (8) are replaced by *quasiparticle coordinates*,

$$\bar{\mathbf{r}}_i = \mathbf{r}_i + \sum_{\substack{j=1 \\ (j \neq i)}}^{N} \eta(r_{ij})(\mathbf{r}_i - \mathbf{r}_j) \,, \tag{19}$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. The optimal function $\eta(r_{ij})$ may be determined variationally. Kwon, Ceperley, and Martin [38] found that the introduction of backflow significantly lowered the VMC and DMC energies of the three-dimensional uniform electron gas at high densities (it is expected to be even more significant in two dimensions). However, the use of backflow wave functions significantly increases the cost of QMC calculations because one can no longer move the electrons one at time, which is a significant saving in the standard algorithm. The basic scaling of the algorithm with backflow is thus N^4 rather than N^3.

In some cases it is necessary to use multi-determinant wave functions to preserve important symmetries of the true wave function. In other cases a single determinant may give the correct symmetry but a significantly better wave function can be obtained by using a linear combination of a few determinants. Multi-determinant wave functions have been used successfully in QMC studies of small molecular systems and even in periodic calculations such as the recent study of the neutral vacancy in diamond due to Hood *et al.* [39]. It is widely believed that a direct expansion in determinants (as used in configuration interaction calculations) converges very slowly because of the difficulty in describing the strong correlations which occur when electrons are close to one another. These correlations result in cusps in the wave function when two electrons are coincident, which are not well approximated by a finite sum of smooth functions [40]. However, this is not the whole story, and Prendergast *et al.* [41] have pointed out that the cusp is energetically less important, and that the slow convergence of determinant expansions has a lot to do with the description of medium-range correlations. In any case the number of determinants required to describe the wave function to some fixed accuracy increases exponentially with the system size. Practically speaking, this rules out the possibility of retrieving a significant extra fraction of the correlation energy with QMC in large systems via an expansion in determinants. Methods in which only local correlations are taken into account might be helpful, but overall an expansion in determinants is not a promising direction to pursue for making QMC trial wave functions for large systems, and the backflow technique, while costly, is more likely to be useful.

3.2   Basis set expansions : how to represent the orbitals?

The importance of using good quality single-particle orbitals in building up the Slater determinants in the trial wave function is clear. The determinant part accounts for by far the most significant fraction of the variational energy. However, the evaluation of the single-particle orbitals and their first and second derivatives can sometimes take up more than half of the total computer time, and consideration must therefore be given to obtaining accurate orbitals which can be evaluated rapidly at arbitrary points in space. It is not difficult to see that the most critical thing is to expand the single-particle orbitals in a basis set of *localized* functions. This ensures that beyond a certain system size, only a fixed number of the localized functions will give a significant contribution to a particular orbital at a particular point. The cost of evaluating the orbitals does not then increase rapidly with the size of the system. Note that 'localized basis functions' can (1) be strictly zero beyond a certain radius, or (2) can decrease monotonically and be pre-screened before the calculation starts, so that only those functions which could be significant in a

particular region are considered for evaluation.

An alternative procedure is to tabulate the orbitals and their derivatives on a grid, and this is feasible for small systems such as atoms, but for periodic solids or larger molecules the storage requirements quickly become enormous. This is an important consideration when using parallel computers as it is much more efficient to store the single-particle orbitals on every node. Historically a very large proportion of condensed matter electronic structure theorists have used plane-wave basis sets in their DFT calculations. However in QMC, plane-wave expansions are normally extremely inefficient because they are not localized in real space ; every basis function contributes at every point, and the required number of functions increases linearly with system size. Only if there is a short repeat length in the problem are plane waves not totally unreasonable. Note that this does not mean that all plane-wave DFT codes are useless for generating trial wave functions for CASINO; a post-processing utility can be used to reexpand a function expanded in plane-waves in another localized basis before the wave function is input into CASINO. The usual thing here is to use some form of localized spline functions on a grid such as those used by the Lawrence Livermore group [15], or the 'blip' functions used by Mike Gillan's group [42] and implemented in CASINO by Dario Alfè [43].

Another pretty good way to do this is to expand the orbitals in a basis of Gaussian-type functions. These are localized, quick to evaluate, and are available from a wide-range of sophisticated software packages. Such a large expertise has been built up within the quantum chemistry community with Gaussians that there is a significant resistance to using any other type of basis. A great many Gaussian-based packages have been developed by quantum chemists for treating molecules. The most well-known of these are the various versions of the GAUSSIAN package [44]. In addition to the regular single determinant methods, these codes include various techniques involving multi-determinant correlated wave functions (although sadly, not QMC!). This makes them very flexible tools for developing accurate molecular trial wave functions. For Gaussian basis sets with periodic boundary conditions, the CRYSTAL program [45] can perform all-electron or pseudopotential Hartree-Fock and DFT calculations both for molecules and for systems with periodic boundary conditions in one, two or three dimensions, which makes it very useful as a tool for generating trial functions for CASINO.

3.3 Pseudopotentials

Pseudopotentials or effective core potentials are commonly used in electronic structure calculations to remove the inert core electrons from the problem and to improve the computational efficiency. Although QMC scales very favourably with system size it has been estimated that the scaling of all-electron calculations with the atomic number Z is approximately $Z^{5.5-6.5}$ which effectively rules out applications to atoms with Z greater than about ten. The use of a pseudopotential serves to reduce the effective value of Z and although errors are inevitably introduced, the gain in computational efficiency is sufficient to make applications to heavy atoms feasible.

Accurate pseudopotentials for single-particle theories such as DFT or Hartree-Fock theory are well developed, but pseudopotentials for correlated wave function techniques such as QMC present additional challenges. The presence of core electrons causes two related problems. The first is that the shorter length scale variations in the wave function near a nucleus of large Z require the use of a small time step. This problem can be significantly reduced (in VMC at least) by the use of acceleration schemes [46, 47]. The second problem is that the fluctuations in the local energy tend to be large near the nucleus because both the kinetic and potential energies are large.

The central idea of pseudopotential theory is to create an effective potential which reproduces the effects of both the nucleus and the core electrons on the valence electrons. This is done separately for each of the different angular momentum states, so the pseudopotential contains angular momentum projectors and is therefore a non-local operator.

It is convenient to divide the pseudopotential for each atom into a local part $V_{\text{loc}}^{\text{ps}}(r)$ common to all angular momenta and a correction, $V_{\text{nl},l}^{\text{ps}}(r)$, for each angular momentum l. The electron-ion potential energy term in the full many-electron Hamiltonian of the atom then takes the form

$$V_{\text{loc}} + \hat{V}_{\text{nl}} = \sum_i V_{\text{loc}}^{\text{ps}}(r_i) + \sum_i \hat{V}_{\text{nl},i}^{\text{ps}} , \tag{20}$$

where $\hat{V}_{\text{nl},i}^{\text{ps}}$ is a non-local operator which acts on an arbitrary function $g(\mathbf{r}_i)$ as follows

$$\hat{V}_{\text{nl},i}^{\text{ps}} g(\mathbf{r}_i) = \sum_l V_{\text{nl},l}^{\text{ps}}(r_i) \sum_{m=-l}^{l} Y_{lm}(\Omega_{\mathbf{r}_i}) \int Y_{lm}^*(\Omega_{\mathbf{r}_i'}) g(\mathbf{r}_i') \, d\Omega_i' , \tag{21}$$

where the angular integration is over the sphere passing through the $\mathbf{r}_i$. This expression can be simplified by choosing the z-axis along $\mathbf{r}_i$, noting that $Y_{lm}(0,0) = 0$ for $m \neq 0$, and using the definition of the spherical harmonics to give

$$\hat{V}_{\text{nl},i}^{\text{ps}} g(\mathbf{r}_i) = \sum_l V_{\text{nl},l}^{\text{ps}}(r_i) \frac{2l+1}{4\pi} \int P_l \left[\cos(\theta_i') \right] g(\mathbf{r}_i') \, d\Omega_i' , \tag{22}$$

where P_l denotes a Legendre polynomial.

It is not currently possible to construct pseudopotentials for heavy atoms entirely within a QMC framework, although progress in this direction was made by Acioli and Ceperley [48]. It is therefore currently necessary to use pseudopotentials generated within some other framework. Possible schemes include HF theory and local density functional theory (DFT), where there is a great deal of experience in generating accurate pseudopotentials. There is evidence to show that HF pseudopotentials give better results within QMC calculations than DFT ones, although DFT ones work quite well in many cases. The problem with DFT pseudopotentials appears to be that they already include a (local) description of correlation which is quite different from the QMC description. HF theory, on the other hand, does not contain any effects of correlation. The QMC calculation puts back the valence-valence correlations but neglects core-core correlations (which have only an indirect and small effect on the valence electrons) and core-valence correlations. Core-valence correlations are significant when the core is highly polarizable, such as in alkali-metal atoms. The core-valence correlations may be approximately included by using a 'core polarization potential' (CPP) which represents the polarization of the core due to the instantaneous positions of the surrounding electrons and ions. Another issue is that relativistic effects are important for heavy elements. It is still, however, possible to use a QMC method for solving the Schrödinger equation with the scalar relativistic effects obtained within the Dirac formalism incorporated within the pseudopotentials. The combination of Dirac HF pseudopotentials and CPPs appears to work well in many QMC calculations. CPPs have been generated for a wide range of elements (see, e.g., Ref. [49]).

Many HF pseudopotentials are available in the literature, mostly in the form of sets of parameters for fits to Gaussian basis sets. Unfortunately many of them diverge at the origin, which can lead to significant time step errors in DMC calculations [50]. We have concluded that none of the available sets are ideal for QMC calculations, and that would be helpful if we generated an on-line periodic table of smooth non-divergent HF pseudopotentials (with relativistic corrections). This project is well on its way to completion.

Recent developments have been made that allow the use of all-electron QMC calculations for much heavier atoms than has previously been the case, based on reducing the fluctuations in the local energy near the nucleus. These developments will be described in the following section.

4 Recent developments

In this section I will describe some recent improvements to the basic algorithms that improve the ability of QMC to (1) treat heavier atoms with all-electron calculations, and (2) to treat larger systems by improving the scaling behaviour. Both these features have recently been implemented in the CASINO code.

4.1 All-electron QMC calculations for heavier atoms

At a nucleus the exact wave function has a cusp so that the divergence in the potential energy is cancelled by an equal and opposite divergence in the kinetic energy. If this cusp is represented accurately in the QMC trial wave function therefore, then the fluctuations in the local energy referred to in the previous section will be greatly reduced. If the wave function is formed from one or more determinants of single-particle orbitals expanded, for example, in a Gaussian basis set, then there can be no cusp in the wave function since Gaussians have zero gradient at $r = 0$. The local energy thus diverges at the nucleus. In practice one finds that the local energy has wild oscillations close to the nucleus which can lead to numerical instabilities in DMC calculations. To solve this problem we can make small corrections to the single particle orbitals close to the nuclei which impose the correct cusp behaviour. Such corrections need to be applied at each nucleus for every orbital which is larger than a given tolerance at that nucleus.

It is likely that a number of other researchers have developed such schemes, but within the literature we are only aware of the scheme developed by Manten and Lüchow [51], which is rather different from ours [52]. Our scheme is based on the idea of making the one-electron part of the local energy for each orbital, $\frac{\hat{H}_{oe}\phi}{\phi}$, finite at the nucleus. $\hat{H}_{oe}$ is given by

$$\hat{H}_{oe} = -\frac{1}{2}\nabla^2 - \frac{Z}{r}, \tag{23}$$

where r is the distance to the nucleus of charge Z. The scheme need only be applied to the s-component of orbitals centred at the nuclear position in question. Inside some radius r_c we replace the orbital expanded in Gaussians by $\phi = sgn[\psi(r = 0)]\exp[p]$, where $sgn[\psi(r = 0)]$ denotes the sign of the Gaussian orbital at $r = 0$ and p is a polynomial in r. Therefore $\ln|\phi| = p$ and the local energy is given by

$$E_L = \frac{\hat{H}_{oe}\phi}{\phi} = -\frac{p'}{r} - \frac{p''}{2} - \frac{p'^2}{2} - \frac{Z}{r}. \tag{24}$$

We impose five constraints, that $p(r_c)$, $p'(r_c)$, and $p''(r_c)$ are continuous, that $p'(0) = -Z$ (to satisfy the cusp condition), and that $E_L(0)$ is chosen to minimize the maximum of the square of the deviation of $E_L(r)$ from an 'ideal curve' of local energy versus radius.

To see the cusp corrections in action, let us first look at a hydrogen atom where the basis set has been made to model the cusp very closely by using very sharp Gaussians with high exponents. Visually (top left in Figure 2) the fact that the orbital does not obey the cusp condition is not immediately apparent. If we zoom in on the region close to the nucleus (top right) we see the problem: the black line is the orbital expanded in Gaussians, the red (or light grey if you've scrimped on the colour printing) line is the cusp-corrected orbital. The effect on the gradient and local energy is clearly significant. This scheme has been implemented within the CASINO code both for finite and for periodic systems, and produces a significant reduction in the computer time required to achieve a specified error bar, as one can appreciate from Figure 3.

In order to understand our capability to do all-electron DMC calculations for heavier atoms, and to understand how the necessary computer time scales with atomic number, we are currently carrying out a series of calculations on the various noble gas atoms [53]. So far it has proved to be perfectly possible to produce converged DMC energies with acceptably small error bars for atoms up to $Z=36$. It seems that even xenon ($Z=54$) can be done (although this is still running as I write this).
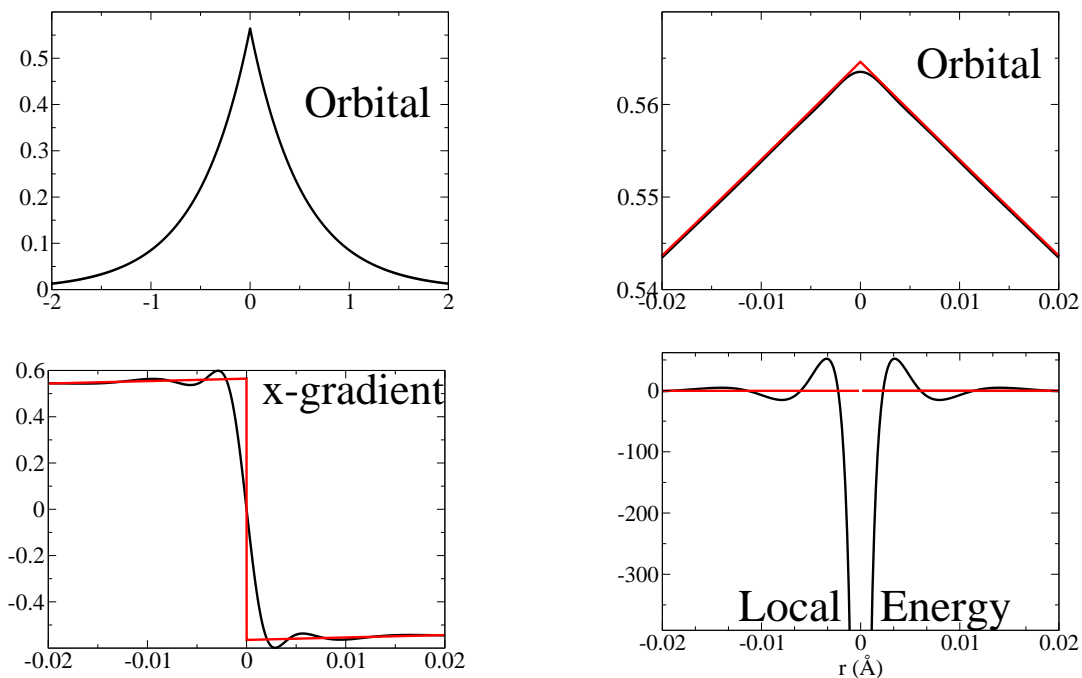
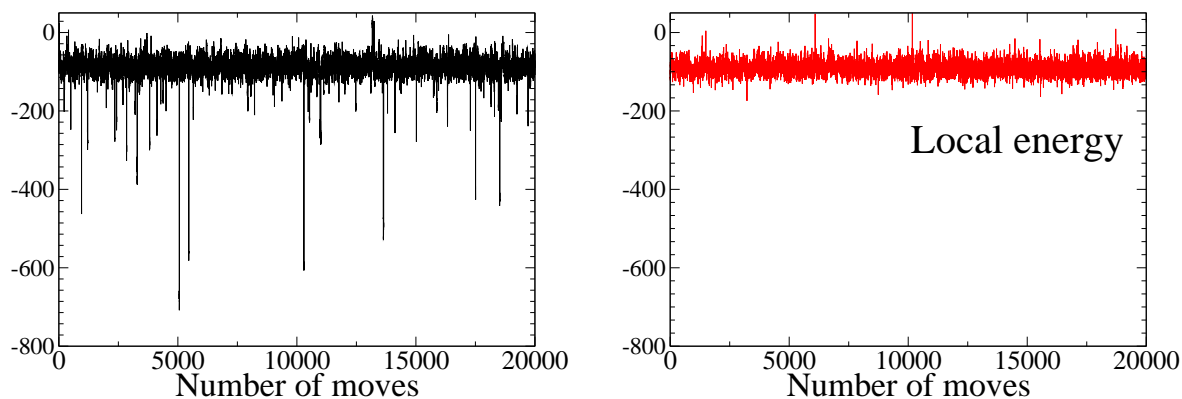Figure 2: *Cusp corrections in the hydrogen atom.*

Figure 3: *Local energy as a function of move number in a VMC calculation for a carbon monoxide molecule with a standard reasonably good Gaussian basis set. The cusp corrections are imposed only in the figure on the right. The reduction in the local energy fluctuations with the new scheme is clearly apparent.*

4.2 Improved scaling algorithms

Let us now consider in more detail how QMC calculations scale with system size, and what one might do in order to improve the scaling behaviour. QMC methods are stochastic and therefore yield mean values with an associated statistical error bar. We might want to calculate the energy of some system and compare it with the energy of a different arrangement of the atoms. The desired result might be a defect formation energy, an energy barrier, or an excitation energy. These are evidently energy differences which become *independent of the system size* when the system is large enough. To perform such a calculation we therefore require an error bar ΔE on the energy of the system which is independent of system size, a feature denoted here by $\Delta E = \mathcal{O}(1)$. There are other quantities such as cohesive energies, lattice constants, and elastic constants, for example, in which both energy and error bar may be defined *per atom* or *per formula unit*, in which case the error bar on the whole system is allowed to scale linearly with system size, i.e., $\Delta E = \mathcal{O}(N)$.

How does the computational cost C of a QMC calculation, yielding an error $\Delta E = \mathcal{O}(1)$, scale with the system size, measured by the number of electrons N? The result for the standard algorithm is $C = AN^3 + \epsilon N^4$, where ϵ is very small [4]. In current solid simulations $N \leq 2000$, and the first term in this expression dominates, giving an N^3 scaling for the standard algorithm: double the system size and the cost goes up eightfold. What is the best scaling we could possibly achieve? As is well known, the best possible scaling for conventional (non-stochastic) single-particle methods such as DFT is $\mathcal{O}(N)$ [54]. A considerable effort has been made over the previous decade to design DFT codes which (a) scale linearly with system size, (b) are faster than the regular cubic scaling algorithm for reasonable system sizes, and (c) are *as accurate* as codes using the regular algorithm, with the latter two problems being the most difficult. In wave function-based QMC, these additional problems do not occur ; with the improved scaling algorithms described here the speed benefit is immediate and there is essentially no loss of accuracy. However, for the scaling one cannot do better than $\mathcal{O}(N^2)$ in general, unless the desired quantity is expressible as an energy per atom. Why is this so? One still has the 'near-sightedness' in the many-body problem which is exploited in linear scaling DFT algorithms, but the difference is the stochastic nature of QMC. The statistical noise in the energy adds incoherently over the particles, so the variance in the mean energy increases as N (and thus the error bar as $\sqrt{N}$). Since the variance is inversely proportional to the number of statistically independent configurations in the calculation, we see that to obtain $\Delta E = \mathcal{O}(1)$ we must therefore evaluate the energy of $\mathcal{O}(N)$ configurations, each of which costs $\mathcal{O}(N)$ operations. This accounts for the 'extra' power of N in the cost of a QMC calculation. However, $\mathcal{O}(N^2)$ scaling is still a vast improvement over $\mathcal{O}(N^3)$ scaling when N can be of the order of a few thousand, and clearly the scaling is improved further for properties which can be expressed in terms of energies per atom. The primary task is thus to reduce the AN^3 term to AN^2. The operations which make up this term are (1) evaluation of the orbitals in the Slater determinants, (2) evaluation of the Jastrow factor, and (3) evaluation of Coulomb interactions between particles.

The first of these operations is by far the most costly. As in $\mathcal{O}(N)$-DFT methods, the solution is to use *localized orbitals* instead of the delocalized single-particle orbitals that arise naturally from standard DFT calculations. The number of such orbitals contributing at a point in space is *independent of N* which leads to the required improvement in scaling. In fact, our collaborators at Lawrence Livermore have recently shown that this approach is extremely effective in molecular QMC calculations using maximally-localized Wannier function orbitals in the CASINO code [15]. We are currently working in collaboration with them to develop these methods and are testing the applicability of the method in various finite and periodic systems. One might expect this method to be more efficient for insulators than metals because the orbitals can be made more localized, but in practice a large efficiency gain in metals also seems to be achievable.

For the Jastrow factor all that is required to achieve the improved scaling is that it be truncated at some distance which is independent of system size. Because the correlations are essentially local it is natural to truncate the Jastrow factor at the radius of the exchange-correlation hole. Of course, truncating the Jastrow factor does not affect the final answer obtained within DMC because it leaves the nodal surface of the wave function unchanged, although if it is truncated at too short a distance the statistical noise increases. The scaling of the Coulomb interactions can be improved using an accurate scheme which exploits the fact that correlation is short-ranged to replace the long-range part by its Hartree contribution (in the style of the Modified Periodic Coulomb (MPC) interaction [23]).

For extremely large systems, the notionally ϵN^4 term might begin to be significant. This arises from N updates of the matrix of cofactors of the inverse Slater matrix (required when computing the ratio of new to old determinants after each electron move), each of which takes a time proportional to N^2, plus the extra factor of N from the statistical noise. In CASINO this operation has been significantly streamlined through the use of sparse matrix techniques and we have not yet found a system where it contributes substantially to the overall CPU time.

Method	Si	Ge	C	BN
LSDA	5.28^a	4.59^a	8.61^a	15.07^e
VMC	$4.38(4)^c$	$3.80(2)^b$	$7.27(7)^d$	$12.85(9)^e$
	$4.82(7)^d$		$7.36(1)^f$	
	$4.48(1)^g$			
DMC	$4.63(2)^g$	$3.85(2)^b$	$7.346(6)^f$	
Exp.	$4.62(8)^a$	3.85^a	7.37^a	12.9^h

Table 1: Cohesive energies of tetrahedrally bonded semiconductors calculated within the LSDA, VMC and DMC methods and compared with experimental values. The energies for Si, Ge, and C are quoted in eV per atom while those for BN are in eV per two atoms. References : *a.* Farid and Needs [55], and references therein. *b.* Rajagopal *et al.* [19], *d.* Fahy, Wang, and Louie [56]. Zero-point energy corrections of 0.18 eV for C and 0.06 eV for Si have been added to the published values for consistency with the other data in the table. *e.* Malatesta, Fahy, and Bachelet [57], *f.* Hood *et al.* [39], *g.* Leung *et al.* [58], *h.* Estimated by Knittle *et al.* [59] from experimental results on hexagonal BN.

Taken together the localization algorithms described above should speed up continuum fermion QMC calculations significantly for large systems, but we can view it in another light - as an *embedding* algorithm in which a QMC calculation could be embedded within a DFT one. The idea is to use the higher accuracy of QMC where it is most needed, such as around a defect site or in the neighbourhood of a molecule attached to a solid surface. Developments along the lines of those described here might allow such QMC/DFT embedding calculations to be performed for the first time. This is quite simple in VMC although a practical DMC embedding scheme would be more difficult.

5   Applications

Time and space preclude me from presenting a long list of applications, but here are some cohesive energies of tetrahedrally-bonded semiconductors to be going on with.

6   The CASINO code

CASINO [1, 2] is a program package developed in Cambridge by the group of Richard Needs for performing quantum Monte Carlo electronic structure calculations for finite and periodic systems. The philosophy behind it involves generality, speed, portability and ease-of-use. Generality in this sense means that one ought to be able to create a trial wave function for any system, expanded in any of a variety of different basis sets, and use it as input to a CASINO QMC calculation. Clearly the wave functions must be generated by an external electronic structure program, and this must in the past have been persuaded to write out the wave function in a format that CASINO understands, either all by itself, or through the transformation of its standard output using a separate CASINO utility. This is one of the main reasons that producing a QMC code is somewhat labour intensive. Maintaining these interfaces as codes evolve, and persuading their owners that this is a good idea in the first place, is a difficult and sometimes frustrating task. It is nevertheless part of the philosophy that CASINO should support a reasonably wide range of the most popular electronic structure codes, and at the present time this list includes CRYSTAL95/98/03 [45], GAUSSIAN94/98/03 [44], CASTEP [60], ABINIT [61], PWSCF [62],

ONETEP [63], TURBOMOLE [64] and JEEP.

The most important current capabilities of CASINO are as follows :

- Variational Monte Carlo (including variance minimization of wave functions).

- Diffusion Monte Carlo (branching DMC and pure DMC).

- Use of Slater-Jastrow wave functions where the Slater part may consist of multiple determinants of spin orbitals.

- Trial wave functions expanded in plane waves or Gaussian basis sets (s, sp, p, d, f or g functions centred on atoms or elsewhere) produced using DFT, HF, or various multideterminant methods).

- Trial wave functions expanded in various kind of spline functions generated by post-processing plane-wave DFT solutions.

- Numerical atomic calculations with the orbitals and their derivatives interpolated from a radial grid.

- 2D/3D electron phases in fluid or crystal wave functions, with arbitrary cell shape/spin polarization/density (including excited state capability).

- 2D/3D electron-hole phases with fluid/crystal/pairing wave functions with arbitrary cell shape/spin polarization/density (including excited state capability). Variable electron-hole mass ratio. 2D layer separation between holes and electrons possible.

- Improved 'linear scaling' mode through use of Wannier orbitals and localized basis functions.

- Computation of excitation energies corresponding to either promotion or addition/subtraction of electrons.

- Computation of distribution functions such as the pair correlation function and density matrices (electron and electron-hole systems only for the moment).

- Calculation of electron-electron interactions using either Ewald and/or our 'modified periodic Coulomb interaction' [23] which is faster and has smaller Coulomb finite size effects.

- Written in strict compliance with the Fortran90 standard using modern software design techniques.

- Parallelized using MPI—tested in parallel on Hitachi SR2201, Cray T3E, SGI Origin 2000, SGI Altix, IBM SP3, Fujitsu Primepower, Alpha servers and SunFire Galaxy, Linux PC clusters. Also set up for workstation use on DEC Alphas, SGI Octane and O2, Linux PC with various compilers. MPI libraries not required on single processor machines.

- Flexible input with full documentation and interactive help system.

It is worth sketching a brief history of the CASINO code. Its development was inspired by a Fortran77 development code (known simply as 'the QMC code') written in the early 1990s in Cambridge by Richard Needs and Guna Rajagopal, assisted by many helpful discussions with Matthew Foulkes. This was later extended by Andrew Williamson up to 1995 and then by Paul Kent and Mike Towler up to 1998. Various different versions of this were able to treat fcc solids, single atoms and the homogeneous electron gas. By the late 1990s it was clear that a modern general code capable of treating arbitrary systems (e.g. at least atoms, molecules, polymers, slabs, crystals, and electron phases) was required, not only for the use of the Cambridge QMC group, but for public distribution. At that time, a user-friendly general publically available code did not exist, at least for periodic systems, and it was felt to be a good thing to create one to allow other researchers to join in the fun. So beginning in 1999 a new Fortran90 code, CASINO, was gradually developed in the group of Richard Needs largely by Mike Towler, considerably assisted from 2002 by Neil Drummond (some routines from the old code were retained, translated and reused, although most were gradually replaced). Various additional contributions have been made over the years by Andrew Porter, Randy Hood, Dario Alfè, Andrew Williamson, Gavin Brown, Chris Pickard, Rene Gaudoin, Ben Wood and others. The main aims of the new code were generality, speed, ease-of-use and

transferability over a wide range of computational hardware. It is hoped that these objectives have been largely attained, but the code continues to be actively developed.

6.1 Availability of CASINO

We are making CASINO available to a number of groups now. However the technology of QMC can hardly be said to be mature, and largely for this reason it has been decided not to distribute the code under a GNU public licence in the admirable way that, for example, the ABINIT [61] project is run. We do not ask for any payment for academic use of the code, but we ask users to sign an agreement concerning its use. The practical upshot of this is that users may not redistribute the code, they may not incorporate any part of it into any other program system, nor may they modify it in any way whatsoever without prior agreement of the Cambridge group.

The CASINO distribution comes with extensive documentation and examples but clearly it remains a research code and learning how to use it is a significant task. This is particularly the case if the user does not have relevant experience such as familiarity with VMC and DMC calculations and knowledge of band structure methods for solids and DFT/molecular orbital theory. Available manpower is sufficiently limited that we find supporting users can take a large amount of our time and so we have to limit the number of such groups that we can work with directly. We find that most people need quite a lot of help and the project turns into a collaboration, but of course we cannot enter into too many projects of this type as our time is limited. We do have people visiting Cambridge to learn about the codes and how to do calculations, and this seems to work well.

7 Discussion and conclusions

In this article I have provided a very brief review of the VMC and DMC methods, highlighted some recent innovations, and introduced our implementation in the CASINO code. I have tried to make the case that QMC is a useful addition to the toolbox of the computational electronic structure theorist.

That said, it's important to be honest about the problems with QMC. If it is such a wonderful technique, why doesn't everyone use it? As an example of why not, see the web site of David Ceperley's group `archive.ncsa.uiuc.edu/Science/CMP/topten/topten.html` where there is a well-known page entitled 'Top Ten List of reasons why quantum Monte Carlo is not generally used in chemistry'. The list contains the following twelve reasons.

1. We need forces, dummy!
2. Try getting O_2 to bind at the variational level.
3. How many graduate students lives have been lost optimizing wavefunctions?
4. It is hard to get 0.01 eV accuracy by throwing dice.
5. Most chemical problems have more than 50 electrons.
6. Who thought LDA or HF pseudopotentials would be any good?
7. How many spectra have you seen computed by QMC?
8. QMC is only exact for energies.
9. Multiple determinants. We can't live with them, we can't live without them.
10. After all, electrons are fermions.
11. Electrons move.
12. QMC isn't included in Gaussian 90. Who programs anyway?

This apparently first appeared on the web in 1996, so it might be worth examining whether any progress has been made in these areas in the last seven years. Here is an attempt at a commentary :

1. *We need forces, dummy!*
Of course it is true that for QMC to be considered a general method, one ought to be able to calculate forces (i.e., derivatives of the energy with respect to nuclear displacements) in order to optimize structures or to perform dynamical simulations. In fact almost all QMC calculations up to the present time have been done within the Born-Oppenheimer approximation. The nuclear positions are thus fixed during the calculation and the wave function depends parametrically on the nuclear coordinates. The fixed nuclear positions are normally taken from geometry optimizations done with alternative methods such as DFT, on the principle that DFT is more reliable for geometries than for total energy differences. Calculating forces using a stochastic algorithm is a difficult thing to do. A straightforward application of the Hellmann-Feynman theorem (where the force is given by the gradient of the potential energy surface with respect to nuclear positions) leads to estimators with a very large variance. While a convincing general algorithm has yet to be demonstrated for QMC calculations, some progress has been made. The literature contains a variety of interesting contributions to this problem, which may be roughly classified into three groups :

- finite differences using correlated sampling techniques which take advantages of correlations between statistical samples to reduce the overall statistical error, e.g. [65].

- methods which use the standard Hellmann-Feynman formula. This is not normally useful in QMC as the Hellmann-Feynman estimator (at least with bare nuclei) has an infinite variance associated with it. One way of getting round this is to use some renormalized expression for the force, see e.g. [66]. Furthermore, the Hellmann-Feynman expression does not give the exact derivative of the DMC energy if the nodal surface is not exact and depends on the nuclear positions. This is due to an additional nodal term rising from the action of the kinetic-energy operator on the discontinuity in the derivative of the wave function at inexact nodal surfaces [67].

- direct attempts to calculate analytic derivatives, see e.g. [68].

Currently one cannot calculate accurate forces for large systems with QMC. However, some progress has been made and one has to be reasonably optimistic that a better general method for calculating forces will be devised in the near future. Also worth mentioning is the interesting recent attempt to design a coupled QMC and DFT molecular dyanmics algorithm by Grossman and Mitas [69].

2. *Try getting O_2 to bind at the variational level.*
I haven't managed to do so either. However, one can choose not to be overly concerned with QMC calculations done at the variational level. The binding energy of the oxygen molecule comes out very accurately in DMC [70, 71].

3. *How many graduate students lives have been lost optimizing wavefunctions?*
To give a feel for the time scale involved in optimizing wave functions, I can tell you about the weekend recently when I added the entire G2-1 set [72, 73] to the examples included with the CASINO distribution. This is a standard set of 55 molecules with various experimentally well-characterized properties intended for benchmarking of different quantum chemistry methods see e.g. Ref.[74]. Grossman has published the results of DMC calculations of these molecules using pseudopotentials [70], while we are doing the same with all-electron calculations [71]. It took a little over three days using only a few single-processor workstations to create all 55 sets of example files from scratch including optimizing the Jastrow factors for each molecule. While if one concentrated very hard on each individual case one might be able to pull a little more energy out of a VMC simulation, the optimized Jastrow factors are all perfectly good enough to be used as input to DMC simulations. I suggest that the process is sufficiently automated these days that graduate students are better employed elsewhere; certainly we have not suffered any fatalities here in Cambridge.

4. *It is hard to get 0.01 eV accuracy by throwing dice.*
With modern computers and efficient computer codes, there are a great many systems where one can get

sufficient accuracy in a reasonable time. Obviously this becomes increasingly difficult for heavier atoms and large systems, but as discussed previously, satisfying the electron-nuclear cusp condition accurately in all-electron calculations or using pseudopotentials helps a lot.

5. *Most chemical problems have more than 50 electrons.*
QMC calculations for several thousand electrons (per simulation cell) have been published, and this number will only increase with the new improved scaling techniques currently being introduced, and with the increasing power of available computational hardware.

6. *Who thought LDA or HF pseudopotentials would be any good?*
Very accurate QMC results using LDA/HF pseudopotentials have been published. Such pseudopotentials seem to work best for *sp*-bonded systems, and it is not clear that particularly good results can be obtained in systems containing, for example, transition elements. As previously explained, we are developing an on-line periodic table containing a new set of smooth non-divergent Dirac-Hartree-Fock pseudopotentials expected to be particularly satisfactory for QMC calculations.

7. *How many spectra have you seen computed by QMC?*
Almost none. The calculation of excited state properties is a difficult problem for most other methods ; one can sometimes compute the energies of individual excited states pretty accurately with QMC, but the calculation of complete spectra is too difficult for the moment. It is very difficult in general to treat frequency-dependent properties.

8. *QMC is only exact for energies.*
The problem here is to find unbiased expectation values of operators that do not commute with the Hamiltonian. In standard DMC, the time evolution generates the mixed distribution $f(\mathbf{R}) = \Psi_T(\mathbf{R})\Psi_0(\mathbf{R})$, where $\Psi_0(\mathbf{R})$ is the best (lowest energy) wave function with the same nodes as the trial function $\Psi_T(\mathbf{R})$. To obtain the exact energy with respect to the distribution $\Psi_0^2(\mathbf{R})$ it is necessary to project by means of a mixed estimator, which in practice just involves summing the local energies over the positions generated in the DMC simulation:

$$E_{DMC} = \frac{\int \Psi_0(\mathbf{R})\hat{H}\Psi_T(\mathbf{R})\,d\mathbf{R}}{\int \Psi_0(\mathbf{R})\Psi_T(\mathbf{R})\,d\mathbf{R}} \approx \frac{1}{M}\sum_{i=1}^{M} E_L(\mathbf{R}_m) \tag{25}$$

If another operator $\hat{O}$ commutes with the Hamiltonian, then any eigenstate of $\hat{H}$ is (or may be chosen to be) an eigenstate of $\hat{O}$ and the same mixed estimator can be used to work out its expectation value. If $\hat{O}$ and $\hat{H}$ do not commute, then the mixed estimator does not give the ground state eigenvalue of $\hat{O}$. It is a good approximation to it however, and normally lies somewhere between the variational estimate and the exact value.

In general, the best thing to do to improve the expectation value of an operator that does not commute with the Hamiltonian is to improve the trial wave function. One may additionally use the extrapolation method [75] which approximates the true estimator as a linear combination of the mixed and variational ones, but it is not aesthetically pleasing and may not be particularly accurate in all cases. Note however that if the operator is just a simple local function of $\mathbf{R}$ and not, for example, a differential operator, then various methods can be used to obtain the ground state eigenvalue exactly. Among them are the future-walking algorithm (clearly explained in Ref. [3]) based on estimating the ratio $\frac{\Psi_0}{\Psi_T}$ using the asymptotic offspring of the DMC branching term, and the related time-correlation method [76]. Both of these methods are somewhat delicate however, and lead to signal-to-noise ratios that decay to zero at large time. Neither are likely to prove useful for large systems. More recently introduced bilinear methods [77] show promise but have not yet been fully developed into a practical general scheme. Probably the best scheme for the moment is the 'reptation QMC' method of Baroni and Moroni [78].

9. *Multiple determinants. We can't live with them, we can't live without them.*

As we have seen, multiple determinants are occasionally required to preserve important wave function symmetries, but more often one sees that using a linear combination of a few determinants can give a significantly lower DMC energy. Clearly this is because the nodal surface of this wave function is in some sense closer to the true one than that of the single determinant trial function. This is a problem not least because most electronic structure codes are not capable of generating optimized multiple determinant wave functions using e.g. the multi-configuration self-consistent field (MCSCF) method. Of the eight codes in the list supported by CASINO, only two (GAUSSIAN and TURBOMOLE) support multideterminant calculations, and to my knowledge there is no periodic code that does so. In an earlier section we observed that the number of determinants likely to be required in large systems effectively rules this out as a general solution. Orbital optimization and the backflow method are likely to prove more useful.

10. *After all, electrons are fermions.*

The solution to the fermion sign problem, which in this context means finding a way to bypass the fixed-node approximation in DMC, is one of the most interesting and important problems in computational physics. A solution to this would allow one to provide exact numerical solutions to the many-electron Schrödinger equation for fermionic systems (such as those containing electrons). It might seem strange to say that a stochastic or Monte Carlo method could solve such a problem exactly, since there is always a statistical error associated with the result, but we rely on the central limit theorem to provide an estimate of the probability that the exact results lies within a given interval. If we need to have the error to be less than this amount, this implies that the computer time will scale proportional to one over the square of the error.

We have seen that in DMC random walks serve to filter out the higher energy modes of some initial distribution, so that for very large imaginary time, the probability distribution of the random walkers is given by the ground state. The problem is that the ground state in question is actually the *bosonic* ground state which has the same sign everywhere in configuration space. It is only by imposing an additional boundary condition for the electronic Hamiltonian in the form of the fixed nodal surface that we are able to prevent collapse to the bosonic state, and maintain the positive and negative regions which are required for an antisymmetric wave function. The simplest methods which allow diffusing configurations to cross nodes and change the nodal surface, such as release-node Monte Carlo [79], are very inefficient because the 'signal' quickly becomes overwhelmed by the 'noise'. Methods which overcome this problem, such as fermion Monte Carlo [7, 8, 9] have been shown to be apparently stable for small test systems, but suffer from scaling problems with system size. It should never be forgotten that a genuine solution to the fermion sign problem in the context of continuum Monte Carlo methods must not only give the exact answer, but must also scale as some low-order polynomial with system size, and work generally rather than just for certain special cases. No proposed solution (and there are many in the literature) has all these characteristics. Recall that in some sense we already know how to solve the Schrödinger equation exactly for any system. One can in principle just do, for example, a full CI expansion in a complete basis set. The only trouble is it would take forever to compute and so this is not a viable solution.

That said, the fixed-node approximation is a good one, and one can normally recover well in excess of 95% of the correlation energy from a DMC calculation. For most problems this turns out to be accurate enough. Another approach - which is not a genuine solution to the sign problem but which might be the best practical thing to do - would be to optimize the nodal surface in some way, e.g., by optimizing the orbitals in the Slater part of the trial wave function (recalling that in some cases more than one determinant might be required for symmetry purposes). This ought certainly to be feasible and work is under way to develop an appropriate algorithm and implement such a facility in our CASINO code.

11. *Electrons move.*

Quite so.

12. *QMC isn't included in Gaussian 90. Who programs anyway?*

Nor in Gaussian94, '98 or '03. Neil Drummond and I and some of our collaborators like programming.

So, quite a lot of progress has been made in the theory and practical implementation of quantum Monte Carlo over the last few years, but certainly many interesting problems remain to be solved. For its most important purpose of calculating highly accurate total energies, the method works well and currently has no serious competitors for medium-sized and large systems. The CASINO program [1, 2] has been designed to allow researchers to explore the potential of QMC in arbitrary molecules, polymers, slabs and crystalline solids and in various electron and electron-hole phases.

Acknowledgements

References

[1] R.J. Needs, M.D. Towler, N.D. Drummond, and P.R.C. Kent, CASINO version 1.7 User Manual, University of Cambridge, Cambridge (2003).

[2] CASINO web site : www.tcm.phy.cam.ac.uk/~mdt26/casino.html

[3] B.L. Hammond, W.A. Lester, Jr., and P.J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, Singapore, 1994).

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

[5] D.M. Ceperley and E.L. Pollock, *Phys. Rev. Lett* **56**, 351 (1986).

[6] C. Pierleoni, D.M. Ceperley, B. Bernu, and W.R. Magro, *Phys. Rev. Lett.* **73**, 2145 (1994).

[7] M.H. Kalos and K.E. Schmidt, *J. Stat. Phys.* **89**, 425 (1997).

[8] F. Pederiva and M.H. Kalos, *Comp. Phys. Commun.* **122**, 445 (1999).

[9] M.H. Kalos and F. Pederiva, *Phys. Rev. Lett.* **85**, 3547 (2000).

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

[11] Y.S.M. Wu, A. Kuppermann, and J.B. Anderson, *Phys. Chem. Chem. Phys.* **1**, 929 (1999).

[12] G. Ortiz, M. Harris, and P. Ballone, Phys. Rev. Lett. **82**, 5317 (1999).

[13] D. Alfè, M.J. Gillan, M.D. Towler and R.J. Needs, to be published.

[14] M.D. Towler, R. Maezono and R.J. Needs, to be published.

[15] A.J. Williamson, R.Q. Hood, and J.C. Grossman, *Phys. Rev. Lett.* **87**, 246406 (2001).

[16] S. Manten and A. Lüchow, *J. Chem. Phys.* **119**, 1307 (2003).

[17] E. Schrödinger, *Ann. Phys.* **79**, 361 (1926).

[18] N.W. Ashcroft and N.D. Mermin, *Solid State Physics*, (Saunders, 1976), p.330.

[19] G. Rajagopal, R.J. Needs, A.J. James, S.D. Kenny, and W.M.C. Foulkes *Phys. Rev. B* **51**, 10591 (1995).

[20] G. Rajagopal, R.J. Needs, S.D. Kenny, W.M.C. Foulkes and A.J. James *Phys. Rev. Lett.* **73**, 1959 (1994).

[21] P.P. Ewald, Ann. Phys. **64**, 25 (1921).

[22] M.P. Tosi, in *Solid State Physics*, Vol. 16, edited by H. Ehrenreich and D. Turnbull (Academic, New York, 1964), p. 1.

[23] L.M. Fraser, W.M.C. Foulkes, G. Rajagopal, R.J. Needs, S.D. Kenny, and A.J. Williamson, *Phys. Rev. B* **53**, 1814 (1996) ; A.J. Williamson, G. Rajagopal, R.J. Needs, L.M. Fraser, W.M.C. Foulkes, Y. Wang, and M.-Y. Chou, *Phys. Rev. B* **55**, R4851 (1997).

[24] W. Kohn, *Rev. Mod. Phys.* **71**, 1253 (1999).

[25] N.D. Drummond, First Year Ph.D. Report, University of Cambridge.

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

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

[28] Z. Radnai, N.D. Drummond, M.D. Towler, and R.J. Needs, unpublished.

[29] S. Fahy, in *Quantum Monte Carlo Methods in Physics and Chemistry*, Nato Science Series C: Mathematical and Physical Sciences, Vol. 525, edited by M.P. Nightingale and C.J. Umrigar (Kluwer Academic, Dordrecht, The Netherlands, 1999), p. 101.

[30] C. Filippi and S. Fahy, *J. Chem. Phys.* **112**, 3523 (2000).

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

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

[33] T. Kato, *Commun. Pure Appl. Math.* **10**, 151 (1957).

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

[35] W.M.C. Foulkes, R.Q. Hood, and R.J. Needs, *Phys. Rev. B* **60**, 4558 (1999).

[36] J.C. Grossman and L. Mitas, *Phys. Rev. Lett.* **74**, 1323 (1995).

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

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

[39] R.Q. Hood, P.R.C. Kent, R.J. Needs, and P.R. Briddon *Phys. Rev. Lett.* **91**, 076403 (2003).

[40] W. Kutzelnigg and J.D. Morgan, III, *J. Phys. Chem.* **96**, 4484 (1992).

[41] D. Prendergast, M. Nolan, C. Filippi, S. Fahy, and J.C. Greer, *J. Chem. Phys.* **115**, 1626 (2001).

[42] E. Hernandez, M.J. Gillan, and C.M. Goringe *Phys. Rev. B* **55**, 13485 (1997).

[43] D. Alfè and M.J. Gillan, to be published.

[44] Gaussian 03, Revision B.03, M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery, Jr., T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J.B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, and J.A. Pople, Gaussian, Inc., Pittsburgh PA, 2003.

[45] V.R. Saunders, R. Dovesi, C. Roetti, R. Orlando, C.M. Zicovich-Wilson, N.M. Harrison, K. Doll, B. Civalleri, I. Bush, Ph. D'Arco and M. Llunell, CRYSTAL2003 User's Manual (University of Torino, Torino, 2003).

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

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

[48] P.H. Acioli and D.M. Ceperley, *J. Chem. Phys.* **100**, 8169 (1994).

[49] E.L. Shirley and R.M. Martin, *Phys. Rev. B* **47**, 15413 (1993).

[50] C.W. Greeff and W.A. Lester, Jr., *J. Chem. Phys.* **109**, 1607 (1998).

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

[52] M.D. Towler and R.J. Needs, unpublished.

[53] M.D. Towler, A. Ma, and R.J. Needs, to be published.

[54] S. Goedecker, *Rev. Mod. Phys.* **71**, 1085 (1999).

[55] B. Farid and R. J. Needs, *Phys. Rev. B* **45**, 1067 (1992).

[56] S. Fahy, X.W. Wang, and S. G. Louie, *Phys. Rev. B* **42**, 3503 (1990).

[57] A. Malatesta, S. Fahy, and G.B. Bachelet, *Phys. Rev. B* **56**, 12201 (1997).

[58] W.-K. Leung, R.J. Needs, G. Rajagopal, S. Itoh, and S. Ihara, *Phys. Rev. Lett.* **83**, 2351 (1999).

[59] E. Knittle, R. Wentzcovitch, R. Jeanloz, and M.L. Cohen, *Nature* **337**, 349 (1989).

[60] M.D. Segall, P.L.D. Lindan, M.J. Probert, C.J. Pickard, P.J. Hasnip, S.J. Clark, and M.C. Payne, *J. Phys.: Cond. Mat.* **14** 2717 (2002).

[61] *First-principles computation of material properties: the ABINIT software project*, X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, Ph. Ghosez, J.-Y. Raty, D.C. Allan. *Computational Materials Science* **25**, 478-492 (2002).

[62] S. Baroni, A. Dal Corso, S. de Gironcoli, and P. Giannozzi, `http://www.pwscf.org`.

[63] C.K. Skylaris, A.A. Mostofi, P.D. Haynes, O. Diéguez, M.C. Payne, *Phys. Rev. B* **66**, 035119 (2002) ; A.A. Mostofi, P.D. Haynes, C.K. Skylaris and M.C. Payne, *J. Chem. Phys.* **119**, 8842 (2003).

[64] R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, *Chem. Phys. Lett.* **162**, 165 (1989).

[65] C. Filippi and C.J. Umrigar, *Phys. Rev. B* **61**, 16291 (2000).

[66] R. Assaraf and M. Caffarel, *J. Chem. Phys.* **113**, 4028 (2000).

[67] K.C. Huang, R.J. Needs and G. Rajagopal, *J. Chem. Phys.* **112**, 4419 (2000).

[68] J. Vrbik and S.M. Rothstein, *J. Chem. Phys.* **96**, 2071 (1992) ; P. Belohorec, S.M. Rothstein and J.Vrbik, *J. Chem. Phys.* **98**, 6401 (1993).

[69] J.C. Grossman and L. Mitas, to be published.

[70] J.C. Grossman, *J. Chem. Phys.* **117**, 1434 (2002).

[71] M.D. Towler, unpublished.

[72] J.A. Pople, M. Head-Gordon, D.J. Fox, K. Raghavachari, and L.A. Curtiss, *J. Chem. Phys.* **90**, 5622 (1989).

[73] L.A. Curtiss, C. Jones, G.W. Trucks, K. Raghavachari, and J.A. Pople, *J. Chem. Phys.* **93**, 2537 (1990).

[74] L.A. Curtiss, K. Raghavachari, P.C. Redfern, and J.A. Pople, *J. Chem. Phys.* **106**, 1063 (1997).

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

[76] D.M. Ceperley and B. Bernu, *J. Chem. Phys.* **89**, 6316 (1988).

[77] F. Arias de Saavedra and M.H. Kalos, *Phys. Rev. E* **67**, 026708 (2003) ; S.W. Zhang and M.H. Kalos, *J. Stat. Phys.* **70**, 515 (1993).

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

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

[80] Royal Society web site : `www.royalsoc.ac.uk`.