# The CASINO program : quantum Monte Carlo in molecular quantum chemistry and condensed matter physics

M.D. Towler and R.J. Needs

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

**Abstract**

The use of diffusion Monte Carlo as a tool for describing the electronic structure of molecules and solids is discussed, including a brief description of a new powerful and generally applicable quantum Monte Carlo software package. Some preliminary results for calculations on a typical strongly-correlated oxide (NiO) are also presented.

## 1 Introduction

Our ability to use computers to study, understand, and ultimately predict the detailed behaviour of electrons in matter, condensed or otherwise, has improved dramatically during the past few decades. Advances in computer technology and the development of new numerical methods have opened up opportunities for solving problems which only recently were considered intractable. Powerful first principles methods are capable of providing accurate 'experimental' data pertaining to atomic scale processes in molecules and solids which can be almost impossible to probe by other means. One problem with such calculations for real materials is the difficulty of finding a consistently accurate way of accounting for the complicated correlated motion of the interacting electrons. The usual approach is to use density functional theory (DFT). This method is exact in principle but relies in practice on occasionally unreliable approximations to an unknown exchange-correlation functional. The standard local density (LDA) and generalized gradient (GGA) approximations to this quantity are very successful on the whole, but fail in a number of important cases. Examples include difficulties in describing weak interactions such as van der Waals' and hydrogen bonds, the treatment of excited states, incorrect energetic ordering of covalently bonded structures with different types of bonds and/or coordination numbers (Which is the smallest stable fullerene? It's difficult to tell with DFT. [1]), and the lack of a satisfactory qualitative description of the physics in 'strongly correlated' systems, which include technologically important materials such as doped manganites and the high-$T_c$ cuprate superconductors. For situations when DFT is not good enough, molecular quantum chemists have in their toolbox a variety of advanced correlated wave function methods such as configuration interaction, CASSCF, and coupled cluster, which are capable of very high accuracy. However, their use has largely been limited to quite small molecules due to the highly unfavourable scaling of computational cost with the number of electrons. Some of these methods suffer from being non-variational or from size consistency problems, and for various reasons, they are also difficult to apply to solids.

It is with issues like this in mind that the quantum Monte Carlo (QMC) method can be a valuable tool to both theoretical chemists and computational physicists. Continuum fermion QMC is a generic term covering several different techniques based on random sampling of the many-particle wave function, which collectively make up a consistent and highly accurate treatment of electron correlation in real materials based on an explicit, non-perturbative description of the many-body quantum effects. Such simulations have several important aspects. First, QMC has been demonstrated to be as accurate as the best of the molecular quantum chemistry techniques. Second, QMC has a reasonable scaling with system size in that computational demands grow roughly as the cube of the system size, $N^3$. This may be favourably compared with the most popular correlated wave function techniques such as coupled cluster theory which scale at least as $N^6$. Furthermore there are no memory or disk bottlenecks even for large systems; with our code small molecular calculations are essentially trivial, and calculations on real crystalline solids with many hundreds of electrons per cell are becoming routine. Third, there is complete freedom in how to represent the trial many-electron wave functions used in QMC without the added complications of problems such as basis set superposition errors or lack of size consistency. Fourth, the diffusion Monte Carlo (DMC) method is essentially immune to finite basis set errors - since it does not represent the wave function in terms of a basis set. The only remaining errors are small and depend weakly on the nodal surface of a trial many-electron wave function. Finally, one can treat both ground and excited states, and all gradations of 'weakly' and 'strongly' correlated systems. With QMC therefore, one can calculate, with high accuracy, ground states, excitation energies, chemical reaction barriers and other properties of essentially any system within a single unified methodological framework.

Another reason QMC methods are now tractable for real systems has been the increasing availability of parallel computers. The QMC algorithm is intrinsically parallel, i.e., one can map independent random walks on individual processors with negligible communication overhead and thus achieve essentially perfect scalability even on massively parallel platforms with thousands of processors. They are thus ideally suited for current and projected innovations in computer technology such as large PC clusters and the proposed new 'Grid' computing models. QMC methods have already been able to provide important results. The most obvious example is Ceperley and Alder's 1980 QMC work on the homogeneous electron gas [2], the results of which were parameterized and made into the first reliable local density approximation to DFT. Clearly this was a vital stimulus to the astonishing growth of *ab initio* electronic structure and quantum molecular dynamics calculations that has occurred over the last twenty years. More recent QMC work [3] on calculating accurate exchange-correlation holes in real materials may yet lead to the development of more accurate density functionals.

For quantum Monte Carlo calculations to become more accessible to the wider community it is clearly essential to have available a good general computer program applicable to all types of problem. Over the last few years, we have been developing just that. Version 1.0 of 'CASINO' [4], as it has come to be called, was released in the summer of 2000 and is already widely used within the QMC research community. CASINO was planned and written (in the modern fortran90 language) using professional modular software design techniques. It is able to treat ground and excited states of (i) finite systems (atoms/molecules/clusters), (ii) systems with periodic boundary conditions in 1D/2D/3D (polymers, slabs/surfaces, crystalline solids), (iii) 2D/3D homogeneous and inhomogeneous electron gases in either fluid or crystalline phases (iv) 2D/3D electron-hole systems [5] with variable mass ratio in fluid, crystalline and excitonic insulator phases. It is able to perform both variational Monte Carlo calculations (with automatic variance minimization) and diffusion Monte Carlo calculations using either workstations, weakly linked clusters or massively parallel computers. The many-electron trial wave function consists of a sum of a Slater part consisting of a small number (often just one) of determinants of one-electron orbitals derived from a Hartree-Fock (HF) or DFT calculation, multiplied by a Jastrow factor depending explicitly on the position and mutual separation of pairs of particles. The orbitals in the Slater part may be represented using basis sets of plane waves, Cartesian Gaussian-type functions of arbitrary angular momentum, or cubic spline functions on a grid ('blip functions'). The code has generic interfaces to allow the transfer of self-consistent orbitals from a range of standard packages such as CRYSTAL [6], GAUSSIAN [7] and CASTEP.[8]

Over the last few years QMC methods have begun to enter the mainstream of electronic structure theory; they are coming to be seen as a standard part of the electronic structure toolbox, and are being used by a wider range of groups interested in a wider range of problems than in the past. To maintain this momentum, the process of doing QMC simulations needs to be made less painful and less time consuming, and work is continuing to develop our QMC methods and programs to the stage where any competent theorist faced with a specific problem for which density functional approaches are not accurate enough can use QMC more or less as a black box.

### 1.1 Diffusion Monte Carlo

In this section we outline briefly the basic principles of the diffusion Monte Carlo method, which is a stochastic projector method for evolving the imaginary-time Schrödinger equation,

$$-\frac{\partial \Psi(\mathbf{R},t)}{\partial t} = -\frac{1}{2}\nabla^2\Psi(\mathbf{R},t) + (V(\mathbf{R}) - E_S)\Psi(\mathbf{R},t) \; . \qquad (1)$$

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, $E_S$ is an energy offset which only affects the normalization of the wave function $\Psi$, 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 (provided they have a non-zero overlap). In DMC the time evolution of Eq. 1 may be followed using a stochastic technique in which $\Psi(\mathbf{R}, t)$ is represented by an ensemble of $3N$-dimensional electron configurations, $\{\mathbf{R}_i\}$. The time evolution of these configurations is governed by the Green function of Eq. 1. Within the short time approximation the propagation divides into two processes: random diffusive jumps of the configurations arising from the kinetic term and removal/addition 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 $\Psi$ 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/adding configurations diverges when the potential energy diverges, which occurs whenever two electrons or an electron and a nucleus are coincident, leading 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* $f = \Phi_T \Psi$, where $\Phi_T$ is known as the trial or guiding wave function, and substitute into Eq. 1 we obtain

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

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

$$\mathbf{v}_D(\mathbf{R}) = \nabla \ln |\Phi_T(\mathbf{R})| = \Phi_T(\mathbf{R})^{-1} \nabla \Phi_T(\mathbf{R}) \ , \tag{3}$$

and

$$E_L(\mathbf{R}) = \Phi_T^{-1} \left( -\frac{1}{2}\nabla^2 + V(\mathbf{R}) \right) \Phi_T \ , \tag{4}$$

is the local energy.

This formulation imposes the *fixed-node approximation*.[9] 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 $\Psi$ is constrained to be the same as that of $\Phi_T$ and therefore $f$ can be interpreted as a probability distribution. The time evolution generates the distribution $f = \Phi_T\Psi$, where $\Psi$ is the best (lowest energy) wave function with the same nodes as $\Phi_T$. The problem of the poor statistical behaviour due to the divergences in the potential energy can also be solved because the term $(V(\mathbf{R})-E_S)$ in Eq. 1 has been replaced by $(E_L(\mathbf{R})-E_S)$ which is much smoother. Indeed, if $\Phi_T$ was an exact eigenstate then $(E_L(\mathbf{R}) - E_S)$ would be independent of position in configuration space. In practice we must use approximate trial wave functions, but we can use our understanding of the physics of many-body wave functions and powerful optimization techniques to make $(E_L(\mathbf{R}) - E_S)$ as smooth as possible.

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 fermion ground states [10, 11] 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.[11]

A DMC simulation proceeds as follows. First we pick an ensemble of a few hundred configurations chosen from the distribution $|\Phi_T|^2$ using the standard Metropolis algorithm. This ensemble is evolved in imaginary time according to the rules of the importance sampled imaginary time Schrödinger equation, 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 $\Phi_T$ is large. After a period of equilibration 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

$$E_D = \frac{\int f(\mathbf{R})E_L(\mathbf{R})\,d\mathbf{R}}{\int f(\mathbf{R})\,d\mathbf{R}} \approx \sum_i E_L(\mathbf{R}_i)\ . \tag{5}$$

This energy expression would be exact if the nodal surface of $\Phi_T$ was exact, and the fixed-node error is second order in the error in the nodal surface of $\Phi_T$ (when a variational theorem exists). 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. 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. It is essential to have a reasonably accurate trial wave function, and we use trial wave functions of the Slater-Jastrow form:

$$\Phi = \mathcal{D}_{\uparrow}\mathcal{D}_{\downarrow} \exp\left(\mathcal{J}\right) \ , \tag{6}$$

where $\exp\left(\mathcal{J}\right)$ is a Jastrow correlation factor and $\mathcal{D}_i^{\uparrow}$ and $\mathcal{D}_i^{\downarrow}$ are Slater determinants of single-particle up-/down-spin orbitals. The Jastrow factor is positive and multiplicative and does not alter the nodes of the wave function which are therefore fixed by the Slater part. The quality of the Slater part of the Slater-Jastrow wave function is crucial because it usually accounts for in excess of 90% of the electronic energy. The Slater part is normally obtained from DFT, HF or similar single-particle methods, so that they include a great deal of chemical information about the system. The Jastrow factor decays with the separation between particles and it is chosen to obey the cusp conditions when two electrons are coincident [12].

This completes our brief introduction to the DMC method. We have left out many technical details of the calculations, and for more information on these aspects the reader is directed to Refs. [13, 14].

## 2 Improved scaling algorithms

It may well be that QMC is the most accurate method available for medium to large systems, and the only such method generally applicable to solids, but it remains relatively expensive compared to DFT. This in itself doesn't matter - the two methods are not in competition. Rather they are designed to complement each other. Nevertheless, if everyone used QMC it would be enough to swamp the supercomputing resources of a medium-sized country. To see how this situation can be improved, we shall now discuss the scaling of computational cost with the number of electrons, $N$.

QMC methods are stochastic and therefore yield mean values with an associated statistical error bar. We normally 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 $\Delta E$ on the energy of the system which is independent of system size, a feature denoted here by $\Delta E = \mathcal{O}(1)$. There are, however, other quantities such as cohesive energies, geometrical properties, and elastic constants, for example, for which the relevant quantity is the energy *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? The result for the standard algorithm is $C = AN^3 + \epsilon N^4$, where $\epsilon$ is small.[14] 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? The best possible scaling for conventional (non-stochastic) single-particle methods such as DFT is $\mathcal{O}(N)$. Much effort has been expended in designing DFT codes which not only scale linearly but additionally are (a) faster than the regular cubic scaling algorithm for reasonable system sizes, and (b) *as accurate* as codes using the regular algorithm. In wave function-based QMC, these latter two problems do not occur; with the improved scaling algorithms that are clearly possible 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, it has recently been shown by collaborators of ours using our code that this approach is extremely effective in molecular QMC calculations using maximally-localized Wannier function orbitals.[15]

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 seems natural to truncate the Jastrow factor at the radius of the exchange-correlation hole, and tests that we have already performed show that this works well. 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. For very large systems, the $\epsilon N^4$ term can be significant, but this can be streamlined through the use of sparse matrix techniques.

# 3 Strongly correlated systems

In the solid state, first principles calculations now underpin our understanding of the electronic structure and bonding in many different types of materials. However, attempts to apply these methods to 'strongly correlated' systems containing transition elements, from simple monoxides to the high-$T_c$ superconducting cuprates, have been fraught with difficulties. For example, when DFT calculations using the standard LDA or GGA functionals are performed on the magnetically ordered, insulating ground state of a material as apparently simple as NiO, a metallic solution is often obtained and many simple properties come out badly wrong. For many years such results were thought to be a failure of the one-electron approximation itself and definitive proof that these systems could not be described with traditional 'band theory' methods. This view has completely permeated the literature, but is now known to be misleading - the failure of such calculations is essentially due to approximations in the design of simple exchange-correlation functionals. The root of the problem is that the eigenvalue spectrum comes out wrong. This in itself is hardly surprising since the DFT eigenvalues are not formally defined as approximations to the energies of one-electron states, but the problem is more serious than that in strongly correlated systems. We can get away with using DFT in silicon, but when the method is required to split a complex manifold of states arising from 'atomic-like' $d$ orbitals in some particular way in order to produce a gap, there are difficulties. Within the LDA, for example, the potential felt by each electron is computed from a functional of the total electron density. This leads to eigenvalues which are relatively weak functions of the particular occupancy, and no gap appears. Furthermore, because of the way the exchange interaction is calculated, the spurious 'self interaction' of an electron with itself is not cancelled, which further scrambles the spectrum.

One might think that this is only really a problem if we wish to calculate spectroscopic properties such as the band gap, and that one still ought to be able to calculate, say, geometrical or elastic properties which depend largely on the electron density. This may sometimes be true, but that is not really the point. In some systems, having the wrong eigenvalue spectrum can seriously affect the density and lead to a solution which has a qualitatively incorrect ground state. Two excellent and important examples are provided by the doped manganites in which colossal magnetoresistance occurs and the high $T_c$ superconducting cuprates. Here the essential part of the problem is 'What happens to these compounds when we dope them with holes?', the answer to which will in general depend on the ordering of the states in the valence band. However, in an LDA calculation of an oxide, the 'metal states' will typically be raised greatly in energy compared to the 'oxygen states' due to the spurious self interaction. This might artificially force the holes to sit largely on the metal sites instead of on the oxygens in seeming contradiction to experiment. So there are many problems, but one can easily imagine designing functionals which reproduce the spectrum of the real system more accurately, and thus the standard view of 'band theory' being inapplicable in some sense for strongly correlated materials is not right, a fact which remains widely unappreciated. That said, the problem is of course a very serious one since most DFT computer codes are only set up to do LDA and GGA calculations, and it is thus rather easy to apply them uncritically.

So what to do about this situation? One simple way to address the problem which appears to lead to excellent trial wave functions for QMC is to use the *unrestricted Hartree-Fock* (UHF) method. The key to this approach is the realization that the magnetically ordered insulating ground state may often be described by a single determinant wave function and something as simple as the UHF approximation which, by definition, contains the exact exchange interaction. Its eigenvalue spectrum is defined, via Koopmans' theorem, as an approximation to the true one-electron energies and may therefore be expected to split the complex manifold of states in the right way and produce a gap. Furthermore, because the non-local exchange exactly cancels the self interaction, the ordering of states in the spectrum is likely to be correct, and indeed the manifold of $d$ states splits in precisely the manner expected from a simple empirical ('Hubbard model') estimate of the on-site interactions between electrons in different $d$ orbitals. It has been shown that UHF calculations (and also hybrid schemes using combinations of DFT and exact exchange such as B3LYP) of strongly correlated magnetic insulators agreed very well with experiment, a fact now documented in a variety of publications. It may seem surprising to molecular quantum chemists that a single determinant HF wave function could be so successful but this is an important characteristic of these *ionic* materials. The highly symmetric environment and long range Coulomb forces tend to separate the orbitals into well defined subsets with a significant gap between occupied and unoccupied states. Hence, the ground state of NiO is rather well described by a single determinant while one could easily imagine a covalently-bonded molecular complex (for example) for which this approximation would be poor.

So there are various ways of doing band theory for strongly correlated systems, and some of them produce qualitatively reasonable approximations to the true ground state. The remaining problem is simply a lack of good quantitative agreement with experiment for certain properties, particularly in HF where we don't have any electron correlation at all (!). But this is OK, because a *reasonable guess at the ground state* is all we need as input to a quantum Monte Carlo calculation. Of course in QMC the wave function is flexible enough to describe any shape of pair-correlation function, there is no such problem as self-interaction, and we do not have a one-electron eigenvalue spectrum to get right or wrong. There is thus little practical distinction between strongly and weakly correlated materials at this level.

From our pseudopotential UHF calculations we obtain a cohesive energy for NiO of 5.89 eV per NiO pair, which is in reasonable agreement with a previous UHF all-electron value [17] of 6.2 eV, but is much smaller than the experimental value of 9.5 eV per NiO pair.[18] The DMC data gives a cohesive energy of 9.442(2) eV, in excellent agreement with experiment (the additional contribution from zero-point effects is expected to be around 0.1 eV using a Debye approximation with $\Theta = 640$ K. [19]). This excellent result clearly suggest there is considerable scope for using DMC to study strongly-correlated systems. A more wide-ranging DMC study of NiO including an analysis of the effectiveness of trial wave functions resulting from different one-electron Hamiltonians will be the subject of a future publication.[20]

## 4 Conclusions

We have given a brief introduction to the DMC method as used in the CASINO code. The crucial requirement for a successful DMC calculation is the availability of a reasonably accurate trial wave function. The problem is particularly acute in strongly correlated systems such as NiO. However, we have found that UHF appears to provide a satisfactory zeroth order description, giving the correct ordering of the energy levels and yielding an insulating state, although as is generally found in Hartree-Fock approaches, the band gap is much too large. DMC calculations with UHF-based trial wave functions work well and we have obtained a cohesive energy in good agreement with experiment.

## 5 Acknowledgments

Financial support was provided by EPSRC (UK). MDT gratefully acknowledges the support of the Lloyd's Tercentenary Foundation. Most of the calculations were performed on the CRAY-T3E administered by CSAR at the University of Manchester and the IBM SP located at the University of Cambridge High Performance Computing Facility.

## References

[1] P. R. C. Kent, M. D. Towler, R. J. Needs, and G. Rajagopal, Phys. Rev. B **62** (2000) 15394.

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

[3] M. Nekovee, W. M. C. Foulkes, and R. J. Needs, Phys. Rev. Lett. **87** (2001) 036401.

[4] R. J. Needs, G. Rajagopal, M. D. Towler, P. R. C. Kent, and A. J. Williamson, CASINO version 1.0 User's Manual, University of Cambridge, Cambridge, 2000 ;
www.tcm.phy.cam.ac.uk/∼mdt26/casino.html .

[5] G. J. Brown, M. D. Towler, P. B. Littlewood, and R. J. Needs, unpublished.

[6] R. Dovesi *et al.*, CRYSTAL98 User's Manual (University of Torino, Torino, 1998).

[7] M. J. Frisch *et al.*, *Gaussian 98*, Revision A.9 (Gaussian Inc., Pittsburgh, PA, 1998).

[8] Molecular Simulations Inc. (1998). CASTEP User's Guide. Molecular Simulations Inc., San Diego, CA (1998)

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

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

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

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

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

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

[15] A. J. Williamson, R. Q. Hood, and J. C. Grossman (unpublished).

[16] A. P. Sutton, p. 227, *Electronic Structure of Materials* (Clarendon, 1993).

[17] M. D. Towler, N. L. Allan, N. M. Harrison, V. R. Saunders, W. C. Mackrodt, and E. Aprà, Phys. Rev. B **50** (1994) 5041.

[18] *CRC Handbook of Chemistry and Physics* 75th Ed., ed. D. R. Lide (CRC Press, Boca Raton, 1994/1995).

[19] R. A. Coy, C. W. Thompson, and E. Gürmen, Solid State Commun. **18** (1976) 845.

[20] M. D. Towler and R. J. Needs (unpublished).