New Avenues for Quantum Monte Carlo Techniques

D. Alfè* and M. J. Gillan†, *Department of Earth Sciences, †Department of Physics and Astronomy, University College London
N. D. Drummond, M. D. Towler and R. J. Needs, Cavendish Laboratory, University of Cambridge

The Holy Grail of condensed matter physics is the ability to solve the 78 year old Schrödinger equation for "real-life" systems. The great difficulty is that the motions of the electrons in molecules and crystals are correlated because of the strong repulsion between their negative charges. The solution of the many-electron Schrödinger equation (the wavefunction) is therefore a complicated function of $3N$ variables, where N is the number of electrons in the system. A boost to the field was given by Hohenberg and Kohn (HK) 40 years ago [1], with the introduction of density functional theory (DFT). They showed that it was possible to reformulate quantum mechanics in such a way that the important physical quantity is the electron density, rather than the electronic wavefunction, reducing the complexity of the problem from $3N$ to 3. Of course, free lunches are seldom available, and the price to pay for the great simplification of HK was the introduction of a new quantity, called exchange-correlation (XC) energy, which is an (as yet) unknown functional of the electron density. To make the theory work, HK suggested a simple form for the XC functional, known as the local-density approximation (LDA). The LDA is exact in a system with a homogeneous electron density, and is only an approximation in (real-life) inhomogeneous systems. Forty years later the LDA is still widely used, and has been the main factor for the great success of DFT, however, there are a number of cases where more accuracy is needed than the LDA can provide. Physicists have struggled for decades to find better approximations for the XC energy, and come up with a number of improvements to the LDA, but there are still several 'difficult' cases for which no available approximation for the XC is really satisfactory.

A completely different approach to the problem is offered by so called quantum Monte Carlo (QMC) techniques. Although their demands on computer power are much greater than those of widely used techniques such as DFT, their accuracy is also much greater for most systems. QMC traditionally embraces two different techniques: variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC). For a detailed descriptions of VMC and DMC see for example Ref. [2], here we only outline rather briefly the main ideas for the two techniques.

The VMC method gives an upper bound on the exact ground-state energy E_0, but the results depend totally on the quality of the chosen trial wavefunction. Given a normalized trial wavefunction $\Psi_T(\boldsymbol{R})$, where $\boldsymbol{R} = (\boldsymbol{r}_1, \boldsymbol{r}_2...,\boldsymbol{r}_N)$ is a $3N$-dimensional vector representing the positions of N electrons, and denoting by $\hat{H}$ the many-electron Hamiltonian, the variational energy $E_v \equiv \langle \Psi_T | \hat{H} | \Psi_T \rangle \geq E_0$ is estimated by sampling the value of the local energy $E_L(\boldsymbol{R}) \equiv \Psi_T^{-1}(\boldsymbol{R}) \, \hat{H} \, \Psi_T(\boldsymbol{R})$ with configurations $\boldsymbol{R}$, distributed according to the probability density $\Psi_T(\boldsymbol{R})^2$. In our calculations the trial wavefunction $\Psi_T(\boldsymbol{r}_1,...\boldsymbol{r}_N)$ consists of a Slater determinant D of single-electron orbitals $\psi_n(\boldsymbol{r}_i)$ multiplied by a parameterized Jastrow correlation factor $J(\boldsymbol{r}_1,... \boldsymbol{r}_N)$ (in the pseudopotential-based QMC of interest here, the $\psi_n(\boldsymbol{r}_i)$ are commonly taken from a plane-wave pseudopotential DFT calculation), and J is "optimized" by varying it's parameters so as to reduce the variance of the "local energy" $\Psi_T^{-1}(\hat{H} \Psi_T)$, where $\hat{H}$ is the many-electron Hamiltonian.

A much more accurate method is DMC, a scheme which in principle yields the exact ground state energy. The main idea of DMC is to solve the Schrödinger equation in imaginary time. This is done by recognizing that this is equivalent to a diffusion equation which includes a potential term, and therefore describes the stochastic motion of Brownian particles (walkers), with the effect of the potential being equivalent to a branching term which has the effect of increasing the number of walkers in regions of low potential and decrease it in regions of high potential. Since there is a one-to-one correspondence between the density of walkers and the value of the wavefunction in any region of space, one can compute the value of the wavefunction by following the motion of these walkers in configuration space. To make the calculations practicable it is necessary to "guide" the walkers using a "trial" wavefunction. The

trial wavefunction should be as close as possible to the real many-body wavefunction for optimal efficiency. Moreover, because of the fermionic nature of the electrons, the many-body wavefunction necessarily has regions of space where it is positive and regions where it is negative. This change of sign is a problem for the diffusion of the walkers. The way this problem is usually addressed is by fixing the surface which separates regions of space where the wavefunction changes sign (nodal surface). The presence of the trial wavefunction allows an easy implementation of this so called "fixed-node" constraint.

With QMC now being applied to large complex systems containing hundreds of atoms, a major issue is the scaling of the required computer effort with system size. The cost of traditional quantum mechanics techniques increases at least with the square of the number of atoms N in the system, and for most of them (including standard QMC) increases as N^3. However, feeding out of ideas about the locality of quantum coherence [3], it has been shown that in fact linear scaling should be achievable, and $O(N)$ techniques have been developed for tight binding (TB) [4], DFT [5, 6] and Hartree-Fock [7]. Recently, a procedure has been suggested [8] for achieving at least partial linear scaling for QMC, based on the idea of "maximally localized Wannier functions" [9]. This is done by performing a unitary transformation which preserves the orthogonality of the single particle orbitals that make up the determinant D, and by using a localized basis set for representing the single particle orbitals. Since the single particle orbitals are mixed by a unitary transformation, the determinant D is unchanged. Linear scaling is achieved then by truncating the orbitals beyond a certain localization region ω. Of course, in order to have accurate calculations, ω must be big enough so that the localization weights of the single particle orbitals, defined as $P = \int_{\omega} d\,r\,|\,\phi\,(r)\,|^2$, is as close to 1 as possible.

We have implemented and tested an alternative method which has two important advantages over the method of Ref. [8]. Firstly, we have developed a technique for representing single-particle orbitals in terms of localized B-splines [10]. These localized functions have previously been used in the context of DFT [11], and they are closely related to plane waves. Our implementation [10] therefore provides a natural interface between plane-wave-DFT codes, from which one can obtain good trial wavefunctions, and QMC codes. Secondly, we have shown that the localization of the single-particle orbitals can be greatly enhanced by removing the orthogonality constraint (i.e. by using a non-unitary

linear transformation on the set of orbitals) [12].

We have tested our method on the prototypical oxide material MgO in the rock-salt structure at ambient pressure. Because of its large band gap (experimental Eg = 7.7 eV), this kind of material should be particularly suited to O(N) methods. The QMC calculations 2 were performed using the appropriately modified CASINO code [13] and a supercell of 64 atoms.

To compare the efficiency of our method with the earlier technique based on Wannier functions [8], we plot Q = 1 - P against the side-length of the cubic localization region for the two methods in Fig. 1. In Fig. 2 the VMC energies obtained using the two methods are plotted against the side-length of the localization region. The energy obtained with untruncated orbitals (the 'standard' QMC result) is also shown. We see that nonorthogonal orbitals are much more localized than Wannier functions, and that the VMC energy evaluated with these orbitals converges to the correct value much more rapidly when the localization region is expanded. Further details can be found in our recent publication, Ref. [12].

In conclusion, we have proposed and tested a new technique for achieving linear scaling in one of the most demanding parts of QMC calculations. This technique is simple and robust, and it appears also to be more efficient than an earlier proposed technique. This new technique already makes it possible to treat large systems that would be out of reach of conventional QMC methods. Research areas where we intend to apply the technique in the near future include defects and surfaces of oxide materials.

References

[1] P. Hohenberg and W. Kohn, Phys. Rev. 136, B864 (1964).
[2] W. M. C. Foulkes, L. Mitas, R. J. Needs and G. Rajagopal, Rev. Mod. Phys. 73, 33 (2001).
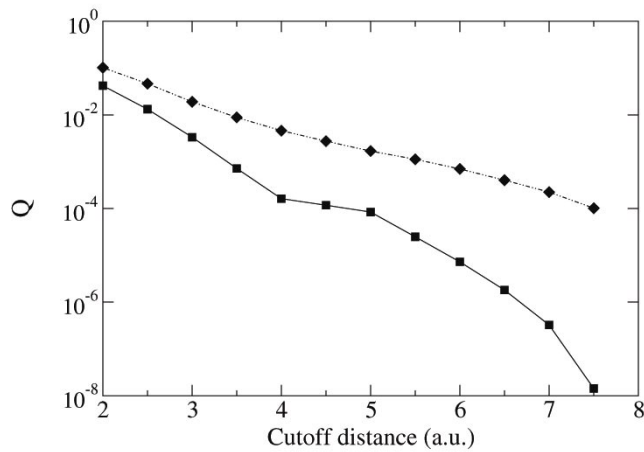
Figure 1: Dependence of localization weight on cut-off distance for bulk MgO. The quantity plotted is $Q \equiv 1 - P$, where P is localization weight defined in the text. Squares: present non-orthogonal orbitals. Diamonds: maximally localized Wannier orbitals.
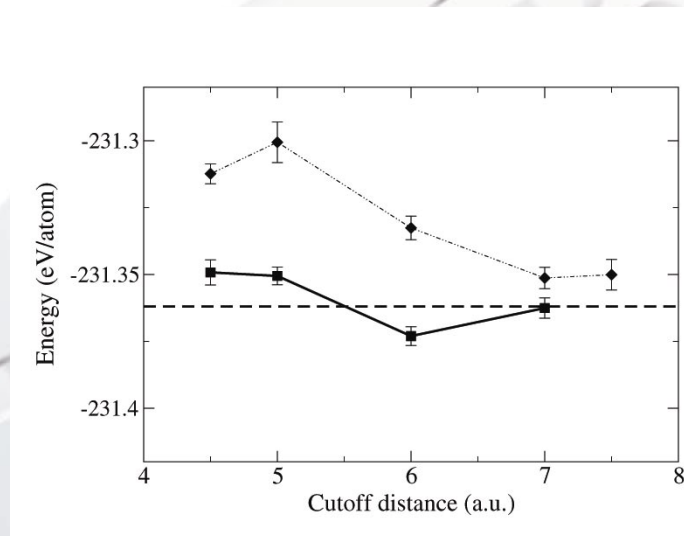

Figure 2: Convergence of linear-scaling VMC total energy per atom to the value obtained with extended orbitals for bulk MgO. Squares: present non-orthogonal orbitals. Diamonds: maximally localized Wannier functions. Horizontal dashed line shows total energy/atom obtained with extended orbitals.

[3] W. Kohn, Phys. Rev. Lett. 76, 3168 (1996).

[4] D. R. Bowler, M. Aoki, C. M. Goringe, A. P. Horsfield and D. G. Pettifor, Modell. Simul. Mater. Sci. Eng. 5, 199 (1997)

[5] D. R. Bowler, T. Miyazaki and M. J. Gillan, J. Phys.: Condens. Matter 14, 2781 (2002).

[6] J. M. Soler, E. Artacho, J. D. Gale, A. Garcia, J. Junquera, P. Ordejon and D. Sanchez-Portal, J. Phys.: Condens. Matter 14, 2745 (2002).

[7] M. Challacombe, J. Chem. Phys. 110, 2332 (1999).

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

[9] N. Marzari and D. Vanderbilt, Phys. Rev. B 56, 12847 (1997).

[10] D. Alfè and M. J. Gillan, Phys. Rev. B, Rapid Comm., in press.

[11] E. Hernández, M. J. Gillan and C. M. Goringe, Phys. Rev. B 55 13485, (1997).

[12] D. Alfè and M. J. Gillan, J. Phys.: Condens. Matter 16, L305 (2004).

[13] R. J. Needs, M. D. Towler, N. D. Drummond and P. R. C. Kent, 'CASINO Version 1.7 User Manual', University of Cambridge, Cambridge (2004).