

# Towards an exact description of electronic wavefunctions in real solids

George H. Booth<sup>1</sup>, Andreas Grüneis<sup>1,2</sup>, Georg Kresse<sup>2</sup> & Ali Alavi<sup>1</sup>

The properties of all materials arise largely from the quantum mechanics of their constituent electrons under the influence of the electric field of the nuclei. The solution of the underlying many-electron Schrödinger equation is a ‘non-polynomial hard’ problem, owing to the complex interplay of kinetic energy, electron–electron repulsion and the Pauli exclusion principle. The dominant computational method for describing such systems has been density functional theory. Quantum-chemical methods—based on an explicit ansatz for the many-electron wavefunctions and, hence, potentially more accurate—have not been fully explored in the solid state owing to their computational complexity, which ranges from strongly exponential to high-order polynomial in system size. Here we report the application of an exact technique, full configuration interaction quantum Monte Carlo to a variety of real solids, providing reference many-electron energies that are used to rigorously benchmark the standard hierarchy of quantum-chemical techniques, up to the ‘gold standard’ coupled-cluster ansatz, including single, double and perturbative triple particle–hole excitation operators. We show the errors in cohesive energies predicted by this method to be small, indicating the potential of this computationally polynomial scaling technique to tackle current solid-state problems.

Although density functional theory has been the workhorse of computational materials science for several decades<sup>1</sup>, systematic routes to improve the crucial but approximate exchange–correlation functionals do not exist<sup>2</sup>. In contrast, for molecular systems, a systematic hierarchy of approximate yet highly successful quantum-chemical techniques, such as coupled-cluster theory, has long been established<sup>3</sup>. This hierarchy has not yet been explored in solids, although initial implementations of its lower levels have been encouraging<sup>4–11</sup>. Part of the reason for this is the high computational complexity of quantum-chemical methods. The computational cost grows rapidly with the number of considered electrons  $N$  and with the basis set size  $M$ . Traditional full configuration interaction (FCI), that is, exact diagonalization, has combinatorial scaling and can cope with at most some ten electrons in a small basis; and even coupled-cluster methods, although requiring only a computational time that is polynomial in  $M$  and  $N$ , are extremely expensive. However, recent developments in methodology, as well as the increase in computer power, mean that it is now possible to address their accuracy and applicability in this domain.

In the development of quantum chemistry, FCI has played an invaluable benchmarking role, by providing exact results within a given basis. This has enabled the electron correlation problem to be addressed in isolation from other complicating factors inherent when comparing to experiment<sup>12</sup>. Moreover, FCI enables us to assess the degree to which the electronic wavefunction is dominated by a single determinant, and therefore which systems are likely to be well described by approximations such as many-body perturbation theory and coupled-cluster theory<sup>13</sup>. In solids, however, the absence of FCI means that the accuracy of such approximations cannot be easily gauged, especially in systems where correlations are expected to be strong. Here we provide, for the first time to our knowledge, FCI-quality energies in a range of realistic solids and unambiguously evaluate the accuracy of high-level quantum-chemical methods.

A further motivation for the application of quantum-chemical methods to solids comes from the multitude of recent developments,

which hold the promise to reduce the computational cost beyond our present consideration. These include optimized virtual spaces<sup>14</sup>, explicit correlation<sup>15,16</sup>, exploitation of locality of correlation<sup>5</sup>, and others<sup>17–19</sup>, and should be directly transferable to the solid state. Combining high accuracy with increased efficiency, quantum-chemistry methods hold the promise to routinely bring high accuracy to computational materials science.

## FCIQMC and the quantum-chemical hierarchy

The FCI quantum Monte Carlo (FCIQMC) method has emerged as a tool to calculate energies that are essentially identical to the true correlation energy captured by the basis set, whilst having a significantly lower computational scaling than a traditional brute force diagonalization of the problem (FCI). This makes it ideally suited for a systematic benchmarking of wavefunction-based methods in the solid state<sup>20–25</sup>. This method involves a stochastic sampling of a Slater determinant space constructed from the basis set—a function space of orthonormal antisymmetrized determinants in which the wavefunctions are expressed. This method has previously been applied to molecular systems<sup>20–24</sup> and the homogeneous electron gas<sup>25</sup>, where energies were calculated that compared favourably, or in some cases surpassed in accuracy, those achieved with state-of-the-art diffusion Monte Carlo (DMC) techniques<sup>26</sup>. This provides the confidence to tackle more realistic solid-state systems here.

Although it is possible for DMC to be used as a benchmark for quantum-chemistry methods and vice versa<sup>27</sup>, DMC does not operate in a Slater determinant space, but rather a real space representation of the wavefunction. As such, it would require quantum-chemical calculations to be converged to high accuracy with respect to the basis set size before any meaningful comparisons could be drawn. Instead, by comparing to values obtained in the same Slater determinant basis, robust comparisons between FCIQMC and more approximate diagrammatic methods can be drawn without the need for absolute convergence. More approximate methods can then also be used to

<sup>1</sup>University of Cambridge, Chemistry Department, Lensfield Road, Cambridge CB2 1EW, UK. <sup>2</sup>University of Vienna, Faculty of Physics and Center for Computational Materials Science, Sensengasse 8/12, A-1090 Vienna, Austria.

extrapolate to the infinite basis set limit<sup>14,17,28</sup>. Additionally, DMC requires an approximation for the nodal surface of the wavefunction. Although this error can be made relatively small<sup>29,30</sup>, releasing the nodal surface is notoriously difficult for solids and greatly increases the computational demand. Similarly, another variant of QMC methods, auxiliary-field quantum Monte Carlo (AFQMC), although now operating in a space of Slater determinants and with favourable scaling, requires analogous constraints within the phaseless approximation in order to go to realistic system sizes and avoid transient energy estimates<sup>31,32</sup>.

Here we illustrate the extension of the FCIQMC method to the solid state. The introduction of translational and crystal momentum symmetries, which arise from working with finite simulation cells with periodic boundary conditions, necessitates a change to the ‘walker’ dynamics and ‘initiator’ rules (see below for descriptions). In order to take advantage of these properties, the method is reformulated for complex wavefunctions. The presence of a ‘phase’ problem, rather than a simpler sign problem, is considered. A range of systems, from the extensively studied lithium hydride, to other ionic, covalent and rare gas solids is considered, while benchmarking the established quantum-chemical methods of second-order Møller–Plesset theory (MP2)<sup>33</sup>, coupled-cluster singles and doubles (CCSD)<sup>34</sup>, and the first implementation of perturbative triples (CCSD(T))<sup>35</sup> for periodic systems. These three methods are generally considered to possess a favourable trade-off between accuracy and cost, and are the most widely used of all quantum-chemical methods. Finally, we study the far more complex electronic structure of the charge-transfer solid NiO, to evaluate the ability of these quantum-chemical methods to handle strong correlation effects, where deficiencies are likely to be exposed.

## Sampling in the solid state

The recently developed FCIQMC method<sup>20–24</sup> involves a discrete sampling of the wavefunction by signed ‘walkers’ which stochastically evolve within a Hilbert space of  $N$ -electron Slater determinants as illustrated in Fig. 1. The method converges rapidly with the number of walkers to the FCI limit, and generally the number of walkers required is a tiny fraction of the Hilbert space. This leads to a huge compression in the wavefunction information, while correctly reproducing exact time-averaged properties. The reason that a sufficient number of walkers is required has to do with overcoming the ‘fermion sign problem’ present in the Monte Carlo sampling of any fermionic

wavefunction<sup>36</sup>. The sign problem in this space manifests itself in the presence of a lower-energy state characterized by a combination of the  $\pm\Psi$  degenerate solutions, which if not suppressed, leads to an exponential increase in noise. The growth of this state is controlled in FCIQMC by annihilation between walkers of opposite signs, which stabilizes the wavefunction to one signed solution<sup>20,37</sup>. The discrete space of Slater determinants allows FCIQMC to implement this annihilation exactly, and provided the walker population is large enough, will directly overcome the sign problem without requiring constraints on the wavefunction.

Because walker annihilation can only occur on occupied determinants, it is important to ensure that the newly occupied space remains sign-coherent to the currently sampled wavefunction. This is the rationale behind the ‘initiator’ rules used in the i-FCIQMC method<sup>22</sup>, which is used exclusively in this work, whereby newly occupied determinants must have originated from a determinant with a population greater than a parameter  $n_{\text{add}}$ . By restricting the growth of the occupied space in this way, the walker density and hence annihilation rate is kept high, ensuring that propagation of noise in the system is kept to a minimum. This biases the dynamic in a small way, but rigorously converges onto exact energies of the Hamiltonian as the walker number increases.

The determinants in this work are composed from antisymmetrized products of one-electron orbitals obtained from a prior Hartree–Fock calculation in a large basis of periodic plane waves within the framework of the projector-augmented wave method, as implemented in VASP. If these orbitals are strictly real then the wavefunction  $\Psi$  can also be real<sup>20</sup>. In this work, however, the orbitals are complex Bloch functions, to account for the translational invariance of the potential. With these we can construct many body wavefunctions and use  $k$ -point sampling to ensure convergence, rather than sampling ever larger unit cells to remove finite-size effects.

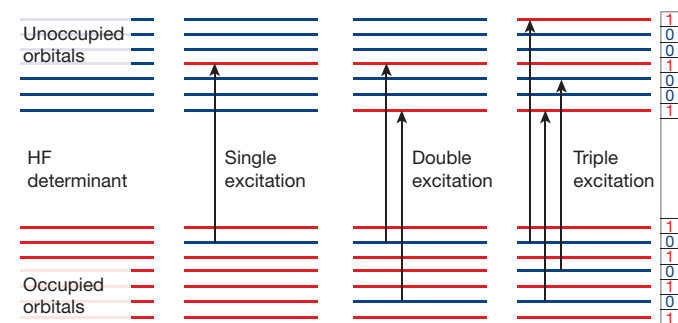
Because it is necessary to correlate between sampled  $k$ -points, the number of explicitly correlated electrons and orbitals increases linearly with the number of sampled  $k$ -points yielding combinatorial scaling in the size of the Hilbert space (essentially exponential with the number of  $k$ -points). However, performing this sampling increases the number of zero Hamiltonian matrix elements between determinants, as crystal momentum must be conserved. By implementing an algorithm to stochastically generate only these momentum-allowed excitations, a saving that grows quadratically with the number of  $k$ -points is achieved, because both the accessible space is reduced, and the magnitude of the time step is increased.

To take advantage of these savings, it is necessary to work with complex orbitals, requiring a complementary set of both ‘real’ and ‘imaginary’ walkers in the FCIQMC dynamic, and a reformulation of the algorithm. The master equations of the FCIQMC method follow naturally from the imaginary-time Schrödinger equation, and are given by

$$-\frac{dN_i}{d\tau} = (H_{ii} - S)N_i + \sum_{j \neq i} H_{ij}N_j \quad (1)$$

where  $N_i$  represents the now complex walker population on determinant  $D_i$ ,  $\tau$  represents imaginary time,  $S$  is a strictly real energy-offset parameter denoted the ‘shift’, which controls population growth, and  $H_{ij} = \langle D_i | \hat{H} | D_j \rangle$  is the many-electron Hamiltonian evaluated between two determinants. In each iteration, for each walker (real and imaginary) on a determinant  $D_i$ , a suitable momentum-allowed excitation,  $D_j$ , is generated. The real ( $\Re$ ) and imaginary ( $\Im$ ) parts of  $H_{ij}$  are considered in turn, and two attempts at generating new walkers on  $D_j$  are stochastically realized. For real parent walkers:

$$p_s^{\Re}(j|i) = \frac{\delta\tau |\Re[H_{ij}]|}{p_{\text{gen}}(j|i)}; \quad \text{sign} = -\text{sign}(\Re[N_i]\Re[H_{ij}]) \quad (2)$$



**Figure 1 | FCIQMC in a nutshell.** The antisymmetric many-electron wavefunction  $\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$  is represented in a space of determinants constructed out of Hartree–Fock (HF) orbitals (occupied as well as unoccupied one electron wavefunctions,  $\phi(\mathbf{r})$ ). Walkers are encoded by a bit string as shown on the right, where each bit corresponds to one orbital. ‘1’ (red) and ‘0’ (black) imply that the orbital is occupied and unoccupied, respectively. Because some Slater determinants (for instance, the Hartree–Fock determinant) have a very high occupation probability, additional bits are reserved to count the signed number of real and imaginary walkers on a determinant. The computational procedure involves selecting a new determinant from an existing walker with a certain sign and transition probability as discussed in the main text. The excitation number of a determinant refers to the numbers of holes which need to be introduced in the Hartree–Fock determinant to generate the determinant.

$$p_s^{\mathcal{I}}(\mathbf{j}|\mathbf{i}) = \frac{\delta\tau |\mathcal{I}[H_{ij}]|}{p_{\text{gen}}(\mathbf{j}|\mathbf{i})}; \quad \text{sign} = -\text{sign}(\Re[N_i]\mathcal{I}[H_{ij}]) \quad (3)$$

and for imaginary parent walkers:

$$p_s^{\Re}(\mathbf{j}|\mathbf{i}) = \frac{\delta\tau |\mathcal{I}[H_{ij}]|}{p_{\text{gen}}(\mathbf{j}|\mathbf{i})}; \quad \text{sign} = -\text{sign}(\mathcal{I}[N_i]\mathcal{I}[H_{ij}]) \quad (4)$$

$$p_s^{\mathcal{I}}(\mathbf{j}|\mathbf{i}) = \frac{\delta\tau |\Re[H_{ij}]|}{p_{\text{gen}}(\mathbf{j}|\mathbf{i})}; \quad \text{sign} = -\text{sign}(\mathcal{I}[N_i]\Re[H_{ij}]) \quad (5)$$

where  $\tau$  is the timestep for the simulation, and  $p_s^{\Re}$  and  $p_s^{\mathcal{I}}$  indicate the probability of creating real and imaginary child walkers, respectively.  $p_{\text{gen}}(\mathbf{j}|\mathbf{i})$  is the probability of generating determinant  $D_j$  from  $D_i$ . After this step, the ‘death’ step is performed for each occupied determinant, with the same death probability for the real and imaginary walkers of  $\delta\tau(H_{ii} - E_0 - S)$  stochastically realized, where  $E_0$  is a reference energy, and  $H_{ii}$  is now strictly real. A final annihilation step occurs every iteration, where real and imaginary walkers are separately considered, and pairs of opposite sign on the same determinant are removed from the simulation.

The value of  $S$  can be used as a strictly real measure of the correlation energy of the problem; however, provided a good overlap of the walker distribution with a reference wavefunction  $D_0$  (generally taken to be the Hartree–Fock determinant) can be found, an averaged projected estimator is often less noisy:

$$E(\tau) = \frac{\langle D_0 | \hat{H} | \Psi(\tau) \rangle}{\langle D_0 | \Psi(\tau) \rangle} \quad (6)$$

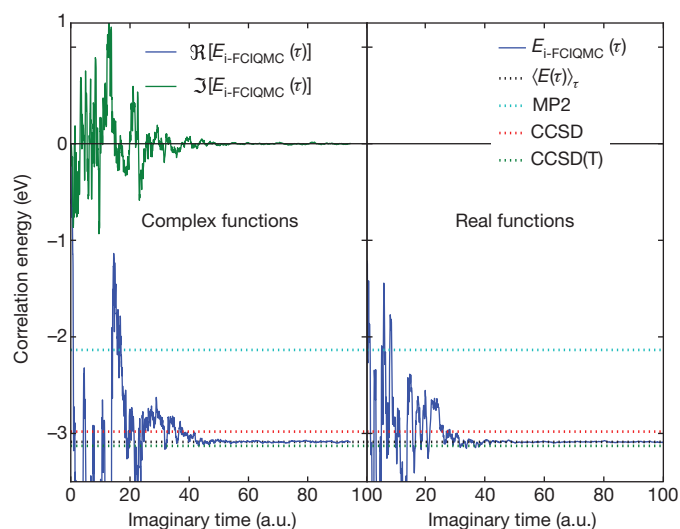
where  $\langle D_0 | \Psi(\tau) \rangle = N_0$ . As opposed to  $S$ ,  $E(\tau)$  is now a complex quantity, where in order to achieve real energies, the imaginary part of the energy must cancel to zero in a non-trivial way. In order to test this, we considered rock-salt-structured LiH sampled using  $2 \times 2 \times 2$   $k$ -points. By choosing all  $k$ -points to lie at the  $\Gamma$ -point or Brillouin zone boundary, it was possible to take linear combinations of the orbitals to give a strictly real basis. This is compared to the complex basis in Fig. 2.

It can be seen that all methods, including FCIQMC, agree exactly between the two bases, and that in the complex basis, the imaginary component of the energy converges to zero within small error bars. Although there is the potential for rotations of the wavefunction in the complex plane, it is observed that the discretization of the wavefunction amplitudes prevents this from happening, and global  $U(1)$  transformations are thus suppressed after an initial arbitrary phase factor is determined. This indicates that there is no more of a sign-issue to overcome with annihilation events than that of the original real formulation of the dynamics.

### Real solids and the accuracy of quantum chemistry

Having established the accuracy and efficiency of the complex FCIQMC walker dynamics, it was initially tested on the most widely studied solid to date, rock-salt-structured LiH (refs 6–9, 11, 17, 27). We first benchmark the accuracy of the MP2, CCSD and CCSD(T) energies for this system, by considering the deviation of each from FCIQMC values. Figure 3 shows an equation of state for a range of volumes, with a  $3 \times 3 \times 3$   $k$ -point sampling, in a minimal basis required to capture any non-dynamic correlation.

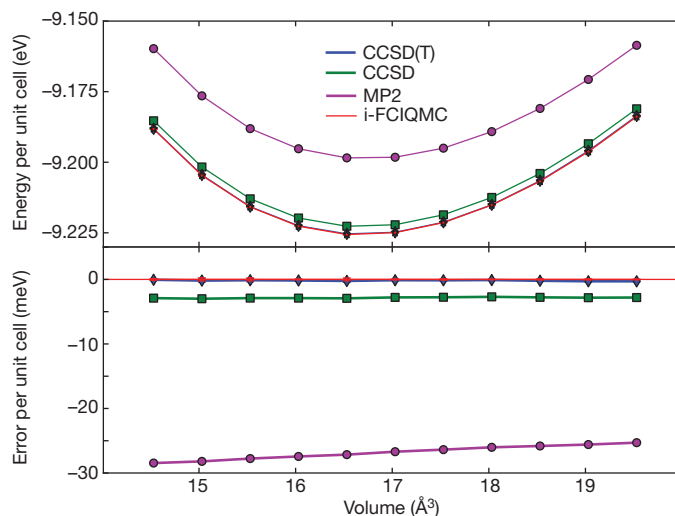
The efficient sampling of the i-FCIQMC method is clear, where the  $3 \times 3 \times 3$   $k$ -point mesh correlates 54 electrons in a space of  $\sim 10^{30}$  determinants, with convergence to the exact energy obtained after



**Figure 2 | Comparison of real and complex i-FCIQMC dynamics for the correlation energy of LiH.** A  $2 \times 2 \times 2$   $\Gamma$ -centred  $k$ -point mesh with 16 electrons, and 40 correlated Hartree–Fock orbitals was employed, at a primitive rock-salt unit cell volume of  $17.03 \text{ \AA}^3$ . Converged energies for 30 million walkers between the two bases agree within small stochastic error bars. The additional overhead for the complex dynamic means that the cost was  $\sim 5$  times that of the real dynamic to converge to equivalent error bars. Also included are MP2, CCSD and CCSD(T) results for comparison.  $\langle E(\tau) \rangle_\tau$  is an imaginary-time average of the projected energy  $E(\tau)$ , taken after a period of equilibration.

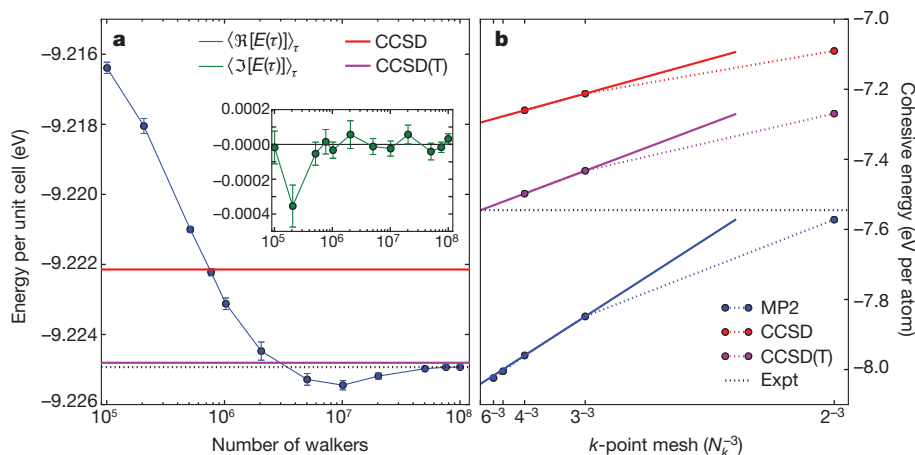
only  $\sim 50$  million walkers, as demonstrated in Fig. 4a. The MP2 values are clearly shifted to higher energies compared to FCIQMC, and because this error changes significantly with volume, the MP2 equilibrium volume and bulk modulus deviate by 3.5% and 6.5%, respectively, from the FCIQMC values.

CCSD energies are virtually parallel to the FCIQMC results, yielding a similar volume and bulk modulus as FCIQMC. Finally, CCSD(T) shows almost exact agreement with FCIQMC in the absolute energies. Overall, these results very much mimic the performance of the standard quantum-chemical hierarchy established for molecular systems and are indicative of their suitability for other similar solids.



**Figure 3 | Equation of state for rock-salt LiH, with the error from the i-FCIQMC value per unit cell at each volume shown below, for MP2, CCSD and CCSD(T).** A  $3 \times 3 \times 3$   $\Gamma$ -centred  $k$ -point mesh was employed, with 54 electrons in 54 Hartree–Fock orbitals. All i-FCIQMC energies were converged with 55 million walkers, and error bars are too small to be seen on the plot ( $\mathcal{O}[0.1 \text{ meV}]$ ).





**Figure 4 | Convergence with respect to the total walker number and  $k$ -point meshes.** **a**, Convergence of the i-FCIQMC energy with total walker number of a  $3 \times 3 \times 3$   $k$ -point sampled LiH calculation. This explicitly correlated 54 electrons in 54 orbitals. Error bars (1 s.d.) were calculated via a Flyvbjerg-Petersen 'blocking' algorithm<sup>43</sup>. The inset shows the imaginary component of the

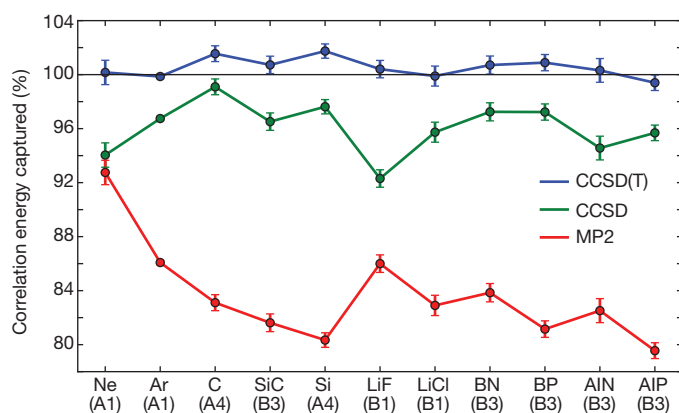
energy on a finer scale. **b**, Convergence of the cohesive energy of diamond with respect to the  $k$ -point mesh using MP2, CCSD and CCSD(T). The solid lines extrapolate to an infinite  $k$ -points mesh using  $3 \times 3 \times 3$  and  $4 \times 4 \times 4$   $k$ -points. MP2 results for  $5 \times 5 \times 5$  and  $6 \times 6 \times 6$   $k$ -points confirm that the finite size error decays as  $1/N_k^3$ , where  $N_k$  is the number of  $k$ -points in each direction.

To cover different bonding situations, in Fig. 5 we consider the relative errors of quantum-chemical methods for the correlation energies of several crystals when compared to FCIQMC. MP2 and CCSD recover between 80% and 98% of the FCIQMC correlation energy for the sample of rare gas, covalent and ionic solids. The dependence of the relative errors on the various systems is most pronounced in the case of MP2 theory. This is not unexpected and reflects the limitations of low-order perturbation theory. MP2 is more accurate for wide-gap insulators, such as Ne and LiF, than for semiconductors with a smaller gap, like Si and AlP. In contrast, CCSD(T) is shown to give a balanced description across the different systems and is in error by at most 2%.

What accuracy, therefore, can one expect from converged CCSD(T) calculations for solids? To answer this question, we computed cohesive energies. These are extremely demanding quantities for any theory, because the correlation effects in solids differ markedly from atoms, potentially leading to large errors in the prediction of the cohesive energy. Furthermore, attention needs to be paid to finite-size scaling (that is,  $k$ -point sampling). For the present study we have limited our attention to four solids, rock-salt LiH, diamond, zinc-blende BN and

AlP, expecting similar results for the other materials. We have used the progressive downsampling technique<sup>14,28</sup>, employing  $k$ -point meshes of up to  $4 \times 4 \times 4$ . Figure 4b shows for diamond that the MP2 cohesive energy converges as  $1/N_k^3$ , where  $N_k$  is the number of  $k$ -points used to sample the Brillouin zone in each direction. By fitting to the MP2, CCSD and CCSD(T) energies for  $3 \times 3 \times 3$  and  $4 \times 4 \times 4$   $k$ -points, we can extrapolate to infinitely dense  $k$ -point meshes. The remaining finite size error on the correlation energy is expected to be less than 20 meV per atom for the considered systems. The computational cost is of the order of 25,000 CPU (central processing unit) hours for diamond, with results reported in Table 1.

As anticipated (owing to the established agreement with i-FCIQMC in smaller bases and supercells), the CCSD(T) results are all in almost exact agreement with the experimental cohesive energy corrected for the zero-point energy<sup>38</sup>. The MP2 cohesive energies generally show substantial error compared to experiment<sup>45,9,17</sup>. MP2 severely underestimates the correlation energy of atoms, and while it also underestimates the correlation energy of the solids as shown in Fig. 5, the perturbative nature of the theory leads to less underestimation of the correlation energy in solids so that the cohesive energy is often—but not always—overestimated compared to the experimental value. This makes MP2 an unreliable method for solids, especially for calculation of cohesive energies. CCSD and CCSD(T), on the other hand, are far more consistent. Although the absolute errors in the CCSD cohesive energies are still quite sizeable, the cohesive energies are always underestimated. This results from the CCSD correlation energy in the solids being always too small (Fig. 5) whereas in atoms it is generally a good approximation. On adding the perturbative triples correction to the CCSD, the cohesive energies dramatically improve, to an error of only 0.03 eV. As for the absolute correlation energies in solids, the (T) correction over-compensates, leading to too-negative correlation energies, and hence an overestimation of the cohesive energies, albeit only mildly so. We finally note that our results indicate that the residual errors are dominated by correlation errors in the

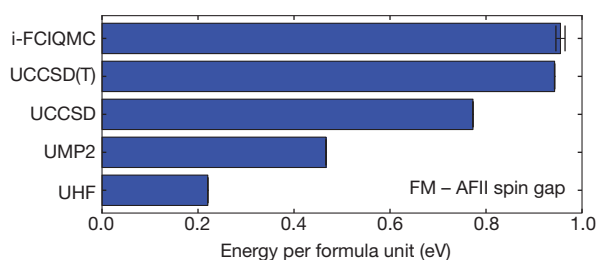


**Figure 5 | Relative errors compared to i-FCIQMC in a range of solids.**  $2 \times 2 \times 2$   $k$ -point sampling was employed, correlating 64 electrons in 64 orbitals, and 80 electrons in 72 orbitals, for LiF and LiCl, respectively. Lattice structures are given as A1 = f.c.c., A4 = diamond, B1 = rock salt, B3 = zinc blende, while lattice constants can be found in ref. 44. Error bars of each method are derived from the random errors of the i-FCIQMC values they are compared to. The i-FCIQMC calculations employed  $\sim 10,000$  walkers on the Hartree-Fock determinant resulting in between 2 million and 300 million total walkers, depending on the system.

**Table 1 | Cohesive energies evaluated within the quantum-chemical hierarchy**

	LiH	C	BN	AlP
MP2	-2.386	-8.039	-7.149	-4.629
CCSD	-2.454	-7.295	-6.572	-4.107
CCSD(T)	-2.483	-7.545	-6.782	-4.347
Experiment	-2.487	-7.545	-6.758	-4.322

Energies are given in eV per atom. Experimental values have been corrected for zero-point vibration energies.



**Figure 6 | Energies of the NiO spin gap, calculated at the level of UHF, UMP2, UCCSD, UCCSD(T) and i-FCIQMC in an UHF basis for both spin species.** Correlating the  $3d$  and  $4s$  of the Ni, and  $2s$  and  $2p$  of the O using a rhombohedral supercell with a lattice constant of  $4.19 \text{ \AA}$  containing 2 formula units, resulted in a sampling of 32 electrons, in 38 Hartree–Fock orbitals. 300 million walkers were required for the i-FCIQMC calculation of the AFII state, requiring  $\sim 50,000$  CPU hours, while only 100 million walkers were needed for the FM state, with random error bar (1 s.d.) shown, and remaining error estimated to be less than 30 meV.

solid, whereas correlation energies for atoms are essentially converged at the CCSD(T) level.

The conclusion is that the accuracy of CCSD(T) has been established for solids to be of the order of  $0.03 \text{ eV}$  or  $1 \text{ kcal mol}^{-1}$ . By contrast, the most widely used density functional (PBE<sup>39</sup>) exhibits a mean-absolute error of  $0.15\text{--}0.2 \text{ eV}$  for a similar range of insulating solids<sup>40</sup>.

### Towards strong correlation

An important and open question is the domain of applicability of CCSD and CCSD(T) as stronger correlation effects set in. An initial indication of this can be found by comparing the performance of the methods for the spin gap between the ground antiferromagnetic state (AFII) and the ferromagnetic state (FM) of nickel oxide in a rhombohedral unit cell. This is a classic charge-transfer insulator, and it is expected to have strong correlation effects for at least one of the states, and as such, a balanced description for the calculation of the spin gap is expected to provide a stern test.

The results for the FM–AFII spin gap are given in Fig. 6 and illustrate the systematic convergence of the quantum-chemical hierarchy with respect to the correlation treatment of the system. Analysis of the FCIQMC wavefunction indicates that the ground state (AFII) is more strongly correlated and multiconfigurational, with a normalized Hartree–Fock weight of only 0.69, compared to 0.86 for the FM state. This lack of a dominant single reference leads to errors in MP2 of over 50%, while CCSD is still 19% in error compared to i-FCIQMC. Despite this, the qualitative behaviour of the quantum chemical hierarchy remains intact, with CCSD(T) providing an excellent approximation to the exact result. Results from unrestricted Hartree–Fock (UHF) calculations of NiO converged to the thermodynamic limit only capture 17% of the experimental spin gap between the two states<sup>41</sup>, inferred from neutron scattering<sup>42</sup>. This agrees well with our restricted  $k$ -point sampled system, where the UHF captures 23% of the spin gap compared to FCIQMC results.

### Conclusions and outlook

We have shown that FCI-quality correlation energies can be obtained for solid-state systems using an extension of the i-FCIQMC method to complex wavefunctions. We have demonstrated that the standard quantum-chemical hierarchy of increasingly accurate polynomially scaling methods holds for a range of materials, including rare gas, ionic and covalent solids and the charge-transfer insulator NiO. As explicitly shown by the cohesive energies of LiH, C, AlP and BN, CCSD(T) is very accurate for the solid state, surpassing  $1 \text{ kcal mol}^{-1}$  accuracy in reproducing experimental results. Considering the proven reliability of CCSD(T) for molecules, we expect a similar precision for insulators and semiconductors in general, with metals

possibly requiring further methodological improvements. In combination with recent developments to reduce the computational cost—for instance with explicit inclusion of the cusp condition for the many-electron wavefunction—as well as further technical, algorithmic and methodological advances, the accuracy of FCIQMC and the quantum-chemistry methods will be brought routinely to solid-state physics and computational materials science. We are witnessing a slow but steady change of our computational paradigm.

Received 17 July; accepted 7 November 2012.

Published online 19 December 2012.

- Kohn, W. Nobel lecture: Electronic structure of matter—wave functions and density functionals. *Rev. Mod. Phys.* **71**, 1253–1266 (1999).
- Cohen, A. J., Mori-Sánchez, P. & Yang, W. Challenges for density functional theory. *Chem. Rev.* **112**, 289–320 (2012).
- Pople, J. A. Nobel lecture: Quantum chemical models. *Rev. Mod. Phys.* **71**, 1267–1274 (1999).
- Müller, C. & Paulus, B. Wavefunction-based electron correlation methods for solids. *Phys. Chem. Chem. Phys.* **14**, 7605–7614 (2012).
- Maschio, L. *et al.* Fast local-MP2 method with density-fitting for crystals. I. Theory and algorithms. *Phys. Rev. B* **76**, 075101 (2007).
- Casassa, S., Halo, M., Maschio, L., Roetti, C. & Pisani, C. Beyond a Hartree–Fock description of crystalline solids: the case of lithium hydride. *Theor. Chem. Acc.* **117**, 781–791 (2007).
- Ayala, P., Kudin, K. & Scuseria, G. Atomic orbital Laplace-transformed second-order Møller–Plesset theory for periodic systems. *J. Chem. Phys.* **115**, 9698–9707 (2001).
- Usvyat, D. *et al.* Approaching the theoretical limit in periodic local MP2 calculations with atomic-orbital basis sets: the case of LiH. *J. Chem. Phys.* **134**, 214105 (2011).
- Marsman, M., Grüneis, A., Paier, J. & Kresse, G. Second-order Møller–Plesset perturbation theory applied to extended systems. I. Within the projector-augmented-wave formalism using a plane wave basis set. *J. Chem. Phys.* **130**, 184103 (2009).
- Stoll, H. & Doll, K. Approaching the bulk limit with finite cluster calculations using local increments: the case of LiH. *J. Chem. Phys.* **136**, 074106 (2012).
- Nolan, S. J., Gillan, M. J., Alfè, D., Allan, N. L. & Manby, F. R. Calculation of properties of crystalline lithium hydride using correlated wave function theory. *Phys. Rev. B* **80**, 165109 (2009).
- Knowles, P. J. & Handy, N. C. A new determinant-based full configuration-interaction method. *Chem. Phys. Lett.* **111**, 315–321 (1984).
- Larsen, H., Olsen, J., Jørgensen, P. & Christiansen, O. Full configuration interaction benchmarking of coupled-cluster models for the lowest singlet energy surfaces of  $\text{N}_2$ . *J. Chem. Phys.* **113**, 6677–6686 (2000).
- Grüneis, A. *et al.* Natural orbitals for wave function based correlated calculations using a plane wave basis set. *J. Chem. Theory Comput.* **7**, 2780–2785 (2011).
- Shiozaki, T. & Hirata, S. Communications: Explicitly correlated second-order Møller–Plesset perturbation method for extended systems. *J. Chem. Phys.* **132**, 151101 (2010).
- Hättig, C., Klopper, W., Köhn, A. & Tew, D. P. Explicitly correlated electrons in molecules. *Chem. Rev.* **112**, 4–74 (2012).
- Nolan, S. J., Bygrave, P. J., Allan, N. L. & Manby, F. R. Comparison of the incremental and hierarchical methods for crystalline neon. *J. Phys. Condens. Matter* **22**, 074201 (2010).
- Kwee, H., Zhang, S. & Krakauer, H. Finite-size correction in many-body electronic structure calculations. *Phys. Rev. Lett.* **100**, 126404 (2008).
- Manby, F. R., Stella, M., Goodpaster, J. D. & Miller, T. F. III. A simple, exact density-functional-theory embedding scheme. *J. Chem. Theory Comput.* **8**, 2564–2568 (2012).
- Booth, G. H., Thom, A. J. W. & Alavi, A. Fermion Monte Carlo without fixed nodes: a game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **131**, 054106 (2009).
- Booth, G. H. & Alavi, A. Approaching chemical accuracy using full configuration-interaction quantum Monte Carlo: a study of ionization potentials. *J. Chem. Phys.* **132**, 174104 (2010).
- Cleland, D., Booth, G. H. & Alavi, A. Survival of the fittest: accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.* **132**, 041103 (2010).
- Cleland, D. M., Booth, G. H. & Alavi, A. A study of electron affinities using the initiator approach to full configuration interaction quantum Monte Carlo. *J. Chem. Phys.* **134**, 024112 (2011).
- Booth, G. H., Cleland, D., Thom, A. J. W. & Alavi, A. Breaking the carbon dimer: the challenges of multiple bond dissociation with full configuration interaction quantum Monte Carlo methods. *J. Chem. Phys.* **135**, 084104 (2011).
- Shepherd, J. J., Booth, G., Grüneis, A. & Alavi, A. Full configuration interaction perspective on the homogeneous electron gas. *Phys. Rev. B* **85**, 081103 (2012).
- López Ríos, P. *et al.* Inhomogeneous backflow transformations in quantum Monte Carlo calculations. *Phys. Rev. E* **74**, 066701 (2006).
- Binnie, S. J. *et al.* Bulk and surface energetics of crystalline lithium hydride: benchmarks from quantum Monte Carlo and quantum chemistry. *Phys. Rev. B* **82**, 165431 (2010).
- Ohnishi, Y.-y. & Hirata, S. Logarithm second-order many-body perturbation method for extended systems. *J. Chem. Phys.* **133**, 034106 (2010).

29. Nemecek, N., Towler, M. D. & Needs, R. J. Benchmark all-electron ab initio quantum Monte Carlo calculations for small molecules. *J. Chem. Phys.* **132**, 034111 (2010).
30. Morales, M. A., McMinis, J., Clark, B. K., Kim, J. & Scuseria, G. E. Multideterminant wave functions in quantum Monte Carlo. *J. Chem. Theor. Comput.* **8**, 2181–2188 (2012).
31. Al-Saidi, W. A., Zhang, S. & Krakauer, H. Bond breaking with auxiliary-field quantum Monte Carlo. *J. Chem. Phys.* **127**, 144101 (2007).
32. Purwanto, W., Krakauer, H. & Zhang, S. Pressure-induced diamond to  $\alpha$ -tin transition in bulk silicon: a quantum Monte Carlo study. *Phys. Rev. B* **80**, 214116 (2009).
33. Møller, C. & Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **46**, 618–622 (1934).
34. Čížek, J. On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods. *J. Chem. Phys.* **45**, 4256–4266 (1966).
35. Raghavachari, K., Trucks, G. W., Pople, J. A. & Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **157**, 479–483 (1989).
36. Troyer, M. & Wiese, U.-J. Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Phys. Rev. Lett.* **94**, 170201 (2005).
37. Spencer, J. S., Blunt, N. S. & Foulkes, W. M. The sign problem and population dynamics in the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **136**, 054110 (2012).
38. Schimka, L., Harl, J. & Kresse, G. Improved hybrid functional for solids: the HSEsol functional. *J. Chem. Phys.* **134**, 024116 (2011).
39. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
40. Csonka, G. I. *et al.* Assessing the performance of recent density functionals for bulk solids. *Phys. Rev. B* **79**, 155107 (2009).
41. Cora, F. *et al.* in *Density Functional Theory in Inorganic Chemistry, Structure and Bonding* (eds McGrady, J. & Kaltsoyannis, N.) 171–232 (Springer, 2004).
42. Hutchings, M. & Samuelsen, E. Measurement of spin-wave dispersion in NiO by inelastic neutron-scattering and its relation to magnetic properties. *Phys. Rev. B* **6**, 3447–3461 (1972).
43. Flyvbjerg, H. & Petersen, H. G. Error estimates on averages of correlated data. *J. Chem. Phys.* **91**, 461–466 (1989).
44. Grüneis, A., Marsman, M. & Kresse, G. Second-order Møller–Plesset perturbation theory applied to extended systems. II. Structural and energetic properties. *J. Chem. Phys.* **133**, 074107 (2010).

**Acknowledgements** G.H.B. acknowledges support from Trinity College, Cambridge. A.G. acknowledges an APART fellowship from the Austrian Academy of Sciences. G.K. acknowledges support from the Austrian Science fund (FWF) within the SFB ViCoM (F41). A.A. acknowledges the support of the EPSRC through grants EP/I014624/1 and EP/J003867/1. The authors thank J. Spencer and A. Thom for technical contributions and discussions. Computer time on the Vienna Scientific Cluster (VSC) and HECToR (under the DEISA Extreme Computing Initiative) are acknowledged.

**Author Contributions** G.H.B. and A.G. contributed equally to this work. G.H.B. and A.A. developed the FCIQMC method, G.H.B. wrote the computer code, and A.G. developed the quantum-chemical methods in VASP. G.K. led the VASP project, and A.A. led the FCIQMC project.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.H.B. ([ghb24@cam.ac.uk](mailto:ghb24@cam.ac.uk)) or A.G. ([andreas.grueneis@univie.ac.at](mailto:andreas.grueneis@univie.ac.at)).