JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Taming the First-Row Diatomics: A Full Configuration Interaction Quantum Monte Carlo Study

Deidre Cleland, George H. Booth, Catherine Overy, and Ali Alavi*

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.

ABSTRACT: The initiator full configuration interaction quantum Monte Carlo (*i*-FCIQMC) method has recently been developed as a highly accurate stochastic electronic structure technique. It has been shown to calculate the exact basis-set ground state energy of small molecules, to within modest stochastic error bars, using tractable computational cost. Here, we use this technique to elucidate an often troublesome series of first-row diatomics consisting of Be_2, C_2, CN, CO, N_2, NO, O_2, and F_2. Using *i*-FCIQMC, the dissociation energies of these molecules are obtained almost entirely to within chemical accuracy of experimental results. Furthermore, the *i*-FCIQMC calculations are performed in a relatively black-box manner, without any a priori knowledge or specification of the wave function. The size consistency of *i*-FCIQMC is also demonstrated with regards to these diatomics at their more multiconfigurational stretched geometries. The clear and simple *i*-FCIQMC wave functions obtained for these systems are then compared and investigated to demonstrate the dynamic identification of the dominant determinants contributing to significant static correlation. The appearance and nature of such determinants is shown to provide insight into both the *i*-FCIQMC algorithm and the diatomics themselves.

INTRODUCTION

Despite their simple appearance, diatomics constructed from first row atoms are notoriously difficult to describe from first principles, due to their varied and sometimes complex electronic structures. As such, they are an excellent testing ground for new electronic structure methods.[1−5] In this paper we investigate a series of such diatomic molecules, at both equilibrium and stretched geometries, which present contrasting challenges with regard to dominating correlation effects.[6] The method applied is a new stochastic method we have developed called full configuration interaction quantum Monte Carlo[7] (FCIQMC), together with its vastly more efficient *initiator* extension[8] (*i*-FCIQMC). These techniques can yield, with tractable computational cost, the full CI energy of molecular Hamiltonians for systems that are too large for conventional FCI methods. Being stochastic methods, these techniques incur random errors, which can be shown to be controllable and modest in magnitude, such that these methods rival in accuracy the most accurate quantum chemical techniques available, while being "black-box" and applicable across a broad range of systems. As such, the (*i*-)FCIQMC methods have considerable promise. In previous studies, we have applied these techniques to the calculation of ionization potentials,[9] electron affinities,[10] and binding curves of C_2,[11] using Dunning correlation consistent basis sets[12] up to aug-cc-pVQZ. These energies can be extrapolated to the complete basis set limit with standard techniques.[13] FCIQMC has also been applied to the homogeneous electron gas,[14] as a first application to condensed matter systems. In this study, we continue our investigations and applications of this new methodology by considering the complex systems which the first row dimers present.

It hardly needs stating that ab initio quantum chemistry is fundamentally concerned with the solutions to the electronic Schrödinger equation, which it seeks to approximate using basis-set expansions together with many-body approximations. Regarding the latter, a vast body of methods exist, including CI methods,[15−30] coupled cluster (CC) methods,[31−41] multireference CC,[42−50] multireference CI and perturbation theory,[51−59] density matrix renormalization group,[60−65] variational density matrix methods,[66,67] and CEEIS methods.[68−72] This list is by no means complete. The ultimate goal of these techniques is to approximate the full CI energy, namely the exact energy of the Schrödinger equation within the given basis set. Coupled with basis set extrapolation methods,[5,13,73−79] or more sophisticated R12/F12 methods[80−93] to remove basis set errors, the full CI energy should approximate the true ground state energy of the Schrödinger equation.

The main goal of this vast and growing body of work has been to reach "chemical accuracy",[94] which depending on the system, has necessitated the different approaches. For example, systems dominated by strong correlation require factorial-scaling multireference methodologies which "single-reference" systems do not. And yet, accuracy should not be the only aim of a theory or method. Robustness and a "black box" approach are also important. The demonstration of convergence onto experiment in the absence of any adjustable parameters is highly desirable, the success of model chemistries notwithstanding.[95−99] But above all, a method should be *simple* enough to enable *insight* to be extracted from the answers. The problem is that "high accuracy" often equates with "high-level theories", the interpretation of which, however, are not simple. The attempt to square this circle, to develop a powerful yet simple method, underlies the theoretical motivation for the *i*-FCIQMC method applied in this study. By demonstrating that *i*-FCIQMC indeed does satisfy both accuracy and simplicity, we hope to encourage further development and use of this method in larger and more complex systems yet to be studied.

Received: June 18, 2012
Published: September 13, 2012

The subset of diatomics investigated here is made up of the homonuclear species Be_2, C_2, N_2, O_2 and F_2, as well as the heteronuclear molecules, CN, CO, and NO. These systems encompass a diverse set of electron correlation effects (such "strong correlation", "dynamical correlation", and chemically significant core−valence correlation) that arise in strong multiple bonds (C_2, N_2, CN, NO), in long bonds (F_2), in stretched bonds, in extremely weak bonds (Be_2), and in open-shell systems (O_2, NO). Such effects are known to present challenges for ab initio electronic structure methods. In particular, the dissociation energy (D_e) of these diatomics is often the most difficult spectroscopic property to converge. This is because an accurate description of each dissociation across the series requires a computational method that can provide a balanced and consistent representation of the varying characteristics, as well as the dissociated atoms. Furthermore, the method must be efficient enough to allow each system to be considered in a basis set of substantial size.

As previously mentioned, FCI provides the exact solution within a given basis set, and thus represents the level of accuracy which quantum chemical methods aim to achieve.[94] Unfortunately, FCI itself is very computationally demanding and is therefore unable to treat most of the systems considered here. While truncated CI methods, and more generally subspace diagonalization methods, exist to reduce the computational cost of such calculations, these often suffer from size consistency errors and the computational cost may still scale unfavorably with the size of the system.

A crucial distinction between subspace diagonalizations and the present *i*-FCIQMC methods[7,8] is that in the latter *the entire Slater determinant space is available to be sampled during a simulation*, each determinant being sampled by a probability proportional to the absolute value of its FCI coefficient in the ground-state wave function. This has some desirable implications: on the one hand, one does not need to store vectors of length equal to the FCI space, which is the primary bottleneck as far as conventional FCI calculations are concerned. What is stored, instead, is a dynamical list of "walkers", the number of which is found to be much much smaller that the size of the underlying Hilbert space. On the other hand, *expectation values* such as energies are computed by post-equilibration averaging over iterations of imaginary time. These expectation values approach their exact value, even though typically only a tiny fraction of the full Hilbert space of the problem would have been sampled during the simulation. Clearly, this is a highly attractive feature of the method, and one purpose of the present paper is to demonstrate this at work. We do so in several ways: for small systems, for which an FCI wave function can be stored in its entirety, time-averaged histograms of the population of walkers visiting each determinant can be stored and subsequently compared with the exact FCI solution. For larger calculations, where histogramming the entire wave function is no longer practical, we show that our *i*-FCIQMC energies for stretched molecules are fully size-consistent. This is an extremely stringent test: truncated CI methods fail this test, whereas full CI does not. The implication is that the *i*-FCIQMC sampling of the FCI space faithfully reflects the ground-state wave function. This leads us to a second very attractive feature of *i*-FCIQMC. Since the populations of walkers directly sample the FCI coefficients, we can simply "read-off" the FCI coefficients for the significantly populated determinants, giving a direct feel for the importance of various configurations in the FCI wave function.

The work presented here begins with a brief recap of the *i*-FCIQMC method applied during this study. At this point, the differences between *i*-FCIQMC and subspace diagonalization methods are discussed through direct comparison of the *i*-FCIQMC energy and wave function to those obtained from an exact diagonalization in the same space. The accuracy of *i*-FCIQMC is then evaluated by calculating the dissociation energies of the above series of dimers and comparing these to both experimental results and values given by some alternative theoretical techniques. It is shown that with tractable computational cost, dissociation energies may be obtained that are generally within chemical accuracy of the experimental result, once basis set errors have been taken into account. Finally, the size consistency of the method is tested, and dynamic identification of dominant determinants during an *i*-FCIQMC simulation is used to provide some chemical insight into these first row diatomics.

(*i*-)FCIQMC. Given the many-electron Schrödinger equation for a molecule:

$$\hat{H}\Psi = E\Psi \tag{1}$$

the FCI wave function represents the exact solution in a given basis set:

$$\Psi = \sum_i C_i |D_i\rangle \tag{2}$$

Here $\{|D_i\rangle\}$ represents the full Hilbert space of determinants, constructable by distributing N electrons in M spatial orbitals. This FCI wave function is obviously highly desirable. Mathematically, the FCI energy is variational, orbitally invariant, spin-pure, and size-consistent. From a chemical perspective, FCI treats dynamical and nondynamical correlation on an equal footing and thus provides a well-balanced description of electron correlation effects when used in conjunction with a suitably large basis set. Furthermore, FCI does not require the user to have any prior knowledge of the wave function. Only the basis-set must be chosen, the errors from which decay systematically, making FCI a "black-box" tool. Of course, the FCI wave function is generally unachievable in practice due to the combinatorial dimension of the problem. The number of determinants in the FCI space is given by

$$N_{\mathrm{FCI}} = \binom{M}{N_\alpha}\binom{M}{N_\beta} \tag{3}$$

where N_α (N_β) is the number of α (β) electrons in the system ($N = N_\alpha + N_\beta$). Therefore, N_{FCI} grows extremely rapidly with N and M, and variational optimization of all C_i coefficients quickly becomes infeasible. To date, the largest space that has been handled by FCI involved 10^{10} determinants.[100] Unfortunately, many problems of chemical interest require vastly larger values of N_{FCI}. For instance, even the calculations performed here involve up to 5×10^{19} determinants.

Subspace diagonalization techniques attempt to reduce the computational cost of FCI, while maintaining its accuracy. The general aim of subspace diagonalization is to decide upon a subset of determinants which is in some sense "optimal". The Hamiltonian is then constructed and diagonalized in this subspace, and the lowest eigenvalue taken as the best variational estimate to the exact FCI ground state energy. Excitation level truncation and complete-active-space (CAS) based methods are some of the most common. As an alternative to this a priori selection of configurations, a class of subspace diagonalization

methods also exist which rely on heuristic iterative algorithms to dynamically search the FCI space for the optimal selection of determinants.[101−103]

Although the procedures to obtain a determinant subset vary substantially, the final aim of these subspace diagonalization methods remains the same. If a determinant which is neglected in the final subspace has precisely zero weight in the FCI wave function, then the absence of this function from the subspace will not affect the energy, and no approximation has been made. However, due to the variationality of any truncated CI result, a neglected nonzero weighted determinant must serve to increase the energy estimate from the true FCI energy, as the flexibility in the description of the final wave function has been reduced. Moreover, this error to the FCI energy is not size-consistent, and thus care must be taken in relative energy contributions when separate fragments are produced. Furthermore, the explicit diagonalization of the Hamiltonian in this subspace still requires storage of at least two vectors, which are the length of the subspace being considered, and so these methods may still become rather computationally demanding. However, in many cases, the FCI wave function is sparse enough, and the selection of determinants sufficiently large and well-chosen, such that very good agreement is found between the results from a subspace and full diagonalization results. The success of this approach has meant that truncated, and especially CAS-based methods have become the first-choice for many investigations into strongly correlated molecular systems.

FCIQMC takes the different approach of projector Monte Carlo methods and represents the wave function in terms of a set of discretized "walkers". One can think of a walker as a particle that carries a sign (± 1) and that "lives" on a determinant, contributing a unit to the amplitude of the determinant it is located on. The purpose of FCIQMC is to stochastically project these walkers according to a set of rules derived from the imaginary-time Schrödinger equation:

$$\frac{\partial \Psi}{\partial \tau} = -(\hat{H} - E_S)\Psi \tag{4}$$

These rules are designed so that the *long-time averaged distribution of the walkers becomes proportional to the FCI wave function*. According to this FCIQMC algorithm, a walker occupying determinant D_i spawns a new walker onto D_j with a probability proportional to $|H_{ij}|$. The sign of the child being spawned is then given by $-\text{sign}(H_{ij})$. The parent walker dies with a probability proportional to $|H_{ii} - E_S|$. (E_S is an energy offset, referred to as the "shift", which can be periodically adjusted to maintain a roughly constant population of walkers). Pairs of walkers of different sign occupying the same determinant are then removed in a series of annihilation events. These annihilations are critical to the emergence of a sign-coherent wave function: that is to say a wave function where the sign of every determinant is correct, up to an overall sign. The precise rules of FCIQMC are described in detail in ref 7.

It turns out that the key parameter which governs the success of FCIQMC is the number of walkers in the simulation (N_w). When N_w is below a system-dependent value, the distribution of walkers fails to settle down on the FCI wave function. This is because a certain density of walkers is required in the population dynamics before annihilation events become sufficiently probable for sign-coherence to emerge.[7,104] On the other hand, above this critical walker number, convergence to the FCI energy is guaranteed. It was found that, in many realistic problems, the required number of walkers scaled with N_{FCI}, obviating any great advantage over traditional FCI methods. Nevertheless, FCIQMC has been used to compute FCI energies for some hitherto unobtainable systems.[7,11]

The initiator method *i*-FCIQMC was developed to counter this problem, and it does so extremely effectively, by making a very small change to the algorithm. An additional *survival of the fittest* criterion is stipulated for walkers spawned onto empty determinants (i.e., determinants not already occupied at that time-step). Such a newly spawned walker is allowed to survive only if its parent resides on a determinant that has been designated an "initiator". Determinants become initiators when their walker population is considered to have a well-established sign, in that it exceeds a critical number n_a.

It should be noted that this initiator rule becomes irrelevant in the limit of a large number of walkers, since all determinants eventually become occupied. In this large N_w limit, *i*-FCIQMC is identical to FCIQMC, and therefore to FCI, irrespective of the value of n_a. However, the *i*-FCIQMC algorithm has been found to achieve energies that are within stochastic errors of the FCI results well before this limit is realized, and with substantially reduced computational cost, even compared to the original FCIQMC method.[8,10,11]

While there will be an optimal set of parameters for a given *i*-FCIQMC simulation, providing the energy is converged with respect to the number of walkers in the system, this method has been found to reliably converge onto the FCIQMC and thus the FCI result. For this reason, the *i*-FCIQMC method is relatively black-box in nature. This refers to the fact that without any orbital specification, extrapolation, selection, or optimization of a trial wave function, or prior knowledge of the final solution, and with minimal input from the user, the errors in converged *i*-FCIQMC energies may be isolated as systematically reducible basis set incompleteness errors.

With the inclusion of an additional initiator spawning criterion, which disallows spawning between some determinant pairs, the *i*-FCIQMC algorithm is effectively applying a slightly modified Hamiltonian operator, in which some of the **H** matrix elements have been set to zero:

$$H_{ij}(\tau) = 0 \qquad \text{if } D_i \notin \{\text{initiators}\} \text{ and } N_j(\tau) = 0 \tag{5}$$

This essentially places restrictions on the determinant space accessible to the *i*-FCIQMC wave function, the extent of which will depend on the number of determinants that have been made initiators (N_{init}) at a particular point in time and, thus, the number of walkers in the system. Furthermore, the affected **H** elements evolve with the distribution of walkers over imaginary time, and as such, these restrictions do not generally behave like a typical truncation of the space (as discussed later). The above modifications can nonetheless introduce a systematic 'initiator error' in the *i*-FCIQMC energy. While such an error can always be reduced by increasing N_w, the number of walkers required to eliminate it is system dependent and must be considered on a case-by-case basis.

There are two methods for computing the energy. The first is the "projected energy":

$$\langle E_{proj}\rangle_\tau = \frac{\langle D_0|\hat{H}|\Psi\rangle}{\langle D_0|\Psi\rangle} \tag{6}$$

$$= \sum_j \langle D_0|\hat{H}|D_j\rangle \frac{\langle N_j\rangle_\tau}{\langle N_0\rangle_\tau} \tag{7}$$

where $|D_0\rangle$ is the reference determinant (usually taken to be the Hartree−Fock (HF) determinant in this study). It is clear that $\langle E_{proj}\rangle_\tau$ is dependent on the population of walkers on $|D_0\rangle$ and the determinants connected to it, namely single and double excitations from the HF determinant. Furthermore, because $\langle N_0\rangle_\tau$ appears in the denominator, it is necessary for the fluctuations in N_0 to be small, to allow for a meaningful estimate of E_{proj}. This in turn requires a large number of walkers to accumulate on D_0, since the fluctuations in N_0 go as ${N_0}^{1/2}$. In practice, we find that $\langle N_0\rangle_\tau \approx 10^4$ walkers is necessary for acceptable fluctuations in $\langle E_{proj}\rangle_\tau$, and if it is affordable, we grow the overall population of walkers so that the reference occupation comfortably exceeds this. This forms the basis with which we select our target population.

The second measure of the energy comes from $\langle E_S\rangle_\tau$, which is accumulated as a time-average of E_S in variable-shift mode. E_S is periodically updated (every 10−20 time-steps) to stabilize the walker growth as detailed in ref 7. Note that, unlike E_{proj}, E_S is determined by the *total* population of walkers. This has two implications. First, the statistical fluctuations in E_S and E_{proj} are generally uncorrelated, providing an independent measure of the energy as the simulation proceeds. As such we have an internal check: a necessary criterion for convergence is that the two measures of energy must agree, to within statistical error bars. The second implication is that, since E_S depends on the distribution of walkers throughout the space, it is a genuinely multireference measure of the energy. As such, for more multireference problems (e.g., stretched geometries), it is easier to obtain small statistical errors using E_S; whereas for systems dominated by a single determinant, E_{proj} can be measured with smaller statistical uncertainty.

Efficient Sampling of the FCI Wave Function with *i*-FCIQMC. An advantage of the stochastic *i*-FCIQMC algorithm is that only the list of instantaneously occupied determinants (N_{occ}) must be stored during a particular iteration. The set of determinants that become occupied at any stage of the simulation defines the space that contributes to the *i*-FCIQMC energy. Clearly as N_w increases, the walkers spread throughout a larger fraction of N_{FCI}. The effect of the initiator adaptation is to reduce the spread of noisy low-weighted walker populations within the wave function, maintain a high annihilation rate, and to promote spawning within a dynamically identified set of configurations that contribute a significant weight to the FCI wave function.

In this sense, *i*-FCIQMC exhibits similarities to some iterative subspace diagonalization techniques;[101−103] determinants are generated at random, added to an occupied determinant subspace according to a stochastically realized criteria, and in the long-time limit *i*-FCIQMC aims to approximate an eigenfunction of the Hamiltonian in the space. However, as we shall see, key aspects of the algorithm mean that *i*-FCIQMC is not simply an efficient way to search the space for an optimal configurational subset, which is subsequently diagonalized, as it may first appear.

To demonstrate this effect, a dramatically undersampled *i*-FCIQMC calculation was considered, using Be_2 in a cc-pVQZ basis set with all electrons correlated. Be_2 is a relatively strongly correlated system, even at the equilibrium geometry considered here, due to the near-degeneracy between the 2s and 2p atomic orbitals. While this system has $N_{FCI} = 1.4 \times 10^{12}$ spin-coupled functions, only 2500 walkers were used to obtain the *i*-FCIQMC projected energies shown in Figure 1. These energies nonetheless oscillate around the energy

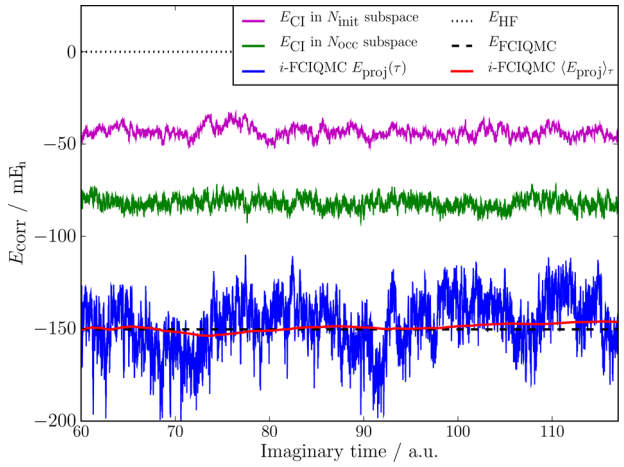

Figure 1. Comparison of the *i*-FCIQMC correlation energy obtained for all electron Be_2 in a cc-pVQZ basis set, to various subspace diagonalizations. This system has $N_{FCI} = 1.4 \times 10^{12}$, but only approximately 2500 walkers were used to obtain the *i*-FCIQMC energy. These occupied $N_{occ} \approx 2140$ determinants, $N_{init} \approx 59$ of which were initiators. The E_{CI} energies in each of these subspaces were obtained by diagonalizing the space that is instantaneously occupied by all *i*-FCIQMC walkers or just the initiator determinants. While the $E_{proj}(\tau)$ values of this *i*-FCIQMC wave function oscillate around FCIQMC energy (E_{FCIQMC}, which is taken to be of FCI accuracy), the variational energies from the diagonalizations capture at most half of the total correlation energy.

obtained using a converged FCIQMC simulation, which is taken to be of FCI accuracy. Included in Figure 1 for comparison are the energies obtained by performing an exact diagonalization of the subspace Hamiltonian at each time step, using the space of determinants that are instantaneously occupied by the *i*-FCIQMC simulation ($N_{occ} \approx 2140$), as well as a separate diagonalization of the smaller subspace of initiators ($N_{init} \approx 59$).

As shown in Figure 1, even at this trivially small number of walkers, the averaged projected energy estimator is already very close to the FCI limit. By comparison, energy estimates taken from the exact diagonalization of the subspace Hamiltonians yield energies which at best capture only about half of the total correlation energy. That these energies are higher than the exact value is unsurprising, since their variationality ensures that any neglected determinants raise the energy, though the magnitude of the error, as well as the quality of the *i*-FCIQMC results with such few walkers is of note. Of course, the instantaneous *i*-FCIQMC wave function is not likely to represent the optimal subspace choice to diagonalize, but this nonetheless demonstrates the effect of the variational estimator in restricted-space calculations. It is certainly not possible that *i*-FCIQMC is sampling the entire space of $O[10^{12}]$ determinants; however, these results suggest this is unnecessary to reach FCI accuracy. Of course, the subspace diagonalization must tend to the exact result as more determinants are included in the expansion, and so as the walker number is increased we should be able to observe this.

The fact that the *i*-FCIQMC energies can achieve such accuracy is indicative of the efficiency of a Monte Carlo sampling of the Hilbert space. Unconstrained by variationality, assuming the sign-problem is controlled, *i*-FCIQMC should return an unbiased estimate of the true energy, with estimable and improvable errorbars. The fact that the whole space has not

been sampled can therefore be reflected in the size of the errorbar, rather than necessarily any systematic deviation. Indeed, in the limit of the absence of a sign-problem in the space, the single walker limit should be sufficient to obtain an exact time-averaged wave function, in common with all Quantum Monte Carlo approaches. What marks out *i*-FCIQMC as different here is the fact that the combination of annihilation events and the initiator approximation seem to efficiently control this sign-problem even within such under-sampled systems, which allows the Monte Carlo sampling to remain unbiased, and achieve accurate time-averaged results, at least for these systems.

In order to analyze the effect of this time-averaging, a smaller Be_2 system in a cc-pVTZ basis set was considered with the core electrons frozen. This smaller system (N_{FCI} = 346 485 spin-coupled functions) makes it possible to monitor the occupations of the entire Hilbert space and compare these to the exact FCI wave function. The *i*-FCIQMC simulation was this time run with 2300 walkers. The upper plot of Figure 2 shows the wave function obtained from a subspace diagonalization on a set of instantaneously occupied determinants (N_w = 2380 and N_{occ} = 1746). This eigenfunction of the truncated space exhibits qualitative agreement with the FCI wave function, and the amplitudes of the most significant determinants are generally correct. Despite this, the variational energy of this solution is still 24.6 mE_h higher than the FCI result.

The second plot of Figure 2 depicts the normalized instantaneous distribution of the *i*-FCIQMC walkers, over the same subspace as the above diagonalization. These occupations bear little resemblance to the FCI coefficients, and the discretization due to the small number of walkers is very evident. However, the instantaneous projected energy of this distribution is only 6.12 mE_h lower than the FCI value, which is a small difference compared to the standard deviation of the data of 9 mE_h observed with this walker number. The determinants of significant amplitude in the FCI wave function are also generally occupied. Nonetheless, the difference between this *i*-FCIQMC distribution and the diagonalization solution in the upper plot further indicates the instantaneous *i*-FCIQMC wave function is not even a qualitative representation of the eigenfunction in the truncated space. The variational energy (i.e., $\langle\psi|H|\psi\rangle/\langle\psi|\psi\rangle$) of this instantaneous distribution would thus have an even higher value than the ground state eigenvalue in the occupied subspace. It is worth noting that when applied to an eigenstate, this variational energy expression will give an identical (and hence also variational) value to that obtained via the projection formula of eq 6.

Finally, using the simulation described above, the *i*-FCIQMC walker distributions were averaged over ~7000 au of imaginary time, and the normalized wave function is presented in the lower frame of Figure 2. Despite occupying only approximately 1750 determinants at a time, this time-averaged wave function nearly perfectly describes the FCI amplitudes on 176 957 determinants, including all the single and double excitations, whose contributions are used to calculate E_{proj}. These results indicate that while the instantaneous *i*-FCIQMC distribution shows little resemblance to an eigenfunction in the FCI, or indeed any truncated space, the *time-averaged* *i*-FCIQMC population dynamics are rather accurately describing the FCI ground state.

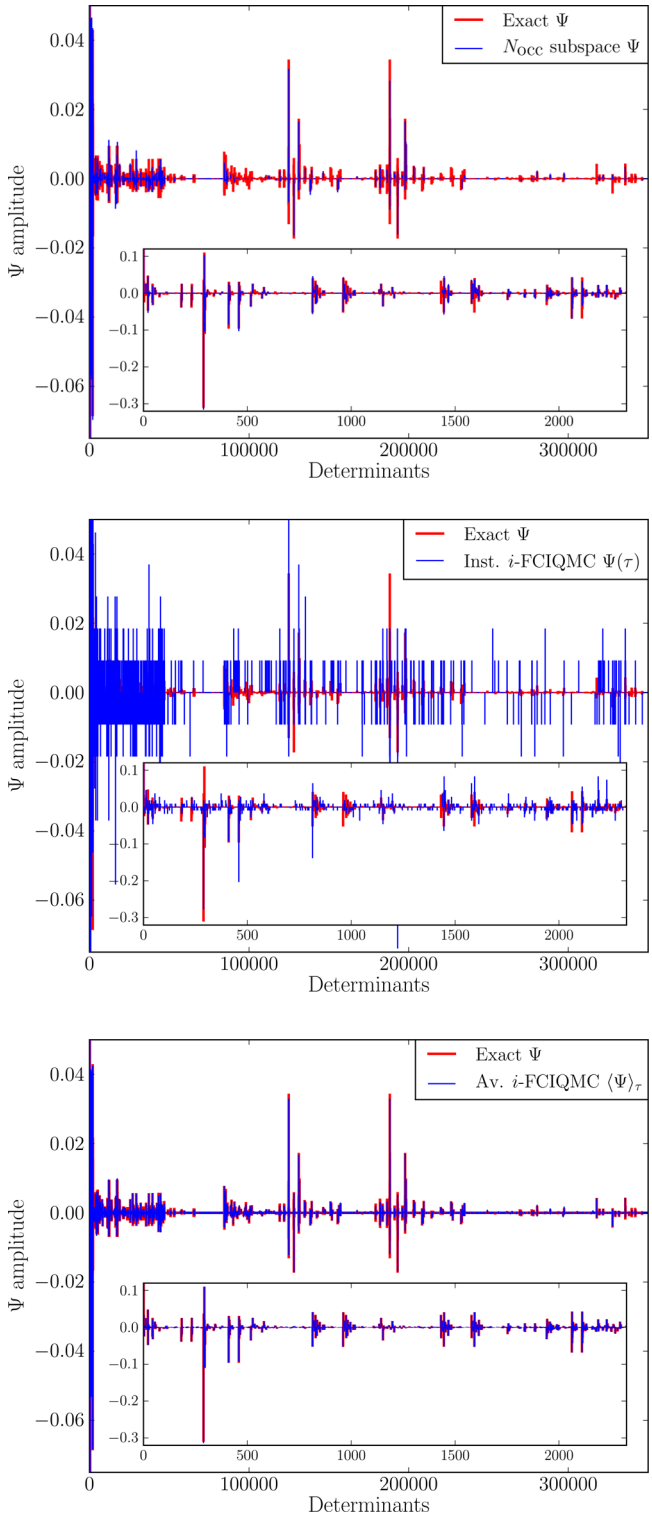

Figure 2. Wave functions obtained from an *i*-FCIQMC simulation of frozen core Be_2 in a cc-pVTZ basis set, compared to the exact FCI wave function. The upper plot shows the relatively good distribution obtained from a subspace diagonalization of the instantaneously occupied space of N_w = 2380 walkers on N_{occ} = 1746 determinants. The second is the normalized instantaneous *i*-FCIQMC distribution of walkers which spans the same subspace as above. This does not agree well with the FCI coefficients, or even the subspace diagonalization, but when the *i*-FCIQMC simulation is averaged over ~7000 au of imaginary time, with N_w ~ 2300 walkers and N_{occ} ~ 1750 walkers, the normalized distribution near perfectly describes the FCI amplitudes on 176 957 determinants (lower plot).

dx.doi.org/10.1021/ct300504f | *J. Chem. Theory Comput.* 2012, 8, 4138−4152

Table 1. *i*-FCIQMC Energies of the Series of First Row Diatomics and Their Constituent Atoms (Hartree)[a]

system	VDZ	VTZ	VQZ	V(TQ)Z	VQZ+$\Delta E_{F12}^{ccsd(T)}$
Be (^1S)[b]	−14.65182(3)	−14.66244(5)	−14.66568(4)	−14.66803(6)	
C (^3P)	−37.76069(1)	−37.78121(1)	−37.786960(9)	−37.79039(1)	−37.788368(9)
N (^4S)	−54.47858(1)	−54.51491(1)	−54.52506(1)	−54.53115(2)	−54.52802(1)
O (^3P)	−74.91010(3)	−74.97414(3)	−74.99388(3)	−75.00602(4)	−75.00103(3)
F (^2P)	−99.52772(4)	−99.6205(1)	−99.65052(7)	−99.6686(2)	−99.66275(7)
Be_2 ($^1\Sigma_g^+$)[b]	−29.30449(8)	−29.32772(7)	−29.3350(1)	−29.3403(1)	
C_2 ($^1\Sigma_g^+$)[b]	−75.7285(1)	−75.7850(1)	−75.8023(3)	−75.8127(3)	−75.8082(3)
CN ($^2\Sigma^+$)	−92.4933(1)	−92.5698(1)	−92.5938(1)	−92.6081(2)	−92.6028(1)
N_2 ($^1\Sigma_g^+$)	−109.2767(1)	−109.3754(1)	−109.4058(1)	−109.4245(1)	−109.4179(1)
CO ($^1\Sigma^+$)	−113.05564(9)	−113.15639(7)	−113.1887(1)	−113.2080(2)	−113.2016(1)
NO ($^2\Pi$)	−129.59995(8)	−129.7185(1)	−129.7562(2)	−129.7793(2)	−129.7713(2)
O_2 ($^3\Sigma_g^-$)	−149.98781(8)	−150.1305(1)	−150.1750(2)	−150.2027(2)	−150.1934(2)
F_2 ($^1\Sigma_g^+$)	−199.09941(9)	−199.2977(1)	−199.3598(2)	−199.3984(2)	−199.3870(2)

[a]Except when noted, these systems had their core electrons frozen and were calculated at the experimental equilibrium bond lengths given by Huber and Herzberg.[105] The VQZ+$\Delta E_{F12}^{ccsd(T)}$ results refer to the *i*-FCIQMC VQZ energy corrected by a CCSD(T)-F12/B contribution, and V(TQ)Z to the basis set extrapolation given by eq 8. The Be_2 experimental bond length was taken from ref 106. The standard F12 basis sets were not available for Be, and so, the corrected energies were omitted for consistency. [b]All electron calculations use the equivalent cc-pCV*X*Z basis sets.

Overall, this analysis suggests that the *i*-FCIQMC algorithm provides a highly compact and accurate representation of the wave function. The *i*-FCIQMC algorithm is effectively identifying the important determinants in the space, without a priori knowledge of the system, such that those neglected from the simulation make negligible contributions to the energy. Furthermore, it is the dynamic nature of the *i*-FCIQMC distribution and its evolution over imaginary time that allows an accurate time-averaged wave function to be obtained while only having to store, in this case, 2 orders of magnitude less memory than would be required to instantaneously consider the entire occupied space. Although not conclusively proven, subspace diagonalization methods with equivalent memory requirements are expected to result in much higher variational energy estimates.

However, for a sufficiently large system, the initiator approximation will result in a non-negligible number of Hamiltonian elements being set to zero at any one time, leading to an error from an effective truncation of the instantaneous available space which does not necessarily reduce to zero with time averaging. It is thus useful to consider the convergence of the *i*-FCIQMC energy with increasing numbers of walkers to establish the value of N_w required to achieve an unbiased estimate, and therefore, FCI accuracy. The severity of the initiator approximation in the presence of multiconfigurational effects is not obvious and will be explored later in this study.

RESULTS AND DISCUSSION

***i*-FCIQMC Energies of the Diatomics at Equilibrium.** The frozen core *i*-FCIQMC total energies of the equilibrium diatomic molecules, and their contributing atoms, are presented in Table 1. The energy of each dimer was obtained using the cc-pVDZ, cc-pVTZ, and cc-pVQZ Dunning correlation consistent basis sets;[12] hereafter referred to as VDZ, VTZ, and VQZ. The equilibrium geometries were taken to be the experimental bond lengths (r_e) given by Huber and Herzberg.[105] For the closed shell molecules Be_2, C_2, N_2, CO, and F_2, as well as O_2 in the VQZ basis set, both angular momentum and spin time-reversal symmetries could be utilized.[11] It should be noted that the energy for the VTZ and VQZ equilibrium geometry for C_2 is different to that published in Tables II and III in ref 11. The equilibrium values in this reference were erroneously averaged over the points shown in Figures 5 and 7, respectively. This hides a sub-error-bar decay of initiator error and results in a systematic error in the published results which is larger than that accounted for by the error bar. Here, we simply use the final energy and error from Figures 5 and 7 in ref 11, which should be taken as correct. The nonequilibrium points are unaffected. For the remaining open shell diatomics, time-reversal symmetry was not used or not applicable, but angular momentum symmetry was employed for the larger VQZ calculations. All calculations were performed using restricted HF (RHF) orbitals, and the core electrons were frozen for all systems except Be and Be_2, for which the corresponding cc-pCV*X*Z basis sets were employed.

Each calculation was performed with the following protocol: a single walker is placed on the HF determinant, and the *i*-FCIQMC algorithm is applied with $\Delta\tau = 10^{-5}-10^{-4}$, $n_a = 3$, and E_S held fixed at zero or ∼0.1 hartree. During this initial stage, the walker population grows exponentially. This walker growth phase is continued until a target population is reached, at which point E_S is allowed to vary to maintain a constant population of walkers at the target value of N_w. During this second phase, various quantities are examined, including the instantaneous values of $N_0(\tau)$ and $E_{proj}(\tau)$. Once these have stabilized, the simulation is deemed to have equilibrated, and the production phase of the calculation can begin (i.e., the accumulators of expectation values are zeroed). This phase is continued, until the statistical errors in $\langle E_{proj}\rangle_\tau$ and $\langle N_0\rangle_\tau$ have been sufficiently reduced. Note that obtaining the statistical error of a quantity being generated in a Monte Carlo simulation requires an estimate of the serial correlation time in the data, which we obtain via a Flyvbjerg−Petersen "blocking" algorithm.[107] For the projected energy measurement, the blocking is performed on the numerator and the denominator separately. $\langle E_{proj}\rangle_\tau$ and its stochastic error are then calculated from the ratio of the averages and the covariance. The final errors are presented in parentheses throughout this paper, as the uncertainty in the preceding digit, and represent one standard deviation in the distribution of energies. This means that, with approximately 68, 95, and 99.7% probability, the stochastically sampled energies will lie within one, two, and three standard deviations of the exact result, respectively.

In order to obtain a reliable *i*-FCIQMC energy, the behavior of the initiator error with walker number must be established, i.e. the rate at which the *i*-FCIQMC energy converges with an increasing number of walkers in the simulation. Two examples of this convergence are shown in Figure 3 for CO and N_2 in

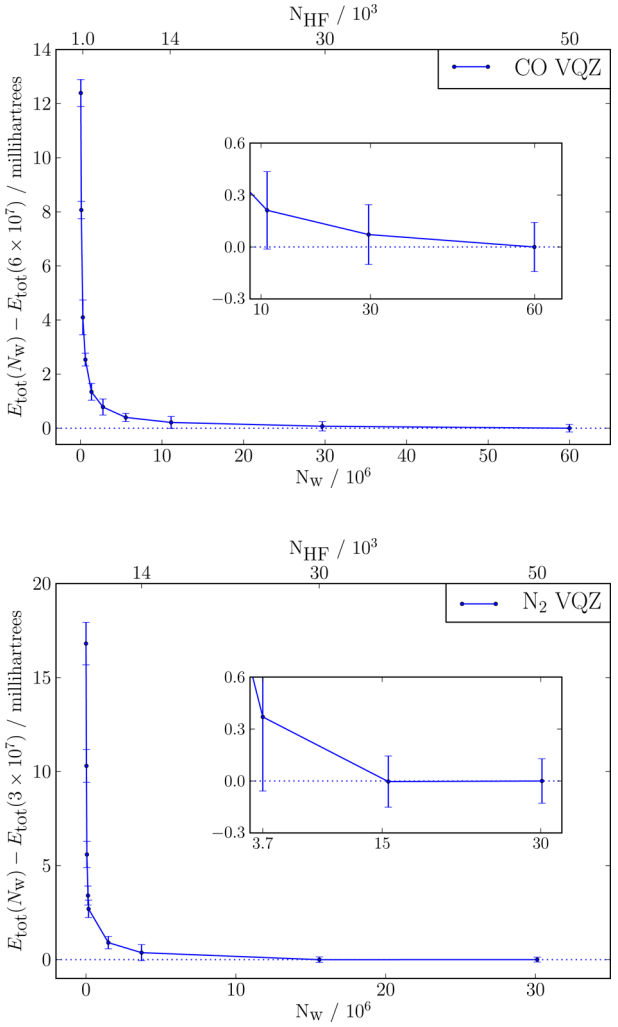

Figure 3. Convergence of the *i*-FCIQMC energies with respect to walker number (N_w), for CO and N_2 in VQZ basis sets. The energies are plot as errors relative to the most converged value (which has 50 000 walkers on the HF determinant). For both systems, the *i*-FCIQMC energy is converged for the last two points, when the population on the HF determinant (shown on the upper *x*-axis) has reached 30 000−50 000 walkers. Taking advantage of spin time-reversal, angular momentum, and point group symmetries, CO and N_2 VQZ have FCI spaces of 4.706×10^{14} and 2.353×10^{14} spin-coupled functions, respectively.

VQZ basis sets. These exemplify the relatively consistent trends that are typical of the diatomic series investigated here. For each of the different systems, the energies are no longer changing beyond the statistical error bars when 30 000−50 000 walkers were accumulated on the HF reference determinant. By using this value of N_0 to determine the total number of walkers required for each calculation, and keeping the other parameters mostly constant, the *i*-FCIQMC simulations performed during this study could be carried out in a relatively black-box manner.

The resulting total energies of the first row diatomics at their equilibrium geometries are shown in Table 1. As an example of the computational cost required for each *i*-FCIQMC calculation, the VDZ, VTZ, and VQZ energies of NO were obtained after approximately 120, 6400, and 11000 CPU hours, respectively.

For the two smallest systems, C_2 and CN in VDZ basis sets, FCI energies could be obtained using MOLPRO.[21,108,109] From these values, the errors in the *i*-FCIQMC results can be calculated as 0.1(1) and 0.0(1) mE_h for C_2 and CN VDZ, respectively ($E_{FCI}(C_2) = -75.72855$ and $E_{FCI}(CN) = -92.49326$ E_h). These verify the absence of any significant initiator errors for these systems.

Basis Set Extrapolation and CCSD(T) F12 Corrections. Any systematic errors from the basis set (FCI) correlation energy, due to the *i*-FCIQMC method itself, are expected to be minimal for the converged total energies presented in Table 1. However, the VDZ, VTZ, and VQZ values are still likely to be hindered by basis set incompleteness errors. In order to obtain results that can be compared to experimental data, these errors can be reduced using either F12 corrections to account for the basis set deficiencies, or extrapolation to the complete basis set (CBS) limit. F12 adapted theories attempt to account for the remaining dynamic correlation which is unable to be captured in the finite basis set, by including explicitly correlated pair products within the wave function ansatz. These have proved extremely successful at achieving basis set converged results. The pseudo F12 corrected results presented here were calculated from the *i*-FCIQMC VQZ values, with an approximate F12 correction provided by the difference between the equivalent basis CCSD(T) and CCSD(T)-F12/B calculations (see ref 10).[109,111] The resulting energies are labeled as VQZ+$\Delta E_{F12}^{CCSD(T)}$, which is later abbreviated to VQZ+F12. Of course, this approach is not ideal, since the basis-set convergence of CCSD(T) and *i*-FCIQMC may differ slightly, and the F12 corrections taken from the CC framework may inherit the difficulties with strong correlation effects. However, for these equilibrium geometries, the approach is expected to be effective.

Alternatively, the extrapolation method of Helgaker et al. was used to calculate the complete basis set correlation energy (E_{corr}^{CBS}) from two data points with different cardinal numbers (X and Y):[13,112]

$$E_{corr}^{CBS} = \frac{E_{corr}^{X} X^3 - E_{corr}^{Y} Y^3}{X^3 - Y^3} \tag{8}$$

This extrapolation technique takes advantage of the systematic improvement in the basis set error, which is provided by the hierarchical sequence of correlation consistent cc-pV*X*Z basis sets used here. It has also been shown to be most effective when calculated using basis sets larger than double-ζ quality.[13] It is therefore applied here to the VTZ and VQZ *i*-FCIQMC results, to obtain estimates of the CBS correlation energy (denoted V(TQ)Z).

The HF energy is known to converge quickly with respect to basis set size, and so, the CBS total energy is simply found by adding E_{corr}^{CBS} to the HF energy obtained in the large aug-cc-pV6Z basis, or aug-cc-pV5Z for the Be systems. The F12 corrected and extrapolated *i*-FCIQMC energies are presented with the finite basis set results for each system in Table 1.

Dissociation Energies. The dissociation energies for this series of first row diatomics can then be calculated from the *i*-FCIQMC total energies presented in Table 1 and compared to experimental results. However, for a fair

comparison to the frozen core nonrelativistic *i*-FCIQMC calculations performed here, the experimentally measured values for the ground state dissociation energy D_0, are first corrected for contributions from the zero point energy ΔE_{ZPE}, scalar relativistic effects ΔE_{SR}, spin−orbit coupling ΔE_{SO}, and core correlation ΔE_{CV}. For simplicity, these effects are assumed to be additive. The exception to this is Be_2, for which all electron calculations were performed and relativistic effects were considered to be negligible. The remaining equilibrium dissociation energies D_e^*(expt), represent the nonrelativistic valence-correlated experimental result which *i*-FCIQMC aims to capture. This calculation of corrected experimental values was previously performed for the homonuclear diatomics by Ruedenberg et al.[71]

The resulting *i*-FCIQMC dissociation energies of each of these diatomics are then presented in Table 2, and their

Table 2. *i*-FCIQMC Dissociation Energies of a Series of Homo- and Heteronuclear First-Row Diatomics (kcal mol^{-1})[a]

basis	Be_2[b]	C_2	CN	CO
VDZ	0.53(3)	129.95(8)	159.40(7)	241.49(6)
VTZ	1.78(6)	139.69(8)	171.71(6)	251.66(5)
VQZ	2.27(9)	143.3(2)	176.80(9)	255.92(9)
V(TQ)Z	2.67(10)	145.5(2)	179.8(1)	258.3(1)
VQZ+$\Delta E_{F12}^{ccsd(T)}$		145.2(2)	179.71(9)	258.68(9)
D_e^*(expt)	2.658(6)	146.9(5)	180.4(2.4)	258.8(2)
basis	NO	N_2	O_2	F_2
VDZ	132.57(5)	200.52(8)	105.17(6)	27.59(7)
VTZ	143.99(6)	216.86(9)	114.35(8)	35.5(1)
VQZ	148.9(1)	223.20(8)	117.5(1)	36.9(1)
V(TQ)Z	151.9(2)	227.3(1)	119.6(1)	38.4(2)
VQZ+$\Delta E_{F12}^{ccsd(T)}$	152.0(2)	227.09(8)	120.1(1)	38.6(1)
D_e^*(expt)	152.63(4)	227.60(5)	120.42(5)	39.0(1)

[a]Only the valence electrons were correlated, and the diatomic geometries were taken to be the experimental bond lengths given by Huber and Herzberg.[105] In angstroms, these are the following: Be_2 2.254,[106] C_2 1.2425, CN 1.1718, N_2 1.0977, CO 1.1283, NO 1.1508, O_2 1.2075, F_2 1.4119. The V(TQ)Z and VQZ+$\Delta E_{F12}^{ccsd(T)}$ energies were calculated from CBS extrapolations and approximate F12 corrected values respectively. The *i*-FCIQMC dissociation energies are compared to corrected experimental results D_e^*(expt). The standard F12 basis sets were not available for Be, and so, the corrected dissociation energy of Be_2 was omitted for consistency. [b]All electron calculations in cc-pCV*X*Z basis sets.

convergence toward the corrected experimental values is depicted in Figure 4. For Be_2, both the all electron cc-pCV*X*Z results and the frozen core cc-pV*X*Z dissociation energies are shown. Since the experimental Be_2 result has not been corrected for core correlation effects, the apparent accuracy of the frozen core cc-pV*X*Z values is fortuitous. However, the extrapolated all electron calculations accurately capture the uncorrected D_e(expt). With the exception of C_2, the *i*-FCIQMC dissociation energies are thus calculated to within chemical accuracy of the experimental results when basis set incompleteness errors are accounted for using either pseudo F12 corrections or extrapolations.

Comparison to Alternative Theoretical Methods. The accuracy of the *i*-FCIQMC dissociation energies can be further evaluated by comparison to some alternative theoretical techniques that have been applied to the same first row diatomics. Such a comparison is made in Table 3, by considering the errors in the calculated D_e values, relative to the corrected experimental results. The focus here is on the accuracy of the electronic structure techniques, rather than on a comparison of the basis set errors and methods of extrapolation. The *i*-FCIQMC results are thus primarily compared to other frozen core nonrelativistic values, which have been obtained using the same basis sets and extrapolation technique.

First, the VDZ, VTZ, and VQZ *i*-FCIQMC results in Table 3 all underestimate the dissociation energy. This suggests these finite basis sets are insufficient to accurately describe the dynamic correlation of the diatomic around equilibrium. This effect has been previously observed for dissociation energies calculated using FCI.[117] It is possible the overly large CEEIS V(TQ)Z value for O_2 is therefore a product of the nonvariationality of this method.

Generally, these CCSD(T) results show rather good agreement with the experimental values, particularly when F12 corrections are included. However, Feller and Sordo have previously identified the more multireference C_2 and CN molecules as the first-row diatomics that present the greatest challenge to CCSD(T).[1] Table 3 demonstrates that more accurate dissociation energies may be achieved for these systems using *i*-FCIQMC, although their experimental values also have the largest errors.

The two QMC methods most commonly applied to molecular systems such as these first-row diatomics are AFQMC[118,119] and DMC.[115,116,120−123] The errors in some DMC results are presented in Table 3 for comparison to the *i*-FCIQMC energies. The values denoted FN-DMC were taken from an older single determinant fixed-node DMC study of the G1 benchmark set of molecules.[115] The errors in these DMC dissociation energies were improved during a more recent study whereby the trial wave functions were constructed from large multideterminant expansions;[116] these are presented in Table 3 as MD FN-DMC. The all-electron DMC results are presented as errors compared to the experimental values without the correction for the core correlation. As well as the use of multiconfigurational wave functions, alternative techniques for optimizing the DMC solution have also been applied to individual first-row dimers to reduce fixed-node errors and improve their DMC dissociation energy.[121−123]

Finally, the *i*-FCIQMC results presented here show relatively similar errors to CEEIS, which has recently been applied to the homonuclear diatomics of this series.[71] Since CEEIS aims to extrapolate truncated calculations to the FCI limit, the agreement between the *i*-FCIQMC and CEEIS results is encouraging, and it suggests that both methods may be effectively and consistently capturing the basis set correlation energy of each system. However, the number of extrapolations required, and the specifics of the orbitals and their partitioning mean that CEEIS is far from a black box method.

Diatomics at Stretched Geometries. Having applied *i*-FCIQMC to the equilibrium diatomics and atoms of this series, and obtained good agreement with CCSD(T) and CEEIS dissociation energies, the next part of this study considers the ability of *i*-FCIQMC to describe the first-row dimers at stretched geometries. The stretched diatomics represent a series of strongly correlated systems that generally require multireference techniques to describe accurately. For example, CCSD(T) has generally diverged by the time it reaches the dissociated bond lengths considered here.[6] However, most multireference methods require some a priori

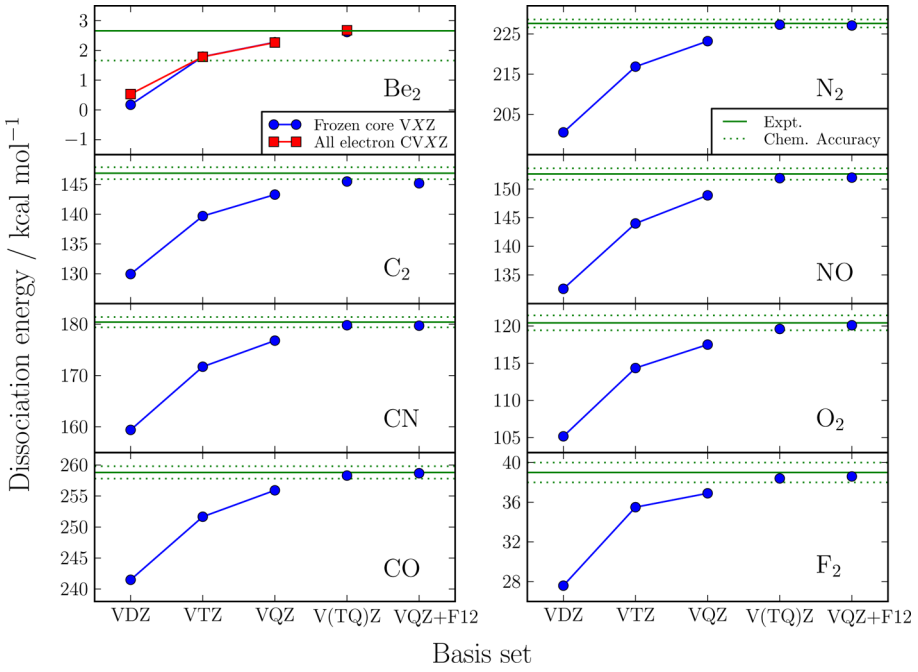

Figure 4. Comparison of calculated *i*-FCIQMC dissociation energies to experimental values. With the exception of Be_2, the nonrelativistic *i*-FCIQMC calculations were performed with the core electrons frozen and so the experimental results have been corrected for relativistic and core correlation effects. These *i*-FCIQMC values have been taken from Table 2, and the VQZ+F12 values refer to the VQZ+$\Delta E_{F12}^{ccsd(T)}$ results. The region of chemical accuracy ($\pm$1 kcal mol^{-1} of D_e^*(expt)) is also marked. Stochastic error bars on the *i*-FCIQMC results are present, but too small to be observed on this plot.

Table 3. Comparison of the Errors in the Series of *i*-FCIQMC Dissociation Energies, to Those Calculated by Other Electronic Structure Methods[a]

method	basis	Be_2[f]	C_2	CN	N_2	CO	NO	O_2	F_2	MAD
i-FCIQMC	VQZ	−0.4	−3.6	−3.6	−4.4	−2.9	−3.7	−2.9	−2.1	3.0
CCSD(T)[b]	VQZ	−1.0	−4.0	−5.6	−5.0	−2.6	−4.5	−3.2	−2.4	3.5
CEEIS[c]	VQZ		−3.4		−4.6			−2.5	−1.9	2.6
i-FCIQMC	V(TQ)Z	0.0	−1.4	−0.6	−0.3	−0.5	−0.7	−0.8	−0.6	0.6
CCSD(T)[b]	V(TQ)Z	−0.6	−1.5	−1.2	−0.6	−0.3	−1.0	−0.7	−0.9	0.9
CEEIS[c]	V(TQ)Z		−1.1		0.0			0.8	−0.4	0.5
FN-DMC[d]				−7.5	−3.2	−2.5	−7.0	−6.5	−5.7	5.4
MD FN-DMC[e]		−0.1	−3.0		−2.5			−3.2	−3.7	2.5
MD FN-DMC-extrap[e]		0.0	−1.5		−2.2			−1.8	−3.3	1.8
i-FCIQMC	VQZ+F12		−1.7	−0.7	−0.5	−0.1	−0.6	−0.3	−0.4	0.6
CCSD(T)-F12	VQZ		−1.8	−1.0	−0.6	0.2	−0.8	−0.3	−0.4	0.5

[a]The errors are given in kilocalories per mole, and these are measured relative to the corrected experimental results. The VQZ+F12 results refer to the VQZ+$\Delta E_{F12}^{ccsd(T)}$ values, and V(TQ)Z, to the CBS extrapolation method described earlier. All coupled-cluster and F12 calculations were performed with MOLPRO.[109−111,113,114] The mean absolute deviation (MAD) across the available results is also presented for each method. [b]Reference 1. [c]Reference 71. [d]Reference 115. [e]Reference 116. [f]All electron calculations in cc-pCV*X*Z basis sets.

understanding of the system, often through the specification of an appropriate multiconfigurational reference state. Contrary to this, *i*-FCIQMC is applied here in an almost identical manner to the equilibrium calculations, starting with only a single walker on the HF determinant. It is hoped that the dynamic identification of the initiator space will be sufficient to still include the determinants required to describe the strong correlation effects. This section begins by examining the accuracy of the *i*-FCIQMC energies for these stretched systems and then considers some comparisons to the more single-reference equilibrium geometries.

Size Consistency. In a sense, the initiator spawning criteria place restrictions on the determinants that may be instantaneously included in the wave function expression. Hence, the loss of size consistency is a potential concern for the *i*-FCIQMC method. For instance, Figure 5 presented later demonstrates the very small fraction of walkers required for each simulation, compared to the number of Slater determinants in the FCI space. The earlier Be_2 results indicate the total sampled space is much larger than the number of instantaneously occupied determinants, but these determinants spanned by the time-averaged simulation will still usually represent only a fraction of each Hilbert space.

The accuracy of the *i*-FCIQMC dissociation energies presented earlier suggests that any size consistency errors for these systems are not significant. However, this can be explicitly tested for the dimers at stretched geometries, by calculating the energy of the molecule at sufficient bond lengths so that the

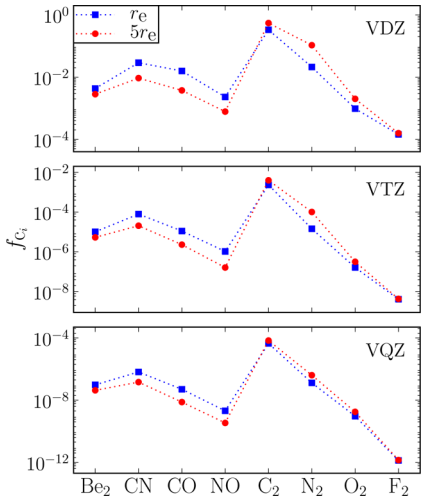

Figure 5. Representation of the difficulty of the *i*-FCIQMC calculations for the series of first-row diatomics at equilibrium and stretched geometries. The values presented are approximate f_{c_i} values for the *i*-FCIQMC simulations. This is given by the number of walkers required to achieve 50 000 walkers on the HF determinant in fixed shift mode, relative to the size of the FCI space. The equilibrium and stretched calculations for each system were performed with the same symmetries and therefore the same value of N_{FCI}. While the stretched calculations of C_2, N_2, O_2, and F_2 appear to be more difficult as expected, the number of walkers required for the CN, CO, and NO simulations is smaller when the internuclear bond is stretched.

atoms are no longer interacting and comparing this result to the sum of the isolated atomic energies. The size consistency error is likely to be system dependent, and so, the entire series of diatomics is investigated here.

The stretched geometries were chosen to be five times the equilibrium bond length of each diatomic ($5r_e$), and the *i*-FCIQMC calculations were run using the same simulation procedure as the equilibrium cases. The time-averaged values of both the energy shift $\langle E_S \rangle_\tau$ and the projected energy $\langle E_{proj} \rangle_\tau$ are presented for each diatomic in Table 4, along with the sum of the constituent atomic energies.

For each basis set, the root-mean-square (rms) deviation of the molecular *i*-FCIQMC energies from these atomic results is within the average stochastic error for both the shift and the projected energy of these systems. The O_2 molecule dissociates through the $^3\Sigma_g^-$ surface into two ^3P oxygen atoms. The stretched calculation was thus performed on O_2 in a singlet state. The VQZ calculation unfortunately required $\sim 9 \times 10^7$ walkers, and the projected energy was slow to converge. Given the size consistency of the shift measurement, the $\langle E_{proj} \rangle_\tau$ value for this system was therefore omitted.

The results in Table 4 suggest there is no significant loss of size consistency for these systems, despite the small fraction of the Hilbert space that is instantaneously occupied during an *i*-FCIQMC simulation. This is encouraging, as it implies that the initiator approximation is effectively sampling the configurations that are important to the wave function and its associated energy, even for these multiconfigurational stretched systems. The determinants that are never occupied likely correspond to "configurational deadwood",[124] in that they have negligible weights in the FCI expression.

Computational Difficulty. While Table 4 demonstrates the ability for *i*-FCIQMC to provide accurate energies for the stretched diatomic systems, these would generally be expected to be more expensive calculations. The well-known multi-configurational nature of the stretched molecules suggests that more walkers would be required to describe the extra highly weighted determinants that are present at the stretched geometry, but are less significant at equilibrium.

The reality of any difference in computational cost between the equilibrium and stretched *i*-FCIQMC calculations was investigated by observing f_{c_i} values for each simulation. These represent the number of walkers required to initially achieve 50 000 walkers on the HF determinant, as a fraction of the Hilbert space of each system ($f_{c_i} = N_w/N_{FCI}$). The observed values are presented in Figure 5, for each of the diatomics in the series considered here.

The f_{c_i} values would be expected to vary slightly with the rate of growth of N_w. To ensure this effect is not responsible for the trends observed here, the CN and N_2 VDZ calculations were performed with identical *effective* shifts, for the equilibrium and stretched geometries. In other words, the initial shift was fixed at values 0.2 (CN) and 0.4 (N_2) E_h above the expected correlation energy, so that the rate of growth of N_w was identical between the two geometries. Despite this, the f_{c_i} values are expected to be only a qualitative comparison of the computational cost, due to the neglect of the relaxation of the walker distribution once the shift can vary.[14]

The C_2, N_2, O_2, and F_2 results in Figure 5 show that more walkers (or at least a similar number) are required for the stretched geometries, to achieve roughly the same population at the HF determinant as its equilibrium counterpart. For example, O_2 in a VTZ basis set required approximately twice the number of walkers when stretched compared to the equilibrium geometry for a similar accuracy. Also, the N_2 calculation at $5r_e$ in a VQZ basis set needed around seven times more walkers than the equilibrium simulation. This behavior is what might be expected as the HF determinant becomes less dominant in the more strongly correlated stretched systems. However, the Be_2, CN, CO, and NO diatomics show the opposite trend. For these molecules, the number of walkers required to initially reach $N_{HF} \approx 50\,000$ actually decreases as each internuclear bond is stretched. For example, the observed N_w for NO in a VQZ basis set is approximately six times smaller in the stretched geometry than at equilibrium.

To attempt to rationalize this observation, the distribution of the *i*-FCIQMC walkers throughout each space was considered. Figure 6 presents a snapshot of the absolute initiator populations for the VDZ wave functions, once the shift has been allowed to change. These are compared between the stretched and equilibrium geometries, for each of the diatomics. The *x*-axis of these plots simply represents each initiator determinant, ordered according to their instantaneous walker populations, which are shown on the *y*-axis.

The systems presented on the right of Figure 6 are those whose stretched geometries need more walkers to accumulate the required population at the HF determinant, than their equilibrium counterparts. The occupations of the initiator determinants show that these extra walkers have been used to describe both the appearance of a few very highly weighted determinants, in accordance with significant static correlation, and an increase in the total number of determinants with large enough weight to become initiators. For instance, N_2 at its equilibrium geometry has $N_{HF} \approx 50\,000$ as required; however, the second largest contribution comes from a determinant with

Table 4. Comparison of the *i*-FCIQMC Diatomic Energies with the Sums of the Constituent Atom Energies (Hartree)[a]

system		VDZ	VTZ	VQZ
Be + Be[b]	$\langle E_{proj}\rangle_\tau$	−29.30364(4)	−29.32488(6)	−29.33137(5)
Be_2 ($5r_e$)[b]	$\langle E_S\rangle_\tau$	−29.3035(6)	−29.3249(6)	−29.3316(9)
	$\langle E_{proj}\rangle_\tau$	−29.30366(7)	−29.32485(9)	−29.3314(3)
C + C	$\langle E_{proj}\rangle_\tau$	−75.52138(2)	−75.56241(1)	−75.57392(1)
C_2 ($5r_e$)	$\langle E_S\rangle_\tau$	−75.5212(2)	−75.5623(1)	−75.5738(2)
	$\langle E_{proj}\rangle_\tau$	−75.5213(5)	−75.5620(3)	−75.5735(2)
C + N	$\langle E_{proj}\rangle_\tau$	−92.23928(1)	−92.29612(2)	−92.31202(2)
CN ($5r_e$)	$\langle E_S\rangle_\tau$	−92.2391(2)	−92.2960(1)	−92.3119(2)
	$\langle E_{proj}\rangle_\tau$	−92.2391(5)	−92.30(1)	−92.3118(4)
N + N	$\langle E_{proj}\rangle_\tau$	−108.95716(1)	−109.02982(2)	−109.05011(2)
N_2 ($5r_e$)	$\langle E_S\rangle_\tau$	−108.9573(2)	−109.0294(4)	−109.0497(3)
	$\langle E_{proj}\rangle_\tau$	−108.957(1)	−109.029(1)	−109.046(5)
C + O	$\langle E_{proj}\rangle_\tau$	−112.67081(3)	−112.75535(3)	−112.78084(3)
CO ($5r_e$)	$\langle E_S\rangle_\tau$	−112.6707(2)	−112.7555(2)	−112.7805(8)
	$\langle E_{proj}\rangle_\tau$	−112.671(1)	−112.755(2)	−112.780(2)
N + O	$\langle E_{proj}\rangle_\tau$	−129.38869(3)	−129.48905(3)	−129.51893(3)
NO ($5r_e$)	$\langle E_S\rangle_\tau$	−129.3885(2)	−129.4891(1)	−129.5185(6)
	$\langle E_{proj}\rangle_\tau$	−129.3886(3)	−129.4890(8)	−129.519(1)
O + O	$\langle E_{proj}\rangle_\tau$	−149.82020(4)	−149.94828(4)	−149.98775(4)
O_2 ($5r_e$)	$\langle E_S\rangle_\tau$	−149.8205(4)	−149.9488(4)	−149.9876(9)
	$\langle E_{proj}\rangle_\tau$	−149.819(1)	−149.949(1)	
F + F	$\langle E_{proj}\rangle_\tau$	−199.05543(6)	−199.2411(2)	−199.3010(1)
F_2 ($5r_e$)	$\langle E_S\rangle_\tau$	−199.0554(6)	−199.2409(8)	−199.302(1)
	$\langle E_{proj}\rangle_\tau$	−199.0556(8)	−199.2404(8)	−199.302(4)
rms	$\langle E_S\rangle_\tau$	0.0001(3)	0.0003(3)	0.0005(6)
rms	$\langle E_{proj}\rangle_\tau$	0.0001(8)	0.001(2)	0.002(2)

[a]The stretched geometries are set at $5r_e$, such that the atoms are no longer interacting. Both the time-averaged shift $\langle E_S\rangle_\tau$ and projected energy $\langle E_{proj}\rangle_\tau$ are shown for the diatomics, because the multiconfigurational nature of the stretched molecules often leads to a decrease of the HF population, and $\langle E_S\rangle_\tau$ becomes the more precise measurement. The O_2 VQZ calculation was computationally expensive, and so the calculation was not run for long enough to obtain a converged projected energy, since the shift was shown to be size consistent. The RMS values correspond to the root mean square deviation of the diatomic energies from the sum of the atoms. These are all within the average stochastic errors. [b]All electron calculations in cc-pCV*X*Z basis sets.

only around 6800 walkers. The populations of the remaining 106 800 initiators continue to decrease from there. In the stretched geometry, the population on the HF determinant has dropped to 35 500 walkers, but 23 other determinants are now present with occupations greater than 20 000. The total number of initiator determinants has also risen by almost a factor of 4. These effects directly demonstrate the typical transition from a predominantly single-reference state at equilibrium, whereby the HF determinant has a large relative weight, to a more multiconfigurational wave function when stretched.

On the other hand, the diatomics on the left of Figure 6 were found to require fewer walkers in their stretched geometries. Their initiator populations suggest that, as the bond is stretched, a few determinants again become apparent with weights comparable to the HF determinant. For example, the largest walker population among the initiators of equilibrium NO (besides from $N_{HF} \approx 50\,000$) is only 3200. Whereas in the stretched system, four more highly weighted determinants are present with occupations larger than 24 000. However, counteracting this effect is a significant reduction in the total number of initiator determinants, which drops from 263 800 to only 93 400 and leads to a net decrease in the number of walkers in this space.

The ability to discern these effects on the plots shown in Figure 6 suggests they are reasonably substantial. As a further test, the stretched CN VDZ simulation was rerun with the total number of walkers increased to match the N_w value used for the equilibrium geometry. The population on the HF determinant was increased to 140 000, yet the total number of initiators remained slightly smaller than that of the equilibrium simulation. This further suggests the trends depicted in Figure 6 are likely to be related to the inherent structure of the Hamiltonian matrix and eigenfunctions, rather than simply to the discretized *i*-FCIQMC sampling. Similar effects were also observed for the larger VTZ and VQZ basis sets, as might be implied by the consistent trends seen in Figure 5.

These results support the appearance of static correlation in stretched diatomics, which would require a multireference method to describe accurately. However, they also show that, in some cases, the total wave function appears to become more compact. Such an effect is favorable for the *i*-FCIQMC method, as it implies a larger fraction of the space may be neglected without introducing a significant initiator error (as demonstrated by Table 4). Furthermore, the *i*-FCIQMC algorithm appears to have little trouble identifying the highly weighted determinants in these first row diatomic systems. On the other hand, the description of many determinants with smaller, but still significant, contributions to the wave function requires more walkers, and thus greater computational cost.

Finally, the presence of many low-lying excited states in the stretched systems meant that occasionally the simulation temporarily converged onto a higher energy symmetry-allowed solution. This will always be corrected eventually but may

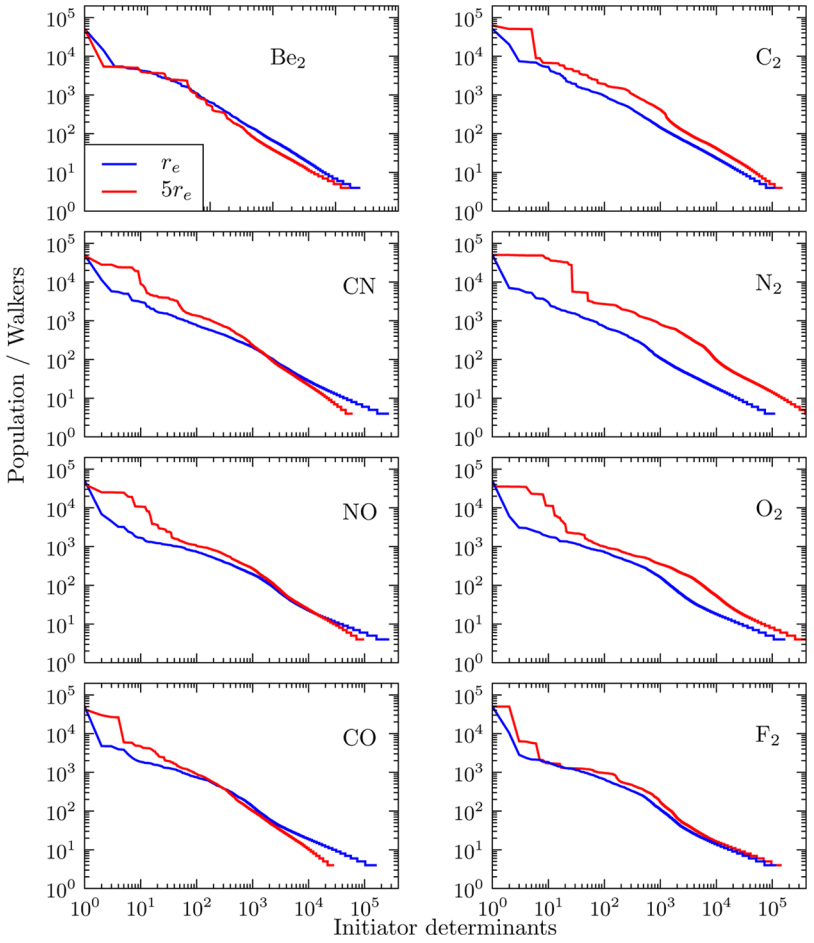

Figure 6. Instantaneous snapshots of the absolute initiator populations, once the *i*-FCIQMC simulation has acquired 50 000 walkers on the HF determinant and the shift has been allowed to change. These are compared between the equilibrium and stretched geometry, for each of the first row diatomics in VDZ basis sets. The initiator determinants have been ordered by their occupations. The systems shown on the left require fewer walkers to achieve the required HF determinant population at the stretched geometry, whereas, for those on the right, N_w is smaller at equilibrium.

require a relatively long simulation. For the systems considered here, these simulations were found to use up to four times more CPU time than would otherwise be expected for a given value of N_w. Therefore, the actual reduction in computational cost for the systems that used fewer walkers in the stretched geometry was not always as significant as implied by the difference in the values of N_w. Nonetheless, the effects described here provide insight into the difficulty of these multiconfigurational systems, with regard to strong correlation and the consequent structure of the *i*-FCIQMC solution.

Interpretation of the Dominant Determinants. As discussed earlier, an advantage of the FCIQMC method is the clear and simple form of the calculated wave function. As a final comment on this study, the instantaneous *i*-FCIQMC wave functions are examined for each molecule in the series of first row diatomics, to demonstrate the kind of insight that can be extracted from such *i*-FCIQMC simulations. In particular, the results considered are those obtained for each of the first row diatomics at their equilibrium geometries, using RHF orbitals in cc-pVDZ basis sets. Shown in Table 5 are the one or two configurations with significant and outstanding instantaneous walker populations, once the *i*-FCIQMC simulation had reached an equilibrium. The approximate coefficients to which these correspond are then presented in parentheses.

Two points are of note with regard to this discussion. First, the dominant determinants and their contributions are not orbitally invariant. However, some comparisons may still be made to previous calculations similarly performed in an RHF basis. Second, only the determinants with outlying dominant populations are presented in Table 5. While these clearly make significant contributions to the wave function, the more abundant determinants with smaller coefficients can also be crucial for the accurate description of a system. Known examples where these are particularly important include Be_2 and N_2. Nonetheless, examination of the highly populated configurations can provide interesting chemical insight and verify the appearance of known characteristics of the wave function.

In general, the highly populated determinants, which are dynamically identified by the *i*-FCIQMC algorithm, are in agreement with previous studies.[71,125−130] For example, the second σ-bound configuration apparent for C_2 is known to be crucial for an accurate description of its dissociation. C_2 presents a particular challenge for CC methods, even at its equilibrium geometry, and this is in part due to its inability to accurately capture the contribution from this second configuration.[71] This second dominant determinant also suggests the C_2 bond may be stronger than it appears when only the double π bond RHF determinant is used as a reference.[131]

Similarly, the coefficient of the RHF determinant for CN is in good agreement with values previously reported.[130] Whereas, to the authors' knowledge, no FCI study of this nature has been

Table 5. Walker Populations ($N_i(\tau)$) and Approximate Coefficients ($C_i(\tau)$) of the Dominant Determinants in a Series of Instantaneous Equilibrated *i*-FCIQMC Wavefunctions[a]

system	dominant determinant (D_i)	$N_i(\tau)$	($C_i(\tau)$)
Be_2[b]	$core^4(2\sigma_g)^2(2\sigma_u^*)^2(\alpha\beta)^4$	50375	(0.89)
Be_2[b]	$core^4(2\sigma_g)^2(3\sigma_g)^2(\alpha\beta)^4$	14212	(0.25)
Be_2[b]	D_3, D_4, ..., $D_{N_{occ}}$	<5500	(<0.10)
C_2	$core^4(2\sigma_g)^2(2\sigma_u^*)^2(1\pi_u)^4(\alpha\beta)^6$	50532	(0.83)
C_2	$core^4(2\sigma_g)^2(3\sigma_g)^2(1\pi_u)^4(\alpha\beta)^6$	20041	(0.33)
C_2	D_3, D_4, ..., $D_{N_{occ}}$	<7500	(<0.12)
CN	$core^4(2\sigma)^2(2\sigma^*)^2(3\sigma)(1\pi)^4(\alpha\beta)^6\alpha$	50221	(0.90)
CN	$core^4(2\sigma)^2(2\sigma^*)(3\sigma)^2(1\pi)^4(\alpha\beta)^6\alpha$	11435	(0.20)
CN	D_3, D_4, ..., $D_{N_{occ}}$	<5800	(<0.10)
CO	$core^4(2\sigma)^2(2\sigma^*)^2(3\sigma)^2(1\pi)^4(\alpha\beta)^7$	50384	(0.94)
CO	D_2, D_3, ..., $D_{N_{occ}}$	<4800	(<0.09)
N_2	$core^4(2\sigma_g)^2(2\sigma_u^*)^2(3\sigma_g)^2(1\pi_u)^4(\alpha\beta)^7$	50361	(0.92)
N_2	D_2, D_3, ..., $D_{N_{occ}}$	<7000	(<0.13)
NO	$core^4(2\sigma)^2(2\sigma^*)^2(3\sigma)^2(1\pi)^4(1\pi^*)(\alpha\beta)^7\alpha$	50287	(0.93)
NO	$core^4(2\sigma)^2(2\sigma^*)^2(3\sigma)^2(1\pi)^2(1\pi^*)^3(\alpha\beta)^7\alpha$	6880	(0.13)
NO	D_3, D_4, ..., $D_{N_{occ}}$	<7000	(<0.08)
O_2	$core^4(2\sigma_g)^2(2\sigma_u^*)^2(3\sigma_g)^2(1\pi_u)^4(1\pi_g^*)^2(\alpha\beta)^7\alpha^2$	49941	(0.94)
O_2	$core^4(2\sigma_g)^2(2\sigma_u^*)^2(3\sigma_g)^2(1\pi_u)^2(1\pi_g^*)^4(\alpha\beta)^7\alpha^2$	6077	(0.11)
O_2	D_3, D_4, ..., $D_{N_{occ}}$	<3100	(<0.06)
F_2	$core^4(2\sigma_g)^2(2\sigma_u^*)^2(3\sigma_g)^2(1\pi_u)^4(1\pi_g^*)^4(\alpha\beta)^9$	50183	(0.94)
F_2	$core^4(2\sigma_g)^2(2\sigma_u^*)^2(3\sigma_u^*)^2(1\pi_u)^4(1\pi_g^*)^4(\alpha\beta)^9$	10463	(0.20)
F_2	D_3, D_4, ..., $D_{N_{occ}}$	<2900	(<0.05)

[a]These are recorded for each of the first row diatomics in the series investigated here, at their equilibrium geometries using a cc-pVDZ basis set. [b]All electron calculations in a cc-pCVDZ basis set.

performed yet for NO. The results presented here suggest that as well as the RHF determinant, a second configuration with the $2\pi y$ pair of electrons excited to the $2\pi y^*$ orbital, also makes a significant contribution to its final wave function.

A system that is notoriously difficult to describe is Be_2. While the RHF determinant alone has a bond order of zero, Table 5 shows that a second bonding configuration makes a substantial contribution when the multireference *i*-FCIQMC method is used to describe the equilibrium wave function. Experimental results likewise find Be_2 to be bound, if somewhat weakly.[132−136] However, previous studies suggest the static correlation captured by the 2s and 2p orbitals alone is insufficient to accurately describe the dissociation energy. The accuracy of the *i*-FCIQMC results shown in Table 2 therefore suggests that, as well as identifying the dominant determinants noted in Table 5, the *i*-FCIQMC algorithm is also effectively capturing the remaining transitions between the 2s and 2p orbitals, as well as the additional dynamic correlation that is crucial to the dissociation energy of Be_2.

Similarly, a UHF description of F_2 finds it to be unbound.[128] The accuracy of the *i*-FCIQMC dissociation energies obtained for F_2 is thus further evidence of the ability to capture dynamic correlation with *i*-FCIQMC. Furthermore, both the dominant determinants shown in Table 5 for F_2 have been previously found to be necessary as a reference state for CEEIS calculations.[129] The second antibonding configuration may contribute to the relatively long equilibrium bond length observed for F_2.

Finally, it should also be noted that although only one configuration is recorded as dominant for N_2, nine determinants are present in the instantaneous *i*-FCIQMC wave function with walker populations between 3000 and 7000.[128] These are single or double excitations from the RHF reference. While such configurations are not included in Table 5 for clarity, they suggest equilibrium N_2 is still a relatively multiconfigurational system, as indicated by previous studies.

CONCLUSION

The recently developed *i*-FCIQMC method was applied here to a series of first row diatomics, which are known to be difficult to accurately describe using ab initio electronic structure techniques. When applied in a relatively black box manner, *i*-FCIQMC was shown to achieve converged energies, which are considered to be within stochastic errors of the FCI limit. It was shown that in the *i*-FCIQMC method, while in an undersampled simulation the instantaneous distribution of walkers can be far from an eigenfunction, the time-averaged distribution of walkers nevertheless accurately reproduces the ground-state wave function and energy. The accuracy of the *i*-FCIQMC results for the equilibrium diatomics was then evaluated through calculation of each dissociation energy. These were generally obtained to within stochastic errors of the experimental results, once basis set errors were accounted for by pseudo F12 corrections or extrapolation to the CBS limit.

The diatomics were then stretched toward dissociation to provide a series of more multiconfigurational systems in which to investigate the applicability of *i*-FCIQMC. The results demonstrate that size consistency is maintained, even when the instantaneous *i*-FCIQMC distribution of walkers sample only a small fraction of the total space. Furthermore, comparison of the stretched diatomic calculations to their equilibrium counterparts reveals different properties of these strongly correlated wave functions, which can be more or less favorable for an *i*-FCIQMC simulation. Finally, consideration

of the *i*-FCIQMC wave functions themselves demonstrates the ability of *i*-FCIQMC to relatively easily identify the most significant determinants that contribute to static correlation. The simple and clear form of the calculated wave function was then used to obtain insight into the structure and difficulty of each diatomic.

To conclude, the *i*-FCIQMC method is shown to provide an accurate and consistent description of both the static and dynamic correlation effects present throughout the series of first row diatomics investigated here. Furthermore, the efficiency and highly parallelizable nature of the *i*-FCIQMC algorithm allows these results to be achieved with tractable computational cost. This study therefore suggests that with continued development and optimization, the *i*-FCIQMC method may be expected to provide valuable insight into a range of interesting systems.

AUTHOR INFORMATION

Corresponding Author

*E-mail: asa10@cam.ac.uk.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank Sandeep Sharma and Garnet Chan for making us aware of an error in ref 11, which is corrected here. The authors would also like to thank the Woolf Fisher Trust, the Cambridge Home and EU Scholarship scheme, EPSRC, and Trinity College Cambridge, for funding. The calculations were performed on the facilities of the Swiss National Supercomputing Centre (CSCS).

REFERENCES

(1) Feller, D.; Sordo, J. A. *J. Chem. Phys.* **2000**, *113*, 485.
(2) Feller, D.; Peterson, K. A. *J. Chem. Phys.* **1998**, *108*, 154.
(3) Bak, K. L.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Gauss, J. *Chem. Phys. Lett.* **2000**, *317*, 116.
(4) Schmidt, M. W.; Lam, M. T. B.; Elbert, S. T.; Ruedenberg, K. *Theor. Chim. Acta* **1985**, *68*, 69.
(5) Peterson, K. A.; Wilson, A. K.; Woon, D. E.; Dunning, T. H., Jr. *Theor. Chim. Acta* **1997**, *97*, 251.
(6) Abrams, M.; Sherrill, C. *J. Chem. Phys.* **2004**, *121*, 9211.
(7) Booth, G. H.; Thom, A. J. W.; Alavi, A. *J. Chem. Phys.* **2009**, *131*, 054106.
(8) Cleland, D.; Booth, G. H.; Alavi, A. *J. Chem. Phys.* **2010**, *132*, 041103.
(9) Booth, G. H.; Alavi, A. *J. Chem. Phys.* **2010**, *132*, 174104.
(10) Cleland, D. M.; Booth, G. H.; Alavi, A. *J. Chem. Phys.* **2011**, *134*, 024112.
(11) Booth, G. H.; Cleland, D.; Thom, A. J. W.; Alavi, A. *J. Chem. Phys.* **2011**, *135*, 084104.
(12) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.
(13) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Let.* **1998**, *286*, 243.
(14) Shepherd, J. J.; Booth, G. H.; Gruneis, A.; Alavi, A. *Phys. Rev. B* **2012**, *85*, 081103.
(15) Shavitt, I. *Mol. Phys.* **1998**, *94*, 3.
(16) Sherrill, C. D.; Schaefer, H. F., III *Adv. Quantum Chem.* **1999**, *34*, 143.
(17) Bauschlicher, C. W.; Taylor, P. R. *J. Chem. Phys.* **1986**, *85*, 2779.
(18) Larsen, H.; Olsen, J.; Jørgensen, P.; Christiansen, O. *J. Chem. Phys.* **2000**, *113*, 6677.
(19) Sherrill, C. D.; Piecuch, P. *J. Chem. Phys.* **2005**, *122*, 124104.
(20) Chaudhuri, R. K.; Freed, K. F. *J. Chem. Phys.* **2005**, *122*, 154310.
(21) Knowles, P. J.; Handy, N. C. *Chem. Phys. Lett.* **1984**, *111*, 3159.
(22) Olsen, J.; Roos, B. O.; Jørgensen, P.; Jensen, H. J. A. *J. Chem. Phys.* **1988**, *89*, 2185.
(23) Harrison, R. *J. Chem. Phys.* **1991**, *94*, 5021.
(24) Povill, A.; Rubio, J.; Illas, F. *Theor. Chim. Acta* **1992**, *82*, 229.
(25) Sherrill, C. D.; Schaefer, H. F., III *J. Phys. Chem.* **1996**, *100*, 6069.
(26) Ivanic, J.; Ruedenberg, K. *Theor. Chem. Acc.* **2001**, *106*, 339.
(27) Shepard, R. *J. Phys. Chem. A.* **2006**, *110*, 8880.
(28) Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9364.
(29) Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9377.
(30) Bunge, C. F. *J. Chem. Phys.* **2006**, *125*, 014107.
(31) Coester, F. *Nucl. Phys.* **1958**, *7*, 421.
(32) Coester, F.; Kummel, H. *Nucl. Phys.* **1960**, *17*, 477.
(33) Cizek, J. *J. Chem. Phys.* **1966**, *45*, 4256.
(34) Bartlett, R. J. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359.
(35) Bartlett, R. J. *J. Phys. Chem.* **1989**, *93*, 1697.
(36) Crawford, T. D.; Schaefer, H. F., III *Rev. Comput. Chem.* **2000**, *14*, 33.
(37) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2000**, *113*, 18.
(38) Piecuch, P.; Kowalski, K.; Pimienta, I. S. O.; Fan, P.-D.; Lodriguito, M.; McGuire, M. J.; Kucharski, S. A.; Ku, T.; Musia, M. *Theor. Chem. Acc.* **2004**, *112*, 349.
(39) Krylov, A. I. *Acc. Chem. Res.* **2006**, *39*, 83.
(40) Krylov, A. I.; Sherrill, C. D. *J. Chem. Phys.* **2002**, *116*, 3194.
(41) Nooijen, M.; Le Roy, R. J. *J. Mol. Struct.: THEOCHEM* **2006**, *768*, 25.
(42) Musia, M.; Bartlett, R. J. *J. Chem. Phys.* **2005**, *122*, 224102.
(43) Bartlett, R. J.; Musia, M. *J. Chem. Phys.* **2006**, *125*, 204105.
(44) Chattopadhyay, S.; Pahari, D.; Mukherjee, D.; Mahapatra, U. S. *J. Chem. Phys.* **2004**, *120*, 5968.
(45) Evangelista, F. A.; Allen, W. D.; Schaefer, H. F., III *J. Chem. Phys.* **2007**, *127*, 024102.
(46) Li, X.; Paldus, J. *J. Chem. Phys.* **1998**, *108*, 637.
(47) Li, X.; Paldus, J. *J. Chem. Phys.* **2006**, *124*, 174101.
(48) Piecuch, P.; Oliphant, N.; Adamowicz, L. *J. Chem. Phys.* **1993**, *99*, 1875.
(49) Ivanov, V. V.; Adamowicz, L.; Lyakh, D. I. *Int. J. Quantum Chem.* **2006**, *106*, 2875.
(50) Lyakh, D. I.; Musial, M.; Lotrich, V. F.; Bartlett, R. J. *Chem. Rev.* **2012**, *112*, 182.
(51) Roos, B. O. *Adv. Chem. Phys.* **1987**, *69*, 399.
(52) Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 41.
(53) Schmidt, M. W.; Gordon, M. S. *Annu. Rev. Phys. Chem.* **1998**, *49*, 233.
(54) Werner, H.-J.; Knowles, P. J. *J. Chem. Phys.* **1988**, *89*, 5803.
(55) Hirao, K. *Chem. Phys. Lett.* **1992**, *190*, 374.
(56) Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983.
(57) Roos, B. O.; Linse, P.; Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Phys.* **1982**, *66*, 197.
(58) Werner, H.-J. *Mol. Phys.* **1996**, *89*, 645.
(59) Celani, P.; Stoll, H.; Werner, H.-J.; Knowles, P. J. *Mol. Phys.* **2004**, *102*, 2369.
(60) White, S. R. *Phys. Rev. Lett.* **1992**, *69*, 2863.
(61) White, S. R.; Martin, R. L. *J. Chem. Phys.* **1999**, *110*, 4127.
(62) Mitrushenkov, A. O.; Linguerri, R.; Palmieri, P.; Fano, G. *J. Chem. Phys.* **2003**, *119*, 4148.
(63) Chan, G. K.-L.; KÃąllay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *121*, 6110.
(64) Legeza, Ö.; Röder, J.; Hess, B. A. *Mol. Phys.* **2003**, *101*, 2019.
(65) Nakata, M.; Ehara, M.; Nakatsuji, H. *J. Chem. Phys.* **2002**, *116*, 5432.
(66) Mazziotti, D. A. *Phys. Rev. Lett.* **2004**, *93*, 213001.
(67) Mazziotti, D. A. *Acc. Chem. Res.* **2006**, *39*, 207.
(68) Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *121*, 10852.
(69) Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *121*, 10905.
(70) Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *121*, 10919.
(71) Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2005**, *122*, 154110.
(72) Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2006**, *124*, 174304.

(73) Truhlar, D. G. *Chem. Phys. Lett.* **1998**, *294*, 45.
(74) Martin, J. M. L. *Chem. Phys. Lett.* **1996**, *259*, 669.
(75) Lee, J. S.; Park, S. Y. *J. Chem. Phys.* **2000**, *112*, 10746.
(76) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104.
(77) Varandas, A. J. C. *J. Chem. Phys.* **2000**, *113*, 8880.
(78) Klopper, W.; Kutzelnigg, W. *J. Mol. Struct.* **1986**, *135*, 339.
(79) Kutzelnigg, W. *Int. J. Quantum Chem.* **1994**, *51*, 447.
(80) Hylleraas, E. A. *Z. Phys.* **1929**, *54*, 347.
(81) Boys, S. F.; Handy, N. C. *Proc. R. Soc. London, Ser. A* **1969**, *310*, 43.
(82) Kutzelnigg, W. *Theor. Chim. Acta* **1985**, *68*, 445.
(83) Klopper, W. *Mol. Phys.* **2001**, *9*, 481.
(84) Noga, J.; Valiron, P.; Klopper, W. *J. Chem. Phys.* **2001**, *115*, 2022.
(85) Valeev, E. F.; Janssen, C. L. *J. Chem. Phys.* **2004**, *121*, 1214.
(86) Gdanitz, R. J. *J. Chem. Phys.* **1998**, *109*, 9795.
(87) Cardoen, W.; Gdanitz, R. J. *J. Chem. Phys.* **2005**, *123*, 024304.
(88) Bukowski, R.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **1999**, *110*, 4165.
(89) Hättig, C.; Tew, D. P.; Köhn, A. *J. Chem. Phys.* **2010**, *132*, 231102.
(90) Klopper, W.; Manby, F. R.; Ten-No, S.; Valeev, E. F. *Int. Rev. Phys. Chem.* **2006**, *25*, 427.
(91) Torheyden, M.; Valeev, E. F. *J. Chem. Phys.* **2009**, *131*, 171103.
(92) Ten-no, S. *Chem. Phys. Lett.* **2007**, *447*, 175.
(93) Yanai, T.; Shiozaki, T. *J. Chem. Phys.* **2012**, *136*, 084107.
(94) Pople, J. A. *Rev. Mod. Phys.* **1999**, *71*, 1267.
(95) Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *J. Chem. Phys.* **1989**, *90*, 5622.
(96) Curtiss, L. A.; Jones, C.; Trucks, G. W.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1990**, *93*, 2537.
(97) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221.
(98) Martin, J. M. L.; De Oliveira, G. *J. Chem. Phys.* **1999**, *111*, 1843.
(99) Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kallay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *120*, 4129.
(100) Rossi, E.; Bendazzoli, G. L.; Evangelisti, S.; Maynau, D. *Chem. Phys. Lett.* **1999**, *310*, 530.
(101) Greer, J. C. *J. Chem. Phys.* **1995**, *103*, 1821.
(102) Troparevsky, M.; Franceschetti, A. *J. Phys.: Condens. Matter* **2008**, *20*, 055211.
(103) Sambataro, M.; Gambacurta, D.; Monaco, L. L. *Phys. Rev. B* **2011**, *83*, 045102.
(104) Spencer, J. S.; Blunt, N. S.; Foulkes, W. M. C. *J. Chem. Phys.* **2012**, *136*, 054110.
(105) Huber, K. P.; Herzberg, G. *Molecular Spectra and Molecular Structure: Constants of Diatomic*; Van Nostrand Reinhold: New York, 1979; pp 112−490.
(106) Merritt, J. M.; Bondybey, V. E.; Heaven, M. C. *Science* **2009**, *324*, 1548.
(107) Flyvbjerg, H.; Petersen, H. G. *J. Chem. Phys.* **1989**, *91*, 461.
(108) Knowles, P. J.; Handy, N. C. *Comput. Phys. Commun.* **1989**, *54*, 75.
(109) Werner, H.-J.; Knowles, P. J.; Manby, F. R.; Schütz, M.; et al. *MOLPRO, version 2010.1, a package of ab initio programs*; 2010; see http://www.molpro.net.
(110) Adler, T. B.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2007**, *127*, 221106.
(111) Knizia, G.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2009**, *130*, 054104.
(112) Helgaker, T.; Klopper, W.; Koch, H.; Nog, J. *J. Chem. Phys.* **1997**, *106*, 9639.
(113) Deegan, M. J. O.; Knowles, P. J. *Chem. Phys. Lett.* **1994**, *227*, 321.
(114) Knowles, P. J.; Hampel, C.; Werner, H.-J. *J. Chem. Phys.* **1993**, *99*, 5219.
(115) Grossman, J. C. *J. Chem. Phys.* **2002**, *117*, 1434.
(116) Morales, M. A.; McMinis, J.; Clark, B. K.; Kim, J.; Scuseria, G. E. *J. Chem. Theory Comput.* **2012**, *8*, 2181.
(117) Abrams, M.; Sherrill, C. *J. Chem. Phys.* **2003**, *118*, 1604.
(118) Al-Saidi, W. A.; Zhang, S.; Krakauer, H. *J. Chem. Phys.* **2006**, *124*, 224101.
(119) Purwanto, W.; Zhang, S.; Krakauer, H. *J. Chem. Phys.* **2009**, *130*, 094107.
(120) López Rios, P.; Ma, A.; Drummond, N. D.; Towler, M. D.; Needs, R. J. *Phys. Rev. E* **2006**, *74*, 066701.
(121) Lu, S.-I. *J. Chem. Phys.* **2003**, *118*, 6152.
(122) Toulouse, J.; Umrigar, C. J. *J. Chem. Phys.* **2008**, *128*, 174101.
(123) Nemec, N.; Towler, M. D.; Needs, R. J. *J. Chem. Phys.* **2010**, *132*, 034111.
(124) Bytautas, L.; Ruedenberg, K. *Chem. Phys.* **2009**, *356*, 64.
(125) Hund, Z. *Physik* **1928**, *51*, 759.
(126) Mulliken, R. S. *Phys. Rev.* **1928**, *32*, 186.
(127) Lennard-Jones, J. E. *Trans. Faraday Soc.* **1929**, *25*, 668.
(128) Laidig, W. D.; Saxe, P.; Bartlett, R. J. *J. Chem. Phys.* **1987**, *86*, 887.
(129) Bytautas, L.; Nagata, T.; Gordon, M. S.; Ruedenberg, K. *J. Chem. Phys.* **2007**, *127*, 164317.
(130) Thørgersen, L.; Olsen, J. *Chem. Phys. Lett.* **2004**, *393*, 36.
(131) Su, P.; Wu, J.; Gu, J.; Wu, W.; Shaik, S.; Hiberty, P. C. *J. Chem. Theory Comput.* **2011**, *7*, 121.
(132) Lengsfield, B. H., III; McLean, A. D.; Yoshimine, M.; Liu, B. *J. Chem. Phys.* **1983**, *79*, 1891.
(133) Harrison, R. J.; Handy, N. C. *Chem. Phys. Lett.* **1983**, *98*, 97.
(134) Bondybey, V. E.; English, J. H. *J. Chem. Phys.* **1984**, *80*, 568.
(135) Bondybey, V. E. *Chem. Phys. Lett.* **1984**, *109*, 436.
(136) Bondybey, V. E. *Science* **1985**, *227*, 125.