Linked Coupled Cluster Monte Carlo

R. S. T. Franklin,[1] J. S. Spencer,[2,3] A. Zoccante,[1] and A. J. W. Thom[1]

[1] *University Chemical Laboratory, Lensfield Road, Cambridge, CB2 1EW, United Kingdom*

[2] *Department of Physics, Imperial College London, Exhibition Road, London, SW7 2AZ, United Kingdom*

[3] *Department of Materials, Imperial College London, Exhibition Road, London, SW7 2AZ, United Kingdom*

We consider a new formulation of the stochastic coupled cluster method in terms of the similarity transformed Hamiltonian. We show that improvement in the granularity with which the wavefunction is represented results in a reduction in the critical population required to correctly sample the wavefunction for a range of systems and excitation levels and hence leads to a substantial reduction in the computational cost. This development has the potential to substantially extend the range of the method, enabling it to be used to treat larger systems with excitation levels not easily accessible with conventional deterministic methods.

Coupled cluster theory is a highly accurate and size consistent method of treating electron correlation.[1,2] It can achieve high accuracy for molecules[3] as well as for extended systems[4,5]. However, the scaling with system size ($O(N^6)$ for the CCSD truncation and worsening for higher truncation) and the complexity of the equations at high truncation levels limit the applicability of the method. It also scales poorly over distributed computer architectures due to the dense linear algebra operations involved. Recently, inspired by the successes of the Full Configuration Interaction Quantum Monte Carlo (FCIQMC) method[4,6–13], a stochastic formulation of the coupled cluster equations (coupled cluster Monte Carlo, CCMC) has been proposed to ameliorate these problems.[14] It uses a set of discrete "excips" to represent the cluster amplitudes, but only converges to the correct ground state when a system dependent critical number of walkers (the "plateau") is exceeded.[14,15] This limits the size of systems accessible to the method. In this paper, we propose a modification of the CCMC algorithm that reduces the plateau heights enabling the application of the method to larger systems.

In coupled cluster theory, the wavefunction is represented using the exponential ansatz $\Psi_{CC} = e^{\hat{T}} |D_{HF}\rangle$, where $\hat{T} = \sum_{\mathbf{i}} t_{\mathbf{i}} \hat{a}_{\mathbf{i}}$ and $\hat{a}_{\mathbf{i}}$ is a string of creation and annihilation operators in second quantization theory, such that $\hat{a}_{\mathbf{i}} |D_{HF}\rangle = |D_{\mathbf{i}}\rangle$, where $|D_{\mathbf{i}}\rangle$ is the Slater determinant with orbitals $i_1, i_2, \ldots, i_N$ occupied. Truncations of the theory are obtained by restricting the sum in $\hat{T}$ to contain excitations of at most the truncation level number of orbitals, but the exponential form ensures the theory remains size consistent. The amplitudes $t_{\mathbf{i}}$ are determined by projecting the Schrödinger equation $\hat{H}\Psi_{CC} = E\Psi_{CC}$ onto the set of substituted determinants generated by $\hat{T}$. In CCMC, after rescaling the amplitudes by t_{HF} to give the modified ansatz $\Psi_{CC} = t_{HF} e^{\hat{T}/t_{HF}} |D_{HF}\rangle$, this is recast into an iterative form:

$$t_{\mathbf{i}} \rightarrow t_{\mathbf{i}} - \delta\tau \left\langle D_{\mathbf{i}} \left| \hat{H} - S \right| \Psi_{CC} \right\rangle \tag{1}$$

where $\delta\tau$ is a small positive number (the timestep)

and S the shift, an energy offset introduced to control the growth of the coefficients. The coefficients of the wavefunction are discretised and represented by populations of fictitious particles called excips on each excitor, where an excitor is a combination of creation and annihilation operators that produces an excited determinant from the Hartree-Fock. The wavefunction and Hamiltonian are then stochastically sampled to update the excitation amplitudes. In outline, the algorithm proceeds as follows (for more details see Ref. 16):

1. Choosing a cluster: s excitors (s is randomly chosen from $0 \leq s \leq l + 2$ where l is the truncation level) are randomly chosen and combined to give an excitation operator which acts on the reference determinant to give a determinant $|D_{\mathbf{n}}\rangle$. This step is equivalent to truncation of the Taylor expansion of $e^{\hat{T}}$ at $l+2$-th power (exact for a 2-body Hamiltonian with up to l-fold excitations in $\hat{T}$) and samples the wavefunction ansatz. The factor of $\frac{1}{n!}$ arising in the Taylor expansion is absorbed into the calculated weight of the cluster.

2. Spawning: A random single or double excitation of $|D_{\mathbf{n}}\rangle$, $|D_{\mathbf{m}}\rangle$, is chosen and a new excip spawned on $|D_{\mathbf{m}}\rangle$ with probability proportional to $|H_{\mathbf{mn}}|$ and sign determined by the sign of $H_{\mathbf{mn}}$, where $H_{\mathbf{mn}} = \langle D_{\mathbf{m}}|\hat{H}|D_{\mathbf{n}}\rangle$. This step samples the action of the Hamiltonian.

3. Death: A new excip is spawned on $|D_{\mathbf{n}}\rangle$ with probability proportional to $|H_{\mathbf{nn}} - E|$ and opposite sign to the parent.

4. Annihilation: Pairs of excips with opposite signs on the same excitor are removed.

In conventional treatments of coupled cluster theory, it is convenient to rewrite the equations in the form

$$\left\langle D_{\mathbf{m}} \left| \bar{H} - E \right| D_{HF} \right\rangle = 0, \tag{2}$$

using the similarity transformed Hamiltonian $\bar{H} = e^{-\hat{T}}\hat{H}e^{\hat{T}}$. This has the advantage that, using the Campbell–Baker–Hausdorff formula and Wick's theorem[17], the expansion of the exponential truncates at fourth order and can be written as

$$\bar{H} = \hat{H} + [\hat{H},\hat{T}]_c + \frac{1}{2}[[\hat{H},\hat{T}],\hat{T}]_c + \frac{1}{3!}[[[\hat{H},\hat{T}],\hat{T}],\hat{T}]_c$$
$$+ \frac{1}{4!}[[[[\hat{H},\hat{T}],\hat{T}],\hat{T}],\hat{T}]_c, \qquad (3)$$

where the subscript c is included to emphasise that only the terms of the commutators in which the Hamiltonian is connected to each cluster operator are needed, and the remaining terms cancel to zero due to the commutators. This decouples the amplitude equations from the energy equation.

The question thus arises as to whether benefits can be derived from using the similarity transformed Hamiltonian in CCMC. The iterative equations used to calculate the amplitudes are, analogously to the original CCMC equations,

$$t_{\mathbf{i}} \rightarrow t_{\mathbf{i}} - \delta\tau t_{\mathrm{HF}}\left\langle D_{\mathbf{i}} \left| \bar{H} \right| D_{\mathrm{HF}}\right\rangle \qquad \text{if } |D_{\mathbf{i}}\rangle \neq |D_{\mathrm{HF}}\rangle$$
$$(4)$$
$$t_{\mathrm{HF}} \rightarrow t_{\mathrm{HF}} - \delta\tau t_{\mathrm{HF}}\left\langle D_{\mathrm{HF}} \left| \bar{H} - S \right| D_{\mathrm{HF}}\right\rangle \qquad (5)$$

Notably, the shift only appears in the equation for the reference population. This means that the population control only acts directly on the reference, and population control for other excitors is achieved indirectly by changes in the rate of spawning, both from the reference and also, due to the change in normalisation, from the other excitors.

Can we sample $\bar{H}$ in a manner analogous to CCMC? In Eq. (3), we may choose any of the five terms at random, and this is equivalent to selecting the size, s, of cluster to consider. Each $\hat{T}$ within the unexpanded commutator is itself a sum over excitors (each with an amplitude). To sample this, for each of the $\hat{T}$ operators, we pick, at random, just one of the excitors within its sum. This results in s (potentially different) excitors, i.e. a cluster of size s.

Given this specific choice of excitors, we must still evaluate the entire nested commutator. Could we instead sample at random just one of the terms in the expansion of the nested commutator rather than the complete expansion? While this is possible, we note that one of the benefits of the similarity transformed approach is that for unlinked clusters all the terms in the nested commutator cancel to zero, and so decomposing the commutator would remove any gains from the cancellation of terms. In order to sample the commutator we must instead evaluate the sum of all the permutations of $\hat{H}$ and $\hat{T}$ operators arising from the commutators for a sampled term from $\bar{H}$. The random sampling of a cluster containing s excitors from the $\hat{T}^s$ term in the power expansion of $e^{\hat{T}}$ in conventional CCMC can also be used to select the excitors included in the s-fold commutator. However, the

allowed excitations in $\hat{H}$ (to give a non-zero projection on to some $D_{\mathbf{i}}$ in Eq. (4)) can differ in each term within the commutator, and so sampling this is not so straightforward.

To solve this, we return briefly to look at the second sampling step in CCMC: $\hat{H}$ is sampled by selecting an allowed excitation of the collapsed cluster $|D_{\mathbf{n}}\rangle$. i.e. we choose $\langle D_{\mathbf{i}}|$ such that matrix element $H_{\mathbf{in}} = \langle D_{\mathbf{i}}|\hat{H}|D_{\mathbf{n}}\rangle$ is non-zero. Once this excitation is chosen, the value of $\mathbf{i}$ has been determined, and this then selects a single t-update, $t_{\mathbf{i}}$, from the set of all possible updates in Eq. (1). We may look at this in a different manner: Let us select a given projectee determinant $\langle D_{\mathbf{i}}|$ in Eq. (1), and see what allowed samplings of $\hat{H}$ may lead to it. For conventional CCMC, the choice is unique.

For linked CCMC, a choice of a given projectee determinant $\langle D_{\mathbf{i}}|$ corresponds to at most one allowed excitation in $\hat{H}$ in each term of the commutator. This excitation is the same in each term, but may have a different corresponding matrix element. Therefore for a given sampling of $\hat{T}^s$ (i.e. cluster), we must carefully select one from all possible allowed projectees, and then we may evaluate the commutator for that projectee.

To use this form of the coupled cluster equations in stochastic coupled cluster requires making three modifications to the steps in conventional CCMC.

1. Due to the truncation of the Hausdorff expansion, only clusters of at most four excitors need to be used, regardless of the truncation level considered.

2. Clusters where two of the operators excite to or from the same orbital (following Bartlett et. al.[18], we refer to these as "conjoint" terms, though they are also known as Exclusion-Principle-Violating terms[19]) must also be considered. For example, take a pair of excitors that excite from the same orbital, say $\hat{a}_i^a$ and $\hat{a}_i^b$, where $\hat{a}_i^a = \hat{c}_i\hat{c}_a^\dagger$ and $\hat{c}_i$ ($\hat{c}_a^\dagger$) is a conventional annihilation (creation) operator. $\hat{a}_i^a\hat{a}_i^b = 0$ and hence $\langle D_{\mathbf{i}}|\hat{H}\hat{a}_i^a\hat{a}_i^b|D_{\mathrm{HF}}\rangle = 0$ for all $|D_{\mathbf{i}}\rangle$ and in the original form of stochastic coupled cluster any such cluster of excitors can be ignored. The argument applies equally to a pair of excitors that excite to the same orbital.

   When the similarity transformed Hamiltonian is used, such a cluster can give a non-zero contribution, however, as terms from the commutators, such as $\langle D_{\mathbf{i}}|\hat{a}_i^a\hat{H}\hat{a}_i^b|D_{\mathrm{HF}}\rangle$ are not necessarily 0. This means that the excitation in $\hat{H}$ must be chosen from the determinant produced by collapsing some subset of the cluster, $|D_{\mathbf{n}}\rangle$, here given by $\hat{a}_i^b|D_{\mathrm{HF}}\rangle = |D_i^b\rangle$, as collapsing the whole cluster just gives zero, not a valid excitor. From this, excitation $\langle\mathbf{m}|$ is chosen with non-zero $H_{\mathbf{mn}}$ to sample $\hat{H}$. The projectee $\langle D_{\mathbf{i}}|$ follows from applying the remaining excitors to $|D_{\mathbf{m}}\rangle$. This overall process corresponds to the product $\langle D_{\mathbf{i}}|\hat{a}_i^a|D_{\mathbf{m}}\rangle \langle D_{\mathbf{m}}|\hat{H}|D_{\mathbf{n}}\rangle \langle D_{\mathbf{n}}|\hat{a}_i^b|D_{\mathrm{HF}}\rangle$

Different partitions may give non-zero contributions to the commutator for different projectees, so to sample all possible spawning events in an unbiased manner requires randomly selecting a partition, using that to choose an allowed projectee, then evaluating the commutator by considering all possible (allowed) partitions.

3. All Hamiltonian matrix elements, in spawning and death, must be replaced by matrix elements of the similarity transformed Hamiltonian. This also means that the shift only applies to death on the reference, and there is no death for conjoint clusters. For instance when a cluster formed of a single excitor $\hat{a}_i^a$ is being considered and the excitation for the Hamiltonian chosen for spawning is $\hat{a}_j^b$, the probability of spawning an excip on $\left|D_{ij}^{ab}\right\rangle$ does not depend on $H_{\mathbf{mn}} = \langle D_{ij}^{ab}|\hat{H}|D_i^a\rangle$ but on $\langle D_{\mathbf{i}}|[\hat{H},\hat{T}]|D_{\mathrm{HF}}\rangle = \langle D_{ij}^{ab}|\hat{H}|D_i^a\rangle - \langle D_j^b|\hat{H}|D_{\mathrm{HF}}\rangle$.

Consider a system of three occupied orbitals, i, j and k, and three virtual orbitals, a, b and c. Some representative spawning attempts are:

1. A cluster consisting of a single excitor $\hat{a}_i^b$ is selected, giving the determinant $\left|D_i^b\right\rangle$ when applied to the reference. The single excitation to $\left\langle D_{ij}^{ab}\right|$ is chosen for the spawning attempt, which succeeds with probability proportional to $|\langle D_{ij}^{ab}|[\hat{H},\hat{a}_i^b]|D_{\mathrm{HF}}\rangle| = |\langle D_{ij}^{ab}|\hat{H}|D_i^b\rangle - \langle D_j^a|\hat{H}|D_{\mathrm{HF}}\rangle|$. Death also occurs on D_i^b with probability proportional to $\langle D_i^b|\hat{H}|D_i^b\rangle - \langle D_{\mathrm{HF}}|\hat{H}|D_{\mathrm{HF}}\rangle$.

2. The cluster composed of the two excitors $\hat{a}_i^a$ and $\hat{a}_j^b$ is chosen, and collapsed onto the reference to give $\left|D_{ij}^{ab}\right\rangle$. The single excitation to $\left|D_{ijk}^{abc}\right\rangle$ is chosen for spawning, but this excitation ($\hat{a}_k^c$) is not linked to the cluster so no spawning is attempted as all the terms in the commutator cancel. Death occurs on $\left|D_{ij}^{ab}\right\rangle$ with probability proportional to $\langle D_{ij}^{ab}|\hat{H}|D_{ij}^{ab}\rangle - \langle D_i^a|\hat{H}|D_i^a\rangle - \langle D_j^b|\hat{H}|D_j^b\rangle + \langle D_{\mathrm{HF}}|\hat{H}|D_{\mathrm{HF}}\rangle$.

3. The cluster composed of two excitors $\hat{a}_i^a$ and $\hat{a}_i^b$ is chosen. These both excite from the same occupied orbital i so do not collapse to give an excitor that can be applied to the reference. To choose an excitation for the spawning step, the subset of the cluster which does produce a determinant must be chosen. The subset containing only $\hat{a}_i^a$ is randomly chosen to be the latter, corresponding to the partitioning $\langle D_{\mathbf{i}}|\hat{a}_i^b\hat{H}\hat{a}_i^a|D_{\mathrm{HF}}\rangle$, giving the determinant $|D_i^a\rangle$ to spawn from. The double excitation ($\hat{a}_{aj}^{ic}$) to $\left|D_j^c\right\rangle$ is chosen in the Hamiltonian, and then the remainder of the original cluster $\hat{a}_i^b$ is applied to give spawning on $\left|D_{ij}^{bc}\right\rangle$ with probability proportional to $|\langle D_j^c|\hat{H}|D_i^a\rangle|$. There is no death attempt for conjoint clusters.
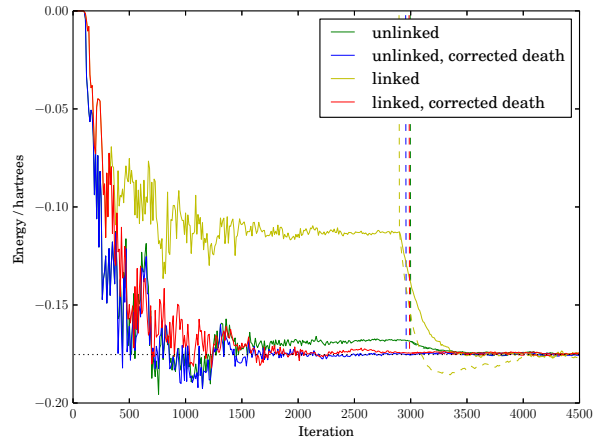

FIG. 1. The correlation energy calculated in CCMC for the pathological LiH_2^+ system[a] at the CCSD level; the projected energy is given by solid lines, the shift by dashed lines and the exact CCSD energy ($-0.175277\mathrm{E}_h$) by the black dotted line. An incorrect projected energy is obtained in the fixed-shift regime using the simple form of death. The shift and projected estimators fluctuate about the exact CCSD energy in the variable-shift regime in all cases.

[a] This was generated for minimal debugging purposes and has geometry $r_{\mathrm{LiH}_1} = 1\,\text{Å}$, $r_{\mathrm{LiH}_2} = 1.1\,\text{Å}$, and $\theta_{\mathrm{HLiH}} = 90\,^\circ$. In the STO-3G basis, unrestricted orbitals were prepared with a custom density functional with purely exchange, $E_{\mathrm{x}} = 2K + X_{\mathrm{Slater}} + X_{\mathrm{Becke88}}$. 2 core electrons and 8 virtual spinorbitals were frozen leaving an active space of 2 electrons in 4 spinorbitals. Q-Chem outputs and molecular orbital integrals for this system are included in the data repository.[20]

When the algorithm described above is used to calculate the coupled cluster energy, convergence to the correct projected energy is only achieved in the variable-shift regime, as shown in Fig. 1. The system shown in Fig. 1 is a pathological system displaying a very pronounced bias but the same behaviour is general to all systems examined, albeit to a lesser degree. Similar behaviour is observed in FCIQMC in a non-orthogonal basis[21]. This can be understood by considering the change in the coefficients when the projector is applied to the correct wavefunction. The desired form once the wavefunction has converged is $\Delta t_{\mathbf{i}} \propto t_{\mathbf{i}}$, as this ensures that the wavefunction being represented does not change, only the normalisation. Substituting the exact wavefunction into the CCMC update step gives for unlinked CCMC

$$\Delta t_{\mathbf{i}} = -\delta\tau t_{\mathrm{HF}}\left\langle D_{\mathbf{i}}\left|\hat{H} - S\right|\Psi_{\mathrm{CC}}\right\rangle \quad (6)$$

$$= -\delta\tau(E_{\mathrm{exact}} - S)\tilde{t}_{\mathbf{i}}, \quad (7)$$

where $\tilde{t}_{\mathbf{i}} = \langle D_{\mathbf{i}}|\Psi_{\mathrm{CC}}\rangle$ is the coefficient of $|D_{\mathbf{i}}\rangle$ in the coupled cluster wavefunction, and for linked CCMC

$$\Delta t_{\mathbf{i}} = -\delta\tau t_{\mathrm{HF}}\left\langle D_{\mathbf{i}}\left|\bar{H} - S\right|D_{\mathrm{HF}}\right\rangle \quad (8)$$

$$= \begin{cases} -\delta\tau t_{\mathrm{HF}}(E_{\mathrm{exact}} - S) & |D_{\mathbf{i}}\rangle = |D_{\mathrm{HF}}\rangle \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

For the original equations, the discrepancy $\tilde{t}_\mathbf{i} - t_\mathbf{i}$ is typically small as the connected amplitudes are the dominant contribution to the wavefunction. However there is a much larger discrepancy when using the similarity transformed Hamiltonian, giving rise to the larger deviation in the projected energy shown in Fig. 1. We emphasise that this is only true in the fixed-shift regime; once the population has grown to its desired size, the shift is varied to control the population and averages to the exact energy and hence the requirement on change in coefficients is trivially met on average.

This can be resolved by modifying the update step Eqs. (1) and (4) to be

$$t_\mathbf{i} \rightarrow t_\mathbf{i} - \delta\tau t_{\mathrm{HF}} \left\langle D_\mathbf{i} \left| \hat{H} - E_{\mathrm{CC}} \right| \Psi_{\mathrm{CC}} \right\rangle - \delta\tau(E_{\mathrm{CC}} - S)t_\mathbf{i}, \tag{10}$$

and for linked CCMC

$$t_\mathbf{i} \rightarrow t_\mathbf{i} - \delta\tau t_{\mathrm{HF}} \left\langle D_\mathbf{i} \left| \bar{H} - E_{\mathrm{CC}} \right| D_{\mathrm{HF}} \right\rangle - \delta\tau(E_{\mathrm{CC}} - S)t_\mathbf{i}, \tag{11}$$

both of which give $\Delta t_\mathbf{i} = -\delta\tau(E_{\mathrm{CC}} - S)t_\mathbf{i}$ when the $t_\mathbf{i}$ represent the exact wavefunction.

When $S = E_{\mathrm{CC}}$, these equations reduce to the original form, so in the variable shift regime the same solution is obtained. The difficulty with using the Eq. (10) and Eq. (11) as written is that the exact coupled cluster energy is not known during the calculation, as it is the quantity we aim to obtain. We therefore use the instantaneous value of the projected energy in its place as it is an easily available estimator. This does not make a difference during the variable shift phase of the calculation, during which statistics are accumulated, but reduces the equilibration time necessary.

This two term form of the propagation can be regarded as effecting a separation between two roles of the dynamics: the first term ensures the excip distribution represents a solution to the coupled cluster equations, while the second allows control of the total population via varying the shift.

The changes to the equations amount in practice to using the projected energy rather than the shift to control the rate of death on the composite clusters. This can be rationalised by noting that no population resides directly on the composite clusters, so such death processes do not contribute directly to population control, thus should not be dependent on the shift.

The population dynamics of the CCMC algorithm are similar to those of FCIQMC.[15] In particular, there is a plateau phase during the initial growth in the number of particles. This plateau occurs at a system-dependent number of excips and is the manifestation of the sign problem in CCMC. During the plateau phase, the total excip population remains constant, but the excip population on the reference continues to grow, so the plateau can be identified by finding the maximum of the ratio of the total to the reference population.[16] In this work we adopt the definition of the plateau height as the average population across the ten times with the largest

ratio of the total population to the reference population. We have found this to be an effective measure of the upper bound to the plateau in these systems, though note that it is somewhat dependent on the initial populations and timesteps[22], so we have kept these the same in the calculations we have compared.
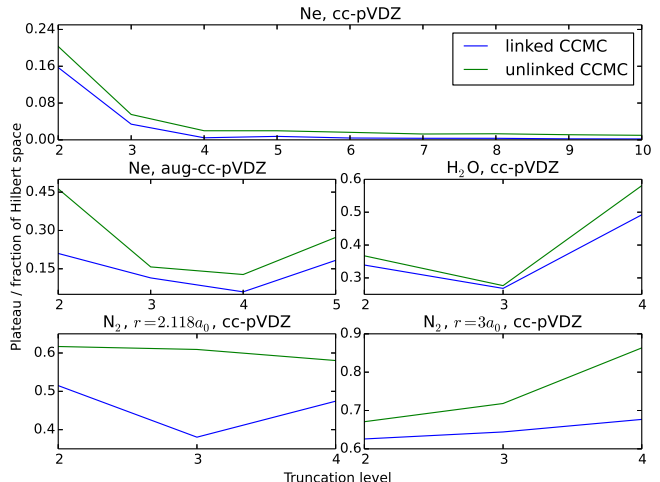

FIG. 2. Plateau heights as a fraction of the size of the Hilbert space for the neon atom and nitrogen and water (geometry from Ref. 23) molecules using linked and unlinked CCMC. In all cases, a lower plateau is obtained with the linked equations.

We have studied some small systems to see the effect of this algorithm. The coupled cluster energies are recovered with, as shown in Fig. 2, lower plateau heights than when using the original formulation of the algorithm. The effect is most pronounced at high truncation levels, which is to be expected as the termination of the expansion of $\bar{H}$ at fourth order will restrict the available clusters more for higher truncations. This means that the smaller number of connections in the effective Hamiltonian are better sampled, increasing the rate of annihilation.

Despite Eq. (1) and Eq. (4) solving equivalent equations for the amplitudes $t_\mathbf{i}$, the proportion of excitors on the reference is much larger when using the linked equations. The instantaneous populations fluctuate, so only the expectation values $\langle N_\mathbf{i} \rangle$ correspond to the correct amplitudes. When considering the total population, however, $N_{\mathrm{tot}} = \sum_\mathbf{i} |N_\mathbf{i}|$ is a non-linear function of $N_\mathbf{i}$, so $\langle N_{\mathrm{tot}} \rangle \geq \sum_\mathbf{i} |\langle N_\mathbf{i} \rangle|$. Excitors on which the sign of the population is not constant raise the value of $\langle N_{\mathrm{tot}} \rangle$. As the linked equations reduce the number of low amplitude excitors that are instantaneously occupied at a given iteration, they give a lower $\langle N_{\mathrm{tot}} \rangle$ for a fixed population of excitors on the reference. We are investigating the impact of this but believe that, due to the presence of the population on the reference in the cluster generation probabilities, this results in more stable calculations.

In conclusion, we have demonstrated that this linked approach makes stochastic coupled cluster calculations

more feasible, especially at the higher truncation levels, which are not typically used but are necessary for an accurate treatment of many chemically interesting problems. We have also presented a modification to the death step for both linked and unlinked formalisms to remove a bias in projected energy estimators while the population is still growing, and we believe that this will be useful in other cases where there is non-linearity in the projector.

By placing CCMC in the same theoretical framework as conventional coupled cluster theory, namely using the similarity transformation of the Hamiltonian, we hope to be able to translate further technical and theoretical developments to our stochastic implementation. We finish by emphasising two important benefits of a stochastic approach to coupled cluster: the substantial reduction in memory demands, by allowing sparsity in the amplitudes to emerge naturally, and access to an efficient coupled cluster algorithm that can handle arbitrary truncation levels.

All CCMC calculations were performed using a development version of HANDE[24,25], with one- and two-body molecular integrals obtained from Hartree-Fock calculations performed in Q-Chem.[26] All data was analysed using pyblock[27], and plots produced using matplotlib[28]. Raw and analysed data and analysis scripts are available at Ref. 20.

ACKNOWLEDGEMENTS

[1] J. Čížek, J. Chem. Phys. **45**, 4265 (1966).

[2] J. Čížek and J. Paldus, Int. J. Quantum Chem. **5**, 359 (1971).

[3] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).

[4] A. Grüneis, G. H. Booth, M. Marsman, J. Spencer, A. Alavi, and G. Kresse, J. Chem. Theor. Comput. **7**, 2780 (2011).

[5] A. Grüneis, The Journal of Chemical Physics **143**, 102817 (2015), http://dx.doi.org/10.1063/1.4928645.

[6] G. H. Booth, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009).

[7] G. H. Booth and A. Alavi, J. Chem. Phys. **132**, 174104 (2010).

[8] G. H. Booth, D. Cleland, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **135**, 084104 (2011).

[9] D. Cleland, G. H. Booth, C. Overy, and A. Alavi, J. Chem. Theor. Comput. **8**, 4138 (2012).

[10] G. H. Booth, A. Grüneis, G. Kresse, and A. Alavi, Nature **493**, 365 (2013).

[11] G. H. Booth, D. Cleland, A. Alavi, and D. P. Tew, J. Chem. Phys. **137**, 1 (2012).

[12] C. Overy, G. H. Booth, N. S. Blunt, J. J. Shepherd, D. Cleland, and A. Alavi, J. Chem. Phys. **141**, 1 (2014).

[13] J. J. Shepherd, G. E. Scuseria, and J. S. Spencer, Phys. Rev. B **90**, 155130 (2014).

[14] A. J. W. Thom, Phys. Rev. Lett. **105**, 263004 (2010).

[15] J. S. Spencer, N. S. Blunt, and W. M. Foulkes, The Journal of Chemical Physics **136**, 054110 (2012).

[16] J. S. Spencer and A. J. W. Thom, (2015), arXiv:1511.05752 [physics.chem-ph].

[17] R. D. Mattuck, *Guide to Feynman diagrams in the many-body problem* (McGraw-Hill Pub. Co, London, New York [etc.] McGraw-Hill Pub. Co. [1967], 1967).

[18] R. J. Bartlett and G. D. Purvis, International Journal of Quantum Chemistry **14**, 561 (1978).

[19] I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory* (Cambridge University Press, 2009).

[20] R. Franklin, J. S. Spencer, A. Zoccante, and A. J. W. Thom, "Data from "Linked Coupled Cluster Monte Carlo"," (2015), http://dx.doi.org/10.5281/zenodo.34211.

[21] S. Smart, *The use of spin-pure and non-orthogonal Hilbert spaces in Full Configuration Interaction Quantum Monte-Carlo*, Ph.D. thesis, University of Cambridge (2013).

[22] W. A. Vigor, J. S. Spencer, M. J. Bearpark, and A. J. W. Thom, (2016), arXiv:1601.00865 [physics.chem-ph].

[23] J. Olsen, P. Jørgensen, H. Koch, A. Balkova, and R. J. Bartlett, The Journal of Chemical Physics **104**, 8007 (1996).

[24] J. S. Spencer, N. S. Blunt, W. A. Vigor, F. D. Malone, W. M. C. Foulkes, J. J. Shepherd, and A. J. W. Thom, in *2nd Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE2)* (2014).

[25] J. S. Spencer, N. S. Blunt, W. A. Vigor, F. D. Malone, W. M. C. Foulkes, J. J. Shepherd, and A. J. W. Thom, J. Open Research Software **3** (2015), http://doi.org/10.5334/jors.bw.

[26] Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Ku, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. W. III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. D. Jr., H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. Hanson-Heine, P. H. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. ONeill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stck, Y.-C. Su, A. J. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. Chan, D. M. Chipman, C. J. Cramer, W. A. G. III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. S. III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. V. Voorhis, J. M. Herbert, A. I. Krylov, P. M. Gill, and M. Head-Gordon, Mol. Phys. **113**, 184 (2015).

[27] J. Spencer, "pyblock," https://github.com/jsspencer/pyblock.

[28] J. Hunter, Computing in Science & Engineering **9**, 90 (2007).