A Stochastic Approach to Unitary Coupled Cluster

Maria-Andreea Filip[1, a)] and Alex J W Thom[1, b)]

Department of Chemistry, University of Cambridge, Cambridge, UK

Unitary coupled cluster (UCC), originally developed as a variational alternative to the popular traditional coupled cluster method, has seen a resurgence as a functional form for use on quantum computers. However, the number of excitors present in the *ansatz* often presents a barrier to implementation on quantum computers. Given the natural sparsity of wavefunctions obtained from Quantum Monte Carlo methods, we consider here a stochastic solution to the UCC problem. Using the Coupled Cluster Monte Carlo framework, we develop cluster selection schemes that capture the structure of the UCC wavefunction, as well as its Trotterized approximation, and use these to solve the corresponding projected equations. Due to the fast convergence of the equations with order in the cluster expansion, this approach scales polynomially with the size of the system. Unlike traditional UCC implementations, our approach naturally produces a non-variational estimator for the energy in the form of the projected energy. For UCCSD in small systems, we find this agrees well with the expectation value of the energy and, in the case of two electrons, with full configuration interaction results. For the larger N_2 system, the two estimators diverge, with the projected energy approaching the coupled cluster result, while the expectation value is close to results from traditional UCCSD.

I. INTRODUCTION

Coupled cluster[1,2] (CC) theory has long been one of the most popular *ab initio* methods in quantum chemistry, due to its balance of high-accuracy, size-consistency, polynomial scaling and systematic improvability, with its truncation at single and double excitations with perturbative triples[3] (CCSD(T)) being considered the industry "gold-standard". However, to obtain these appealing features, one has to sacrifice the variationality of the method. Unlike energies obtained from methods such as Hartree–Fock (HF) theory, configuration interaction (CI) or even density functional theory (DFT), coupled cluster projected energies do not obey the variational principle, so there is no guarantee that they will be an upper bound on the ground state energy. At times, this allows for catastrophic behaviour, such as the well known collapse of the CCSD(T) energy in the case of strongly correlated systems, such as N_2.[4]

A variational formulation of coupled cluster is therefore a tempting proposition and many attempts have been made to develop such formulations,[5] leading to methods such as variational coupled cluster,[6,7] extended coupled cluster[8] and unitary coupled cluster.[9–15] However, these often involve non-terminating series for the energy, which cannot necessarily be truncated intuitively. Computational scaling is also increased, becoming exponential for methods like variational and unitary coupled cluster.

While these various issues have stopped alternative CC methods from becoming widely adopted in the quantum chemistry community, recently there has been renewed interest in the unitary coupled cluster formalism, due to its applicability as a wavefunction *ansatz* for quantum computers, which are emerging as a promising tool for computational chemistry. The qubit model of computation naturally encodes the exponentially sized Hilbert space of a molecule in a linear number of qubits. By mapping each qubit to a spin-orbital and appropriately entangling them, one can encode a FCI wavefunction in a number of qubits equal to the number of spin-orbitals in the basis, rather than requiring the storage of $\binom{N_{\text{spinorb}}}{N_{\text{elec}}}$ determinant coefficients. In principle, on an ideal quantum computer, one can obtain a wavefunction with good overlap with the true ground state wavefunction using adiabatic state preparation[16] and trotterized time evolution[17,18], where a guess wavefunction is propagated to the ground state, in an approach reminiscent of DMC[19,20] and FCIQMC[21]. The true ground state energy can then be computed using Quantum Phase Estimation[22,23] (QPE). However, current quantum computers, known as Noisy Intermediate-Scale Quantum (NISQ) machines, are limited in both number of qubits and qubit decoherence time, as well as being subject to significant noise in the results, making the application of such a resource–intensive algorithm infeasible. However, alternatives tailored to such machines have been developed. One such algorithm is the Variational Quantum Eigensolver[24,25] (VQE) where an *ansatz*-based wavefunction is prepared on a quantum computer, its energy is measured and a classical optimisation algorithm is used to minimise the energy and get optimal wavefunction coefficients. As all operations available to a quantum computer must be representable by a unitary transformation, UCC has resurfaced as an appropriate *ansatz* for this algorithm.[25–27]

The number of excitation operators in the expansion can be a limiting factor in the use of UCC *ansatze*, as it increases the required quantum circuit depth. Therefore, it would be useful to develop a method to

―――――

a) Electronic mail: maf63@cam.ac.uk
b) Electronic mail: ajwt3@cam.ac.uk

pre-select the most important cluster operators before devising the circuit. Quantum Monte Carlo (QMC) methods, such as FCIQMC[21] and CCMC[28] produce naturally sparse representations of the wavefunction, as only significant contributors are meaningfully sampled by the propagation scheme. Therefore, they seem ideal candidates for the screening of cluster amplitudes. This idea has been suggested before as a means to only include the most important triple and quadruple excitations in a conventional CC calculation, with good results.[29–31] We would be interested in assessing its applicability in screening UCC amplitudes as well, and will pursue this in further work.

In this paper, we set out to detail the algorithmic steps involved in the implementation of a projective variant of UCC, primarily in a stochastic framework, as well as discussing the results obtained from this approach on a series of test systems. In the following section, we review theoretical aspects of coupled cluster theory and its unitary formalism. We then revisit the CCMC algorithm in section III and discuss its expansion to UCC in section IV. Section V presents a series of benchmark results and Section VI brings together our conclusions.

II. COUPLED CLUSTER THEORY

A. Traditional Coupled Cluster

In traditional coupled cluster, the wavefunction is given by an exponential *ansatz*

$$\Psi_{\mathrm{CC}} = e^{\hat{T}} |D_0\rangle, \tag{1}$$

where $|D_0\rangle$ is the Hartree–Fock reference wavefunction and the cluster operator

$$\hat{T} = \sum_i \hat{T}_i \tag{2}$$

where operators $\hat{T}_i$ generate determinants of excitation level i, *i.e.*

$$\hat{T}_1 = \sum_{i,a} t_i^a \hat{a}^\dagger \hat{i} \tag{3}$$

$$\hat{T}_2 = \frac{1}{4} \sum_{i,j,a,b} t_{ij}^{ab} \hat{a}^\dagger \hat{b}^\dagger \hat{i} \hat{j} \tag{4}$$

and so on. In this context, $\hat{p}^\dagger$ and $\hat{p}$ represent creation and annihilation operators for orbital p respectively. In the summations, $i, j, \ldots$ range over the occupied orbitals in the reference and $a, b, \ldots$ range over the virtual orbitals. The CC wavefunction is equivalent to the FCI wavefunction if all possible excitors are included.

However, one can truncate $\hat{T}$ to a certain excitation level, giving a progression of increasingly accurate methods: CCSD ($\hat{T} = \hat{T}_1 + \hat{T}_2$), CCSDT ($\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3$), CCSDTQ ($\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \hat{T}_4$) and so on.

In order to obtain t_i, the Schrödinger equation is projected onto each of the determinants $|D_i\rangle$, leading to a series of coupled cluster equations to be solved:

$$\langle D_i | \hat{H} - E | \Psi_{\mathrm{CC}} \rangle = 0, \tag{5}$$

where E is the energy of Ψ_{CC}. One can instead use a similarity transformed Hamiltonian $\bar{H} = e^{-\hat{T}} \hat{H} e^{\hat{T}}$, giving an equivalent set of equations:

$$\langle D_i | \bar{H} - E | D_0 \rangle = 0, \tag{6}$$

The Campbell–Baker–Hausdorff (CBH)[32–34] expansion of the similarity transformed Hamiltonian

$$\bar{H} = \hat{H} + [\hat{H}, \hat{T}] + \frac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{3!}[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}] + \\ + \frac{1}{4!}[[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}], \hat{T}] \tag{7}$$

terminates at fourth order, guaranteeing a finite number of terms in the projected CC equations. The time required to computationally solve these equations scales as $\mathcal{O}(N^{2n+2})$ where N is the size of the system and n is the truncation level. Therefore, CCSD scales as $\mathcal{O}(N^6)$, CCSDT as $\mathcal{O}(N^8)$ and so on.

Having solved these projected equations, one typically uses the projected energy

$$E_{\mathrm{proj}} = \langle D_0 | \hat{H} | \Psi_{\mathrm{CC}} \rangle \tag{8}$$

as an estimator for the energy of the wavefunction. Where coupled cluster is able to capture the true wavefunction, this should be equal to the expectation value of the energy,

$$\langle E \rangle_{\mathrm{CC}} = \frac{\langle \Psi_{\mathrm{CC}} | \hat{H} | \Psi_{\mathrm{CC}} \rangle}{\langle \Psi_{\mathrm{CC}} | \Psi_{\mathrm{CC}} \rangle}. \tag{9}$$

However, where coupled cluster is not exact, this equality is not guaranteed.

B. Unitary Coupled Cluster

Consider the anti-Hermitian operator

$$\hat{\tau} = \hat{T} - \hat{T}^\dagger, \tag{10}$$

where as before $\hat{T}$ can be truncated in excitation level. $e^{\hat{\tau}}$ is then a unitary operator and can be used to define a UCC wavefunction

$$|\Psi_{\mathrm{UCC}}\rangle = e^{\hat{\tau}} |D_0\rangle \tag{11}$$

The expectation value of the energy is then given by

$$\langle E \rangle_{\text{UCC}} = \frac{\langle \Psi_{\text{UCC}} | \hat{H} | \Psi_{\text{UCC}} \rangle}{\langle \Psi_{\text{UCC}} | \Psi_{\text{UCC}} \rangle} = \frac{\langle D_0 | e^{-\hat{\tau}} \hat{H} e^{\hat{\tau}} | D_0 \rangle}{\langle D_0 | e^{-\hat{\tau}} e^{\hat{\tau}} | D_0 \rangle} \quad (12)$$
$$= \langle D_0 | e^{-\hat{\tau}} \hat{H} e^{\hat{\tau}} | D_0 \rangle$$

The cluster coefficients are then usually obtained by setting

$$\frac{\partial \langle E \rangle_{\text{UCC}}}{\partial t_{\mathbf{i}}} = 0 \quad (13)$$

for all $t_{\mathbf{i}}$ in the expansion.

For $\bar{H} = e^{-\hat{\tau}} \hat{H} e^{\hat{\tau}}$, the CBH expansion

$$\bar{H} = \hat{H} + [\hat{H}, \hat{T}] + [\hat{T}^{\dagger}, \hat{H}] \\ + \frac{1}{2}([[\hat{H}, \hat{T}], \hat{T}] + [T^{\dagger}, [\hat{T}^{\dagger}, \hat{H}]] + [\hat{H}, [\hat{T}, \hat{T}^{\dagger}]]) + ... \quad (14)$$

no longer terminates at a finite order. This is due to the presence of mixed terms such as $[\hat{H}, [\hat{T}, \hat{T}^{\dagger}]]$ in eq. (14), which have no termination point,[5] leading to infinite series for the energy and amplitude equations. Most implementations of UCC rely on some truncation of these expressions, either to a particular order in perturbation theory[15] or to a particular power of $\hat{\tau}$.

While this has not been the most common approach taken in the past, it is also possible[35] to solve a set of projected UCC equations,

$$\langle D_{\mathbf{i}} | \hat{H} - E | \Psi_{\mathbf{UCC}} \rangle = 0, \quad (15)$$

or their similarity transformed counterpart,

$$\langle D_{\mathbf{i}} | \bar{H} - E | D_0 \rangle = 0. \quad (16)$$

Much like traditional coupled cluster and variational coupled cluster do not in general lead to the same energy or indeed wavefunction, we expect that projective and variational UCC will also generate different results. We will focus our attention on the projective method as it is naturally more amenable to the QMC algorithms described below. An example of the differences between these methods can be found in section V.

Deterministically solving the projected UCC equations also requires a truncation of the exponential to make the computation tractable. In our case, we have chosen to truncate at a given order of $\hat{\tau}$ and have found that the results converge rapidly with increasing truncation order, as is shown in one of the following sections. Using the truncated exponential, one can naively solve the residual equations iteratively, by starting from a Hartree–Fock wavefunction ($\forall \mathbf{i} : t_{\mathbf{i}} = 0$) and taking steps of the form

$$t_{\mathbf{i}}(\beta + \delta\beta) = t_{\mathbf{i}}(\beta) - \delta\beta \langle D_{\mathbf{i}} | \hat{H} - E_{\text{proj}} | \Psi_{\text{UCC}} \rangle, \quad (17)$$

where E is the current projected energy estimate, $E_{\text{proj}} = \langle D_0 | \hat{H} | \Psi_{\text{UCC}} \rangle / \langle D_0 | \Psi_{\text{UCC}} \rangle$. We find that this approach converges well, albeit slowly. This could likely be improved by using more involved optimisation algorithms.

Another approach to simplify the UCC *ansatz* involves the Suzuki–Trotter decomposition[36,37] of the exponential,

$$e^{\hat{\tau}} = e^{\sum_{\mathbf{i}} \hat{\tau}_{\mathbf{i}}} \approx (\prod_{\mathbf{i}} e^{\frac{\hat{\tau}_{\mathbf{i}}}{\rho}})^{\rho}, \quad (18)$$

where the index $\mathbf{i}$ in eq. (18) runs over all possible excitations. This approximation recovers the original *ansatz* in the limit that $\rho \to \infty$, but recently it has been shown that a value of $\rho = 1$ is sufficient to give an exact parametrisation of the wavefunction, provided the operators are arranged in a particular order.[38] Having replaced the full UCC *ansatz* with a trotterized form, the resulting equations can then be solved either variationally or by projection, as discussed above.

III. COUPLED CLUSTER MONTE CARLO

The CC equations can be equivalently written as

$$\langle D_{\mathbf{i}} | 1 - \delta\beta(\hat{H} - E) | \Psi_{\text{CC}} \rangle = \langle D_{\mathbf{i}} | \Psi_{\text{CC}} \rangle \quad (19)$$

and recast in an iterative form as[39]

$$t_{\mathbf{i}}(\beta + \delta\beta) = t_{\mathbf{i}}(\beta) - \delta\beta \langle D_{\mathbf{i}} | \hat{H} - E | \Psi_{\text{CC}} \rangle \quad (20)$$

The iterative equations describe the population dynamics of a set of particles in the Hilbert space. This dynamics is controlled by two processes, corresponding to the action of diagonal and off-diagonal Hamiltonian elements respectively:

1. death/cloning of particles on $|D_{\mathbf{i}}\rangle$.

2. spawning of a particle from $|D_{\mathbf{i}}\rangle$ onto another $|D_{\mathbf{j}}\rangle$ coupled to it by the action of the Hamiltonian

A third process, annihilation, where pairs of particles of opposite signs on the same excitor are removed, is required to mitigate the sign problem in the algorithm.[21] These processes allow for the variation of the normalisation of the wavefunction, so one can write the CCMC wavefunction as $\Psi_{\text{CCMC}} = N_0 e^{\sum_{\mathbf{i}} \frac{N_{\mathbf{i}}}{N_0} \hat{a}_{\mathbf{i}}}$, where N_0 corresponds to the population on the reference and $N_{\mathbf{i}}$ to the population on the excitor $\mathbf{i}$.

Such a calculation provides two estimators for the correlation energy of a system, which should agree once the population dynamics have reached a steady–state

corresponding to the ground state wavefunction. Firstly, one can compute the instantaneous projected energy

$$E_{\text{proj}} = \frac{\langle D_0 | \hat{H} | \Psi_{\text{CCMC}} \rangle}{\langle D_0 | \Psi_{\text{CCMC}} \rangle}. \tag{21}$$

Secondly, a shift S is introduced to replace the unknown E in eq. (20) and act as a population-control parameter. If it is allowed to vary such that a stable particle population is maintained,[21] the shift will converge onto the value of the correlation energy.

The population dynamics described above can be performed stochastically, by sampling the first two processed with probabilities given by the values of the relevant Hamiltonian terms. In the original implementation of CCMC, the sampling was carried out in the following way:[40]

1. a cluster size s is selected with probability

$$p(s) = \frac{1}{2^{s+1}} \tag{22}$$

2. a particular cluster of s distinct excitors is selected with probability

$$p(e|s) = s! \prod_{i=1}^{s} \frac{|N_i|}{|N_{\text{ex}}|} \tag{23}$$

where N_{ex} is the total population on excitors. The total selection probability is therefore

$$p_{\text{sel}}(e) = p(e|s)p(s) \tag{24}$$

3. spawning from $|D_{\mathbf{i}}\rangle$ to $|D_{\mathbf{j}}\rangle$ creates a particle with probability

$$p_{\text{spawn}} = \delta\beta \frac{|w_e|}{n_{\text{a}} p_{\text{sel}}(e)} \frac{|H_{\mathbf{ij}}|}{p_{\text{gen}}}, \tag{25}$$

where w_e is the total amplitude on the cluster which collapses to $|D_{\mathbf{i}}\rangle$, n_{a} is the number of spawning attempts and p_{gen} is the probability of generating $|D_{\mathbf{j}}\rangle$

4. death occurs with probability

$$p_{\text{death}} = \delta\beta \frac{|w_e|}{n_{\text{a}} p_{\text{sel}}(e)} |H_{\mathbf{ii}} - S| \tag{26}$$

More recently, an importance sampling based method has been developed,[40] which allows the term $\frac{|w_e|}{n_{\text{a}} p_{\text{sel}}(e)}$ to be provably equal to 1 for all cluster sizes, thereby decreasing the time spent sampling large clusters that are unlikely to contribute.

Further improvements of the algorithm have been implemented, including an initiator approximation,[39] a linked approach,[41] efficient parallelisation[42] and excitation generators.[43,44]

A. Variants of stochastic coupled cluster

What we have described above is the initial, unlinked implementation of Coupled Cluster Monte Carlo. Since then a linked CCMC formalism[41], as well as a diagrammatic version of CCMC[45] have been developed.

In linked CCMC, one follows a similar procedure to that described above, but sampling the action of the similarity transformed Hamiltonian $\bar{H} = e^{-\hat{T}} \hat{H} e^{\hat{T}}$, rather than $\hat{H}$. This involves sampling the four commutators in Equation (7), which requires some significant changes to the selection algorithm described above.[41] Firstly, and helpfully, clusters may have at most four excitors. Secondly, excitors which act on some of the same orbitals may give a non-zero contribution to the energy and must therefore be considered. Thirdly, once a particular set of excitors has been selected, all possible orderings of the excitors and the Hamiltonian in the commutator must be considered simultaneously, to maintain the benefit of only sampling connected terms. Finally, for spawning and death, one must build the relevant terms of the similarity transformed Hamiltonian. The final three steps all induce significant added complexity to the CCMC algorithm, however this is offset by the reduction in size of the expansion that needs to be considered.

In contrast, in diagrammatic CCMC, one considers the similarity transformed normal ordered Hamiltonian,

$$\begin{aligned} \bar{H}_{\text{N}} &= e^{-\hat{T}} H_{\text{N}} e^{\hat{T}} \\ &= \hat{H}_{\text{N}} + \overline{H_{\text{N}}T} + \overline{H_{\text{N}}TT} + \overline{H_{\text{N}}TTT} + \overline{H_{\text{N}}TTTT} \end{aligned} \tag{27}$$

for which the CBH expansion reduces to the connected contributions shown above, which can be represented in diagrammatic form.[46] Since the cluster operator $\hat{T}$ only appears to the right of the Hamiltonian in Equation (27), these terms are simpler to sample than the corresponding commutators in Equation (7). The cluster amplitudes can be found by sampling the update equation

$$t_{\mathbf{i}}(\beta + \delta\beta) = t_{\mathbf{i}}(\beta) - \delta\beta \langle D_{\mathbf{i}} | \bar{H}_{\text{N}} | D_0 \rangle. \tag{28}$$

This can be done by selecting particular diagrams relevant to the second term on the right hand side and computing their contributions, as described in Ref. 45. While the selection scheme required is still more involved than the one employed in unlinked CCMC, it avoids some of the complications of the linked formalism, while also more strongly imposing connectivity constraints on the considered diagrams, leading to promising performance improvements over the unlinked formalism.

The main benefit these approaches have over the unlinked CCMC formalism comes from their ability to encode the terminating CBH expansion of the similarity transformed Hamiltonian, therefore guaranteeing that fewer clusters need to be considered. For UCC this expansion is non-terminating and therefore the benefits of employing the linked or diagrammatic formalisms

are much diminished, while the additional computational complexities remain. Therefore, in our initial stochastic approach to the UCC problem, we will focus on the original, unlinked formalism, whose relative computational simplicity provides an ideal testing ground for the feasibility of such an endeavour.

IV. STOCHASTIC UNITARY COUPLED CLUSTER

The stochastic implementation of UCC (herein referred to as UCCMC) is based around the same principles as that of traditional CCMC. One must select a cluster amplitude and allow it to undergo spawning, death and annihilation. Modifications that must be made to the original algorithm to accommodate for the change in the cluster operator are detailed below, for both full and trotterized UCCMC.

A. Full unitary coupled cluster

The presence of deexcitation operators $\hat{T}^\dagger$ in the full UCC *ansatz* substantially changes the structure of the allowed clusters. In traditional CCMC, the largest allowed excitation level of any considered cluster is $n + 2$, where n is the considered truncation level, as this is the highest order excitation that couples to the CCMC wavefunction through the Hamiltonian. As such a cluster could be formed from at most $n+2$ single excitors, this is also the largest size of cluster one must consider. In UCC however, the inclusion of deexcitation operators can lower the overall excitation level of the cluster while increasing its size, so this heuristic no longer holds. Therefore, in principle, for an implementation of UCCMC, one must consider clusters of up to infinite size. However, as in the deterministic case, one can truncate the expansion to a finite size of cluster. Indeed, in the original implementation of CCMC in HANDE-QMC,[47] the maximum allowed cluster size is 12, due to technical limitations in computing a factorial. This size limitation has been preserved, and we have found that even for larger systems like N_2, valid clusters of size 12 are sampled extremely rarely, so increasing the polynomial truncation level would not improve the precision of our current algorithm. Having taken this into consideration, the selection scheme for UCCMC is as follows:

1. select a cluster size s with probability $p(s) = \frac{1}{2^{s+1}}$. Other distributions, such as the uniform distribution or distributions where $p(s)$ increases with s, have been attempted here, with little to no effect on the quality of the results. However, as in the case of CCMC, it is possible that tweaking the selection probabilities may improve the efficiency of the algorithm.

2. for all but the first excitor in the cluster, decide with probability $\frac{1}{2}$ whether it will be an excitation

or deexcitation operator

3. a particular cluster is selected as before, with probability given by eq. (23)

Having selected the cluster, it undergoes spawning and death as before. The final aspect one must be careful of is the normalisation of the HF reference. In the case of CCMC and FCIQMC, the population on the reference was equal to the wavefunction projection onto the reference determinant, $\langle D_0 | \Psi_{CCMC} \rangle = N_0$. Consequently, other excitor populations can be normalised relative to this to give an intermediately normalised wavefunction as naturally arises from the exponential form of the *ansatz* and the correlation energy can be computed as

$$E_{\text{proj}} = \sum_{i \neq 0} \frac{N_i^{\text{CI}}}{N_0} H_{i0} \quad (29)$$

where N_i^{CI} is the equivalent CI population on a determinant, obtained by sampling all combinations of different-sized clusters collapsing onto that determinant. However, for UCCMC, it is trivial to show that $\langle D_0 | \Psi_{UCC} \rangle \neq 1$. Consider the Taylor expansion of the exponential

$$e^{\hat{T}-\hat{T}^\dagger} = 1+\hat{T}-\hat{T}^\dagger+\frac{1}{2}(\hat{T}\hat{T}-\hat{T}\hat{T}^\dagger-\hat{T}^\dagger\hat{T}+\hat{T}^\dagger\hat{T}^\dagger)+... \quad (30)$$

If we further expand $\hat{T} = \sum_i t_i \hat{a}_i$, where i indexes over all allowed excitors and $\hat{a}_i$ is the corresponding excitation operator, then the mixed term $\hat{T}^\dagger\hat{T}$ becomes

$$\hat{T}^\dagger\hat{T} = \sum_i t_i^2 \hat{a}_i^\dagger \hat{a}_i + \sum_{i \neq j} t_i t_j \hat{a}_i^\dagger \hat{a}_j = \sum_i t_i^2 \hat{I} + \sum_{i \neq j} t_i t_j \hat{a}_i^\dagger \hat{a}_j \quad (31)$$

Therefore, this mixed squared term contributes to the projection onto the reference determinant with $\frac{1}{2}\sum_i t_i^2$. Similarly, any even powered term of the form $\hat{T}^\dagger\hat{T}\hat{T}^\dagger\hat{T}...\hat{T}^\dagger\hat{T}$ will have a similar contribution, leading to

$$\langle D_0 | \Psi_{UCC} \rangle = 1 + \sum_i \left(\sum_{p=1}^\infty (-1)^p \frac{t_i^{2p}}{2p!} \right) \quad (32)$$

Considering a UCCMC wavefunction $\Psi_{UCCMC} = N_0 e^{\sum_i \frac{N_i}{N_0}(\hat{a}_i - \hat{a}_i^\dagger)}$,

$$\langle D_0 | \Psi_{UCCMC} \rangle = N_0 \left\{ 1 + \sum_i \left[\sum_{p=1}^\infty (-1)^p \left(\frac{N_i}{N_0} \right)^{2p} \right] \right\} \quad (33)$$

Therefore, when one normalises the cluster coefficients relative to N_0, the projected energy becomes:

$$E_{\text{proj}} = \sum_{i \neq 0} \frac{N_i^{\text{CI}} H_{i0}}{N_0 \left\{ 1 + \sum_i \left[\sum_{p=1}^\infty (-1)^p \left(\frac{N_i}{N_0} \right)^{2p} \right] \right\}} \quad (34)$$

The denominator does not need to be explicitly computed and can be sampled stochastically during the course of the calculation, concurrently with N_0 and $N_i^{\text{CI}} H_{i0}$.

B. Trotterized unitary coupled cluster

Consider the trotterized UCC *ansatz* with $\rho = 1$.

$$|\Psi_{\text{tUCC}}\rangle = \prod_{\mathbf{i}} e^{\hat{\tau}_{\mathbf{i}}} |D_0\rangle \qquad (35)$$

A particular term $e^{\hat{\tau}_{\mathbf{i}}}$ can be expanded as

$$e^{\hat{\tau}_{\mathbf{i}}} = 1 + t_{\mathbf{i}}(\hat{a}_{\mathbf{i}} - \hat{a}_{\mathbf{i}}^{\dagger}) + \frac{t_{\mathbf{i}}^2}{2!}(\hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}} - \hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}}^{\dagger} - \hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}} + \hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}}^{\dagger})$$
$$+ \frac{t_{\mathbf{i}}^3}{3!}(\hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}} - \hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}}^{\dagger} - \hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}} - \hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}} + \hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}}^{\dagger}$$
$$+ \hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}}^{\dagger} + \hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}} - \hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}}^{\dagger}) + ... \qquad (36)$$

Any term that that applies two of the same operator sequentially vanishes when applied to any wavefunction, so this can be rewritten

$$e^{\hat{\tau}_{\mathbf{i}}} = 1 + t_{\mathbf{i}}(\hat{a}_{\mathbf{i}} - \hat{a}_{\mathbf{i}}^{\dagger}) + \frac{t_{\mathbf{i}}^2}{2!}(-\hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}}^{\dagger} - \hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}})$$
$$+ \frac{t_{\mathbf{i}}^3}{3!}(-\hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}} + \hat{a}_{\mathbf{i}}^{\dagger}\hat{a}_{\mathbf{i}}\hat{a}_{\mathbf{i}}^{\dagger}) + ... \qquad (37)$$

Consider applying $e^{\hat{\tau}_{\mathbf{i}}}$ to an arbitrary single determinant wavefunction $|\Psi\rangle$. There are then three possibilities:

CASE 1: $\hat{a}_{\mathbf{i}}^{\dagger} |\Psi\rangle = 0$ and $\hat{a}_{\mathbf{i}} |\Psi\rangle \neq 0$

$$e^{\hat{\tau}_{\mathbf{i}}} |\Psi\rangle = \sum_{p=0}^{\infty} (-1)^p \Big(\frac{t_{\mathbf{i}}^{2p}}{2p!} \hat{I} + \frac{t_{\mathbf{i}}^{2p+1}}{(2p+1)!} \hat{a}_{\mathbf{i}} \Big) |\Psi\rangle$$
$$= \cos(t_{\mathbf{i}}) |\Psi\rangle + \sin(t_{\mathbf{i}}) |\Psi_{\mathbf{i}}\rangle \qquad (38)$$

where $|\Psi_{\mathbf{i}}\rangle$ is the result of applying the excitation to $|\Psi\rangle$.

CASE 2: $\hat{a}_{\mathbf{i}}^{\dagger} |\Psi\rangle \neq 0$ and $\hat{a}_{\mathbf{i}} |\Psi\rangle = 0$

$$e^{\hat{\tau}_{\mathbf{i}}} |\Psi\rangle = \sum_{p=0}^{\infty} (-1)^p \Big(\frac{t_{\mathbf{i}}^{2p}}{2p!} \hat{I} - \frac{t_{\mathbf{i}}^{2p+1}}{(2p+1)!} \hat{a}_{\mathbf{i}}^{\dagger} \Big) |\Psi\rangle$$
$$= \cos(t_{\mathbf{i}}) |\Psi\rangle - \sin(t_{\mathbf{i}}) |\Psi^{\mathbf{i}}\rangle \qquad (39)$$

where $|\Psi^{\mathbf{i}}\rangle$ is the result of applying the deexcitation to $|\Psi\rangle$.

CASE 3: $\hat{a}_{\mathbf{i}}^{\dagger} |\Psi\rangle = 0$ and $\hat{a}_{\mathbf{i}} |\Psi\rangle = 0$

$$e^{\hat{\tau}_{\mathbf{i}}} |\Psi\rangle = |\Psi\rangle \qquad (40)$$

To translate this into a stochastic algorithm, an ordering of excitors is defined. The consequences of this choice are discussed in the following section. By default, a choice consistent with Ref. 38 is used, with excitors applied in decreasing order of highest orbital excited from and increasing order of excitation level. For each excitor present in the wavefunction, the algorithm assesses which of the cases listed above is appropriate. If the excitor cannot be applied (case 3), the next excitor is considered. If the excitor can be applied, the probability of doing so is computed as

$$p_{\text{excit}} = \frac{|\sin(t)|}{|\sin(t)| + |\cos(t)|} \qquad (41)$$

With probability p_{excit}, the excitor is applied and the cluster amplitude is multiplied by $\sin(t)$. With probability $1 - p_{\text{excit}}$, the operator is not applied and the cluster amplitude is multiplied by $\cos(t)$. Once a cluster is selected in this way it undergoes the same spawning, death and annihilation steps as in the traditional CCMC algorithm. As in the case of full UCCMC, this *ansatz* modifies the normalisation of the wavefunction with respect to the reference. In this case,

$$\langle D_0 | \Psi_{\text{tUCC}} \rangle = \prod_{\mathbf{i}} \cos(t_{\mathbf{i}}). \qquad (42)$$

Therefore, in the stochastic representation,

$$t_{\mathbf{i}} = \frac{N_{\mathbf{i}}}{N_0} \qquad (43)$$

and

$$E_{\text{proj}} = \sum_{i \neq 0} \frac{N_{\mathbf{i}}^{\text{CI}} H_{\mathbf{i}0}}{N_0 \prod_{\mathbf{i}} \cos\left(\frac{N_{\mathbf{i}}}{N_0}\right)} \qquad (44)$$

C. Energy estimators

To end this section, we go through the available energy estimators and how they are obtained in each case.

1. The shift S is output directly from the calculation at each report loop. The final value is obtained by a reblocking analysis.[48]

2. The projected energy E_{proj}, given by eq. (29), eq. (34) or eq. (44), is computed as a ratio of the average values of the numerator and denominator. Estimates of these are output at each report loop and averages obtained by a reblocking analysis.

3. The expectation value of the energy $\langle E \rangle$ is not generally computed for QMC methods, or indeed for conventional CC. However, to obtain a fairer comparison with the variational UCC method, we have implemented the following procedure. At the end of each calculation a list of average cluster coefficients is output. These are used to generate the corresponding CI wavefunction and the expectation value of the energy is computed. This value is averaged over 5 independent calculations to obtain the final estimate.

V. NUMERICAL RESULTS

The following section presents results obtained using a development version of HANDE-QMC.[47] Values for the projected energy are obtained from a reblocking analysis,[48] while the expectation value of the energy is obtained as described in Section IV C. In all cases, error-bars correspond to a single standard deviation, σ. In the case of FCIQMC and CCMC, the error bars decrease with imaginary time as $\sigma \propto \sqrt{\beta^{-1}}$ and the same behaviour is observed in UCCMC (see Supplementary Information).

A. Two-electron systems

The natural starting point for the investigation of UCCMC energies and wavefunctions is H_2. As it only has two electrons, UCCSD should be exact in this case. Particularly, in the STO-3G basis set,[49] only two determinants contribute to the wavefunction, which may be written in the UCC form as

$$|\Psi_{\mathrm{UCC}}\rangle = e^{t\hat{a}_{12}^{34}} |D_0\rangle = \cos(t) |D_0\rangle + \sin(t) |D_{01}^{23}\rangle \quad (45)$$

It is therefore trivial to analytically solve either the variational or the projected equations to obtain t and the exact UCC energy, which is in this case equivalent to FCI. Stochastic estimates of the energy and the coefficient t have been obtained using both full UCCMCSD and trotterized UCCMCSD (tUCCMCSD). The differences relative to the exact values are given in Figure 1 and Figure 2. Both show good agreement between the stochastic estimates and the exact values. In particular, the expectation-value estimator for the energy, which has been implemented for a fairer comparison to conventional UCC, shows excellent agreement with the exact energy.
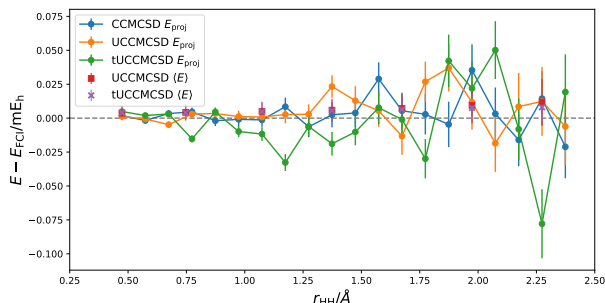

FIG. 1: Error in stochastic coupled cluster and unitary coupled cluster energy estimates relative to the exact UCCSD (FCI) energy for H_2 in a STO-3G basis set.

Increasing the basis to 6-31G[50] leads to a system with 7 symmetry-allowed excited determinants. Figure 3 shows the energies of stochastic UCC approaches relative to the exact value. The agreement is good, but one can observe the increase in the size of the error bars of the projected energy as the bond length increases. This is caused by the
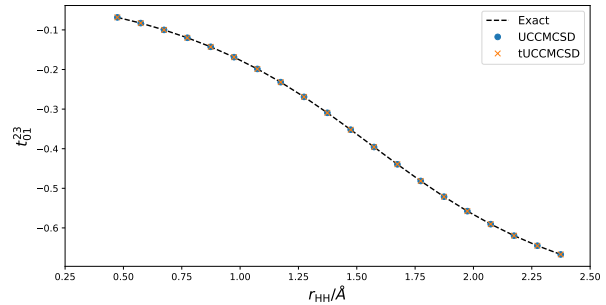

FIG. 2: Error in stochastic UCCSD and tUCCSD estimates of the coefficient t on $|D_{01}^{23}\rangle$ for H_2 in a STO-3G basis set.

increase in static correlation as the molecule approaches dissociation. In simple cases such as this, the size of the error bars can be decreased by simply running longer calculations at higher bond lengths.

The coefficients obtained from UCCMCSD are also in good agreement with the ones obtained by solving the projected UCCMCSD equations deterministically, as can be seen in Figure 4. It is worth noting that in both cases the results were obtained with an expansion truncated at $\hat{\tau}^{12}$, but deterministic tests suggest that convergence with respect to polynomial order is reached at $\hat{\tau}^4$. For tUCCMCSD however, the ordering dependence becomes obvious. In UCC, the coefficients on spin-flipped excitors are equal up to a sign change. However, when using the order of Evangelista et al[38] (see Figure 4 caption), these excitors acquire different amplitudes. An alternative ordering, where all single excitations are applied after all double excitations, recovers the equivalence of spin-flipped pairs, however these coefficients do not necessarily correspond to the UCC values. Clearly, the ordering of excitors in the Trotter expansion is a significant parameter of such a calculation and must be chosen carefully. Further, from Figure 4, the observation can be made that the default ordering used for tUCCMC ensures one coefficient of each pair agrees with the full UCCMC result, while the other is modified by the ordering. This is observed in larger systems as well (see Supplementary Information).
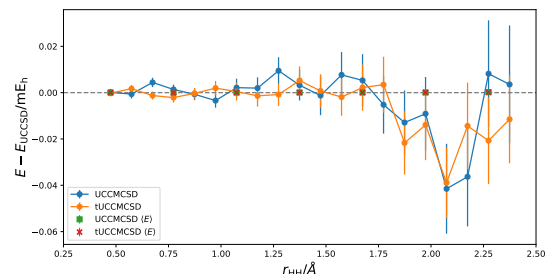

FIG. 3: Error in stochastic UCCSD and tUCCSD energy estimates relative to the exact UCCSD (FCI) energy for H_2 in a 6-31G basis set.
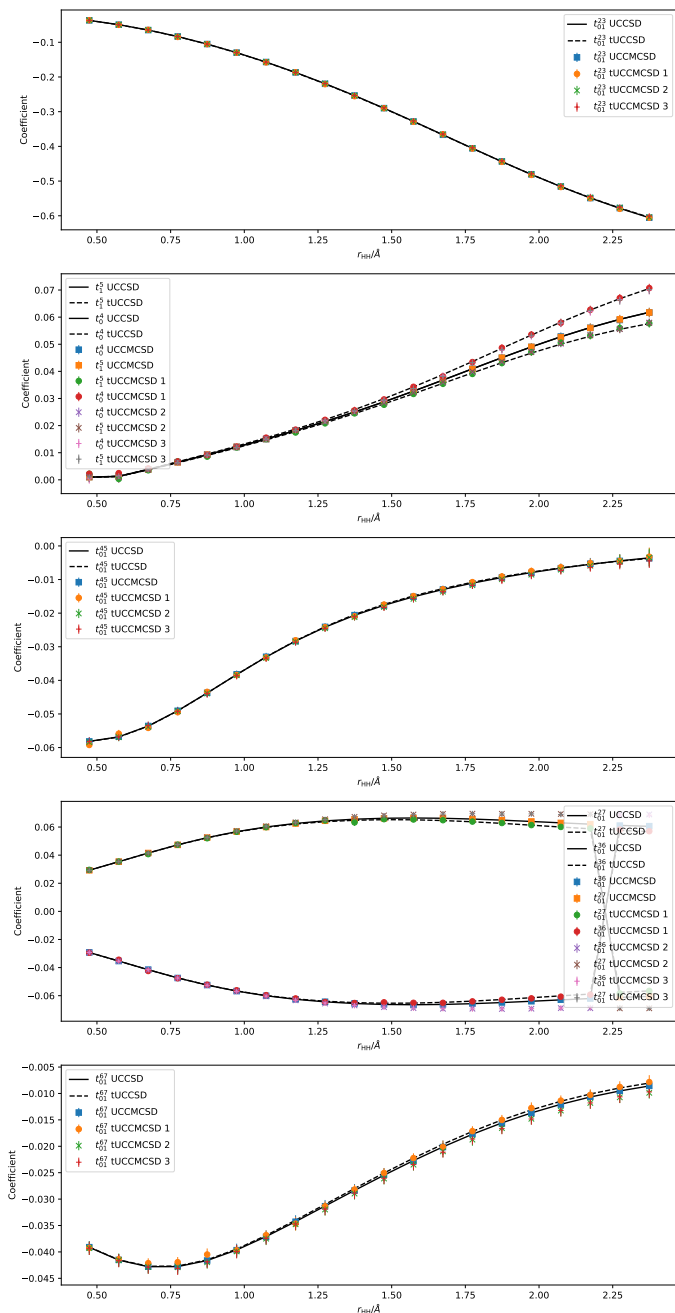
FIG. 4: UCCMCSD and tUCCMCSD coefficients for H_2 in a 6-31G basis set. The first set of tUCCMCSD results use an order consistent with[38], giving
$|\Psi_{\text{tUCCSD}}\rangle = e^{\hat{T}_0^4} e^{\hat{T}_{01}^{23}} e^{\hat{T}_{01}^{43}} e^{\hat{T}_{01}^{25}} e^{\hat{T}_{01}^{45}} e^{\hat{T}_{01}^{67}} e^{\hat{T}_1^5} |D_0\rangle$. The second and third set use the alternative orders with single excitations applied first and last respectively. The orbital labels are chosen to reflect the physical nature of the orbitals, rather than the energy. Consequently, at $0.895 \leq r_{\text{HH}} \leq 1.395$, spin-orbitals 4 and 5 are higher in energy than 6 and 7.

B. Beyond two electrons

Taking a small step away from the simplicity of two-electron systems, we look at LiH in a STO-3G basis set,

which has 4 electrons in 12 spin-orbitals.

For LiH with all electrons correlated, UCCSD is no longer expected to be exact, and indeed the deterministic projected results show an increasing deviation from FCI as the bond length increases (see Figure 5). The energies from stochastic UCCSD agree well with those from its deterministic counterpart, as can be seen from Figure 5. Coefficients show similar behaviour to that observed for two-electron systems, with spin-flipped coefficients acquiring different values in tUCCMCSD (see Supplementary Information).
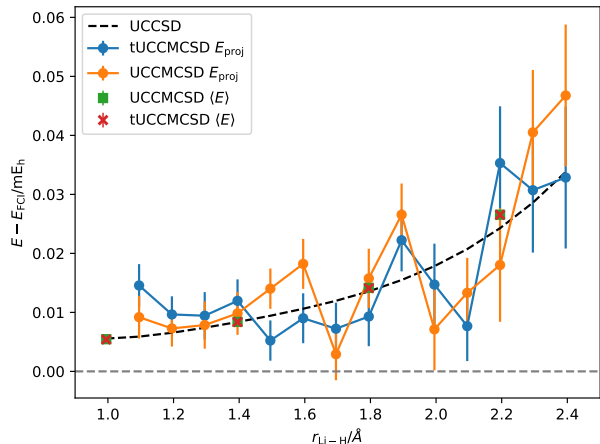

FIG. 5: Error in stochastic UCCSD energy estimates relative to the exact FCI energy for LiH in a STO-3G basis set.

While small systems like H_2 and LiH are good models to test the fundamental behaviour of new algorithms, they are hardly representative of the kind of problems of interest in electronic structure today. To approach these, we will look at the dissociation of the nitrogen molecule, in which a triple bond must be broken. The system is therefore characterised by important strong correlation effects, which cause failures of both conventional and stochastic CC methods at large bond-lengths.[4]

Variational UCCSD has been benchmarked for this system by Cooper and Knowles,[51] showing an improvement over traditional CCSD. By comparison, using the projected energy estimator from projected UCCSD (or tUCCSD) gives results that are very similar to CCSD. However, if one computes the expectation value of the energy instead, this recovers almost all of the correlation energy of the variational approach, as can be seen from Figure 6. Additionally, the method converges beyond the point reported in Ref. 51. At bond lengths beyond those shown in Figure 6, the stochastic method is difficult to converge and the projected energy dips below the FCI results, as it does for traditional CCSD, however the expectation value of the energy remains variationally above the FCI result.

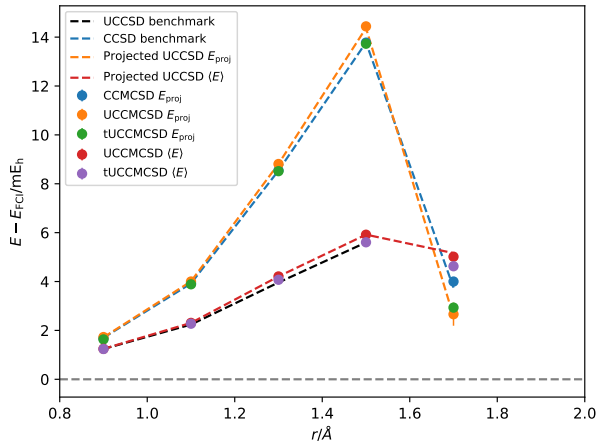The N_2 system is sufficiently large to study the efficiency of UCCMC relative to traditional CCMC. We find

FIG. 6: Error in stochastic CCSD and UCCSD energy estimates relative to the FCI energy for N_2 in a STO-3G basis set. While the E_{proj} estimator for UCCMCSD and tUCCMCSD approaches that for CCSD, the expectation value $\langle E \rangle$ remains close to the variational UCCSD value. CCSD and UCCSD benchmark values are from Ref. 51.

that both the population plateau and the convergence of the projected energy estimator with imaginary time behave very similarly between CCMCSD, UCCMCSD and tUCCMCSD (see Supplementary Information).

We also investigate the effect the polynomial truncation of the UCCSD expansion has on the quality of the obtained energy estimators. Consider truncating the exponential expansion at a polynomial order o:

$$e^{\hat{\tau}} \approx \sum_{i=0}^{o} \frac{1}{i!} \hat{\tau}^i \qquad (46)$$

Table I gives the values of the energy computed using the projected UCCSD method truncated at orders $o = 2 - 12$. One finds that system, $o = 4$ provides a sub miliHartree approximation of the final result and $o = 8$ appears converged to within $10^{-8} E_{\text{h}}$. Therefore we are confident that the truncation at $o = 12$ generally used in our calculations does not introduce any significant error into the results, in either the deterministic or stochastic case.

Finally, we note that the stochastic approach we have implemented can be directly applied to higher orders of coupled cluster, with no additional complexity. For example, Figure 7 shows projected energy results for stochastic methods including up to fourth order excitations. These are significantly more accurate than their CCSD counterparts, but once again we observe that variational energy would provide a better quality estimator than the projected energy. It is worth noting that, due to the linear scaling of cluster selection with number of excitors in tUCCMC, we observe a slowing down of the trotterized method relative to full UCCMC, for identical calculation parameters.

Truncation order o	$E_{\text{proj}}/E_{\text{h}}$	$\langle E \rangle /E_{\text{h}}$
2	-0.21526093	-0.22010447
3	-0.21712594	-0.22135307
4	-0.21649549	-0.22109101
5	-0.21646238	-0.22107111
6	-0.21646951	-0.22107423
7	-0.21646976	-0.22107439
8	-0.21646972	-0.22107437
9	-0.21646972	-0.22107437
10	-0.21646972	-0.22107437
11	-0.21646972	-0.22107437
12	-0.21646972	-0.22107437

TABLE I: Projected and expectation value UCCSD correlation energy for STO-3G N_2 at $r = 1.3$, as a function of truncation order. The results are converged to within $10^{-8} E_{\text{h}}$ by $o = 8$.
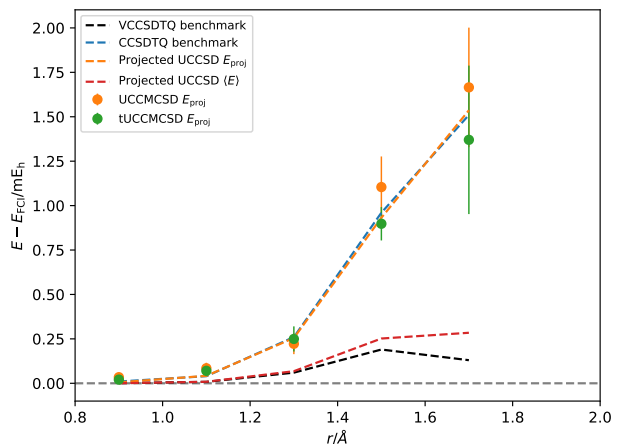

FIG. 7: Error in stochastic CCSDTQ and UCCSDTQ energy estimates relative to the FCI energy for N_2 in a STO-3G basis set. CCSDTQ and VCCSDTQ benchmark values are from Ref. 51.

VI. CONCLUSIONS

In this paper, we have developed a projective approach to the unitary coupled cluster method, based on solving the residual equations for a polynomially truncated unitary exponential ansatz or its Trotter approximation. We have further implemented a stochastic version of this method, within the framework of CCMC.

For two-electron systems, the UCCMCSD method shows good agreement with FCI, as expected. For larger systems, we find that the stochastic and deterministic results computed at the same polynomial truncation level agree, implying that no bias is introduced by our selection schemes. Further, we have shown that for N_2, the results quickly converge with polynomial truncation level, guaranteeing that this truncation does not introduce any meaningful error in the results. Finally, we have observed that, in the context of UCCSD, the expectation value of the energy appears to provide a better

estimator for the correlation energy than the projected energy, approaching the value obtained by variational UCCSD, without requiring the explicit variational optimisation of the wavefunction with respect to the cluster coefficients, which leads to significantly more involved equations. Both computing this quantity and obtaining an unbiased estimator for it is more expensive than the corresponding procedure for the projected energy, so it is satisfying to note that while less accurate, the projected energy we obtain is comparable to CCSD and we therefore expect it to improve in a similar way with increased cluster orders. Furthermore, unlike its conventional counterpart, the stochastic UCCMC approach and its trotterized approximation naturally extends beyond singles and doubles, allowing one to access higher accuracy methods in a unitary fashion. As such, we are optimistic that this approach could become a viable alternative to traditional coupled cluster, when high accuracy is required and a unitary *ansatz* is preferable. UCCMC is expected to scale well with increasing cluster truncation levels, provided the polynomial truncation level is preserved. However, for very high cluster truncations, the error due to the finite polynomial order used may become higher and this would need to be increased appropriately. While tUCCMC suffers from no such errors, the current sampling algorithm used scales linearly with the number of excitors in the expansion and we therefore expect this step to become limiting in large enough Hilbert spaces.

Given the renewed interest in unitary coupled cluster as a functional form in the quantum computing community, we believe that our stochastic method may be of interest as a means to provide a better–than–Hartree–Fock initial guess for the wavefunction or to screen the cluster amplitudes, as CCMC has been used before for conventional algorithms,[29,30] helping to streamline the quantum algorithm, by decreasing the complexity of the circuits required in a system-dependent, physically justified way. These ideas will be explored in a further publication.

VII. SUPPLEMENTARY MATERIAL

See supplementary material for details on the convergence of UCCMC and tUCCMC, as well as cluster coefficients for the LiH and N_2 systems.

VIII. ACKNOWLEDGEMENTS

IX. DATA AVAILABILITY

The data that support the findings of this study are openly available in the Apollo - University of Cambridge Repository at https://doi.org/10.17863/CAM.60037.

[1] J. Čížek, J. Chem. Phys **45**, 4256 (1966).
[2] J. Čížek, Adv. Chem. Phys. **24**, 35 (1969).
[3] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
[4] G. K.-L. Chan, M. Kállay, and J. Gauss, J. Chem. Phys. **121**, 6110 (2004).
[5] P. G. Szalay, M. Nooijen, and R. J. Bartlett, J. Chem. Phys. **103**, 281 (1995).
[6] R. J. Bartlett and J. Noga, Chem. Phys. Lett. **150**, 29 (1988).
[7] T. Van Voorhis and M. Head-Gordon, J. Chem. Phys. **113**, 8873 (2000).
[8] J. Arponen, Ann. Phys. **151**, 311 (1983).
[9] W. Kutzelnigg, J. Chem. Phys. **77**, 3081 (1982).
[10] W. Kutzelnigg and S. Koch, J. Chem. Phys. **79**, 4315 (1983).
[11] W. Kutzelnigg, J. Chem. Phys. **80**, 822 (1984).
[12] K. Tanaka and H. Terashima, Che. Phys. Lett. **106**, 558 (1984).
[13] M. R. Hoffmann and J. Simons, Chem. Phys. Lett. **142**, 451 (1987).
[14] M. R. Hoffmann and J. Simons, J. Chem. Phys. **88**, 993 (1988).
[15] R. J. Bartlett, S. A. Kucharski, and J. Noga, Chem. Phys. Lett. **155**, 133 (1989).
[16] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, Science **292**, 472 (2001), 0104129 [quant-ph].
[17] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Science **309**, 1704 (2005).
[18] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, Mol. Phys. **109**, 735 (2011).
[19] J. B. Anderson, J. Chem. Phys. **63**, 1499 (1975).
[20] D. Ceperley and B. Alder, Science **231**, 555 (1986).
[21] G. H. Booth, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009).
[22] A. Y. Kitaev, "Quantum measurements and the abelian stabilizer problem," (1995), arXiv:quant-ph/9511026.
[23] D. S. Abrams and S. Lloyd, Phys. Rev. Lett. **83**, 5162 (1999).
[24] A. Peruzzo, J. McClean, P. Shadbolt, M. H. Yung, X. Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nature Commun. **5**, 1 (2014).
[25] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New J. Phys **18**, 023023 (2016).
[26] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik, Chem. Rev. **119**, 10856 (2019).
[27] J. Romero, R. Babbush, J. R. Mcclean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Quantum Sci. Technol. **4**, 014008 (2019).
[28] A. J. W. Thom, Phys. Rev. Lett. **105**, 1 (2010).
[29] J. E. Deustua, J. Shen, and P. Piecuch, Phys. Rev. Lett. **119**, 223003 (2017).
[30] J. E. Deustua, I. Magoulas, J. Shen, and P. Piecuch, J. Chem. Phys. **149**, 151101 (2018).
[31] J. E. Deustua, S. H. Yuwono, J. Shen, and P. Piecuch, J. Chem. Phys **150**, 111101 (2019).
[32] J. E. Campbell, Proc. London Math. Soc. **28**, 381 (1897).
[33] H. F. Baker, Proc. London Math. Soc. **s2-3**, 24 (1905).
[34] F. Hausdorff, Ber. Verh. Kgl. Sächs. Ges. Wiss. Leipzig., Math.-phys. Kl. **58**, 19 (1906).
[35] S. Pal, Theoretica Chimica Acta **66**, 207 (1984).
[36] H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959).
[37] M. Suzuki, Commun. Math. Phys. **51**, 183 (1976).
[38] F. A. Evangelista, G. K. L. Chan, and G. E. Scuseria, J. Chem. Phys. **151**, 244112 (2019).
[39] J. S. Spencer and A. J. W. Thom, J. Chem. Phys. **144**, 084108 (2016).

[40] C. J. C. Scott and A. J. W. Thom, J. Chem. Phys. **147**, 124105 (2017).

[41] R. S. T. Franklin, J. S. Spencer, A. Zoccante, and A. J. W. Thom, J. Chem. Phys. **144**, 044111 (2016).

[42] J. S. Spencer, V. A. Neufeld, W. A. Vigor, R. S. T. Franklin, and A. J. W. Thom, J. Chem. Phys. **149**, 204103 (2018).

[43] A. A. Holmes, H. J. Changlani, and C. J. Umrigar, J. Chem. Theory Comput. **12**, 1561 (2016).

[44] V. A. Neufeld and A. J. Thom, J. Chem. Theory Comput. **15**, 127 (2019).

[45] C. J. C. Scott, R. Di Remigio, T. D. Crawford, and A. J. W. Thom, J. Phys. Chem. Lett. **10**, 925 (2019).

[46] I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics* (Cambridge University Press, 2009).

[47] J. S. Spencer, N. S. Blunt, S. Choi, J. Etrych, M.-A. Filip, W. M. C. Foulkes, R. S. T. Franklin, W. J. Handley, F. D. Malone, V. A. Neufeld, *et al.*, J. Chem. Theory Comput. **15**, 1728 (2019).

[48] H. Flyvbjerg and H. G. Petersen, J. Chem. Phys. **91**, 461 (1989).

[49] W. J. Hehre, R. F. Stewart, and J. A. Pople, J. Chem. Phys. **51**, 2657 (1969).

[50] R. Ditchfield, W. J. Hehre, and J. A. Pople, J. Chem. Phys. **54**, 724 (1971).

[51] B. Cooper and P. J. Knowles, J. Chem. Phys. **133** (2010).