Coupled Cluster Externally Corrected by Adaptive Configuration Interaction

Gustavo J.R. Aroeira, Madeline M. Davis, Justin M. Turney, and Henry F. Schaefer III*

Center for Computational Quantum Chemistry, University of Georgia, Athens

E-mail: ccq@uga.edu

Phone: +1 706 542 0406. Fax: +1 706 542 2067

Abstract

An externally corrected coupled cluster (CC) method, where an adaptive configuration interaction (ACI) wave function provides the external cluster amplitudes, named ACI-CC, is presented. By exploiting the connection between configuration interaction and coupled cluster through cluster analysis, the higher-order T_3 and T_4 terms obtained from ACI are used to augment the T_1 and T_2 amplitude equations from traditional coupled cluster. These higher-order contributions are kept frozen during the coupled cluster iterations and do not contribute to an increased cost with respect to CCSD. We have benchmarked this method on three closed-shell systems: beryllium dimer, carbonyl oxide, and cyclobutadiene, with good results compared to other corrected coupled cluster methods. In all cases, the inclusion of these external corrections improved upon the "gold standard" CCSD(T) results, indicating that ACI-CCSD(T) can be used to assess strong correlation effects in a system and as an inexpensive starting point for more complex external corrections.

1

Introduction

Since its introduction in quantum chemistry, coupled-cluster theory (CC)[1-4] has become a method routinely employed to compute reliable correlation energies. The single reference (SR) coupled-cluster including single and double excitations (CCSD)[5-7] method can provide satisfactory results for various molecular properties. These results are often superior to those from alternative methods, such as configuration interaction with single and double excitations (CISD), or second-order Møller–Plesset perturbation theory (MP2).[8] The contribution of high-order excitations can be included in an economical way using energy correcting methods such as the CCSD with perturbative triples [CCSD(T)] approach.[9,10] This method has become a popular option to compute accurate energies, in particular for nondegenerate closed-shell systems near their equilibrium geometries. Thus, the CCSD(T) method is often regarded as the "gold standard" of quantum chemistry.[11] However, many properties require energies at different points on the potential energy surface (PES). Despite the effectiveness of the CCSD(T) at equilibrium geometries, the method may break down, for instance, when bonds are significantly stretched.[12] For so-called multireference (MR) or strongly correlated systems, such as the bond-breaking scenario, the reference determinant loses its dominance. As a consequence, excitations beyond double become increasingly important and the inclusion of high-order contributions via perturbative methods becomes less efficient.

In order to improve the description of strongly correlated systems while maintaining the SR framework one can explicitly account for high-order excitations. Methods such as CCSDT,[13-15] CCSDTQ,[16-18] CCSDTQP,[19] and higher can successfully solve many multireference problems, such as quasidegeneracy. Unfortunately, these approaches are computationally very demanding and, generally, can only be used with relatively small basis sets. Alternative approaches, such as CCSDt and CCSDtq,[20] aim to reduce this cost by defining active spaces for the higher-order cluster operators. For example, the CCSDt method uses the same equations as CCSDT, but include only a small subset of the full $\hat{T}_3$ amplitudes (thus

"t" instead of "T"). However, even with this augmentation, the description of systems like N_2 with triple bonds can still be less than desirable. To properly account for these quaside-generacies, a MR CC formulation is needed. Efforts to formulate such methods have given rise to a family of MR CC approaches each with different achievements and shortcomings.[8,21]

Externally corrected coupled-cluster (ecCC) has been proposed as a strategy to recover higher-order excitation contributions to the CC energy. The premise is to use a separate method from which three- and four-body cluster amplitudes can be extracted and introduced into the CC equations. An early application of this strategy used unrestricted Hartree–Fock wave functions to estimate quadruple-excitation contributions to the coupled-pair many-electron theory.[22] Later, Paldus and Planelles employed valence bond wave functions as the source of the higher-order amplitudes.[23-25] The usage of complete active space self-consistent field (CASSCF) as the source for the external amplitudes was suggested by Stolarczyk[26] in his proposal of a complete active space coupled-cluster (CASCC) theory. However, this strategy was only implemented and explored by Peris and coworkers, two years later.[27] This idea was further pursued by Li and Paldus[28] in their reduced multireference (RMR) CC method, which can be described as a step-by-step procedure: (a) a multireference configuration interaction (MRCI) wave function is constructed within an active space; (b) using cluster analysis, T_3 and T_4 amplitudes are extracted; (c) a CCSD computation is performed, including the three- and four-body contributions generated previously, in a CCSDTQ-like framework. This approach provided results close to the full configuration interaction (FCI) limit for a few test cases they reported.[28] They also used RMR CCSD to study the potential energy surfaces of HF, F_2, and H_2O.[29] In all cases the RMR CC outperformed MRCI and CCSD. Peris et $al.$[30] also proposed the use of perturbative selected configuration interaction (CIPSI)[31] to reduced the dimension of the CI wave function. Their results for model systems were within 1 mE_h of the RMR CCSD methods while considerably reducing the cost of the computation. Xu and Li[32] demonstrated that this approach can be extended to even higher-order external amplitudes. They used CASSCF to generate $\hat{T}_4$ and $\hat{T}_5$ terms that were introduced into a CCSDt routine

yielding results in good agreement with experiment for a selected set of strongly correlated systems.

The split amplitude strategy of the ecCC methods was implemented in a different way by Kinoshita and coworkers.[33] In their tailored coupled-cluster method (TCC), a complete active space configuration interaction (CASCI) computation generates the external correction. However, unlike the RMR CC method, only T_1 and T_2 amplitudes are extracted from the CASCI computation. These amplitudes are then constrained to be unchanged during the CC iterations. The method relies on the assumption that the CAS amplitudes (T_1 and T_2) will retain their multireference information and will continue to describe the strong correlation aspects of the system. The remaining T_1 and T_2 amplitudes (outside the CAS space) are then used to obtain the dynamical component of the correlation energy. Despite its simplicity, potential energy surfaces and dissociation energies obtained with this method are in good agreement with the more demanding MRCI approach.[33] However, due to the decoupling of the two set of amplitudes, TCC presents a large nonparallelity error (NPE).[34]

More recently, Veis and coworkers[35] combined the TCC split amplitude strategy with density matrix renormalization group (DMRG). This version of TCC uses DMRG as the source of T_1 and T_2 amplitudes, this approach has the advantage that DMRG can handle large active spaces with a more favorable scaling than CASCI or CASSCF. Further work shows that DMRG-TCCSD performs well for difficult multireference systems and also reports a version of the method employing local pair natural orbital.[36,37] The authors suggest that the inclusion of high-order external amplitudes, as done in the RMR CC method, could lead to even more accurate results.[35]

In this research, we reexamine the simplest type of size-extensive externally corrected coupled cluster, referred here as CASCI-CC. In this approach, the external correction comes from a small FCI computation within an active space. A shortcoming of this strategy is the prohibitive cost for large active spaces. Therefore, we investigate the substitution of CASCI for the adaptive configuration interaction (ACI) method developed by Schriber and

4

Evangelista.[38,39] ACI is a modern selected CI algorithm that is suitable for extended active spaces. In his review of externally and internally corrected coupled cluster methods, Paldus[40] highlights desirable properties for the external source of high-order clusters. These properties are: (i) universal availability: the ACI method is a general approach that relies on known CI technology, thus it can be extended for different systems; (ii) size consistent: as an approximation to FCI, the ACI method can numerically converge to a size consistent answer; (iii) account for non dynamic correlation: ACI selects the most important determinants to build the wave function, including higher-than-pair excitations, thus accounting for nondynamical correlation; (iv) be systematically improvable: the strongest suit of the ACI method is its dependency on a pre selected σ parameter that approximately controls the precision of the method (see Equation 1).

$$|E_{\text{FCI}} - E_{\text{ACI}}| \approx \sigma \tag{1}$$

In this light, one can tune the ACI method arbitrarily towards the FCI limit. The ACI externally corrected CCSD is henceforth denoted as ACI(σ)-CCSD and, when perturbative triples are included, ACI(σ)-CCSD(T), where σ is on the energy scale of mE_{h}. For example, ACI(10) denotes a σ of 10 mE_{h}.

Theory

Adaptive Configuration Interaction

The following is a brief description of the ACI algorithm; a more thorough discussion of the method can be found in the previous literature.[38,39] We start with set of determinants P, which has an associated wave function $|\Psi_P\rangle = \sum c_p |\Phi_p\rangle$ for $p \in P$, where c_p are the CI expansion coefficients. A new set, F, is formed by taking single and double excitations of

P. For each $|\Phi_f\rangle$, where $f \in F$, we associate an energy contribution ϵ determined by

$$\epsilon_f = \frac{\Delta}{2} - \sqrt{\frac{\Delta^2}{4} + V^2} \tag{2}$$

$$\Delta = E_f - E_P = \langle \Phi_f | H | \Phi_f \rangle - \langle \Psi_P | H | \Psi_P \rangle \tag{3}$$

$$V = \langle \Phi_f | H | \Psi_P \rangle = \sum_k c_p \langle \Phi_f | H | \Phi_k \rangle \quad k \in P \tag{4}$$

The determinants with smallest $|\epsilon|$ are removed from F until the sum of their contributions reaches a predefined parameter σ. After this removal, the remaining determinants form the space Q. The updated model space (M) is created as $M = P \cup Q$. For this new space, the Hamiltonian matrix is constructed and diagonalized. These steps are repeated until convergence is reached. However, before each new iteration, the set M undergoes a coarse graining procedure, where the determinants with the lowest contribution to the wave function (c^2) are removed until

$$\sum_m c_m |\Phi_m\rangle < 1 - \gamma\sigma \tag{5}$$

where γ is set to 1 E_h as recommended by the authors.[39] Figure 1 schematically summarizes the ACI procedure.

Externally Corrected Coupled-Cluster

In CC theory, the energy of the system is determined utilizing only the T_1 and T_2 amplitudes. However, despite higher-order amplitudes not being present in the energy expression, they are coupled with T_1 and T_2 though the full CC equations. In looking at the amplitude equations from the CCSDTQ level of theory, it is readily seen that T_3 and T_4 amplitudes
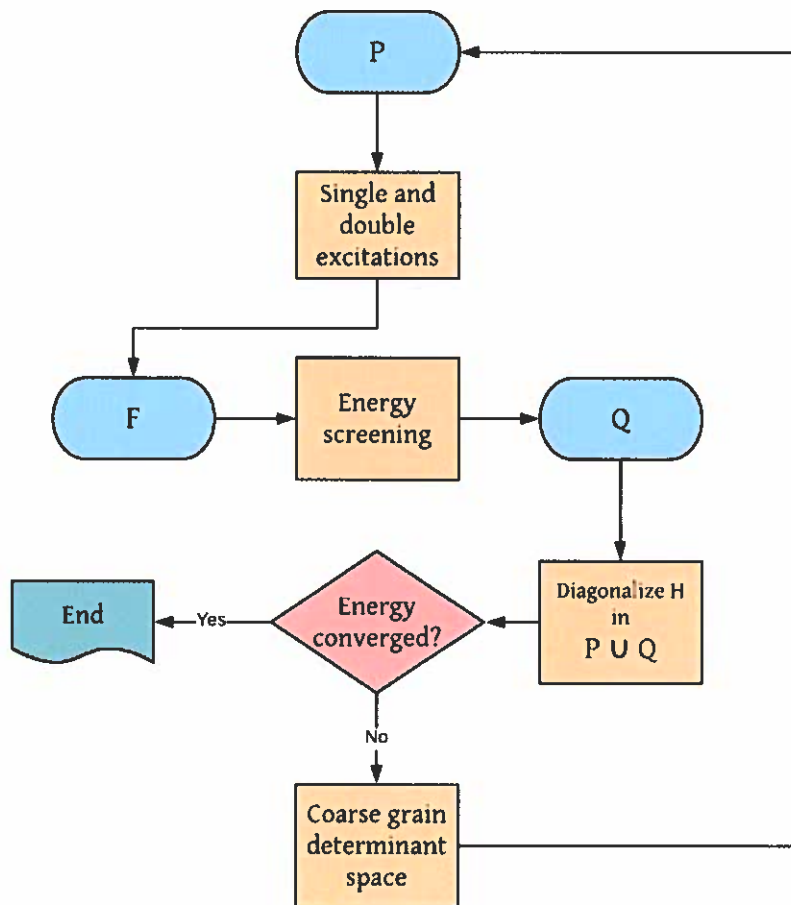
Figure 1: Summary of the ACI procedure. Determinants in the set F are screened using Equation 2. Coarse-graining is applied until Equation 5 is satisfied.

interact directly with T_1 and T_2 as follows

$$t_i^a D_i^a \leftarrow \langle \Phi_i^a | H_N \hat{T}_3 | \Phi_0 \rangle_c \tag{6}$$

$$t_{ij}^{ab} D_{ij}^{ab} \leftarrow \langle \Phi_{ij}^{ab} | H_N (\hat{T}_3 + \hat{T}_4 + \hat{T}_1 \hat{T}_3) | \Phi_0 \rangle_c \tag{7}$$

The traditional CCSD approach can be viewed as an approximation to the full CC where $T_3 = T_4 = 0$. As discussed in the literature,[40] any reasonable approximation for those amplitudes would yield better results than simply setting them to zero. Therefore, if we can estimate the contributions of $\hat{T}_3$ and $\hat{T}_4$ onto $\hat{T}_1$ and $\hat{T}_2$, we can recover some nondynamical correlation energy. In this work, the ACI method is utilized as the external method to approximate the the T_3 and T_4 amplitudes

$$\hat{T}_3 \approx \hat{T}_3^{\mathrm{ACI}} \tag{8}$$

$$\hat{T}_4 \approx \hat{T}_4^{\mathrm{ACI}} \tag{9}$$

where $\hat{T}^{\mathrm{ACI}}$ is obtained from a cluster analysis of the ACI wave function. For example, the equation relating c_{ij}^{ab} to cluster amplitudes is obtained as

$$c_{ij}^{ab} = \langle \Phi_{ij}^{ab} | e^{\hat{T}} | \Phi_0 \rangle \tag{10}$$

$$= \langle \Phi_{ij}^{ab} | \hat{T}_2 + \frac{1}{2} T_1^2 | \Phi_0 \rangle \tag{11}$$

$$= t_{ij}^{ab} + t_i^a t_j^b - t_j^a t_i^b \tag{12}$$

Recursive equations for determining t_i^a, t_{ij}^{ab}, t_{ijk}^{abc}, and t_{ijkl}^{abcd} can be found in the literature.[41] With amplitudes available, the external corrections for T_1 and T_2 are computed. Equations

6 and 7 give rise to the following algebraic expressions

$$t_i^a D_i^a \leftarrow \frac{1}{4} t_{imn}^{aef} v_{ef}^{mn} \tag{13}$$

$$t_{ij}^{ab} D_{ij}^{ab} \leftarrow f_e^m t_{ijm}^{abe} + t_m^e t_{ijn}^{abf} v_{ef}^{mn} + \frac{1}{2} \hat{P}(ab)[t_{ijm}^{aef} v_{ef}^{bm}] - \frac{1}{2} \hat{P}(ij)[t_{imn}^{abe} v_{je}^{mn}]$$
$$+ \frac{1}{2} \hat{P}(ij)[t_i^e t_{jmn}^{abf} v_{ef}^{mn}] + \frac{1}{2} \hat{P}(ab)[t_m^a t_{ijn}^{bef} v_{ef}^{mn}] + \frac{1}{4} t_{ijmn}^{abef} v_{ef}^{mn} \tag{14}$$

where the summation over repeated indices is implied and $\hat{P}(pq)$ is the antisymmetric permutation operator. Naturally, these expressions are the same as the ones found in the CCSDT and CCSDTQ equations. The difference is that T_3 and T_4 are known *a priori*. These terms are obtained once from the ACI wave function and their contributions (Equations 13 and 14) are fixed throughout the computation. In principle, the $T_1 T_3$ terms can be updated for each iteration, as the t_i^a amplitudes change. However, as typically done in ecCC, we compute this term one time using the T_1^{ACI} and neglect its changes. The difference between the two strategies seems to be insignificant.[28]

Perturbative Triples Correction (T)

The inclusion of perturbative energy corrections has been done in ecCC by treating connected triples outside the external correction in the same way it is done in the CCSD(T) method.[42] In our implementation, the perturbative correction for triple amplitudes is included using a standard *ijk* algorithm.[43] However, the energy correction associated with each amplitude t_{ijk}^{abc} is included only if the corresponding determinant is not part of the ACI wave function.

$$E_{(T)} \leftarrow t_{ijk}^{abc} \ \text{if} \ \left| \Phi_{ijk}^{abc} \right\rangle \notin \text{ACI} \tag{15}$$

Computational Details

The ACI, CC and ecCC methods were implemented using our under development, open-source, electronic structure code FERMI,[44] written in the programming language JULIA.[45] CCSDT and CCSDTQ computations were performed using the MRCC package.[46–48] Potential energy curves were interpolated and vibrational energy levels were obtained using the matrix Numerov method.[49] Dunning's correlation consistent cc-pVXZ basis set were employed for all computations.[50]

Results and Discussion

Beryllium Dimer

Be_2 is a challenging system for both theory and experiment. For a complete and recent discussion on this topic, we refer the reader to Lesiuk *et al.*[51] and Merritt *et al.*[52] Past work has been concluded that the elucidation of the Be_2 bond must include a description of strong correlation.[53] We examine the potential energy curve of the beryllium dimer using CC and ecCC methods. Computed low lying vibrational levels ($\nu = 0$–4) using the cc-pVQZ basis set and the active space of eight electrons in ten orbitals ($8e,10o$) are shown in Table 1, and corresponding selected potential energy curves are depicted in Figure 2. The CCSD(T) method successfully describes the fundamental transition, being only 1.5 cm^{-1} away from the experimental value. However, the quality of the results decays rapidly for the higher vibrational energy levels; already for ν_3 the deviation is greater than 26 cm^{-1}. The ACI(0.1)-CCSD(T) result for the fundamental is 4 cm^{-1} away from the experimental value; however it improves upon CCSD(T) for higher energy levels; for ν_2 and ν_3 errors with respect to experiment are only 1 cm^{-1} and 5 cm^{-1}. The overall effect of post-CCSD(T) terms is to raise vibrational energy levels. In fact, the CCSDT method seems to perform poorly when compared to CCSD(T), as all energies predicted at this level are above the experimental

results. Moreover, it can be seen in Figure 2 that utilizing an adaptive method as the source of external correction has not introduced noticeable discontinuities in the potential energy curve.
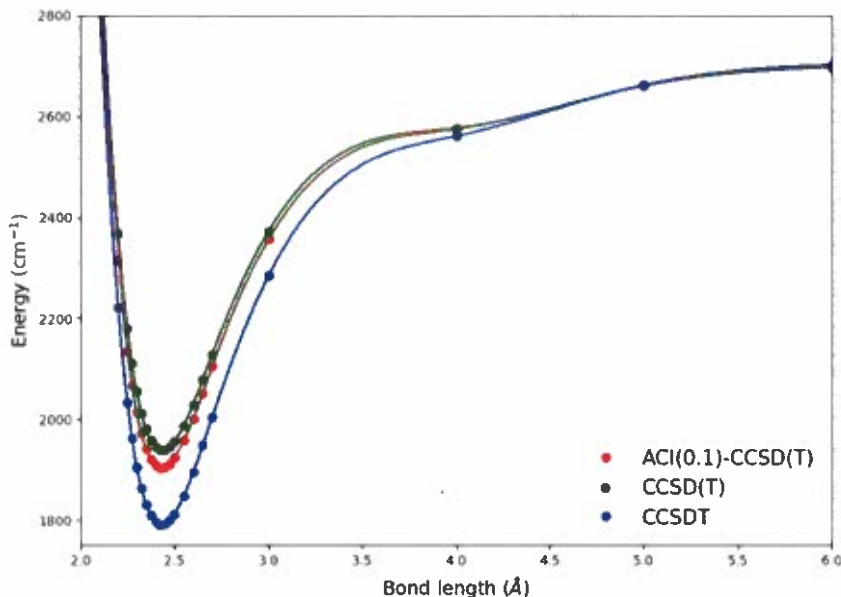

Figure 2: Potential energy curve for the beryllium dimer. Energies are shown with respect to the dissociation limit at the corresponding level of theory. All electrons are correlated. ACI computed with a $(8e,10o)$ active space.

Table 2 presents results for the Be_2 well depth using the cc-pV5Z basis set. These values were computed as the difference between the energy of the dimer at the minimum (determined using the cc-pVQZ basis set, see Table 1) using the $(8e,10o)$ active space and two isolated beryllium atoms using $(4e,5o)$ as the active space. It can be seen that the external correction brings the CCSD(T) result down by 25 cm^{-1} outperforming the CCSDT method, which seems to overestimate the well depth. When compared to experiment CCSD(T) and ACI-CCSD(T) have similar errors, around 12 cm^{-1}. However, the expanded Morse oscillator model of the potential employed by Merritt et al.[52] to determine D_e has been criticised by Lesiuk et al.[51] whose exhaustive theoretical work predicted a slightly deeper potential well. Similarly, many other theoretical studies found well depths greater than 935

cm^{-1}.[54-56] Thus, we believe the shift in the energy due to the external correction can be seen as an improvement. The ability of ACI to produce CASCI quality corrections is also illustrated in this example. While the $\sigma = 1$ mE_h case does not seem to contain all significant information, a tighter $\sigma = 0.1$ mE_h reproduces the full CASCI correction very well. We highlight that using the CASCI method, the active space size contains 44100 determinants (no symmetry considered). Whereas the ACI(0.1) optimized model space comprises less than 90 determinants when $R = 2.425$ Å. More interesting, for the beryllium atoms the ACI wave function does not contain any determinant which excitation greater than 2, even when we tried a very tight ACI(0.01). This means that no external correction is added to the beryllium atoms, which seems problematic suggesting the atoms are not being treated on the same footing as the molecule. However, a comparison with CASCI results, which explicitly include all determinants within the active space, confirms that, when computing D_e, the external correction for the Be atom is negligible.

Table 1: Beryllium dimer bond length (R_e, in angstrom) and low lying vibrational frequencies ($\Delta\nu = \nu_i - \nu_0$ in cm^{-1} for i = 1–4) for different methods. All electrons are correlated and the active space used included eight electrons and the ten lowest orbitals (8e,10o). The cc-pVQZ basis set was employed in all cases. Values in parenthesis are errors with respect to experimental values.

Method	R_e/Å	ν_1	ν_2	ν_3	ν_4
CCSD(T)	2.436	221.1 (*1.5*)	386.3 (*10.8*)	492.4 (*25.7*)	535.3 (*59.5*)
ACI(1)-CCSD(T)	2.432	224.5 (*1.9*)	393.5 (*3.6*)	505.5 (*12.6*)	553.4 (*41.4*)
ACI(0.1)-CCSD(T)	2.430	226.6 (*4*)	398.1 (*1*)	512.9 (*5.2*)	563.2 (*31.6*)
CASCI-CCSD(T)	2.428	228.0 (*5.4*)	400.22 (*3.1*)	515.3 (*2.8*)	566.2 (*28.6*)
CCSDT	2.432	235.0 (*12.4*)	412.0 (*14.9*)	555.8 (*37.7*)	636.1 (*41.3*)
ecCCSDt-CASSCF[a]	2.423	242.5 (*19.9*)			
Exp.[b]	2.454	222.6	397.1	518.1	594.8

[a] Xu and Li;[32] [b] Merritt *et al.*[52]

Table 2: Well depth (D_e) for the beryllium dimer. All electrons are correlated and the active space included four electrons and the ten lowest orbitals ($8e,10o$). The cc-pV5Z basis set was employed in all cases. Geometries were determined at the cc-pVQZ level of theory.

Method	D_e/cm^{-1}
CCSD(T)	919
CCSDT	1082
ACI(1)-CCSD(T)	933
ACI(0.1)-CCSD(T)	943
CASCI-CCSD(T)	943
Reference Value (Theory)[a]	934.6 ± 2.5
Reference Value (Exp.)[b]	929.7 ± 2.5

[a] Lesiuk et al.;[51] [b] Merrritt et al.[52]

Carbonyl Oxide Heat of Formation

Carbonyl oxide is the simplest Criegee intermediate, it has received great attention due to the role it might play in atmospheric chemistry.[57] Since this molecule possesses mild multireference character, many composite approaches applied to these systems included additive corrections such as $E_{CCSDT(Q)} - E_{CCSD(T)}$ with a relatively small basis set.[58,59]

To study the role of post-CCSD(T) contributions on the carbonyl oxide energy profile, we examine the following elementary reaction

$$H_2(g) + CO_2(g) \longrightarrow H_2COO(g) \tag{R1}$$
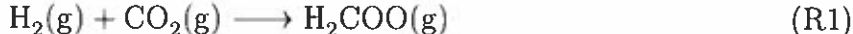
Reliable data for the enthalpy of formation of CO_2 is available;[60] hence this reaction gives information that leads to the heat of formation of the carbonyl oxide. Figure 3 depicts errors in the electronic energies for reaction R1 employing the cc-pVDZ basis set with respect to the CCSDTQ result. At the CCSD level the error with respect to the CCSDTQ result is 3.7 kcal mol^{-1}; this number is greatly improved with the inclusion of perturbative triple corrections; CCSD(T) is 0.84 kcal mol^{-1} away from the CCSDTQ result. If chemical accuracy is desired and assuming CCSDTQ is a good approximation for FCI, then CCSD(T) seems to suffice.

The CCSDT(Q) is the only method to present a negative deviation, which indicates that the perturbative quadruple correction may be favoring disproportionately the Criegee intermediate. This raises problems when CCSDT(Q) is used to assess post-CCSD(T) contributions. For this system, the full T plus (Q) correction to the CCSD(T) energy $[E_{\text{CCSDT(Q)}} - E_{\text{CCSD(T)}}]$ is -1.45 kcal mol^{-1} which, contrary to the result discussed above, suggests that CCSD(T) is not within 1.0 kcal mol^{-1} of accuracy.

The ACI computation was performed with a full valence active space $(18e,14o)$ and $\sigma = 10$ and $1\ mE_{\text{h}}$. The external correction pushed the CCSD results further away from the reference data. However, the inclusion of perturbative triples improves the results dramatically; the ACI(1)-CCSD(T) result is only 0.24 kcal mol^{-1} above the CCSDTQ value. The fact that the ACI-CCSD results are somewhat worse than the traditional CCSD suggests some imbalance in the recovery of correlation energy. The correlation energy increases by 0.7% for CH_2OO and 1.1% for CO_2 when external correction is included into CCSD. The subsequent inclusion of perturbative triples further increases the correlation energies by 4.4% for CH_2OO and 3.4% for CO_2. Therefore, the external T_3^{ACI} correction appears to be improving the CO_2 absolute energy more rapidly than that of the carbonyl oxide. The inclusions of all triples (some through external corrections and the remaining through the (T) strategy) gives the two molecules a more balanced description. It is worth reinforcing that the perturbative triples energy corrections in the ecCC framework is not the same as in the tradition coupled cluster. Since ecCC is an amplitude corrected method, the final t_i^a and t_{ij}^{ab} are different, and therefore the (T) energy will also be different. Moreover, perturbative corrections associated with t_{ijk}^{abc} inside the ACI are not included to avoid double counting. It has been long known that CCSD needs to be augmented with corrections for triples in order to achieve high accuracy. This system provides an example that this could also be the case for externally corrected methods. Finally, ACI-CCSD(T) can be used successfully to assess higher-order contributions to the CCSD(T) energy. For reaction R1, $E_{\text{ACI(1)-CCSD(T)}} - E_{\text{CCSD(T)}}$ is -0.60 kcal mol^{-1}, which, in this case, is a more realistic assessment than the one from CCSDT(Q)
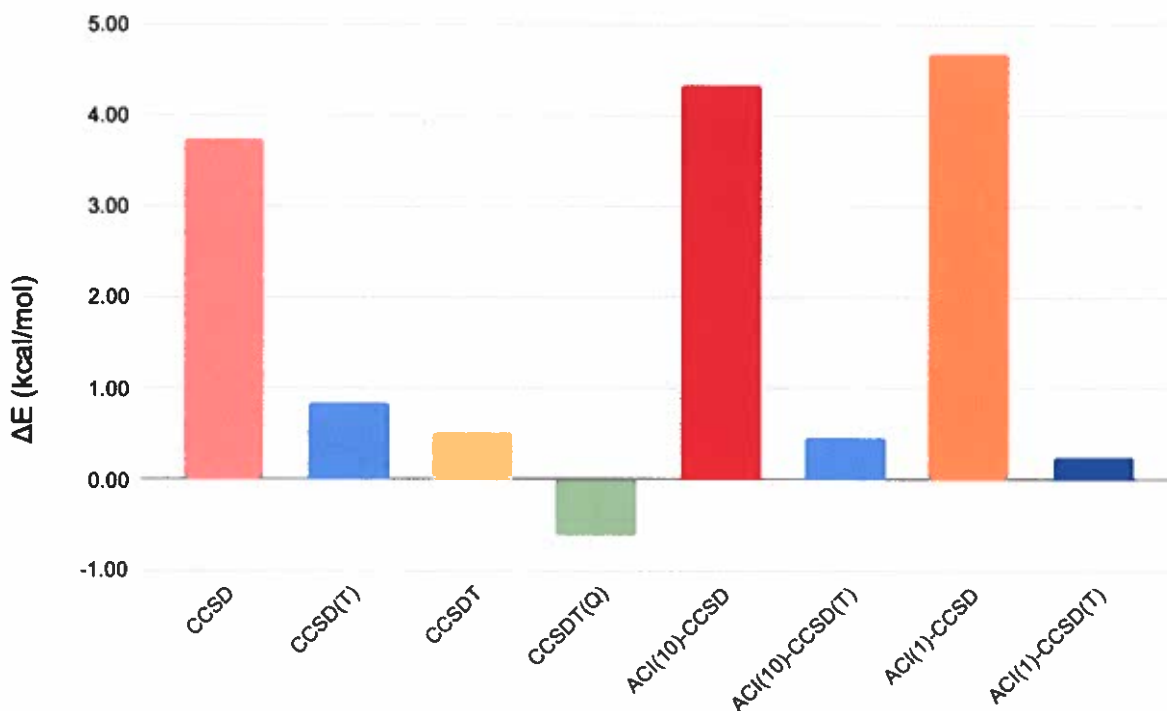
14

at lower cost.

Figure 3: Errors in the $H_2 + CO_2 \longrightarrow H_2COO$ reaction energy compared to a CCSDTQ computation in kcal mol^{-1} using the cc-pVDZ basis set. ACI utilized a full valence active space. Core electrons were frozen for all methods. Geometries optimized at the CCSD(T)/ANO2 level of theory.

Cyclobutadiene Automerization

The automerization of cyclobutadiene, depicted in Figure 4, is a classic example of a multireference transition state. The reaction starts at the rectangular D_{2h} geometry reaching a square D_{4h} transition state and finishes at a new rectangular geometry, degenerate to the first point on the PES. The experimental value is estimated to lie within the wide range of 1.6–10 kcal mol^{-1}.[61]

Table 3 presents the results for the automerization barrier height for the a few selected methods including ACI(10)-CCSD(T). These values do not include zero-point vibrational energy corrections. For the ACI computations, we employed a full valence active space

with 20 electrons in 20 orbitals ($20e,20o$). While it is clear that the ACI(10) captured some important contributions, improving CCSD(T) results by more than 3 kcal mol^{-1}, the external correction is only half as much as the one produced by RMR CCSD(T).[62] The ACI wave function contains two dominant configurations, one of which is the Hartree–Fock determinant. To properly account for the correlation energy, one must include all single and double excitations from these determinants as it is done in MRCI. Thus, it is not surprising that the RMR CCSD(T), which uses MRCI as a source of external corrections, describes this system better. We believe an augmentation of the ACI wave function, allowing it to include important determinants outside the initial active space, would produce competitive results. This is similar to the proposed PSCI CCSD proposed by Peris *et al.*[30] It is important to note that the ACI-CCSD(T) method is a simpler and less demanding approach than the RMR CCSD(T). Nevertheless, ACI-CCSD(T) can provide a natural starting point for more complex external corrections.
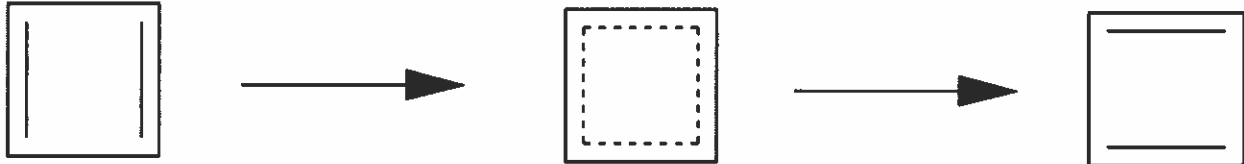

Figure 4: Automerization reaction of the cyclobutadiene

Table 3: Selected MR CC results for the pure electronic barrier height of the cyclobutadiene automerization using the cc-pVDZ basis set. ACI utilized a full valence active space ($20e,20o$)

Method	Barrier height/kcal mol^{-1}	Reference
CCSD(T)	16.2	This work
ACI(10)-CCSD(T)	12.7	This work
RMR-CCSD(T)	7.2	62
RMR-CCSD(T)^a	5.9	62
MR-AQCCSD	7.3	63

^{a} MCSCF orbitals.

Conclusions

We introduce in this work the ACI(σ)-CCSD(T) method. It may be seen as an approximation to the simplest size-consistent externally corrected method, using a small FCI within an active space as the source of external corrections. However, employing the ACI allows us to compute corrections with extended active spaces. We anticipate that a production level code could be used on actives spaces much larger than the ($20e$,$20o$) utilized here. In the three systems examined the ACI(σ)-CCSD(T) performed satisfactory, particularly well for Be_2 and CH_2OO, where the results were comparable or superior to the CCSDT method.

Despite the fact that the ACI(σ)-CCSD(T) method improved upon CCSD(T) results in all cases studied, the computed cyclobutadiene automerization energy fell short when compared to more involved methods, such as RMR CCSD (which is itself an ecCC method). This highlights the limitation of using a restricted active space to compute external corrections. Future work should explore a way to screen determinants without the limitations of an active space, keeping the computational cost manageable. Furthermore, using ACI as the FCI solver, orbital optimization can also considered. In future work, we hope to extend the ACI-CCSD(T) method to open-shell systems.

Acknowledgement

Supporting Information Available

Raw electronic energies and Cartesian coordinates of all molecules.

References

(1) Čížek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* **1966**, *45*, 4256–4266.

(2) Čížek, J. On the Use of the Cluster Expansion and the Technique of Diagrams in Calculations of Correlation Effects in Atoms and Molecules. *Adv. Chem. Phys.* **1969**, *14*, 35–89.

(3) Čížek, J.; Paldus, J. Correlation problems in atomic and molecular systems III. Rederivation of the coupled-pair many-electron theory using the traditional quantum chemical methodst. *Int. J. Quantum Chem.* **1971**, *5*, 359–379.

(4) Paldus, J.; Čížek, J.; Shavitt, I. Correlation Problems in Atomic and Molecular Systems. IV. Extended Coupled-Pair Many-Electron Theory and Its Application to the BH3Molecule. *Phys. Rev. A* **1972**, *5*, 50–67.

(5) Purvis, G. D.; Bartlett, R. J. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

(6) Scuseria, G. E.; Scheiner, A. C.; Lee, T. J.; Rice, J. E.; Schaefer, H. F. The closed-shell coupled cluster single and double excitation (CCSD) model for the description of electron correlation. A comparison with configuration interaction (CISD) results. *J. Chem. Phys.* **1987**, *86*, 2881–2890.

(7) Scuseria, G. E.; Janssen, C. L.; Schaefer, H. F. An efficient reformulation of the closed-shell coupled cluster single and double excitation (CCSD) equations. *J. Chem. Phys.* **1988**, *89*, 7382–7387.

(8) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291–352.

(9) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order per-turbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(10) Bartlett, R. J.; Watts, J.; Kucharski, S.; Noga, J. Non-iterative fifth-order triple and quadruple excitation energy corrections in correlated methods. *Chem. Phys. Lett.* **1990**, *165*, 513–522.

(11) Gauss, J. In *The Encyclopedia of Computational Chemistry*; v. R. Schleyer; N. L. Allinger; T. Clark; J. Gasteiger; P. A. Kollman; H. F. Schaefer III;, P., Scheiner, P. R., Eds.; Wiley: Chichester, 1998; pp 615–636.

(12) Ghose, K. B.; Piecuch, P.; Adamowicz, L. Improved computational strategy for the state-selective coupled-cluster theory with semi-internal triexcited clusters: Potential energy surface of the HF molecule. *J. Chem. Phys.* **1995**, *103*, 9331–9346.

(13) Scuseria, G. E.; Schaefer, H. F. A new implementation of the full CCSDT model for molecular electronic structure. *Chem. Phys. Lett.* **1988**, *152*, 382–386.

(14) Noga, J.; Bartlett, R. J. The full CCSDT model for molecular electronic structure. *J. Chem. Phys.* **1987**, *86*, 7041–7050.

(15) Watts, J. D.; Bartlett, R. J. The coupled-cluster single, double, and triple excitation model for open-shell single reference functions. *J. Chem. Phys.* **1990**, *93*, 6104–6105.

(16) Oliphant, N.; Adamowicz, L. Coupled-cluster method truncated at quadruples. *J. Chem. Phys.* **1991**, *95*, 6645–6651.

(17) Kucharski, S. A.; Bartlett, R. J. Recursive intermediate factorization and complete computational linearization of the coupled-cluster single, double, triple, and quadruple excitation equations. *J. Theoret. Chim. Acta* **1991**, *80*, 387–405.

(18) Kucharski, S. A.; Bartlett, R. J. The coupled-cluster single, double, triple, and quadruple excitation method. *J. Chem. Phys.* **1992**, *97*, 4282–4288.

(19) Musiał, M.; Kucharski, S. A.; Bartlett, R. J. Formulation and implementation of the full coupled-cluster method through pentuple excitations. *J. Chem. Phys.* **2002**, *116*, 4382–4388.

(20) Piecuch, P.; Kucharski, S. A.; Špirko, V. Coupled-cluster methods with internal and semi-internal triply excited clusters: Vibrational spectrum of the HF molecule. *J. Chem. Phys.* **1999**, *111*, 6679–6692.

(21) Lyakh, D. I.; Musiał, M.; Lotrich, V. F.; Bartlett, R. J. Multireference Nature of Chemistry: The Coupled-Cluster View. *Chemical Reviews* **2011**, *112*, 182–243.

(22) Paldus, J.; Čížek, J.; Takahashi, M. Approximate account of the connected quadruply excited clusters in the coupled-pair many-electron theory. *Physical Review A* **1984**, *30*, 2193–2209.

(23) Paldus, J.; Planelles, J. Valence bond corrected single reference coupled cluster approach. *J. Theoret. Chim. Acta* **1994**, *89*, 13–31.

(24) Planelles, J.; Paldus, J.; Li, X. Valence bond corrected single reference coupled cluster approach. *J. Theoret. Chim. Acta* **1994**, *89*, 33–57.

(25) Planelles, J.; Paldus, J.; Li, X. Valence bond corrected single reference coupled cluster approach. *J. Theoret. Chim. Acta* **1994**, *89*, 59–76.

(26) Stolarczyk, L. Z. Complete active space coupled-cluster method. Extension of single-reference coupled-cluster method using the CASSCF wavefunction. *Chem. Phys. Lett.* **1994**, *217*, 1–6.

(27) Peris, G.; Planelles, J.; Paldus, J. Single-reference CCSD approach employing three-

and four-body CAS SCF corrections: A preliminary study of a simple model. *Int. J. Quantum Chem.* **1997**, *62*, 137–151.

(28) Li, X.; Paldus, J. Reduced multireference CCSD method: An effective approach to quasidegenerate states. *J. Chem. Phys.* **1997**, *107*, 6257–6269.

(29) Li, X.; Paldus, J. Reduced multireference couple cluster method. II. Application to potential energy surfaces of HF, F2, and H2O. *J. Chem. Phys.* **1998**, *108*, 637–648.

(30) Peris, G.; Planelles, J.; Malrieu, J.-P.; Paldus, J. Perturbatively selected CI as an optimal source for externally corrected CCSD. *J. Chem. Phys.* **1999**, *110*, 11708–11716.

(31) Huron, B.; Malrieu, J. P.; Rancurel, P. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. *J. Chem. Phys.* **1973**, *58*, 5745–5759.

(32) Xu, E.; Li, S. The externally corrected coupled cluster approach with four- and five-body clusters from the CASSCF wave function. *J. Chem. Phys.* **2015**, *142*, 094119.

(33) Kinoshita, T.; Hino, O.; Bartlett, R. J. Coupled-cluster method tailored by configuration interaction. *J. Chem. Phys.* **2005**, *123*, 074106.

(34) Melnichuk, A.; Bartlett, R. J. Relaxed active space: Fixing tailored-CC with high order coupled cluster. I. *J. Chem. Phys.* **2012**, *137*, 214103.

(35) Veis, L.; Antalík, A.; Brabec, J.; Neese, F.; Örs Legeza,; Pittner, J. Coupled Cluster Method with Single and Double Excitations Tailored by Matrix Product State Wave Functions. *J. Phys. Chem. Lett.* **2016**, *7*, 4072–4078.

(36) Faulstich, F. M.; Máté, M.; Laestadius, A.; Csirik, M. A.; Veis, L.; Antalik, A.; Brabec, J.; Schneider, R.; Pittner, J.; Kvaal, S.; Örs Legeza, Numerical and Theoretical Aspects of the DMRG-TCC Method Exemplified by the Nitrogen Dimer. *J. Chem. Theory Comput.* **2019**, *15*, 2206–2220.

(37) Antalík, A.; Veis, L.; Brabec, J.; Demel, O.; Örs Legeza,; Pittner, J. Toward the efficient local tailored coupled cluster approximation and the peculiar case of oxo-Mn(Salen). *J. Chem. Phys.* **2019**, *151*, 084112.

(38) Schriber, J. B.; Evangelista, F. A. Communication: An adaptive configuration interaction approach for strongly correlated electrons with tunable accuracy. *J. Chem. Phys.* **2016**, *144*, 161106.

(39) Schriber, J. B.; Evangelista, F. A. Adaptive Configuration Interaction for Computing Challenging Electronic Excited States with Tunable Accuracy. *J. Chem. Theory Comput.* **2017**, *13*, 5354–5366.

(40) Paldus, J. Externally and internally corrected coupled cluster approaches: an overview. *J. Math. Chem.* **2016**, *55*, 477–502.

(41) Lehtola, S.; Tubman, N. M.; Whaley, K. B.; Head-Gordon, M. Cluster decomposition of full configuration interaction wave functions: A tool for chemical interpretation of systems with strong correlation. *J. Chem. Phys.* **2017**, *147*, 154105.

(42) Li, X.; Paldus, J. Reduced multireference coupled cluster method with singles and doubles: Perturbative corrections for triples. *J. Chem. Phys.* **2006**, *124*, 174101.

(43) Rendell, A. P.; Lee, T. J.; Komornicki, A. A parallel vectorized implementation of triple excitations in CCSD(T): application to the binding energies of the AlH3, AlH2F, AlHF2 and AlF3 dimers. *Chem. Phys. Lett.* **1991**, *178*, 462–470.

(44) Davis, M. M.; Aroeira, G. J. R.; Turney, J. M.; Schaefer, H. F. Fermi Quantum Chemistry Package. See https://github.com/FermiQC/Fermi.jl.

(45) Bezanson, J.; Edelman, A.; Karpinski, S.; Shah, V. B. Julia: A Fresh Approach to Numerical Computing. *SIAM Rev.* **2017**, *59*, 65–98.

(46) Kállay, M. et al. The MRCC program system: Accurate quantum chemistry from water to proteins. *J. Chem. Phys.* **2020**, *152*, 074107.

(47) MRCC, a quantum chemical program suite written by M. Kállay, P. R. Nagy, D. Mester, Z. Rolik, G. Samu, J. Csontos, J. Csóka, P. B. Szabó, L. Gyevi-Nagy, B. Hégely, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P. D. Mezei, and Á. Ganyecz. See www.mrcc.hu.

(48) Kállay, M.; Surján, P. R. Higher excitations in coupled-cluster theory. *J. Chem. Phys.* **2001**, *115*, 2945–2954.

(49) Pillai, M.; Goglio, J.; Walker, T. G. Matrix Numerov method for solving Schrödinger's equation. *Am. J. Phys.* **2012**, *80*, 1017–1019.

(50) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(51) Lesiuk, M.; Przybytek, M.; Balcerzak, J. G.; Musiał, M.; Moszynski, R. Ab initio Potential Energy Curve for the Ground State of Beryllium Dimer. *J. Chem. Theory Comput.* **2019**, *15*, 2470–2480.

(52) Merritt, J. M.; Bondybey, V. E.; Heaven, M. C. Beryllium Dimer–Caught in the Act of Bonding. *Science* **2009**, *324*, 1548–1551.

(53) Khatib, M. E.; Bendazzoli, G. L.; Evangelisti, S.; Helal, W.; Leininger, T.; Tenti, L.; Angeli, C. Beryllium Dimer: A Bond Based on Non-Dynamical Correlation. *J. Phys. Chem. A* **2014**, *118*, 6664–6673.

(54) Martin, J. M. The ground-state spectroscopic constants of Be2 revisited. *Chem. Phys. Lett.* **1999**, *303*, 399–407.

(55) Patkowski, K.; Podeszwa, R.; Szalewicz, K. Interactions in Diatomic Dimers Involving Closed-Shell Metals†. *J. Phys. Chem. A* **2007**, *111*, 12822–12838.

(56) Koput, J. The ground-state potential energy function of a beryllium dimer determined using the single-reference coupled-cluster approach. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20311.

(57) Khan, M. A. H.; Percival, C. J.; Caravan, R. L.; Taatjes, C. A.; Shallcross, D. E. Criegee intermediates and their impacts on the troposphere. *Environ. Sci. Process Impacts* **2018**, *20*, 437–453.

(58) Long, B.; Bao, J. L.; Truhlar, D. G. Atmospheric Chemistry of Criegee Intermediates: Unimolecular Reactions and Reactions with Water. *J. Am. Chem. Soc.* **2016**, *138*, 14409–14422.

(59) Misiewicz, J. P.; Elliott, S. N.; Moore, K. B.; Schaefer, H. F. Re-examining ammonia addition to the Criegee intermediate: converging to chemical accuracy. *Phys. Chem. Chem. Phys.* **2018**, *20*, 7479–7491.

(60) Ruscic, B.; Pinzon, R. E.; von Laszewski, G.; Kodeboyina, D.; Burcat, A.; Leahy, D.; Montoy, D.; Wagner, A. F. Active Thermochemical Tables: thermochemistry for the 21st century. *J. Phys. Conf. Ser.* **2005**, *16*, 561–570.

(61) Whitman, D. W.; Carpenter, B. K. Limits on the activation parameters for automerization of cyclobutadiene-1, 2-d2. *J. Am. Chem. Soc.* **1982**, *104*, 6473–6474.

(62) Li, X.; Paldus, J. Accounting for the exact degeneracy and quasidegeneracy in the automerization of cyclobutadiene via multireference coupled-cluster methods. *J. Chem. Phys.* **2009**, *131*, 114103.

(63) Eckert-Maksić, M.; Vazdar, M.; Barbatti, M.; Lischka, H.; Maksić, Z. B. Automerization reaction of cyclobutadiene and its barrier height: Anab initiobenchmark multireference average-quadratic coupled cluster study. *J. Chem. Phys.* **2006**, *125*, 064310.

Graphical TOC Entry

Some journals require a graphical entry for the Table of Contents. This should be laid out "print ready" so that the sizing of the text is correct.

Inside the `tocentry` environment, the font used is Helvetica 8 pt, as required by *Journal of the American Chemical Society*.

The surrounding frame is 9 cm by 3.5 cm, which is the maximum permitted for *Journal of the American Chemical Society* graphical table of content entries. The box will not resize if the content is too big: instead it will overflow the edge of the box.

This box and the associated title will always be printed on a separate page at the end of the document.