Alternative single-reference coupled cluster approaches for multireference problems: The simpler, the better

Francesco A. Evangelista

View Online   Export Citation

ARTICLES YOU MAY BE INTERESTED IN

Alternative single-reference coupled cluster approaches for multireference problems: The simpler, the better

Francesco A. Evangelista[a)]
Institut für Physikalische Chemie, Universität Mainz, D-55099 Mainz, Germany

We report a general implementation of alternative formulations of single-reference coupled cluster theory (extended, unitary, and variational) with arbitrary-order truncation of the cluster operator. These methods are applied to compute the energy of Ne and the equilibrium properties of HF and C_2. Potential energy curves for the dissociation of HF and the BeH_2 model computed with the extended, variational, and unitary coupled cluster approaches are compared to those obtained from the multireference coupled cluster approach of Mukherjee *et al.* [J. Chem. Phys. **110**, 6171 (1999)] and the internally contracted multireference coupled cluster approach [F. A. Evangelista and J. Gauss, J. Chem. Phys. **134**, 114102 (2011)]. In the case of Ne, HF, and C_2, the alternative coupled cluster approaches yield almost identical bond length, harmonic vibrational frequency, and anharmonic constant, which are more accurate than those from traditional coupled cluster theory. For potential energy curves, the alternative coupled cluster methods are found to be more accurate than traditional coupled cluster theory, but are three to ten times less accurate than multireference coupled cluster approaches. The most challenging benchmark, the BeH_2 model, highlights the strong dependence of the alternative coupled cluster theories on the choice of the Fermi vacuum. When evaluated by the accuracy to cost ratio, the alternative coupled cluster methods are not competitive with respect to traditional CC theory, in other words, the simplest theory is found to be the most effective one. © *2011 American Institute of Physics.* [doi:10.1063/1.3598471]

I. INTRODUCTION

The defining feature of single-reference (SR) coupled cluster (CC) theories is the parameterization of the exact electronic wave function Ψ in terms of a reference Slater determinant (Φ) and an exponential operator $\exp(\hat{S})$ which maps Φ into Ψ,

$$|\Psi\rangle = e^{\hat{S}} |\Phi\rangle . \tag{1}$$

The exponential structure of the SRCC wave function naturally leads to a set of connected equations that ensure the correct scaling of the energy and other properties with respect to the size of the system[1–3] and the correct energy separability of noninteracting fragments.[4]

The coupled cluster wave function ansatz does not define a unique computational approach because there are several possible parameterizations of the cluster operator $\hat{S}$ and different approaches for finding solutions to the Schrödinger equation of the form of Eq. (1). Among the several variants of single-reference coupled cluster theory, the initial formulation introduced by Coester, Kümmel, and Čížek[5–7] – here referred to as traditional CC (TCC) theory[8] – has emerged as a *de facto* standard in quantum chemistry and is often used in high-accuracy computations of properties of both atoms and small to medium sized molecules.[9,10]

A key factor in determining the success of TCC theory is the high accuracy of schemes in which the operator $\hat{S}$ is truncated to include only low-order excitations. Coupled cluster theory with single and double substitutions (TCCSD) (Ref. 11) and CCSD with perturbative triple excitation corrections [TCCSD(T)] (Ref. 12) are very accurate theoretical models, and can be used in conjunction with basis set extrapolation[13,14] or explicitly correlated approaches[15] to determine thermochemical parameters approaching chemical accuracy, here defined as 1 kcal mol^{-1}.

Notwithstanding the success of TCC theory, there are two unsatisfactory aspects of this approach. The first problem is the high computational costs required to achieve chemical accuracy. Indeed, despite the fact that TCC energy converges quickly to the full-configuration-interaction (FCI) energy with respect to the truncation of the excitation rank,[16] triple and quadruple excitations are often necessary to achieve chemical accuracy.[14] This observation naturally leads to the question: Is traditional coupled cluster theory the most efficiency parameterization of the electronic wave function?

Nooijen has conjectured that it is possible to parameterize the exact wave function with a general two-body operator containing as many parameters as the number of matrix elements of the Hamiltonian, which grows only with the quartic power of the number of electrons.[17] Although this conjecture was disproved, Nooijen's generalized coupled cluster approach was shown to be surprisingly accurate.[18,19] More recently, the adaptive[20] and stochastic[21] coupled cluster approaches have also been successful in achieving accuracy close to that of the truncated TCC theory but using a smaller number of parameters.

The second drawback of TCC theory is the failure to describe multireference problems, which require a

---

[a)]Electronic mail: evangeli@uni-mainz.de.

zeroth-order wave function that is a linear combination of Slater determinants. In principle, the best solution to multireference problems is a genuine multireference coupled cluster (MRCC) approach that starts from a multideterminantal zeroth-order wave function.[10,22–24] However, several problems – most importantly intruder states[25,26] – have prevented the wide spread use of genuine MRCC theories in chemical applications. Only recently, with the introduction of state-specific MRCC methods[27–36] and other theories that avoid intruder states,[37,38] MRCC theories have been successfully applied to realistic chemical systems.[39–41] Nevertheless, MRCC approaches are still being developed and can currently be applied only to systems with a small number of active orbitals.

In principle, one can still treat multireference problems in the framework of TCC theory by including higher-than-double excitations. However, because of its high computational costs, this solution is only viable for small systems. Several strategies based on the TCC approach have been advanced to deal with multireference problems, including: (i) introducing selected higher excitations,[42–48] (ii) taking into consideration high-order moments of the coupled cluster equations, (iii) applying renormalization techniques,[49] and (iv) employing equation-of-motion approaches.[50–53] Since these approaches depend on the choice of a particular reference determinant, they can incur problems when the reference determinant is not the leading contribution to the wave function.[34,54]

Can one solve the problems of traditional CC theory by exploiting the flexibility of the exponential parameterization? Several alternative formulations of CC methods have been considered in the literature (see Ref. 55 for a review), including expectation value and variational CC theory (XCC,VCC),[56–58] the unitary CC (UCC) approach,[59–62] the extended CC (ECC) method,[19,63–67] and the generalized coupled cluster ansatz (GCC).[17–19,68–71] Among these approaches, the extended coupled cluster theory of Arponen[63] is perhaps the most promising since, contrary to the other methods, its equations contain a finite number of terms. However, at a given truncation level, the computational costs of ECC theory are higher than those of the traditional CC method. For example, the computational cost of ECCSD scales as N^{10}, where N is the size of the system,[55] while TCCSD has a cost proportional to N^6.

Most of the arguments in favor of alternative formulations of the SRCC theory are based on theoretical considerations. In the case of extended CC theory, Bishop and Arponen[65] claim that it is robust enough to deal with symmetry breaking problems and describe global potential energy surfaces. Relying on formal analyses, Kutzelnigg suggests that variational, unitary, and extended coupled cluster theories are more accurate than the traditional CC theory at low order of truncation.[62,72] Numerical implementations of alternative CC methods have been considered by van Voorhis and Head-Gordon,[58] Fan et al.,[73] and Cooper and Knowles.[74] These pioneering studies have provided evidence that the alternative SRCC approaches are more accurate than the TCC theory and that they may also be used to treat multireference problems.

In this paper, we investigate the extended, unitary, and variational CC theories with excitation levels ranging from doubles to hextuples, and benchmark the performance of these methods with respect to traditional single-reference coupled cluster theory and full-configuration-interaction (FCI) results. Of particular interest is the possibility of using alternative single-reference coupled cluster methods as black-box approaches for treating multireference problems. Therefore, we also computed the potential energy curve of HF and of the BeH_2 model, comparing the results with those obtained from the state-specific and internally contracted multireference coupled cluster theory.[28,36] Our results suggest that while the alternative methods are more robust and accurate with respect to the traditional CC theory, they cannot be considered general black-box solutions to the multireference problem.

II. THEORY

A. Parametrization of the cluster operator

For all the approaches considered here, we assume that the reference Slater determinant Φ entering Eq. (1) is built from a set of orthonormal spin orbitals $\{\phi_p, p = 1, \ldots, N\}$. We use the convention of labeling spin orbitals occupied in Φ with $i, j, \ldots$ while unoccupied (virtual) spin orbitals are labeled with $a, b, \ldots$. The cluster operator ($\hat{S}$) may be parameterized in several ways. In the traditional CC theory, $\hat{S}$ is a pure excitation operator ($\hat{T}$), truncated to a given excitation level n,

$$\hat{T} = \sum_{k=1}^{n} \hat{T}_k, \qquad (2)$$

where each $\hat{T}_k$ is a k-fold excitation operator defined as

$$\hat{T}_k = \frac{1}{(k!)^2} \sum_{ij\cdots}^{occ} \sum_{ab\cdots}^{vir} t_{ij\cdots}^{ab\cdots} \hat{\tau}_{ij\cdots}^{ab\cdots}. \qquad (3)$$

The cluster amplitudes ($t_{ij\cdots}^{ab\cdots}$) entering Eq. (3) are $2k$-dimensional tensors, antisymmetric with respect to separate permutations among the upper and lower indices, while $\hat{\tau}_{ij\cdots}^{ab\cdots}$ are k-fold tensor excitation operators, defined using second quantization creation ($\hat{a}^\dagger$) and annihilation ($\hat{a}$) operators as $\hat{\tau}_{ij\cdots}^{ab\cdots} = \hat{a}_a^\dagger \hat{a}_b^\dagger \cdots \hat{a}_j \hat{a}_i$.

Alternatively, $\hat{S}$ can be chosen to be an anti-Hermitian operator ($\hat{A}$), truncated analogously to $\hat{T}$ to a given rank n,

$$\hat{A} = \sum_{k=1}^{n} \hat{A}_k, \qquad (4)$$

where each component $\hat{A}_k$ is defined as

$$\hat{A}_k = \frac{1}{(k!)^2} \sum_{ij\cdots}^{occ} \sum_{ab\cdots}^{vir} a_{ij\cdots}^{ab\cdots} [\hat{\tau}_{ij\cdots}^{ab\cdots} - \hat{\tau}_{ab\cdots}^{ij\cdots}], \qquad (5)$$

and the amplitudes $a_{ij\cdots}^{ab\cdots}$ have the same permutational symmetry as the $t_{ij\cdots}^{ab\cdots}$ amplitudes. Being $\hat{A}$ anti-Hermitian, $\hat{A}^\dagger = -\hat{A}$, and consequently, $\exp(\hat{A})$ is a unitary operator.

For convenience, $\hat{T}$ and $\hat{A}$ can be written in a compact form by introducing general excitation operators ($\hat{\tau}_q$) and the corresponding amplitudes (t_q, a_q) labeled by a collective index $q = 1, \ldots, n_{\mathrm{ex}}$ ($q \equiv {}^{ab\cdots}_{ij\cdots}$),

$$\hat{T} = \sum_{q=1}^{n_{\mathrm{ex}}} t_q \hat{\tau}_q, \tag{6}$$

$$\hat{A} = \sum_{q=1}^{n_{\mathrm{ex}}} a_q (\hat{\tau}_q - \hat{\tau}_q^\dagger). \tag{7}$$

Using this notation, a generic excited determinant obtained from Φ can be written as $|\Phi_q\rangle = \hat{\tau}_q |\Phi\rangle$ and the set of all excited determinants spanned at the linear level by the cluster operator is Q, defined as

$$Q = \{\Phi_q : |\Phi_q\rangle = \hat{\tau}_q |\Phi\rangle, q = 1, \ldots, n_{\mathrm{ex}}\}. \tag{8}$$

B. Projective approaches

In traditional coupled cluster theory, the cluster operator is determined by requiring that projections of the Schrödinger equation onto the set Q are satisfied. This is achieved by inserting the wave function ansatz [Eq. (1)] into the Schrödinger equation

$$\hat{H} e^{\hat{S}} |\Phi\rangle = E e^{\hat{S}} |\Phi\rangle, \tag{9}$$

premultiplying[75] the Schrödinger equation on the left by $e^{-\hat{S}}$, and projecting with $\langle\Phi|$ and $\langle\Phi_q|$,

$$\langle\Phi| e^{-\hat{S}} \hat{H} e^{\hat{S}} |\Phi\rangle = E, \tag{10}$$

$$\langle\Phi_q| e^{-\hat{S}} \hat{H} e^{\hat{S}} |\Phi\rangle = 0 \quad \forall \Phi_q \in Q. \tag{11}$$

Equation (10) gives the energy (E) in terms of the cluster operator $\hat{S}$, while Eq. (11) is a system on nonlinear equations that determines the cluster amplitudes. The quantity $\bar{H} = e^{-\hat{S}} \hat{H} e^{\hat{S}}$ is the similarity-transformed Hamiltonian and can be expressed as a series via the Baker-Campbell-Hausdorff (BCH) expansion

$$e^{-\hat{S}} \hat{H} e^{\hat{S}} = \hat{H} + [\hat{H}, \hat{S}] + \frac{1}{2}[[\hat{H}, \hat{S}], \hat{S}] + \ldots$$
$$= \hat{H} + \sum_{k=1}^{\infty} \frac{1}{k!} \underbrace{[[\hat{H}, \hat{S}], \ldots \hat{S}]}_{k-\text{nested commutators}}. \tag{12}$$

The traditional and unitary coupled cluster theories are both based on the projective approach and differ in the choice of cluster operator. In the traditional CC approach $\hat{S} = \hat{T}$, while in the projective unitary CC (UCC) theory the cluster operator is anti-Hermitian ($\hat{S} = \hat{A}$). The former approach has the advantage that the Baker-Campbell-Hausdorff expansion of $\bar{H}$ terminates after the term containing four commutators, while for unitary CC theory the series does not terminate.[62] Yanai and Chan[76] and Taube and Bartlett[77] suggested ways to approximate $\bar{H}$ in unitary coupled cluster theory.

C. Variational approaches

Variational couples cluster approaches seek to minimize the Rayleigh-Ritz energy functional

$$E[\hat{S}] = \frac{\langle\Phi| e^{\hat{S}^\dagger} \hat{H} e^{\hat{S}} |\Phi\rangle}{\langle\Phi| e^{\hat{S}^\dagger} e^{\hat{S}} |\Phi\rangle}. \tag{13}$$

The variational approaches have the advantage that the energy is an upper bound of the full-configuration-interaction value, a condition that is not necessarily satisfied by the projective approaches. In the variational approaches, the cluster amplitudes are determined by imposing the stationarity condition

$$\frac{\partial E[\hat{S}]}{\partial s_q} = 0 \quad q = 1, \ldots, n_{\mathrm{ex}}. \tag{14}$$

In variational coupled cluster (VCC) theory, the cluster operator is a pure excitation operator. It follows then that the derivative of the Rayleigh-Ritz energy functional with respect to the variation of the cluster amplitudes can be evaluated directly as

$$\frac{\partial E[\hat{T}]}{\partial t_q} = \frac{2\mathrm{Re}\langle\Phi_q| e^{\hat{T}^\dagger} (\hat{H} - E) e^{\hat{T}} |\Phi\rangle}{\langle\Phi| e^{\hat{T}^\dagger} e^{\hat{T}} |\Phi\rangle}. \tag{15}$$

For a finite system, the VCC equations truncate and contain terms up to $\hat{T}^{n_{\mathrm{el}}}$, where n_{el} is the number of electrons.

One can also define a variational unitary CC approach[62] (VUCC, often referred to in the literature with the acronym UCC) by formulating the variational energy functional,

$$E[\hat{A}] = \langle\Phi| e^{\hat{A}^\dagger} \hat{H} e^{\hat{A}} |\Phi\rangle, \tag{16}$$

and minimizing $E[\hat{A}]$ with respect to $\hat{A}$. In this case, one evaluates the derivative of $\exp(\hat{A})$ via the Wilcox identity.[18,78] The VUCC was recently implemented by Cooper and Knowles.[74] Because in the projective unitary CC approach, the energy is evaluated via the same functional used in the VUCC theory, it follows that the UCC energy is also variational.

D. The coupled cluster energy functional and extended coupled cluster theory

The equations that define traditional coupled cluster theory can be cast into an energy functional

$$\tilde{E}_{\mathrm{TCC}} = \langle\Phi| (1 + \hat{\Lambda}) e^{-\hat{T}} \hat{H} e^{\hat{T}} |\Phi\rangle, \tag{17}$$

where $\hat{\Lambda}$ is a de-excitation operator analogous to $\hat{T}^\dagger$, defined as

$$\hat{\Lambda} = \sum_q \lambda_q \hat{\tau}_q^\dagger. \tag{18}$$

Here, $\tilde{E}_{\mathrm{TCC}}$ is a functional of the cluster (t_q) and lambda amplitudes (λ_q) and its extremum gives a set of amplitudes which solves Eqs. (10) and (11).

In the extended coupled cluster (ECC) approach of Arponen and co-workers[63–65] the linear operator $1 + \hat{\Lambda}$ is replaced with the exponential operator $1 + \hat{\Lambda} = \exp(\hat{\Sigma})$, where $\hat{\Sigma}$ is a de-excitation operator analogous to $\hat{\Lambda}$,

$$\hat{\Sigma} = \sum_q \sigma_q \hat{\tau}_q^\dagger, \tag{19}$$

and the corresponding sets of amplitudes are related via the condition $\lambda_q = \langle\Phi| e^{\hat{\Sigma}} |\Phi_q\rangle$. Substituting the exponential

TABLE I. Definition of the single-reference coupled cluster approaches implemented in this work.

CC method	Energy equation	Amplitudes equation
Traditional	$E = \langle\Phi|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi\rangle$	$\langle\Phi_q|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi\rangle = 0$
Unitary	$E = \langle\Phi|e^{-\hat{A}}\hat{H}e^{\hat{A}}|\Phi\rangle$	$\langle\Phi_q|e^{-\hat{A}}\hat{H}e^{\hat{A}}|\Phi\rangle = 0$
Variational	$E = \dfrac{\langle\Phi|e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}|\Phi\rangle}{\langle\Phi|e^{\hat{T}^\dagger}e^{\hat{T}}|\Phi\rangle}$	$\mathrm{Re}\langle\Phi_q|e^{\hat{T}^\dagger}(\hat{H}-E)e^{\hat{T}}|\Phi\rangle = 0$
Extended	$E = \langle\Phi|e^{\hat{\Sigma}}e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi\rangle$	$\langle\Phi_q|e^{\hat{\Sigma}}e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi\rangle = 0$
		$\langle\Phi|e^{\hat{\Sigma}}e^{-\hat{T}}[\hat{H},\hat{\tau}_q]e^{\hat{T}}|\Phi\rangle = 0$

parameterization of $\hat{\Lambda}$ in $\tilde{E}_{\mathrm{TCC}}$ yields the extended coupled cluster energy functional ($\tilde{E}_{\mathrm{ECC}}$),

$$\tilde{E}_{\mathrm{ECC}} = \langle\Phi|\,e^{\hat{\Sigma}}e^{-\hat{T}}\hat{H}e^{\hat{T}}\,|\Phi\rangle. \qquad (20)$$

Since Eq. (20) is equivalent to the TCC energy functional, its extremum will also yield (when no truncation is enforced) the exact solution to the Schrödinger equation.[79] The ECC amplitude equations result from imposing stationary conditions on $\tilde{E}_{\mathrm{ECC}}$. This yields a set of equations for the cluster amplitudes and one for the sigma amplitudes, that contrary to traditional CC theory, are coupled and must be solved simultaneously (see Table I). Interestingly, the ECC amplitude equation obtained by differentiation of $\tilde{E}_{\mathrm{ECC}}$ with respect to σ_q can be written as $\langle\Phi|e^{\hat{\Sigma}}\hat{\tau}_q^\dagger\bar{H}|\Phi\rangle = 0, \forall\hat{\tau}_q\Phi \in Q$. This expression may be viewed as a modified TCC amplitude equation, where instead of the bare projection manifold $\{\langle\Phi_q| = \langle\Phi|\hat{\tau}_q^\dagger\}$, one uses a dynamical projection manifold $\{\langle\tilde{\Phi}|\hat{\tau}_q^\dagger = \langle\Phi|e^{\hat{\Sigma}}\hat{\tau}_q^\dagger\}$, which effectively introduces projections on higher order excited determinants. This interpretation of the ECC equations shows some similarity with the active-space CC amplitude equations considered by Adamowicz, Oliphant, and Piecuch.[42,43] This method considers selected higher-that-double excitations with the aim of simulating single and double excitations out of determinants that are nearly degenerate with respect to the Hartree-Fock reference. However, there is a subtle difference with respect to the ECC theory. In the case of the active-space CC approach, one introduces both a set of amplitudes and equations related to higher excitations, while in the ECC theory the number of amplitudes is not augmented.

III. IMPLEMENTATION

The traditional, extended, unitary, and variational coupled cluster theories are implemented in a pilot code capable of performing computations with arbitrary truncation of the excitation level. The code is interfaced with the PSI3 quantum chemistry package in order to read one- and two-electron integrals in the molecular orbital basis set.[80] All methods were implemented expanding the corresponding equations in the full-configuration-interaction basis set of determinants $\Phi_{\mathcal{P}}$, where $\mathcal{P}$ is a string that specifies the spin-orbital occupancy of $\Phi_{\mathcal{P}}$. Using this notation, a generic wave function $|\Omega\rangle$ is written as

$$|\Omega\rangle = \sum_{\mathcal{P}}|\Phi_{\mathcal{P}}\rangle\,C_{\mathcal{P}}, \qquad (21)$$

where the coefficients $C_{\mathcal{P}} = \langle\Phi_{\mathcal{P}}|\Omega\rangle$, are given by the projection of $|\Omega\rangle$ onto $|\Phi_{\mathcal{P}}\rangle$.

To compute the action of the Hamiltonian operator on a generic state vector $|\Omega\rangle$, that is $\hat{H}|\Omega\rangle$, we use the FCI algorithm of Bendazzoli and Evangelisti.[81] The action of $\exp(\hat{S})$ on a state vector $|\Omega\rangle$ is computed using the Taylor series expansion of the exponential operator

$$\exp(\hat{S})|\Omega\rangle = |\Omega\rangle + \hat{S}|\Omega\rangle + \frac{1}{2!}\hat{S}^2|\Omega\rangle + \ldots, \qquad (22)$$

and the algorithm advanced by Hirata and Bartlett.[82] The summation in Eq. (22) is truncated when the largest element of the vector $(1/k!)\hat{S}^k|\Omega\rangle$ is less than 10^{-12}. A generalization of the Bendazzoli-Evangelisti algorithm is used to compute the effect of the cluster operator $\hat{S}$ on a generic state vector. In order to compute the BCH expansions of the similarity-transformed Hamiltonian truncated at a certain number of commutators we use the combinatorial identity

$$\underbrace{[[\hat{H},\hat{S}],\ldots,\hat{S}]}_{k-\text{nested commutators}} = \sum_{l=0}^{k}(-1)^l\binom{k}{l}\hat{S}^l\hat{H}\hat{S}^{k-l}. \qquad (23)$$

For all the methods implemented in this work, the corresponding amplitude equations (see Table I) are solved iteratively by a general procedure. From an initial guess of the amplitudes, we compute the residual vector $|r\rangle$ according to

$$|r\rangle = \begin{cases} e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi\rangle & \text{TCC} \\ e^{-\hat{A}}\hat{H}e^{\hat{A}}|\Phi\rangle & \text{UCC} \\ e^{\hat{T}^\dagger}(\hat{H}-E)e^{\hat{T}}|\Phi\rangle & \text{VCC} \\ e^{\hat{\Sigma}}e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi\rangle & \text{ECC.} \end{cases} \qquad (24)$$

The coupled cluster equations for the various methods are satisfied when the projection of the residual on the excited determinants is null, $r_q = \langle\Phi_q|r\rangle = 0$ for $q = 1,\ldots,n_{\mathrm{ex}}$.[83] The residuals r_q are used to update the amplitudes s_q (t_q or a_q) via the equation

$$s_q(\text{new}) = s_q(\text{old}) + \frac{r_q}{\Delta_q}, \qquad (25)$$

where Δ_q is a Møller-Plesset denominator computed from the canonical Hartree-Fock orbital energies (ε_i),

$$\Delta_q \equiv \Delta_{ij\ldots}^{ab\cdots} = \varepsilon_a + \varepsilon_b + \ldots - \varepsilon_i - \varepsilon_j - \ldots. \qquad (26)$$

In the case of extended CC theory, an additional set of equations is solved. The residual corresponding to the second set of equations (r'_q) is computed according to

$$r'_q = \langle\Phi|\,e^{\hat{\Sigma}}e^{-\hat{T}}[\hat{H},\hat{\tau}_q]e^{\hat{T}}|\Phi\rangle, \qquad (27)$$

$$= \langle\Phi|\,e^{\hat{\Sigma}}e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_q\rangle - \langle\Phi|\,e^{\hat{\Sigma}}e^{-\hat{T}}\hat{\tau}_q\hat{H}e^{\hat{T}}|\Phi\rangle. \qquad (28)$$

For the computation of this term we form three vectors

$$|A\rangle = e^{\hat{T}^\dagger}\hat{H}e^{-\hat{T}^\dagger}e^{\hat{\Sigma}^\dagger}|\Phi\rangle,$$

$$|B\rangle = e^{-\hat{T}^\dagger}e^{\hat{\Sigma}^\dagger}|\Phi\rangle, \qquad (29)$$

$$|C\rangle = \hat{H}e^{\hat{T}}|\Phi\rangle,$$

and obtain r'_q as

$$r'_q = \langle \Phi_q | A \rangle^* - \langle B | \hat{\tau}_q | C \rangle. \qquad (30)$$

The residual r'_q is then used to update the sigma amplitudes σ_q via

$$\sigma_q(\text{new}) = \sigma_q(\text{old}) + \frac{r'_q}{\Delta_q}. \qquad (31)$$

After the amplitude update step, the energy is computed using the equations reported in Table I. The convergence of the iterative procedure is monitored via the energy change between iterations $\Delta E = E(\text{new}) - E(\text{old})$ and the norm of the residual vector r_q (together with the norm of r'_q in the case of ECC theory). In examples with quasi-degenerate frontier molecular orbitals, we found it useful to shift the Møller-Plesset denominators by a factor of η. The direct inversion in the iterative subspace (DIIS) method is applied to reduce the number of iterative cycles by extrapolation of the amplitudes.[84]

The coupled cluster approaches implemented in this work were validated against the TCC results of Hirata and Bartlett,[82] the extended coupled cluster benchmark data published by Fan *et al.*,[73] and the variational CC result of Cooper and Knowles.[74] The benchmark cases considered in this work are reproduced in the supplementary material.[85] To the best of our knowledge, the projective UCC approach considered in this work was never implemented before.

IV. COMPUTATIONAL DETAILS

All single-reference coupled cluster results were obtained using restricted closed-shell Hartree-Fock (RHF) molecular orbitals while the Mukherjee and internally contracted multireference coupled cluster computations employed complete-active-space self-consistent-field (CASSCF) orbitals. For Ne, HF, and C_2, we employed the standard correlation-consistent polarized-valence basis sets series of Dunning (cc-pVXZ, with X = D, T).[86] The potential energy curves of HF and BeH_2 were computed using a double-zeta valence (DZV) and a [Be $(10s3p/3s2p)$,H$(4s/2s)$] basis set, respectively (provided in the supplementary material[85]). In all computations only the pure spherical harmonic components of the d orbitals were used. The $1s$-like orbitals of C, F, and Ne were frozen in the correlated treatments. Mk-MRCCSD and Mk-MRCCSD with perturbative triple excitations [Mk-MRCCSD(T)] (Ref. 87) results were obtained using the PSIMRCC (Ref. 88) implementation in PSI3.[80]

Equilibrium geometries, harmonic vibrational frequencies, and anharmonic force constants for the diatomic molecules were computed from a sixth-order polynomial fit of the potential energy curve. The fit was obtained by sampling the potential energy curve at seven points equally spaced by 0.01 Å and centered around the equilibrium bond length. The harmonic vibrational frequencies and anharmonic force constants were computed using the equations reported in Ref. 89.

V. RESULTS

A. Ne

The total energy of the neon atom computed using various single-reference coupled cluster approaches is reported in Table II. The cc-pVDZ basis set results suggest that at a given truncation level the alternative coupled cluster approaches are more accurate than traditional coupled cluster theory. For example, TCCSDT recovers 99.916% of the correlation energy, while at the SDT truncation level ECC, UCC, and VCC recover ~99.995% of the correlation energy, which is comparable with the TCCSDTQ result (99.995%).

However, this trend appears to be an artifact, attributable to the size of the cc-pVDZ basis set. Indeed, the energies computed with the cc-pVTZ basis set show that at each truncation level, the correlation energy recovered by traditional coupled cluster theory is comparable to the alternative approaches. This reversal of trend is particularly striking at the SDT truncation level: the TCCSDT approach recovers more correlation energy than the alternative approaches, 99.975% vs. 99.967%, respectively.

All the alternative CC approaches yield energies that are very close. At the single and double (SD) truncation level, the ECC, UCC, and VCC energies differ by less than 10 μE_h. By the time full triple excitations are included, this deviation is less than 1 μE_h, and becomes smaller as the truncation level increases. The similarity of the alternative coupled cluster results at a give truncation level is also observed by Cooper and Knowles[74] and can be understood via the theoretical arguments advanced by Kutzelnigg.[72] By analyzing the case of doubles-only truncated schemes (CCD), Kutzelnigg shows that the traditional CCD energy differs from the VCCD and ECCD energies by a term of order $O(\hat{S}^4)$, while the VCCD and ECCD energies differ by a factor of $O(\hat{S}^6)$. This formal relationship is rendered in Fig. 1. Thus, when the magnitude of $\hat{S}$ is small, one expects the VCC and ECC theories to yield energies that are similar, but still distinct from the TCC energy. A similar analysis can be extended to the case of unitary CC theory.[72]

B. Equilibrium properties of HF

To assess the accuracy of molecular properties computed with the alternative coupled cluster methods, we have computed the bond length, the harmonic vibrational frequency, and the anharmonic constant of hydrogen fluoride. Table III reports these properties for the ground state ($X\,^1\Sigma^+$) of HF. Like in the case of neon, the alternative CC approaches yield results that are almost identical, with the values of r_e, ω_e, and $\omega_e x_e$ computed with the ECC, UCC, and VCC theories differing by less than 10^{-5} Å, 0.2 cm^{-1}, and 0.1 cm^{-1}, respectively. At low orders of truncation, the TCC equilibrium properties are appreciably distinct from the UCC, ECC, and VCC results, particularly in the case of ω_e (4169.0 cm^{-1} for TCCSD vs. 4166.0 cm^{-1} for ECCSD). However, when quadruple excitations are included, the properties computed with all coupled cluster approaches are almost identical, a sign that the respective wave functions are nearly converged to the FCI solution.

TABLE II. Total energy (in E_h) of the ground state state of neon (1^1S_0), computed using the cc-pVDZ and cc-pVTZ basis sets. The value in parentheses is the percentage of correlation energy recovered. The core $1s$ atomic orbital was frozen in all coupled cluster theory computations.

Excitation level	TCC	ECC	UCC	VCC
		cc-pVDZ		
SD	−128.677792257	−128.678002556	−128.677999887	−128.677996583
	(99.352)	(99.463)	(99.461)	(99.459)
SDT	−128.678864848	−128.679017462	−128.679016412	−128.679015512
	(99.916)	(99.996)	(99.995)	(99.995)
SDTQ	−128.679014931	−128.679023992	−128.679023855	−128.679023727
	(99.995)	(99.999)	(99.999)	(99.999)
SDTQP	−128.679025538	−128.679025037	−128.679025029	−128.679025025
	(100.000)	(100.000)	(100.000)	(100.000)
SDTQPH	−128.679025033	−128.679025053	−128.679025053	−128.679025054
	(100.000)	(100.000)	(100.000)	(100.000)
		cc-pVTZ		
SD	−128.798208605	−128.798408977	−128.798405997	−128.798402267
	(98.402)	(98.476)	(98.475)	(98.474)
SDT	−128.802467607	−128.802446722	−128.802446369	−128.802445903
	(99.975)	(99.968)	(99.968)	(99.967)
SDTQ	−128.802533625	−128.802531517	−128.802531496	−128.802531462
	(100.000)	(99.999)	(99.999)	(99.999)
SDTQP	−128.802534189	−128.802533974	−128.802533970	−128.802533967
	(100.000)	(100.000)	(100.000)	(100.000)
SDTQPH	−128.802534058	−128.802534041	−128.802534041	−128.802534040
	(100.000)	(100.000)	(100.000)	(100.000)

The last column of Table III reports the energy error for a given method X with respect to the FCI value, computed at the FCI equilibrium bond length (r_{FCI}),

$$\epsilon = E_X(r_{FCI}) - E_{FCI}(r_{FCI}). \tag{32}$$

As in the case of the properties, the energy error converges quickly to the FCI value with respect to the truncation level. When the alternative CC methods are truncated at low orders – from SD to SDTQ – they recover a larger portion of correlation energy with respect to the traditional CC theory. For example, at the SDT and SDTQ truncation levels, the UCC, ECC, and VCC errors with respect to FCI is approximately five times less than the TCCSDT and TCCSDTQ errors, respectively. However, the alternative coupled cluster

FIG. 1. Formal relationships among the TCCD, ECCD, and VCCD energies.

approaches with n-fold excitations do not perform better than the TCC method with excitations truncated to $n + 1$.

C. Equilibrium properties of C_2

C_2 is well known to be a challenging benchmark case because, even at the equilibrium geometry, the electronic wave function has considerable multireference character.[90–93] Indeed, near the equilibrium geometry (experimental[94] $r_e \approx$ 1.2425 Å), the zeroth-order wave function of C_2 is dominated by two electronic configurations

$$|\Phi_1\rangle = |(1\sigma_g)^2(1\sigma_u)^2(2\sigma_g)^2(2\sigma_u)^2(1\pi_u)^4\rangle,$$
$$|\Phi_2\rangle = |(1\sigma_g)^2(1\sigma_u)^2(2\sigma_g)^2(1\pi_u)^4(3\sigma_g)^2\rangle, \tag{33}$$

which contribute to the zeroth-order wave function with weights of $\sim 87\%$ and $\sim 13\%$, respectively. For this reason, Table IV reports both single-reference and multireference coupled cluster results for the ground state ($X\ ^1\Sigma_g^+$) of C_2.

When compared to the HF case, the equilibrium properties of C_2 computed with the single-reference approaches show a remarkably slow convergence toward the exact value with respect to the truncation level of the cluster operator. Only the inclusion of full quadruple excitations brings the agreement between the single-reference coupled cluster methods and the FCI answer within less than 0.001 Å for the bond length, 2 cm^{-1} for the harmonic vibrational frequency, and 0.1 cm^{-1} for the anharmonic constant. Moreover, as in the case of HF, properties computed with the ECC, UCC,
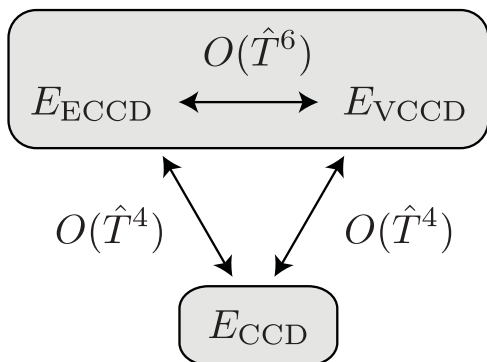
TABLE III. Equilibrium distance (r_e), harmonic vibrational frequency (ω_e), anharmonic constant ($\omega_e x_e$), and energy error with respect to FCI at the FCI geometry (ϵ) for the $X\,^1\Sigma^+$ state of HF computed at the TCC, UCC, ECC, and VCC levels of theory using the cc-pVDZ basis set and RHF orbitals. The fluorine 1s-like core molecular orbital was frozen in all computations. The FCI r_e, ω_e, $\omega_e x_e$, and total energy at r_e are 0.92025 Å, 4143.5 cm^{-1}, 92.9 cm^{-1}, and $-100.228652\ E_h$, respectively.

Excitation level	TCC	ECC	UCC	VCC
		r_e (Å)		
SD	0.91895	0.91914	0.91914	0.91913
SD(T)	0.91988	…	…	…
SDT	0.91995	0.92018	0.92018	0.92018
SDTQ	0.92024	0.92024	0.92024	0.92024
SDTQP	0.92024	0.92025	0.92025	0.92025
SDTQPH	0.92025	0.92025	0.92025	0.92025
		ω_e (cm^{-1})		
SD	4169.0	4166.0	4166.1	4166.3
SD(T)	4150.2	…	…	…
SDT	4148.5	4144.6	4144.7	4144.7
SDTQ	4143.7	4143.6	4143.6	4143.6
SDTQP	4143.6	4143.5	4143.5	4143.5
SDTQPH	4143.5	4143.5	4143.5	4143.5
		$\omega_e x_e$ (cm^{-1})		
SD	92.5	92.4	92.4	92.4
SD(T)	93.0	…	…	…
SDT	93.0	92.9	92.9	93.0
SDTQ	92.9	92.9	92.9	92.9
SDTQP	93.1	92.9	92.8	92.9
SDTQPH	92.9	92.9	92.9	92.9
		ϵ (mE_h)		
SD	2.423	2.017	2.023	2.028
SD(T)	0.496	…	…	…
SDT	0.406	0.068	0.072	0.075
SDTQ	0.011	0.002	0.002	0.002
SDTQP	0.001	−0.000	−0.000	−0.000
SDTQPH	−0.000	−0.000	−0.000	−0.000

TABLE IV. Equilibrium distance (r_e), harmonic vibrational frequency (ω_e), anharmonic constant ($\omega_e x_e$), and energy error with respect to FCI at the FCI geometry (ϵ) for the $X\,^1\Sigma^+$ state of C_2 computed with the TCC, UCC, ECC, VCC, and Mk-MRCC theories using the cc-pVDZ basis set and RHF orbitals. All Mk-MRCC computations used CASSCF(2,2) natural orbitals. The FCI r_e, ω_e, $\omega_e x_e$, and total energy at r_e using RHF orbitals are 1.27273 Å, 1812.9 cm^{-1}, 13.5 cm^{-1}, and $-75.729853\ E_h$, respectively. The FCI r_e, ω_e, $\omega_e x_e$, and total energy at r_e using CASSCF(2,2) orbitals are 1.27276 Å, 1812.7 cm^{-1}, 13.3 cm^{-1}, and $-75.729875\ E_h$, respectively.

Excitation level	TCC	ECC	UCC	VCC	Mk-MRCC[a]
			r_e (Å)		
SD	1.26616	1.26646	1.26640	1.26636	1.26073
SD(T)	1.27045	…	…	…	1.27153
SDT	1.27073	1.27104	1.27094	1.27085	1.27011
SDTQ	1.27221	1.27247	1.27245	1.27242	1.27239
SDTQP	1.27258	1.27268	1.27267	1.27267	…
SDTQPH	1.27272	1.27273	1.27273	1.27273	…
			ω_e (cm^{-1})		
SD	1861.5	1856.6	1856.7	1856.6	1889.7
SD(T)	1828.1	…	…	…	1828.4
SDT	1828.9	1825.0	1825.6	1826.1	1828.2
SDTQ	1815.8	1814.8	1814.9	1815.1	1812.7
SDTQP	1813.8	1813.3	1813.3	1813.4	…
SDTQPH	1812.9	1812.9	1812.9	1812.9	…
			$\omega_e x_e$ (cm^{-1})		
SD	12.3	12.6	12.7	12.7	12.4
SD(T)	12.7	…	…	…	12.4
SDT	12.8	13.2	13.2	13.2	13.8
SDTQ	13.5	13.5	13.5	13.5	13.4
SDTQP	13.5	13.5	13.5	13.5	…
SDTQPH	13.5	13.5	13.5	13.5	…
			ϵ (mE_h)		
SD	29.957	27.095	27.326	27.621	22.965
SD(T)	2.042	…	…	…	2.804
SDT	3.371	2.682	2.749	2.817	4.153
SDTQ	0.658	0.250	0.263	0.280	0.739
SDTQP	0.115	0.033	0.034	0.036	…
SDTQPH	0.007	0.002	0.002	0.002	…

[a]Mk-MRCC results are based on a CAS(2,2) active space composed of two electrons distributed in the $(2\sigma_u, 3\sigma_g)$ molecular orbitals.

and VCC methods are consistent among each other and are not substantially more accurate than those from the TCC theory.

Interestingly, when comparison is available, our results show that the Mk-MRCC approach does not offer any advantage in describing the equilibrium properties of C_2 with respect to the traditional and alternative single-reference coupled cluster approaches. In the case of ic-MRCCSD theory, r_e, ω_e, $\omega_e x_e$, ϵ are, respectively, 1.2616 Å, 1887.6 cm^{-1}, 12.2 cm^{-1}, and 16.0 mE_h, in slightly better agreement with the FCI results. When compared to the SRCC approaches, MRCC methods are found to underestimate r_e and overestimate ω_e. This effect is particularly substantial in the case of Mk-MRCCSD theory ($\Delta\omega_e = 28$ cm^{-1}, with respect to TCCSD) and it is likely to be a consequence of overestimating the weight of Φ_2 in the wave function, with the overall effect of reducing the population of the antibonding $2\sigma_u$ orbital in favor of the $3\sigma_g$ bonding orbital. A similar result can be found in the work of Li and Paldus.[93] These authors report r_e and ω_e for C_2 computed via the reduced-MRCC approach with singles and doubles (RMR-CCSD) that deviate considerably from the TCCSD results (for example, from Table II of Ref. [93], $\Delta\omega_e = 41$ cm^{-1} using a cc-pV5Z basis set).

The discrepancy between single-reference and multireference results and the absolute errors with respect to FCI are greatly mitigated by the inclusion of triple excitations. For example, perturbative triples already reduce the discrepancy between the harmonic frequency computed with TCCSD(T) and Mk-MRCCSD(T) theory to only 0.3 cm^{-1}, while the error with respect to FCI for ω_e computed with the Mk-MRCCSD(T) approach is 15.5 cm^{-1}. Similarly, the harmonic vibrational frequency computed with the RMR-CCSD method with perturbative triples, RMR-CCSD(T), differs from the CCSD(T) result by only 11 cm^{-1}. However, as shown by Li and Paldus, basis set effects are equally

important when comparing r_e and ω_e with experimental results as they amount to 0.025 Å and 30 cm^{-1}, respectively.[93]

These results suggest that it might be desirable to develop an equivalent perturbative correction for the alternative CCSD schemes. Kucharski and Bartlett[95] have suggested different routes for obtaining noniterative triples corrections to the CCSD energy, several of which are based on the extended and variational CC energy functionals and may be straightforwardly generalized to the alternative CCSD methods. Perturbative triples corrections for the alternative CC methods may also be obtained via approaches based on the lambda equations such as the Λ-CCSD(T) method of Kucharski and Bartlett[95] and the equivalent asymmetric-CCSD(T) correction of Crawford and Stanton.[96] In the case of ECCSD theory, Fan *et al.* have successfully developed noniterative triples corrections using a generalization of the method-of-moments approach.[73]

The problematic character of C_2 can be also appreciated from the slow convergence of the correlation energy with respect to the truncation of the excitation operator (reported in last column of Table IV). For all the single-reference coupled cluster approaches, the error in the correlation energy at the SD level of truncation is of the order of 27–30 mE_h, significantly higher than in the case of hydrogen fluoride (less than 3 mE_h). The deviation of the Mk-MRCC energy from the FCI result is slightly smaller only at the SD level ($\sim$23 mE_h), while at the Mk-MRCCSDT and Mk-MRCCSDTQ levels ϵ is twice as large as the single-reference approaches.

D. Potential energy curve of HF

To test the accuracy of alternative CC approaches when applied to multireference problems, we computed the potential energy curve of hydrogen fluoride.[97–99] At the equilibrium distance, the wave function of HF is dominated by the closed-shell configuration

$$|\Phi_1\rangle = |(1\sigma)^2(2\sigma)^2(3\sigma)^2(1\pi)^4\rangle, \qquad (34)$$

while at stretched geometries the two additional configurations generated by excitations from the 3σ to the 4σ MOs
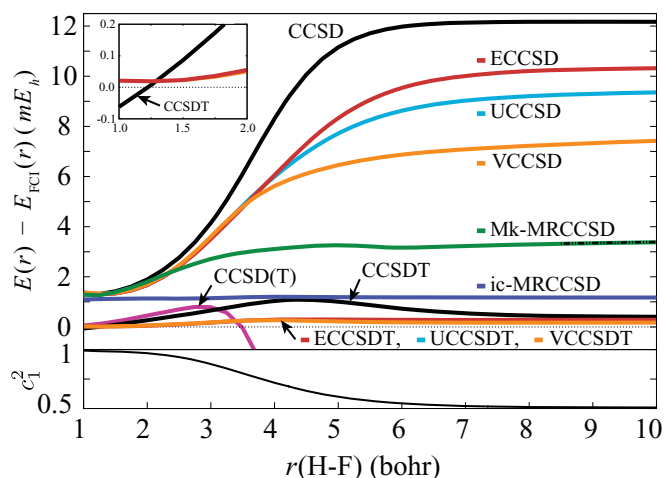

FIG. 2. Potential energy curve errors (with respect to FCI) for the $^1A_1$ ground state of HF using the DZV basis set and restricted Hartree-Fock orbitals. The lower panel shows the square of the CI coefficient for the reference determinant $|\Phi\rangle$ vs. the HF bond distance r(H-F) (in bohr) computed using a CAS(2,2) wave function. The inset shows the full triples energy in the range r(H-F) $\in$ [1, 2]. All electrons were correlated.

contribute to the zeroth-order wave function,

$$|\Phi_2\rangle = |(1\sigma)^2(2\sigma)^2(1\pi)^4(4\sigma)^2\rangle,$$

$$|\Phi_3\rangle = \frac{1}{\sqrt{2}}[|(1\sigma)^2(2\sigma)^2(1\pi)^4\,3\sigma_\alpha\,4\sigma_\beta\rangle \qquad (35)$$

$$- |(1\sigma)^2(2\sigma)^2(1\pi)^4\,3\sigma_\beta\,4\sigma_\alpha\rangle].$$

Figure 2 plots the energy error with respect to the FCI energy for a given method [$\Delta E(r)$] vs. the H-F bond distance r, where $\Delta E(r)$ is defined as

$$\Delta E(r) = E(r) - E_{\text{FCI}}(r). \qquad (36)$$

To quantify the deviation from the FCI potential energy curve, Table V reports the nonparallelism error (NPE), defined over the range of geometries $R = [1, 10]$ bohr,

$$\text{NPE} = \max_{r \in R}[\Delta E(r)] - \min_{r \in R}[\Delta E(r)]. \qquad (37)$$

At the singles and doubles level, alternative CC approaches are more accurate than traditional CC theory (NPE 10.9 mE_h) along the entire potential energy curve, in particular, in the case of variational CC (NPE 6.1 mE_h).

TABLE V. Nonparallelism error in mE_h for the potential energy curve of HF and the BeH$_2$ model.

Excitation level	TCC	ECC	UCC	VCC	Mk-MRCC	ic-MRCC
				HF $\rightarrow$ H + F		
SD (Φ_1)	10.88	9.01	8.04	6.10	2.11[a]	0.11[a]
SDT (Φ_1)	1.15	0.28	0.27	0.25	...	...
				BeH$_2$ $\rightarrow$ Be + H$_2$		
SD (Φ_1)	10.61	21.19	21.55	21.32[b]	2.24[a]	1.04[a]
SD (Φ_2)	19.59[b]	9.34[b]	18.25	7.26[b]	...	...
SDT (Φ_1)	18.90	2.74	3.02	3.36[b]	0.20[a]	...
SDT (Φ_2)	20.61[b]	5.43[b]	1.49	1.34[b]	...	...

[a] All MRCC results are based on CASSCF(2,2) natural orbitals and a CAS(2,2) model space.
[b] Because of convergence problems, the result is from a smaller subset of the original range.

However, when the NPE for the alternative CC approaches are compared to those for the genuine MRCC methods (NPE 2.11 and 0.11 mE_h for the Mk- and ic-MRCC approaches, respectively) it is evident that the latter are better suited to describe the potential energy curve of HF. Interestingly, near the equilibrium geometry the accuracy of the multiference methods and the SRCC approaches is comparable. In general, MRCC approaches tend to underperform in the single-reference regime, as it can be appreciated in the C_2 example (Sec. V C) and other results in the literature.[100]

The accuracy of single-reference coupled cluster approaches is greatly increased with the addition of triple excitations, which can be seen in the drastic decrease in nonparallelism error for the ECC, UCC, and VCC theories to less than 0.3 mE_h (from a max of 9.01 mE_h at the ECCSD level). On the contrary, for traditional CCSDT the NEP is slightly higher than the alternative methods, 1.15 mE_h. The TCCSD(T) scheme rapidly diverges to the lower part of the plots and reaches an error of ~ -50 mE_h (not shows in the plot) as the molecule is dissociated. A similar, but less pronounced deviation from the FCI energy is also found for the Λ-CCSD(T) approach (~ -8 mE_h, using the 6–31G* basis set),[101] while more robust approaches such as the CCSD(2)$_T$ and the completely renormalized CCSD method [CR-CC(2,3)](Refs. 102 and 103) deviate from FCI by less than 2 mE_h (using the DZ basis set).

The reader might also observe that in the entire range sampled, the energy of the alternative SRCC approaches remains above the variational limit, while the traditional CCSDT energy falls below the FCI value for bond distances close to 1 bohr (see inset in Fig. 2). While one expects the VCC and UCC energies to be strictly variational at all levels of truncation, it is rather intriguing that the ECC theory also numerically satisfies this condition because it is not guaranteed to yield an energy that is an upper bound to the FCI energy.

E. Potential energy curve of BeH$_2$

The second multireference problem that we consider is the insertion of a beryllium atom into H$_2$ to form BeH$_2$. Purvis *et al.*[104] introduced a model that samples this insertion reaction along a C_{2v} pathway at a fixed number of points. This model has been used extensively to benchmark new multiference methods.[24, 28, 54, 105–109]

For this study, we adopt a variant[110] of the original model that consists in a beryllium atom placed at the center of a two-dimensional coordinate system and two hydrogen atoms that lie on the lines described by the equation

$$y(x) = \pm(2.54 - 0.46x), \qquad (38)$$

where x is a continuous parameter that lies in the range 0–4 bohr (see the supplementary material[85] for an illustration of the BeH$_2$ model). Along the C_{2v} insertion pathway defined by Eq. (38), two electronic configurations predominate the zeroth-order wave function

$$\begin{aligned} |\Phi_1\rangle &= |(1a_1)^2(2a_1)^2(1b_2)^2\rangle, \\ |\Phi_2\rangle &= |(1a_1)^2(2a_1)^2(3a_1)^2\rangle, \end{aligned} \qquad (39)$$
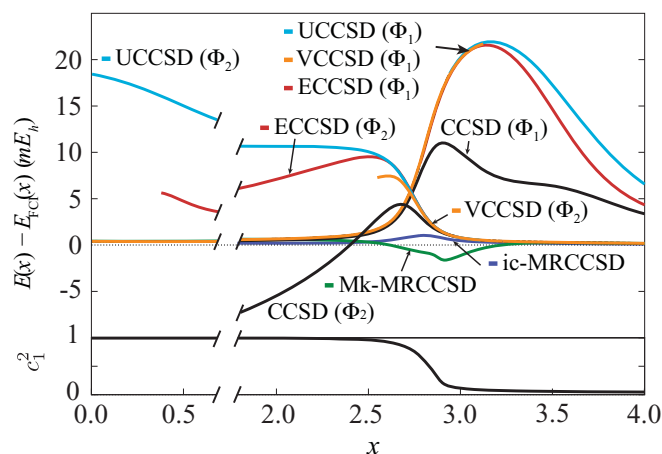

FIG. 3. Coupled cluster theory with singles and doubles potential energy curve errors (in mE_h with respect to FCI) for the $^1A_1$ ground state of the BeH$_2$ model system using a Be(3s2p)/H(2s) basis set and correlating all electrons. For all SRCC computations we also report the reference determinant used to obtain the Hartree-Fock orbitals. Multireference coupled cluster results were obtained using a CAS(2,2) model space and CASSCF(2,2) molecular orbitals. The lower panel shows the square of the CI coefficient for the reference determinant $|\Phi_1\rangle$ vs. the parameter x (in bohr) computed using a CAS(2,2) wave function.

where Φ_1 and Φ_2 are, respectively, the leading contributions for molecular BeH$_2$ ($x = 0$) and Be + H$_2$ ($x = 4$). In the region $2.5 < x < 3$, the zeroth-order wave function switches from Φ_1 to Φ_2, with a maximum multireference character ($c_1^2 = c_2^2 = 1/2$) when x is approximately equal to $x_{MR} = 2.825$.

Figures 3 and 4 depict $\Delta E(r)$ for the BeH$_2$ model computed using the SRCC and MRCC methods, while Table V reports the respective nonparallelism errors. Because the zeroth-order wave function switches from Φ_1 to Φ_2, for all
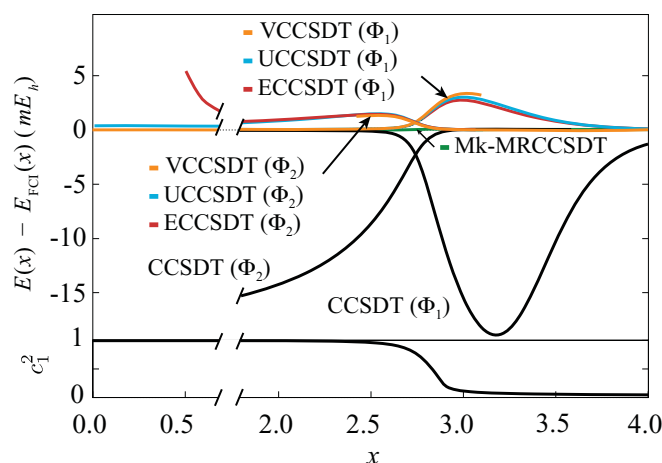

FIG. 4. Coupled cluster theory with singles, doubles, and triples potential energy curve errors (in mE_h with respect to FCI) for the $^1A_1$ ground state of the BeH$_2$ model system using a Be(3s2p)/H(2s) basis set and correlating all electrons. For all SRCC computations we also report the reference determinant used to obtain the Hartree-Fock orbitals. Multireference coupled cluster results were obtained using a CAS(2,2) model space and CASSCF(2,2) molecular orbitals. The lower panel shows the square of the CI coefficient for the reference determinant $|\Phi_1\rangle$ vs. the parameter x (in bohr) computed using a CAS(2,2) wave function.

the SRCC methods we have computed curves using Hartree-Fock orbitals based on both reference functions. The drastic change in the zeroth-order wave function that occurs around x_{MR} makes the BeH_2 model system a particularly challenging problem for all the single-reference methods. Indeed, Fig. 3 shows that both the traditional and the alternative SRCC approaches break down near the multireference region. The quality of all potential energy curves computed using Φ_1 as a reference function diminishes when $x > x_{MR}$, as in this region Φ_2 is the leading contribution to the wave function. Similarly, all SRCC computations based on the Φ_2 reference determinant start to deviate considerably from the FCI result when $x < x_{MR}$.

When comparison among the alternative SRCC methods is possible, one observes little deviation in the performance when using Φ_1 as a reference determinant, while the relative performance among ECC, VCC, and UCC differs considerably when Φ_2 is chosen as a reference determinant. Interestingly, the UCC theory is the most robust and can be converged along the entire potential energy curve. On the contrary, the TCC, ECC and VCC methods could not always be converged even when recurring to DIIS and shifting the energy denominators. The MRCCSD approaches are superior in accuracy to all the SRCCSD methods and do not suffer from convergence problems along the entire potential energy curve. Gauging the relative accuracy with the NPE, the MRCCSD results are found to be at least three times more accurate than all the SRCCSD approaches, for example, the NPE for Mk- and ic-MRCCSD is 2.4, 1.0 mE_h, respectively, vs. 7.3 mE_h for VCCSD(Φ_2).

Figure 4 reports the energy error computed with the singles, doubles, and triples approximation. At this level, traditional coupled cluster theory still shows large deviations from the FCI energy (NPE ~ 20 mE_h) while the UCCSDT approach deviates from FCI by less than 4 mE_h when using either Φ_1 or Φ_2 as a reference determinant. The ECCSDT equations can be converged on the entire potential energy curve only when using Φ_1 as a reference, while VCCSDT fails to converge with both references Φ_1 and Φ_2. The Mk-MRCCSDT approach yields the lowest NPE

(0.2 mE_h) and its accuracy is unmatched by any of the SR-CCSDT schemes.

The BeH_2 model provides the opportunity to investigate some of the features of the alternative SRCC approaches. The first aspect that we consider is the dynamical projection manifold in the ECC theory. Figure 5 illustrates the reference state $\langle\Phi|\exp(\hat{\Sigma})$, upon which one projects the ECC sigma amplitude equations, for the BeH_2 model system using a minimal basis set and a model space composed of $|\Phi_1\rangle$ and $|\Phi_2\rangle$. The panels depict the magnitude of the vector $\langle\Phi|\exp(\hat{\Sigma})$, represented in the basis of alpha and beta string $(\mathcal{P}_\alpha, \mathcal{P}_\beta)$ as a matrix $C_{\mathcal{P}_\alpha,\mathcal{P}_\beta} = \langle\Phi|\exp(\hat{\Sigma})|\mathcal{P}_\alpha, \mathcal{P}_\beta\rangle$. In Fig. 5, we plot $C_{\mathcal{P}_\alpha,\mathcal{P}_\beta}$ as a tiled matrix with each tile representing a determinant and the size and color proportional to $|C_{\mathcal{P}_\alpha,\mathcal{P}_\beta}|$. The top panels depict $\langle\Phi|\exp(\hat{\Sigma})$ for computations that use Φ_1 as a reference. In this case, when $x = 2.0$ and Φ_1 is the leading contribution to Ψ, the projection manifold is also dominated by excitations almost exclusively out of Φ_1 (large red tile in the top left panel). As we reach the point of maximum multireference character ($x = 2.825$), the projection manifold has a large contribution from Φ_2 (large blue tile in the top center panel), which is almost comparable in magnitude to Φ_1. Interestingly, even when the exact wave function is no longer dominated by Φ_1 ($x = 4.0$), the projection manifold still has a large contribution from Φ_1. The same conclusions can be drawn from the lower panels, which show the results computed using Φ_2 as a reference.

A second aspect that we consider is the convergence of the energy in approximate versions of the UCC theory with a truncated Baker-Campbell-Hausdorff expansion of $\bar{H}$. Bartlett $et\ al.$[60] have considered a hierarchy of finite-order truncations of the variational unitary CC energy functional based on a perturbative analysis. Here, we explore a different scheme in which the infinite BCH expansion of $\bar{H}$ is truncated to include up to a given number of commutators n_{tr},

$$e^{-\hat{A}}\hat{H}e^{\hat{A}} \approx \hat{H} + \sum_{k=1}^{n_{tr}} \frac{1}{k!} \underbrace{[[\hat{H}, \hat{A}], \ldots, \hat{A}]}_{k-\text{nested commutators}}. \tag{40}$$

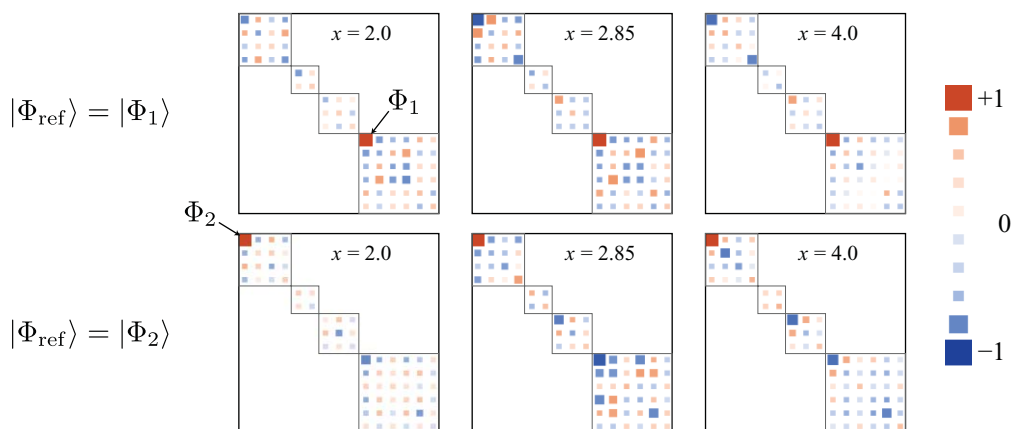FIG. 5. Plots of the vector $\exp(\Sigma^\dagger)|\Phi\rangle$ for the BeH_2 model system using a minimal basis set. The expansion coefficients $C_{\mathcal{I}_\alpha\mathcal{I}_\beta} = \langle\Phi_{\mathcal{I}_\alpha\mathcal{I}_\beta}|\exp(\Sigma^\dagger)|\Phi\rangle$ (corresponding to the determinant $|\Phi_{\mathcal{I}_\alpha\mathcal{I}_\beta}\rangle$ formed by the pair of strings $\mathcal{I}_\alpha$ and $\mathcal{I}_\beta$) are depicted with tiles using a color scale where red and blue indicate positive and negative values, respectively. The magnitude of each matrix element is proportional to the size and color intensity of the corresponding tile.

TABLE VI. Approximate UCCSD approach using a Baker-Campbell-Hausdorff expansion of $\bar{H}$ truncated after n_{tr} commutators. Total energy (E) and error (ϵ) with respect to untruncated UCCSD (in E_h).

n_{tr}	$x = 2$, Φ_1		$x = x_{MR}$, Φ_1		$x = x_{MR}$, Φ_2	
	E	ϵ	E	ϵ	E	ϵ
1	-15.763684331	-6.9×10^{-2}	...	...	-15.702012862	-7.9×10^{-2}
2	-15.695942575	-1.3×10^{-3}	...	...	-15.637941541	-1.5×10^{-2}
3	-15.694224787	4.1×10^{-4}	-15.611576165	2.6×10^{-3}	-15.618342604	5.0×10^{-3}
4	-15.694617317	1.5×10^{-5}	-15.609739823	4.5×10^{-3}	-15.620738227	2.6×10^{-3}
5	-15.694633891	-1.7×10^{-6}	-15.613712839	5.0×10^{-4}	-15.623625571	-2.6×10^{-4}
6	-15.694632255	-9.9×10^{-8}	-15.614495257	-2.8×10^{-4}	-15.623458183	-9.1×10^{-5}
7	-15.694632150	5.0×10^{-9}	-15.614255675	-3.9×10^{-5}	-15.623360835	6.6×10^{-6}
8	-15.694632155	4.7×10^{-10}	-15.614207294	9.4×10^{-6}	-15.623365363	2.1×10^{-6}
9	-15.694632156	-1.2×10^{-11}	-15.614214858	1.8×10^{-6}	-15.623367567	-1.1×10^{-7}
10	-15.694632155	-2.0×10^{-12}	-15.614216883	-2.1×10^{-7}	-15.623367485	-3.2×10^{-8}
11	-15.694632155	$< 10^{-12}$	-15.614216721	-5.1×10^{-8}	-15.623367451	1.4×10^{-9}
12	-15.694632155	$< 10^{-12}$	-15.614216667	3.5×10^{-9}	-15.623367452	3.5×10^{-10}
13	-15.694632155	$< 10^{-12}$	-15.614216669	9.9×10^{-10}	-15.623367453	-1.4×10^{-11}
14	-15.694632155	$< 10^{-12}$	-15.614216670	-4.3×10^{-11}	-15.623367453	-4.0×10^{-12}
15	-15.694632155	$< 10^{-12}$	-15.614216670	-1.5×10^{-11}	-15.623367453	$< 10^{-12}$

Table VI reports the UCCSD energy computed using the BCH expansion truncation in the residual and the energy expressions, for two geometrical configurations of the BeH$_2$ model. When $x = 2$, the UCCSD energy quickly converges to the untruncated value and each additional commutator reduces the error by roughly one order of magnitude. By the time five commutators are included, the error with respect to the untruncated UCCSD energy is of the order of 1 μE_h. At the maximum multireference point ($x = 2.825$), the convergence of the UCCSD energy is much slower than the one observed when $x = 2$. In Table VI we present results using both Φ_1 and Φ_2 as reference determinants. Depending on the choice of the reference determinant, eight to nine commutators are required to achieve an error of $\sim$1 μE_h.

Let us take a closer look at the convergence of UCC theory along the potential energy curve errors of the BeH$_2$ model. Figure 6 depicts the energy error of the truncated UCCSD schemes ($n_{tr} = 2, \ldots, 8$) with respect to FCI, using Φ_1 as reference determinant. For $x \ll x_{MR}$, all the approximate schemes yield similar results, with the exception of the case $n_{tr} = 2$. Near the point of maximum multireference character, however, most approximate schemes have convergence problems. Only the schemes with $n_{tr} = 7, 8$ converge along the entire range, but surprisingly they struggle to approach the untruncated UCCSD energy when $x > x_{MR}$. Clearly, what is happening here is a consequence of large cluster amplitudes, which cause the BCH expansion of the UCC similarity-transformed Hamiltonian to converge very slowly. Indeed, an analysis of the cluster amplitudes a_{ij}^{ab}, reveals that the amplitude with the largest magnitude has value $\sim$$-0.715$ (corresponding to the $\Phi_1 \rightarrow \Phi_2$ excitation), followed by amplitudes with values 0.214 and -0.128.

VI. SUMMARY

We report a new arbitrary-order string-based implementation of the extended (ECC), variational (VCC), and projective unitary (UCC) coupled cluster theories. These al-

ternative single-reference methods were used to compute the total energy of Ne, the equilibrium properties of HF and C$_2$ and the potential energy curves of HF and the BeH$_2$ model. Results were benchmarked against FCI and TCC theory, the state-specific MRCC method of Mukherjee et al. (Mk-MRCC),[28] and the internally contracted MRCC approach (ic-MRCC).[36]

The equilibrium properties of HF and C$_2$ in Tables III and IV reveal that alternative coupled cluster approaches are only slightly more accurate than the traditional CC theory. Specifically, the results from alternative CC approaches truncated at n-fold excitations are always inferior to the results from traditional CC computations truncated at order $n + 1$. Even in the case of a multireference molecule such as C$_2$, the bond length, harmonic vibrational frequency, and the anharmonic constant obtained from alternative CC approaches do not improve considerably with respect to the TCC results. For example, r_e
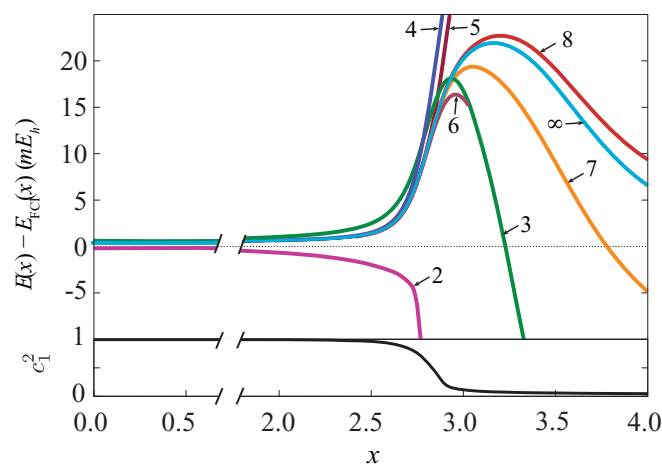

FIG. 6. Convergence of the approximate unitary CCSD energy vs. the number of commutators (n_{tr}) included in the Baker-Campbell-Hasdorff expansion. Potential energy curve errors (in $m E_h$ with respect to FCI) for the $^1A_1$ ground state of the BeH$_2$ model system using a Be(3s2p)/H(2s) basis set and Φ_1 as reference determinant.

and ω_e for ECCSD and TCCSD are, respectively, (1.2665 Å, 1857 cm^{-1}) and (1.2662 Å, 1862 cm^{-1}). Contrast this with the TCCSDT and FCI values, (1.2707 Å, 1829 cm^{-1}) and (1.2727 Å, 1813 cm^{-1}), respectively. It is somewhat surprising that in the case of C_2, the Mk-MRCCSD approach gives results that are slightly worse than traditional coupled cluster theory (1.2607 Å, 1890 cm^{-1}).

Several problems arise when alternative SRCC methods are used to compute potential energy curves. When the deviation from the FCI curve is quantified by the NPE reported in Table V, the alternative CC approaches in general outperform traditional CC theory, but yield results that are significantly less accurate than the MRCC methods. In the case of hydrogen fluoride, the NPE values for TCCSD, VCCSD, and Mk-MRCCSD are 10.9, 6.1, and 2.1 mE_h, respectively. In the more challenging case of the BeH_2 model, once again the superiority of the MRCC approaches is demonstrated, while problems with the convergence of traditional, extended, and variational coupled cluster theories are observed. Only the unitary CC approach is robust enough to converge on the entire range of the potential energy curve. Furthermore, the BeH_2 model exposes the strong dependence of the alternative CC methods on the choice of the Fermi vacuum: the energy quickly deteriorates when the reference determinant is a poor representation of the zeroth-order wave function.

Unitary coupled cluster theory stands out for its robustness when applied to pronounced multireference problems. However, since the UCC equations do not truncate at a finite number of terms, a practical approach requires some form of approximation. The BeH_2 model results in Fig. 6 demonstrate that the UCC commutator series converges slowly when some of the cluster amplitudes become sizable and that one must be careful when truncating the Baker-Campbell-Hausdorff expansion of the UCC equations. It is unclear whether other approximations of the unitary similarity-transformed Hamiltonian advanced in the literature[76,77] suffer from the same problems. Our unitary CC results also speak to the issue of what is the best cluster parameterization for multireference CC approaches. Unitary MRCC theories have been advocated by several authors because this choice leads to an Hermitian effective Hamiltonian.[76,111] In light of our results, the increase in complexity that results from the unitary cluster ansatz does not seem to justify this choice.

In conclusion, we find that in the case of molecular systems, the extended, variational, and projective unitary coupled cluster approaches offer only a marginal improvement in accuracy with respect to the traditional formulation of coupled cluster theory, an advantage that is rendered irrelevant by the additional computational costs of the alternative SRCC methods. While the general conclusions drawn here may be viewed as disappointing, we view its consequences as positive: the simplest theory (traditional CC theory) is also the most effective.

[1] R. J. Bartlett and G. D. Purvis III, Int. J. Quantum Chem. **14**, 561 (1978).
[2] R. J. Bartlett, Annu. Rev. Phys. Chem. **32**, 359 (1981).
[3] M. Nooijen, K. R. Shamasundar, and D. Mukherjee, Mol. Phys. **103**, 2277 (2005).
[4] J. A. Pople, J. S. Binkley, and R. Seeger, Int. J. Quantum Chem., Quantum Chem. Symp. **10**, 1 (1976).
[5] F. Coester, Nucl. Phys. **7**, 421 (1958).
[6] F. Coester and H. Kümmel, Nucl. Phys. **17**, 477 (1960).
[7] J. Čížek, J. Chem. Phys. **45**, 4256 (1966).
[8] The acronym TCC should not be confused with the one used by Kinoshita *et al.* for their tailored CC method (Ref. 45).
[9] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).
[10] I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory* (Cambridge University Press, Cambridge, England, 2009).
[11] G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).
[12] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
[13] M. S. Schuurman, S. R. Muir, W. D. Allen, and H. F. Schaefer, J. Chem. Phys. **120**, 11586 (2004).
[14] A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, and J. F. Stanton, J. Chem. Phys. **121**, 11599 (2004).
[15] W. Klopper, F. R. Manby, S. Ten-no, and E. F. Valeev, Int. Rev. Phys. Chem. **25**, 427 (2006).
[16] M. Kállay and P. R. Surján, J. Chem. Phys. **115**, 2945 (2001).
[17] M. Nooijen, Phys. Rev. Lett. **84**, 2108 (2000).
[18] T. van Voorhis and M. Head-Gordon, J. Chem. Phys. **115**, 5033 (2001).
[19] P.-D. Fan and P. Piecuch, Adv. Quantum Chem. **51**, 1 (2006).
[20] D. I. Lyakh and R. J. Bartlett, J. Chem. Phys. **133**, 244112 (2010).
[21] A. J. W. Thom, Phys. Rev. Lett. **105**, 263004 (2010).
[22] D. Mukherjee, R. Moitra, and A. Mukhopadhyay, Mol. Phys. **33**, 955 (1977).
[23] B. Jeziorski and H. J. Monkhorst, Phys. Rev. A **24**, 1668 (1981).
[24] U. S. Mahapatra, B. Datta, B. Bandyopadhyay, and D. Mukherjee, Adv. Quantum Chem. **30**, 163 (1998).
[25] J. Paldus, P. Piecuch, L. Pylypow, and B. Jeziorski, Phys. Rev. A **47**, 2738 (1993).
[26] K. Kowalski and P. Piecuch, Phys. Rev. A **61**, 052506 (2000).
[27] U. S. Mahapatra, B. Datta, and D. Mukherjee, Mol. Phys. **94**, 157 (1998).
[28] U. S. Mahapatra, B. Datta, and D. Mukherjee, J. Chem. Phys. **110**, 6171 (1999).
[29] D. Datta and D. Mukherjee, J. Chem. Phys. **134**, 054122 (2011).
[30] J. Mášik and I. Hubač, Adv. Quantum Chem. **31**, 75 (1999).
[31] J. Pittner, P. Nachtigall, P. Čársky, J. Mášik, and I. Hubač, J. Chem. Phys. **110**, 10275 (1999).
[32] J. Pittner, J. Chem. Phys. **118**, 10876 (2003).
[33] M. Hanrath, J. Chem. Phys. **123**, 084102 (2005).
[34] M. Hanrath, J. Chem. Phys. **128**, 154118 (2008).
[35] S. Li, J. Chem. Phys. **120**, 5017 (2004).
[36] F. A. Evangelista and J. Gauss, J. Chem. Phys. **134**, 114102 (2011).
[37] X. Li and J. Paldus, J. Chem. Phys. **119**, 5320 (2003).
[38] X. Li and J. Paldus, Chem. Phys. Lett. **496**, 183 (2010).
[39] F. A. Evangelista, W. D. Allen, and H. F. Schaefer, J. Chem. Phys. **127**, 024102 (2007).
[40] X. Li and J. Paldus, J. Chem. Phys. **132**, 114103 (2010).
[41] T. Saito, N. Yasuda, S. Nishihara, S. Yamanaka, Y. Kitagawa, T. Kawakami, M. Okumura, and K. Yamaguchi, Chem. Phys. Lett. **505**, 11 (2011).
[42] N. Oliphant and L. Adamowicz, J. Chem. Phys. **94**, 1229 (1991).
[43] P. Piecuch, N. Oliphant, and L. Adamowicz, J. Chem. Phys. **99**, 1875 (1993).
[44] X. Li and J. Paldus, J. Chem. Phys. **107**, 6257 (1997).
[45] T. Kinoshita, O. Hino, and R. J. Bartlett, J. Chem. Phys. **123**, 074106 (2005).
[46] V. A. Rassolov and F. Xu, J. Chem. Phys. **127**, 044104 (2007).
[47] D. W. Small and M. Head-Gordon, J. Chem. Phys. **130**, 084103 (2009).
[48] J. A. Parkhill and M. Head-Gordon, J. Chem. Phys. **133**, 124102 (2010).
[49] P. Piecuch, K. Kowalski, I. S.O. Pimienta, and M. J. McGuire, Int. Rev. Phys. Chem. **21**, 527 (2002).

[50] M. Nooijen and R. J. Bartlett, J. Chem. Phys. **107**, 6812 (1997).

[51] A. I. Krylov, Annu. Rev. Phys. Chem. **59**, 433 (2008).

[52] L. Kong, K. R. Shamasundar, O. Demel, and M. Nooijen, J. Chem. Phys. **130**, 114101 (2009).

[53] M. Musiał, A. Perera, and R. J. Bartlett, J. Chem. Phys. **134**, 114108 (2011).

[54] M. Kállay, P. G. Szalay, and P. R. Surján, J. Chem. Phys. **117**, 980 (2002).

[55] P. G. Szalay, M. Nooijen, and R. J. Bartlett, J. Chem. Phys. **103**, 281 (1995).

[56] R. J. Bartlett and J. Noga, Chem. Phys. Lett. **150**, 29 (1988).

[57] N. Vaval and S. Pal, Phys. Rev. A **54**, 250 (1996).

[58] T. van Voorhis and M. Head-Gordon, J. Chem. Phys. **113**, 8873 (2000).

[59] W. Kutzelnigg, J. Chem. Phys. **77**, 3081 (1982).

[60] R. J. Bartlett, S. A. Kucharski, and J. Noga, Chem. Phys. Lett. **155**, 133 (1989).

[61] J. D. Watts, G. W. Trucks, and R. J. Bartlett, Chem. Phys. Lett. **157**, 359 (1989).

[62] W. Kutzelnigg, Theor. Chim. Acta **80**, 349 (1991).

[63] J. S. Arponen, Ann. Phys. **151**, 311 (1983).

[64] J. S. Arponen, R. F. Bishop, and E. Pajanne, Phys. Rev. A **36**, 2519 (1987).

[65] R. F. Bishop and J. S. Arponen, Int. J. Quantum Chem., Quantum Chem. Symp. **38**, 197 (1990).

[66] A. B. Kumar, N. Vaval, and S. Pal, Chem. Phys. Lett. **295**, 189 (1998).

[67] P. Piecuch and R. J. Bartlett, Adv. Quantum Chem. **34**, 295 (1999).

[68] P. Piecuch, K. Kowalski, P.-D. Fan, and K. Jedziniak, Phys. Rev. Lett. **90**, 113001 (2003).

[69] E. R. Davidson, Phys. Rev. Lett. **91**, 123001 (2003).

[70] D. A. Mazziotti, Phys. Rev. A **69**, 012507 (2004).

[71] W. Kutzelnigg and D. Mukherjee, Phys. Rev. A **71**, 022502 (2005).

[72] W. Kutzelnigg, Mol. Phys. **94**, 65 (1998).

[73] P.-D. Fan, K. Kowalski, and P. Piecuch, Mol. Phys. **103**, 2191 (2005).

[74] B. Cooper and P. J. Knowles, J. Chem. Phys. **133**, 234102 (2010).

[75] Alternatively, one can avoid the premultiplication of the Schrödinger equation by $\exp(-\hat{S})$. When the cluster operator is truncated, in the case of traditional CC theory the premultiplication does not have any effect, however, in unitary CC theory the premultiplication leads to a different result with respect to the original Schrödinger equation.

[76] T. Yanai and G. K.-L. Chan, J. Chem. Phys. **124**, 194106 (2006).

[77] A. G. Taube and R. J. Bartlett, Int. J. Quantum Chem. **106**, 3393 (2006).

[78] R. M. Wilcox, J. Math. Phys. **8**, 962 (1967).

[79] A more elegant formulation of extended coupled cluster theory uses a modified energy functional $\tilde{E}_{ECC'} = \langle \Phi | (e^{\hat{\Sigma}} \bar{H} e^{-\hat{\Sigma}})_{dc} | \Phi \rangle$, where the symbol dc indicates that only *doubly connected* terms are included, consisting in all the diagrams with $\hat{\Sigma}$ connected to either $\hat{H}$ or at least two cluster operators. Double connectedness guarantees that both the energy functional and the equations resulting from the stationary condition contain only connected terms.

[80] T. D. Crawford, C. D. Sherrill, E. F. Valeev, J. T. Fermann, R. A. King, M. L. Leininger, S. T. Brown, C. L. Janssen, E. T. Seidl, J. P. Kenny, and W. D. Allen, J. Comput. Chem. **28**, 1610 (2007).

[81] G. L. Bendazzoli and S. Evangelisti, J. Chem. Phys. **98**, 3141 (1993).

[82] S. Hirata and R. J. Bartlett, Chem. Phys. Lett. **321**, 216 (2000).

[83] We assume the residual to be a real number, in the general case it is necessary to take the real part of the VCC theory residual ($r_q = \text{Re}\langle \Phi_q | r \rangle$).

[84] T. P. Hamilton and P. Pulay, J. Chem. Phys. **84**, 5728 (1986).

[85] See supplementary material at http://dx.doi.org/10.1063/1.3598471 for a pdf file containing the definition of the BeH_2 model, gaussian basis sets, code benchmarks, and the energies for the BeH_2 model and the potential energy curves of HF.

[86] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).

[87] F. A. Evangelista, E. Prochnow, J. Gauss, and H. F. Schaefer, J. Chem. Phys. **132**, 074107 (2010).

[88] F. A. Evangelista and A. C. Simmonett, PSIMRCC, a computer code written at the University of Georgia and capable of performing multireference coupled cluster computations, 2007; see http://www.ccc.uga.edu/psimrcc.

[89] T. D. Crawford and H. F. Schaefer, J. Chem. Phys. **104**, 6259 (1996).

[90] C. W. Bauschlicher, Jr., and S. R. Langhoff, J. Chem. Phys. **87**, 2919 (1987).

[91] M. L. Abrams and C. D. Sherrill, J. Chem. Phys. **121**, 9211 (2004).

[92] C. D. Sherrill and P. Piecuch, J. Chem. Phys. **122**, 124104 (2005).

[93] X. Li and J. Paldus, Chem. Phys. Lett. **431**, 179 (2006).

[94] K. P. Huber and G. Herzberg, "Constants of diatomic molecules," (data prepared by J. W. Gallagher and R. D. Johnson, III) in *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, Eds. P. J. Linstrom and W. G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899, http://webbook.nist.gov (retrieved June 1, 2011).

[95] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. **108**, 5243 (1998).

[96] T. D. Crawford and J. F. Stanton, Int. J. Quantum Chem. **70**, 601 (1998).

[97] V. V. Ivanov, L. Adamowicz, and D. I. Lyakh, J. Mol. Struct. **768**, 97 (2006).

[98] A. Engels-Putzka and M. Hanrath, J. Mol. Struct. **902**, 59 (2009).

[99] S. Das, D. Mukherjee, and M. Kállay, J. Chem. Phys. **132**, 074103 (2010).

[100] F. A. Evangelista, W. D. Allen, and H. F. Schaefer, J. Chem. Phys. **125**, 154113 (2006).

[101] A. G. Taube and R. J. Bartlett, J. Chem. Phys. **128**, 044110 (2008).

[102] S. R. Gwaltney and M. Head-Gordon, J. Chem. Phys. **115**, 2014 (2001).

[103] P. Piecuch and M. Włoch, J. Chem. Phys. **123**, 224105 (2005).

[104] G. D. Purvis III, R. Shepard, F. B. Brown, and R. J. Bartlett, Int. J. Quantum Chem. **23**, 835 (1983).

[105] R. J. Gdanitz and R. Ahlrichs, Chem. Phys. Lett. **143**, 413 (1988).

[106] S. B. Sharp and G. I. Gellene, J. Phys. Chem. A **104**, 10951 (2000).

[107] P. J. A. Ruttink, J. H. van Lenthe, and P. Todorov, Mol. Phys. **103**, 2497 (2005).

[108] J. Pittner, H. V. Gonzalez, R. J. Gdanitz, and P. Čársky, Chem. Phys. Lett. **386**, 211 (2004).

[109] D. I. Lyakh, V. V. Ivanov, and L. Adamowicz, Theor. Chem. Acc. **116**, 427 (2006).

[110] Here, we adopt a variant of the model that interpolates the path from points A to H of the original model by Purvis *et al.* (Ref. 104). Moreover, we use a Be(10s3p/3s2p), H(4s/2s) basis set (included in the Supplementary Material) which differs slightly from the original double-ζ basis set used in previous studies. In our basis set the tight beryllium p primitive function was decontracted to better describe core correlation of the Be atom.

[111] M. R. Hoffmann and J. Simons, J. Chem. Phys. **88**, 993 (1988).