Canonical transformation theory for multireference problems

Takeshi Yanai and Garnet Kin-Lic Chan

Citation: J. Chem. Phys. **124**, 194106 (2006); doi: 10.1063/1.2196410
View online: http://dx.doi.org/10.1063/1.2196410
View Table of Contents: http://aip.scitation.org/toc/jcp/124/19
Published by the American Institute of Physics

Canonical transformation theory for multireference problems

Takeshi Yanai[a] and Garnet Kin-Lic Chan
Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301

(Received 27 January 2006; accepted 22 March 2006; published online 16 May 2006)

We propose a theory to describe dynamic correlations in bonding situations where there is also significant nondynamic character. We call this the canonical transformation (CT) theory. When combined with a suitable description of nondynamic correlation, such as given by a complete-active-space self-consistent Field (CASSCF) or density matrix renormalization group wave function, it provides a theory to describe bonding situations across the entire potential energy surface with quantitative accuracy for both dynamic and nondynamic correlation. The canonical transformation theory uses a unitary exponential ansatz, is size consistent, and has a computational cost of the same order as a single-reference coupled cluster theory with the same level of excitations. Calculations using the CASSCF based CT method with single and double operators for the potential energy curves for water and nitrogen molecules, the BeH_2 insertion reaction, and hydrogen fluoride and boron hydride bond breaking, consistently yield quantitative accuracies typical of equilibrium region coupled cluster theory, but across all geometries, and better than obtained with multireference perturbation theory. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2196410]

I. INTRODUCTION

In the chemical bond, the electronic correlation can be divided into two types: the nondynamic correlation, associated with the overlap of near-degenerate atomic states, and the dynamic correlation, associated with the scattering of electrons at short distances. We know that nondynamic correlation is well described by a multireference theory in an active space. Examples of multireference theories include the complete-active-space self-consistent field[1,2] (CASSCF) and density matrix renormalization group (DMRG) methods.[3–12] An exact treatment of correlation in the active space yields potential energy surfaces that are qualitatively, but not quantitatively, accurate across the entire surface, due to the neglect of dynamic correlation. Dynamic correlation is well described by high-order perturbative approaches, and in particular coupled cluster (CC) theory.[13–16] Coupled cluster theory in its usual form yields quantitatively accurate results near equilibrium, but not at stretched geometries, due to the neglect of nondynamic correlation. Multireference perturbation theories[17–20] present an approach to include both nondynamic and dynamic correlations, but they do not attain quantitative "chemical" accuracy—say equal to that of the coupled cluster theory in the equilibrium region—because the perturbation theory can only practically be applied to low order. Thus quantum chemistry is still lacking a general purpose theory that describes bonding situations across the entire potential energy surface with quantitative accuracy for both dynamic and nondynamic correlations.

Here we propose a theory that describes dynamic correlations in bonding situations where there is also significant nondynamic character. We shall assume that a suitable description of the nondynamic correlation, e.g., through a CASSCF or DMRG calculation, can first be obtained. When such a treatment is then combined with the proposed theory, a complete and quantitative description of a potential energy surface from equilibrium to stretched geometries is achieved. The current theory is size consistent and uses a unitary exponential description of dynamic correlation. The computational cost is of the same order as a *single-reference* coupled cluster method with the same level of excitations. We use the theory of cumulants to characterize the nondynamic correlation in the reference, and to simplify the working equations for the energy and amplitude equations. We shall call this theory *canonical transformation* theory, to reflect the essential ideas in its construction, and the link with the earlier canonical diagonalization theory of White.[21]

There is an extensive literature on the problem of constructing theories of dynamic correlation, using an exponential ansatz on top of multireference wave functions. One main approach is that taken by multireference coupled cluster theory, which is reviewed in the article by Paldus and Li.[16] These are often complicated by the need to handle very complicated multireference wave functions in the coupled cluster equations. In our current theory, we avoid a direct manipulation of the complex multireference wave function and instead characterize the nondynamic correlation in the reference using only the one and two particle reduced density matrices. A recent multireference CC ansatz with a similar emphasis on simplicity, but which uses the $T1$ and $T2$ amplitudes to characterize the nondynamic correlation in the reference, is the tailored CC theory of Kinoshita and co-workers.[22,23] The unitary exponential employed in our theory has previously been investigated in a multireference setting by Freed,[24] Kirtman,[25] and Hoffman and Simons.[26] We mention also the single-reference unitary coupled cluster work of Kutzelnigg,[27,28] Bartlett and co-workers,[29,30] and Pal

[a]Electronic mail: yanait@gmail.com

et al.[31,32] As recognized by Freed,[33] the unitary exponential can be seen as a kind of renormalization transformation and our work may be viewed from this perspective. In this context, we mention related work on the flow-renormalization group (Flow-RG) by Wegner[34] and Glazek and Wilson.[35] The Flow-RG was independently reinvented and further developed for quantum chemical problems by White,[21] who named it canonical diagonalization. The canonical diagonalization theory has been the most direct influence on this current work.

In Sec. II we present the basic equations and ideas of the canonical transformation theory. We introduce the linearized canonical transformation [linearized canonical transformation with doubles and linearized canonical transformation with singles and doubles (L-CTD and L-CTSD)] models. We carry out a perturbative analysis of the theory, and demonstrate that it is a familiar generalization of Hartree-Fock theory to a two-particle mean-field theory that includes electron correlation. In Sec. III we describe the computational implementation of the canonical transformation theory. In Sec. IV, we report ground-state calculations on the water and nitrogen potential energy curves, the BeH_2 insertion reaction, and hydrogen fluoride and boron hydride bond breaking. Finally, our conclusions and future directions of the theory are presented in Sec. V.

II. THEORY

A. Basic formulation

The generic chemical problem involving both dynamic and nondynamic correlations is illustrated in Fig. 1. The orbitals are divided into two sets: the active orbitals, usually the valence orbitals, which display partial occupancies (assuming spin-orbitals) very different from 0 or 1 for the state of interest, and the external orbitals, which are divided into the core (largely occupied in the target state) or virtual (largely unoccupied in the target state) orbitals. The asymmetry between the core and virtual orbitals can formally be removed by transforming to a core Fermi vacuum where all core states are filled (although in our numerical work, we retain the distinction between core and virtuals for reasons of efficiency). Nondynamic correlation is associated with active-active space correlations, while dynamic correlation is associated with correlations between the active-external and external-external spaces.

Define a reference wave function Ψ_0 (not in general a single determinant) which accounts for the nondynamic correlation, and which (relative to the core Fermi vacuum) exists only in the active space. The electronic Hamiltonian is written as

$$\hat{H} = \hat{H}_{\text{act}} + \hat{H}_{\text{act-ext}} + \hat{H}_{\text{ext}}. \tag{1}$$

An exact eigenfunction Ψ of $\hat{H}$ which incorporates the remaining dynamical correlations out of the active space, can be obtained by an appropriate canonical transformation of Ψ_0,

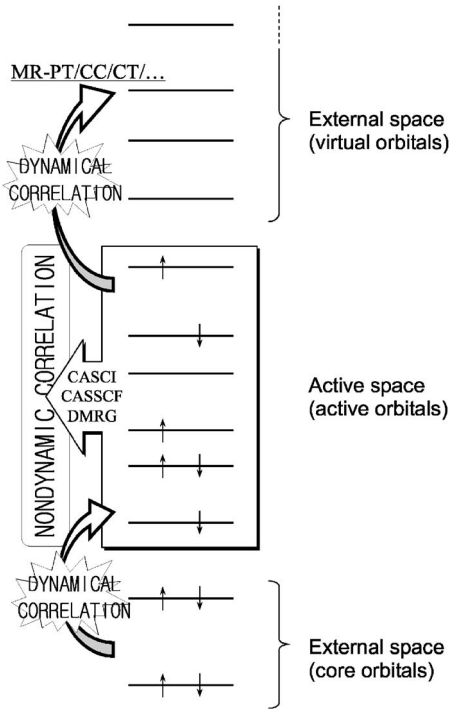

FIG. 1. Multireference problems involve both dynamical and nondynamical correlations. The nondynamical correlation is accounted for by the CASCI/CASSCF/DMRG wave function, which is made of multiple configurations generated in the active space with a fixed number of active electrons. The dynamical correlation is recovered on top of the multiconfigurational reference by correlating the active orbitals with orbitals in the external space, i.e., core and virtual orbitals.

$$\Psi = \hat{U}\Psi_0. \tag{2}$$

A canonical transformation, which may be single particle or many particle in nature, is one which preserves the commutation relations of the particles involved. Strictly speaking, it need not be unitary (it need only be isometric, see, e.g., Ref. 36), but this distinction is less important for calculational purposes and we shall henceforth consider only unitary canonical transformations where $\hat{U}$ satisfies $\hat{U}\hat{U}^\dagger = \hat{1}$.

If Ψ_0 is itself an eigenfunction of $\hat{H}_{\text{act}}$, then no active-active rotations are necessary and $\hat{U}$ rotates only between the active-external and external-external spaces. Without loss of generality, we can write $\hat{U}$ in exponential form, namely,

$$\hat{U} = e^{\hat{A}}, \tag{3}$$

$$\begin{aligned}\hat{A} = &\sum_{ae} A^a_e c^\dagger_a c_e + \sum_{e_1 e_2} A^{e_1}_{e_2} c^\dagger_{e_1} c_{e_2} + \sum_{a_1 a_2 e_3 e_4} A^{a_1 a_2}_{e_3 e_4} c^\dagger_{a_1} c^\dagger_{a_2} c_{e_3} c_{e_4} \\ &+ \sum_{a_1 a_2 a_3 e_4} A^{a_1 a_2}_{a_3 e_4} c^\dagger_{a_1} c^\dagger_{a_2} c_{a_3} c_{e_4} + \sum_{a_1 e_2 e_3 e_4} A^{a_1 e_2}_{e_3 e_4} c^\dagger_{a_1} c^\dagger_{e_2} c_{e_3} c_{e_4} \\ &+ \sum_{e_1 e_2 e_3 e_4} A^{e_1 e_2}_{e_3 e_4} c^\dagger_{e_1} c^\dagger_{e_2} c_{e_3} c_{e_4} + \cdots + h.c.,\end{aligned} \tag{4}$$

where a,e denote active, external indices, respectively, and all amplitudes A are antihermitian to ensure unitarity. $\hat{A}$ may be decomposed as the difference of an operator and its

hermitian conjugate, $\hat{A}=\hat{T}-T^{\dagger}$. The terms in Eq. (4) include not only the usual single and double excitations, but also semi-internal excitations that couple relaxation in the active space with external excitations.

The unitary operator can be pictured as acting either on the reference state, or on the reference Hamiltonian: in the latter case it gives rise to an effective Hamiltonian which has Ψ_0 as an eigenfunction and the exact eigenenergy E of the target state as its eigenvalue. Thus

$$\hat{\bar{H}} = e^{-\hat{A}}\hat{H}e^{\hat{A}}, \tag{5}$$

$$\hat{\bar{H}}\Psi_0 = E\Psi_0, \tag{6}$$

$$E = \langle\Psi_0|\hat{\bar{H}}|\Psi_0\rangle. \tag{7}$$

Note that in contrast to a general similarity transformation (as found for example in the usual coupled cluster theory) the canonical transformation produces a Hermitian effective Hamiltonian, which is computationally very convenient. When $\hat{U}$ is expressed in exponential form, the effective Hamiltonian can be constructed termwise via the formally infinite Baker-Campbell-Hausdorff (BCH) expansion,

$$\hat{\bar{H}} = \hat{H} + [\hat{H},\hat{A}] + \tfrac{1}{2}[[\hat{H},\hat{A}],\hat{A}] + \cdots. \tag{8}$$

The eigenvalue equation (6) defines the amplitudes A that determine the unitary operator. These can be reduced to nonlinear amplitude equations by projection on excited determinants, e.g.,

$$\langle\Psi_0|\hat{\bar{H}}\hat{\gamma}_\alpha|\Psi_0\rangle = 0, \tag{9}$$

where $\hat{\gamma}_\alpha$ denotes the operator $c^{\dagger}_{e_2}c^{\dagger}_{e_1}c_{a_2}c_{a_1}$ and α denotes the indices a_1a_2,e_1e_2. From the hermiticity of $\hat{\bar{H}}$, and the fact that (relative to the core Fermi vacuum) Ψ_0 exists only in the active space, it follows that

$$\langle\Psi_0|[\hat{\bar{H}}, \hat{\gamma}_\alpha - \hat{\gamma}^{\dagger}_\alpha]|\Psi_0\rangle = 0. \tag{10}$$

In this form, the amplitude equations (10) have been previously studied by Kutzelnigg, and named as the generalized Brillouin conditions.[37]

B. Cumulant formulation

To evaluate the energy and amplitude equations (7) and (10), we need to (i) construct $\hat{\bar{H}}$, and (ii) have some means of evaluating expectation values of operators $\hat{O}$ with the reference Ψ_0.

Let us first discuss (i). The primary difficulty associated with the infinite BCH expansion comes from the fact that each term in the expansion generates operators of greater particle rank (i.e., involving a longer string of creation and annihilation operators) than the previous term. Thus it is necessary to assume some closure or truncation when constructing $\hat{\bar{H}}$. This is commonly cited as an obstacle to the adoption of unitary coupled cluster theory, but as we shall see, a simple yet accurate closure can be found. More recently this concern has also arisen in studies of generalized two-body exponential theories of correlation.[38–41]

The approach typically taken in unitary coupled cluster theories is to truncate either by stopping the BCH expansion at some low order (e.g., second) (Ref. 26) or by keeping terms in the effective Hamiltonian such that it is correct through a given order of perturbation theory.[29] The accuracy of such a truncation is therefore tied to the accuracy of the underlying perturbation series. We will start with a different nonperturbative approach used in the theory of canonical diagonalization, that is, to restrict the *form* of $\hat{\bar{H}}$ to contain only certain classes of operators.[21] (In our numerical work, we shall restrict $\hat{\bar{H}}$ to contain only one- and two-particle operators.) If we then neglect all the higher particle-rank operators, we obtain the approximation used in canonical diagonalization. In the CT theory, we go one step further and account for the higher particle-rank operators appearing in $\hat{\bar{H}}$ in an approximate way. We achieve this by using an analog of the *cumulant decomposition* to express high particle-rank operators in terms of lower particle-rank operators and effective fields. This may be regarded as generalizing Hartree-Fock theory, where the effective Hamiltonian (the Fock operator), contains an average over the two-body interaction with a density field (see also Sec. II E).

The procedure is clearest with the aid of an example. Consider the first commutator in the BCH expansion, $[\hat{H},\hat{A}]$. Let $\hat{A}$ be the two-particle operator, $\Sigma_{a_1a_2e_1e_2}A^{a_1a_2}_{e_1e_2}c^{\dagger}_{a_1}c^{\dagger}_{a_2}c_{e_1}c_{e_2}$, and consider a two-particle term in the Hamiltonian, $V^{g_1g_2}_{g_3g_4}c^{\dagger}_{g_1}c^{\dagger}_{g_2}c_{g_3}c_{g_4}$ where g denotes a general index (i.e., a or e). Then, the commutator of the two terms yields both two-particle (through double contraction) and three-particle (through single contraction) operators. We wish to decompose the new three-particle operators in terms of one and two-particle quantities. Recall that the cumulant decomposition in statistical mechanics offers the best statistical decomposition of a high-particle rank correlation function in terms of lower-particle rank functions. In the context of reduced density matrices, it has been studied extensively by Colmenero and Valdemoro,[42,43] Yasuda and Nakatsuji,[44] Nakatsuji and Yasuda,[45] and Mazziotti.[46,47] The cumulant decomposition for a three-particle density matrix element $\langle c^{\dagger}_i c^{\dagger}_j c^{\dagger}_k c_l c_m c_n\rangle$ is

$$\langle c^{\dagger}_i c^{\dagger}_j c^{\dagger}_k c_l c_m c_n\rangle \Rightarrow 9\langle c^{\dagger}_i c^{\dagger}_n\rangle \wedge \langle c^{\dagger}_j c^{\dagger}_k c_l c_m\rangle - 12\langle c^{\dagger}_i c_n\rangle \wedge \langle c^{\dagger}_j c_m\rangle \wedge \langle c^{\dagger}_k c_l\rangle. \tag{11}$$

This is not quite what we need, as in the current context, we require a cumulant decomposition of a three-particle *operator*. We can construct an operator cumulant expansion by (i) requiring it to yield the same expectation value as the cumulant expansion for a reduced density matrix element, and (ii) by keeping two-particle operators rather than single-particle operators, when the choice arises. This gives

$$c_i^\dagger c_j^\dagger c_k^\dagger c_l c_m c_n \Rightarrow 9\langle c_i^\dagger c_n^\dagger\rangle \wedge (c_j^\dagger c_k^\dagger c_l c_m) - 12\langle c_i^\dagger c_n\rangle \wedge \langle c_j^\dagger c_m\rangle \wedge (c_k^\dagger c_l)$$

$$\begin{aligned}
&= \langle c_i^\dagger c_n\rangle c_j^\dagger c_k^\dagger c_l c_m + \langle c_j^\dagger c_m\rangle c_k^\dagger c_i^\dagger c_n c_l + \langle c_k^\dagger c_l\rangle c_i^\dagger c_j^\dagger c_m c_n - \langle c_i^\dagger c_l\rangle c_j^\dagger c_k^\dagger c_n c_m - \langle c_j^\dagger c_n\rangle c_k^\dagger c_i^\dagger c_m c_l - \langle c_k^\dagger c_m\rangle c_i^\dagger c_j^\dagger c_l c_n \\
&- \langle c_i^\dagger c_m\rangle c_j^\dagger c_k^\dagger c_l c_n - \langle c_j^\dagger c_l\rangle c_k^\dagger c_i^\dagger c_n c_m - \langle c_k^\dagger c_n\rangle c_i^\dagger c_j^\dagger c_m c_l - \tfrac{2}{3}\{(\langle c_i^\dagger c_n\rangle\langle c_j^\dagger c_m\rangle - \langle c_i^\dagger c_m\rangle\langle c_j^\dagger c_n\rangle)c_k^\dagger c_l + (\langle c_j^\dagger c_m\rangle\langle c_k^\dagger c_l\rangle \\
&- \langle c_j^\dagger c_l\rangle\langle c_k^\dagger c_m\rangle)c_i^\dagger c_n + (\langle c_k^\dagger c_l\rangle\langle c_i^\dagger c_n\rangle - \langle c_k^\dagger c_n\rangle\langle c_i^\dagger c_l\rangle)c_j^\dagger c_m - (\langle c_i^\dagger c_l\rangle\langle c_j^\dagger c_m\rangle - \langle c_i^\dagger c_m\rangle\langle c_j^\dagger c_l\rangle)c_k^\dagger c_n - (\langle c_j^\dagger c_n\rangle\langle c_k^\dagger c_l\rangle \\
&- \langle c_j^\dagger c_l\rangle\langle c_k^\dagger c_n\rangle)c_i^\dagger c_m - (\langle c_k^\dagger c_m\rangle\langle c_i^\dagger c_n\rangle - \langle c_k^\dagger c_n\rangle\langle c_i^\dagger c_m\rangle)c_j^\dagger c_l - (\langle c_i^\dagger c_m\rangle\langle c_j^\dagger c_n\rangle - \langle c_i^\dagger c_n\rangle\langle c_j^\dagger c_m\rangle)c_k^\dagger c_l - (\langle c_j^\dagger c_l\rangle\langle c_k^\dagger c_m\rangle \\
&- \langle c_j^\dagger c_m\rangle\langle c_k^\dagger c_l\rangle)c_i^\dagger c_n - (\langle c_k^\dagger c_n\rangle\langle c_i^\dagger c_l\rangle - \langle c_k^\dagger c_l\rangle\langle c_i^\dagger c_n\rangle)c_j^\dagger c_m\},
\end{aligned} \tag{12}$$

where $\wedge$ denotes an antisymmetrization over all indices with an associated factor $1/(P!)^2$ (P is the particle rank of the original operator) and $\langle\cdots\rangle$ denotes an average with the reference wave function Ψ_0 (this yields McWeeny normalization for the density matrices, i.e., $\mathrm{Tr}\langle c_i^\dagger c_j^\dagger c_m c_n\rangle = N(N-1)$, N is the number of particles).

Unlike the density cumulant expansion, which can in principle be exact for certain states (such as the Slater determinants), the operator cumulant expansion is never exact, in the sense that we cannot reproduce the full spectrum of a three-particle operator faithfully by an operator of reduced particle rank. However, if the density cumulant expansion is good for the state of interest, we expect the operator cumulant expansion to also be good for that state and also for states nearby.

With the above decomposition, the commutator $[\hat{H},\hat{A}]$ is reduced to an expression containing only terms of the form we wish to keep, i.e., one- and two-particle operators. Let us denote this approximate form of the commutator as $[\hat{H},\hat{A}]_{(1,2)}$, to indicate that the cumulant decomposition retains only one- and two-particle operators. Then, we can apply the procedure recursively, and thus the next commutator in the BCH expansion is approximated as

$$[[\hat{H},\hat{A}],\hat{A}] = [[\hat{H},\hat{A}]_{(1,2)},\hat{A}]_{(1,2)}. \tag{13}$$

Consequently, we can carry out the BCH expansion to arbitrarily high order without any increase in the complexity of the terms in the effective Hamiltonian. In practice, the expansion is carried out until convergence in a suitable norm of the operator coefficients is achieved. Note that through the decomposition (12), the effective Hamiltonian depends on the one- and two-particle density matrices and therefore becomes state specific, much like the Fock operator in Hartree-Fock theory.

Size consistency is a desirable feature of any approximate theory. Since we truncate in the operator space (as opposed to the Hilbert space of wave functions), the current approximation is naturally size consistent. Consider two widely separated systems X and Y. Then, we can construct two bases of creation/annihilation operators, $c_X^\dagger, c_Y^\dagger$ that generate/destroy the Fock spaces of X and Y, respectively, and which commute by virtue of separation. In terms of these operators, the starting Hamiltonian is separable into components that act only on X and Y, respectively, $\hat{H}=\hat{H}_X+\hat{H}_Y$, and so too is the exponential operator $\exp(\hat{A})=\exp(\hat{A}_X+\hat{A}_Y)=\exp(\hat{A}_X)\exp(\hat{A}_Y)$. Consequently the effective Hamiltonian is also separable $\hat{\bar{H}}=\hat{\bar{H}}_X+\hat{\bar{H}}_Y$. The amplitude equations (10) are solvable separately in the X and Y spaces. Consequently, the total effective hamiltonian is the sum of the corresponding effective hamiltonians for the systems X and Y considered in isolation, the total wave function is a product, and the total energy is additive, as is required in a size-consistent theory.

We now discuss (ii), the evaluation of operator expectation values with the reference Ψ_0. We are interested in multireference problems where Ψ_0 may be extremely complicated, i.e., a very long Slater determinant expansion, or a compact but complex wave function, such as the DMRG wave function. By using the cumulant decomposition, we limit the terms that appear in the effective Hamiltonian to only low order, e.g., one- and two-particle operators, and thus we only need the one- and two-particle density matrices of the reference wave function to evaluate the expectation value of the energy in the energy expression (7). To solve the amplitude equations, we further require the commutator of $[\hat{\bar{H}},\hat{\gamma}_\alpha]$, which, for a two-particle effective hamiltonian and two-particle operator $\hat{\gamma}_\alpha$ again involves the expectation value of three-particle operators. We therefore invoke the cumulant decomposition once more, and solve instead the modified amplitude equation

$$\langle\Psi_0|[\hat{\bar{H}}_{(1,2)},\hat{\gamma}_\alpha]_{(1,2)}|\Psi_0\rangle = 0. \tag{14}$$

This modified amplitude equation does not correspond to the minimization of the energy functional (7), and thus the generalized Hellmann-Feynman theorem[48] does not apply.

Consequently, with the simplifications above, all the working equations of the canonical transformation theory can be evaluated entirely in terms of a limited number of reduced density matrices, e.g., one- and two-particle density matrices, and *no* explicit manipulation of the complicated reference function is required.

C. The linearized model

To summarize the theory, dynamic correlations are described by the unitary operator $\exp\hat{A}$ acting on a suitable reference function, where $\hat{A}$ consists of excitation operators of the form (4). We employ a cumulant decomposition to evaluate all expressions in the energy and amplitude equa-

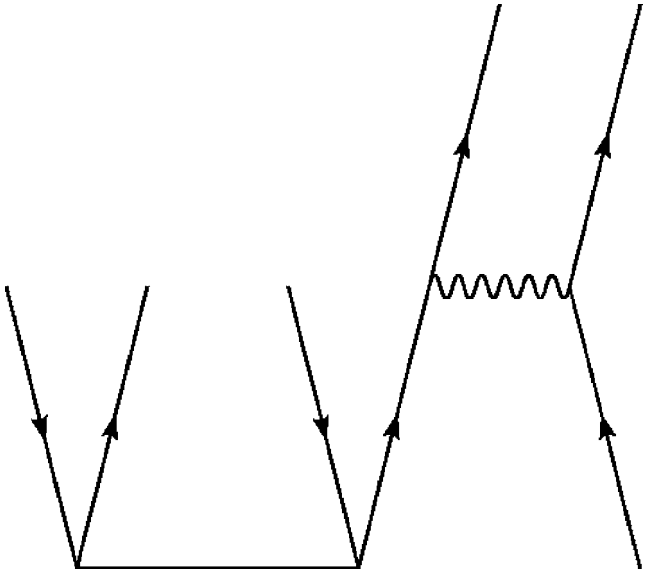

FIG. 2. An example of a diagram of the three-particle operator appearing in $[\hat{W},\hat{A}_2]$.

tions. Since we are applying the cumulant decomposition after the first commutator (the term "linear" in the amplitudes), we shall call this theory *linearized* canonical transformation theory, by analogy with the coupled cluster usage of the term. (Further links between this linearization approximation and that in coupled cluster theory are shown in the perturbative analysis of Sec. II D.)

In the current work, we shall consider primarily two theoretical models: the L-CTD L-CTSD theories. These are defined by the choice of operators in $\hat{A}$. The L-CTD theory contains only two-particle operators (including also the two-particle semi-internal excitations) and the L-CTSD theory contains both one and two-particle operators, thus

$$\hat{A}(\text{L-CTSD}) = \hat{A}_1 + \hat{A}_2,$$

$$\hat{A}(\text{L-CTD}) = \hat{A}_2,$$

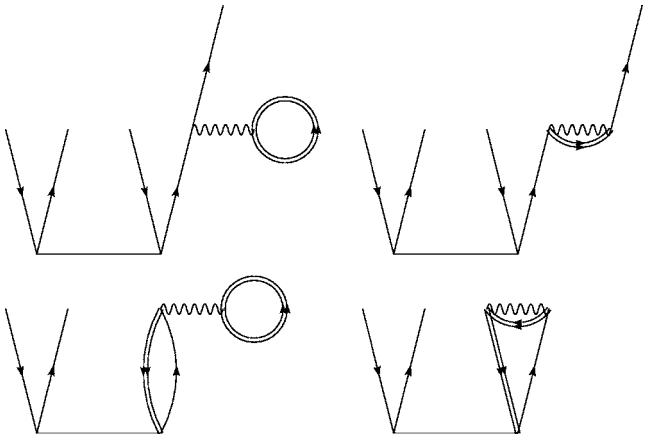

FIG. 3. A diagrammatic representation of the cumulant decomposition $([\hat{W},\hat{A}_2]_{(1,2)})$ for the three-particle operator drawn in Fig. 2. Four kinds of one- and two-particle operators are obtained. The double line is the contraction for the particle rank reduction (closure) where the correlation is averaged with the effective field, i.e., density matrices.

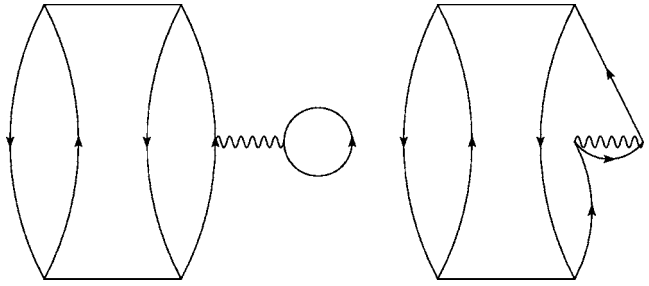

FIG. 4. Two connected diagrams in the term $\langle[[\hat{W},\hat{A}_2],\hat{A}_2]\rangle$, which contribute to E_3 [Eq. (20)].

$$\hat{A}_1 = \sum_{ae} A_e^a c_a^\dagger c_e + \sum_{e_1 e_2} A_{e_2}^{e_1} c_{e_1}^\dagger c_{e_2},$$

$$\hat{A}_2 = \sum_{a_1 a_2 e_3 e_4} A_{e_3 e_4}^{a_1 a_2} c_{a_1}^\dagger c_{a_2}^\dagger c_{e_3} c_{e_4} + \sum_{a_1 a_2 a_3 e_4} A_{a_3 e_4}^{a_1 a_2} c_{a_1}^\dagger c_{a_2}^\dagger c_{a_3} c_{e_4} + \sum_{a_1 e_2 e_3 e_4} A_{e_3 e_4}^{a_1 e_2} c_{a_1}^\dagger c_{e_2}^\dagger c_{e_3} c_{e_4} + \sum_{e_1 e_2 e_3 e_4} A_{e_3 e_4}^{e_1 e_2} c_{e_1}^\dagger c_{e_2}^\dagger c_{e_3} c_{e_4} + \cdots + h.c.$$

Although L-CTD theory does not include explicit one-particle single excitations (it does include two-particle semi-internal single excitations), in most of the applications in this work it is combined with a CASSCF reference, which is already based on optimized orbitals.

D. Perturbative analysis and relation to coupled cluster theory

Perturbative analyses have yielded many insights into single-reference coupled cluster theory. Although we will generally be using the canonical transformation theory together with a *multireference* wave function, it is informative to carry out the analogous perturbative analysis for the single-reference canonical transformation theory, to highlight the connections with existing coupled cluster methods. The analysis in the section follows that of Bartlett *et al.*[29,30,49]

First consider a Hartree-Fock reference function, and transform to the Fermi vacuum (all occupied orbitals are in the vacuum). Then all particle density matrices are zero and

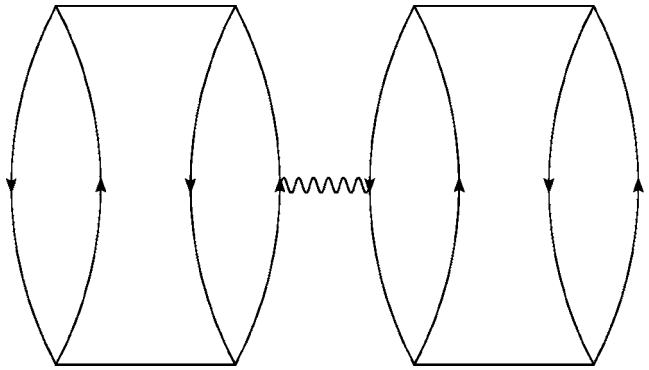

FIG. 5. A diagram in $\langle[[[\hat{W},\hat{A}_2],\hat{A}_2],\hat{A}_2]\rangle$ that yields nonzero energy in E_4 [Eq. (21)] and which is missed in the cumulant decomposition in L-CTSD theory. In this diagram, the three-particle operator arising from $[\hat{W},\hat{A}_2]$ contracts successively with two other $\hat{A}_2$ terms.

the cumulant decomposition Eq. (12) based on this reference corresponds to simply neglecting all three and higher particle-rank operators generated by commutators. This type of operator truncation is used in the canonical diagonalization theory of White.[21]

Now write the Hamiltonian as

$$\hat{H} = E_{\mathrm{HF}} + \hat{F} + \hat{W}, \tag{15}$$

where $\hat{F}$ is the one-particle Fock operator and $\hat{W}$ is the two-particle fluctuation potential. From Brillouin's theorem, we recognize that $\hat{A}_2$ is first order in $\hat{W}$, while $\hat{A}_1$ is second order in $\hat{W}$. [To make contact with the analysis of unitary coupled cluster theory in Refs. 29 and 30, write $\hat{A}_1$ as $(\hat{T}_1 - \hat{T}_1^\dagger)$ and $\hat{A}_2 = (\hat{T}_2 - \hat{T}_2^\dagger)$.] Then consider the expectation value of the energy $E = \langle \exp \hat{A}^\dagger \hat{H} \exp \hat{A} \rangle$ *without* using any cumulant decomposition. Expanding in powers of the fluctuation operator, we have

$$E = E^0 + E^1 + E^2 + E^3 + E^4 + \cdots, \tag{16}$$

where these are defined as

$$E^0 = \langle E_{\mathrm{HF}} + \hat{F} \rangle, \tag{17}$$

$$E^1 = \langle \hat{W} \rangle, \tag{18}$$

$$E^2 = \langle [\hat{W}, \hat{A}_2] + [\hat{F}, \hat{A}_1] + [[\hat{F}, \hat{A}_2], \hat{A}_2] \rangle, \tag{19}$$

$$E^3 = \left\langle \tfrac{1}{2}[[\hat{W}, \hat{A}_2], \hat{A}_2] + [\hat{W}, \hat{A}_1] + \tfrac{1}{2}[[\hat{F}, \hat{A}_2], \hat{A}_1] + \tfrac{1}{2}[[\hat{F}, \hat{A}_1], \hat{A}_2] \right\rangle, \tag{20}$$

$$E^4 = \left\langle \tfrac{1}{6}[[[\hat{W}, \hat{A}_2], \hat{A}_2], \hat{A}_2] + \tfrac{1}{2}[[\hat{W}, \hat{A}_1], \hat{A}_2] + \tfrac{1}{2}[[W, \hat{A}_2], \hat{A}_1] + \tfrac{1}{2}[[\hat{F}, \hat{A}_1], \hat{A}_1] + \tfrac{1}{24}[[[[\hat{F}, \hat{A}_2], \hat{A}_2], \hat{A}_2], \hat{A}_2] \right\rangle. \tag{21}$$

Now consider the effect of the cumulant decomposition on the different orders of energy contribution. Firstly, no decomposition is involved in computing E^0, E^1. For E^2, the cumulant decomposition corresponds to

$$E^2 \Rightarrow \langle [\hat{W}, \hat{A}_2]_{(1,2)} + [\hat{F}, \hat{A}_1] + [[\hat{F}, \hat{A}_2], \hat{A}_2]_{(1,2)} \rangle. \tag{22}$$

We have used the subscript (1,2) only when the commutator generates three-particle terms, e.g., $[\hat{F}, \hat{A}_2]$ generates only two-particle terms and thus no decomposition is applied. We can illustrate the different terms diagrammatically using the coupled cluster-type diagrams popularized by Bartlett.[50] Figure 2 illustrates a three-particle term that appears in $[\hat{W}, \hat{A}_2]$. Additional "double" lines are used to indicate contractions with a reduced density matrix. When these double lines are cut and rotated, one recovers the usual CC-type diagram. (The cumulant decomposition of $[\hat{W}, \hat{A}_2]_{(1,2)}$ yields four kinds of diagrams for one- and two-particle operators, shown in Fig. 3, but for the single reference case we are considering, all these terms vanish since all particle density matrices are zero from the Fermi vacuum.) Now $\langle [\hat{F}, \hat{A}_1] \rangle$ vanishes (from Brillouin's theorem). Both $[\hat{W}, \hat{A}_2]$ and $[[\hat{F}, \hat{A}_2], \hat{A}_2]$ generate three-particle operators that are approximated in the cumulant decomposition, but these have no expectation value with the Fermi vacuum, and thus do not contribute to the energy. Thus no error is made in Eq. (22) for E^2.

In the expression for E^3, we apply the cumulant decomposition twice for the double commutator $[[\hat{W}, \hat{A}_2]_{(1,2)}, \hat{A}_2]_{(1,2)}$. Once again, only the fully contracted term contributes to the energy. The only way fully contracted terms arise is from double contractions in $[\hat{W}, \hat{A}_2]$ to produce a two-particle operator, which then doubly contracts with the final $\hat{A}_2$ commutator, to contribute to the energy. Since double contractions are involved in each step, no cumulant decomposition is involved for this term. There is no contribution from the three-particle operators generated by either commutator, and the cumulant decomposition approximation is exact for E^3 (see Fig. 4).

In the expression for E^4 we find our first error from using the cumulant decomposition. Here, the three-particle operator arising from the first commutator $[\hat{W}, \hat{A}_2]$, which is dropped in the cumulant decomposition, can contract successively with two other $\hat{A}_2$ terms, in $[[[\hat{W}, \hat{A}_2], \hat{A}_2], \hat{A}_2]$ to yield a fully contracted term and a contribution to the energy. Although the cumulant decomposition misses this contribution, it *does*, however, contain the contribution that arises from contracting the two-particle operators generated in the first commutator $[\hat{W}, \hat{A}_2]$. A diagrammatic illustration of the same result is shown in Fig. 5. By a similar analysis, we find that the cumulant decomposition also provides an incomplete evaluation of $[[[[\hat{F}, \hat{A}_2], \hat{A}_2], \hat{A}_2], \hat{A}_2]$, arising from intermediate three-particle operators.

In the usual coupled cluster hierarchy $\Sigma_{i=0}^2 E_i$ is the MP2 energy functional, while $\Sigma_{i=0}^3 E_i$ is the linearized coupled cluster single-doubles (L-CCSD) energy functional. $\Sigma_{i=0}^4 E_i$ is the unitary CCSD energy functional. The linearized CTSD energy is correct up to third order in perturbation theory, such as the linearized CCSD theory. However, unlike the linearized CCSD theory, fourth order terms (such as $[[[\hat{W}, \hat{A}_2], \hat{A}_2], \hat{A}_2]$) are not completely neglected but partly included as discussed above. From this, we might expect the single-reference L-CTSD theory to perform intermediate between linearized CCSD and the full CCSD theories. But, in fact, there are an *infinite* number of additional diagrams that are included in linearized CTSD theory as compared to the usual CC and UCC(n) theories, because the energy functional does not terminate at finite order, but contains further partial contributions from E^5, E^6 and indeed to infinite order. For example, all terms involving pure orbital rotations (i.e., $\hat{A}_1$) are included to *all* orders in the energy functional. Terms involving $\hat{A}_2$ where all $\hat{A}_2$ operators are at least doubly contracted with one other operator also included to all orders. Example of these additional diagrams contained in L-CTSD, but not in the usual CC theories, are shown in Fig. 6. One might speculate that these additional diagrams would yield an improved theory, but in the general case, and certainly when we extend the discussion to cases where a multideterminantal reference wave function is used, the significance of

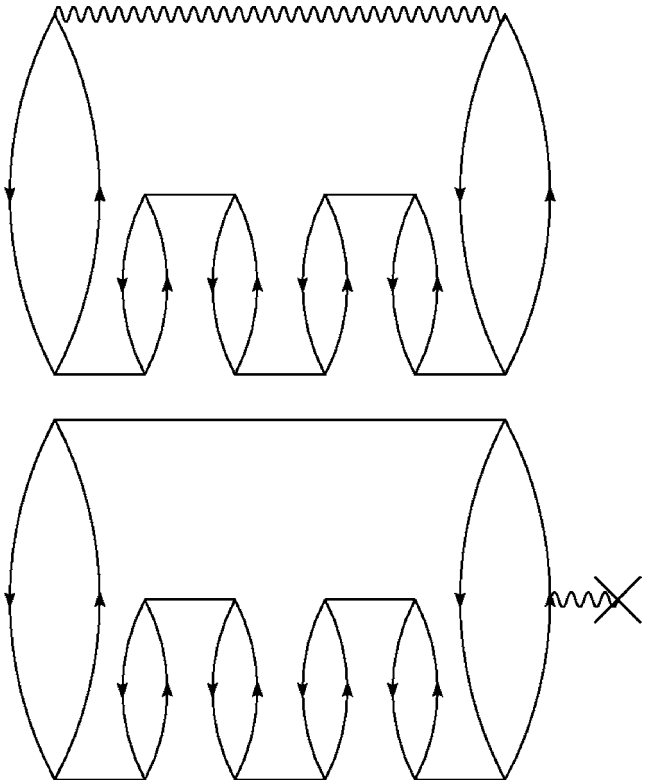

FIG. 6. An example of two diagrams in $[\cdots[[\hat{W},\hat{A}_2],\hat{A}_2]\cdots],\hat{A}_2]$ (upper) and $[\cdots[\hat{F},\hat{A}_2],\hat{A}_2]\cdots],\hat{A}_2]$ (lower) that appear at higher orders in linearized CTD and CTSD. The diagrams involve $\hat{A}_2$ where all $\hat{A}_2$ operators (here six $\hat{A}_2$) are at least doubly contracted with one other operator.

these additional contributions can only be assessed numerically.

E. Variation of the reference

A further point is of interest in the formal discussion of the canonical transformation theory. So far, we have assumed that the reference function is fixed, and have considered only solving for the amplitudes in the excitation operator. We may also consider optimization of the reference function itself in the presence of the excitation operator $\hat{A}$. This consideration is useful in understanding the nature of the cumulant decomposition in the canonical transformation theory.

Using the energy functional (7) and the cumulant decomposition, and making the energy stationary with respect to variations in Ψ_0^*, we find that the optimal reference Ψ_0 satisfies

$$\hat{\bar{H}}_{(1,2)}\Psi_0 + \langle\Psi_0|\frac{\delta\hat{\bar{H}}_{(1,2)}}{\delta\Psi_0^*}|\Psi_0\rangle = \hat{\bar{F}}\Psi_0 = 0, \tag{23}$$

where the second term arises because the effective Hamiltonian is state dependent through the usage of the cumulant decomposition. Thus the optimal reference function is an eigenfunction not of the effective Hamiltonian, but a correlated two-particle "Fock" operator $\hat{\bar{F}}$.

To understand this more clearly, consider a simpler model where $\hat{A}$ consists of single excitations, only single-particle operators are retained in the effective Hamiltonian, and we choose the reference function Ψ_0 to be a single determinant. Then, from a cumulant decomposition of the two-particle terms, the effective Hamiltonian becomes

$$\hat{\bar{H}}_{(1)} = \sum_{ij} T_{ij}c_i^\dagger c_j + \tfrac{1}{2}\sum_{ijkl} V_{ijkl}[\langle c_i^\dagger c_l\rangle c_j^\dagger c_k - \langle c_j^\dagger c_l\rangle c_i^\dagger c_k]. \tag{24}$$

Note this resembles the (N-particle) Fock operator that appears in the Hartree-Fock theory, but the contribution of the two-electron term is only half the normal contribution in the Fock operator. However, if we consider making the energy stationary with respect to variations in the reference, we must also consider the second term in Eq. (23), where we find

$$\frac{\delta\hat{\bar{H}}_{(1)}}{\delta\Psi_0^*} = \tfrac{1}{2}\sum_{ijkl}[V_{ijkl}(\langle c_i^\dagger c_l\rangle c_j^\dagger c_k - \langle c_j^\dagger c_l\rangle c_i^\dagger c_k)]\Psi_0, \tag{25}$$

and thus the final Fock operator $\hat{\bar{F}}$ that determines the optimal reference function is identical to the usual Hartree-Fock operator

$$\hat{\bar{F}} = \sum_{ij} T_{ij}c_i^\dagger c_j + \sum_{ijkl} V_{ijkl}[\langle c_i^\dagger c_l\rangle c_j^\dagger c_k - \langle c_j^\dagger c_l\rangle c_i^\dagger c_k]. \tag{26}$$

Thus we see that the Hartree-Fock theory is identical to a canonical transformation theory retaining only one-particle operators with an optimized reference, and the canonical transformation model retaining one- and two-particle operators employed in the current work, if employed with an optimized reference, is a natural extension of the Hartree-Fock theory to a two-particle theory of correlation.

Finally, we note that if we retain two-particle operators in the effective Hamiltonian, but restrict $\hat{A}$ to single-particle form, we recover exactly the orbital rotation formalism of the multiconfigurational self-consistent field. Indeed, this is the way in which we obtain the CASSCF wave functions used in this work.

III. IMPLEMENTATION OF THE LINEARIZED CANONICAL TRANSFORMATION THEORY

A. Computational algorithm

Considering the one- and two-particle operators in Eq. (14) separately, we obtain equations for the L-CTSD and L-CTD models,

$$R_q^p = \langle[\hat{\bar{H}}_{(1,2)}, \hat{\gamma}_q^p - \hat{\gamma}_q^{p\dagger}]\rangle = 0 \quad \text{(L-CTSD)}, \tag{27}$$

$$R_{rs}^{pq} = \langle[\hat{\bar{H}}_{(1,2)}, \hat{\gamma}_{rs}^{pq} - \hat{\gamma}_{rs}^{pq\dagger}]_{(1,2)}\rangle = 0 \quad \text{(L-CTD/L-CTSD)}, \tag{28}$$

where $\hat{\gamma}_q^p = c_p^\dagger c_q$ and $\hat{\gamma}_{rs}^{pq} = c_p^\dagger c_q^\dagger c_s c_r$. These nonlinear equations must be solved for the amplitudes A which define the effective Hamiltonian. A sketch of our implementation is as follows:

(1) Given the electronic Hamiltonian $\hat{H}$ [Eq. (1)] determine the reference function in the active space (e.g., CASSCF, CASCI, or HF). Compute the one- and two-particle density matrices $\gamma_q^p = \langle c_p^\dagger c_q \rangle$ and $\gamma_{rs}^{pq} = \langle c_p^\dagger c_q^\dagger c_s c_r \rangle$ of the reference function.

(2) Compute the preconditioner for the amplitude equations, given by the diagonal linear terms of the amplitude equations [Eqs. (27) and (28)],

$$D_q^p = \langle [[\hat{H}, \hat{\gamma}_q^p - \hat{\gamma}_q^{p\dagger}], \hat{\gamma}_q^p - \hat{\gamma}_q^{p\dagger}] \rangle, \tag{29}$$

$$D_{rs}^{pq} = \langle [[\hat{H}, \hat{\gamma}_{rs}^{pq} - \hat{\gamma}_{rs}^{pq\dagger}]_{(1,2)}, \hat{\gamma}_{rs}^{pq} - \hat{\gamma}_{rs}^{pq\dagger}]_{(1,2)} \rangle. \tag{30}$$

(3) Choose the initial amplitudes A (which we set as zero).

(4) Compute the transformed Hamiltonian, $\hat{\bar{H}}$ via the BCH equation [Eq. (8)]. This is done by iterating a subroutine that computes the cumulant-decomposed commutator $\hat{\bar{H}}^{(n+1)} = [\hat{\bar{H}}^{(n)}, \hat{A}]_{(1,2)}$, which contains only one- and two-particle operators, and where $\hat{\bar{H}}^{(0)} = \hat{H}$. The full $\hat{\bar{H}}$ is then obtained as a sum of one- and two-particle operators $\hat{\bar{H}} = \Sigma_{n=0} \hat{\bar{H}}^{(n)}/n!$. The sum is truncated after the norm of coefficients of the nth term $\hat{\bar{H}}^{(n)}/n!$ is less than a given threshold, which we set as 10^{-9}. The energy is then computed as $E = \langle \hat{\bar{H}} \rangle$ with the density matrix elements γ_q^p and γ_{rs}^{pq} of the reference wave function.

(5) Compute the new residuals of the amplitude equations, R_q^p [Eq. (27)] and R_{rs}^{pq} [Eq. (28)].

(6) Correct the amplitudes by adding the preconditioned residuals,

$$A_q^p \leftarrow A_q^p + R_q^p / D_q^p, \tag{31}$$

$$A_{rs}^{pq} \leftarrow A_{rs}^{pq} + R_{rs}^{pq} / D_{rs}^{pq}. \tag{32}$$

(7) Repeat (4)–(6) until convergence.

Within the above scheme, we implemented the generalized minimal residual (GMRES) method,[51] which is a robust linear solver that ensures convergence of the iterative solution.

B. Computational scaling

In active space calculations, the total orbital space is usually partitioned into external core orbitals (c), active orbitals (a), and unoccupied virtual (external) orbitals (v). (There can additionally be some frozen core orbitals which remain doubly occupied throughout the calculation.)

In the iterative algorithm outlined in Sec. III A the computational scalings are $n^2 n_{ca}^2 n_{av}^2$ for step (4), and $n n_{ca} n_{av}^4$ and $n n_{ca}^2 n_{av}^3$ for step (5), where $n = n_c + n_a + n_v$, $n_{ca} = n_c + n_a$, and $n_{av} = n_a + n_v$. Note that, unlike the conventional multireference methods, these scalings do not depend on the number of configurations in the expansion of the reference wave function. In fact, the scaling is roughly $\sim O(n_a^2 n_v^4)$ which is *essentially* the same as that of single-reference coupled cluster theory.

TABLE I. Classification of one-particle excitation operators [c, a, and v denote active core, active, virtual (external) orbitals].

Classified operator	$\hat{A}_q^p = \hat{\gamma}_q^p - \hat{\gamma}_q^{p\dagger}$
ca	$\hat{\gamma}_{c_1}^{a_1'} = c_{a_1'}^\dagger c_{c_1}$
cv	$\hat{\gamma}_{c_1}^{v_1'} = c_{v_1'}^\dagger c_{c_1}$
av	$\hat{\gamma}_{a_1}^{v_1'} = c_{v_1'}^\dagger c_{a_1}$

C. Classes of excitations for the exponential operator

To improve the convergence in the iterative solution of the CT amplitude equations, we have adopted a simple scheme, where we classify the amplitude equations by the type of excitation operators involved and solve the classified sets of the amplitude equations in successive steps.

First number the orbitals thus, core orbitals range from 1 through the number of core orbitals N_{core}, i.e., $c = 1, \ldots, N_{\mathrm{core}}$. Active orbitals range from $N_{\mathrm{core}}+1$ to $N_{\mathrm{core}} + N_{\mathrm{active}}$, and the virtual orbitals start from $N_{\mathrm{core}} + N_{\mathrm{active}} + 1$. Then, we divide the equations into eight classes (see also Tables I and II: (i) internal double excitations $ccaa$, (ii) external double excitations ($ccvv$, $cavv$, and $aavv$) where the indices of ca, ca, and aa are $\leqslant$ the number of electrons N_{elec}, (iii) external double excitations where either of the indices of ca, ca, and aa is $\leqslant N_{\mathrm{elec}}$ and the other is $> N_{\mathrm{elec}}$, (iv) external double excitations where the indices of ca, ca, and aa are $> N_{\mathrm{elec}}$, (v) single excitations (ca, cv, and av), (vi) semi-internal excitations ($ccav$, $caav$, and $aaav$) where the indices of ca, ca, and aa are $\leqslant N_{\mathrm{elec}}$, (vii) semi-internal excitations where either of the indices of ca, ca, and aa is $\leqslant N_{\mathrm{elec}}$ and the other is $> N_{\mathrm{elec}}$, and (viii) semi-internal excitations where the indices of ca, ca, and aa are $> N_{\mathrm{elec}}$. When solving for a given class of amplitudes, the amplitudes in all previous classes are also allowed to vary, while the amplitudes in all later classes are 0. The excitation operators for $caaa$ are entirely neglected, and those for (i), (vi), (vii), and (viii) are

TABLE II. Classification of two-particle excitation operators [c, a, and v denote active core, active, virtual (external) orbitals].

Classified operator	$\hat{A}_{rs}^{pq} = \hat{\gamma}_{rs}^{pq} - \hat{\gamma}_{rs}^{pq\dagger}$
Internal excitation	
$ccaa$	$\hat{\gamma}_{c_1 c_2}^{a_1' a_2'} = c_{a_1'}^\dagger c_{a_2'}^\dagger c_{c_2} c_{c_1}$
$caaa$	$\hat{\gamma}_{c_1 a_2}^{a_1' a_2'} = c_{a_1'}^\dagger c_{a_2'}^\dagger c_{a_2} c_{c_1}$
Semi-internal excitation	
$ccav$	$\hat{\gamma}_{c_1 c_2}^{v_1' a_2'} = c_{v_1'}^\dagger c_{a_2'}^\dagger c_{c_2} c_{c_1}$
$caav$	$\hat{\gamma}_{c_1 a_2}^{v_1' a_2'} = c_{v_1'}^\dagger c_{a_2'}^\dagger c_{a_2} c_{c_1}$
$aaav$	$\hat{\gamma}_{a_1 a_2}^{v_1' a_2'} = c_{v_1'}^\dagger c_{a_2'}^\dagger c_{a_2} c_{a_1}$
Double excitation	
$ccvv$	$\hat{\gamma}_{c_1 c_2}^{v_1' v_2'} = c_{v_1'}^\dagger c_{v_2'}^\dagger c_{c_2} c_{c_1}$
$cavv$	$\hat{\gamma}_{c_1 a_2}^{v_1' v_2'} = c_{v_1'}^\dagger c_{v_2'}^\dagger c_{a_2} c_{c_1}$
$aavv$	$\hat{\gamma}_{a_1 a_2}^{v_1' v_2'} = c_{v_1'}^\dagger c_{v_2'}^\dagger c_{a_2} c_{a_1}$

TABLE III. Total energies (E_h) for the simultaneous bond breaking of the H_2O molecule with 6-31G basis sets. The value in parentheses is the difference from the FCI total energy in mE_h. The bond angle is fixed at $\angle$HOH=109.57°. R_e=0.9929 Å.

r_{OH}	Method reference CAS orbitals	HF	FCI	CASSCF (6e,5o)	L-CTSD CASSCF (6e,5o)	L-CTSD CASSCF (6e,5o) NOs	L-CTSD(2) CASSCF (6e,5o) NOs	L-CTD CASSCF (6e,5o)	L-CTD CASSCF (6e,5o) NOs	L-CTD CASCI (6e,5o) NOs	MRMP CASSCF (6e,5o)	CCSD	CCSDT
1.0R_e		−75.981 92	−76.121 02	−76.038 02	−76.120 65	−76.120 63	−76.119 91	−76.119 63	−76.119 45	−76.118 03	−76.114 23	−76.119 35	−76.120 54
		(139.09)		(83.00)	(0.37)	(0.38)	(1.11)	(1.39)	(1.57)	(2.99)	(6.79)	(1.66)	(0.48)
1.4R_e		−75.816 77	−76.004 54	−75.877 86	−76.003 77	−76.003 85	−76.003 24	−76.003 10	−76.002 70	−75.998 37	−76.000 44	−75.999 56	−76.003 47
		(187.77)		(80.06)	(0.77)	(0.69)	(1.30)	(1.44)	(1.84)	(6.17)	(4.10)	(4.98)	(1.07)
1.8R_e		−75.636 59	−75.899 26	−75.827 79	−75.905 43	−75.905 22	−75.904 60	−75.903 24	−75.901 69	−75.897 38	−75.896 01	−75.889 54	−75.898 68
		(262.67)		(71.47)	(−6.17)	(−5.95)	(−5.33)	(−3.97)	(−2.42)	(1.89)	(3.25)	(9.72)	(0.59)
2.2R_e		−75.500 43	−75.852 77	−75.788 37	−75.854 97	−75.853 77	−75.853 25	−75.854 53	−75.852 64	−75.857 70	−75.848 54	−75.830 50	−75.858 13
		(352.34)		(64.40)	(−2.20)	(−1.00)	(−0.49)	(−1.76)	(0.13)	(−4.93)	(4.23)	(22.27)	(−5.36)
2.6R_e		−75.439 89	−75.839 55	−75.778 09	−75.838 19	−75.839 21	−75.838 57	−75.838 52	−75.839 42	−75.848 42	−75.834 47		
		(399.66)		(61.46)	(1.36)	(0.34)	(0.98)	(1.03)	(0.13)	(−8.87)	(5.08)		
3.0R_e		−75.415 73	−75.836 32	−75.775 68	−75.834 14	−75.835 95	−75.835 24	−75.835 17	−75.836 61	−75.848 23	−75.830 99		
		(420.59)		(60.64)	(2.18)	(0.37)	(1.08)	(1.15)	(−0.29)	(−11.91)	(5.33)		
3.4R_e		−75.403 84	−75.835 44	−75.775 03	−75.832 36	−75.834 94	−75.834 30	−75.833 58	−75.835 82	−75.848 36	−75.830 05		
		(431.60)		(60.41)	(3.07)	(0.50)	(1.14)	(1.85)	(−0.39)	(−12.93)	(5.39)		
3.8R_e		−75.397 02	−75.835 14	−75.774 80	−75.831 41	−75.834 75	−75.834 07	−75.833 63	−75.835 68	−75.848 54	−75.829 74		
		(438.12)		(60.34)	(3.73)	(0.40)	(1.08)	(1.51)	(−0.54)	(−13.39)	(5.40)		

TABLE IV. Total energies (E_h) for the simultaneous bond-breaking of the H_2O molecule with cc-pVDZ basis sets. The value in parentheses is the difference from the FCI total energy in mE_h. The bond angle is fixed at $\angle$HOH=109.57°. R_e=0.9929 Å.

r_{OH}	Method reference CAS orbitals	HF	FCI	CASSCF (6*e*,5*o*)	L-CTSD CASSCF (6*e*,5*o*) NOs	L-CTSD(2) CASSCF (6*e*,5*o*) NOs	L-CTD CASSCF (6*e*,5*o*) NOs	MRMP CASSCF (6*e*,5*o*)	CCSD	CCSDT
1.0R_e		−76.021 67 (217.18)	−76.238 85	−76.075 86 (162.99)	−76.238 19 (0.66)	−76.238 12 (0.73)	−76.236 60 (2.26)	−76.224 07 (14.78)	−76.235 01 (3.84)	−76.238 34 (0.51)
1.4R_e		−75.841 12 (257.89)	−76.099 02	−75.945 57 (153.45)	−76.098 80 (0.22)	−76.098 75 (0.27)	−75.097 14 (1.88)	−76.089 46 (9.56)	−76.090 46 (8.55)	−76.097 78 (1.23)
1.8R_e		−75.651 86 (326.28)	−75.978 14	−75.840 02 (138.12)	−75.984 11 (−5.97)	−75.984 01 (−5.88)	−75.980 47 (−2.34)	−75.970 98 (7.16)	−75.960 19 (17.94)	−75.976 80 (1.34)
2.2R_e		−75.510 38 (416.84)	−75.927 22	−75.799 46 (127.76)	−75.926 65 (0.58)	−75.926 49 (0.74)	−75.924 89 (2.34)	−75.919 86 (7.37)	−75.903 63 (23.60)	−75.938 71 (−11.48)
2.6R_e		−75.408 77 (504.64)	−75.913 41	−75.789 38 (124.04)	−75.910 66 (2.75)	−75.910 31 (3.10)	−75.910 24 (3.17)	−75.905 47 (7.94)	−75.896 46 (16.95)	−75.943 97 (−30.56)
3.0R_e		−75.336 38 (573.65)	−75.910 03	−75.787 02 (123.01)	−75.907 02 (3.01)	−75.906 60 (3.43)	−75.907 06 (2.97)	−75.901 90 (8.13)	−75.900 36 (9.67)	−75.951 09 (−41.06)
3.4R_e		−75.285 62 (623.46)	−75.909 08	−75.786 37 (122.71)	−75.905 80 (3.28)	−75.905 47 (3.61)	−75.906 05 (3.03)	−75.900 91 (8.17)	−75.903 81 (5.27)	−75.955 33 (−46.25)
3.8R_e		−75.250 43 (658.36)	−75.908 78	−75.786 17 (122.61)	−75.905 60 (3.18)	−75.905 23 (3.55)	−75.905 92 (2.86)	−75.900 60 (8.18)	−75.905 98 (2.80)	−75.957 69 (−48.91)

partially neglected when the corresponding preconditioner element D_{rs}^{pq}, [Eq. (30)], is smaller than the truncation threshold 0.5. Note that this separation scheme may break the orbital invariance of the CT theory.

IV. NUMERICAL RESULTS

A. Simultaneous bond breaking of water molecule with 6-31G and cc-pVDZ basis sets

As a prototype multireference application, we performed calculations of potential curves for the symmetric breaking of the water molecule H_2O in which the two O–H bonds are stretched simultaneously. We used both the 6-31G (Ref. 52) and cc-pVDZ (Ref. 53) basis sets. The results of L-CTD and L-CTSD calculations, together with a number of conventional methods—Hartree-Fock (HF), full configuration interaction (FCI), multireference second-order perturbation theory (MRMP),[18,19] and coupled cluster theory (CCSD/CCSDT) presented in Tables III and IV. All multireference calculations used a CAS with six active electrons in five active orbitals [denoted by (6*e*,5*o*)]. The 1*s* orbital in O atom was held frozen in all calculations. The FCI and MRMP calculations were carried out using GAMESS,[54] and the CC calculations using the TCE (Ref. 55) implemented in UTCHEM.[56]

Figures 7 and 8 plot deviations of total energies from FCI results for the various methods. It is clear that the CASSCF/L-CTD theory performs best out of all the methods studied. (We recall that although the canonical transformation operator $\exp \hat{A}$ does not explicitly include single excitations, the main effects are already included via the orbital relaxation in the CASSCF reference.) The absolute error of the CASSCF/L-CTD theory at equilibrium: 1.57 mE_h (6-31G), 2.26 mE_h (cc-pVDZ) is slightly better than that of the CCSD theory: 1.66 mE_h (6-31G) and 3.84 mE_h (cc-pVDZ), but unlike for the CCSD and CCSDT theories, the CASSCF/L-CTD error stays quite constant as the molecule is pulled apart while the CC theories exhibit a nonphysical turnover and a qualitatively incorrect dissociation curve. The largest error for the CASSCF/L-CTD method occurs at the intermediate bond distance of 1.8R_e with an error of −2.34 mE_h (6-31G) and −2.42 mE_h (cc-pVDZ). Although the MRMP curve is qualitatively correct, it is not quantitatively correct especially in the equilibrium region, with an error of 6.79 mE_h (6-31G) and 14.78 mE_h (cc-pVDZ). One measure of the quality of a dissociation curve is the nonparallelity error (NPE), the absolute difference between the maximum and minimum deviations from the FCI energy. For MRMP

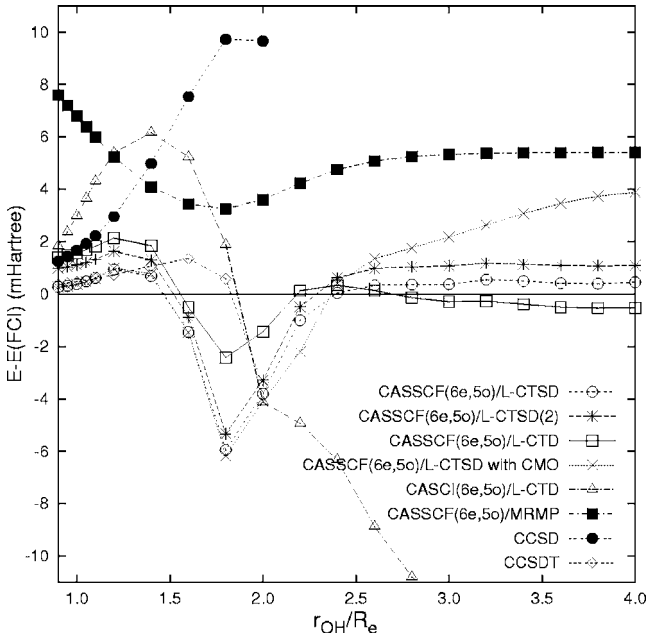

FIG. 7. Energy differences $E-E$(FCI) for the simultaneous bond breaking of H_2O molecule with 6-31G basis sets.

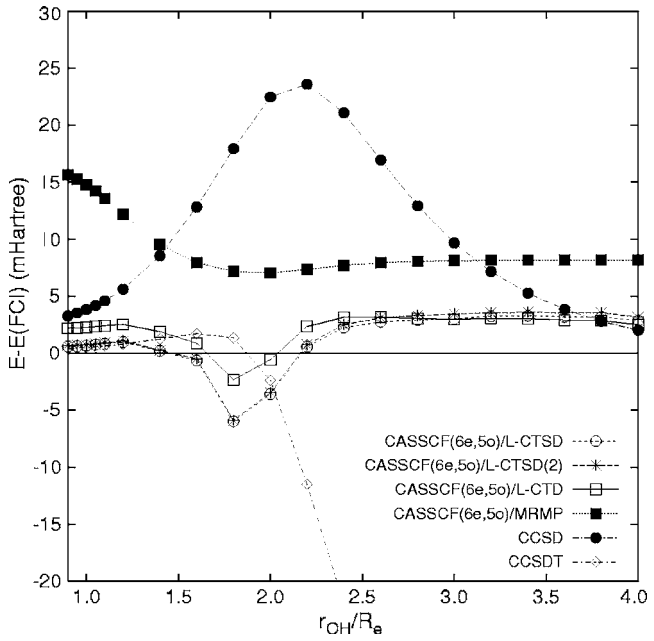

FIG. 8. Energy differences $E-E(\mathrm{FCI})$ for the simultaneous bond breaking of H_2O molecule with cc-pVDZ basis sets.

the NPE is 4 mE_h (6-31G) and 9 mE_h (cc-pVDZ), whereas for CASSCF/L-CTD the NPE is 5 mE_h (6-31G) and 6 mE_h (cc-pVDZ), showing that the CASSCF/L-CTD provides a quantitative description of the bond breaking with a nonparallelity error competitive with that of MRMP.

We now discuss the other CT calculations on water that are presented here. The CASSCF/L-CTSD method incorporates an additional orbital rotation on top of those contained in the CASSCF optimization, by the inclusion of one-particle operators in the excitation operator $\hat{A}$. Comparison of the CASSCF/L-CTSD with the CASSCF/L-CTD results shows that although the broad features of the potential energy curves are similar (small errors near in and far out, larger errors in the intermediate region), the absolute errors of CASSCF/L-CTSD are often larger than that of CASSCF/L-CTD. We suggest that this may arise from a lack of balance between the one-particle single excitations (which are always treated exactly) and the semi-internal single excitations, which are treated approximately both from the cumulant decomposition (Sec. II B) and due to the use of the numerical cutoff in solving the amplitude equations (Sec. III C). This cutoff also makes the CT calculation not orbital invariant (with respect to active-active, external-external rotations), and this is probably the reason for the observed differences between calculations using a CASSCF density matrix expressed in the CASSCF natural orbital basis, and a density matrix expressed in the canonical CASSCF orbital basis. Generally speaking, the canonical CASSCF orbital based calculations perform less well at longer bond lengths. Finally, the importance of choosing a reasonable set of active-space orbitals is reinforced by the CASSCF/L-CTD curve, which does not include any direct mechanism for orbital relaxation. The energy computed with this method breaks down in the intermediate region, past $1.8R_e$.

Table V lists the norms of internal, semi-internal and external excitation amplitudes obtained in the CASSCF$(6e,5o,\mathrm{NO})$/L-CTSD calculations with the 6-31G basis set. The percentage of retained amplitudes for internal and semi-internal excitations are also shown in the table. The norm of the external amplitudes, which primarily contribute to dynamic correlation, does not fluctuate much across the potential curve. The maximum of the internal and semi-internal amplitudes are found at the intermediate bond region. More internal and semi-internal excitation operators are retained as the bond length R_{OH} is increased.

We also measured the energy contributions from the different classes of excitation operators used in solving the amplitude equations (Sec. III C). The changes in the total energy occurring during the solution of the eight classes of amplitudes are shown in Table VI. The external excitation operators give the largest contribution to the correlation energy. The contributions from steps (iii) and (iv), which correlate non-HF configurations with the external orbitals, grow larger with the longer OH bond. This reflects the importance of a multiconfigurational description for dissociation. The energy contribution from semi-internal operators is significant, and is largest at the equilibrium structure.

Finally, to assess the convergence of the commutator expansion in the effective Hamiltonian as the bond is stretched, we computed a "second-order" energy using the L-CTSD amplitudes, denoted CASSCF/L-CTSD (2). Here the energy expression is evaluated as $\langle \hat{H}^{(0)}+\hat{H}^{(1)}+\hat{H}^{(2)}\rangle$. As seen from Tables III and IV and Figs. 7 and 8, the second-order energy curve faithfully follows the parent CASSCF/L-CTSD curve. This is promising for the development of hybrid CT-perturbation theories, along the lines of CC(2) theory.[57,58]

TABLE V. Norms and nontruncation ratios of amplitudes in CASSCF$(6e,5o)$/L-CTSD with NOs for H_2O molecule with 6-31G basis sets.

	norm $\times 10^2$			Nontruncation ratio	
r_{OH}	Internal	Semi-internal	External	Internal	Semi-internal
$1.0R_e$	0.57	2.88	5.07	16%	34%
$1.4R_e$	5.76	3.31	4.84	32%	65%
$1.8R_e$	3.70	6.41	4.77	64%	79%
$2.2R_e$	3.60	3.52	4.77	64%	80%
$3.8R_e$	1.87	1.87	4.69	64%	80%

TABLE VI. Energy changes (in millihartree) in the eight steps used in solving the CT equations. Values are from CASSCF($6e$,$5o$)/L-CTSD calculations with NOs for the H_2O molecule with 6-31G basis sets.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
r_{OH}	$ccaa$[a]	$ccvv+cavv+aavv$[b]			$ca+cv+av$[c]	$ccav+caav+aaav$[d]		
$1.0R_e$	−1.38	−43.88	−0.39	+0.01	+0.13	−37.11	+0.00	+0.00
$1.4R_e$	−2.08	−45.41	−1.91	−0.04	+0.16	−29.98	+0.05	−0.15
$1.8R_e$	−1.08	−44.82	−5.53	−0.19	+0.59	−19.61	−1.88	−4.90
$2.2R_e$	−0.40	−43.16	−9.85	−0.27	+0.80	−10.67	−0.50	−1.35
$3.8R_e$	−0.00	−41.05	−14.12	−0.29	+0.95	−5.48	+0.00	+0.04

[a]Step (i) solves the equations for *ccaa*.
[b]Steps (ii)–(iv) solve the equations for *ccvv*, *cavv*, and *aavv*.
[c]Step (v) solves the equations for one-particle operators.
[d]Steps (vi)–(viii) solve the equations for *ccav*, *caav*, and *aaav*.

B. Bond breaking of nitrogen molecules with 6-31G basis sets

The second application focuses on the dissociation of the triple bond in the nitrogen molecule N_2, which has been chosen by many multireference studies as a simple multireference model that is difficult to solve. For these calculations, we used the 6-31G basis set and froze the 1*s* orbitals. We chose two types of CAS space, ($6e$,$6o$) and ($10e$,$8o$). The triple bonding in N_2 is primarily formed by three valence orbitals, one σ and two π orbitals. The bond breaking is basically a chemical reaction that involves these three bonding orbitals and the corresponding higher-lying antibonding orbitals. The CAS($6e$,$6o$) is thus the smallest active space that allows for a qualitatively correct treatment of the bond breaking. We also look at CAS($10e$,$8o$), which includes the valence 2*s* occupied orbitals in addition to CAS($6e$,$6o$). This allows a correct description of the primary 2*s*-2*p* relaxation effects during bond breaking.

TABLE VII. Total energies (E_h) for the bond breaking of the N_2 molecule with 6-31G basis sets. The value in parentheses is the difference from the FCI total energy in mE_h.

r_{NN}	Method reference CAS orbitals	HF	CASSCF ($6e$,$6o$)	L-CTSD CASSCF ($6e$,$6o$)	L-CTSD CASSCF ($6e$,$6o$) NOs	L-CTSD(2) CASSCF ($6e$,$6o$) NOs	L-CTD CASSCF ($6e$,$6o$) NOs	L-CTD CASCI ($6e$,$6o$) NOs	MRMP CASSCF ($6e$,$6o$)
1.15 Å		−108.857 17	−109.017 61	−109.098 86	−109.098 86	−109.098 44	−109.099 93	−109.095 25	−109.090 51
		(248.75)	(88.31)	(7.06)	(7.06)	(7.48)	(5.99)	(10.67)	(15.41)
1.40 Å		−108.699 62	−108.929 73	−109.014 14	−109.014 16	−109.013 66	−109.014 65	−109.002 08	−109.007 56
		(322.44)	(92.33)	(7.92)	(7.90)	(8.40)	(7.41)	(19.98)	(14.50)
1.80 Å		−108.420 80	−108.796 15	−108.885 71	−108.885 71	−108.885 22	−108.881 90	−108.873 88	−108.878 82
		(467.23)	(91.88)	(2.32)	(2.32)	(2.81)	(6.13)	(14.15)	(9.21)
2.20 Å		−108.216 46	−108.766 57	−108.846 20	−108.846 20	−108.845 78	−108.844 69	−108.853 23	−108.840 01
		(631.10)	(80.99)	(1.36)	(1.36)	(1.78)	(2.87)	(−5.67)	(7.55)
2.60 Å		−108.076 48	−108.764 50	−108.839 61	−108.839 61	−108.838 95	−108.839 66	−108.859 87	−108.832 49
		(764.11)	(76.09)	(0.98)	(0.98)	(1.64)	(0.93)	(−19.28)	(8.10)
3.00 Å		−107.982 60	−108.764 36	−108.838 73	−108.838 73	−108.837 95	−108.839 34	−108.866 29	−108.830 82
		(855.29)	(73.53)	(−0.84)	(−0.84)	(−0.06)	(−1.45)	(−28.40)	(7.07)

r_{NN}	Method reference CAS orbitals	FCI	CASSCF ($10e$,$8o$)	L-CTSD CASSCF ($10e$,$8o$)	L-CTSD CASSCF ($10e$,$8o$) NOs	L-CTSD(2) CASSCF ($10e$,$8o$) NOs	MRMP CASSCF ($10e$,$8o$)	CCSD	CCSDT
1.15 Å		−109.105 92	−109.031 25	−109.102 40	−109.102 40	−109.101 85	−109.097 25	−109.094 89	−109.103 55
			(74.67)	(3.52)	(3.52)	(4.07)	(8.67)	(11.03)	(2.37)
1.40 Å		−109.022 06	−108.944 10	−109.017 23	−109.017 24	−109.016 71	−109.013 89	−108.999 78	−109.015 75
			(77.96)	(4.83)	(4.82)	(5.35)	(8.17)	(22.28)	(6.31)
1.80 Å		−108.888 03	−108.804 16	−108.887 24	−108.887 19	−108.886 64	−108.882 03	−108.852 93	−108.887 70
			(83.87)	(0.79)	(0.84)	(1.39)	(6.00)	(35.10)	(0.33)
2.20 Å		−108.847 56	−108.768 65	−108.846 65	−108.846 68	−108.846 25	−108.841 24	−108.920 53	−108.961 90
			(78.91)	(0.91)	(0.88)	(1.31)	(6.32)	(−72.97)	(−114.34)
2.60 Å		−108.840 59	−108.765 12	−108.839 34	−108.839 66	−108.838 87	−108.832 96		
			(75.47)	(1.25)	(0.93)	(1.72)	(7.63)		
3.00 Å		−108.837 89	−108.764 56	−108.838 15	−108.838 87	−108.837 62	−108.831 03		
			(73.33)	(−0.26)	(−0.98)	(0.27)	(6.86)		

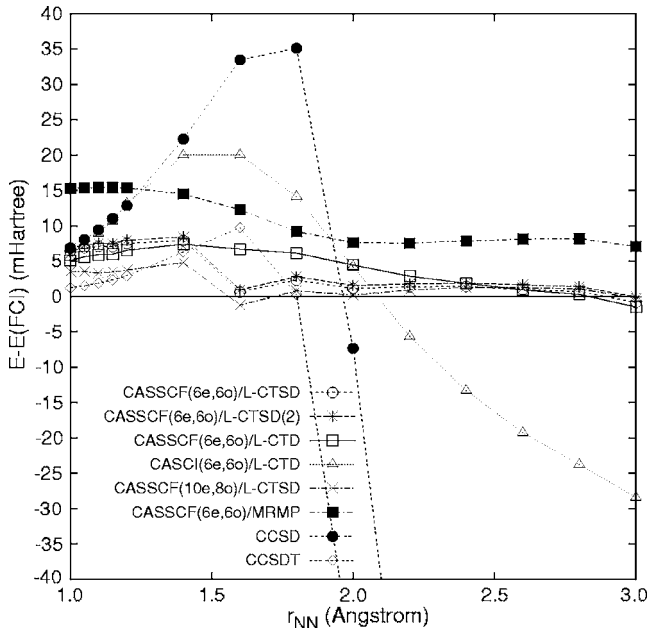

FIG. 9. Energy differences $E-E(\mathrm{FCI})$ for the bond-breaking curve of the N_2 molecule with 6-31G basis sets.

Table VII shows the total energies as a function of bond length R_{NN} using several CT methods, as well as those using the HF, FCI, CASSCF, MRMP, CCSD, and CCSDT methods. Figure 9 plots the energy differences from the FCI results.

Comparing the different methods we see once again that the CASSCF/L-CTD method yields the most accurate description of the potential energy curve out of all the theories. The error at equilibrium (5.99 mE_h) is better than that of CCSD (11.03 mE_h) and once again this error stays roughly constant across the curve, while that of the CC based approaches exhibit a nonphysical turnover. For comparison, the MRMP error at equlibrium is 15.41 mE_h. The nonparallelity errors for CASSCF/L-CTD and MRMP are 8.9 and 8.3 mE_h, respectively, demonstrating again that CASSCF/L-CTD yields quantitatively accurate curves with NPEs competitive with that of MRMP theory.

A source of error in the CASSCF$(6e,6o)$ based methods is an incomplete treatment of the active-core relaxation. Although some effects of active-core relaxation are incorporated via the exponential operator in the CT calculations, this is incomplete due to the truncation of some operators in the *ccaa* class as explained in Sec. III C. Comparing CAS$(10e,8o)$ with CAS$(6e,6o)$ shows us the effects of the truncation. At the equilibrium structure (R_{NN}=1.15 Å), we observe that the L-CTSD energy with CAS$(6e,6o)$ is 3.5 mE_h higher than that with CAS$(10e,8o)$. For comparison the MRMP energy with CAS$(6e,6o)$ is 6.7 mE_h higher than that with CAS$(10e,8o)$. Thus the truncated *ccaa* operators capture much, but not all of the active-core relaxation. In the region far from the equilibrium structure, the discrepancy in energies between the two types of CAS disappears, as the active-core relaxation is less important.

In the nitrogen molecule, we found no significant difference in the CT energies between using NOs and canonical CAS orbitals over the potential curve, unlike in the H_2O case. However, it is still important to optimize the orbitals, as the CASCI-based L-CTD does not yield a correct potential curve. Finally, as in the water calculations, the L-CTSD(2) approximation recovers most of L-CTSD correlation energy across the potential energy curve.

C. Comparison with MR-CISD and MR-LCCM on the two-configuration reference insertion of Be in H_2 molecule

Several authors have studied the insertion of Be in H_2 as an example of a true two-configuration multireference problem.[19,26,59] Laidig and Bartlett presented a multireference coupled cluster method which they called MR-LCCM, in which they applied a linearized form of coupled cluster theory to a two-configuration reference.[59] Table VIII shows the multireference results obtained in that work using the MR-LCCM and multireference configuration interaction (MR-CISD) methods at three structures in the C_{2v} insertion of Be into H_2. The details of MR-CISD and MR-LCCM calculations are described in Ref. 59. For comparison, the corresponding CASSCF-based linearized CT calculations with the same Gaussian basis set are also presented. Figure 10 plots the total energies.

Comparing the different calculations, we see that MR-LCCM generally overestimates the correlation energy, while MR-CISD generally underestimates the correlation energy. The CASSCF/L-CTD method yields the best energies out of

TABLE VIII. BeH_2 energies ($^1A_1$ state) at three selected geometries (Bohr). The value in parentheses is the difference from the FCI total energy in mE_h.

Points	r_{Be-H_2}	r_{HH}	Method reference CAS orbitals	HF	FCI	CASSCF $(2e,2o)$	L-CTSD CASSCF $(2e,2o)$ NOs	L-CTD CASSCF $(2e,2o)$ NOs	MR-CISD[c] CASSCF $(2e,2o)$	MR-LCCM[c] CASSCF $(2e,2o)$
(1)	2.5	2.78		−15.562 68[a] (60.20)	−15.622 88	−15.569 57 (53.31)	−15.622 20 (0.68)	−15.622 10 (0.78)	−15.622 04 (0.84)	−15.625 50 (−2.62)
(2)	2.75	2.55		−15.521 19[a] (81.73)	−15.602 92	−15.538 57 (64.35)	−15.598 06 (4.86)	−15.603 07 (−0.15)	−15.600 91 (2.01)	−15.605 32 (−2.40)
(3)	3.0	2.32		−15.536 47[b] (88.49)	−15.624 96	−15.558 28 (66.68)	−15.627 23 (−2.27)	−15.626 19 (−1.23)	−15.621 89 (3.01)	−15.630 46 (−5.50)

[a]Configuration: $1a_1^2 2a_1^2 1b_2^2$.
[b]Configuration: $1a_1^2 2a_1^2 1a_1^2$.
[c]Reference 59.

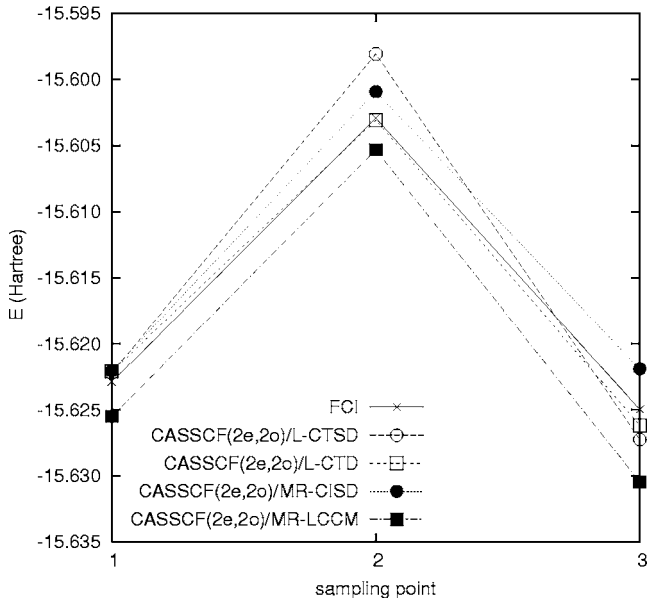

FIG. 10. Plots of total energies of BeH_2 ($^1A_1$ state) at three structures.

all the methods, at all points, with a maximum absolute error of 1.23 mE_h. In agreement with our previous findings, the CASSCF/L-CTSD method is less accurate, with a maximum error of 4.86 mE_h. The nonparallelity errors are 2.0 mE_h (CASSCF/L-CTD), 2.2 mE_h (MR-CISD), and 3.1 mE_h (MR-LCCM). From this (admittedly small) sample of results, we can say that the CASSCF/L-CTD theory outperforms both the MR-LCCM and MR-CISD methods.

D. Single- and multireference linearized CT for HF and BH molecules

We have so far examined the performance of the canonical transformation theory when paired with a suitable multireference wave function, such as the CASSCF wave function. As we have argued, because the exponential operator describes dynamic correlation, this hybrid approach is the way in which the theory is intended to be used in general bonding situations. However, we can also examine the behavior of the single-reference version of the theory (i.e., using the Hartree-Fock reference). In this way, we can compare in detail with the related single-reference coupled cluster theory, and make contact with the perturbative analysis of the cumulant decomposition in Sec. II D.

In Tables IX and X we present calculations using single-reference linearized (HF/CTD) and (HF/CTSD) theories on several geometries of the HF and BH molecules. Note the HF/CTD theory does not include any semi-internal two-particle excitations, since the active space is completely occupied. The results of calculations using both standard CC theories (CCSD, CCSDT, and linearized CCSD), as well as two alternative coupled-cluster theories called the expectation value CC (*X*CC) method and the unitary CC (UCC) method (introduced in Refs. 29, 48, and 59), are also shown. The CC energies were obtained from Ref. 29. CASSCF/L-CTD and CASSCF/L-CTSD calculations are presented for comparison. All calculations used the Dunning DZP basis sets and the geometries described in Ref. 29

TABLE IX. Energy differences $E-E$(FCI) for the HF molecule. The units are mE_h. FCI energy: $E(1.0R_e)=-100.250\,97\ E_h$, $E(1.5R_e)=-100.160\,39\ E_h$, and $E(2.0R_e)=-100.081\,11\ E_h$.

r_{HF}	Method reference CAS orbitals	HF	CCSD[a]	CCSDT[a]	L-CCSD	XCCSD(4)	XCC(4)[a]	UCCSD(4)	UCC(4)[a]	L-CTSD HF	L-CTD HF	CASSCF $(2e,2o)$	L-CTSD CASSCF $(2e,2o)$	L-CTD CASSCF $(2e,2o)$
$1.0R_e$[b]		203.88	3.01	0.26	0.81	2.65	−0.38	−0.96	−4.11	2.87	0.46	180.78	1.67	2.01
$1.5R_e$[b]		227.17	5.10	0.65	−2.45	4.62	−0.51	2.25	−3.00	7.00	0.66	173.97	1.85	2.30
$2.0R_e$[b]		263.54	10.18	1.12	−47.34	8.60	−3.00	8.51	−2.69	11.93	3.59	160.12	−0.93	1.42

[a]The numbers for CCSD, CCSDT, XCC, and UCC are from Ref. 29.
[b]$R_e=1.733$ bohr.

TABLE X. Energy differences $E-E(\mathrm{FCI})$ for the BH molecule. The units are mE_h. FCI energy: $E(1.0R_e)=-25.227\,63\ E_h$, $E(1.5R_e)=-25.175\,98\ E_h$, and $E(2.0R_e)=-25.127\,35\ E_h$.

r_{HF}	Method reference CAS orbitals	HF	CCSD[a]	CCSDT[a]	L-CCSD	XCCSD(4)	XCC(4)[a]	UCCSD(4)	UCC(4)[a]	L-CTSD HF	L-CTD HF	CASSCF $(4,3o)$	L-CTSD CASSCF $(4e,3o)$	L-CTD CASSCF $(4e,3o)$
$1.0R_e$[b]		102.37	1.78	0.06	−3.86	1.99	0.65	1.92	0.57	−4.87	−2.92	95.98	−2.43	−0.21
$1.5R_e$[b]		113.76	2.64	0.02	−6.44	2.94	0.82	2.81	0.67	−8.33	−3.38	80.42	−3.38	−1.72
$2.0R_e$[b]		139.15	5.04	0.02	−26.74	5.60	0.91	5.01	0.17	−35.32	−11.89	77.08	−5.48	−3.62

[a]The numbers for CCSD, CCSDT, XCC, and UCC are from Ref. 29.
[b]$R_e=2.329$ bohr.

In both the HF molecule and BH molecule, the HF/L-CTSD and HF/L-CTD theories give energies comparable to CCSD at the equilibrium geometry. However, as the bond is stretched, they both display significantly increased errors, typical of a single reference theory. At stretched geometries, the errors of the HF/L-CTSD method are worse than those of CCSD and comparable to those of the linearized CCSD theory.

The perturbative analysis in Sec. II D showed that single-reference L-CTSD is exact through third order in the fluctuation potential, much like the L-CCSD theory, and our results are consistent with this analysis. This suggests that one of the things we can do to improve the linearized CT results would be to treat properly some of the higher-order particle terms in E^4 [see Eq. (21)] which the present linearized theories approximate using the cumulant expansion. The importance of these high-order terms is lessened with a multi-reference starting point, as illustrated by the CASSCF/L-CTSD and CASSCF/L-CTD results, where even a very small active space restores the correct quantitative behavior across the entire potential energy curve, as we have found in our earlier calculations.

V. CONCLUSIONS

We have proposed a canonical transformation (CT) theory to describe the dynamic correlation in bonding situations where there is also significant nondynamic character. By pairing this theory with a suitable multireference description of the nondynamic correlation, such as provided by the CASSCF wave function, we have obtained consistently quantitative descriptions for a variety of molecules over a wide range of different geometries. The best performing method we have found is the linearized CT with doubles model with a CASSCF reference (CASSCF/L-CTD). Using this method, the accuracy obtained is comparable to or better than that of CCSD theory in the equilibrium region of a potential energy curve, but unlike in coupled cluster theories, this accuracy persists all the way out to bond dissociation. The CASSCF/L-CTD nonparallelity errors are competitive with, or better than those obtained with multireference perturbation theory.

In addition to the encouraging numerical results, the canonical transformation theory has a number of appealing formal features. It is based on a unitary exponential, and is therefore a Hermitian theory; it is size consistent; and it has a cost comparable to that of single-reference coupled cluster theory. Cumulants are used in two places in the theory, to close the commutator expansion of the unitary exponential, and to decouple the complexity of the multireference wave function from the treatment of dynamic correlation.

There are a number of clear directions to improve the current theory. A perturbative analysis demonstrates that the linearized CT method used here is an approximation to a "full" CT theory, since it neglects certain higher-order terms, in much the same way that linearized coupled cluster theory neglects certain terms in the full coupled cluster equations. In addition, the numerical solution of the CT equations can be challenging due to the presence of low energy intruder excitations. These and other topics are currently under active investigation.

ACKNOWLEDGMENTS

We would like to thank S. R. White for useful discussions and Cornell University for financial support.

[1] B. O. Roos, Adv. Chem. Phys. **69**, 399 (1987).
[2] K. Ruedenberg, M. W. Schmidt, M. M. Gilbert, and S. T. Elbert, Chem. Phys. **71**, 41 (1982).
[3] S. R. White and R. L. Martin, J. Chem. Phys. **110**, 4127 (1998).
[4] A. O. Mitrushenkov, G. Fano, F. Ortolani, R. Linguerri, and P. Palmieri, J. Chem. Phys. **115**, 6815 (2001).
[5] A. O. Mitrushenkov, R. Linguerri, P. Palmieri, and G. Fano, J. Chem. Phys. **119**, 4148 (2003).
[6] G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002).
[7] G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **118**, 8551 (2003).
[8] G. K.-L. Chan, J. Chem. Phys. **120**, 3172 (2004).
[9] G. K.-L. Chan, M. Ka'llay, and J. Gauss, J. Chem. Phys. **121**, 6110 (2004).
[10] G. K.-L. Chan and T. V. Voorhis, J. Chem. Phys. **122**, 204101 (2005).
[11] Ö. Legeza, J. Röder, and B. A. Hess, Mol. Phys. **101**, 2019 (2003).
[12] Ö. Legeza, J. Röder, and B. A. Hess, Phys. Rev. B **67**, 125114 (2003).
[13] J. Cizek, J. Chem. Phys. **45**, 4256 (1969).
[14] J. Paldus, J. Cizek, and I. Shavitt, Phys. Rev. A **5**, 50 (1972).
[15] R. J. Bartlett, Annu. Rev. Phys. Chem. **32**, 359 (1981).
[16] J. Paldus and X. Li, Adv. Chem. Phys. **110**, 1 (1999).
[17] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990).
[18] K. Hirao, Chem. Phys. Lett. **190**, 374 (1992).
[19] H. Nakano, J. Chem. Phys. **99**, 7983 (1993).
[20] R. K. Chaudhuri, K. F. Freed, G. Hose *et al.*, J. Chem. Phys. **122**, 134105 (2005).
[21] S. R. White, J. Chem. Phys. **117**, 7472 (2002).
[22] T. Kinoshita, O. Hino, and R. J. Bartlett, J. Chem. Phys. **123**, 074106 (2005).
[23] O. Hino, T. Kinoshita, G. K.-L. Chan, and R. J. Bartlett, J. Chem. Phys. **124**, 114311 (2006).
[24] K. F. Freed, *Many-Body Methods in Quantum Chemistry*, edited by U. Kaldor (Springer, Berlin, 1989), p. 1.
[25] B. Kirtman, J. Chem. Phys. **75**, 798 (1981).

[26] M. R. Hoffmann and J. Simons, J. Chem. Phys. **88**, 993 (1988).
[27] W. Kutzelnigg, J. Chem. Phys. **77**, 3081 (1982).
[28] W. Kutzelnigg, J. Chem. Phys. **80**, 822 (1984).
[29] J. D. Watts, G. W. Trucks, and R. J. Bartlett, Chem. Phys. Lett. **157**, 359 (1989).
[30] R. J. Bartlett, S. A. Kucharski, and J. Noga, Chem. Phys. Lett. **155**, 133 (1989).
[31] S. Pal, M. D. Prasad, and D. Mukherjee, Theor. Chim. Acta **62**, 523 (1983).
[32] S. Pal, Theor. Chim. Acta **66**, 207 (1984).
[33] K. F. Freed, *Renormalization Group Theory of Macromolecules* (Wiley Interscience, New York, 1987).
[34] F. Wegner, Ann. Phys. **3**, 77 (1994).
[35] S. D. Glazek and K. G. Wilson, Phys. Rev. D **49**, 4214 (1994).
[36] A. Anderson, Phys. Lett. B **319**, 157 (1993).
[37] W. Kutzelnigg, Chem. Phys. Lett. **64**, 383 (1979).
[38] M. Nooijen, Phys. Rev. Lett. **84**, 2108 (2000).
[39] H. Nakatsuji, J. Chem. Phys. **113**, 2949 (2000).
[40] T. V. Voorhis and M. Head-Gordon, J. Chem. Phys. **115**, 5033 (2001).
[41] P. Piecuch, K. Kowalski, P.-D. Fan, and K. Jedziniak, Phys. Rev. Lett. **90**, 113001 (2003).
[42] F. Colmenero and C. Valdemoro, Phys. Rev. A **47**, 979 (1993).
[43] F. Colmenero and C. Valdemoro, Int. J. Quantum Chem. **51**, 369 (1994).
[44] K. Yasuda and H. Nakatsuji, Phys. Rev. Lett. **76**, 1039 (1996).
[45] H. Nakatsuji and K. Yasuda, Phys. Rev. A **56**, 2648 (1997).
[46] D. A. Mazziotti, Phys. Rev. A **57**, 4219 (1998).
[47] D. A. Mazziotti, Chem. Phys. Lett. **289**, 419 (1998).
[48] R. J. Bartlett and J. Noga, Chem. Phys. Lett. **150**, 29 (1988).
[49] R. J. Bartlett, S. A. Kucharski, J. Noga, J. D. Watts, and G. W. Trucks, in *Many-Body Methods in Quantum Chemistry*, edited by U. Kaldor (Springer, Berlin, 1989), p. 125.
[50] R. J. Bartlett, Adv. Quantum Chem. **18**, 281 (1986).
[51] Y. Saad and M. Schultz, SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **7**, 856 (1986).
[52] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).
[53] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).
[54] M. W. Schmidt, K. K. Boldridge, J. Å. Boatz *et al.*, J. Comput. Chem. **14**, 1347 (1993).
[55] S. Hirata, J. Phys. Chem. A **107**, 4940 (2003).
[56] T. Yanai *et al.*, Utchem—a program for *ab initio* quantum chemistry, in *Lecture Notes in Computer Science* (Springer-Verlag, Berlin, 2003).
[57] S. Hirata, M. Nooijen, I. Grabowski, and R. J. Bartlett, J. Chem. Phys. **114**, 3919 (2001).
[58] S. R. Gwaltney and M. Head-Gordon, J. Chem. Phys. **115**, 2014 (2001).
[59] W. D. Laidig and R. J. Bartlett, Chem. Phys. Lett. **104**, 424 (1984).