Canonical transformation theory from extended normal ordering

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

The canonical transformation theory of Yanai and Chan [J. Chem. Phys. **124**, 194106 (2006)] provides a rigorously size-extensive description of dynamical correlation in multireference problems. Here we describe a new formulation of the theory based on the extended normal ordering procedure of Mukherjee and Kutzelnigg [J. Chem. Phys. **107**, 432 (1997)]. On studies of the water, nitrogen, and iron oxide potential energy curves, the linearized canonical transformation singles and doubles theory is competitive in accuracy with some of the best multireference methods, such as the multireference averaged coupled pair functional, while computational timings (in the case of the iron oxide molecule) are two to three orders of magnitude faster and comparable to those of the complete active space second-order perturbation theory. The results presented here are greatly improved both in accuracy and in cost over our earlier study as the result of a new numerical algorithm for solving the amplitude equations. © *2007 American Institute of Physics.*
[DOI: 10.1063/1.2761870]

I. INTRODUCTION

While nondynamic correlation between electrons establishes the qualitative features of chemical bonding, it is the accurate description of dynamic correlation, associated with the short-range cusp behavior of the wave function, which is necessary to obtain quantitative agreement with experiment. Starting from a suitable reference function, the exponential ansatz provides an accurate and economical description of dynamic correlation. For example, in systems that are qualitatively described by a single determinant reference, coupled cluster (CC) theory paired with a large basis set yields predictions with chemical accuracy.[1–3] However, for the many chemical problems which require a multireference characterization, a practical theory for dynamic correlation with the desirable qualities of the exponential ansatz—size extensivity, chemical accuracy, and moderate computational cost—has yet to be widely established.

In an earlier article,[4] we presented a canonical transformation (CT) theory which is based on an exponential ansatz, is rigorously size extensive, and which may easily be combined with any multireference starting wave function. In the form implemented in that work, the computational cost is $O(a^2 e^4)$, where a is the number of active orbitals and e is the number of external orbitals. In calculations of bond-breaking potential energy curves, the linearized canonical transformation doubles (L-CTD) theory performed significantly better than the multireference perturbation theory and obtained the accuracy of coupled cluster single doubles (CCSD) at the equilibrium geometry across the *entire* potential energy curve. Our work was directly motivated by the canonical diagonalization theory of White,[5] although there are earlier related contributions, as we will describe below.

The purpose of the current work is to improve on our initial contribution in several areas. A central feature of the canonical transformation theory is the use of an *operator decomposition*, both to close the infinite expansions associated with an exponential ansatz and to reduce the complexity of the energy and amplitude equations that arise when working with a complicated reference function. In our earlier work, we introduced a cumulant-type operator decomposition by analogy to the cumulant decomposition of density matrices found in reduced density matrix theories.[6–11] However, this choice of operator decomposition is not unique and here we explore an alternative operator decomposition, with some formal advantages, that is based on the concept of *extended normal ordering* as introduced by Mukherjee and Kutzelnigg.[12–14] Indeed, examination of the articles by these authors shows that they anticipated the utility of their results in multireference correlation theories, and in this context our current theory is in part a realization along such directions.

A second focus of this work is to investigate in detail the behavior of the canonical transformation theory in a variety of chemical problems. For example, we study, with a range of basis sets, the bond-breaking potential energy curves of water, nitrogen, and iron oxide and compare our results against state-of-the-art multireference configuration interaction and perturbation theories. In addition, we examine numerically the size-extensivity and density-scaling properties of the canonical transformation energies. The results in the present study are much improved over our earlier work, in large part, because of improvements we have made to our numerical algorithms, and we describe in detail the numerical aspects of efficiently implementing and converging the CT equations.

---

a)Present address: Department of Theoretical and Computational Molecular Science, Institute for Molecular Science, Okazaki, Aichi 444-8585, Japan; Electronic mail: yanait@ims.ac.jp

II. CANONICAL TRANSFORMATION THEORY

A. Recapitulation

In multireference problems, we divide the orbitals into *active orbitals*, which describe the nondynamic correlation and *external orbitals* which describe the dynamic correlation. The external orbitals may further be divided into core and virtual orbitals; core orbitals are those which remain doubly occupied in all the reference configurations.

We will assume that a reference wave function Ψ_0 is available that describes the nondynamic correlation in the problem. This may be obtained, for example, from a complete active space self-consistent field (CASSCF) calculation that exactly correlates electrons within the active orbitals.[15,16] Alternatively, and especially for larger active spaces, a density matrix renormalization group wave function may be used.[17,18] We then incorporate the remaining dynamic correlation on top of the reference wave function Ψ_0 via an exponential operator that generates excitations between the active and external spaces, yielding

$$\Psi = e^A \Psi_0. \tag{1}$$

We will be concerned with a *unitary* formulation, where $A^\dagger = -A$. The excitations are understood to be both of external and semi-internal forms,

$$A = A_i^a(a_i^a - a_a^i) + A_{ij}^{ab}(a_{ij}^{ab} - a_{ab}^{ij}) + A_{ij}^{ak}(a_{ij}^{ak} - a_{ak}^{ij}) + \cdots, \tag{2}$$

where $ijk\cdots$ denote active indices, $abc\cdots$, external indices, $a_i^a = a_a^\dagger a_i$, $a_{ij}^{ab} = a_a^\dagger a_b^\dagger a_j a_i$, and the summation convention is assumed. For example, the first two terms are the usual external single and double excitations, while the third term (with three active indices) is a semi-internal single excitation, which captures the coupling between singles relaxation in the active space and singles excitation to the external space.

In a related picture, we can also view e^A as generating an effective *canonically transformed* Hamiltonian $\bar{H}$ that acts only in the active space, but which has dynamic correlation folded in from the external space, where

$$\bar{H} = e^{-A} H e^A, \tag{3}$$

$$\bar{H}\Psi = E\Psi. \tag{4}$$

The exponential ansatz combined with a multireference wave function Ψ_0, as shown in Eq. (1), has a long history and we necessarily can only give an incomplete account here. Such an ansatz is used in some forms of multireference coupled cluster theory, as discussed in the review by Paldus and Li.[19] In particular, an early example of a complete theoretical scheme for a related multireference coupled cluster method was given by Mukherjee in Ref. 12. While CC theory is usually formulated in terms of similarity rather than canonical (i.e., unitary) transforms, unitary exponentials have previously been explored in a multireference setting by Freed,[20] Kirtman,[21] and Hoffman and Simons.[22] We mention also the single-reference unitary coupled cluster work by Kutzelnigg,[23,24] Bartlett and co-workers,[25–27] and Pal.[28,29] The general concept of effective Hamiltonians and canonical transformations is of course very old, dating back to van Vleck.[30] We note, in particular, some modern theories that emphasize an effective Hamiltonian language similar to our own such as the effective valence hamiltonian theory of Freed[20] and the generalized van Vleck theory of Kirtman.[21] As recognized by Freed, the folding in of dynamic correlation into the active-space effective Hamiltonian is a form of renormalization transformation. This picture was pursued by White in his theory of canonical diagonalization[5] and, as described previously, this is the primary precursor to our work.

In the exponential ansatz of single-reference coupled cluster theory, the commutativity of the excitation operators in the single-reference form of A, i.e., $A = A_i^a a_i^a + A_{ij}^{ab} a_{ij}^{ab}$, allows the Baker-Campbell-Hausdorff expansion of $\bar{H}$ to terminate at low order for low-particle rank in A. The difficulty in working with the multireference exponential ansatz arises from the noncommuting excitations in the multireference form of A in Eq. (2), which leads to a nonterminating expansion for the effective Hamiltonian $\bar{H}$. (In fact, this difficulty already arises if we use a *unitary* e^A with the single-reference form of A.)

In our earlier CT theory, we introduced a new route to a tractable and computationally efficient formulation for the multireference ansatz [Eq. (1)]. Starting from the Baker-Campbell-Hausdorff expansion of the exact effective Hamiltonian,

$$\bar{H} = H + [H,A] + \tfrac{1}{2}[[H,A],A] + \cdots, \tag{5}$$

we replace each commutator by an approximate *decomposed* commutator to yield an approximate effective Hamiltonian,

$$\bar{H}_{1,2,\ldots} = H + [H,A]_{1,2,\ldots} + \tfrac{1}{2}[[H,A]_{1,2,\ldots},A]]_{1,2,\ldots} + \cdots. \tag{6}$$

Each subscript denotes a decomposition, and the numbers $1,2,\ldots$ denote the particle ranks of the operators that remain after the decomposition. Note that if *all* particle ranks were included in the decomposition (i.e., the subscripts ranged from 1 to n, where n is the number of particles), then Eqs. (5) and (6) would be identical. If in addition to including all particle ranks in Eq. (6) A contained up to n-body excitations, then the CT ansatz [Eq. (1)] would be exact in the sense of full configuration interaction, and indeed Eq. (4) would hold exactly. The two relevant approximations thus arise from restricting the excitations in A (wave function ansatz) as well as the form of the operator decomposition (operator ansatz).[55]

As an example, let us consider the linearized CT single and doubles theory (L-CTSD) introduced in our earlier work. Here A is restricted to contain only one- and two-particle excitations as in Eq. (2), and we restrict all decomposed commutators to contain at most one- and two-body operators (i.e., subscripts 1,2). Since $[H,A]$ generates a three-body operator, this requires some decomposition of a three-body operator into lower-body operators. We proposed an explicit decomposition into one- and two-body operators based on an analogy to the cumulant decomposition of density matrices,

$$a_{stu}^{pqr} \Rightarrow 9(\gamma_s^p \wedge a_{tu}^{qr}) - 12(\gamma_s^p \wedge \gamma_t^q \wedge a_u^r), \tag{7}$$

where in the above $\wedge$ denotes an antisymmetrization over all upper and lower indices with an associated factor of $1/(p)^2$, i.e., $1/36$ in the above case, where p is the particle rank of the original operator. Here we will present the explicit steps leading to the above decomposition. Our notation follows closely that of Kutzelnigg and Mukherjee.[31] Recall that the cumulant decomposition provides a way to rewrite reduced density matrices γ in terms of products of cumulants λ, via

$$\gamma_s^p = \langle a_s^p \rangle = \lambda_s^p, \tag{8}$$

$$\gamma_{st}^{pq} = \langle a_{st}^{pq} \rangle = \lambda_{st}^{pq} + \gamma_s^p \gamma_t^q - \gamma_t^p \gamma_s^q, \tag{9}$$

$$\begin{aligned}
\gamma_{stu}^{pqr} = \langle a_{stu}^{pqr} \rangle = &\lambda_{stu}^{pqr} + \gamma_s^p \lambda_{tu}^{qr} - \gamma_t^p \lambda_{su}^{qr} + \gamma_u^p \lambda_{st}^{qr} - \gamma_s^q \lambda_{tu}^{pr} \\
&+ \gamma_t^q \lambda_{su}^{pr} - \gamma_u^q \lambda_{st}^{pr} + \gamma_s^r \lambda_{tu}^{pq} - \gamma_t^r \lambda_{su}^{pq} + \gamma_u^r \lambda_{st}^{pq} \\
&+ \gamma_s^p \gamma_t^q \gamma_u^r - \gamma_s^p \gamma_t^r \gamma_u^q + \gamma_s^q \gamma_t^r \gamma_u^p - \gamma_s^q \gamma_t^p \gamma_u^r \\
&+ \gamma_s^r \gamma_t^p \gamma_u^q - \gamma_s^r \gamma_t^q \gamma_u^p.
\end{aligned} \tag{10}$$

For the three-particle density matrix, by dropping the three-particle cumulant λ_{stu}^{pqr} and substituting expressions (8) and (9) in Eq. (10), we obtain an approximate decomposition in terms of one- and two-particle density matrices only,

$$\begin{aligned}
\gamma_{stu}^{pqr} \Rightarrow &\gamma_s^p \gamma_{tu}^{qr} - \gamma_t^p \gamma_{su}^{qr} + \gamma_u^p \gamma_{st}^{qr} - \gamma_s^q \gamma_{tu}^{pr} + \gamma_t^q \gamma_{su}^{pr} - \gamma_u^q \gamma_{st}^{pr} \\
&+ \gamma_s^r \gamma_{tu}^{pq} - \gamma_t^r \gamma_{su}^{pq} + \gamma_u^r \gamma_{st}^{pq} - 2(\gamma_s^p \gamma_t^q \gamma_u^r - \gamma_s^p \gamma_t^r \gamma_u^q \\
&+ \gamma_s^q \gamma_t^r \gamma_u^p - \gamma_s^q \gamma_t^p \gamma_u^r + \gamma_s^r \gamma_t^p \gamma_u^q - \gamma_s^r \gamma_t^q \gamma_u^p) \\
= &9(\gamma_s^p \wedge \gamma_{tu}^{qr}) - 12(\gamma_s^p \wedge \gamma_t^q \wedge \gamma_u^r).
\end{aligned} \tag{11}$$

To obtain our operator decomposition, we simply replaced expectation values in the above terms by the corresponding operators, i.e., $\lambda_{st}^{pq} \rightarrow a_{st}^{pq}$ and $\lambda_s^p \rightarrow a_s^p$, yielding Eq. (7). Note that by construction, the expectation value of the operator decomposition reproduces the three-particle density matrix cumulant decomposition [Eq. (11)].

By using this decomposition recursively, i.e., by constructing the double commutator by first using the decomposed single commutator $[H,A]_{1,2}$ as in Eq. (6), the full effective Hamiltonian $\bar{H}_{1,2}$ at the L-CTSD level *contains only one-and two-body operators*. Evaluation of the energy then only requires the one- and two-particle density matrices of the reference function. As discussed in our initial work, this fulfills one of the criteria for an efficient multireference theory, namely, we do not need to explicitly manipulate the complicated reference function. From a different perspective, the canonical transformations can also be viewed as providing a parametrization of a two-particle density matrix theory. Recently, such connections have been explored from a different direction by Mazziotti[32,33] and while interesting, we shall not dwell further on these matters here.

We call the above formulation a linearized theory because the operator decomposition is applied at the first commutator. Then, at the L-CTSD level, the energies and amplitudes are evaluated via

$$E = \langle \Psi_0 | \bar{H}_{1,2} | \Psi_0 \rangle, \tag{12}$$

$$0 = \langle \Psi_0 | [\bar{H}_{1,2}, a_i^a - a_a^i]_{1,2} | \Psi_0 \rangle, \tag{13}$$

$$0 = \langle \Psi_0 | [\bar{H}_{1,2}, a_{ij}^{ab} - a_{ab}^{ij}]_{1,2} | \Psi_0 \rangle, \tag{14}$$

$$0 = \langle \Psi_0 | [\bar{H}_{1,2}, a_{ij}^{ak} - a_{ak}^{ij}]_{1,2} | \Psi_0 \rangle. \tag{15}$$

The resulting computational cost of the theory is $O(a^2 e^4)$ and is thus comparable to that of a *single-reference* coupled cluster calculation.

B. Accuracy of the operator decomposition

As presented above, the accuracy of the canonical transformation theory rests on the accuracy of operator decomposition, given at the L-CTSD level by Eq. (7). However, although our operator decomposition was chosen so that its expectation value would reproduce the density matrix cumulant decomposition, *this choice is not unique*. For example, we could add to the right-hand side of Eq. (7) any term with vanishing expectation value with Ψ_0 and still preserve the correspondence with the density matrix cumulant decomposition [Eq. (11)]. This simply reflects the fact that a decomposition for expectation values (i.e., the cumulant decomposition) does not contain sufficient information to specify a corresponding operator decomposition.

In our earlier work, we examined the accuracy of the operator decomposition through a perturbative analysis of CT theory starting from a single determinantal wave function Ψ_D and using a single-reference single-doubles excitation operator $A = A_i^a(a_i^a - a_a^i) + A_{ij}^{ab}(a_{ij}^{ab} - a_{ab}^{ij})$. This analysis showed that the L-CTSD theory was accurate through third order in the fluctuation potential $W = H - F$, where F is the Fock operator, i.e.,

$$\begin{aligned}
\langle \Psi_D | \bar{H} | \Psi_D \rangle = &\langle \Psi_D | \bar{H}_{1,2} | \Psi_D \rangle + O(W^4) \\
= &\langle \Psi_D | H + [H,A]_{1,2} \\
&+ [[H,A]_{1,2}, A]_{1,2} | \Psi_D \rangle + O(W^4).
\end{aligned} \tag{16}$$

However, consider what happens if we use the more general *multireference* form of A in Eq. (2) that includes semi-internal excitations such as $A_{ij}^{ak}(a_{ij}^{ak} - a_{ak}^{ij})$, together with a single-reference wave function $|\Psi_D\rangle$. Such excitations should not contribute as they destroy the single-reference wave function, and thus all expectation values of exact commutators containing only semi-internal excitations, e.g., $\langle \Psi_D | [H, A_{ij}^{ak}(a_{ij}^{ak} - a_{ak}^{ij})] | \Psi_D \rangle$, $\langle \Psi_D | [[H, A_{ij}^{ak}(a_{ij}^{ak} - a_{ak}^{ij})], A_{lm}^{bn}(a_{lm}^{bn} - a_{bn}^{lm})] | \Psi_D \rangle$, must vanish. However, using the cumulant based operator decomposition [Eq. (7)], we find that although the expectation value of the first commutator $\langle \Psi_D | [H, A_{ij}^{ak}(a_{ij}^{ak} - a_{ak}^{ij})]_{1,2} | \Psi_D \rangle$ correctly vanishes, it does not do so for the second commutator. Nonvanishing terms arise, e.g., from

$$\langle \Psi_D | [[H, A_{ij}^{ak}(a_{ij}^{ak} - a_{ak}^{ij})]_{1,2}, A_{lm}^{bn}(a_{lm}^{bn} - a_{bn}^{lm})] | \Psi_D \rangle. \tag{17}$$

Writing H and the two A operators as $g^\dagger g^\dagger g g$, $o^\dagger o^\dagger o v$, $v^\dagger o^\dagger o o$, respectively, using g, o, v to denote general, occupied, and virtual indices, respectively, we can see a nonzero contribution arising from

$$\langle \Psi_D | (g^\dagger g^\dagger \underline{gg})(o^\dagger \underline{\overline{o^\dagger o}}v)(v^\dagger \underline{\overline{o^\dagger o}}o) | \Psi_D \rangle \neq 0, \tag{18}$$

where the underbracket denotes contraction and the over-bracket denotes a replacement by a density matrix in the operator decomposition.

In a multireference situation, we use the *same* extended excitation operator A (with semi-internal excitations) through the entire potential energy surface, even when the underlying reference wave function is largely of a single-reference nature, as is sometimes the case near the equilibrium geometry. Thus, the above deficiency of the cumulant operator decomposition for single-reference wave functions motivates us to examine other possible decompositions, as we will describe now.

III. EXTENDED NORMAL ORDERING

A. Normal ordering for a multireference wave function

Normal ordering provides a standard way to decompose an operator into a sum of zero-, one-, two-, and higher body contributions that are ordered with respect to a given vacuum. In many-body theory, it is common to use normal ordering not with respect to the physical vacuum but rather with respect to a single determinant state or Fermi vacuum. With respect to the Fermi vacuum, normal ordering of the operators a_s^p, a_{rs}^{pq}, a_{stu}^{pqr} yields

$$\tilde{a}_s^p = a_s^p - \delta_s^p n_s, \tag{19}$$

$$\tilde{a}_{st}^{pq} = a_{st}^{pq} - \delta_s^p n_s a_t^q - \delta_t^q n_t a_s^p + \delta_t^p n_t a_s^q + \delta_s^q n_s a_t^p + \delta_{st}^{pq} n_p n_q, \tag{20}$$

$$\begin{aligned}\tilde{a}_{stu}^{pqr} = {} & a_{stu}^{pqr} - \delta_s^p n_p a_{tu}^{qr} + \delta_s^q n_q a_{ru}^{pt} + \delta_s^r n_r a_{tu}^{qp} + \delta_t^p n_p a_{su}^{qr} \\ & - \delta_t^q n_q a_{su}^{pr} + \delta_t^r n_r a_{su}^{pq} + \delta_u^p n_p a_{ts}^{qr} + \delta_u^q n_q a_{st}^{pr} - \delta_u^r n_r a_{st}^{pq} \\ & + n_p n_q \delta_{st}^{pq} a_u^r + n_p n_r \delta_{su}^{pr} a_t^q + n_q n_r \delta_{tu}^{qr} a_s^p - n_p n_q \delta_{su}^{pq} a_t^r \\ & - n_p n_q \delta_{ut}^{pq} a_s^r - n_p n_r \delta_{st}^{pr} a_u^q - n_p n_r \delta_{tu}^{pr} a_s^q - n_q n_r \delta_{ts}^{qr} a_u^p \\ & - n_q n_r \delta_{su}^{qr} a_t^p - n_p n_q n_r \delta_{stu}^{pqr}, \end{aligned} \tag{21}$$

where the tilde represents operators normal ordered with respect to the Fermi vacuum (quasiparticle operators), n_p is the occupation number (0 or 1) of the pth orbital, and $\delta_{rs}^{pq} = \delta_r^p \delta_s^q - \delta_s^p \delta_r^q$, $\delta_{stu}^{pqr} = \delta_s^p \delta_t^q \delta_u^r + \delta_t^p \delta_u^q \delta_s^r + \delta_u^p \delta_s^q \delta_t^r - \delta_u^p \delta_t^q \delta_s^r - \delta_s^p \delta_u^q \delta_t^r - \delta_t^p \delta_s^q \delta_u^r$. Note that all normal-ordered operators (other than the "zero-body" constant term) yield a vanishing expectation value with the Fermi vacuum, e.g., $\langle \tilde{a}_s^p \rangle = 0$. If we are interested in a state which is well approximated by the Fermi vacuum, the higher-particle rank quasiparticle operators such as $\tilde{a}_{stu}^{pqr}$ are less relevant to its properties than the lower-rank ones since they represent multiple simultaneous excitations away from the state. Thus, the Fermi-vacuum normal ordering presents a natural way to approximate high-particle rank operators in terms of simpler lower-body terms by simply neglecting the high-particle-rank quasiparticle operators that appear in the normal-ordered form. For example, to approximate a_{stu}^{pqr} in terms of one- and two-body operators alone, we would neglect $\tilde{a}_{stu}^{pqr}$ in Eq. (21).

In the canonical transformation theory, however, we are often interested in reference states which cannot be represented well by any Fermi vacuum. Recently, Mukherjee and Kutzelnigg proposed an elegant generalization of normal ordering with respect to such multireference states.[12–14] By examining the form of the above normal ordering equations when rotated into an arbitrary one-particle basis, they arrived at the generalized relations

$$a_s^p = \tilde{a}_s^p + \gamma_s^p, \tag{22}$$

$$\begin{aligned} a_{st}^{pq} &= \tilde{a}_{st}^{pq} + \gamma_s^p \tilde{a}_t^q + \gamma_t^q \tilde{a}_s^p - \gamma_t^p \tilde{a}_s^q - \gamma_s^q \tilde{a}_t^p + \gamma_{st}^{pq} \\ &= \tilde{a}_{st}^{pq} + 4(\gamma_s^p \wedge \tilde{a}_t^q) + \gamma_{st}^{pq}, \end{aligned} \tag{23}$$

$$\begin{aligned} a_{stu}^{pqr} = {} & \tilde{a}_{stu}^{pqr} + \gamma_s^p \tilde{a}_{tu}^{qr} - \gamma_s^q \tilde{a}_{tu}^{pr} - \gamma_s^r \tilde{a}_{tu}^{qp} - \gamma_t^p \tilde{a}_{su}^{qr} + \gamma_t^q \tilde{a}_{su}^{pr} \\ & - \gamma_t^r \tilde{a}_{su}^{pq} - \gamma_u^p \tilde{a}_{ts}^{qr} - \gamma_u^q \tilde{a}_{st}^{pr} + \gamma_u^r \tilde{a}_{st}^{pq} + \gamma_{st}^{pq} \tilde{a}_u^r + \gamma_{su}^{pr} \tilde{a}_t^q \\ & + \gamma_{tu}^{qr} \tilde{a}_s^p - \gamma_{su}^{pq} \tilde{a}_t^r - \gamma_{ut}^{pq} \tilde{a}_s^r - \gamma_{st}^{pr} \tilde{a}_u^q - \gamma_{tu}^{pr} \tilde{a}_s^q - \gamma_{ts}^{qr} \tilde{a}_u^p \\ & - \gamma_{su}^{qr} \tilde{a}_t^p + \gamma_{stu}^{pqr} \\ = {} & \tilde{a}_{stu}^{pqr} + 9(\gamma_n^p \wedge \tilde{a}_{tu}^{qr}) + 9(\gamma_{st}^{pq} \wedge \tilde{a}_u^r) + \gamma_{stu}^{pqr}. \end{aligned} \tag{24}$$

Let us examine the physical meaning of the above expressions, taking Eq. (23) as an example. Here, we see that the original two-body operator a_{st}^{pq} is written in terms of an average over the reference state (the zero-body operator γ_{st}^{pq}), a product of a one-body average with a one-body quasiparticle operator (the terms like $\gamma_s^p \tilde{a}_t^q$), and a two-body quasiparticle operator $\tilde{a}_{st}^{pq}$. The quasiparticle operators describe fluctuations about the reference because just as in the usual form of normal ordering, their expectation values with the reference vanish, e.g., $\langle \tilde{a}_s^p \rangle = 0$, $\langle \tilde{a}_{st}^{pq} \rangle = 0$.

B. Application to canonical transformation theory

The extended normal ordering provides a systematic operator decomposition which is well suited to the canonical transformation theory. At the linearized CTSD level, we wish to decompose the three-body operators, arising from the commutator $[H, A]$, into lower-body terms. We can do so by neglecting the effects of the simultaneous three-body fluctuations described by the operator $\tilde{a}_{stu}^{pqr}$. For consistency, we should also remove the fully connected three-body cumulant λ_{stu}^{pqr}. First, let us rewrite a_{stu}^{pqr} in terms of a_s^p, a_{st}^{pq} be rearranging Eq. (24), and substituting in the cumulant decomposition of γ_{stu}^{pqr} [Eq. (10)], we find

$$\begin{aligned} a_{stu}^{pqr} = {} & \tilde{a}_{stu}^{pqr} - \gamma_s^p [a_{tu}^{qr} - \gamma_t^q (a_u^r - \gamma_u^r) + \cdots \gamma_{tu}^{qr}] \\ & + \cdots - \gamma_{st}^{pq} (a_u^r - \gamma_u^r + \cdots + \lambda_{stu}^{pqr} + \gamma_s^p \lambda_{tu}^{qr} \\ & + \cdots \gamma_s^p \gamma_t^q \gamma_u^r + \cdots \\ = {} & \tilde{a}_{stu}^{pqr} + 9(\gamma_s^p \wedge a_{tu}^{qr}) - 36(\gamma_s^p \wedge \gamma_t^q \wedge a_u^r) + 9(\gamma_{st}^{pq} \wedge a_u^r) \\ & + 24(\gamma_s^p \wedge \gamma_t^q \wedge \gamma_u^r) - 9(\gamma_s^p \wedge \gamma_{tu}^{qr}) + \lambda_{stu}^{pqr}. \end{aligned} \tag{25}$$

Now, dropping $\tilde{a}_{stu}^{pqr}$ and λ_{stu}^{pqr}, we obtain the extended normal-ordered decomposition, which we name the MK decomposition after Mukherjee and Kutzelnigg,

$$a_{stu}^{pqr} \Rightarrow 9(\gamma_s^p \wedge a_{tu}^{qr}) - 36(\gamma_s^p \wedge \gamma_t^q \wedge a_u^r) + 9(\gamma_{st}^{pq} \wedge a_u^r)$$
$$+ 24(\gamma_s^p \wedge \gamma_t^q \wedge \gamma_u^r) - 9(\gamma_{st}^{pq} \wedge \gamma_u^r). \tag{26}$$

Comparing the MK decomposition to our earlier cumulant-type decomposition [Eq. (7)], we see that they yield the same expectation value with the reference function Ψ_0 and thus differ only by terms whose expectation values vanish. In addition to some different factors, the MK decomposition includes additional operators: a constant term and the term $\gamma_{st}^{pq} \wedge a_u^r$. Computationally, both these terms are easily implemented without affecting the scaling of the original L-CTSD algorithm.

To better understand the differences between the MK and cumulant-type (CU) decompositions, it is instructive to compare the two for a simpler example, namely, the decomposition of the two-particle operator a_{st}^{pq}. These are

$$a_{st}^{pq} \Rightarrow 2(\gamma_s^p \wedge a_t^q) = \tfrac{1}{2}(\gamma_s^p a_t^q + \gamma_t^q a_s^p - \gamma_t^p a_s^q - \gamma_s^q a_t^p) \quad \text{(CU)}, \tag{27}$$

$$a_{st}^{pq} \Rightarrow \gamma_s^p(a_t^q - \gamma_t^q) + \gamma_t^q(a_s^p - \gamma_s^p) - \gamma_t^p(a_s^q - \gamma_s^q)$$
$$- \gamma_s^q(a_t^p - \gamma_t^p) + \gamma_s^p \gamma_t^q - \gamma_t^p \gamma_s^q \quad \text{(MK)}. \tag{28}$$

Here we see that the MK decomposition is expressed in terms of *fluctuations*, e.g., $a_s^q - \gamma_s^q$ in the presence of the field γ_t^p, while the cumulant decomposition involves the bare operators a_s^q directly. The neglected term $\tilde{a}_{st}^{pq}$ in the MK decomposition has the conceptual meaning of a simultaneous two-particle fluctuation operator, and we consider this to be conceptually appealing.

Returning to the earlier example that motivated our examination of alternative operator decompositions, let us now look at the normal-product decomposition of commutators involving semi-internal excitation operators, as in Eq. (17). Starting from a single determinantal reference, the extended normal ordering reduces to the usual normal ordering with respect to a Fermi vacuum described by Eqs. (19)–(21). Then, the operator decomposition corresponds to dropping the three-particle normal-ordered operators $\tilde{a}_{stu}^{pqr}$ in Eq. (21). By construction, the remaining normal-ordered operators, e.g., $\bar{a}_{rs}^{pq}$, all have vanishing expectation value with the Fermi vacuum and, consequently using the MK decomposition, the expectation values of all commutators of the form of Eq. (17) with single determinant references *vanish as they should*, in contrast to the cumulant-type decomposition.

Thus, we see that the extended normal-ordered MK decomposition offers some conceptual and formal advantages over our earlier cumulant-type CU decomposition. Encouraged by these aspects, we have implemented this decomposition and we now proceed to the numerical results.

IV. CALCULATIONS

A. Water and nitrogen potential energy curves

We performed prototype multireference CT calculations for the simultaneous bond-breaking curve of the water molecule and the bond-breaking curve of the nitrogen molecule. We chose these molecules to allow a direct comparison with the results in our previous paper with the CU

TABLE I. Total energies of FCI and differences of various methods from FCI for the simultaneous bond breaking of H_2O molecule with CAS(6e, 5o) and cc-pVDZ basis sets. The units are E_h. The bond angle is fixed at $\angle$HOH = 109.57°. R_e = 0.9929 Å. $\tau_s = 10^{-2}$ and $\tau_d = 10^{-2}$ (described in the Appendix) were used in the L-CT calculations (see Ref. 4 for the previous L-CT results).

	$1R_e$	$2R_e$	$3R_e$	$4R_e$
FCI	−76.238 85	−75.945 58	−75.910 03	−75.908 72
RHF	0.217 18	0.370 02	0.573 65	0.671 59
CASSCF	0.162 99	0.131 96	0.123 02	0.122 59
CASPT2	0.013 30	0.008 43	0.008 48	0.008 52
CASPT3	0.003 77	0.003 83	0.001 74	0.001 58
MR-CI	0.005 56	0.003 78	0.002 96	0.002 90
MR-CI+Q	−0.000 56	−0.000 53	−0.000 66	−0.000 68
MR-ACPF	0.000 93	0.000 54	0.000 20	0.000 17
MR-AQCC	0.002 31	0.001 50	0.001 02	0.000 98
CCSD	0.003 84	0.022 48	0.009 67	0.002 00
CCSDT	0.000 51	−0.002 38	−0.041 06	−0.049 73
L-CTSD(CU)	0.000 29	−0.000 97	−0.001 71	−0.001 72
L-CTSD(MK)	−0.000 77	−0.001 28	−0.001 92	−0.001 92
Previous L-CTD(CU)	0.002 19	−0.000 56	0.002 97	0.002 51
Previous L-CTSD(CU)	0.000 61	−0.003 58	0.003 01	0.002 87

decomposition.[4] Here we have used a wider range of basis sets, including the cc-pVDZ and cc-pVTZ basis sets for water and 6-31G, cc-pVDZ, and cc-pVTZ basis sets for nitrogen.[34,35] For assessment, we carried out calculations with state-of-the-art internally contracted multireference methods—second- and third-order perturbation theory (CASPT2 and CASPT3),[36–39] configuration interaction (MR-CI),[40–42] the *a posteriori* size-extensivity corrected configuration interaction due to Davidson (MR-CI+Q),[43,44] averaged coupled pair functional (MR-ACPF),[45,46] and averaged quadratic coupled cluster theory[47] (MR-AQCC) (both *a priori* size-extensivity modifications of configuration interaction), as well as single-reference coupled cluster calculations at the CCSD and CCSDT level[48,49]). Full configuration interaction (FCI) energies were also used for comparison where available. The CAS space for the multireference calculations was six active electrons in five active orbitals [denoted (6e, 5o)] for the water calculations and (6e, 6o) for the nitrogen calculations. The 1s orbitals in O and N atoms were held frozen in all calculations. For the L-CTSD calculations, we employed both the cumulant (CU) and normal-ordering (MK) operator decompositions described in Sec. II. The internally contracted multireference calculations were executed using MOLPRO,[56] the CC calculations using TCE (Ref. 49) in UTCHEM,[50] and the CT calculations using our own computer program.

Tables I–V present the errors in the total energies of various methods as measured from FCI or (in the larger basis sets) MR-CI+Q at several points across the potential curve. These errors are plotted in Figs. 1–5.

Comparing all the different methods, in the calculations where FCI energies were available, MR-CI+Q provided the smallest maximum absolute error (MAE) and nonparallelity error (NPE) and for this reason was used as the benchmark method when FCI energies could not be obtained. The general order of accuracy in terms of MAE from most to least

TABLE II. Total energies of MR-CI+Q and differences of various methods from MR-CI+Q for the simultaneous bond breaking of H_2O molecule with CAS(6e,5o) and cc-pVTZ basis sets. The units are E_h. The bond angle is fixed at $\angle HOH = 109.57°$. $R_e = 0.9929$ Å. $\tau_s = 10^{-1}$ and $\tau_d = 10^{-2}$ (described in the Appendix) were used in the L-CT calculations.

	$1R_e$	$2R_e$	$3R_e$	$4R_e$
MR-CI+Q	−76.328 47	−76.015 91	−75.974 84	−75.973 45
RHF	0.276 79	0.416 97	0.608 95	0.705 00
CASSCF	0.222 28	0.182 79	0.168 80	0.168 15
CASPT2	0.015 45	0.009 15	0.009 04	0.009 11
CASPT3	0.005 74	0.006 46	0.003 07	0.002 86
MR-CI	0.010 05	0.007 61	0.006 11	0.006 03
MR-ACPF	0.002 32	0.001 76	0.001 39	0.001 37
MR-AQCC	0.004 67	0.003 53	0.002 80	0.002 77
CCSD	0.007 42	0.029 95	0.027 24	0.019 99
CCSDT	−0.000 55	−0.001 47	−0.039 65	−0.048 66
L-CTSD(CU)	0.002 14	0.000 58	−0.000 81	−0.000 84
L-CTSD(MK)	0.001 86	−0.000 04	−0.001 02	−0.001 03

accurate was MR-CI+Q ≈ MR-ACPF ≈ L-CTSD(CU), L-CTSD(MK) ≈ MR-AQCC > CASPT3 ≈ MR-CI > CASPT2. While the MAE of L-CTSD(CU) and L-CTSD(MK) was comparable to that of MR-ACPF and MR-AQCC, the NPE was larger; in the intermediate region, the shapes of the curves somewhat resembled the CASPT3 curve. In the equilibrium region, the L-CTSD energies were similar in accuracy to CCSDT.

The MAE and NPE for the two CT operator decompositions L-CTSD(CU) and L-CTSD(MK) are compared in Table VI. We find that the two operator decompositions performed quite similarly in these systems, although the MAE of L-CTSD(CU) was slightly smaller. For comparison, we have also included the L-CTSD(CU) energies from our calculations in our earlier work.[4] We note that our new L-CTSD(CU) energies are *significantly improved*, particularly in the intermediate dissociation region. This is a result

TABLE III. Total energies of FCI and difference of various methods from FCI for the bond breaking of N_2 molecule with CAS(6e,6o) and 6-31G basis sets. The units are E_h. $\tau_s = 10^{-1}$ and $\tau_d = 10^{-2}$ (described in the Appendix) were used in the L-CT calculations (see Ref. 4 for the previous L-CT results).

	1 Å	2 Å	3 Å
FCI	−109.046 67	−108.859 68	−108.839 05
RHF	0.211 43	0.550 08	0.856 49
CASSCF	0.085 51	0.086 23	0.074 72
CASPT2	0.013 72	0.008 34	0.008 30
CASPT3	0.005 58	0.007 69	0.004 09
MR-CI	0.002 68	0.003 03	0.002 10
MR-CI+Q	−0.000 12	−0.000 14	−0.000 16
MR-ACPF	0.000 92	0.000 71	0.000 27
MR-AQCC	0.001 33	0.001 25	0.000 69
CCSD	0.006 85	−0.007 31	
CCSDT	0.001 22	−0.052 20	
L-CTSD(CU)	0.001 42	−0.001 65	−0.001 73
L-CTSD(MK)	0.000 82	−0.001 87	−0.002 50
Previous L-CTD(CU)	0.005 10	0.004 47	
Previous L-CTD(CU)	0.006 46	0.001 12	

TABLE IV. Total energies of MR-CI+Q and differences of various methods from MR-CI+Q for the bond breaking of N_2 molecule with CAS(6e,6o) and cc-pVDZ basis sets. The units are E_h. $\tau_s = 10^{-1}$ and $\tau_d = 10^{-2}$ (described in the Appendix) were used in the L-CT calculations.

	1 Å	2 Å	3 Å
MR-CI+Q	−109.228 91	−108.983 76	−108.960 35
RHF	0.299 07	0.653 17	0.966 27
CASSCF	0.184 53	0.194 13	0.183 16
CASPT2	0.022 43	0.015 58	0.016 16
CASPT3	0.007 00	0.007 81	0.003 75
MR-CI	0.009 26	0.011 61	0.010 11
MR-ACPF	0.002 62	0.002 45	0.001 73
MR-AQCC	0.004 19	0.004 65	0.003 74
CCSD	0.011 12	0.074 24	
CCSDT	0.001 77	−0.043 82	
L-CTSD(CU)	0.001 18	0.000 24	−0.000 45
L-CTSD(MK)	0.001 17	0.001 62	0.000 26

of the new numerical algorithm, described in the Appendix, which allowed us to significantly reduce the truncation of the operator manifold that we used in our previous work. However, the curves of the new L-CTSD in the figures are not completely smooth due to some remaining operator truncation effects in the numerical solution and removal of this nonsmooth behavior will be addressed in future work.

Table VII shows the spectroscopic constants of N_2 computed by fitting the potential curves. Compared to the available FCI results in the 6-31G basis, MR-CI+Q once again came closest for all spectroscopic parameters (R_e, ω_e, D_e), while the related MR-ACPF and MR-AQCC methods behaved very similarly to MR-CI+Q. Comparing CT against the other methods, different trends were observed for different quantities. For the dissociation energies, we found that MR-CI+Q > MR-ACPF ≈ MR-AQCC ≈ L-CTSD(CU) > L-CTSD(MK) ≈ CCSDT > CASPT3 > CASPT2 > CCSD. For frequencies, in the cc-pVDZ and cc-pVTZ basis, L-CTSD was comparable in accuracy to MR-ACPF/MR-AQCC (though with errors in the opposite direction) and better than those of CCSDT, while the equilibrium bond dis-

TABLE V. Total energies of MR-CI+Q and differences of various methods from MR-CI+Q for the bond breaking of N_2 molecule with CAS(6e,6o) and cc-pVTZ basis sets. The units are E_h. $\tau_s = 10^{-1}$ and $\tau_d = 10^{-2}$ (described in the Appendix) were used in the L-CT calculations.

	1 Å	2 Å	3 Å
MR-CI+Q	−109.337 74	−109.058 71	−109.030 45
RHF	0.369 72	0.701 19	1.004 58
CASSCF	0.254 75	0.250 35	0.235 71
MR-CI	0.014 52	0.017 25	0.015 16
MR-ACPF	0.003 63	0.003 24	0.002 34
MR-AQCC	0.006 25	0.006 69	0.005 48
CASPT2	0.023 99	0.010 94	0.011 95
CASPT3	0.007 53	0.010 77	0.003 96
CCSD	0.015 32	0.095 93	
CCSDT	0.000 21	−0.032 76	
L-CTSD(CU)	0.004 53[a]	−0.000 06	−0.000 61
L-CTSD(MK)	0.002 49	0.001 62	0.000 35

[a] $\tau_s = 5 \times 10^{-1}$ and $\tau_d = 2 \times 10^{-2}$ were used because of convergence problems.
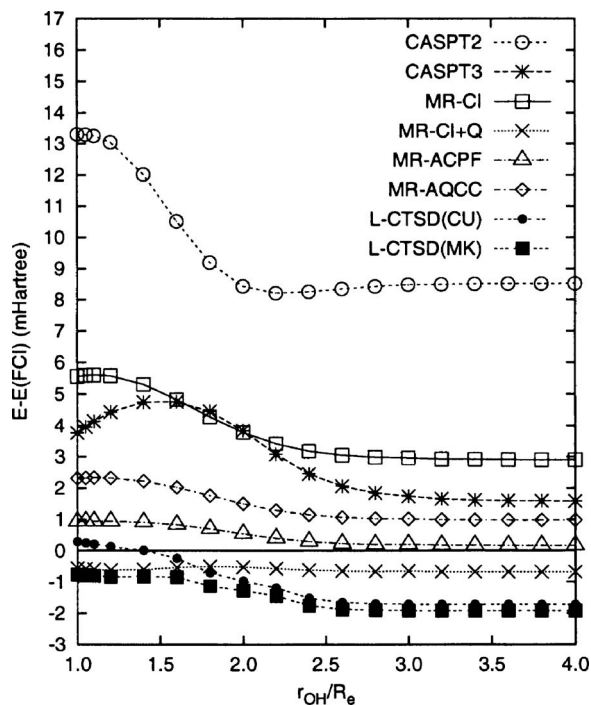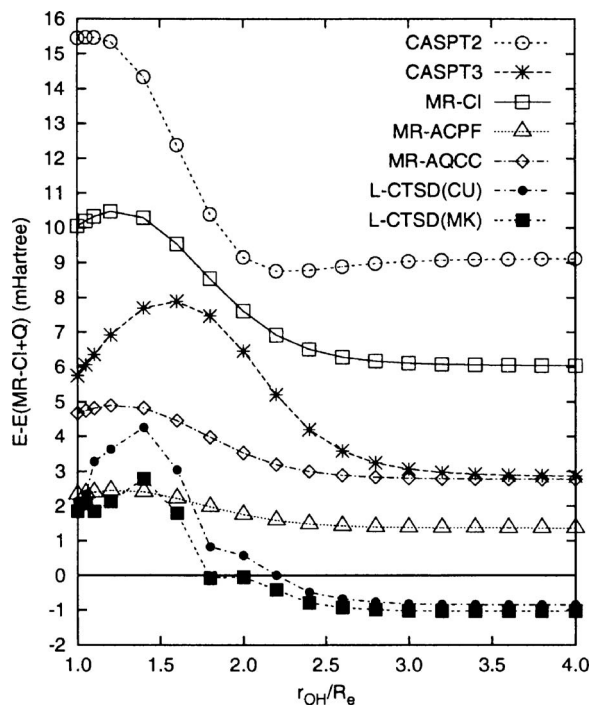
FIG. 1. Energy differences $E-E$(FCI) for the simultaneous bond breaking of H_2O molecule with CAS$(6e,5o)$ and cc-pVDZ basis sets.

FIG. 3. Energy differences $E-E$(FCI) for the bond breaking of N_2 molecule with CAS$(6e,6o)$ and 6-31G basis sets.

tances were less accurate than MR-ACPF/MR-AQCC though still comparable to CCSDT. L-CTSD(CU) and L-CTSD(MK) generally performed similarly, although the spectroscopic constants for L-CTSD(CU) with cc-pVTZ could not be obtained because of convergence problems at the fitting geometries. The small nonsmoothness in the potential energy curves resulting from the numerical approximations in solving the CT equations may also be a factor in the less systematic errors of the CT methods for R_e and ω_e.

Thus, to summarize, the overall performance of L-CTSD(CU) and L-CTSD(MK) for these potential energy curves was competitive with the best multireference methods, such as MR-ACPF and MR-AQCC, particularly for energetic quantities such as the MAE and D_e. The shapes of the curves in the intermediate regions looked somewhat like the CASPT3 curves, though with significantly smaller absolute errors. The spectroscopic constants ω_e, R_e and the non-

FIG. 2. Energy differences $E-E$(MR-CI+Q) for the simultaneous bond breaking of H_2O molecule with CAS$(6e,5o)$ and cc-pVTZ basis sets.

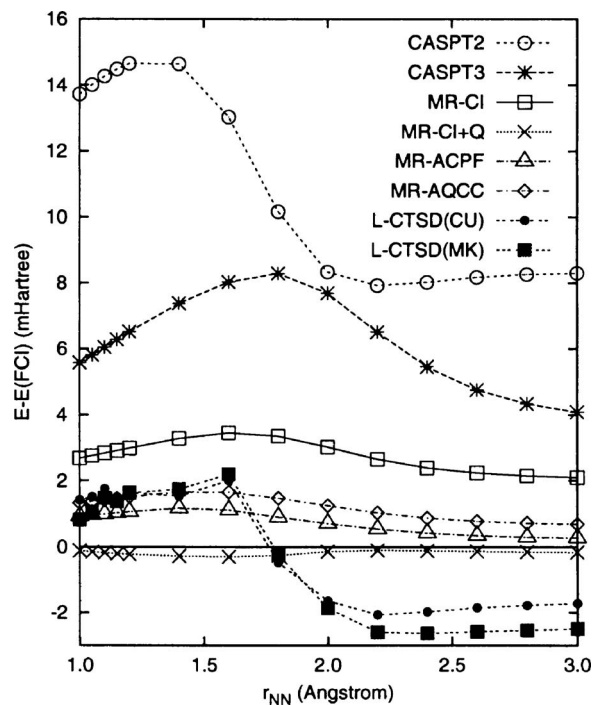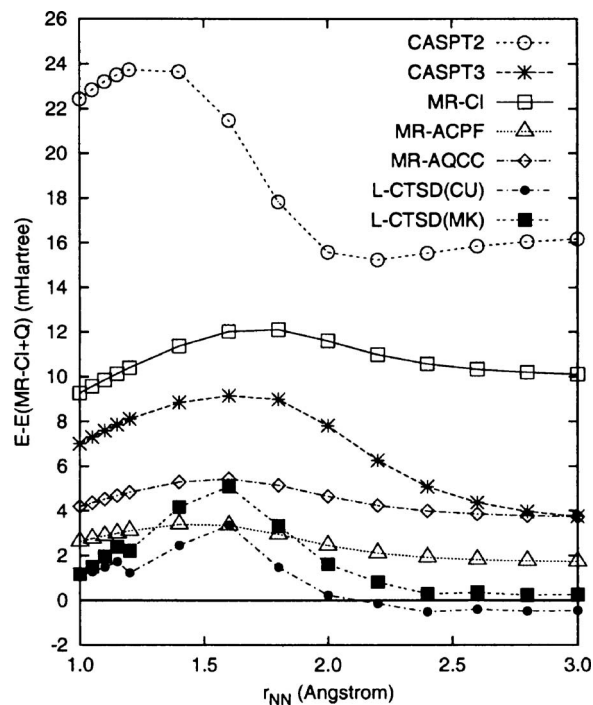FIG. 4. Energy differences $E-E$(MR-CI+Q) for the bond breaking of N_2 molecule with CAS$(6e,6o)$ and cc-pVDZ basis sets.

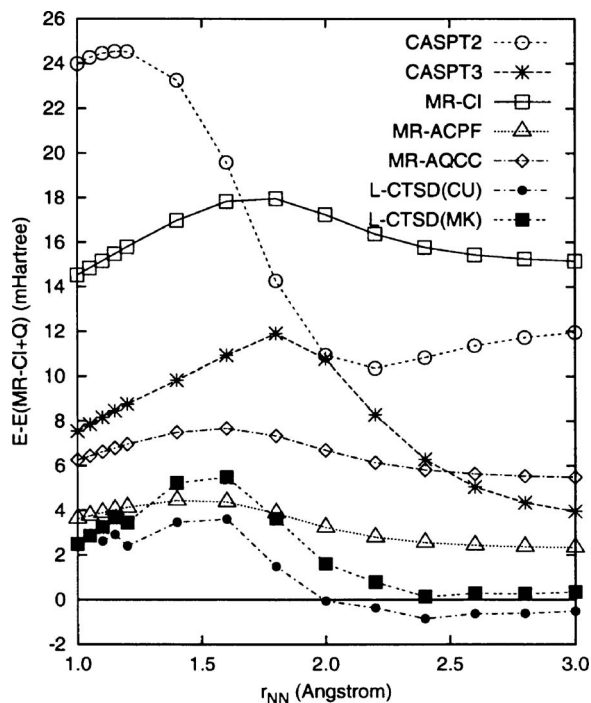FIG. 5. Energy differences $E-E$(MR-CI+Q) for the bond breaking of N_2 molecule with CAS($6e,6o$) and cc-pVTZ basis sets.

parallelity error from L-CTSD were slightly less accurate than from MR-ACPF and MR-AQCC and this was in part related to our numerical approximations in solving the CT equations.

B. Size consistency

As is well recognized, size consistency is a crucial requirement for any correlation model to obtain chemically accurate results in systems with many correlated electrons. As discussed in our initial work,[4] the L-CT theory is naturally size consistent. One way to see this is to observe that the energy is obtained as the expectation value of an effective Hamiltonian that contains only connected contributions by virtue of its construction via a commutator expansion [Eq. (6)]. Here we verify the size consistency property of the L-CT theory through explicit numerical calculations on supermolecules. We have chosen to use supermolecules that contain more than one type of molecule since certain approximate size-extensive theories such as the ACPF and

TABLE VI. Maximum absolute error (MAE) and nonparallelity error (NPE) of L-CTSD(CU) and L-CTSD(MK). The units are mE_h.

	L-CTSD(CU)		L-CTSD(MK)	
	MAE	NPE	MAE	NPE
H_2O/cc-pVDZ	1.72	2.01	1.92	1.11
H_2O/cc-pVTZ	4.25	5.10	2.79	3.82
N_2/6-31G	2.07	4.08	2.54	4.74
N_2/cc-pVDZ	3.35	3.83	5.09	4.86
N_2/cc-pVTZ	3.62	4.46	5.49	5.35

TABLE VII. Spectroscopic constants for N_2 molecule by various methods with 6-31G, cc-pVDZ, and cc-pVTZ basis sets. The dissociation energy D_e was obtained with additional atomic calculations for the nitrogen atom..

	R_e (Å)	ω_e (cm^{-1})	D_e (kcal/mol)
	6-31G		
FCI	1.134 86	2208.27	168.45
RHF	−0.045 74	452.94	−103.25
CCSD	−0.006 77	77.53	−6.27
CCSDT	−0.002 06	26.93	−1.38
CASSCF	−0.003 77	29.35	−8.87
CASPT2	−0.000 91	3.71	−3.85
CASPT3	−0.001 01	8.80	−1.62
MR-CI	−0.000 33	3.14	−1.24
MR-CI+Q	0.000 10	−0.15	−0.02
MR-ACPF	−0.000 16	1.68	−0.49
MR-AQCC	−0.000 20	1.99	−0.55
L-CTSD(CU)	0.001 11	−15.35	−0.67
L-CTSD(MK)	0.000 53	−10.34	−1.21
	cc-pVDZ		
MR-CI+Q	1.120 36	2321.25	200.59
RHF	−0.043 06	436.76	−88.43
CCSD	−0.007 54	87.11	−8.62
CCSDT	−0.001 87	24.90	−1.79
CASSCF	−0.005 89	43.95	−3.58
CASPT2	−0.001 21	4.48	−4.11
CASPT3	−0.001 07	7.77	−2.40
MR-CI	−0.001 10	8.71	−2.83
MR-ACPF	−0.000 47	3.34	−0.48
MR-AQCC	−0.000 62	4.39	−0.48
L-CTSD(CU)	−0.001 25	−4.77	−0.42
L-CTSD(MK)	−0.002 07	−6.20	−1.13
	cc-pVTZ		
MR-CI+Q	1.104 76	2332.40	216.00
RHF	−0.037 59	398.68	−95.61
CCSD	−0.008 05	90.91	−8.30
CCSDT	−0.001 65	22.60	−0.15
CASSCF	−0.001 10	18.63	−12.24
CASPT2	−0.000 45	−4.75	−6.99
CASPT3	−0.001 23	10.06	−2.41
MR-CI	−0.001 24	10.58	−4.22
MR-ACPE	−0.000 43	3.09	−0.37
MR-AQCC	−0.000 63	4.87	−0.43
L-CTSD(MK)	−0.002 49	2.02	−1.68
Expt.	1.107 68	2358.57	228.4

AQCC methods (which modify the non-size-consistent CISD method) are rigorously size consistent only in the special case when the supermolecule is made of n noninteracting *identical* subsystems.

Table VIII gives the size consistency errors of L-CTSD, CISD, CCSD, ACPF, and AQCC calculations for the Be $+n$He and N_2+nHe, respectively. All calculations used the Hartree-Fock (HF) wave function as the reference and the molecules/atoms were each separated by a distance of 1000 bohr. Size consistency implies the condition $E(A+nB)=E(A)+nE(B)$. As can be seen, the L-CTSD and

TABLE VIII. Size consistency errors (mE_h) of CISD, ACPF, AQCC, CCSD, L-CTSD(CU), and L-CTSD(MK) calculations.

	Be+He	Be+2He	Be+3He	Be+4He
CISD	3.10	6.23	9.46	12.83
ACPF	−0.56	−0.76	−0.86	−0.92
AQCC	1.98	2.38	2.66	2.91
CCSD	0.00	0.00	0.00	0.00
L-CTSD(CU)	0.00	0.00	0.00	0.00
L-CTSD(MK)	0.00	0.00	0.00	0.00
	N_2+He	N_2+2He	N_2+3He	N_2+4He
CISD	1.96	4.00	6.12	8.31
ACPF	−0.41	−0.71	−0.94	−1.11
AQCC	−0.57	−0.98	−1.28	−1.50
CCSD	0.00	0.00	0.00	0.00
L-CTSD(CU)	0.00	0.00	0.00	0.00
L-CTSD(MK)	0.00	0.00	0.00	0.00

CCSD calculations are rigorously size consistent, while those of CISD, ACPF, and AQCC steadily increase.

Consider now the ACPF energy functional of the noninteracting system $A+nB$, given by

$$F_{A+nB}^{ACPF} = \frac{\langle H_A \rangle + n\langle H_B \rangle}{1 + (2/N)\langle \delta_A | \delta_A \rangle + (2n/N)\langle \delta_B | \delta_B \rangle}, \quad (29)$$

where $\langle H_A \rangle = \langle \Psi_A | H_A | \Psi_A \rangle$, N is the total number of electrons, which is equal to $N_A + nN_B$, and δ denotes the orthogonal correlation component of Ψ, e.g., $\Psi_A = \Psi_{0A} + \delta_A$. If $A=B$ in Eq. (29), we readily confirm that $F_{(1+n)B}^{ACPF} = (1+n)F_B^{ACPF}$ and the energy is size consistent. The size consistency error in the functional is obtained as

$$\varepsilon(n) = F_{A+nB}^{ACPF} - F_A^{ACPF} - nF_B^{ACPF}$$

$$= \frac{R_A R_B(N_A\langle H_B \rangle + N_B\langle H_A \rangle) - R_A^2 N_B\langle H_B \rangle - R_B^2 N_A\langle H_A \rangle}{R_A R_B^2 + R_A^2 R_B/n}, \quad (30)$$

where $R_A = N_A + 2\langle \delta_A | \delta_A \rangle$. The errors of ACPF and AQCC indeed appear to behave as the above function. [Note that $\epsilon(n=\infty)$ does not vanish.]
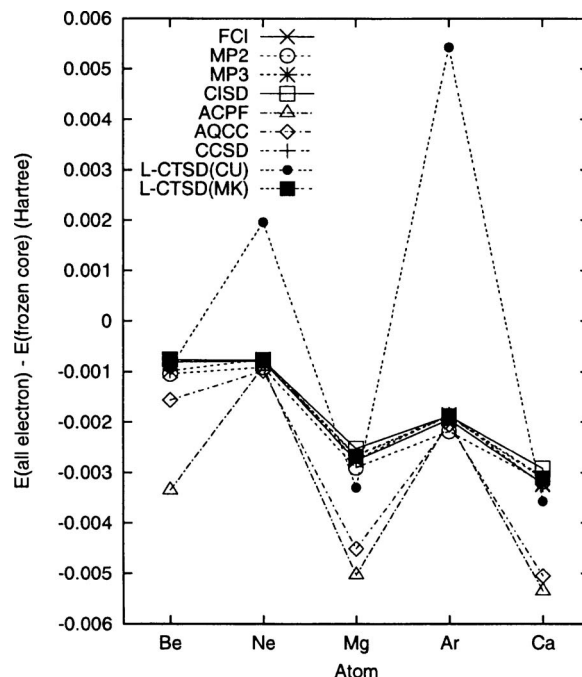
FIG. 6. Density scaling: Energy difference E(all electron)$-E$(frozen core) for atomic calculations shown in Table IX.

C. Density dependence

Rather than considering the scaling behavior of the energy as we increase the number of molecules, we can also consider the complementary trend of going to atoms with larger and larger nuclear charge Z (and consequently more and more electrons in the same region of space). In essence, this measures the density dependence of the energy. To study the behavior of the CT and other methods under this condition, we chose five closed-shell atoms, two rare gas atoms (Ne[$10e$] and Ar[$18e$]), and three alkaline earth metals (Be[$4e$], Mg[$12e$], and Ca[$20e$]).

Table IX and Fig. 6 present the core correlation energies, defined as the energy difference between all-electron and frozen-core atomic calculations, using 6-31G basis sets. The frozen-core calculations correlate eight and two electrons in the valence orbitals for the rare gas atoms and alkaline earth metals, respectively, and represent 98.3% (Be), 99.3% (Ne), 92.1% (Mg), 95.3% (Ar), and 88.7% (Ca) of the all-electron

TABLE IX. Energy difference of all-electron and frozen-core atomic calculations, i.e., E(all electron) $-E$(frozen core), by various method with 6-31G basis sets.

|	Be	Ne	Mg	Ar	Ca
FCI	−0.000 81	−0.000 78	−0.002 76	−0.001 95	−0.003 22
CISD	−0.000 75	−0.000 77	−0.002 52	−0.001 87	−0.002 92
CCSD(T)	−0.000 80	−0.000 78	−0.002 77	−0.001 96	−0.003 22
CCSD	−0.000 78	−0.000 77	−0.002 65	−0.001 89	−0.003 08
AQCC	−0.001 56	−0.000 98	−0.004 52	−0.002 05	−0.005 05
ACPF	−0.003 35	−0.000 90	−0.005 03	−0.001 99	−0.005 35
MP3	−0.000 99	−0.000 76	−0.002 75	−0.001 85	−0.003 24
MP2	−0.001 05	−0.000 91	−0.002 89	−0.002 19	−0.003 18
L-CTSD(CU)	0.053 77	0.113 00	0.036 12	0.038 01	0.028 55
L-CTSD(MK)	0.000 04	0.004 60	−0.002 89	0.006 60	−0.003 20
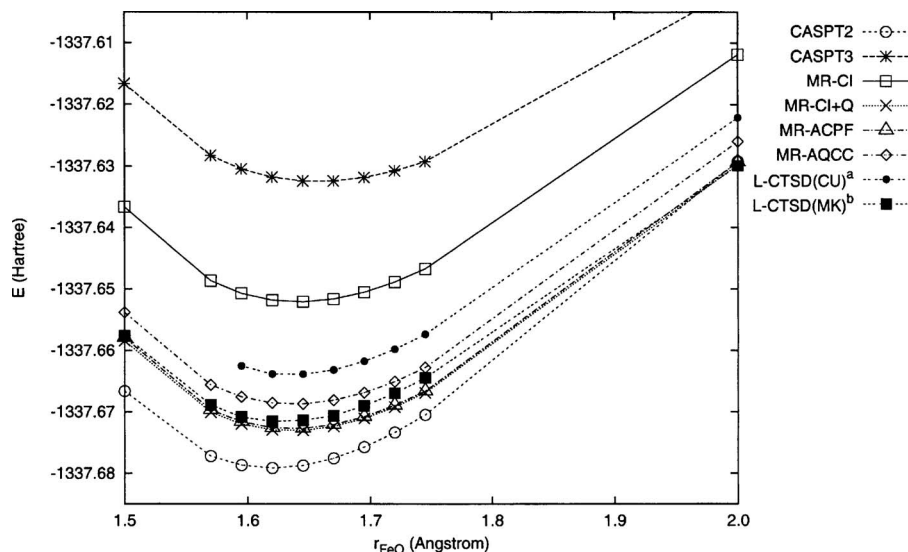
FIG. 7. Potential curve for the ground $1\,{}^5\Delta$ state of FeO molecule. (a) $\tau_s = 3.0 \times 10^{-1}$ and $\tau_d = 5.0 \times 10^{-2}$. (b) $\tau_s = 1.5 \times 10^{-1}$ and $\tau_d = 5.0 \times 10^{-2}$.

correlation energies. The energy difference between the all-electron and frozen-core calculations is the core correlation energy from the core-valence and core-external excitations. Since we can regard valence-electron correlation as a size-intensive quantity that is described by a fixed, i.e., $O(1)$, small number of valence electrons at a roughly constant valence electron density,[57] the rest of the correlation for the bulk of the electrons, i.e., the core correlation, must contain the main density dependence as we change the number of electrons and nuclear charge Z.

Compared to the exact FCI core correlation energies, it is clear that ACPF, AQCC, and L-CTSD(CU) have difficulty reproducing the correct behavior. In particular, large errors are found in the ACPF and AQCC calculations for the alkaline earth metals and in the L-CTSD(CU) calculations of the rare gas atoms. By contrast, the size-inconsistent CISD method as well as the MP2 and MP3 methods are able to capture the correct behavior of the core correlation. This illustrates the difficulty in finding an *ad hoc* size consistency correction, as employed in ACPF and AQCC, that works under all conditions. Most interestingly, the new operator decomposition in L-CTSD(MK) behaves much better than L-CTSD(CU) and reproduces the correct behavior.

D. FeO binding curve

As a realistic example of a difficult multireference problem, we calculated the potential curve for the ground $1\,{}^5\Delta$ state of the FeO molecule. ANO basis sets[51,52] of DZP quality were used, $[21s15p10d6f]/(5s4p3d1f)$ and $[14s9p4d]/(3s2p1d)$ for the Fe and O basis, respectively. To facilitate the setup for the multireference calculations, the initial orbitals were obtained from closed-shell restricted Hartree-Fock (RHF) calculations for the $1\,{}^1\Sigma$ state. The ten lowest lying orbitals for 20 electrons,

$$(1\sigma)^2(2\sigma)^2(3\sigma)^2(4\sigma)^2(5\sigma)^2(6\sigma)^2(1\pi)^4(2\pi)^4, \quad (31)$$

were held frozen for the CASSCF and subsequent dynamic correlation calculations. We verified that the errors made by this orbital restriction were almost constant within 1 mE_h and thus would not affect the shapes of the potential curves.

The orbital $(7\sigma)^2$ was treated as an external core orbital, which was optimized by CASSCF and then correlated. The remaining 12 electrons were fully correlated with 12 active orbitals,

$$(8\sigma)^2(9\sigma)^1(10\sigma)^0(11\sigma)^0(3\pi)^4(4\pi)^2(5\pi)^0(1\delta)^3 \quad (32)$$

[the occupations are based on the restricted open-shell Hartree-Fock (ROHF) configuration of the $^5\Delta$ state] for CAS denoted as $(12e, 12o)$. This CAS is derived from Fe $3d$ and $4s$ orbitals, oxygen $2p$ orbitals, and the third bonding and antibonding π orbitals, which are formed from the oxygen $2p_\sigma$ orbital mixing with some Fe $4p_\sigma$.[53,54]

Figure 7 shows the potential curves of FeO computed by various multireference methods. As exact energies are not available for this system, we report the differences from MR-CI+Q energies in Table X. Clearly, the MR-ACPF, MR-AQCC, and L-CTSD(MK) curves are all very close to each other, while the MR-CI and CASPT2/CASPT3 curves are significantly further away. The MR-ACPF curve follows the MR-CI+Q curve with deviations of less than 0.5 mE_h, while the MR-AQCC curve is also nearly parallel with deviations of 3.8–4.6 mE_h. The L-CTSD(CU) and L-CTSD(MK) curves were shifted relative to each other; the L-CTSD(MK) energies were significantly closer to the MRCI+Q energies, with deviations of less than 2.3 mE_h. CASPT2 seemed to overestimate the correlation energy, while going to the third-order CASPT3 overcorrected too much in the opposite direction and strongly underestimated the correlation energy.

Table XI shows the spectroscopic constants measured from the potential curves. While the basis used is probably too small for direct comparison to experiment, we see that in relation to the experimental results, MR-CI and related modifications MR-CI+Q, MR-ACPF, and MR-AQCC give frequencies that are too low and bond lengths that are too long, while L-CTSD(MK) gives frequencies that are too high and slightly improved bond lengths. As we have already seen in the difference between the CASPT2 and CASPT3 curves, multireference perturbation theory seemed to break down in this molecule.

TABLE X. Total energies of MR-CI+Q (E_h) and differences (mE_h) of various methods from MR−CI+Q for the ground $1\,^5\Delta$ state of FeO molecule.

	1.50 Å	1.57 Å	1.65 Å	1.72 Å	2.00 Å
MRCI+Q	−1337.658 43	−1337.670 07	−1337.673 02	−1337.669 23	−1337.629 80
CASSCF	299.22	294.87	290.00	284.82	265.74
CASPT2	−8.20	−7.14	−5.73	−4.12	0.57
CASPT3	41.79	41.78	40.57	38.46	28.73
MRCI	21.78	21.41	20.93	20.37	17.92
MRACPF	0.51	0.40	0.32	0.29	0.45
MRAQCC	4.61	4.46	4.30	4.16	3.79
L-CTSD(CU)[a]	[b]	[b]	9.17	9.43	7.67
L-CTSD(CU)[c]	[b]	[b]	6.85	6.85	3.51
L-CTSD(MK)[a]	3.09	2.73	3.00	3.85	3.52
L-CTSD(MK)[c]	0.83	1.21	1.66	2.25	−0.16

[a]$\tau_s = 3.0 \times 10^{-1}$ and $\tau_d = 5.0 \times 10^{-2}$.
[b]Not converged.
[c]$\tau_s = 1.5 \times 10^{-1}$ and $\tau_d = 5.0 \times 10^{-2}$.

E. Timings

It is our intention that the CT theory should be practically applicable to problems of reasonable size, and let us now examine the computational timings for the multireference calculations on the FeO molecule we have just discussed. These are shown in Table XII. All timings were obtained on a single CPU of the Altex system (Itanium 1.5 GHz) at the Research Center for Computational Science, Okazaki. As can be seen, the MR-CI based methods were two to three orders of magnitude more expensive than CASPT2. L-CTSD(MK) displayed very satisfactory performance. Even in our protoypte CT implementation, which did not use point-group symmetry, the single point energy calculation *took less time than even the CASPT2 calculation, while providing a significantly better accuracy competitive with MR-ACPF*.

V. SUMMARY AND CONCLUSIONS

We have been developing the canonical transformation theory to describe dynamic correlation in multireference problems. The theory uses a size-extensive unitary exponential acting on a multireference function. In our initial work, we introduced a central approximation that rendered the manipulation of this ansatz practical, namely, a cumulant-based operator decomposition. This choice of decomposition is not unique, however, and in the current work we introduced a new operator decomposition, based on the extended normal ordering of Mukherjee and Kutzelnigg,[12–14] which possesses attractive formal and conceptual features.

We carried out calculations at the L-CTSD level using both our earlier cumulant-based and current Mukherjee-Kutzelnigg operator decompositions. In studies of the water, nitrogen, and iron oxide binding curves, we found the accuracy of L-CTSD to be competitive with some of the best existing multireference methods such as the multireference averaged coupled pair functional, while the computational cost was two to three orders of magnitude less and comparable to that of complete active space second-order perturbation theory. Compared to our earlier work, our results and computational timings were greatly improved, in part due to the use of a new numerical algorithm for converging the canonical transformation equations.

TABLE XI. Spectroscopic constants for the ground $1\,^5\Delta$ state of FeO molecule.

	R_e (Å)	ω_e (cm^{-1})
CASSCF	1.703	691.1
CASPT2	1.620	913.8
CASPT3	1.657	755.2
MR-CI	1.641	844.4
MR-CI+Q	1.635	863.2
MR-ACPF	1.636	863.5
MR-AQCC	1.637	858.7
L-CTSD(MK)[a]	1.631	914.0
L-CTSD(MK)[b]	1.630	911.1
Expt.	1.616	880

[a]$\tau_s = 3.0 \times 10^{-1}$ and $\tau_d = 5.0 \times 10^{-2}$
[b]$\tau_s = 1.5 \times 10^{-1}$ and $\tau_d = 5.0 \times 10^{-2}$.

TABLE XII. Timings for different multireference methods for a single point calculation on the FeO curve. Note that the L-CTSD calculation did not use point-group symmetry, while C_{2v} symmetry was used in all the other calculations. The time for the CASSCF calculation is not included.

	Time (s)
CASPT2	5 900
CASPT3	17 000
MR-CI+Q	158 000
MR-ACPF	168 000
L-CTSD(MK)[a]	4 500

[a]The time for constructing density matrices is not included.

APPENDIX: IMPLEMENTING CANONICAL TRANSFORMATION THEORY

1. Recapitulation

In our previous implementation of the CT algorithm,[4] we solved the residual equations using the following skeletal algorithm:

(1) Set up the electronic Hamiltonian H and the one- and two-particle density matrices of a reference wave function.

(2) Compute the transformed Hamiltonian $\bar{H}_{1,2}$ [Eq. (6)].

(3) Compute the residuals of CT amplitude equations,

$$R_s^p = \langle [\bar{H}_{1,2}, a_s^p - a_p^s]_{1,2} \rangle, \tag{A1}$$

$$R_{st}^{pq} = \langle [\bar{H}_{1,2}, a_{st}^{pq} - a_{pq}^{st}]_{1,2} \rangle, \tag{A2}$$

(4) Update the amplitudes by adding the preconditioned residuals,

$$A_s^p \leftarrow A_s^p - R_s^p / D_s^p, \tag{A3}$$

$$A_{st}^{pq} \leftarrow A_{st}^{pq} - R_{st}^{pq} / D_{st}^{pq}, \tag{A4}$$

where the factors $1/D_s^p$ and $1/D_{st}^{pq}$ are the diagonal preconditioners.

(5) Repeat (2)–(4) until convergence.

In addition, we employed a somewhat complicated division of the optimization process into different steps involving different classes of excitations in the A operator.

2. Preconditioning and orthogonalization

Our primary concern in the current implementation was to improve the convergence of the CT equations. To achieve this, instead of using a diagonal preconditioner as in Eqs. (A3) and (A4), we updated the amplitudes through an exact Newton step. The simplest way to define the Newton update is through the linear equation

$$D_{s,y}^{p,v} \Delta A_y^v = -R_s^p, \tag{A5}$$

$$D_{st,yz}^{pq,vw} \Delta A_{yz}^{vw} = -R_{st}^{pq}, \tag{A6}$$

with

$$D_{s,y}^{p,v} = \langle [[\bar{H}_{1,2}, a_y^v - a_v^y]_{1,2}, a_s^p - a_p^s]_{1,2} \rangle, \tag{A7}$$

$$D_{st,yz}^{pq,vw} = \langle [[\bar{H}_{1,2}, a_{yz}^{vw} - a_{vw}^{yz}]_{1,2}, a_{st}^{pq} - a_{pq}^{st}]_{1,2} \rangle. \tag{A8}$$

We can interpret the D matrices as the derivatives of the residual or Hessians of the energy. However, Eqs. (A7) and (A8) are nonoptimal as the search directions (i.e., the components of A) within the first-order interacting space, namely, the singly external, doubly external, and semi-internal excitations,

$$(a_a^i - a_i^a)\Psi_0, \tag{A9}$$

$$(a_{ij}^{ab} - a_{ab}^{ij})\Psi_0, \tag{A10}$$

$$(a_{ij}^{ak} - a_{ak}^{ij})\Psi_0, \tag{A11}$$

generate a nonorthogonal and even linearly dependent basis. The large spread in eigenvalues of the overlap of the first-order interacting basis [Eqs. (A9)–(A11)] can then cause poor convergence of the linear equations [Eqs. (A5) and (A6)].

To remedy this, we first orthogonalize the first-order interacting basis by diagonalizing the overlap matrix S made of the one- two- and three-particle density matrices,

$$S_{i,j} = \langle (a_i^a - a_a^i)^\dagger (a_j^a - a_a^j) \rangle = \gamma_j^i, \tag{A12}$$

$$S_{i,jkl} = \langle (a_i^a - a_a^i)^\dagger (a_{jk}^{al} - a_{al}^{jk}) \rangle = \gamma_{jk}^{jl}, \tag{A13}$$

$$S_{ijk,lmn} = \langle (a_{ij}^{ak} - a_{ak}^{ij})^\dagger (a_{lm}^{an} - a_{an}^{lm}) \rangle = \delta_{kn} \gamma_{lm}^{ij} - \gamma_{lmk}^{jin}, \tag{A14}$$

$$S_{ij,kl} = \langle (a_{ij}^{ab} - a_{ab}^{ij})^\dagger (a_{kl}^{ab} - a_{ab}^{kl}) \rangle \quad (a \neq b) = \gamma_{kl}^{jj}, \tag{A15}$$

and change to the orthogonalized excitation operators a_μ^a and a_ν^{ab},

$$a_\mu^a = S_{\mu,i}^{-1/2} (a_i^a - a_a^i) + S_{\mu,ijk}^{-1/2} (a_{ij}^{ak} - a_{ak}^{ij}), \tag{A16}$$

$$a_\nu^{ab} = S_{\nu,ij}^{-1/2} (a_{ij}^{ab} - a_{ab}^{ij}). \tag{A17}$$

We can then solve the Newton equations [Eqs. (A5) and (A6)] in this orthogonalized representation. To do so, the quantities A, R, and D are reexpressed in terms of a_μ^a and a_ν^{ab},

$$A = \tilde{A}_\mu^a a_\mu^a + \tilde{A}_\nu^{ab} a_\nu^{ab}, \tag{A18}$$

$$\tilde{R}_\mu^\alpha = \langle [\bar{H}_{1,2}, a_\mu^a - a_a^\mu]_{1,2} \rangle, \tag{A19}$$

$$\tilde{R}_\mu^{ab} = \langle [\bar{H}_{1,2}, a_\mu^{ab} - a_{ab}^\mu]_{1,2} \rangle, \tag{A20}$$

$$\tilde{D}_{\mu,\nu}^{a,b} = \langle [[\bar{H}_{1,2}, a_\nu^b - a_b^\nu]_{1,2} a_\mu^a - a_a^\mu]_{1,2} \rangle, \tag{A21}$$

$$\tilde{D}_{\mu,\nu}^{ab,cd} = \langle [[\tilde{H}_{1,2}, a_\nu^{cd} - a_{cd}^\nu]_{1,2}, a_\mu^{ab} - a_{ab}^\mu]_{1,2} \rangle. \tag{A22}$$

The numbers of operators a_μ^a and a_ν^{ab} are $O(a^3 e)$ and $O(a^2 e^2)$, respectively. Thus, the additional cost of the transformation is $O(a^6 e)$ for the terms involving a_μ^a and $O(a^4 e^2)$ for the terms involving a_ν^{ab}. The diagonalization of the overlap matrix S for the semi-internal and singly external [i.e., Eqs. (A12)–(A14)] requires a cost of $O(a^9)$. While the scaling of these steps is relatively high, they are not expected to be a bottleneck for systems where conventional CASSCF calculations can be performed (for example, internally contracted CASPT2 also contains steps with such cost.[36,37]) However, if we were to use a large active space arising from, e.g., a DMRG calculation, a different algorithm should be used.

Let us consider now the condition number of $\tilde{D}$ and the convergence characteristics of the Newton equations in the orthogonalized representation. If D were formed *without* any operator decomposition approximation, then $\tilde{D}$ would represent the true Hessian of the energy with respect to an orthogonal set of directions in the first-order interacting space. The condition number of $\tilde{D}$ would then be governed by the

excitation energy between the reference and excited states, which could be expected to be reasonable in most systems. Loosely speaking, we can regard the improved condition number of $\tilde{D}$ as arising from the cancellation of small eigenvalues of D by the large eigenvalues of $S^{-1/2}$. However, such a cancellation is unstable, if we approximate D using the operator decomposition. Therefore, to further improve the condition number of $\tilde{D}$, we discarded those operators a_μ^a and a_ν^{ab} which corresponded to small eigenvalues of S. The eigenvalue truncation thresholds are denoted hereafter as τ_s for the singly external and semi-internal and τ_d for the doubly external excitations. This limits the largest linear combination amplitude coefficients (e.g., $S_{\mu,i}^{-1/2}$) appearing in Eqs. (A16) and (A17) to $O(\tau_s^{-1/2})$ and $O(\tau_d^{-1/2})$, respectively, preserving numerical stability in the amplitude equations. Typically, we used $\tau_s < 10^{-1}$ and $\tau_d < 10^{-2}$. These cutoffs appear large because of the extreme degeneracy of the first-order interacting space near equilibrium and because of the incomplete removal of the poorly conditioned components, due to the slight incompatibility (unstable cancellation) between the approximate Hessian and the overlap matrix in this space. Linear dependency is particularly strong near equilibrium because some of the active orbitals which are being excited by A have nearly zero occupancy. However, the contribution of the neglected terms to the energy is small. For N_2 with the cc-pVDZ basis, the size of the effective orthogonalized operator space was 410 ($R_{NN}=1.0$), 698 ($R_{NN}=1.6$), and 938 ($R_{NN}=3.0$), indicating that over 50% of the operator basis was truncated near equilibrium. In the dissociation region of the potential energy curves studied here, truncation did not occur.

In our previous work, we encountered numerical difficulties in using singly external excitation operators in conjunction with doubles, i.e., for L-CTSD. We now see that the reason is the linear dependency between singly external and semi-internal excitations, which appears as nonzero overlap in $S_{i,jkl}$ [Eq. (A13)]. The orthogonalization fixes this issue, and thus we have used L-CTSD as the standard L-CT model in this work. This should be naturally superior to L-CTD as it includes orbital relaxation and extra correlation such as three- or higher-particle excitations from the direct product of singles and doubles.

Using the Newton update as described above, we observed efficient convergence in the CT amplitude equations. Typically, only ten Newton steps would be required to converge the amplitudes in multireference calculations. Convergence behavior of the amplitude equations in L-CTSD(CU) and L-CTSD(MK) was generally similar, but there were some cases where we could converge the L-CTSD(MK) but not the L-CTSD(CU) calculations with the standard truncation thresholds, for example, at $R_{NN}=1$ Å for N_2 (cc-pVTZ), as discussed in Sec. IV A.

3. Operator orthogonalization with cumulant density matrix

Rather than using the exact three-particle density matrix for the orthogonalization procedure described above, we could also imagine using its cumulant decomposition in

TABLE XIII. The total energies (E_h) of L-CTSD(MK) for the bond breaking of the N_2 molecule with CAS($6e,6o$) and various basis sets using the exact and approximate (cumulant) three-particle density matrices for orthogonalization.

	1 Å	2 Å	3 Å
6-31G			
Exact orthog.	−109.045 85	−108.861 56	−108.841 55
Cumulant orthog.	−109.045 66	−108.862 20	−108.841 55
Diff. (mE_h)	+0.19	−0.64	00
cc-pVDZ			
Exact orthog.	−109.227 74	−108.982 14	−108.960 09
Cumulant orthog.	−109.227 52	−108.983 05	−108.960 09
Diff. (mE_h)	+0.22	−0.91	0.0
cc-pVTZ			
Exact orthog.	−109.335 25	−109.057 09	−109.030 11
Cumulant orthog.	Not conv.	−109.058 24	−109.030 11
Diff. (mE_h)		−1.15	0.0

terms of the one- and two-particle density matrices [Eq. (10)]. Table XIII shows the differences of the total energies computed using the amplitude operators that are orthogonalized with the exact and approximate (cumulant) three-particle density matrices for the N_2 potential energy curves discussed in Sec. IV A. With the truncation threshold $\tau_s=10^{-1}$ and $\tau_d=10^{-2}$ and using various basis sets, the energies from both orthogonalizations were generally in good agreement within a deviation of 1.2 mE_h. However, at $R_{NN}=1.0$ Å with the cc-pVTZ basis set, the L-CTSD(MK) calculation with the cumulant based orthogonalization did not converge.

[1] R. J. Bartlett and J. F. Stanton, Rev. Comput. Chem. **5**, 65 (1994).
[2] T. Helgaker, T. A. Ruden, P. Jørgensen, J. Olsen, and W. Klopper, J. Phys. Org. Chem. **17**, 913 (2004).
[3] T. J. Lee and G. E. Scuseria, in *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, edited by S. R. Langhoff (Kluwer Academic, Dordrecht, 1995), Vol. 2, pp. 47–108.
[4] T. Yanai and G. K.-L. Chan, J. Chem. Phys. **124**, 194106 (2006).
[5] S. R. White, J. Chem. Phys. **117**, 7472 (2002).
[6] F. Colmenero and C. Valdemoro, Phys. Rev. A **47**, 979 (1993).
[7] F. Colmenero and C. Valdemoro, Int. J. Quantum Chem. **51**, 369 (1994).
[8] H. Nakatsuji and K. Yasuda, Phys. Rev. Lett. **76**, 1039 (1996).
[9] K. Yasuda and H. Nakatsuji, Phys. Rev. A **56**, 2648 (1997).
[10] D. A. Mazziotti, Phys. Rev. A **57**, 4219 (1998).
[11] D. A. Mazziotti, Chem. Phys. Lett. **289**, 419 (1998).
[12] D. Mukherjee, in *Recent Progress in Many-Body Theories*, edited by E. Schachinger, H. Mitter, and H. Sormann (Plenum, New York, 1995), Vol. 4, p. 127.
[13] D. Mukherjee, Chem. Phys. Lett. **274**, 561 (1997).
[14] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **107**, 432 (1997).
[15] B. O. Roos, Adv. Chem. Phys. **69**, 399 (1987).
[16] K. Ruedenberg, M. W. Schmidt, M. M. Gilbert, and S. T. Elbert, Chem. Phys. **71**, 41 (1982).
[17] G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002).
[18] J. Hachmann, W. Cardoen, and G. K.-L. Chan, J. Chem. Phys. **125**, 144101 (2006).
[19] J. Paldus and X. Li, Adv. Chem. Phys. **110**, 1 (1999).
[20] K. F. Freed, in *Many-Body Methods in Quantum Chemistry*, edited by U. Kaldor (Springer, Berlin, 1989), p. 1.
[21] B. Kirtman, J. Chem. Phys. **75**, 798 (1981).
[22] M. R. Hoffmann and J. Simons, J. Chem. Phys. **88**, 993 (1988).
[23] W. Kutzelnigg, J. Chem. Phys. **77**, 3081 (1982).
[24] W. Kutzelnigg, J. Chem. Phys. **80**, 822 (1984).
[25] J. D. Watts, G. W. Trucks, and R. J. Bartlett, Chem. Phys. Lett. **157**, 359

(1989).

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

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

[28] S. Pal, M. D. Prasad, and D. Mukherjee, Theor. Chim. Acta **62**, 523 (1983).

[29] S. Pal, Theor. Chim. Acta **66**, 207 (1984).

[30] J. H. van Vleck, Phys. Rev. **33**, 467 (1929).

[31] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **110**, 2800 (1999).

[32] D. A. Mazziotti, Phys. Rev. Lett. **97**, 143002 (2006).

[33] D. A. Mazziotti, Phys. Rev. A **75**, 022505 (2007).

[34] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).

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

[36] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990).

[37] K. Andersson, P.-Å. Malmqvist, and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).

[38] H.-J. Werner, Mol. Phys. **89**, 645 (1996).

[39] P. Celani and H.-J. Werner, J. Chem. Phys. **112**, 5546 (2000).

[40] H.-J. Werner and E. A. Reinsch, J. Chem. Phys. **76**, 3144 (1982).

[41] H.-J. Werner and P. J. Knowles, J. Chem. Phys. **89**, 5803 (1988).

[42] P. J. Knowles and H.-J. Werner, Chem. Phys. Lett. **145**, 514 (1988).

[43] S. R. Langoff and E. R. Davidson, Int. J. Quantum Chem. **8**, 61 (1974).

[44] E. R. Davidson and D. W. Silver, Chem. Phys. Lett. **52**, 403 (1977).

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

[46] H.-J. Werner and P. J. Knowles, Theor. Chim. Acta **78**, 175 (1990).

[47] P. G. Szalay and R. J. Bartlett, Chem. Phys. Lett. **214**, 481 (1993).

[48] R. J. Bartlett, in *Modern Electronic Structure Theory*, Advanced Series in Physical Chemistry, edited by D. R. Yarkony (World Scientific, River Edge, NJ, 1995), pp. 1047–1131.

[49] S. Hirata, J. Phys. Chem. A **107**, 4940 (2003).

[50] T. Yanai, H. Nakano, T. Nakajima, T. Tsuneda, S. Hirata, Y. Kawashima, Y. Nakao, M. Kamiya, H. Sekino, and K. Hirao, Utchem-a program for ab initio quantum chemistry. *Lecture Notes in Computer Science* (Springer-Verlag, Berlin, 2003), Vol. 2660, pp. 84–95.

[51] P. O. Widmark, P. Å. Malmqvist, and B. O. Roos, Theor. Chim. Acta **77**, 291 (1990).

[52] R. Pou-Amérigo, M. Merchán, I. Nebot-Gil, P. O. Widmark, and B. O. Roos, Theor. Chim. Acta **92**, 149 (1995).

[53] C. W. Bauschlicher, S. R. Langhoff, and A. Komornicki, Theor. Chim. Acta **77**, 263 (1990).

[54] J. F. Harrison, Chem. Rev. (Washington, D.C.) **100**, 679 (2000).

[55] Coupled cluster theory may similarly be viewed as involving two sets of approximations: truncation of the excitation operators and approximate solution of the Schrödinger equation in a restricted projected space of excitations.

[56] H.-J. Werner, P. J. Knowles, R. Lindh *et al.*, MOLPRO, Version 2006.1, a package of *ab initio* programs.

[57] This separation of core and valence densities can be made precise in the Thomas-Fermi theory, see, e.g., E. H. Lieb, Rev. Mod. Phys. **48**, 553 (1976).