A Canonical Transformation Theory from Extended Normal Ordering

Takeshi Yanai* and Garnet Kin-Lic Chan
Department of Chemistry and Chemical Biology
Cornell University, Ithaca, NY 14853-1301
(Dated: February 5, 2008)

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 Linearised 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-three orders of magnitude faster and comparable to those of 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.

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 behaviour of the wavefunction, 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 describable by a single determinant reference, Coupled Cluster (CC) theory paired with a large basis set yields predictions with chemical accuracy [1, 2, 3]. However, for the many chemical problems which require a multireference characterisation, 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 wavefunction. In the form implemented in that work the computational cost is $O(a^2e^4)$, 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 Linearised Canonical Transformation Doubles (L-CTD) theory performed significantly better than 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 Diagonalisation theory of White [5] although there are earlier related contributions as we 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, 7, 8, 9, 10, 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, 13, 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 realisation along such directions.

A second focus of this work is to investigate in detail the behaviour 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 multi-reference 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.

*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 wavefunction $\Psi_0$ 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 wavefunction may be used \[17,18\]. We then incorporate the remaining dynamic correlation on top of the reference wavefunction $\Psi_0$ via an exponential operator that generates excitations between the active and external spaces yielding

$$\Psi = e^A\Psi_0$$

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

$$A = A_i^a (a_i^a - a_i^a) + A_{ij}^{ab}(a_{ij}^{ab} - a_{ij}^{ab}) + A_{ijk}^{abc}(a_{ijk}^{abc} - a_{ijk}^{abc}) + \ldots$$

where $ijk \ldots$ denote active indices, $abc \ldots$, external indices, $a_i^a = a_i^a a_i^a$, $a_{ij}^{ab} = a_i^a a_j^b a_i^a$, 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 $\tilde{H}$ that acts only in the active space, but which has dynamic correlation folded in from the external space, where

$$\tilde{H} = e^{-A}He^A$$

$$\tilde{H}\Psi = E\Psi$$

The exponential ansatz combined with a multireference wavefunction $\Psi_0$ as shown in eqn. (1) has a long history and we necessarily can only give an incomplete account here. Such an ansatz is used in some forms of multi-reference coupled cluster theory (MRCC) as discussed in the review by Paldus and Li \[19\]. In particular, an early example of a complete theoretical scheme for a related multi-reference 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 multi-reference setting by Freed et al \[20\], Kirtman et al \[21\], and Simons et al \[22\]. We mention also the single-reference unitary coupled cluster work by Kutzelnigg \[23,24\], Bartlett et al \[25,26,27\], 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 recognised by Freed, the folding in of dynamic correlation into the active-space effective Hamiltonian is a form of renormalisation transformation. This picture was pursued by White in his theory of Canonical Diagonalisation \[31\], 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 $H$ to terminate at low-order for low-particle rank in $A$. The difficulty in working with the multireference exponential ansatz arises from the non-commuting excitations in the multi-reference form of $A$ in eqn. (2), which leads to a non-terminating expansion for the effective Hamiltonian $\tilde{H}$. (In fact this difficulty already arises if we use a unitary $e^A$ with the single-reference form of $A$).

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

$$\tilde{H} = H + [H, A] + 1/2[[H, A], A] + \ldots$$

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

$$\tilde{H}_{1,2,\ldots} = H + [H, A]_{1,2,\ldots} + 1/2[[H, A]_{1,2,\ldots}, A]_{1,2,\ldots} + \ldots$$

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, 2, \ldots, $n$, where $n$ is the number of particles), then eqns. (5) and (6) would be identical. If in addition to including all particle ranks in eqn. (6) $A$ contained up to $n$-body excitations, then the CT ansatz (1) would be exact in the sense of full configuration interaction, and indeed eqn. (1) would hold exactly. The two relevant approximations thus arise from restricting the excitations in $A$ (wavefunction ansatz) as well as the form of the operator decomposition (operator ansatz)

As an example, let us consider the linearised 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 eqn. (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,

$$a_{stu}^{PQR} \Rightarrow 9(\gamma_p^p \wedge a_{tu}^q) - 12(\gamma_p^p \wedge \gamma_q^q \wedge a_u^r)$$

where in the above $\wedge$ denotes an antisymmetrisation 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 $\gamma$ in terms of products of cumulants $\lambda$, via

$$\gamma_p^p = (a_p^p)^2 = \lambda_p^p$$

$$\gamma_{st} = \langle a_{st}^P \rangle = \lambda_{st}^P + \gamma_{t}^P \gamma_{s}^P$$

$$\gamma_{stu} = \langle a_{stu}^{PQR} \rangle = \lambda_{stu}^{PQR} + \gamma_{tu}^{PQR} - \gamma_{t}^{PQR} \gamma_{u}^{PQR} + \gamma_{tu}^{PQR} - \gamma_{t}^{PQR} \gamma_{u}^{PQR}$$

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

$$\gamma_{stu}^{PQR} \Rightarrow \gamma_{tu}^{PQR} - \gamma_{tu}^{PQR} + \gamma_{tu}^{PQR}$$

$$\gamma_{stu}^{PQR} - \gamma_{tu}^{PQR} + \gamma_{tu}^{PQR} - \gamma_{tu}^{PQR} + \gamma_{tu}^{PQR}$$

$$-2(\gamma_{tu}^{PQR} - \gamma_{tu}^{PQR} + \gamma_{tu}^{PQR} - \gamma_{tu}^{PQR} + \gamma_{tu}^{PQR})$$

$$= 9(\gamma_p^p \wedge \gamma_q^q \wedge a_u^r)$$

To obtain our operator decomposition, we simply replaced expectation values in the above terms by the decomposed commutators to contain at most one and two-body operators, based on an analogy to the cumulant decomposition of density matrices, one- and two-body operators based on an analogy to the cumulant decomposition of density matrices, respectively.

$$E = \langle \Psi_0 | H_{1,2} | \Psi_0 \rangle$$

$$0 = \langle \Psi_0 | [H_{1,2}, a_{ij}^{ab} - a_{ij}^{ab}] | \Psi_0 \rangle$$

The resulting computational cost of the theory is $O(a_2^2 e^4)$, 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 eqn. (11). 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 r.h.s. of eqn. (11) any term with vanishing expectation value with $\Psi_0$ and still preserve the correspondence with the density matrix cumulant decomposition. 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 wavefunction $\Psi_D$ and using a single-reference single-doubles excitation operator $A = A^d_0 (a_{ij}^a - a_{ij}^a) + A^b (a_{ij}^a - a_{ij}^a)$. 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.

$$\langle \Psi_D | H | \Psi_D \rangle = 0$$

$$\langle \Psi_D | H | [H, A]_{1,2} | \Psi_D \rangle + O(W^2)$$

$$\langle \Psi_D | H | [H, A]_{1,2} | \Psi_D \rangle + O(W^2)$$

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

\[
\langle \Psi_D | [H, A_{ij}^{ak}(a_{aj}^k - a_{aj}^k), A_{im}^{ln}(a_{im}^l - a_{im}^l)] | \Psi_D \rangle
\]

Writing \( H \) and the two \( A \) operators as \( g^i g^j g^q \), \( o^i o^j o^v \), \( v^i v^j v^{o} \) respectively, using \( g, o, v \) to denote general, occupied, and virtual indices respectively, we can see a non-zero contribution arising from

\[
\langle \Psi_D | \left( g^i g^j g^q \right) \left( o^i o^j o^v \right) \left( v^i v^j v^{o} \right) | \Psi_D \rangle \neq 0
\]

where the underbracket denotes contraction and the overbracket 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 wavefunction 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 wavefunctions motivates us to examine other possible decompositions, as we describe now.

III. EXTENDED NORMAL ORDERING

A. Normal ordering for a Multireference wavefunction

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 \( a^q_s \), \( a^{pq}_{st} \), \( a^{pq}_{stu} \) yields

\[
\tilde{a}^q_s = a^q_s - \delta^q_s n_s
\]

\[
\tilde{a}^{pq}_{st} = a^{pq}_{st} - \delta^{pq}_{st} n_s a^1_s - \delta^{pq}_{st} n_t a^1_t + \delta^{pq}_{st} n_s a^p_s + \delta^{pq}_{st} n_t a^p_t
\]

\[
\tilde{a}^{pq}_{stu} = a^{pq}_{stu} - \delta^{pq}_{stu} n_s a^1_s - \delta^{pq}_{stu} n_t a^1_t + \delta^{pq}_{stu} n_u a^{pq}_{su} + \delta^{pq}_{stu} n_v a^{pq}_{sv} + \delta^{pq}_{stu} n_s a^{pq}_{ts} + \delta^{pq}_{stu} n_t a^{pq}_{tu} + \delta^{pq}_{stu} n_u a^{pq}_{ut} + \delta^{pq}_{stu} n_v a^{pq}_{vt} + \delta^{pq}_{stu} n_s a^{pq}_{st} + \delta^{pq}_{stu} n_t a^{pq}_{st} + \delta^{pq}_{stu} n_u a^{pq}_{st} + \delta^{pq}_{stu} n_v a^{pq}_{st}
\]

where the tilde represents operators normal-ordered w.r.t. the Fermi vacuum (quasi-particle operators), \( n_p \) is the occupation number (0 or 1) of the \( p \)-th orbital, and \( \delta^n_{pq} = \delta^n_{pq} - \delta^n_{qs} \delta^n_{ts} + \delta^n_{qs} \delta^n_{tp} - \delta^n_{qs} \delta^n_{qs} - \delta^n_{ts} \delta^n_{tp} + \delta^n_{qs} \delta^n_{qs} - \delta^n_{ts} \delta^n_{tp} - \delta^n_{qs} \delta^n_{qs} \).

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 generalisation of normal-ordering w.r.t. such multireference states \[12, 13, 14\]. By examining the form of the above normal-ordering equations when rotated into an arbitrary one-particle basis, they arrived at the generalised relations

\[
a^{pq}_{st} = \tilde{a}^{pq}_{st} + \gamma^{pq}_{st}
\]

\[
a^{pq}_{stu} = \tilde{a}^{pq}_{stu} + \gamma^{pq}_{stu} - \gamma^{pq}_{st} - \gamma^{pq}_{su}
\]

\[
a^{pq}_{stu} = \tilde{a}^{pq}_{stu} + \gamma^{pq}_{stu} - \gamma^{pq}_{st} - \gamma^{pq}_{su}
\]

\[
\tilde{a}^{pq}_{stu} = \tilde{a}^{pq}_{stu} + \gamma^{pq}_{stu} - \gamma^{pq}_{st} - \gamma^{pq}_{su}
\]

\[
\tilde{a}^{pq}_{stu} = \tilde{a}^{pq}_{stu} + \gamma^{pq}_{stu} - \gamma^{pq}_{st} - \gamma^{pq}_{su}
\]

\[
\tilde{a}^{pq}_{stu} = \tilde{a}^{pq}_{stu} + \gamma^{pq}_{stu} - \gamma^{pq}_{st} - \gamma^{pq}_{su}
\]

B. Application to Canonical Transformation Theory

The extended normal ordering provides a systematic operator decomposition which is well suited to Canonical Transformation theory. At the linearised 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 \(\tilde{a}_{\text{str}}\). For consistency, we should also remove the fully connected three-body cumulant \(\lambda_{\text{str}}\). First let us rewrite \(a_{\text{str}}^{\text{pp}}\) in terms of \(a_{\text{str}}^{\gamma}, a_{\text{str}}^{\tau}\), by rearranging eqn. (24), and substituting in the cumulant decomposition of \(a_{\text{str}}^{\text{pp}}\) from Table I, we find

\[
a_{\text{str}}^{\text{pp}} = \tilde{a}_{\text{str}}^{\text{pp}} - \gamma_{\text{str}}^{\gamma}, a_{\text{str}}^{\gamma} - \gamma_{\text{str}}^{\tau}, a_{\text{str}}^{\tau} + \ldots
\]

\[
- \gamma_{\text{str}}^{\gamma}, a_{\text{str}}^{\gamma} - \gamma_{\text{str}}^{\tau}, a_{\text{str}}^{\tau} + \ldots
\]

\[
+ \lambda_{\text{str}}^{\text{pp}} + \gamma_{\text{str}}^{\gamma}, a_{\text{str}}^{\gamma} + \gamma_{\text{str}}^{\tau}, a_{\text{str}}^{\tau} + \ldots
\]

\[
= \tilde{a}_{\text{str}}^{\text{pp}} + 9(\gamma_{\text{str}}^{\gamma} \wedge a_{\text{str}}^{\gamma} - 36(\gamma_{\text{str}}^{\gamma} \wedge \gamma_{\text{str}}^{\tau} \wedge a_{\text{str}}^{\tau}) + 9(\gamma_{\text{str}}^{\gamma} \wedge a_{\text{str}}^{\tau})
\]

\[
+ 24(\gamma_{\text{str}}^{\gamma} \wedge \gamma_{\text{str}}^{\tau} \wedge a_{\text{str}}^{\tau}) - 9(\gamma_{\text{str}}^{\gamma} \wedge \gamma_{\text{str}}^{\tau}) + \lambda_{\text{str}}^{\text{pp}}
\]

(25)

Now dropping \(\tilde{a}_{\text{str}}^{\text{pp}}\) and \(\lambda_{\text{str}}^{\text{pp}}\) we obtain the extended normal-ordered decomposition, which we name the MK decomposition after Mukherjee and Kutzelnigg,

\[
a_{\text{str}}^{\text{pp}} \Rightarrow 9(\gamma_{\text{str}}^{\gamma} \wedge a_{\text{str}}^{\gamma} - 36(\gamma_{\text{str}}^{\gamma} \wedge \gamma_{\text{str}}^{\tau} \wedge a_{\text{str}}^{\tau}) + 9(\gamma_{\text{str}}^{\gamma} \wedge a_{\text{str}}^{\tau})
\]

\[
+ 24(\gamma_{\text{str}}^{\gamma} \wedge \gamma_{\text{str}}^{\tau} \wedge a_{\text{str}}^{\tau}) - 9(\gamma_{\text{str}}^{\gamma} \wedge \gamma_{\text{str}}^{\tau}) + \lambda_{\text{str}}^{\text{pp}}
\]

(26)

Comparing the MK decomposition to our earlier cumulant-type decomposition (7) we see that they yield the same expectation value with the reference function \(\Psi_0\) and thus differ only by terms whose expectation values vanish. In addition to some different factors, the MK decomposition include additional operators: a constant term, and the term \(\gamma_{\text{str}}^{\gamma} \wedge a_{\text{str}}^{\gamma}\). 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 \(a_{st}^{\text{pp}}\). These are

\[
a_{st}^{\text{pp}} \Rightarrow 2(\gamma_{st}^{\gamma} \wedge a_{st}^{\gamma})
\]

\[
= \frac{1}{2}(\gamma_{st}^{\gamma} a_{st}^{\gamma} + \gamma_{st}^{\gamma} a_{st}^{\gamma} - \gamma_{st}^{\gamma} a_{st}^{\gamma} - \gamma_{st}^{\gamma} a_{st}^{\gamma})
\]

CU

(27)

\[
a_{st}^{\text{pp}} \Rightarrow \gamma_{st}^{\gamma}(a_{st}^{\gamma} - \gamma_{st}^{\gamma}) + \gamma_{st}^{\gamma}(a_{st}^{\gamma} - \gamma_{st}^{\gamma})
\]

\[
- \gamma_{st}^{\gamma}(a_{st}^{\gamma} - \gamma_{st}^{\gamma}) - \gamma_{st}^{\gamma}(a_{st}^{\gamma} - \gamma_{st}^{\gamma}) + \gamma_{st}^{\gamma} \gamma_{st}^{\gamma} - \gamma_{st}^{\gamma} \gamma_{st}^{\gamma}
\]

MK

(28)

Here we see that the MK decomposition is expressed in terms of fluctuations e.g. \(\gamma_{st}^{\gamma} - \gamma_{st}^{\tau}\) in the presence of the field \(\gamma_{st}^{\gamma}\), while the cumulant decomposition involves the bare operators \(a_{st}^{\gamma}\) directly. The neglected term \(a_{st}^{\text{pp}}\) 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 eqn. (14). Starting from a single determinant reference, the extended normal ordering reduces to the usual normal ordering with respect to a Fermi vacuum described by eqns (19)-(21). Then, the operator decomposition corresponds to dropping the three-particle normal-ordered operators \(\tilde{a}_{\text{str}}^{\text{pp}}\) in eqn. (21). By construction, the remaining normal-ordered operators e.g. \(a_{str}^{\gamma}\) 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 eqn. (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 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].

| Method       | 1R_e  | 2R_e  | 3R_e  | 4R_e  |
|--------------|-------|-------|-------|-------|
| FCID         | -76.23885 | -75.94558 | -75.91003 | -75.90872 |
| MR-HF        | 0.21718 | 0.37002 | 0.57365 | 0.67159 |
| CASCCF       | 0.16299 | 0.13196 | 0.12302 | 0.12259 |
| CASPT2       | 0.01330 | 0.00843 | 0.00848 | 0.00852 |
| CASPT3       | 0.00377 | 0.00383 | 0.00174 | 0.00018 |
| MR-CI        | 0.00556 | 0.00378 | 0.00296 | 0.00290 |
| MR-CI+Q      | -0.00056 | -0.00053 | -0.00066 | -0.00068 |
| MR-ACPF      | 0.00093 | 0.00054 | 0.00020 | 0.00017 |
| MR-AQCC      | 0.00231 | 0.00150 | 0.00102 | 0.00098 |
| CCSD         | 0.00384 | 0.02248 | 0.00967 | 0.00200 |
| CCSDT        | 0.00051 | -0.00238 | -0.00106 | -0.00073 |
| L-CTSD(CU)   | 0.00029 | -0.00097 | -0.00171 | -0.00172 |
| L-CTSD(MK)   | -0.00077 | -0.00128 | -0.00192 | -0.00192 |
| previous L-CTSD(CU) | 0.00219 -0.00056 | 0.00297 | 0.00251 |
| previous L-CTSD(CU) | 0.00061 -0.00358 | 0.00301 | 0.00287 |

TABLE I: Total energies of FCID and differences of various methods from FCID for the simultaneous bond breaking of H2O molecule with CAS(6e, 5o) and cc-pVDZ basis sets. The units are E_h. The bond angle is fixed at \(<\text{HOH} = 109.57^\circ\). \(R_e = 0.9299\). \(\tau_0 = 10^{-2}\) and \(\tau_d = 10^{-2}\) (described in Sec. VIII) were used in the L-CT calculations. See Ref. [4] for the previous L-CT results.
TABLE II: Total energies of MR-CI+Q and differences of various methods from MR-CI+Q for the simultaneous bond breaking of H$_2$O molecule with CAS(6e, 5o) and cc-pVTZ basis sets. The units are $E_h$. The bond angle is fixed at $<\text{HOH}>=109.57^\circ$. $R_e=0.9929$ Å, $\tau_s=10^{-1}$ and $\tau_d=10^{-2}$ (described in Sec. VII B) were used in the L-CT calculations.

|           | $R_e$ | $2R_e$ | $3R_e$ | $4R_e$ |
|-----------|-------|--------|--------|--------|
| MR-CI+Q   | -76.32847 | -76.01591 | -75.97484 | -75.97345 |
| RHF       | 0.27679 | 0.41697 | 0.60895 | 0.70500 |
| CASSCF    | 0.22228 | 0.18279 | 0.16880 | 0.16815 |
| CASPT2    | 0.01545 | 0.00915 | 0.00904 | 0.00911 |
| CASPT3    | 0.00574 | 0.00646 | 0.00307 | 0.00286 |
| MR-CI     | 0.01005 | 0.00611 | 0.00603 | 0.00603 |
| MR-ACPF   | 0.00232 | 0.00176 | 0.00137 | 0.00137 |
| MR-AQCC   | 0.00467 | 0.00353 | 0.00277 | 0.00277 |
| CCSD      | 0.00742 | 0.02995 | 0.02724 | 0.01999 |
| CCSDT     | -0.00855 | -0.00147 | -0.03965 | -0.04866 |
| L-CTSD(CU)| 0.00214 | 0.00058 | -0.00081 | -0.00084 |
| L-CTSD(MK)| 0.00186 | -0.00004 | -0.00102 | -0.00103 |

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, 37, 38, 39], configuration interaction (MR-CI) [40, 41, 42], the a posteriori size-extensivity corrected configuration interaction due to Davidson (MR-CI+Q) [43, 44], averaged coupled pair functional (MR-

TABLE III: Total energies of FCI and differences 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 Sec. VII B) were used in the L-CT calculations. See Ref. [4] for the previous L-CT results.

|           | 1Å    | 2Å    | 3Å    |
|-----------|-------|-------|-------|
| FCI       | -109.04667 | -108.85968 | -108.83905 |
| RHF       | 0.21143 | 0.55008 | 0.85649 |
| CASSCF    | 0.08551 | 0.08623 | 0.07472 |
| CASPT2    | 0.01372 | 0.00834 | 0.00830 |
| CASPT3    | 0.00558 | 0.00769 | 0.00409 |
| MR-CI     | 0.00268 | 0.00303 | 0.00210 |
| MR-CI+Q   | -0.00012 | -0.00014 | -0.00016 |
| MR-ACPF   | 0.00092 | 0.00071 | 0.00027 |
| MR-AQCC   | 0.00133 | 0.00125 | 0.00069 |
| CCSD      | 0.00685 | -0.00731 | |
| CCSDT     | 0.00122 | -0.05220 | |
| L-CTSD(CU)| 0.00142 | -0.00165 | -0.00173 |
| L-CTSD(MK)| 0.00082 | -0.00187 | -0.00250 |
| previous L-CTD(CU) | 0.00510 | 0.00447 | |
| previous L-CTSD(CU) | 0.00646 | 0.00112 | |

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, 37, 38, 39], configuration interaction (MR-CI) [40, 41, 42], the a posteriori size-extensivity corrected configuration interaction due to Davidson (MR-CI+Q) [43, 44], averaged coupled pair functional (MR-

FIG. 1: Energy differences $E-E(\text{FCI})$ for the simultaneous bond breaking of H$_2$O molecule with CAS(6e, 5o) and cc-pVDZ basis sets.

FIG. 2: Energy differences $E-E(\text{MR-CI+Q})$ for the simultaneous bond breaking of H$_2$O molecule with CAS(6e, 5o) and cc-pVTZ basis sets.
### 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 Sec. VII B) were used in the L-CT calculations.

| Method       | 1Å E$_\text{MR-CI+Q}$ | 2Å E$_\text{MR-CI+Q}$ | 3Å E$_\text{MR-CI+Q}$ |
|--------------|------------------------|------------------------|------------------------|
| RHF          | -109.22891             | -108.98376             | -108.96035             |
| CASCF        | 0.29907                | 0.65317                | 0.96627                |
| CASPT2       | 0.18453                | 0.19413                | 0.18316                |
| CASPT3       | 0.00700                | 0.00781                | 0.00375                |
| MR-CI        | 0.00926                | 0.01161                | 0.01011                |
| MR-ACPF      | 0.00262                | 0.00245                | 0.00173                |
| MR-AQCC      | 0.00419                | 0.00465                | 0.00374                |
| CCSD         | 0.01112                | 0.07424                |                          |
| CCSDT        | 0.00177                | -0.04382               |                          |
| L-CTSD(CU)   | 0.00118                | 0.00024                | -0.00045               |
| L-CTSD(MK)   | 0.00117                | 0.00162                | 0.00026                |

ACPF) [45, 46] and averaged quadratic coupled-cluster theory (MR-AQCC) [47] (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

### 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 Sec. VII B) were used in the L-CT calculations.

| Method       | 1Å E$_\text{MR-CI+Q}$ | 2Å E$_\text{MR-CI+Q}$ | 3Å E$_\text{MR-CI+Q}$ |
|--------------|------------------------|------------------------|------------------------|
| RHF          | -109.33774             | -109.05871             | -109.03045             |
| CASCF        | 0.36972                | 0.70119                | 1.00458                |
| MR-CI        | 0.01452                | 0.01725                | 0.01516                |
| MR-ACPF      | 0.00636                | 0.00324                | 0.00234                |
| MR-AQCC      | 0.00625                | 0.00669                | 0.00548                |
| CASPT2       | 0.02399                | 0.01094                | 0.01195                |
| CASPT3       | 0.00753                | 0.00107                | 0.00396                |
| CCSD         | 0.01532                | 0.09593                |                          |
| CCSDT        | 0.00021                | -0.03276               |                          |
| L-CTSD(CU)   | 0.00453a               | -0.00006               | -0.00051               |
| L-CTSD(MK)   | 0.00249                | 0.00162                | 0.00035                |

a) $\tau_s = 5 \times 10^{-1}$ and $\tau_d = 2 \times 10^{-2}$ were used because of convergence problems.

### FIG. 3: Energy differences E-E(FCI) for the bond breaking of N$_2$ molecule with CAS(6e, 6o) and 6-31G 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.
L-CTSD calculations, we employed both the cumulant (CU) and normal-ordering (MK) operator decompositions described in section III. The internally-contracted multireference calculations were executed using MOLPRO [49], the CC calculations using TCE [42] in UTCHEM [50], and the CT calculations using our own computer program.

Tables I, II, III, IV, and 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 Figures 1, 2, 3, 4, and 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 non-parallelity 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 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 of the new numerical algorithm, described in VII B 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 non-smooth behaviour will be addressed in future work.

Table VII shows the spectroscopic constants of N2 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 (\( R_e \), \( \omega_e \), \( D_e \)) 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 distances 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 non-smoothness 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 \( \omega_e \).

Thus to summarise, the overall performance of L-CTSD(CU) and L-CTSD(MK) for these potential energy curves was competitive with the best multirefer-

![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.](image-url)
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.

|        | $R_e$  | $\omega_e$ | $D_e$  |
|--------|--------|------------|--------|
|        | Å      | cm$^{-1}$  | kcal/mol |
| 6-31G  |        |            |         |
| FCI    | 1.134  | 2208.27    | 168.45  |
| RHF    | -0.045 | 452.94     | -103.25 |
| CCSD   | -0.006 | 77.53      | -6.27   |
| CCSDT  | -0.002 | 26.93      | -1.38   |
| CASSCF | -0.003 | 29.35      | -8.87   |
| CASPT2 | -0.000 | 3.71       | -3.85   |
| CASPT3 | -0.001 | 8.80       | -1.62   |
| MR-CI  | -0.000 | 3.14       | -1.24   |
| MR-CI+Q| 0.000  | -0.15      | -0.02   |
| MR-ACPF| -0.000 | 1.68       | -0.49   |
| MR-AQCC| -0.000 | 1.99       | -0.55   |
| L-CTSD(CU)| 0.001 | -15.35    | -0.67   |
| L-CTSD(MK)| 0.000 | -10.34    | -1.21   |

| cc-pVDZ        |        |            |         |
| MR-CI+Q        | 1.120  | 2321.25    | 200.59  |
| RHF            | -0.043 | 436.76     | -88.43  |
| CCSD           | -0.007 | 87.11      | -8.62   |
| CCSDT          | -0.001 | 24.90      | -1.79   |
| CASSCF         | -0.005 | 43.95      | -3.58   |
| CASPT2         | -0.001 | 4.48       | -4.11   |
| CASPT3         | -0.001 | 7.77       | -2.40   |
| MR-CI          | -0.001 | 8.71       | -2.83   |
| MR-ACPF        | -0.000 | 3.34       | -0.48   |
| MR-AQCC        | -0.000 | 4.39       | -0.48   |
| L-CTSD(CU)     | -0.001 | 4.77       | -0.42   |
| L-CTSD(MK)     | -0.002 | -6.20      | -1.13   |

| cc-pVTZ        |        |            |         |
| MR-CI+Q        | 1.104  | 2332.40    | 216.00  |
| RHF            | -0.037 | 398.68     | -95.61  |
| CCSD           | -0.008 | 90.91      | -8.30   |
| CCSDT          | -0.001 | 22.60      | -0.15   |
| CASSCF         | -0.001 | 18.63      | -12.24  |
| CASPT2         | -0.001 | -4.75      | -6.99   |
| CASPT3         | -0.001 | 10.06      | -2.41   |
| MR-CI          | -0.001 | 10.58      | -4.22   |
| MR-ACPF        | -0.000 | 3.09       | -0.37   |
| MR-AQCC        | -0.000 | 4.87       | -0.43   |
| L-CTSD(MK)     | -0.002 | 2.02       | -1.68   |

exptl          1.107 2358.57 228.4

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        |

ence 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 $\omega_e$, $R_e$ and the non-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 (6). Here we verify the size consistency property of 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 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 sub-systems.

Table VIII gives the size consistency errors of L-CTSD, CISD, CCSD, ACPF and AQCC calculations for the Be + $n$ He and N$_2$ + $n$ He, respectively. All calculations used the HF wavefunction 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 CCSD calculations are rigorously size-consistent while those of CISD, ACPF and AQCC steadily increase.

Consider now the ACPF energy functional of the non-interacting system $A + nB$, given by

$$F_{A+nB}^{ACPF} = \frac{\langle H_A \rangle + n\langle H_B \rangle}{1 + (2/N)\langle \delta_A | \delta_A \rangle + (2n/N)\langle \delta_B | \delta_B \rangle}$$ (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 $\delta$ denotes the orthogonal correlation component of $\Psi$, e.g. $\Psi_A = \Psi_{0A} + \delta_A$. If $A = B$ in eqn. (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

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

$$= \frac{R_A R_B (N_A \langle H_B \rangle + N_B \langle H_A \rangle) - R_A^2 N_B \langle H_B \rangle - R_B^2 N_A \langle H_A \rangle}{R_A R_B + R_A^2 + R_B^2}$$ (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).

### C. Density dependence

Rather than considering the scaling behaviour 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 behaviour 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 figure 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 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 behaviour. 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 behaviour 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 behaviour.

### D. FeO binding curve

As a realistic example of a difficult multireference problem, we calculated the potential curve for the ground $^1\Delta$ 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,
TABLE X: Total energies of MR-CI+Q (\(E_h\)) and differences (\(\Delta E_h\)) of various methods from MR-CI+Q for the ground \(^1\Delta\) state of FeO molecule.

| Method       | 1.50Å | 1.57Å | 1.65Å | 1.72Å | 2.00Å |
|--------------|-------|-------|-------|-------|-------|
| MR-CI+Q      | -1337.65843 | -1337.67007 | -1337.67302 | -1337.66923 | -1337.62980 |
| 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 | c     | c     | 9.17  | 9.43  | 7.67  |
| L-CTSD(CU)^b | c     | c     | 6.85  | 6.85  | 3.51  |
| L-CTSD(MK)^a | 3.09  | 2.73  | 3.00  | 3.85  | 3.52  |
| L-CTSD(MK)^b | 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) \(\tau_s = 1.5 \times 10^{-1}\) and \(\tau_d = 5.0 \times 10^{-2}\).

c) Not converged.

TABLE XI: Spectroscopic constants for the ground \(^1\Delta\) state of FeO molecule.

| Method       | \(R_e(\text{Å})\) | \(\omega_e(\text{cm}^{-1})\) |
|--------------|-------------------|-----------------------------|
| 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                       |
| exptl.       | 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 \(\tau_d = 5.0 \times 10^{-2}\).
are formed from the oxygen 2pσ orbital mixing with some Fe 4pσ [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 mEh, while the MR-AQCC curve is also nearly parallel with deviations of 3.8-4.6 mEh. 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 mEh. CASPT2 seemed to overestimate the correlation energy, while going to the third order CASPT3 over-corrected 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 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.

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 1.5GHz) 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 prootypte 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 multiref-
reference 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, 13, 14], which possesses attractive formal and conceptual features.

We carried out calculations at the Linearised Canonical Transformation Theory Singles and Doubles (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 we also acknowledge a grant of computer time at the Research Center for Computational Science, Okazaki, Japan, with which some of these calculations were performed.

VI. ACKNOWLEDGEMENTS

This work was supported by Cornell University, the National Science Foundation CAREER program CHE-0645380, and the David and Lucile Packard Foundation. We also acknowledge a grant of computer time at the Research Center for Computational Science, Okazaki, Japan, with which some of these calculations were performed.

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.

| Method | Time/s |
|--------|--------|
| CASPT2 | 5900   |
| CASPT3 | 17000  |
| MR-Cl+Q | 158000 |
| MR-ACPF | 168000 |
| L-CTSD(MK)* | 4500  |

a) The time for constructing density matrices is not included.

VII. APPENDIX: IMPLEMENTING CANONICAL TRANSFORMATION THEORY

A. 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 wavefunction.

2. Compute the transformed Hamiltonian $\hat{H}_{1,2}$ (eqn. (2)).

3. Compute the residuals of CT amplitude equations.

\[ R_p = \langle [\hat{H}_{1,2}, a_p^s - a_p^s]|_{1,2} \rangle \]  (33)

\[ R_{pq} = \langle [\hat{H}_{1,2}, a_{pq}^s - a_{pq}^s]|_{1,2} \rangle \]  (34)

4. Update the amplitudes by adding the preconditioned residuals

\[ A_p^s \leftarrow A_p^s - R_p^s / D_{p}^s \]  (35)

\[ A_{pq}^s \leftarrow A_{pq}^s - R_{pq}^s / D_{pq}^s \]  (36)

where the factors $1/D_p^s$ and $1/D_{pq}^s$ are the diagonal preconditioners.

5. Repeat (2)-(4) until convergence.

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

B. Preconditioning and orthogonalisation

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 [33, 36], we updated the amplitudes through an exact Newton step. The simplest way to define the Newton update is through the linear equation

\[ D_{s,y}^{p,v} \Delta A_{y}^{p,v} = -R_{p}^{s} \]  (37)

\[ D_{s,t,yz}^{p,vw} \Delta A_{yz}^{vw} = -R_{pq}^{s} \]  (38)

with

\[ D_{s,y}^{p,v} = \langle [\hat{H}_{1,2}, a_{y}^{p} - a_{y}^{s}]|_{1,2}, a_{p}^{s} \rangle \]  (39)

\[ D_{s,t,yz}^{p,vw} = \langle [\hat{H}_{1,2}, a_{yz}^{vw} - a_{yz}^{vw}]|_{1,2}, a_{s}^{p} \rangle \]  (40)

We can interpret the $D$ matrices as the derivatives of the residual or Hessians of the energy. However, eqns. [35, 36] are non-optimal as the search directions (i.e. the components of $A$) within the first-order interacting...
The condition number of $\tilde{D}$ 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 $a^a_\mu$ and $a^{ab}_\mu$ which corresponded to small eigenvalues of $S$. The eigenvalue truncation thresholds are denoted hereafter as $\tau_s$ for the singly external and semi-internal and $\tau_d$ for the doubly external excitations. This limits the largest linear combination amplitude coefficients (e.g. $S^{\mu,1/2}_{\mu,i}$) appearing in eqns. (48), (49) 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 orthogonalised operator space was 410 ($R_{NN} = 1.0$), 698 ($R_{NN} = 1.6$), and 983 ($R_{NN} = 3.0$) 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 non-zero overlap in $S_{i,jk}$ (eqn. (15)). The orthogonalisation 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 equa-
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 orthogonalisation.

|               | 1Å   | 2Å   | 3Å   |
|---------------|------|------|------|
| 6-31G         |      |      |      |
| exact orthog. | -109.04585 | -108.86156 | -108.84155 |
| cumulant orthog. | -109.04566 | -108.86220 | -108.84155 |
| diff (m$E_h$) | +0.19 | -0.64 | 0.0  |
| cc-pVDZ       |      |      |      |
| exact orthog. | -109.22774 | -108.98214 | -108.96009 |
| cumulant orthog. | -109.22752 | -108.98305 | -108.96009 |
| diff (m$E_h$) | +0.22 | -0.91 | 0.0  |
| cc-pVTZ       |      |      |      |
| exact orthog. | -109.33525 | -109.05709 | -109.03011 |
| cumulant orthog. | not conv. | -109.05824 | -109.03011 |
| diff (m$E_h$) | -1.15 | 0.0  |      |

with the standard truncation thresholds, for example at $R_{NN} = 1\text{Å}$ for $N_2$ (cc-pVTZ), as discussed in Sec. [IV.A]

C. Operator orthogonalisation with cumulant density matrix

Rather than using the exact three-particle density matrix for the orthogonalisation procedure described above, we could also imagine using its cumulant decomposition in terms of the one- and two-particle density matrices, eqn. [10].

Table [XIII] shows the differences of the total energies computed using the amplitude operators that are orthogonalised with the exact and approximate (cumulant) three-particle density matrices for the $N_2$ potential energy curves discussed in section [IV.A]. With the truncation threshold $\tau_s = 10^{-1}$ and $\tau_d = 10^{-2}$ and using various basis sets, the energies from both orthogonalisations were generally in good agreement within a deviation of 1.2 m$E_h$. However, at $R_{NN} = 1\text{Å}$ with the cc-pVTZ basis set, the L-CTSD(MK) calculation with the cumulant based orthogonalisation did not converge.
[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. Langhoff 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. Gdnitz 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, edited by D. R. Yarkony, Advanced Series in Physical Chemistry, pages 1047–1131, World Scientific, River Edge, NJ, 1995.
[49] S. Hirata, J. Phys. Chem. A 107, 4940 (2003).
[50] T. Yanai et al., Utchem – a program for ab initio quantum chemistry., in Lecture Notes in Computer Science, pages 84–95, Berlin, 2003, Springer-Verlag.
[51] P. O. Widmark, P. Ä. Malmqvist, and B. O. Roos, Theor. Chim. Acta 77, 291 (1990).
[52] R. Pou-Amérgio, 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. 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] MOLPRO, version 2006.1, a package of ab initio programs, H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, and others.
[57] This separation of core and valence densities can be made precise in Thomas-Fermi theory, see e.g. E. H. Lieb, Rev. Mod. Phys. 48, 553 (1976).
