

Strongly contracted canonical transformation theory

Eric Neuscamman¹, Takeshi Yanai, and Garnet Kin-Lic Chan

Citation: *J. Chem. Phys.* **132**, 024106 (2010); doi: 10.1063/1.3274822

View online: <http://dx.doi.org/10.1063/1.3274822>

View Table of Contents: <http://aip.scitation.org/toc/jcp/132/2>

Published by the American Institute of Physics

Strongly contracted canonical transformation theory

Eric Neuscamman,^{1,a)} Takeshi Yanai,² and Garnet Kin-Lic Chan¹

¹Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853, USA

²Institute of Molecular Science, 38 Nishigo-Naka, Myodaiji, Okazaki 444-8585, Japan

(Received 3 September 2009; accepted 25 November 2009; published online 13 January 2010)

Canonical transformation (CT) theory describes dynamic correlation in multireference systems with large active spaces. Here we discuss CT theory's intruder state problem and why our previous approach of overlap matrix truncation becomes infeasible for sufficiently large active spaces. We propose the use of strongly and weakly contracted excitation operators as alternatives for dealing with intruder states in CT theory. The performance of these operators is evaluated for the H₂O, N₂, and NiO molecules, with comparisons made to complete active space second order perturbation theory and Davidson-corrected multireference configuration interaction theory. Finally, using a combination of strongly contracted CT theory and orbital-optimized density matrix renormalization group theory, we evaluate the singlet-triplet gap of free base porphyrin using an active space containing all 24 out-of-plane 2p orbitals. Modeling dynamic correlation with an active space of this size is currently only possible using CT theory. © 2010 American Institute of Physics.

[doi:10.1063/1.3274822]

I. INTRODUCTION

Systems containing a large number of strongly correlated electrons represent a particular challenge for current theoretical methods in quantum chemistry. Historically, these systems have been modeled in a two step process. First the *static* electron correlation, which involves interactions between nearly degenerate electronic configurations and is responsible for the qualitative shape of the wave function, is established through an active space method, usually complete active space self-consistent field (CASSCF) theory.^{1,2} Second, the *dynamic* correlation, which can often be viewed as a small perturbation to the static correlation which is necessary for chemical accuracy, is treated with a method such as complete active space second order perturbation (CASPT2) theory,³⁻⁵ multireference Moller-Plesset theory,⁶ n-electron valence perturbation (NEVPT2) theory,⁷⁻⁹ or a corrected configuration interaction (CI) method such as Davidson-corrected multireference configuration interaction (MRCI+Q) theory.¹⁰⁻¹⁴ While these methods have been successful in modeling systems with a small number of strongly correlated electrons, their formulation in terms of the exponentially numerous active space CI coefficients (for uncontracted methods) or the three- and four-body reduced density matrices (for contracted methods) prevents them from treating systems with more than 16 active orbitals. This limitation has until recently been of marginal concern, as the CASSCF method upon which these theories rely is itself limited to 16 active orbitals. However, with the development of methods such as density matrix renormalization group (DMRG) theory,¹⁵⁻²¹ restricted active space (RAS) theory,^{22,23} and generalized valence bond theories,^{24,25} it has become possible to model static correlation in systems containing more than 16 active orbitals.

In light of these developments, it becomes necessary to re-examine the treatment of dynamic correlation so that it too can be modeled in systems with large active spaces. One recent approach is RASPT2 theory, which has been used to model dynamic correlation in *restricted* active spaces of up to 32 orbitals.²⁶ By restricting the occupations of the high- and low-energy orbitals in the active space, the RASPT2 method greatly reduces the number of electronic configurations used to treat static correlation. While this restriction improves efficiency and marks a significant advance relative to CASPT2 theory, it also produces a number of difficulties (such as a loss of size extensivity) that are not present when using a complete active space. Another recent development is cumulant-approximated n-electron valence perturbation theory (cu-NEVPT2),²⁷ which can be applied to large active spaces but suffers from poor accuracy. This paper deals with a third approach, canonical transformation theory, which can be applied to *complete* active spaces of more than 30 orbitals.

Canonical transformation (CT) theory²⁸⁻³¹ is a rigorously size-extensive method for treating dynamic correlation in multireference systems. The theory is based on a unitary exponential ansatz similar to that of unitary coupled cluster (CC) theory³²⁻³⁸ and some multireference CC theories.³⁹⁻⁴² By employing operator^{29,31,43,44} and cumulant^{29,31,45-47} decompositions, CT theory produces a two-body effective Hamiltonian through an approximate Baker-Campbell-Hausdorff expansion. Unlike traditional multireference dynamic correlation methods, CT theory can be applied efficiently in conjunction with complicated wave functions such as those from DMRG theory because the only wave function information it requires are the one- and two-body reduced density matrices (RDMs) in the active space. Furthermore, CT theory has a lower cost scaling (n^6) than either CASPT2 theory (n^8) or MRCI+Q theory (n^{10}) when the number of strongly correlated electrons is assumed to be proportional to

^{a)}Electronic mail: eric.neuscamman@gmail.com.

the total number of electrons. Previous studies^{28,29,31} on H₂O, N₂, and FeO have shown CT theory to be more accurate than CASPT2 and competitive in accuracy with MRCI+Q.

CT theory is not without difficulties, however, the most prominent of which are intruder states. Previous work has demonstrated that as written, the CT equations are often too poorly conditioned to be solved numerically. A new type of intruder state, which arises from the cumulant and operator decomposition approximations, is the source of this difficulty. These decomposition intruder states differ from the traditional intruder states of CASPT2 theory, which are caused by inadequate zeroth order Hamiltonians.⁴⁸ They are also distinct from the problem of redundant states, in which a multireference dynamic correlation theory may have a linearly degenerate first order interacting basis. While these three types of problematic states have different origins, the numeric problem that they lead to is the same; they each produce unphysically small eigenvalues in the system's Jacobian matrix. Previously, the problem of intruder states in CT theory was addressed by the overlap truncation method,^{28,29,31} which is an aggressive truncation of the first order interacting basis. While this approach has been successful in small systems, it relies on two components that become intractable in large systems. The first is the n_{act}^9 (n_{act} =number of active orbitals) cost of orthogonalizing the first order interacting basis. The second is the arbitrary choice of a truncation threshold, which in difficult systems such as NiO can influence the CT energy.

In this work we propose the use of strongly contracted excitation operators, first introduced by Malrieu in the context of NEVPT2 theory,^{7,8} and the closely related weakly contracted excitation operators as alternatives to the overlap truncation method. These operators intelligently restrict the first order interacting basis in a way that helps to avoid intruder states while retaining the excitations most important for capturing dynamic correlation. While these operators do not necessarily avoid all intruder states (as seen in our small molecule results), the simplicity they bring to the first order interacting basis allows any intruders that remain to be efficiently detected and removed. In the large active spaces explored in this paper, those of NiO and free base porphyrin, the strongly contracted operators are so effective at avoiding intruder states that this subsequent removal becomes unnecessary and the potential energy surfaces are smooth. In smaller molecules, removal of the remaining intruder states creates discontinuities in the potential energy surface. Although these discontinuities are likely to lead to difficulties in the evaluation of gradients, they are small enough that they do not affect the accuracy of energy differences. Finally and most critically for their application to large active spaces, the strongly contracted operators accomplish the above tasks without the expensive orthogonalization step present in the overlap truncation method, and thus produce a method with an n^6 cost scaling regardless of how large the active space is.

We begin by describing the nature of intruder states in CT theory (Sec. II A), followed by the three methods for dealing with them: overlap truncation (Sec. II B), strong contraction (Sec. II C), and weak contraction (Sec. II D). These methods are then applied to three benchmark systems to as-

sess their capabilities: H₂O (Sec. III A), N₂ (Sec. III B), and NiO (Sec. III C). Finally, as a demonstration of CT theory's ability to model large numbers of strongly correlated electrons, we apply a combination of orbital-optimized DMRG (DMRG-SCF) theory^{18,19} and CT theory with strongly contracted operators to calculate the singlet-triplet gap in free base porphyrin using a complete active space of 24 orbitals (Sec. III D). A more detailed discussion of the combination of DMRG-SCF and CT is given elsewhere.⁴⁹

II. THEORY

A. Intruder states in CT theory

We begin our theoretical discussion by describing the difficulties encountered when solving the CT amplitude equations and comparing them to intruder states in second order perturbation (PT2) theory. In CT theory, we attempt to find an approximate effective Hamiltonian \bar{H} for which the reference wave function $|\Psi_0\rangle$ is an eigenstate,

$$\bar{H} = e^{-A} H e^A = H + [H, A]_{1,2} + \frac{1}{2!} [[H, A]_{1,2}, A]_{1,2} + \dots, \quad (1)$$

$$\bar{H}|\Psi_0\rangle = E|\Psi_0\rangle. \quad (2)$$

Here the notation $[H, A]_{1,2}$ indicates that we have approximated the three-body operators resulting from the commutator using a combination of one- and two-body operators. This approximation is defined by neglecting the spin free three-body cumulant and spin free three-body extended-normal-ordered (ENO) operator from Eq. 65c of Ref. 43. The three-body cumulant and ENO operator can be interpreted as irreducible three-body fluctuations from one- and two-body behavior that CT theory neglects. The explicit form for the approximation of three-body spin free operators is given in the supplemental information.⁵⁰ The operator A that defines \bar{H} consists of a sum of antisymmetric excitation operators \hat{o}_i ,

$$A = -A^\dagger = \sum_i C_i \hat{o}_i, \quad (3)$$

which for example could include one- and two-body operators such as $(E_{a_1}^{v_1} - E_{v_1}^{a_1})$ and $(E_{a_1 a_2}^{v_1 v_2} - E_{v_1 v_2}^{a_1 a_2})$. For now we will not define the exact structure of the excitation operators \hat{o}_i except to say that they consist of one- and two-body operators that when applied to the reference wave function create at least one hole in the core (doubly occupied) space or one particle in the virtual (unoccupied) space. Note that the effective Hamiltonian \bar{H} depends on the coefficients C_i , which can be found by solving the nonlinear projected amplitude equations,

$$R_i = \langle \Psi_0 | [\bar{H}, \hat{o}_i]_{1,2} | \Psi_0 \rangle = 0, \quad (4)$$

using the iterative Newton-Raphson method. This approach requires solving the linear equation

$$\sum_j J_{ij} X_j = -R_i, \quad (5)$$

where X_j are the corrections to the current coefficients C_j and J is the CT Jacobian matrix,

$$J_{ij} = \frac{\partial R_i}{\partial C_j} = \langle \Psi_0 | [[\bar{H}, \hat{\delta}_j]_{1,2}, \hat{\delta}_i]_{1,2} | \Psi_0 \rangle + O(A). \quad (6)$$

For the typical initial guess $C_i=0$, the Newton–Raphson equation becomes

$$\begin{aligned} \sum_j J_{ij} X_j &= \sum_j \langle \Psi_0 | [[H, \hat{\delta}_j]_{1,2}, \hat{\delta}_i]_{1,2} | \Psi_0 \rangle X_j \\ &= -\langle \Psi_0 | [H, \hat{\delta}_i]_{1,2} | \Psi_0 \rangle, \end{aligned} \quad (7)$$

where the effective Hamiltonian \bar{H} has simplified to the original H . A connection to perturbation theory is now apparent, for if we replace the Hamiltonian on the left side of Eq. (7) with a zeroth order Hamiltonian satisfying $H_0 | \Psi_0 \rangle = E_0 | \Psi_0 \rangle$ and remove the decomposition approximations, we obtain the defining equation for PT2 theory,

$$\begin{aligned} \sum_j (J_0)_{ij} X_j &= \sum_j \langle \Psi_0 | \hat{\delta}_i^\dagger (H_0 - E_0) \hat{\delta}_j | \Psi_0 \rangle X_j \\ &= -\langle \Psi_0 | \hat{\delta}_i^\dagger H | \Psi_0 \rangle. \end{aligned} \quad (8)$$

Here the coefficients X_j are the weights of the first order interacting states $\hat{\delta}_j | \Psi_0 \rangle$ in the first order wave function of PT2 theory. Ignoring operator decomposition, we see that the first iteration of the Newton–Raphson method in CT theory differs from PT2 theory by only the Hamiltonian used to define the Jacobian matrix. In this way the relationship between Eqs. (7) and (8) resembles the relationship between the amplitude equations of linearized CC theory and second order Moller–Plesset theory.

In both CT and PT2 theory, we can identify intruder states as those with large values of X_i , which result from unphysically small eigenvalues in J and J_0 . It is important to differentiate intruder states from redundancies that may occur in the first order interacting basis $\hat{\delta}_i | \Psi_0 \rangle$, which for a multiconfigurational reference function is not necessarily linearly independent. Such redundancies create small Jacobian eigenvalues in both CT theory and PT2 theory. These redundant states aside, both PT2 theory and CT theory suffer from true intruder states, which create small Jacobian eigenvalues due to the methods’ approximations and have nothing to do with an overcompleteness of the basis. To see the source of intruders in PT2 theory, recognize that in Eq. (8) the Jacobian eigenvalues are excitation energies of the zeroth order Hamiltonian corresponding to transitions from $| \Psi_0 \rangle$ to states in the first order interacting basis. Thus “traditional” intruder states are associated with a zeroth order Hamiltonian H_0 that incorrectly predicts near-degeneracy between the reference state and states in the first order interacting basis. In CT theory, the Jacobian is defined using the exact Hamiltonian, and so intruder states must instead be caused by the decomposition approximations present in Eq. (6). (Strictly speaking, the Jacobian is defined using \bar{H} , but for the initial guess $C_i=0$ this reduces to the exact electronic Hamiltonian H .) We identify these states as “decomposition” intruder states, and

note that their origin is different from the “traditional” intruders found in PT2 theory. A particularly troublesome quality of these new intruder states is that they cannot be removed by expanding the active space, as doing so does not make the decomposition approximations any more accurate. As shown by the investigation of Zgid *et al.* of cu-NEVPT2 theory,²⁷ the decomposition intruder states are not unique to CT theory. Indeed, the use of cumulant-approximations in cu-NEVPT2 led to intruder state problems that were every bit as severe as those encountered in CT theory.

In order for CT theory to be tractable, these decomposition intruder states must be avoided. The following three sections discuss how we attempt to do so using special definitions for the excitation operators $\hat{\delta}_i$. These definitions seek to restrict the first order interacting basis in such a way as to avoid intruder states while retaining sufficient freedom to accurately describe the system’s dynamic correlation. The degree to which this balance can be achieved is difficult to predict in advance and can only be satisfactorily determined through applications of these operator definitions to real systems. Such applications will be presented in Sec. III, in which we will analyze the definitions’ results in three benchmark systems.

B. Overlap matrix truncation

In previous presentations of CT theory, the problem of small Jacobian eigenvalues (intruder states and redundant states) was circumvented by defining operators $\hat{\delta}_i$ that were orthonormalized and truncated with respect to the first order interacting basis $\hat{\delta}_i | \Psi_0 \rangle$. For an arbitrary choice of excitation operators \hat{p}_i , the basis $\hat{p}_i | \Psi_0 \rangle$ is not orthonormal and there is a dense overlap matrix $S_{ij} = \langle \Psi_0 | \hat{p}_i^\dagger \hat{p}_j | \Psi_0 \rangle$. In this case, the eigenvalues of J are defined by the generalized eigenvalue equation,

$$\sum_j J_{ij} B_{jk} = \epsilon_k \sum_l S_{il} B_{lk}, \quad (9)$$

where B is the matrix of Jacobian eigenvectors and ϵ_k is the k th eigenvalue. If we instead work in the orthonormal basis $\hat{\delta}_i = \sum_j (S^{-1/2})_{ij} \hat{p}_j | \Psi_0 \rangle$, we obtain a simple eigenvalue equation in terms of the orthonormalized matrices $\tilde{J} = S^{-1/2} J S^{-1/2}$ and $\tilde{B} = S^{1/2} B$,

$$\sum_j \tilde{J}_{ij} \tilde{B}_{jk} = \epsilon_k \tilde{B}_{ik}. \quad (10)$$

Assuming that the Jacobian is similar to that of PT2 theory as discussed in Sec. II A, the eigenvalues ϵ_k should be no smaller than the energy cost to excite from the reference state to the lowest lying state in the first order interacting basis. This condition requires that the small eigenvalues present in S and J cancel in the product $S^{-1/2} J S^{-1/2}$. Error in either of these matrices can prevent this cancellation and produce unphysically small values of ϵ_k . In CT theory, operator decomposition essentially guarantees that such errors will be present. We therefore define our excitation operators as $\hat{\delta}_i = \sum_j (S^{-1/2})_{ij} \hat{p}_j | \Psi_0 \rangle$, where S has been truncated by removing its eigenvalues below some threshold τ . This choice avoids the miscancellation of small eigenvalues and produces an

orthonormal Jacobian without intruder states or redundant states. In practice, two different thresholds are used: τ_1 for the overlap matrices of single and semi-internal excitation operators and τ_2 for the overlap matrices of double excitation operators. (Semi-internal excitations are double excitations in which one electron is excited from one active orbital to another while the other electron is excited into a virtual orbital and/or out of a core orbital.)

Although the overlap truncation method has proven successful in systems such as N_2 , H_2O , and FeO ,²⁹ it possesses significant disadvantages. One problem is that the choice of τ is arbitrary, and indeed for some systems the choice is not obvious (see Sec. III C). A second problem is that the set of excitation operators \hat{o}_i may change discontinuously with system geometry as overlap eigenvalues pass above or below the truncation threshold, which can produce discontinuous potential energy surfaces (see Sec. III). Finally, with regard to large active spaces, there is the challenge of computing and diagonalizing the three-body RDM. While the cumulant decomposition can be used to avoid computing the true three-body RDM, this produces errors in S which make it more difficult to select good values for τ_1 and τ_2 . With or without the exact three-body RDM, diagonalizing the semi-internal overlap matrices has an n_{act}^9 cost scaling. For small active spaces this diagonalization is trivial, but for the large active spaces common to DMRG wave functions it becomes unaffordable. In an attempt to address these difficulties, we turn to strongly contracted excitation operators.

C. Strong contraction

Strongly contracted excitation operators^{7,8} are a compact yet powerful organization of a multireference system's excitation degrees of freedom. The basic idea is to group together all excitation operators sharing the same external (nonactive) indices into a single excitation operator, in which each individual excitation is given some weight. Formally, they are defined by projecting the Hamiltonian onto subspaces corresponding to different sets of external orbitals. The number and type (core or virtual) of external orbitals determine the type of excitation, while the specific orbitals chosen identify the unique strongly contracted operator. For example, the external orbital set (v_1, v_2) defines a strongly contracted double excitation operator that promotes two electrons from the active space to occupy the virtual orbitals v_1 and v_2 . The explicit form of this operator in CT theory is

$$\hat{o}^{v_1 v_2} = \sum_{a_1 a_2} g_{a_1 a_2}^{v_1 v_2} (E_{a_1 a_2}^{v_1 v_2} - E_{v_1 v_2}^{a_1 a_2}), \quad (11)$$

an antisymmetrized projection of the Hamiltonian onto a space of operators that promote from active orbitals into v_1 and v_2 . Here g is the tensor of two-electron integrals appearing in the Hamiltonian,

$$H = \sum_{p_1 q_1} t_{q_1}^{p_1} E_{q_1}^{p_1} + \frac{1}{2} \sum_{p_2 p_3} g_{q_2 q_3}^{p_2 p_3} E_{q_2 q_3}^{p_2 p_3}, \quad (12)$$

and E is a spin-free excitation operator,

$$E_{q_1}^{p_1} = \sum_{\sigma=\alpha}^{\beta} a_{p_1 \sigma}^{\dagger} a_{q_1 \sigma}, \quad (13)$$

$$E_{q_1 q_2}^{p_1 p_2} = \sum_{\sigma, \tau=\alpha}^{\beta} a_{p_1 \sigma}^{\dagger} a_{p_2 \tau}^{\dagger} a_{q_2 \sigma} a_{q_1 \tau}. \quad (14)$$

In this paper we use the indices c , a , and v to represent core (doubly occupied), active, and virtual (unoccupied) orbitals, respectively, while p and q represent general orbitals. For single and double excitations, there are a total of eight types of strongly contracted operators, the precise definitions of which are given in the Appendix.

As seen in Eq. (11), strongly contracted operators combine many excitations out of (or into) the active space with regard to a certain set of external orbitals. This grouping of excitations would be exactly justified if the energies of states in the first order interacting basis depended only on their external indices, which is approximately true if the energy gaps between the core, active, and virtual spaces are large relative to the energy spread within the active space. We therefore expect strongly contracted operators to be most effective in systems for which such an active space can be chosen, such as the π -orbital active spaces commonly used in polyenes. For ground state calculations, the use of strongly contracted operators can also be justified by considering a Krylov expansion of the reference wave function, $\{H|\Psi_0\rangle, H^2|\Psi_0\rangle, H^3|\Psi_0\rangle, \dots\}$. This expansion is commonly used as a basis in which to solve the eigenvalue problem $H|\Psi\rangle = E|\Psi\rangle$ for the extremal eigenvalue because each vector in the series is increasingly similar to the desired eigenvector. It is this preference toward the ground state that makes strong contraction effective. If we were to formulate a theory in which only a single excitation operator was used, then by a Krylov argument H itself would be an excellent choice. Strong contraction restores some flexibility by separating the components of H in each external subspace. Certainly this description is more limited than if we did not group excitation operators together, and we will show in our results that CT theory's accuracy can suffer as a result. As discussed earlier, however, intruder states prevent the use of the complete set of single and double excitations. From this perspective, strong contraction can be viewed as a practical formulation of CT theory that seeks to restrict the first order interacting basis in a way that avoids intruder states but retains the excitations important for describing dynamic correlation.

An important property of strongly contracted operators is that they are orthogonal by construction, as each operator corresponds to a different set of external orbitals (e.g., $\langle\Psi_0|\hat{o}_{c_1}^{\dagger}\hat{o}_{c_2}|\Psi_0\rangle=0$ if $c_1 \neq c_2$.) This property has two key consequences. First, no redundant states will be present in the strongly contracted first order interacting basis. Second, no diagonalization of the basis' overlap matrix is necessary, removing what for large active spaces is the most expensive step in the overlap truncation method. We must still evaluate the operators' norms, which formally require the three-body RDM for the semi-internal operators \hat{o}^{v_1} and \hat{o}_{c_1} . However,

we have found that the cumulant approximation produces sufficiently accurate norms in most cases. These approximate operator norms,

$$\|\hat{o}_i\|^2 \simeq \langle \Psi_0 | \hat{o}_i^\dagger \hat{o}_i | \Psi_0 \rangle_{1,2}, \quad (15)$$

can be evaluated in n^6 time using only the one- and two-body RDMs because the three-body spin free cumulant has been neglected (as indicated by the $_{1,2}$ subscript).

In some systems, particularly NiO and free base porphyrin, we have observed that the use of strongly contracted excitation operators produces a first order interacting basis free of intruder states. This likely arises due to both the orthogonality of the strongly contracted operators and the fact that they severely restrict the basis' freedom, making it more difficult to form a linear combination of excited states with intruder character. However, in some systems, particularly those with smaller active spaces in which strong contraction is a less drastic constraint on the basis' freedom, the use of strongly contracted operators does not by itself remove all intruder states. To detect and remove from the basis any remaining intruders, we recall that intruder states are eigenstates of the Jacobian with unphysically small eigenvalues. Thus if we knew the eigenvalues, we could identify and remove the intruders directly. Unfortunately, evaluating these eigenvalues exactly is too expensive, so we instead approximate them by replacing the Hamiltonian present in the Jacobian's definition with Dyall's zeroth order Hamiltonian,⁵¹

$$H_0 = C + \sum_{c_1} \bar{t}_{c_1}^1 E_{c_1}^{c_1} + \sum_{v_1} \bar{t}_{v_1}^1 E_{v_1}^{v_1} + \sum_{a_1 a_2} \bar{t}_{a_2}^{a_1} E_{a_2}^{a_1} + \frac{1}{2} \sum_{a_1 a_2} g_{a_3 a_4}^{a_1 a_2} E_{a_3 a_4}^{a_1 a_2}, \quad (16)$$

in which C is a constant, \bar{t} is a set of effective one-body integrals, and g is the usual two-body integral tensor. Note that C and \bar{t} are defined such that $H_0 |\Psi_0\rangle = E_0 |\Psi_0\rangle$. The reason we choose this form for H_0 is that it makes the Jacobian matrix diagonal in the strongly contracted operators, as H_0 cannot connect operators with different sets of external orbital indices. This simplification gives us a set of approximate Jacobian eigenvalues ϵ_i that may be evaluated in n^6 time,

$$\epsilon_i = \frac{1}{\|\hat{o}_i\|^2} \langle \Psi_0 | [[H_0, \hat{o}_i]_{1,2}, \hat{o}_i]_{1,2} | \Psi_0 \rangle. \quad (17)$$

Note that Dyall's zeroth order Hamiltonian is also used in NEVPT2 theory, and that our approximate Jacobian eigenvalues can be seen as operator-decomposed approximations of the NEVPT2 denominators.

As discussed in Sec. II A, the eigenvalues ϵ_i should be at least as large as the Dyall Hamiltonian's lowest excitation energy between the reference function and the first order interacting basis. Thus for an active space large enough to prevent traditional intruder states (which could arise from the use of H_0), we can use Eq. (17) to measure how severe the decomposition error is for a given strongly contracted operator. If an approximate eigenvalue ϵ_i is unphysically small, the decomposition error for the corresponding operator \hat{o}_i is

large and a decomposition intruder state may occur. We therefore remove from our excitation operator basis all operators \hat{o}_i for which ϵ_i is below some threshold τ_ϵ . Unlike the truncation thresholds employed in the overlap truncation method, τ_ϵ is not arbitrary but rather is chosen to be a reasonable value for the excitation energy between the reference function and the lowest state in the first order interacting basis. It is important at this point to note that because we formulate the semi-internal operator norms approximately, it is possible that for some operators $\|\hat{o}_i\|^2$ will be negative. These operators are also excluded from our excitation basis, as they are clearly affected adversely by the cumulant approximation. Together, Eqs. (15) and (17) are used to remove any remaining intruder states from the strongly contracted first order interacting basis. The explicit tensor contractions required by these equations for the various types of strongly contracted operators have been derived automatically³¹ and are included in the supplemental material.⁵⁰ To summarize, the strongly contracted operator basis is constructed via the following steps:

- (1) construct strongly contracted operators using Eqs. (A1)–(A8),
- (2) remove operators with negative approximate norms, Eq. (15), and
- (3) remove operators with approximate Jacobian eigenvalues below τ_ϵ , Eq. (17).

In conclusion, strongly contracted CT theory (SC-CTSD) requires only the one- and two-body RDMs of the reference wave function, can be evaluated in n^6 time, and should have no intruder states when a sufficiently large active space is used.

D. Weak contraction

There exists a simple extension of strong contraction in the semi-internal subspaces (containing double excitations with one active-to-active excitation and one external excitation) which we term weak contraction (WC-CTSD). Rather than summing over all the active indices in these subspaces, we can fix one index, creating the following weakly contracted excitation operators,

$$\hat{o}_{a_1}^{v_1} = \left(t_{a_1}^{v_1} + \sum_{c_2} (2g_{a_1 c_2}^{v_1 c_2} - g_{a_1 c_2}^{c_2 v_1}) \right) (E_{a_1}^{v_1} - E_{v_1}^{a_1}) + \sum_{a_2 a_3} g_{a_1 a_2}^{v_1 a_3} (E_{a_1 a_2}^{v_1 a_3} - E_{v_1 a_3}^{a_1 a_2}), \quad (18)$$

$$\hat{o}_{c_1}^{a_1} = \left(t_{c_1}^{a_1} + \sum_{c_2} (2g_{c_1 c_2}^{a_1 c_2} - g_{c_1 c_2}^{c_2 a_1}) \right) (E_{c_1}^{a_1} - E_{a_1}^{c_1}) + \sum_{a_2 a_3} g_{c_1 a_3}^{a_1 a_2} (E_{c_1 a_3}^{a_1 a_2} - E_{a_1 a_2}^{c_1 a_3}). \quad (19)$$

An important property of this partitioning of the Hamiltonian's operators is that each bare excitation operator (e.g., $E_{c_1}^{a_1}$ or $E_{c_1 a_3}^{a_1 a_2}$) appears exactly once among the weakly contracted excitation operators. This ensures that all of the Hamiltonian's semi-internal excitations are included while preventing explicit redundancies between the components of

the different weakly contracted excitation operators. These operators are related to their strongly contracted counterparts by the fact that Eqs. (A7) and (A8) can be recovered by summing over the a_1 index in Eqs. (18) and (19). Thus we recognize that weakly contracted operators are simply a further partitioning of the Hamiltonian to increase flexibility in the semi-internal subspaces.

The same general procedures for avoiding intruder states in strong contraction are followed when using weakly contracted excitation operators, except that the overlap and approximate Jacobian matrices of Eqs. (20) and (21) are now block diagonal with respect to the virtual or core index, with block dimension equal to the number of active orbitals,

$$S_{a_1 a_2}^{v_1} \simeq \langle \Psi_0 | (\hat{\delta}_{a_1}^{v_1})^\dagger \hat{\delta}_{a_2}^{v_1} | \Psi_0 \rangle_{1,2}, \quad (20)$$

$$J_{a_1 a_2}^{v_1} = \langle \Psi_0 | [[H_0, \hat{\delta}_{a_1}^{v_1}]_{1,2}, \hat{\delta}_{a_2}^{v_1}]_{1,2} | \Psi_0 \rangle. \quad (21)$$

Building and diagonalizing these matrices has a cost that is at most n^7 . We do not expect this step to be the bottleneck in CT theory, however, as it is performed once while the n^6 cost commutator and residual evaluations must be performed repeatedly during the iterative solution of the amplitude equations. The explicit tensor contractions required to build these matrices have been derived automatically³¹ and are available in the supplemental information.⁵⁰ Note that when using weak contraction we only apply it to the core-to-active and active-to-virtual semi-internal subspaces. The other six subspaces are treated with strong contraction.

III. RESULTS

Here we present results for H₂O, N₂, NiO, and free base porphyrin. For CASSCF, CASPT2, complete active space third order perturbation theory (CASPT3) and MRCI+Q calculations we have used the MOLPRO program package.⁵² Before getting into the details of the results, it is important to explain the origin and consequences of discontinuities in the potential energy surfaces produced by CT theory. Unlike CASPT2 and other perturbation theories, which use a level shift to remove intruder states, the decomposition intruder states in CT theory are removed by ejecting states with eigenvalues below some threshold from the first order interaction basis. In the case of overlap truncation (CTSD), this ejection is based on the eigenvalues of the basis' overlap matrix, while in strong contraction (SC-CTSD) and weak contraction (WC-CTSD) it is based on the eigenvalues of the approximate Jacobian matrix. In either case, smoothly varying the geometry of a molecule may cause an eigenvalue to cross the truncation threshold, which produces a discontinuous change in the first order interacting basis and thus the CT energy. The size of this energy discontinuity is controlled by the importance of the state in question in describing the system's dynamic correlation. As we shall see, the size of these discontinuities is small for SC-CTSD when compared to the energy scale of the potential energy surface (see Figs. 2 and 4) and will therefore not affect properties based on energy differences, such as dissociation energies, reaction barriers, and vibrational frequencies. Nonetheless, any discontinuity can cause difficulties when performing a geometry optimization,

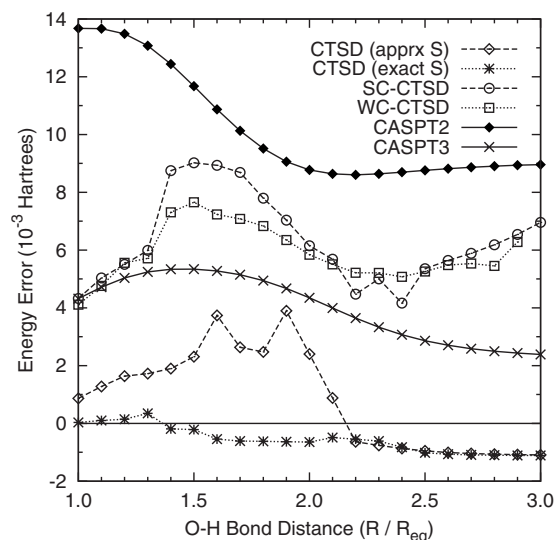


FIG. 1. Ground state energy errors relative to MRCI+Q for the symmetric stretch of H₂O in the cc-pVDZ basis set. $R_{\text{eq}}=0.9929$ Å. Bond angle = 109.57°. CTSD employed truncation thresholds of ($\tau_1=0.2$, $\tau_2=0.1$) and ($\tau_1=0.1$, $\tau_2=0.01$) when using the approximate and exact semi-internal overlap matrices, respectively. Both SC-CTSD and WC-CTSD used an energy threshold of $\tau_e=0.1 E_h$.

especially when numeric derivatives are employed. Significantly, in the larger active spaces of NiO and free base porphyrin, the strongly contracted operator basis is restrictive enough that intruder states are avoided entirely without the need for any thresholds or basis set truncation. Thus in these systems, SC-CTSD produces completely smooth potential energy surfaces. The WC-CTSD method (and CTSD in the case of NiO) has more difficulty with intruder states and is therefore less capable of producing smooth potential energy surfaces.

A. Water

We have computed the energy curve for the symmetric stretch of H₂O in the Dunning cc-pVDZ basis set⁵³ with spherical d orbitals. The active space consists of five orbitals: the oxygen 2p orbitals and the hydrogen 1s orbitals. All orbitals were correlated. The results are shown in Figs. 1 and 2 and Table I, in which MRCI+Q is used as a benchmark.

The CTSD method shows a relative error of $1.5 mE_h$ when the semi-internal overlap matrix S_{int} is evaluated exactly using the three-body RDM. (Note that other than for the truncation of S_{int} , all other terms in CTSD use only the one- and two-body RDMs. This is the standard CTSD method with overlap truncation introduced in Ref. 29 and identified as LCTSD in Ref. 31.) However, when S_{int} is approximated with the cumulant decomposition, CTSD has a significantly larger relative error of $5.0 mE_h$. The accuracy suffers because the truncation thresholds τ_1 and τ_2 must be increased to prevent the cumulant decomposition approximations in S_{int} from creating intruder states. This more aggressive truncation results in a more limited first order interacting basis which reduces accuracy. The SC-CTSD and WC-CTSD methods, which do not require the three-body RDM, show relative errors of 4.9 and $3.5 mE_h$, respectively. All of the CT methods are more accurate than CASPT2, whose relative

TABLE I. Results for H₂O in the cc-pVDZ basis. MRCI+Q is reported in E_h, with other methods reported as the difference from MRCI+Q in mE_h. All methods use a (6e,5o) active space. All orbitals were correlated. The H–O–H bond angle was fixed at 109.57°.

R/R _{eq} ^a	MRCI+Q	CASSCF	CASPT2	CASPT3	CTSD ^b	CTSD ^c	SC-CTSD ^d	WC-CTSD ^d
1.0	-76.241 466	165.601	13.674	4.321	0.866	0.034	4.326	4.110
1.2	-76.182 723	161.881	13.487	5.038	1.638	0.145	5.505	5.558
1.4	-76.101 419	155.832	12.442	5.336	1.900	-0.187	8.753	7.309
1.6	-76.031 316	148.035	10.874	5.278	3.741	-0.541	8.937	7.236
1.8	-75.980 391	140.359	9.514	4.945	2.477	-0.620	7.796	6.831
2.0	-75.947 828	134.194	8.774	4.350	2.402	-0.643	6.150	5.840
2.2	-75.929 506	130.039	8.606	3.646	-0.633	-0.549	4.475	5.216
2.4	-75.920 239	127.638	8.697	3.071	-0.869	-0.824	4.165	5.074
2.6	-75.915 748	126.371	8.820	2.705	-1.007	-1.064	5.641	5.483
2.8	-75.913 526	125.713	8.908	2.499	-1.065	-1.102	6.185	5.454
3.0	-75.912 381	125.364	8.959	2.387	-1.088	-1.109	6.960	No convergence
Relative error		40.237	5.068	2.956	4.988	1.458	4.855	3.548

^aR_{eq}=0.9929 Å.

^bτ₁=0.2, τ₂=0.1, approximate S_{int}.

^cτ₁=0.1, τ₂=0.01, exact S_{int}.

^dτ_ε=0.1 E_h.

error is 5.1 mE_h. The discontinuity problem discussed above is more severe for SC-CTSD and WC-CTSD than for CTSD, although this is a somewhat unfair comparison as the latter uses the exact three-body RDM to evaluate its overlap matrix. Indeed, the SC-CTSD and WC-CTSD discontinuities are less severe than those that occur in CTSD when the cumulant-approximated overlap matrix is used. Fortunately, as shown in Fig. 2, the discontinuities present in SC-CTSD are small enough that they do not affect the shape of the potential energy curve.

B. Nitrogen

We have computed the energy curve for the bond breaking of N₂ in the cc-pVDZ basis set with spherical d orbitals. The results, shown in Figs. 3 and 4 and Table II, are based on a 6-orbital active space consisting of the 2p orbitals. All orbitals were correlated.

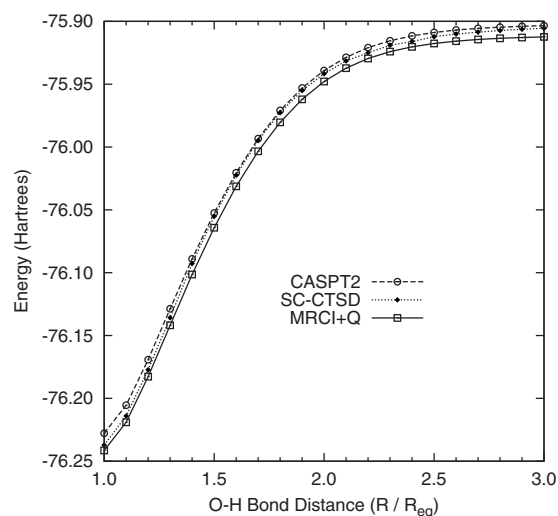


FIG. 2. Ground state potential energy curves for H₂O in the cc-pVDZ basis set. R_{eq}=0.9929 Å. Bond angle=109.57°. SC-CTSD used an energy threshold of τ_ε=0.1 E_h.

The standard CTSD method with exact overlap was again the closest to MRCI+Q, with a relative error of 1.4 mE_h. Also as in H₂O, CTSD is less accurate when S_{int} is approximated via the cumulant decomposition, showing a relative error of 7.2 mE_h. In the SC-CTSD method, the 2s* → 3s* excitation operator caused an intruder state at a bond distance of 1.3 Å that was not detected using the approximate Jacobian eigenvalues. After manually disabling this operator at all geometries, SC-CTSD produces a potential energy curve with a relative error of 3.9 mE_h. This result shows that SC-CTSD can be more accurate than CTSD when the exact three-body RDM is not available. WC-CTSD, on the other hand, encounters multiple intruder states that are not easily removed and consequently has an accuracy inferior to SC-CTSD, with a relative error of 8.0 mE_h. All of the

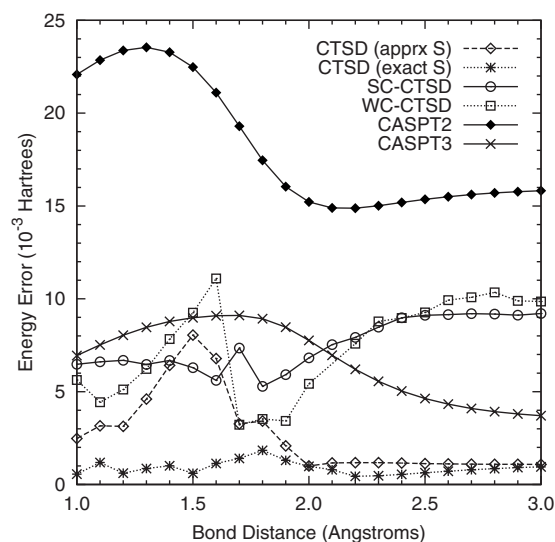


FIG. 3. Ground state energy errors relative to MRCI+Q for N₂ in the cc-pVDZ basis set. CTSD employed truncation thresholds of (τ₁=0.3, τ₂=0.1) and (τ₁=0.1, τ₂=0.01) when using the approximate and exact semi-internal overlap matrices, respectively. Both SC-CTSD and WC-CTSD used an energy threshold of τ_ε=0.1 E_h.

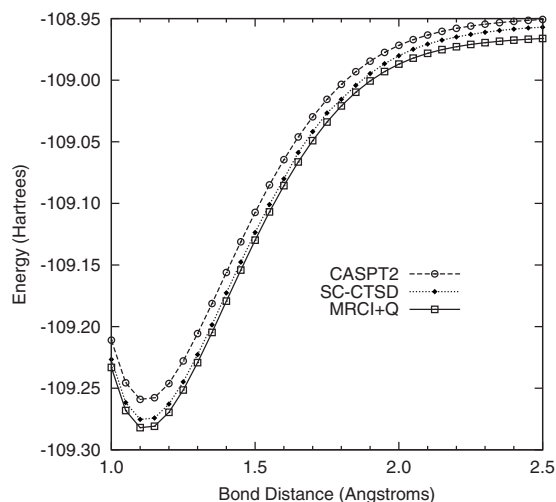


FIG. 4. Ground state potential energy curves for N_2 in the cc-pVDZ basis set. SC-CTSD used an energy threshold of $\tau_\epsilon=0.1 E_h$.

CT methods are more accurate than CASPT2, whose relative error is $8.7 mE_h$. As for H_2O , the discontinuity problem was more severe for SC-CTSD than for CTSD with exact overlap, but the discontinuities were small enough that the shape of the potential energy curve was unaffected, as shown in Fig. 4.

If the 2s orbitals are added to the active space for the CASPT2 calculation, the relative error improves to $5.3 mE_h$, which is still larger than the error of SC-CTSD using six active orbitals. Curiously, the relative error of SC-CTSD gets worse ($7.9 mE_h$) when the 2s orbitals are included in the active space. This unusual behavior can be explained by the structure of the strongly contracted excitations. When the 2s orbitals are in the core, they have individual excitations into each virtual orbital. When they are included in the active space, however, their excitations to the virtuals are strongly contracted with those of the 2p orbitals, reducing the degrees of freedom available for treating excitations from the 2s to 3s and 3p orbitals.

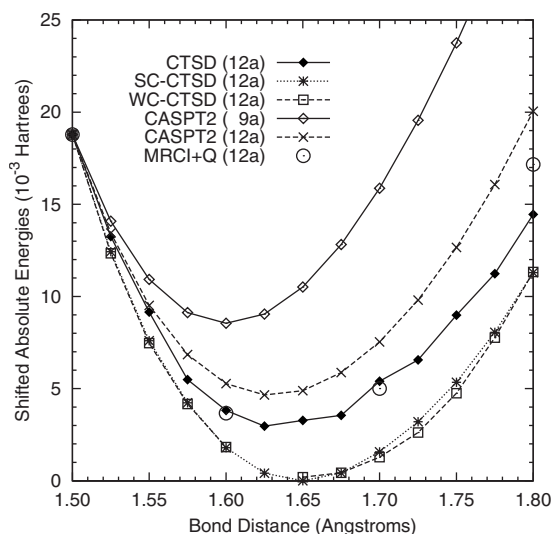


FIG. 5. Ground state potential energy curves for NiO in an ANO basis with (4s,3p,2d) and (6s,5p,4d,2f) contractions for O and Ni, respectively. The curves have been shifted so that the energies are equal at 1.5 Å. Methods labeled (9a) and (12a) use 9 and 12 active orbitals, respectively. CTSD used exact overlap matrices with $\tau_1=0.2$ and $\tau_2=0.1$, while both SC-CTSD and WC-CTSD used an energy threshold of $\tau_\epsilon=0.1 E_h$.

C. Nickel oxide

The potential energy curve of NiO was computed near the equilibrium geometry using various methods. All calculations employed an ANO basis set^{54,55} using spherical d and f orbitals, with (4s,4p,2d) and (6s,5p,4d,2f) contractions for O and Ni, respectively. All orbitals were optimized during CASSCF calculations. For dynamic correlation calculations, the Ni 1s, 2s, and 2p and the O 1s orbitals were not correlated. An active space of 12 orbitals (Ni 4s,3d,4p and O 2p) was employed unless stated otherwise. It is critical to include the Ni 4p orbitals in the active space, as demonstrated by the improvement in the CASPT2 energy shown in Fig. 5 and Table III. In this analysis we use MRCI+Q as our benchmark method, although we note that in NiO there are a suf-

TABLE II. Results for N_2 in the cc-pVDZ basis. MRCI+Q is reported in E_h , with other methods reported as the difference from MRCI+Q in mE_h . All methods use a (6e,6o) active space. All orbitals were correlated.

R^a	MRCI+Q	CASSCF	CASPT2	CASPT3	CTSD ^b	CTSD ^c	SC-CTSD ^d	WC-CTSD ^d
1.0	-109.233 082	188.680	22.082	6.943	2.474	0.558	6.476	5.631
1.2	-109.269 517	194.453	23.372	8.037	3.127	0.609	6.691	5.116
1.4	-109.179 349	199.244	23.278	8.780	6.414	1.004	6.679	7.834
1.6	-109.085 632	202.348	21.104	9.092	6.779	1.133	5.610	11.096
1.8	-109.020 859	201.686	17.460	8.932	3.419	1.839	5.280	3.524
2.0	-108.986 923	197.262	15.220	7.762	0.997	0.979	6.821	5.423
2.2	-108.972 848	192.612	14.880	6.203	1.175	0.440	7.919	7.581
2.4	-108.967 470	189.554	15.186	5.033	1.156	0.543	8.982	8.963
2.6	-108.965 211	187.813	15.500	4.325	1.109	0.716	9.151	9.926
2.8	-108.964 093	186.826	15.704	3.925	1.090	0.855	9.178	10.349
3.0	-108.963 464	186.249	15.822	3.700	1.083	0.951	9.206	9.855
Relative error		16.367	8.688	5.411	7.181	1.400	3.927	7.977

^aRadius in angstroms.

^b $\tau_1=0.3$, $\tau_2=0.1$, approximate S_{int} .

^c $\tau_1=0.1$, $\tau_2=0.01$, exact S_{int} .

^d $\tau_\epsilon=0.1 E_h$.

TABLE III. Results for the NiO potential energy curve. MRCI+Q is reported in E_h , with other methods reported as the difference from MRCI+Q in mE_h . Unless stated otherwise, a (14e,12o) active space is used.

R^a	MRCI+Q	CASSCF	CASPT2 ^b	CASPT2	CTSD ^c	CTSD ^d	CTSD ^e	SC-CTSD ^f	WC-CTSD ^f	$ c_{\max} ^2$ ^g
1.5	-1582.540 991	751.290	-89.379	-74.859	-6.571	-3.105	-1.722	12.703	0.991	0.829
1.6	-1582.556 108	746.607	-84.500	-73.253	-6.418	-2.907	-0.653	10.831	-0.842	0.806
1.7	-1582.554 764	742.195	-78.515	-72.334	-6.177	-2.792	-0.830	9.264	-2.737	0.776
1.8	-1582.542 605	736.472	-73.218	-71.976	-9.286	-5.195	-1.629	6.830	-4.842	0.728
Relative error		14.818	16.162	2.882	3.109	2.403	1.069	5.874	5.834	N/A

^aRadius in angstroms.

^bThe Ni 4p orbitals were excluded to create a (14e,9o) active space.

^c $\tau_1=0.2$, $\tau_2=0.1$, exact S_{int} .

^d $\tau_1=0.3$, $\tau_2=0.1$, exact S_{int} .

^e $\tau_1=0.4$, $\tau_2=0.1$, exact S_{int} .

^f $\tau_\epsilon=0.1 E_h$.

^gWeight of the leading determinant in the CASSCF CI expansion using natural orbitals.

ficient number of electrons to cast doubt on the accuracy of MRCI+Q's size extensivity correction, which is only exact for noninteracting electron pairs.

The CTSD method with exact overlap produces a curve whose relative error with respect to MRCI+Q is smaller than that of SC-CTSD (5.9 mE_h) or WC-CTSD (5.8 mE_h) and comparable to that of CASPT2 (2.9 mE_h). However, unlike for H_2O and N_2 , there is no obviously correct value for τ_1 in NiO, making a precise definition of the CTSD energy difficult. As seen in Table III, varying τ_1 between 0.2 and 0.4 changes CTSD's relative error from 3.1 mE_h to 1.1 mE_h . Additionally, Fig. 6 shows that the CTSD potential energy curve has unphysical characteristics for any of these values for τ_1 . These results indicate that some important excitation operators' overlap eigenvalues are crossing the truncation threshold as the bond is stretched. These crossings create discontinuous changes in the energy that produce features such as the one seen at 1.7 Å in Figs. 5 and 6.

For SC-CTSD, the smallest approximate semi-internal Jacobian eigenvalue for any geometry was 1.1 E_h [see Eq. (17)]. We typically use $\tau_\epsilon=0.1 E_h$ as our cutoff threshold to

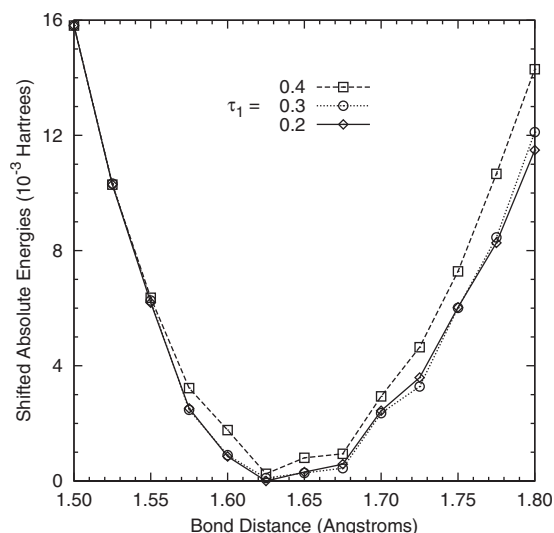


FIG. 6. The effect of τ_1 on the CTSD potential energy curve for NiO. The curves have been shifted so that the energies are equal at 1.5 Å. All curves use a 12 orbital active space, $\tau_2=0.1$, and exact overlap matrices.

prevent intruder states, so in NiO none of the strongly contracted excitation operators showed any intruder state character. Furthermore, the smallest eigenvalue was large enough that the SC-CTSD result was insensitive to our choice of τ_ϵ , in contrast to CTSD's strong dependence on τ_1 . An important result of this insensitivity is that the SC-CTSD potential energy curve is completely smooth for NiO, in contrast to the results for H_2O and N_2 . The situation for WC-CTSD is not as fortunate, with an intruder state appearing at 1.625 Å for $\tau_\epsilon=0.1 E_h$. This intruder state can be removed by raising τ_ϵ to 0.5 E_h , but this reduces WC-CTSD's accuracy at bond distances above 1.7 Å.

In terms of overall relative accuracy when compared to MRCI+Q, the tested methods rank as follows: CASPT2 \approx CTSD > SC-CTSD > WC-CTSD. We note that while CASPT2 is more accurate than SC-CTSD when the Ni 4p orbitals are included in the active space, its accuracy suffers severely if the active space is restricted to the Ni 3d and 4s and the O 2p orbitals. This limitation will be important in compounds with two transition metals, where SC-CTSD can afford to keep the 4p orbitals in the active space while CASPT2 cannot.

D. Free base porphin

We have evaluated the singlet-triplet gap of free base porphin ($C_{20}H_{14}N_4$) in two basis sets: 6-31G⁵⁶ and an ANO basis^{54,55} with spherical d orbitals and contractions of (3s,2p,1d) for C and N and (2s) for H. The active space was taken as the 24 out-of-plane 2p orbitals of C and N. For the DMRG-SCF calculations, a Pipek-Mezey⁵⁷ localization was applied to the out-of-plane 2p orbitals obtained from a Hartree-Fock calculation in PSI3,⁵⁸ which were then arranged on the orbital lattice as shown in Fig. 7. The orbitals were then optimized using 1200 DMRG states, after which the final energies and one- and two-body RDMs were evaluated using 2400 states. In the SC-CTSD calculations, the C and N 1s orbitals were not correlated, and the strongly contracted excitation operators were defined using the Hamiltonian in the DMRG-SCF natural orbital basis. A threshold of $\tau_\epsilon=0.1 E_h$ was employed, although it proved unnecessary as none of the strongly contracted excitation operators displayed intruder state character (their approximate Jacobian

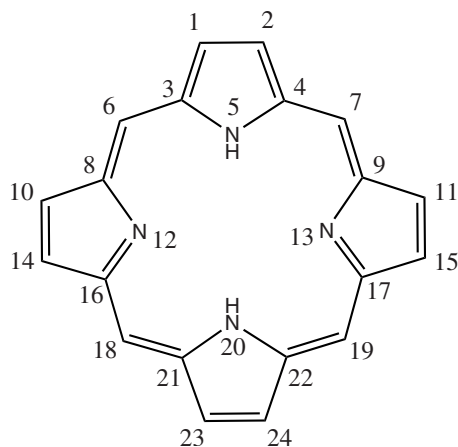


FIG. 7. Ordering of the out-of-plane 2p active orbitals on the DMRG orbital lattice for free base porphyrin.

eigenvalues were all larger than the threshold). Thus, as for NiO, we expect the SC-CTSD potential energy surface to be smooth near the equilibrium geometry. In both the SC-CTSD and CASPT2 results discussed, the lowest lying triplet state was of B_{2u} symmetry.

The DMRG and SC-CTSD calculations were performed using the geometry optimized by Haeser *et al.*⁵⁹ through density functional theory. These calculations, like the CASPT2 results of Roos *et al.*,⁶⁰ correspond to a vertical excitation in which the triplet state's geometry is not allowed to relax. However, the measurement of the experimental gap was performed by observing phosphorescence emission,⁶¹ which, due to the millisecond time scale separating excitation and emission, measures the nonvertical gap (the gap after the triplet geometry has relaxed). Therefore, to compare to experiment, it would have been more appropriate to calculate the nonvertical singlet-triplet gap. To approximately correct for this disparity, we optimized the geometry of both the singlet and triplet states with spin-unrestricted B3LYP density functional theory⁶² in the 6-31G* basis set⁶³ using the GAUSSIAN03 program package.⁶⁴ The change in the B3LYP singlet-triplet gap due to geometry relaxations was then combined with the vertical gaps of the other methods to produce approximate nonvertical gaps, which can be more appropriately compared with experiment. The results of these calculations are shown in Table IV, while the geometries involved can be found in the supplemental information.⁵⁰

The approximate nonvertical SC-CTSD singlet-triplet gaps were 1.65 and 1.73 eV for the 6-31G and ANO basis sets, respectively. These gaps are both within 0.15 eV of the 1.58 eV experimental value. After accounting for geometry relaxation, the CASPT2 gap is 1.30 eV, which has an error of 0.28 eV when compared to experiment. Finally, B3LYP density functional theory produced a 1.53 eV nonvertical gap, in error by only 0.05 eV. Note that none of the theoretical methods account for solvent effects, which should be kept in mind when comparing to the experiment.

An important difference between the SC-CTSD and CASPT2 results is found in the role of intruder states. Limited to a 14-orbital active space, the CASPT2 calculation required a 0.4 Hartree level shift in order to avoid intruder

TABLE IV. Singlet-triplet gaps for free base porphyrin. See Sec. III D.

Method	DMRG states	Basis set	Gap (eV)
Vertical gaps			
DMRG-SCF	1200	6-31G	1.56
SC-CTSD	1200	6-31G	1.85
DMRG-SCF	2400	6-31G	1.57
SC-CTSD	2400	6-31G	1.87
DMRG-SCF	1200	ANO	1.63
DMRG-SCF	2400	ANO	1.65
SC-CTSD	2400	ANO	1.95
CASPT2 ^a	...	ANO	1.52
B3LYP	...	6-31G*	1.74
Nonvertical gaps			
SC-CTSD ^b	2400	6-31G	1.65
SC-CTSD ^b	2400	ANO	1.73
CASPT2 ^{a,b}	...	ANO	1.30
B3LYP	...	6-31G*	1.53
Experiment ^c	1.58

^aReference 60.

^bApproximated using the B3LYP geometry relaxation.

^cReference 61.

states, which the authors cautioned could produce up to 0.2 eV of error in the excitation energy.⁶⁰ In contrast, the strongly contracted operators used in CT theory showed no intruder state characteristics, allowing an unambiguous singlet-triplet gap to be obtained.

TABLE V. Orbital occupations for the active space natural orbitals of free base porphyrin when using 2400 DMRG states and the ANO basis set. See Sec. III D.

Orbital	Singlet occupation	Triplet occupation
1	1.9859	1.9877
2	1.9854	1.9876
3	1.9636	1.9567
4	1.9612	1.9548
5	1.9488	1.9427
6	1.9456	1.9393
7	1.9405	1.9369
8	1.9301	1.9238
9	1.9167	1.9069
10	1.9149	1.8933
11	1.9147	1.8754
12	1.8197	1.7229
13	1.7955	1.1608
14	0.2286	0.8502
15	0.2069	0.3104
16	0.1075	0.1436
17	0.0779	0.1013
18	0.0649	0.0724
19	0.0620	0.0719
20	0.0591	0.0625
21	0.0444	0.0504
22	0.0429	0.0498
23	0.0424	0.0497
24	0.0407	0.0490

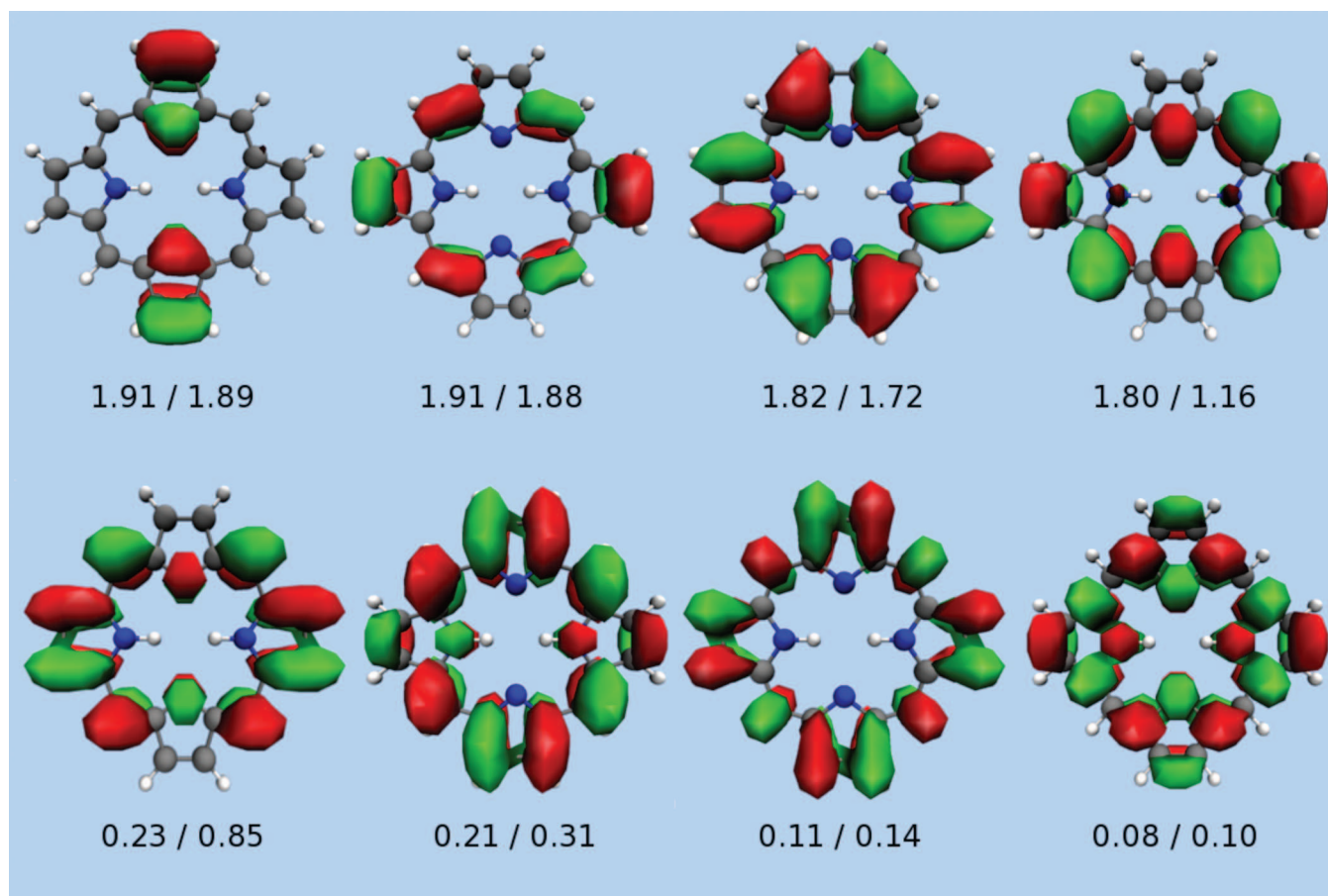


FIG. 8. Isosurface plots of the eight natural orbitals of free base porphin surrounding the dominant singlet configuration's highest occupied and lowest unoccupied molecular orbitals. The orbitals plotted are the singlet natural orbitals for the 2400 state DMRG solution in the ANO basis. The triplet orbitals have the same qualitative shapes. The numbers represent the singlet/triplet natural orbital occupations.

We have included in Table V the singlet and triplet natural orbital occupations for the 2400 state DMRG solutions in the ANO basis set. All orbitals of the singlet and triplet states had occupations differing from single reference behavior by more than 0.01, while 16 of the singlet and 17 of the triplet orbitals had occupations differing by more than 0.05. Isosurface plots of eight of the singlet state's orbitals are shown in Fig. 8. As a final note, we observe that the contribution of correlations between the active (out-of-plane 2p) and external (everything else) orbitals to the SC-CTSD singlet-triplet gap was not sensitive to either the number of DMRG states retained or the presence of polarization functions in the basis set. This can be seen by recognizing that the difference between the DMRG-SCF and SC-CTSD vertical energy gaps (0.30 eV) changes by less than 0.01 eV across the three SC-CTSD calculations that were performed. Future work should investigate the performance of both WC-CTSD and (when the DMRG three-body RDM becomes available) CTSD in order to assess the effects of excitation operator contraction.

IV. CONCLUSIONS

CT theory is a new and promising method for modeling dynamic correlation in multireference systems with large active spaces. It is rigorously size extensive, produces a two-body effective Hamiltonian, and has a lower cost scaling

than either CASPT2 or MRCI+Q. Like many other multireference dynamic correlation methods, however, the theory must overcome the challenge of intruder states, which in CT theory are caused by the operator and cumulant decomposition approximations. Previously, these intruder states were circumvented using a truncated, orthonormal excitation operator basis. This approach requires diagonalizing the first order interacting basis' overlap matrix, which has an n_{act}^9 cost and requires the reference wave function's three-body RDM in order to be accurate. As a result, the standard CTSD method becomes intractable in systems with sufficiently large active spaces. In this work, we have proposed using strongly and weakly contracted excitation operators as alternatives for addressing the intruder state problem and have introduced the SC-CTSD and WC-CTSD methods. Strongly contracted operators are orthogonal by construction, and their approximate Jacobian eigenvalues (which are used to detect and avoid intruder states) can be evaluated without changing the overall n^6 cost scaling of CT theory.

Tests on H₂O, N₂, and NiO show that SC-CTSD is at least as accurate as CASPT2. Furthermore, when the exact three-body RDM of the reference wave function is unavailable, SC-CTSD is more accurate than even overlap truncation based CTSD, making it the most accurate method available for treating dynamic correlation with a DMRG reference wave function. In the large active spaces of NiO (12 orbitals)

and free base porphyrin (24 orbitals), SC-CTSD managed to avoid intruder states without the use of a truncation threshold, making its potential energy surfaces completely smooth. This is especially significant in porphyrin, for which CASPT2 calculations in smaller active spaces require level shifts as large as 0.4 Hartrees. WC-CTSD was less successful at removing intruder states, and as a result was less reliable and less accurate in most systems tested. In summary, we recommend that CTSD with overlap truncation be attempted in systems for which the exact three-body RDM can be diagonalized, while SC-CTSD should be used when the three-body RDM is unavailable or when overlap truncation has difficulty removing intruder states. For non-experts, SC-CTSD is recommended for all systems because it is less sensitive to the choice of truncation threshold than the overlap truncation method.

We have demonstrated that SC-CTSD is both accurate and reliable, and that in combination with the DMRG-SCF method it can be applied to systems with complete active spaces of unprecedented size. Future work should focus on further applications of these methods in order to better understand their strengths and weaknesses.

ACKNOWLEDGMENTS

This work was supported by the Department of Energy Office of Science Award No. DE-FG02-07ER46432. Eric Neuscamman would like to acknowledge the support of the National Science Foundation Graduate Research Fellowship Program.

APPENDIX: SC OPERATOR DEFINITIONS

Here we present precise definitions for the eight types of strongly contracted excitation operators. See Sec. II C for a description of how we arrive at these definitions,

$$\hat{\delta}_{c_1v_2}^{v_1v_2} = \sum_{a_1a_2} g_{a_1a_2}^{v_1v_2} (E_{a_1a_2}^{v_1v_2} - E_{v_1v_2}^{a_1a_2}), \quad (\text{A1})$$

$$\hat{\delta}_{c_1c_2}^{a_1a_2} = \sum_{a_1a_2} g_{c_1c_2}^{a_1a_2} (E_{c_1c_2}^{a_1a_2} - E_{a_1a_2}^{c_1c_2}), \quad (\text{A2})$$

$$\hat{\delta}_{c_1c_2}^{v_1} = \sum_{a_1} g_{c_1c_2}^{v_1a_1} (E_{c_1c_2}^{v_1a_1} - E_{v_1a_1}^{c_1c_2}), \quad (\text{A3})$$

$$\hat{\delta}_{c_1}^{v_1v_2} = \sum_{a_1} g_{a_1c_1}^{v_1v_2} (E_{a_1c_1}^{v_1v_2} - E_{v_1v_2}^{a_1c_1}), \quad (\text{A4})$$

$$\hat{\delta}_{c_1c_2}^{v_1v_2} = g_{c_1c_2}^{v_1v_2} (E_{c_1c_2}^{v_1v_2} - E_{v_1v_2}^{c_1c_2}), \quad (\text{A5})$$

$$\begin{aligned} \hat{\delta}_{c_1}^{v_1} = & \left(t_{c_1}^{v_1} + \sum_{c_2} (2g_{c_1c_2}^{v_1c_2} - g_{c_1c_2}^{c_2v_1}) \right) (E_{c_1}^{v_1} - E_{v_1}^{c_1}) \\ & + \sum_{a_1a_2} g_{a_1a_2}^{v_1a_2} (E_{c_1a_1}^{v_1a_2} - E_{v_1a_2}^{c_1a_1}) + \sum_{a_1a_2} g_{c_1a_1}^{a_2v_1} (E_{c_1a_1}^{a_2v_1} - E_{a_2v_1}^{c_1a_1}), \end{aligned} \quad (\text{A6})$$

$$\begin{aligned} \hat{\delta}^{v_1} = & \sum_{a_1} \left(t_{a_1}^{v_1} + \sum_{c_2} (2g_{a_1c_2}^{v_1c_2} - g_{a_1c_2}^{c_2v_1}) \right) (E_{a_1}^{v_1} - E_{v_1}^{a_1}) \\ & + \sum_{a_1a_2a_3} g_{a_1a_2}^{v_1a_3} (E_{a_1a_2}^{v_1a_3} - E_{v_1a_3}^{a_1a_2}), \end{aligned} \quad (\text{A7})$$

$$\begin{aligned} \hat{\delta}_{c_1}^{a_1} = & \sum_{a_1} \left(t_{c_1}^{a_1} + \sum_{c_2} (2g_{c_1c_2}^{a_1c_2} - g_{c_1c_2}^{c_2a_1}) \right) (E_{c_1}^{a_1} - E_{a_1}^{c_1}) \\ & + \sum_{a_1a_2a_3} g_{c_1a_3}^{a_1a_2} (E_{c_1a_3}^{a_1a_2} - E_{a_1a_2}^{c_1a_3}). \end{aligned} \quad (\text{A8})$$

- ¹K. Ruedenberg, M. W. Schmidt, M. M. Gilbert, and S. T. Elbert, *Chem. Phys.* **71**, 41 (1982).
- ²B. O. Roos, *Adv. Chem. Phys.* **69**, 399 (1987).
- ³K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, *J. Phys. Chem.* **94**, 5483 (1990).
- ⁴K. Andersson, P.-Å. Malmqvist, and B. O. Roos, *J. Chem. Phys.* **96**, 1218 (1992).
- ⁵P. Celani and H.-J. Werner, *J. Chem. Phys.* **112**, 5546 (2000).
- ⁶K. Hirao, *Chem. Phys. Lett.* **190**, 374 (1992).
- ⁷C. Angeli, R. Cimraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu, *J. Chem. Phys.* **114**, 10252 (2001).
- ⁸C. Angeli, R. Cimraglia, and J.-P. Malrieu, *J. Chem. Phys.* **117**, 9138 (2002).
- ⁹C. Angeli, M. Pastore, and R. Cimraglia, *Theor. Chem. Acc.* **117**, 743 (2007).
- ¹⁰H.-J. Werner and E. A. Reinsch, *J. Chem. Phys.* **76**, 3144 (1982).
- ¹¹H.-J. Werner and P. J. Knowles, *J. Chem. Phys.* **89**, 5803 (1988).
- ¹²P. J. Knowles and H.-J. Werner, *Chem. Phys. Lett.* **145**, 514 (1988).
- ¹³S. R. Langhoff and E. R. Davidson, *Int. J. Quantum Chem.* **8**, 61 (1974).
- ¹⁴E. R. Davidson and D. W. Silver, *Chem. Phys. Lett.* **52**, 403 (1977).
- ¹⁵S. R. White and R. L. Martin, *J. Chem. Phys.* **110**, 4127 (1999).
- ¹⁶G. K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.* **116**, 4462 (2002).
- ¹⁷J. Hachmann, W. Cardoen, and G. K.-L. Chan, *J. Chem. Phys.* **125**, 144101 (2006).
- ¹⁸D. Zgid and M. Nooijen, *J. Chem. Phys.* **128**, 144116 (2008).
- ¹⁹D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan, *J. Chem. Phys.* **128**, 144117 (2008).
- ²⁰K. H. Marti, I. M. Ondík, G. Moritz, and M. Reiher, *J. Chem. Phys.* **128**, 014104 (2008).
- ²¹J. J. Dorando, J. Hachmann, and G. K.-L. Chan, *J. Chem. Phys.* **130**, 184111 (2009).
- ²²J. Olsen, B. O. Roos, P. Jørgensen, and H. J. Aa. Jensen, *J. Chem. Phys.* **89**, 2185 (1988).
- ²³P. A. Malmqvist, A. Rendell, and B. O. Roos, *J. Phys. Chem.* **94**, 5477 (1990).
- ²⁴W. A. Goddard III and R. C. Ladner, *J. Am. Chem. Soc.* **93**, 6750 (1971).
- ²⁵W. J. Hunt, P. J. Hay, and W. A. Goddard III, *J. Chem. Phys.* **57**, 738 (1972).
- ²⁶P.-Å. Malmqvist, K. Pierloot, A. R. M. Shahi, C. J. Cramer, and L. Gagliardi, *J. Chem. Phys.* **128**, 204109 (2008).
- ²⁷D. Zgid, D. Ghosh, E. Neuscamman, and G. K.-L. Chan, *J. Chem. Phys.* **130**, 194107 (2009).
- ²⁸T. Yanai and G. K.-L. Chan, *J. Chem. Phys.* **124**, 194106 (2006).
- ²⁹T. Yanai and G. K.-L. Chan, *J. Chem. Phys.* **127**, 104107 (2007).
- ³⁰G. K.-L. Chan and T. Yanai, *Adv. Chem. Phys.* **134**, 343 (2007).
- ³¹E. Neuscamman, T. Yanai, and G. K.-L. Chan, *J. Chem. Phys.* **130**, 124102 (2009).
- ³²W. Kutzelnigg, *J. Chem. Phys.* **77**, 3081 (1982).
- ³³W. Kutzelnigg, *J. Chem. Phys.* **80**, 822 (1984).
- ³⁴J. D. Watts, G. W. Trucks, and R. J. Bartlett, *Chem. Phys. Lett.* **157**, 359 (1989).
- ³⁵R. J. Bartlett, S. A. Kucharski, and J. Noga, *Chem. Phys. Lett.* **155**, 133 (1989).
- ³⁶A. G. Taube and R. J. Bartlett, *Int. J. Quantum Chem.* **106**, 3393 (2006).
- ³⁷S. Pal, M. D. Prasad, and D. Mukherjee, *Theor. Chim. Acta* **62**, 523 (1983).
- ³⁸S. Pal, *Theor. Chim. Acta* **66**, 207 (1984).
- ³⁹J. Paldus and X. Li, *Adv. Chem. Phys.* **110**, 1 (1999).
- ⁴⁰M. R. Hoffmann and J. Simons, *J. Chem. Phys.* **88**, 993 (1988).
- ⁴¹D. Mukherjee, in *Recent Progress in Many-Body Theories*, edited by E.

- Schachinger (Plenum, New York, 1995), Vol. 4.
- ⁴²U. S. Mahapatra, B. Datta, B. Bandyopadhyay, and D. Mukherjee, *Advances in Quantum Chemistry* (Academic, New York, 1995), Vol. 30.
- ⁴³W. Kutzelnigg and D. Mukherjee, *J. Chem. Phys.* **107**, 432 (1997).
- ⁴⁴D. Mukherjee, *Chem. Phys. Lett.* **274**, 561 (1997).
- ⁴⁵F. Colmenero and C. Valdemoro, *Phys. Rev. A* **47**, 979 (1993).
- ⁴⁶W. Kutzelnigg and D. Mukherjee, *J. Chem. Phys.* **110**, 2800 (1999).
- ⁴⁷H. Nakatsuji and K. Yasuda, *Phys. Rev. Lett.* **76**, 1039 (1996).
- ⁴⁸N. Forsberg and P.-Å. Malmqvist, *Chem. Phys. Lett.* **274**, 196 (1997).
- ⁴⁹T. Yanai, Y. Kurashige, E. Neuscamman, and G. K.-L. Chan, *J. Chem. Phys.* **132**, 024105 (2010).
- ⁵⁰See supplementary material at <http://dx.doi.org/10.1063/1.3274822> for tensor contractions, the porphin geometry, and the operator decomposition definition.
- ⁵¹K. G. Dyall, *J. Chem. Phys.* **102**, 4909 (1995).
- ⁵²H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, and M. Schütz, MOLPRO, version 2006, a package of *ab initio* programs, see <http://www.molpro.net>.
- ⁵³T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- ⁵⁴P. Widmark, P. Malmqvist, and B. Roos, *Theor. Chim. Acta* **77**, 291 (1990).
- ⁵⁵R. Pou-Amérgigo, M. Merchán, I. Nebot-Gil, P. Widmark, and B. Roos, *Theor. Chim. Acta* **92**, 149 (1995).
- ⁵⁶W. J. Hehre, R. Ditchfield, and J. A. Pople, *J. Chem. Phys.* **56**, 2257 (1972).
- ⁵⁷J. Pipek and P. G. Mezey, *J. Chem. Phys.* **90**, 4916 (1989).
- ⁵⁸T. D. Crawford, C. D. Sherill, E. F. Valeev, J. T. Fermann, R. A. King, M. L. Leininger, S. T. Brown, C. L. Janssen, E. T. Seidl, J. P. Kenny, and W. D. Allen, *J. Comput. Chem.* **28**, 1610 (2007).
- ⁵⁹J. Almlöf, T. H. Fischer, P. G. Gassman, A. Ghosh, and M. Haeser, *J. Phys. Chem.* **97**, 10964 (1993).
- ⁶⁰L. Serrano-Andrés, M. Merchán, M. Rubio, and B. O. Roos, *Chem. Phys. Lett.* **295**, 195 (1998).
- ⁶¹M. Gouterman and G.-E. Khalil, *J. Mol. Spectrosc.* **53**, 88 (1974).
- ⁶²A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
- ⁶³P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta* **28**, 213 (1973).
- ⁶⁴M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.