

Multireference quantum chemistry through a joint density matrix renormalization group and canonical transformation theory

Takeshi Yanai, Yuki Kurashige, Eric Neuscamman, and Garnet Kin-Lic Chan*

Citation: *J. Chem. Phys.* **132**, 024105 (2010); doi: 10.1063/1.3275806

View online: <http://dx.doi.org/10.1063/1.3275806>

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

Published by the [American Institute of Physics](#)

Multireference quantum chemistry through a joint density matrix renormalization group and canonical transformation theory

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

¹*Institute for Molecular Science, 38 Nishigo-Naka, Myodaiji, Okazaki 444-8585, Japan*

²*Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301, USA*

(Received 3 October 2009; accepted 30 November 2009; published online 13 January 2010)

We describe the joint application of the density matrix renormalization group and canonical transformation theory to multireference quantum chemistry. The density matrix renormalization group provides the ability to describe static correlation in large active spaces, while the canonical transformation theory provides a high-order description of the dynamic correlation effects. We demonstrate the joint theory in two benchmark systems designed to test the dynamic and static correlation capabilities of the methods, namely, (i) total correlation energies in long polyenes and (ii) the isomerization curve of the $[\text{Cu}_2\text{O}_2]^{2+}$ core. The largest complete active spaces and atomic orbital basis sets treated by the joint DMRG-CT theory in these systems correspond to a $(24e, 24o)$ active space and 268 atomic orbitals in the polyenes and a $(28e, 32o)$ active space and 278 atomic orbitals in $[\text{Cu}_2\text{O}_2]^{2+}$. © 2010 American Institute of Physics. [doi:10.1063/1.3275806]

I. INTRODUCTION

Multireference quantum chemical systems, that is, molecules where the wave function is not qualitatively described by a single determinant, remain a major challenge for quantum chemistry. These problems arise in transition metal chemistry, excited states, and at stretched regions of the potential energy surface. Traditional quantum chemical methods for multireference problems start by describing the wave function qualitatively in an active space, a selected set of orbitals exhibiting variable occupancy which is correlated exactly at the full configuration interaction level. The electron correlation in the active space is termed *static* correlation. Quantitative corrections to the active space wave function arise from correlations between active space orbitals and the remaining external (core and virtual) orbitals. These effects are termed *dynamic correlation* and their description (on top of the static correlation) is usually obtained with a lower level theory, such as configuration interaction singles and doubles, coupled pair functionals,^{1,2} second-order perturbation theory,^{3,4} or density functional theory.^{5,6}

The standard approach outlined above, while successful for small molecules, suffers from a number of drawbacks for larger molecules both in the treatment of static correlation and in the treatment of dynamic correlation. In the case of static correlation, the obvious problem is the exponential dependence of the full configuration interaction theory on the size of the active space. For this reason, many groups have explored approximate nonexponential cost theories of static correlation, including restricted active space configuration interaction,^{7,8} perfect pairing and other valence-bond and geminal type theories,⁹⁻¹⁴ high-order active-space coupled cluster methods,^{15,16} and variational density matrix methods.^{17,18} In the treatment of the remaining dynamic correlation, while traditional low-order corrections based on

multireference configuration interaction and perturbation theory are formally polynomial in cost, they do not possess the same satisfactory combination of utility, accuracy, and formal correctness (e.g., size-extensivity) as the standard coupled cluster method in single reference problems. Consequently, much effort has been devoted to finding suitable extensions of coupled cluster theory to treat residual dynamic correlation in multireference problems.¹⁹

We have been implementing a program of research to tackle the aforementioned twin challenges of static and dynamic correlation in the multireference quantum chemistry of larger molecules. For the static correlation problem, we have been exploring the density matrix renormalization group (DMRG).²⁰⁻²² While the DMRG is not yet a black-box method, our investigations,²³⁻³⁴ as well as those of many others,^{22,35-45} indicate that the DMRG provides a promising approach to static correlation. In particular, very accurate DMRG descriptions of active space energies and energy differences—considerably beyond the requirements of chemical accuracy—can be obtained for active space sizes ranging at 30–40 active orbitals in compact molecules^{24,26,34} to more than 100 active orbitals in the optimal case of long chains.^{28,29,31} For the remaining dynamic correlation problem, we have been developing the canonical transformation (CT) theory.⁴⁶⁻⁴⁸ In common with coupled cluster methods, CT theory is based on an exponential ansatz, but it uses operator decompositions to avoid the high computational cost that is typical of multireference versions of coupled cluster theories. CT theory exhibits an accuracy on par with the best multireference configuration interaction approaches, but shares the same favorable sixth-power computational scaling as the single-reference coupled cluster theory.

In the current work we describe the joint application of the DMRG and CT theories to large multireference quantum chemical problems. While the combination of a DMRG treatment of static correlation and an exponential based treat-

^{a)}Electronic mail: gc238@cornell.edu.

ment of dynamic correlation can be found already in White's work on canonical diagonalization⁴⁹ (the predecessor of CT theory) our particular realization of this strategy for large multireference quantum chemical problems relies on several additional contributions. In the case of the DMRG, such relevant recent developments include orbital optimization and algorithms to obtain reduced density matrices as developed by Zgid and Nooijen in Refs. 44 and 45 and by our group in Ref. 31, as well as efficient parallel quantum chemical DMRG algorithms for larger active spaces.^{25,34} In the case of the CT theory, we employ a systematic operator decomposition derived from the generalized normal ordering of Mukherjee and Kutzelnigg,^{50–52} as well as the cumulant decomposition of density matrices,^{53–56} to formulate the CT equations. On the algorithmic front, the large CT calculations described here have used a parallelized implementation to handle large numbers of basis functions, and better techniques (including the use of strong contraction as introduced in Ref. 57 to reduce the variational degrees of freedom in the excitation basis) to converge the CT equations. Details of the parallelized algorithm and the strongly contracted formulation of the CT theory will be presented elsewhere,^{58,59} although the former builds off the parallelization strategy established for orbital optimization in Ref. 31, and the latter grows out of our experiences with cumulant approximated strongly contracted n -electron valence perturbation theory.^{57,60,61} We cannot here list completely also the numerous connections to other earlier works on internally contracted multireference configuration interaction,^{62–64} perturbation theory,^{3,65} coupled cluster theory,^{52,66–69} and irreducible or anti-Hermitian contracted Schrödinger equation density matrix methods.^{70–72}

This paper now proceeds as follows. In Sec. II we recall key elements of the DMRG and CT theory. Here, we focus in particular on the ideas and quantities that specify a joint DMRG-CT calculation. From there we proceed to applications to two systems to understand the capabilities of the theory. The first is a calculation of the total correlation energies of long polyenes for which the accurate reproduction is a test of the dynamic correlation component of the theory. The second is a calculation of the isomerization curve of the $[\text{Cu}_2\text{O}_2]^{2+}$ cluster that forms a simple model of the tyrosinase core. This has previously been identified as requiring a very large active space for a balanced multireference description^{42,73,74} and is a test of the static correlation component of the theory. We finish then with our summary and conclusions.

II. THEORY

A. Density matrix renormalization group

The DMRG is based on a matrix product state wave function. It may be viewed as working with a direct factorization of the wave function expressed in an occupation number (Fock-space) representation. A more complete introduction to the DMRG wave function can be found in Ref. 75; see also the general review article.⁷⁶

The form of the DMRG wave function is

$$|\Psi\rangle = \sum_{\{n\}} \mathbf{A}^{n_1} \mathbf{A}^{n_2} \cdots \mathbf{A}^{n_k} |n_1 n_2 \cdots n_k\rangle, \quad (1)$$

where $|n_1 n_2 \cdots n_k\rangle$ denotes a determinant in occupation number representation, $\{n\} = \{|\text{vac}\rangle, |\alpha\rangle, |\beta\rangle, |\alpha\beta\rangle\}$, and the various \mathbf{A}^n are matrices, i.e., the elements are A_{ij}^n with indices i, j ranging from 1 to a maximum of M . The matrix dimension M controls the accuracy of the DMRG wave function; the wave function becomes exact as M approaches the square root of the full Hilbert space dimension. \mathbf{A}^{n_1} and \mathbf{A}^{n_k} are row and column vectors, respectively, thus the product of \mathbf{A}^n in Eq. (1) produces a single scalar coefficient $\Psi^{n_1 n_2 \cdots n_k}$. This coefficient multiplies the determinant $|n_1 n_2 \cdots n_k\rangle$ in a complete determinantal expansion of the wave function. We see that the DMRG assumes a factorized form for the full-configuration interaction amplitudes, as indexed in the occupation number form.

Note that the DMRG wave function makes no reference to a Hartree–Fock (HF) state and supports nonvanishing coefficients for all determinants in the space $\{n_1 \cdots n_k\}$. This property together with the compactness of the ansatz is what allows the DMRG to capture a wide variety of multireference correlations in a balanced and efficient way. Furthermore, the product structure of the wave function allows the DMRG to be exactly size-consistent, at least in a local basis.

For reasons of numerical stability, it is convenient to choose a canonical form of the DMRG wave function, where all matrices but one satisfies an orthogonality property. In this form, the DMRG wave function is written as

$$|\Psi\rangle = \sum_{\{n\}} \mathbf{L}^{n_1} \cdots \mathbf{L}^{n_{p-1}} \mathbf{C}^{n_p} \mathbf{R}^{n_{p+1}} \cdots \mathbf{R}^{n_k} |n_1 \cdots n_k\rangle, \quad (2)$$

where \mathbf{L}^n and \mathbf{R}^n satisfy orthogonality conditions

$$\sum_n \mathbf{L}^{n\dagger} \mathbf{L}^n = \mathbf{1}, \quad (3)$$

$$\sum_n \mathbf{R}^n \mathbf{R}^{n\dagger} = \mathbf{1}. \quad (4)$$

Formally, the string of \mathbf{L}^n and \mathbf{R}^n , respectively, define sets of many-particle basis states $\{l\}$, $\{r\}$, and \mathbf{C}^{n_p} gives the coefficients of the wave function in the product basis $\{l\} \otimes \{n_p\} \otimes \{r\}$. Thus the determinant expansion (2) is equivalent to an expansion in the smaller orthonormal renormalized basis

$$|\Psi\rangle = \sum_{lnr} C_{lr}^{n_p} |ln_p r\rangle. \quad (5)$$

The energy of the DMRG wave function is obtained by variationally minimizing with respect to each of the components \mathbf{L}^n , \mathbf{R}^n , and \mathbf{C}^n in Eq. (2) one at a time. This is known as the DMRG sweep algorithm. The component-wise optimization resembles Hartree's self-consistent algorithm for HF theory where one orbital is updated at a time. Many other useful parallels with HF theory can be drawn, for example, Eq. (5) is a kind of Fock equation in the DMRG.³³ During the optimization care must be taken to avoid spurious local minima by applying a gradually decreasing perturbative "noise." This has recently been discussed in detail by Zgid and Nooijen.⁴⁵

1. Orbital ordering

The DMRG wave function is expressed as a product of matrices, each associated with a given orbital (1). Since matrices do not generally commute, a DMRG calculation requires a specified ordering of the orbitals. Formally we view this as mapping the orbitals $1 \cdots k$ to a fictitious one-dimensional (1D) lattice. The DMRG wave function supports a limited amount of correlation between orbitals that are widely separated on this 1D lattice, and it is important to choose an ordering which minimizes such separations. For long chains (such as the polyene molecules studied later) there is a natural ordering in a basis that is localized in real-space. For compact molecules (such as the Cu_2O_2 core studied later) it is sometimes better to use a basis that is more localized in the energy domain. In this case, it can be advantageous to place bonding and antibonding orbitals adjacent to each other on the lattice.

2. Orbital optimization and reduced density matrices

The shape of the active orbitals used in static correlation must often be adapted to the molecular environment. Orbital optimization of the active space is the basis of the fully optimized reaction space⁷⁷ and complete active space self-consistent field (CASSCF) (Ref. 78) models. Orbital optimization of the DMRG active space was introduced by Zgid and Nooijen⁴⁴ followed by Ghosh *et al.*,³¹ giving, respectively, the DMRG-SCF/DMRG-CASSCF models. In each case orbital optimization is based on calculating the DMRG active space one- and two-body density matrices. Note that the one- and two-body DMRG density matrices form the input into the subsequent CT theory treatment of dynamic correlation in a joint DMRG-CT calculation.

Unlike more sophisticated orbital optimization algorithms used in modern CASSCF implementations, current DMRG orbital optimization implementations decouple the optimization of the orbital parameters and the wave function coefficients in Eq. (2). Acceleration of this two-step procedure can be achieved through direct inversion in the iterative subspace techniques.^{44,79}

B. Canonical transformation theory

The CT theory was developed in Refs. 46–48 as a way to incorporate dynamical correlation on top of a general multireference wave function. It uses a canonical (i.e., unitary) exponential ansatz together with an operator decomposition based on Mukherjee–Kutzelnigg normal ordering^{50–52} and density matrix cumulant decompositions^{55,56} to achieve a high-order, size-consistent treatment of dynamic correlation in a computationally efficient way. Starting from a reference $|\Psi_0\rangle$, the CT ansatz for $|\Psi\rangle$ is

$$|\Psi\rangle = \exp \hat{A} |\Psi_0\rangle, \quad (6)$$

where \hat{A} is an anti-Hermitian operator, and $\exp \hat{A}$ is a unitary transformation. The orbital space is partitioned into core (a, b, \dots), active (p, q, \dots), and virtual spaces (v, w, \dots) such that the reference $|\Psi_0\rangle$ consists of doubly occupied orbitals in the core space, arbitrary occupancy in the active space,

and zero occupancy in the virtual space. \hat{A} is a linear combination of operators \hat{O}_μ which excite into or out of the active space, with the general anti-Hermitian form

$$\hat{A} = \sum_{\mu} A_{\mu} (\hat{O}_{\mu} - \hat{O}_{\mu}^{\dagger}). \quad (7)$$

At the singles and doubles (CTSD) level of theory, the operators \hat{O} create one or two holes or particles in the core or virtual spaces, i.e.,

$$\hat{A} = \hat{A}_1 + \hat{A}_2, \quad (8)$$

$$\hat{A}_1 = \sum_{ap} A_a^p (a_a^p - a_p^a) + \sum_{av} A_a^v (a_a^v - a_v^a) + \sum_{pv} A_p^v (a_p^v - a_v^p), \quad (9)$$

$$\begin{aligned} \hat{A}_2 = & \sum_{pqab} A_{pq}^{ab} (a_{pq}^{ab} - a_{ab}^{pq}) + \sum_{pqra} A_{pq}^{ra} (a_{pq}^{ra} - a_{ra}^{pq}) \\ & \text{(active-core),} \\ & + \sum_{pqvw} A_{pq}^{vw} (a_{pq}^{vw} - a_{vw}^{pq}) + \sum_{pqrv} A_{pq}^{rv} (a_{pq}^{rv} - a_{rv}^{pq}) \\ & \text{(active-virtual),} \\ & + \sum_{pqva} A_{pa}^{qv} (a_{pa}^{qv} - a_{qv}^{pa}) + \sum_{pvwa} A_{pa}^{vw} (a_{pa}^{vw} - a_{vw}^{pa}) \\ & + \sum_{pvab} A_{ab}^{pv} (a_{ab}^{pv} - a_{pv}^{ab}) \quad \text{(active-core-virtual),} \\ & + \sum_{abvw} A_{ab}^{vw} (a_{ab}^{vw} - a_{vw}^{ab}) \quad \text{(core-virtual).} \end{aligned} \quad (10)$$

The CT ansatz in Eqs. (6) and (8)–(10) is the same as that used in an internally contracted unitary multireference coupled cluster theory. The distinction between the CT theory and such a coupled cluster theory lies in how the CT effective Hamiltonian, energy, and amplitudes are defined. The exact effective Hamiltonian corresponding to the ansatz (6) (as would be used in a coupled cluster theory) is formally

$$\bar{H} = \exp(-\hat{A}) \hat{H} \exp(\hat{A}) = \hat{H} + [\hat{H}, \hat{A}] + \frac{1}{2} [[\hat{H}, \hat{A}], \hat{A}] + \dots \quad (11)$$

Equation (11) is impractical due to the increasing complexity of the commutators, i.e., $[\hat{H}, \hat{A}]$ gives rise to three-body operators, $[[\hat{H}, \hat{A}], \hat{A}]$ to four-body operators, and so on. To eliminate this complexity CT theory defines a modified expansion of the effective Hamiltonian

$$\bar{H}_{1,2} = \hat{H} + [\hat{H}, \hat{A}]_{1,2} + \frac{1}{2} [[\hat{H}, \hat{A}]_{1,2}, \hat{A}]_{1,2} + \dots, \quad (12)$$

where 1,2 denotes that the commutator is replaced by its operator approximation in terms of one- and two-body operators only, using the generalized normal ordering of Mukherjee and Kutzelnigg.^{47,48,52} From the recursive application of this decomposition in Eq. (12) it is possible to obtain the *full* effective CT Hamiltonian $\bar{H}_{1,2}$ in terms of only one- and two-body operators.

Note that there is some freedom in when to apply the recursive operator decomposition. For example, the quadratic term in Eq. (12) could be approximated applying the decomposition once $\frac{1}{2}[[\hat{H}, \hat{A}], \hat{A}]_{1,2}$ rather than applying it twice $\frac{1}{2}[[\hat{H}, \hat{A}]_{1,2}, \hat{A}]_{1,2}$. The relative merits of such different approximations were explored in Ref. 48. In practice, the decomposition in Eq. (12), referred to in our earlier work^{47,48} as L-CT (“linear” CT, since the recursive decomposition is applied immediately after the linear commutator) provides good accuracy with the least computational effort, and we consider this to be the standard form of the CT approximation. Note that this use of the word “linear” is quite different from the linearization procedure sometimes used in coupled cluster theories! To avoid this conceptual confusion, and also to reflect that this is the standard definition of the CT theory, we will henceforth refer to the CT theory defined from Eq. (11) with the singles and doubles form of \hat{A} as simply CTSD.

The CT energy follows from the CT effective Hamiltonian as

$$E = \langle \Psi_0 | \bar{H}_{1,2} | \Psi_0 \rangle, \quad (13)$$

and the amplitudes in the CT theory are determined from, e.g.,

$$\langle \Psi_0 | [\bar{H}_{1,2}, a_{pq}^{vw}]_{1,2} | \Psi_0 \rangle = 0. \quad (14)$$

The CTSD amplitudes and energies may be obtained with $O(n_{\text{act}}^2 n_{\text{ext}}^4)$ cost where n_{act} denotes the size of the active space and n_{ext} the size of the external space (though an additional one-time $O(n_{\text{act}}^9)$ cost is required if the optional overlap truncation procedure is used, see later).

Note that the energy and amplitudes are defined using *only* the one- and two-body active-space density matrices of the reference wave function. There are thus connections to work on reduced density matrix theories. Indeed the CT amplitude equations correspond (after a simple transformation) to the Brillouin conditions of the irreducible contracted Schrödinger equation of Mukherjee and Kutzelnigg in Ref. 70. The combination of Brillouin conditions, unitary parametrization, and cumulant approximations found in CT theory has also recently been used by Mazziotti in the anti-Hermitian contracted Schrödinger equation,⁷¹ and consequently that theory has a structure that closely resembles CT theory.

Our CT calculations in this work use a new parallelized implementation of the CTSD equations. The details of the parallelization are presented elsewhere.⁵⁹

1. Convergence stabilization

When $|\Psi_0\rangle$ has multireference character, nontrivial convergence issues can arise when solving the CT amplitude equations. While these issues are of a technical nature, the elimination of such problems is necessary for practical CT calculations. The origins of the convergence issues are twofold.

- (1) Internally contracted dynamical correlation methods (including internally contracted multireference CI, perturbation theory, and coupled cluster theory) generally

suffer from near linear-dependencies in the first-order interacting basis $\{\hat{O}_\mu |\Psi_0\rangle\}$, since, e.g., $a_{pq}^{vw} |\Psi_0\rangle$ is not necessarily orthogonal to $a_{p'q'}^{vw} |\Psi_0\rangle$.^{3,64}

- (2) When solving the nonlinear CT amplitude equation $\langle \Psi_0 | [\bar{H}_{1,2}, (\hat{O}_\mu - \hat{O}_\mu^\dagger)]_{1,2} | \Psi_0 \rangle$ [where \hat{O}_μ is the generic excitation operator in Eq. (7)], the exact Jacobian to use in a Newton–Raphson algorithm is defined by

$$J_{\mu\nu} = \frac{\partial}{\partial A_\nu} \langle \Psi_0 | [\bar{H}_{1,2}, (\hat{O}_\mu - \hat{O}_\mu^\dagger)]_{1,2} | \Psi_0 \rangle. \quad (15)$$

The Jacobian matrix can have spuriously small eigenvalues of a nonphysical nature (arising from the operator decompositions) leading to failure of the Newton–Raphson scheme.⁵⁸

In recent CT calculations (Refs. 47 and 48) a scheme based on overlap truncation, detailed in Ref. 47, has been used to stabilize the convergence and remove spuriously small eigenvalues of the Jacobian (15). This relies on there being some alignment between the (near) null spaces of the first-order interacting basis overlap matrix $S_{\mu\nu} = \langle \hat{O}_\mu^\dagger \hat{O}_\nu \rangle$ and the approximate Jacobian, i.e., that $S^{-1}J$ is well-conditioned. In this scheme, the overlap matrix of the first-order interacting basis is computed from the one-, two- and three-body density matrices of the reference wave function and the excitation operators \hat{O}_μ are transformed to generate orthogonal excitations in the first-order interacting space. All operators generating excitations that correspond to eigenvalues of the overlap matrix below a certain threshold (τ_1 for singles and semi-internal excitations, τ_2 for double excitations) are truncated.

Since computing the three-body density matrix of the reference wave function is impractical for the large active space problems targeted by this work, here we perform overlap truncation constructing the first-order interacting basis overlap matrix using the one-, two- and *cumulant approximated* three-body density matrices of the reference wave function. Compared with the exact overlap truncation scheme, higher truncation thresholds τ_1, τ_2 are necessary with this approximate scheme to ensure convergence of the CT equations.

2. Strongly contracted canonical transformation theory

The overlap truncation method described above formally involves diagonalizing the overlap matrix of the first-order interacting basis, which requires $O(n_{\text{act}}^9)$ cost. This is not a problem for smaller active spaces, but becomes a bottleneck in larger systems that we are targeting with the joint DMRG-CT theory.

An alternative approach to enforce orthogonality of the excitation basis is to use *strongly contracted* excitation operators as introduced in *n*-electron valence perturbation theory.^{57,60,61} This leads to a so-called strongly contracted CT theory. The details of this theory are presented in the accompanying paper (Ref. 58). Here we present only the essential ideas.

In strong contraction, rather than allowing many different kinds of excitations from the active space to a set of external orbitals, e.g., $a_{pq}^{vw}, a_{p'q'}^{vw}$, only a single linear combination of active space states is allowed to excite to a given set of external orbitals, e.g., for the external orbitals vw an excitation of the form $\hat{O}^{vw} = \sum_{pq} X_{pq}^{vw} a_{pq}^{vw}$ is used, where the contraction coefficients X_{pq}^{vw} are determined *a priori* from the electronic integrals. (For the spin-orbital form see, e.g., Ref. 61, the spin-free form used here is given in the accompanying paper, Ref. 58.)

While the energy denominators in perturbation theory normally depend on both the active space and external labels of the excitation, strong contraction makes the assumption that the energy denominators vary only with the external orbitals, there being some average energy associated with removing electrons from the active space. In this way, strong contraction defines a very significant reduction in the total number of excitation operators. Most importantly, because there is only one excitation operator to each or each pair of external orbitals, the first-order interacting space in the strongly contracted theory is orthogonal by construction and we do not need the $O(n_{\text{act}}^2)$ diagonalization cost for the stabilization of the CT equations. The resulting strongly contracted CT theory (SC-CT) is instead stabilized on the basis of the matrix elements of the Dyll Hamiltonian that is the zeroth order Hamiltonian used in NEVPT2,^{57,60,61} and which is diagonal in the strongly contracted first-order interacting space. This scheme requires a single truncation threshold τ with units of energy, corresponding to the smallest out-of-active-space excitation energy in the Dyll Hamiltonian. For further details on strongly contracted CT theory, we refer to Ref. 58.

Finally, we note that while both the overlap truncation and the energy truncation in strong contraction formally break the size consistency of the CT theory, in practice we have observed negligible size-consistency errors (below $0.01mE_h$).⁵⁹

C. Combining DMRG and CT

Let us summarize the key ingredients that make up a joint DMRG and CT calculation. We assume the one- and two-electron integrals in an orthogonal basis (such as the canonical or localized HF molecular orbitals) are available and that a partitioning into core, active, and external spaces has been decided.

- (1) For the DMRG calculation we specify the ordering of the active-space orbitals and the number of states M [the dimension of the matrices in the wave function (2)]. We also specify a level of perturbative noise (that is gradually reduced) to bypass local minima in the DMRG sweeps.
- (2) The DMRG sweeps with orbital optimization are converged producing one- and two-body density matrices in the active space.
- (3) We choose a stabilization strategy for the CT equations: either overlap truncation (requiring thresholds τ_1 for the singles and semi-internal singles and τ_2 for the ex-

ternal doubles), or strong contraction together with truncation according to the Dyll Hamiltonian (a single threshold τ with units of energy).

- (4) The CTSD or SC-CTSD amplitude equations are converged using the active-space one- and two-body density matrices of the DMRG calculation.

III. RESULTS

A. Benchmark correlation energies of polyenes

The performance of the CT method together with a traditional CASSCF description of static correlation (which can be regarded as producing results identical to the joint DMRG-CT theory) has previously been assessed.^{46–48} Here we focus on evaluating the joint DMRG-CT theory in problems where the complete valence active space is not accessible to traditional CASSCF, but instead must be treated using DMRG-SCF. The natural question then arises as to what is a suitable benchmark. In such larger systems accurate reference correlation energies are only available for single-reference electronic structure (where we can compare to results from coupled cluster theory). This limit is, however, an important test of a multireference theory. If we use a complete valence active space approach to study a chemical transformation that changes between multireference and single-reference electronic structure, an accurate treatment of the single-reference limit is essential. (Note that the single-reference limit here refers to the description of a single-reference type problem within a complete valence active space multireference method).

As a benchmark model to assess the joint DMRG-CT theory we have therefore performed single-point energy calculations on *all-trans* polyenes, C_nH_{n+2} ($n=4, 8, \dots, 24$) with the 6-31G (Ref. 80) basis set. The ground-state of the neutral polyenes is considered to be accurately described by the single-reference CCSD(T) method, so we will regard the CCSD(T) correlation energies as a reliable benchmark for comparison.

The geometries were first optimized using the B3LYP method with the 6-31G basis set. Restricted HF calculations were then carried out in the 6-31G basis to determine a starting set of orbitals. DMRG-SCF calculations were carried out in the complete π valence space [corresponding to an n electron, n orbital ($[n]e, [n]o$) CAS] using up to $M=700$ states. [In Ref. 28, we showed that in these systems, this number of states was sufficient for the DMRG to obtain near-exact (beyond $0.1 mE_h$) active-space correlation energies.] Subsequent CT calculations were performed using the overlap truncation method with $\tau_1=0.5$, $\tau_2=0.1$ using only one- and two-body density matrix information.

For comparison, traditional single-reference coupled cluster CCSD(T) and multireference complete active space perturbation theory (CASPT2)^{3,65} and multireference averaged coupled pair functional (MRACPF) (Ref. 2) calculations were also performed, using the MOLPRO package.⁸¹ The multireference calculations used a full π valence space optimized CAS reference. Because of the exponential blow-up of the CAS space, the largest polyene that could be studied with

TABLE I. Correlation energies E_h of neutral polyenes (C_nH_{n+2} , $n=4,8,\dots,24$) as obtained from CCSD(T), DMRG-SCF, DMRG-CTSD, CASPT2, and MRACPF calculations using a 6-31G basis. Total energies E_h of the uncorrelated restricted HF calculations are included. The multireference calculations employ full π valence references corresponding to $([n]e,[n]o)$ complete active spaces.

	Correlation energy					HF total energy (uncorrelated)
	CCSD(T)	DMRG-SCF	DMRG-CTSD	CASPT2	MRACPF	
C_4H_6	-0.40957	-0.06029	-0.40965	-0.35989	-0.40333	-154.86348
C_8H_{10}	-0.80641	-0.11963	-0.80348	-0.71768	-0.79158	-308.58396
$C_{12}H_{14}$	-1.20393	-0.17919	-1.19831			-462.30442
$C_{16}H_{18}$	-1.60176	-0.23893	-1.59218			-616.02462
$C_{20}H_{22}$	-1.99975	-0.29878	-1.98750			-769.74462
$C_{24}H_{26}$	-2.39778	-0.35868	-2.38252			-923.46451

CASPT2 and MRACPF using the full π valence space was C_8H_{10} , corresponding to a CAS(8e,8o) reference.

Table I shows the correlation energies (i.e., the difference from the restricted HF energy) obtained with the 6-31G basis, using the DMRG-SCF and DMRG-CTSD methods, as well as the comparison methods. We observe that DMRG-CTSD correlation energies are very close to those of the reference CCSD(T) results, and MRACPF similarly reproduces the CCSD(T) correlation energies well. CASPT2 appears to underestimate the correlation energies due to its low order treatment of dynamic correlation. Unlike in the case of CASPT2/MRACPF, the joint DMRG-CT theory allows us to continue to use the full π -valence space for the larger polyenes, including the (24e,24o) active space for the $C_{24}H_{26}$ molecule. In these longer systems, the DMRG-CT correlation energies continue to closely track the CCSD(T) correlation energies. (Interestingly, although the CCSD(T) energies are always lower than the DMRG-CTSD energies, we have previously shown that CCSD(T) provides a slight overestimate of the correlation energy in these systems, at least in a minimal basis²⁸). The maximum percentage error of the DMRG-CTSD correlation energy compared with the CCSD(T) correlation energy was about 0.7% (as compared to 1.8% and 12.1% for MRACPF and CASPT2, respectively). In addition to demonstrating the quality of the CT treatment of the dynamic correlation, this also illustrates the scalability and size-extensivity of the DMRG-CT theory in larger problems.

B. Isomerization of $[Cu_2O_2]^{2+}$

Transition metal chemistry is a natural area of application for the joint DMRG-CT theory, due to the large active spaces typically involved. Here, we apply the DMRG-CT theory to the $[Cu_2O_2]^{2+}$ molecule, whose derivatives have been extensively studied as model catalysts, and which form the active site of certain enzymes that perform O_2 activation.^{82,83} One of the primary postulated reaction pathways utilized by $[Cu_2O_2]^{2+}$ cores is an isomerization between the bis(μ -oxo) and μ - η^2 : η^2 peroxo isomers, shown schematically in Fig. 1. One can construct a simple reaction coordinate for this isomerization that linearly interpolates between the two structures.⁷³ From a theoretical point of view, the resulting isomerization curve has been considered quite challenging. This is because the state is very sensitive to the

balance between static and dynamic correlation.⁸³ Consequently, any complete active space treatment is likely to fail unless a very large active space is used. For example, while it is argued that there is no minimum along the isomerization coordinate,^{73,83} previous CASPT2 studies that used a moderate active space (e.g., 16e and 14o) uniformly produced such a minimum!^{73,84,85} Only a very recent RASPT2 study⁷⁴ with a large (24e,28o) active space gave a reasonable isomerization profile without an unphysical minimum.

The joint DMRG-CT theory is a natural candidate to treat the static and dynamic correlation associated with the large active spaces in these kinds of problems. Marti *et al.*⁴² studied relative energy differences of the $[Cu_2O_2]^{2+}$ core using the DMRG, although unfortunately their calculations were not fully converged. In an earlier study, two of us³⁴ obtained a qualitatively correct isomerization curve using the DMRG alone in a very large (32e,62o) active space. This active space was sufficiently large to describe some dynamic correlation effects in addition to the static correlation. Nonetheless, a full account of the dynamic correlation requires a still much larger correlated space which cannot be efficiently treated through only using the DMRG. The efficient separate description of static and dynamic correlation in the joint DMRG-CT method allows us to re-examine the isomerization of $[Cu_2O_2]^{2+}$ with a proper account of the dynamic correlation using large basis sets and correlation spaces.

For our calculations, geometries along the isomerization coordinate were taken from the study of Cramer.⁷³ Note that the isomerization is parametrized by a single variable F that corresponds to the degree of isomerization, i.e., $F=0.0$ and $F=1.0$ represent the bis(μ -oxo) and the μ - η^2 : η^2 peroxo isomers, respectively.

Two different basis sets were used. In the small basis (BS1), we used the Stuttgart pseudopotential along with the associated basis functions⁸⁶ for the Cu atoms, and the atomic natural orbital (ANO) basis set⁸⁷ for O, corresponding to $Cu[8s7p6d/6s5p3d] + ECP$ and $O:[14s9p4d/4s3p2d]$ con-

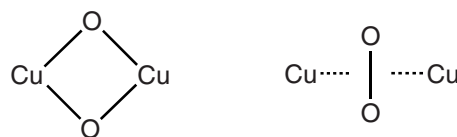


FIG. 1. Left: bis(μ -oxo). Right: μ - η^2 : η^2 -peroxo forms of $[Cu_2O_2]^{2+}$.

TABLE II. DMRG energies and energy differences (E_h) for the bis(μ -oxo) ($F=0.0$) and μ - η^2 : η^2 peroxo ($F=1.0$) isomers in the BS1 basis set with various number of states (M).

M	bis(μ -oxo) ($F=0.0$)	μ - η^2 : η^2 peroxo ($F=1.0$)	ΔE (kcal mol $^{-1}$)
512	-541.757 548	-541.801 050	27.3
1024	-541.763 670	-541.805 239	26.1
2048	-541.766 586	-541.807 191	25.5
∞^a	-541.768 414	-541.809 226	25.6
RASSCF	-541.735 787	-541.792 831	35.8

^aThe large M limit energies are estimated by the extrapolation procedure in Ref. 34.

tractions. This basis was the same as that used in earlier CASPT2/RASPT2 studies.^{73,74} In the large basis (BS2), we used an ANO basis set^{87,88} corresponding to Cu: $[21s15p10d6f4g/8s7p5d3f2g]$ and O: $[14s9p4d3f/5s4p3d2f]$ contractions.

The DMRG-SCF calculations were performed in a $(28e, 32o)$ active space consisting of all the molecular orbitals originating from the relevant valence and second-order correlating shell Cu: $3d, 4d$ and O: $2p, 3p$ orbitals. The molecular orbitals were ordered on the DMRG 1D lattice sites grouped by symmetry $\{a_g, b_{1u}, b_{2u}, b_{3g}, b_{3u}, b_{2g}, b_{1g}, \text{ and } a_u\}$ and by orbital energy within each symmetry block, where the D_{2h} labels correspond to setting the molecule in the yz -plane with the two Cu atoms on the z -axis. The RASSCF calculations were also performed in the same active space but all configurations with more than four excited electrons relative to the HF configuration were neglected. (This RAS model, which corresponds essentially to a CISDTQ calculation in the RAS1/RAS2/RAS3 spaces, follows the choice made by Malmqvist *et al.* in Ref. 74.) To evaluate the dynamic correlation, strongly contracted CT calculations ($\tau=0.1E_h$) were carried out on top of the DMRG-SCF reference wave function. The correlations of the inactive orbitals Cu: $4s$ and O: $2s$ were included in the CT calculations, and the other inactive orbitals were defined as frozen core orbitals.

To assess the quality of our DMRG calculations and the consequent description of the static correlation in the active space, we examined the convergence of our DMRG-SCF calculations as a function of the DMRG number of states M . Table II shows DMRG-SCF energies (E_h) for the bis(μ -oxo) and μ - η^2 : η^2 peroxo isomers using BS1 with various number of states (M). The energy at the infinite M (untruncated)

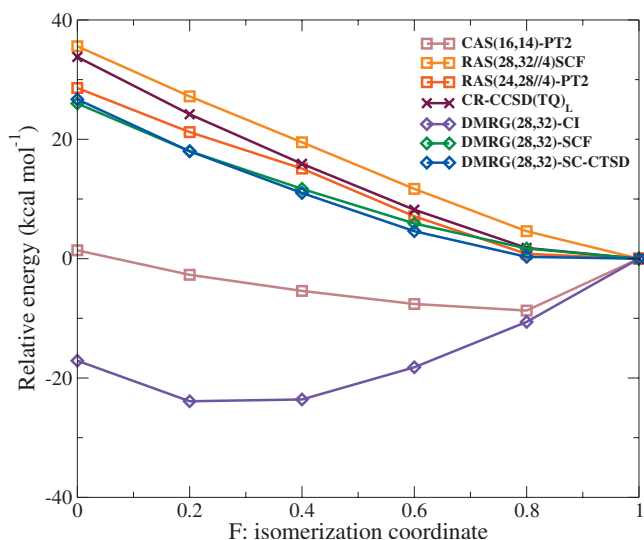


FIG. 2. $[\text{Cu}_2\text{O}_2]^{2+}$ energies along the isomerization coordinate using basis set BS1. All energies (in kcal mol $^{-1}$) are measured relative to that of the μ - η^2 : η^2 peroxo isomer. The isomerization coordinate linearly connects the bis(μ -oxo) ($F=0.0$) and μ - η^2 : η^2 peroxo ($F=1.0$) isomers. Geometries and CASPT2, CR-CC energies taken from Ref. 73. RASPT2 energies taken from Ref. 74.

limit was evaluated by a linear extrapolation from the DMRG-SCF energies and total discarded weights.²⁴ The DMRG-SCF energies with $M=512, 1024$, and 2048 were above the extrapolated infinite M limit energy by $10.8, 4.7$, and $1.8mE_h$ for the bis(μ -oxo) isomer and $8.1, 3.9$, and $2.0mE_h$ for the μ - η^2 : η^2 peroxo isomer, respectively. As compared to the total energies, the relative energies were even more rapidly converged due to the well balanced nature of the DMRG wave function for the different isomers, and are within chemical accuracy (0.5 kcal mol $^{-1}$ error) even with $M=1024$. By contrast, since the RASSCF wave function neglects a large number of excited configurations in the CAS($28e, 32o$) space, the RASSCF energies were much higher than the estimated exact CASSCF energies (corresponding to the large M limit of the DMRG-SCF) by 32.6 and $16.3mE_h$ for the bis(μ -oxo) and μ - η^2 : η^2 peroxo isomer, respectively, and the relative energy of the isomers showed a large error (10.2 kcal mol $^{-1}$) as compared with the DMRG-SCF results.

Table III and Fig. 2 show the relative energies computed by various methods, using the small basis BS1, for the isomerization coordinate that linearly connects the

TABLE III. $[\text{Cu}_2\text{O}_2]^{2+}$ energies along the isomerization coordinate using basis set BS1. All energies (in kcal mol $^{-1}$) are measured relative to that of the μ - η^2 : η^2 peroxo isomer. The isomerization coordinate linearly connects the bis(μ -oxo) ($F=0.0$) and μ - η^2 : η^2 peroxo ($F=1.0$) isomers.

$F=$	0.0	0.2	0.4	0.6	0.8	1.0
DMRG-CI	-17.1	-23.9	-23.6	-18.2	-10.6	0.0
DMRG-SCF	26.1	18.0	11.7	5.9	1.7	0.0
DMRG-SC-CTSD	27.1	18.6	11.0	3.9	0.2	0.0
CR-CCSD(TQ) _L ^a	33.8	24.2	15.9	8.2	1.8	0.0
RASPT2 ^b	28.6	21.2	15.1	7.1	0.8	0.0

^aReference 73, supporting information, computed as CR-CCSD(T)_L+[CR-CCSD(TQ)-CR-CCSD(T)].

^bReference 74.

TABLE IV. $[\text{Cu}_2\text{O}_2]^{2+}$ energies along the isomerization coordinate using basis set BS2. All energies (in kcal mol⁻¹) are measured relative to that of the $\mu\text{-}\eta^2\text{:}\eta^2$ peroxy isomer. The isomerization coordinate linearly connects the bis($\mu\text{-oxo}$) ($F=0.0$) and $\mu\text{-}\eta^2\text{:}\eta^2$ peroxy ($F=1.0$) isomers.

$F=$	0.0	0.2	0.4	0.6	0.8	1.0
DMRG-CI	-12.8	-20.9	-21.5	-16.7	-10.0	0.0
DMRG-SCF	26.4	17.9	11.0	5.1	1.1	0.0
DMRG-SC-CTSD	37.4	29.0	22.0	14.4	6.1	0.0
CCSD(T)	38.6	29.8	21.5	12.9	4.7	0.0

bis($\mu\text{-oxo}$) and $\mu\text{-}\eta^2\text{:}\eta^2$ peroxy isomers. We used $M=1024$ for all these DMRG calculations. Comparing the DMRG calculation without orbital optimization (denoted DMRG-CI) and the DMRG-SCF results, we see that orbital optimization drastically shifts the relative energies, resulting in a monotonically decreasing curve as the isomerization progresses and a most stable structure at $F=1.0$. As we have described earlier, this is believed to be the physically correct isomerization profile. As compared with the joint DMRG-SC-CT calculations, the CR-CC and RASPT2 calculations as reported in Ref. 73 and 74 appear to slightly overestimate the stability of the $\mu\text{-}\eta^2\text{:}\eta^2$ peroxy isomers. We note that in basis BS1 the dynamic correlation treatment through the CT theory only very slightly changes the relative energies from the DMRG-SCF treatment (less than 2 kcal mol⁻¹). The small contribution of dynamic correlation here is due to the small size of the basis.

In Table IV we show the relative energies computed using the large basis BS2, which allows for a more complete description of dynamic correlation. Here we see that incorporating dynamic correlation through CT theory increases the stability of the $\mu\text{-}\eta^2\text{:}\eta^2$ peroxy isomers by a remarkable 10 kcal mol⁻¹. Overall, this illustrates, as always, the importance of coupling a high-order treatment of dynamic correlation to large basis sets. We see also that the CCSD(T) energy gap between the two isomers and the DMRG-SC-CT energy gap is in this case in close agreement. We argue that this is because the particular multireference nature (predominantly biradical) of this molecule is sometimes correctly captured within the CCSD(T) formalism. However, previous studies on ligated $[\text{Cu}_2\text{O}_2]^{2+}$ species indicate that these related systems present multireference character for which CCSD(T) gives unphysical results.⁷³ It is naturally to these and more difficult challenges that we look to in the future with the joint DMRG-CT theory.

IV. CONCLUSIONS

In this work we described the joint application of the DMRG and CT theory to obtain an accurate description of both static and dynamic correlation in multireference problems. The combined DMRG-CT theory represents the logical culmination of ideas contained in earlier works on the DMRG and CT. As we have demonstrated, the combination of the DMRG and CT methods achieves a good description in both problems that require a high-order treatment of the dynamic correlation (accurate total correlation energies of long polyenes) as well as those that require the ability to treat large active spaces (isomerization of the $[\text{Cu}_2\text{O}_2]^{2+}$ core).

Further investigations using the joint DMRG-CT theory into more challenging systems are currently underway.

ACKNOWLEDGMENTS

This work was supported by the Department of Energy Office of Science Award No. DE-FG02-07ER46432, the David and Lucile Packard Foundation, the Alfred P. Sloan Foundation, and the Camille and Henry Dreyfus Foundation.

- ¹R. Ahlrichs, P. Scharf, and C. Ehrhard, *J. Chem. Phys.* **82**, 890 (1985).
- ²R. J. Gdanitz and R. Ahlrichs, *Chem. Phys. Lett.* **143**, 413 (1988).
- ³K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, *J. Phys. Chem.* **94**, 5483 (1990).
- ⁴K. Hirao, *Chem. Phys. Lett.* **190**, 374 (1992).
- ⁵S. Grimme and M. Waletzke, *J. Chem. Phys.* **111**, 5645 (1999).
- ⁶J. Gräfenstein and D. Cremer, *Chem. Phys. Lett.* **316**, 569 (2000).
- ⁷J. Olsen, B. O. Roos, P. Jørgensen, and H. J. A. Jensen, *J. Chem. Phys.* **89**, 2185 (1988).
- ⁸P.-Å. Malmqvist, A. Rendell, and B. O. Roos, *J. Phys. Chem.* **94**, 5477 (1990).
- ⁹F. Faglioni and W. Goddard, *Int. J. Quantum Chem.* **73**, 1 (1999).
- ¹⁰W. J. Hunt, P. J. Hay, and I. W. A. Goddard, *J. Chem. Phys.* **57**, 738 (1972).
- ¹¹J. Cullen, *Chem. Phys. Lett.* **202**, 217 (1996).
- ¹²T. Van Voorhis and M. Head-Gordon, *J. Chem. Phys.* **117**, 9190 (2002).
- ¹³V. A. Rassolov and F. Xu, *J. Chem. Phys.* **127**, 044104 (2007).
- ¹⁴J. A. Parkhill, K. Lawler, and M. Head-Gordon, *J. Chem. Phys.* **130**, 084101 (2009).
- ¹⁵P. Piecuch, N. Oliphant, and L. Adamowicz, *J. Chem. Phys.* **99**, 1875 (1993).
- ¹⁶P. D. Fan and S. Hirata, *J. Chem. Phys.* **124**, 104108 (2006).
- ¹⁷M. Nakata, H. Nakatsuji, M. Ehara, M. Fukuda, K. Nakata, and K. Fujisawa, *J. Chem. Phys.* **114**, 8282 (2001).
- ¹⁸D. A. Mazziotti, *Phys. Rev. Lett.* **93**, 213001 (2004).
- ¹⁹J. Paldus and X. Li, *Adv. Chem. Phys.* **110**, 1 (1999).
- ²⁰S. R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).
- ²¹S. R. White, *Phys. Rev. B* **48**, 10345 (1993).
- ²²S. R. White and R. L. Martin, *J. Chem. Phys.* **110**, 4127 (1999).
- ²³G. K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.* **116**, 4462 (2002).
- ²⁴G. K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.* **118**, 8551 (2003).
- ²⁵G. K.-L. Chan, *J. Chem. Phys.* **120**, 3172 (2004).
- ²⁶G. K.-L. Chan, M. Kállay, and J. Gauss, *J. Chem. Phys.* **121**, 6110 (2004).
- ²⁷G. K.-L. Chan and T. Van Voorhis, *J. Chem. Phys.* **122**, 204101 (2005).
- ²⁸J. Hachmann, W. Cardoen, and G. K.-L. Chan, *J. Chem. Phys.* **125**, 144101 (2006).
- ²⁹J. Hachmann, J. J. Dorando, M. Avilés, and G. K.-L. Chan, *J. Chem. Phys.* **127**, 134309 (2007).
- ³⁰J. J. Dorando, J. Hachmann, and G. K.-L. Chan, *J. Chem. Phys.* **127**, 084109 (2007).
- ³¹D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan, *J. Chem. Phys.* **128**, 144117 (2008).
- ³²J. J. Dorando, J. Hachmann, and G. K.-L. Chan, *J. Chem. Phys.* **130**, 184111 (2009).
- ³³G. K.-L. Chan, *Phys. Chem. Chem. Phys.* **10**, 3454 (2008).
- ³⁴Y. Kurashige and T. Yanai, *J. Chem. Phys.* **130**, 234114 (2009).
- ³⁵S. Daul, I. Ciofini, C. Daul, and S. R. White, *Int. J. Quantum Chem.* **79**, 331 (2000).

- ³⁶ A. O. Mitrushenkov, G. Fano, F. Ortolani, R. Linguerri, and P. Palmieri, *J. Chem. Phys.* **115**, 6815 (2001).
- ³⁷ A. O. Mitrushenkov, R. Linguerri, P. Palmieri, and G. Fano, *J. Chem. Phys.* **119**, 4148 (2003).
- ³⁸ Ö. Legeza, J. Röder, and B. A. Hess, *Phys. Rev. B* **67**, 125114 (2003).
- ³⁹ Ö. Legeza, J. Röder, and B. A. Hess, *Mol. Phys.* **101**, 2019 (2003).
- ⁴⁰ G. Moritz, B. A. Hess, and M. Reiher, *J. Chem. Phys.* **122**, 024107 (2005).
- ⁴¹ G. Moritz, A. Wolf, and M. Reiher, *J. Chem. Phys.* **123**, 184105 (2005).
- ⁴² K. H. Marti, I. M. Ondík, G. Moritz, and M. Reiher, *J. Chem. Phys.* **128**, 014104 (2008).
- ⁴³ D. Zgid and M. Nooijen, *J. Chem. Phys.* **128**, 014107 (2008).
- ⁴⁴ D. Zgid and M. Nooijen, *J. Chem. Phys.* **128**, 144116 (2008).
- ⁴⁵ D. Zgid and M. Nooijen, *J. Chem. Phys.* **128**, 144115 (2008).
- ⁴⁶ T. Yanai and G. K.-L. Chan, *J. Chem. Phys.* **124**, 194106 (2006).
- ⁴⁷ T. Yanai and G. K.-L. Chan, *J. Chem. Phys.* **127**, 104107 (2007).
- ⁴⁸ E. Neuscamman, T. Yanai, and G. K.-L. Chan, *J. Chem. Phys.* **130**, 124102 (2009).
- ⁴⁹ S. R. White, *J. Chem. Phys.* **117**, 7472 (2002).
- ⁵⁰ D. Mukherjee, in *Recent Progress in Many-Body Theories*, edited by E. Schachinger (Plenum, New York, 1995), Vol. 4, p. 127.
- ⁵¹ D. Mukherjee, *Chem. Phys. Lett.* **274**, 561 (1997).
- ⁵² W. Kutzelnigg and D. Mukherjee, *J. Chem. Phys.* **107**, 432 (1997).
- ⁵³ F. Colmenero and C. Valdemoro, *Phys. Rev. A* **47**, 979 (1993).
- ⁵⁴ H. Nakatsuji and K. Yasuda, *Phys. Rev. Lett.* **76**, 1039 (1996).
- ⁵⁵ D. A. Mazziotti, *Phys. Rev. A* **57**, 4219 (1998).
- ⁵⁶ W. Kutzelnigg and D. Mukherjee, *J. Chem. Phys.* **110**, 2800 (1999).
- ⁵⁷ C. Angeli, R. Cimирaglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu, *J. Chem. Phys.* **114**, 10252 (2001).
- ⁵⁸ E. Neuscamman, T. Yanai, and G. K.-L. Chan, *J. Chem. Phys.* **132**, 024106 (2010).
- ⁵⁹ E. Neuscamman, T. Yanai, and G. K.-L. Chan, "A Review of Canonical Transformation Theory," *Int. Rev. Phys. Chem.* (submitted).
- ⁶⁰ C. Angeli, M. Pastore, and R. Cimирaglia, *Theor. Chem. Acc.* **117**, 743 (2007).
- ⁶¹ D. Zgid, D. Ghosh, E. Neuscamman, and G. K.-L. Chan, *J. Chem. Phys.* **130**, 194107 (2009).
- ⁶² R. J. Buenker and S. D. Peyerimhoff, *Theor. Chim. Acta* **35**, 33 (1974).
- ⁶³ P. E. M. Siegbahn, *J. Chem. Phys.* **72**, 1647 (1980).
- ⁶⁴ H.-J. Werner and P. J. Knowles, *J. Chem. Phys.* **89**, 5803 (1988).
- ⁶⁵ P. Celani and H. Werner, *J. Chem. Phys.* **112**, 5546 (2000).
- ⁶⁶ A. Banerjee and J. Simons, *Int. J. Quantum Chem.* **19**, 207 (1981).
- ⁶⁷ A. Banerjee and J. Simons, *J. Chem. Phys.* **76**, 4548 (1982).
- ⁶⁸ M. R. Hoffmann and J. Simons, *J. Chem. Phys.* **88**, 993 (1988).
- ⁶⁹ L. G. Kong, K. R. Shamasundar, O. Demel, and M. Nooijen, *J. Chem. Phys.* **130**, 114101 (2009).
- ⁷⁰ D. Mukherjee and W. Kutzelnigg, *J. Chem. Phys.* **114**, 2047 (2001).
- ⁷¹ D. A. Mazziotti, *Phys. Rev. Lett.* **97**, 143002 (2006).
- ⁷² A. E. DePrince III and D. A. Mazziotti, *J. Chem. Phys.* **127**, 104104 (2007).
- ⁷³ C. J. Cramer, M. Włoch, P. Piecuch, C. Puzzarini, and L. Gagliardi, *J. Phys. Chem. A* **110**, 1991 (2006).
- ⁷⁴ P.-Å. Malmqvist, K. Pierloot, A. R. M. Shahi, C. J. Cramer, and L. Gagliardi, *J. Chem. Phys.* **128**, 204109 (2008).
- ⁷⁵ G. K.-L. Chan, J. J. Dorando, D. Ghosh, J. Hachmann, E. Neuscamman, H. Wang, and T. Yanai, *Prog. Theor. Chem. Phys.* **18**, 49 (2008); G. K.-L. Chan, J. J. Dorando, D. Ghosh, J. Hachmann, E. Neuscamman, H. Wang, and T. Yanai, e-print arXiv:0711.1398.
- ⁷⁶ U. Schollwöck, *Rev. Mod. Phys.* **77**, 259 (2005).
- ⁷⁷ K. Ruedenberg, M. W. Schmidt, M. M. Gilbert, and S. T. Elbert, *Chem. Phys.* **71**, 41 (1982).
- ⁷⁸ B. O. Roos, *Adv. Chem. Phys.* **69**, 399 (1987).
- ⁷⁹ T. Yanai, Y. Kurashige, D. Ghosh, and G. K.-L. Chan, *Int. J. Quantum Chem.* **109**, 2178 (2009).
- ⁸⁰ W. J. Hehre, R. Ditchfield, and J. Pople, *J. Chem. Phys.* **56**, 2257 (1972).
- ⁸¹ MOLPRO, a package of *ab initio* programs designed by H.-J. Werner and P. J. Knowles, version 2006, R. D. Amos, A. Bernhardsson, A. Berning *et al.*
- ⁸² E. I. Solomon, U. M. Sundaram, and T. E. Machonkin, *Chem. Rev. (Washington, D.C.)* **96**, 2563 (1996).
- ⁸³ B. F. Gherman and C. J. Cramer, *Coord. Chem. Rev.* **253**, 723 (2009).
- ⁸⁴ M. Flock and K. Pierloot, *J. Phys. Chem. A* **103**, 95 (1999).
- ⁸⁵ M. F. Rode and H.-J. Werner, *Theor. Chem. Acc.* **114**, 309 (2005).
- ⁸⁶ M. Dolg, U. Wedig, H. Stoll, and H. Preuss, *J. Chem. Phys.* **86**, 866 (1987).
- ⁸⁷ P. O. Widmark, P.-Å. Malmqvist, and B. O. Roos, *Theor. Chim. Acta* **77**, 291 (1990).
- ⁸⁸ R. Pou-Amérigo, M. Merchan, I. Nebot-Gil, P. O. Widmark, and B. O. Roos, *Theor. Chim. Acta* **92**, 149 (1995).