JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. 2019, 15, 4399−4414

pubs.acs.org/JCTC

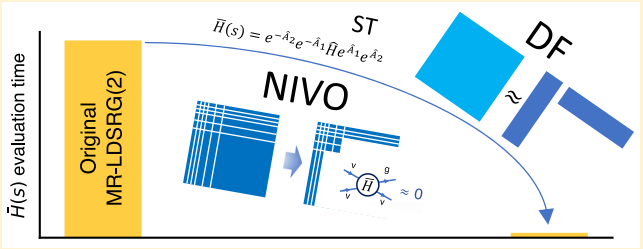
Improving the Efficiency of the Multireference Driven Similarity Renormalization Group via Sequential Transformation, Density Fitting, and the Noninteracting Virtual Orbital Approximation

Tianyuan Zhang,* Chenyang Li,* and Francesco A. Evangelista*

Department of Chemistry and Cherry L. Emerson Center for Scientific Computation, Emory University, Atlanta, Georgia 30322, United States

Supporting Information

ABSTRACT: This study examines several techniques to improve the efficiency of the linearized multireference driven similarity renormalization group truncated to one- and two-body operators [MR-LDSRG(2)]. We propose a sequential MR-LDSRG(2) [sq-MR-LDSRG(2)] scheme, in which one-body substitutions are folded exactly into the Hamiltonian. This new approach is combined with density fitting (DF) to reduce the storage cost of two-electron integrals. To further avoid storage of large four-index intermediates, we propose a noninteracting virtual orbital (NIVO) approximation of the Baker−Campbell−Hausdorff series that neglects commutators terms with three and four virtual indices. The NIVO approximation reduces the computational prefactor of the MR-LDSRG(2), bringing it closer to that of coupled cluster with singles and doubles (CCSD). We test the effect of the DF and NIVO approximations on the MR-LDSRG(2) and sq-MR-LDSRG(2) methods by computing properties of eight diatomic molecules. The diatomic constants obtained by DF-sq-MR-LDSRG(2)+NIVO are found to be as accurate as those from the original MR-LDSRG(2) and coupled cluster theory with singles, doubles, and perturbative triples. Finally, we demonstrate that the DF-sq-MR-LDSRG(2)+NIVO scheme can be applied to chemical systems with more than 550 basis functions by computing the automerization energy of cyclobutadiene with a quintuple-ζ basis set. The predicted automerization energy is found to be similar to the value computed with Mukherjee's state-specific multireference coupled cluster theory with singles and doubles.



1. INTRODUCTION

The failure of conventional many-body methods to describe near-degenerate electronic states has motivated the development of many efficient and practical multireference approaches, including multireference perturbation theories (MRPTs)[1−7] and multireference configuration interaction (MRCI) schemes.[8−12] Considerable efforts have been dedicated to the development of multireference coupled cluster (MRCC) methods,[13−27] with the goal of creating nonperturbative theories that are both size extensive and systematically improvable. Analogous many-body methods based on unitary transformations have received less attention.[27−34] Unitary theories yield Hermitian transformed Hamiltonians and, therefore, can be interfaced with many exact and approximated methods for Hamiltonian diagonalization. This property is important both for the formulation of multireference theories and in new applications of unitary methods to quantum computing.[35−47]

One of the main obstacles in the formulation of both single- and multireference unitary coupled cluster theories is that they lead to nonterminating equations. The central quantity evaluated in these approaches is the similarity transformed Hamiltonian ($\bar{H}$) defined as

$$\hat{H} \rightarrow \bar{H} = \hat{U}^{\dagger}\hat{H}\hat{U} = e^{-\hat{A}}\hat{H}e^{\hat{A}} \tag{1}$$

where $\hat{H}$ is the bare Hamiltonian and $\hat{U}$ is a unitary operator. In writing this transformation, we have expressed $\hat{U}$ as the exponential of the anti-Hermitian operator $\hat{A}$ ($\hat{A}^{\dagger} = -\hat{A}$), which is commonly written in terms of the coupled cluster excitation operator $\hat{T}$ as $\hat{A} = \hat{T} - \hat{T}^{\dagger}$. Using the Baker−Campbell−Hausdorff (BCH) identity,[29,31,33] the transformed Hamiltonian may be computed as the following commutator series

$$\bar{H} = \hat{H} + [\hat{H}, \hat{A}] + \frac{1}{2!}[[\hat{H}, \hat{A}], \hat{A}] + \frac{1}{3!}[[[\hat{H}, \hat{A}], \hat{A}], \hat{A}] + \cdots \tag{2}$$

Since the operator $\hat{A}$ contains both excitations and de-excitations, contractions are possible among components of $\hat{A}$, and as a consequence, the BCH series given in eq 2 is nonterminating.

Received: April 15, 2019
Published: July 3, 2019

Downloaded via UNIV TOULOUSE III - PAUL SABATIER on February 10, 2021 at 15:33:52 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

© 2019 American Chemical Society

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

Various approximations have been proposed to evaluate the unitarily transformed Hamiltonian. Perhaps the simplest approximation is truncating the BCH expansion after a certain number of commutators.[29,31,48] Proof-of-principle studies on unitary coupled cluster (CC) theory[48] suggest that for a series containing up to n-nested commutators, the error decays as 10^{-n}, and about four commutators are necessary to achieve sub-milliHartree accuracy. Taube and Bartlett[33] have suggested tractable approximations to unitary CC theory based on the Zassenhaus expansion that are exact for a given number of electrons. A convenient way to truncate the unitary BCH series was suggested by Yanai and Chan.[49] In their linear truncation scheme, each single commutator $[\cdot, \hat{A}]$ in the BCH series is approximated with its scalar and one- and two-body components, which we indicate as $[\cdot, \hat{A}]_{0,1,2}$. Since in this truncation scheme, the commutator $[\cdot, \hat{A}]_{0,1,2}$ preserves the many-body rank (number of creation and annihilation operators) of the Hamiltonian, the full BCH series can then be evaluated via a recursive relation. Consequently, the linear truncation scheme is equivalent to a summation of a subset of the diagrams that contribute to the BCH series, some of which enter with incorrect prefactors.[50] An advantage of this approach is that closed-form expressions for terms like $[\hat{O}, \hat{A}]_{0,1,2}$ can be easily derived, where $\hat{O}$ is an operator containing up to two-body terms. This truncation scheme has been employed in canonical transformation (CT) theory[49,51,52] and has been used to truncate normal-ordered equations in the flow renormalization group of Wegner.[53,54]

We have recently developed a multireference driven similarity renormalization group (MR-DSRG)[55−60] approach that avoids the multiple-parentage problem[21,23,61−64] and numerical instabilities[63−69] encountered in other nonperturbative multireference methods. In MR-DSRG, we perform a unitary transformation of the Hamiltonian controlled by a flow parameter, which determines to which extent the resulting effective Hamiltonian is band diagonal.[56,60] The simplest nonperturbative approximation, the linearized MR-DSRG truncated to two-body operators [MR-LDSRG(2)],[57] assumes that $\hat{A}$ contains up to two-body operators (singles and doubles) and employs the linear commutator approximation as described in the previous paragraph. Preliminary benchmarks indicate that the MR-LDSRG(2) method is more accurate than CCSD around equilibrium geometries, and this accuracy is preserved along potential energy curves, especially for single-bond breaking processes.[58] The cost to evaluate a single commutator in the MR-LDSRG(2) scales as $O(N_C^2 N_V^2 N^2) = O(N_C^2 N_V^4 + N_C^3 N_V^3 + \cdots)$ where N_C, N_V, and N are the numbers of core, virtual, and total orbitals, respectively. This scaling is identical to that of CC with singles and doubles (CCSD) $[O(N_C^2 N_V^4)]$. However, while in CCSD the most expensive term is evaluated only once, in the MR-LDSRG(2) the same term must be evaluated for each nested commutator in the BCH series. Consequently, if the BCH series is truncated after $n + 1$ terms, the computational cost of the MR-LDSRG(2) is roughly n times that of CCSD, where approximately 10 or more terms are usually required to converge the Frobenius norm of $\bar{H}$ to 10^{-12} E_h.[57] Moreover, computing the BCH series requires storing large intermediate tensors of size that scales as $O(N^4)$. When these intermediates are stored in memory, practical computations are limited to 200−300 basis functions on a single computer node.

In this work, we combine a series of improvements and approximations to reduce the computational and memory requirements of the MR-LDSRG(2) down to a small multiple of the cost of CCSD. To begin with, we consider an alternative ansatz for the MR-DSRG based on a sequential similarity transformation.[70] This sequential ansatz reduces the complexity of the MR-DSRG equations and, when combined with integral factorization techniques, reduces significantly the cost to evaluate singles contributions. Second, we apply density fitting (DF)[71−73] to reduce the memory requirements and the I/O cost by avoiding the storage of two-electron repulsion integrals. Together with Cholesky decomposition (CD)[74−78] and other tensor decomposition schemes,[79] these techniques have been crucial in enabling computations with 1000 or more basis functions and have been applied in numerous electronic structure methods,[80−91] including coupled cluster methods.[92−96] Third, we reduce the cost of MR-LDSRG(2) computations by neglecting operators that involve three or more virtual electrons. We term this truncation scheme the noninteracting virtual orbital (NIVO) approximation. A perturbative analysis of the NIVO approximation shows that the errors introduced appear at order three in the amplitudes and order four in the energy. To the best of our knowledge, such approximation has not been introduced in multireference theories, but it is analogous to other truncation schemes used in CCSD in which certain diagrams have modified coefficients[50,97−99] or are completely removed.[100,101]

This paper is organized in the following way. In Section 2, we present an overview of the MR-DSRG theory, discuss the sequential MR-DSRG, and introduce the NIVO approximation. Details of the implementation together with a discussion of timings are given in Section 3. In Section 4, we assess the accuracy of several MR-LDSRG(2) schemes on a benchmark set of several diatomic molecules and determine the automerization barrier of cyclobutadiene. Finally, we conclude this work with a discussion of the main results in Section 5.

2. THEORY

We first define the orbital labeling convention employed in this work. The set of molecular spin orbitals $\mathbf{G} \equiv \{\phi^p, p = 1, 2, ..., N\}$ is partitioned into core ($\mathbf{C}$), active ($\mathbf{A}$), and virtual ($\mathbf{V}$) components of sizes N_C, N_A, and N_V, respectively. We use the indices m, n to label core orbitals, u, v, x, y to label active orbitals, and e, f, g, h to label virtual orbitals. For convenience, we also define the set of hole ($\mathbf{H} = \mathbf{C} \cup \mathbf{A}$) and particle ($\mathbf{P} = \mathbf{A} \cup \mathbf{V}$) orbitals with dimensions $N_H = N_C + N_A$ and $N_P = N_A + N_V$, respectively. Hole orbitals are denoted by the indices i, j, k, l, and those of particle by the indices a, b, c, d. General orbitals ($\mathbf{G}$) are labeled by the indices p, q, r, s.

The MR-DSRG assumes a multideterminantal reference wave function (Ψ_0) given by a linear combination of a set of determinants $\{\Phi^\mu; \mu = 1, 2, ..., d\}$ that spans a complete active space (CAS)

$$|\Psi_0\rangle = \sum_{\mu=1}^{d} c_\mu |\Phi^\mu\rangle \tag{3}$$

The orbitals and expansion coefficients (c_μ) that enter in eq 3 are assumed to come from a CAS self-consistent field (CASSCF) computation.[102] All operators are then normal ordered with respect to Ψ_0 according to the scheme of Mukherjee and Kutzelnigg.[103] For example, the bare Hamiltonian is expressed as

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

$$\hat{H} = E_0 + \sum_{pq} f_p^q \{\hat{a}_q^p\} + \frac{1}{4}\sum_{pqrs} v_{pq}^{rs}\{\hat{a}_{rs}^{pq}\} \tag{4}$$

where $E_0 = \langle\Psi_0|\hat{H}|\Psi_0\rangle$ is the reference energy and $\{\hat{a}_{rs\cdots}^{pq\cdots}\} = \{\hat{a}^p\hat{a}^q\cdots\hat{a}_s\hat{a}_r\}$ is a product of creation ($\hat{a}^p \equiv \hat{a}_p^\dagger$) and annihilation ($\hat{a}_p$) operators in its normal-ordered form, as indicated by the curly braces. The generalized Fock matrix (f_p^q) introduced in eq 4 is defined as

$$f_p^q = h_p^q + \sum_{rs} v_{pr}^{qs}\gamma_s^r \tag{5}$$

where $h_p^q = \langle\phi_p|\hat{h}|\phi^q\rangle$ and $v_{pq}^{rs} = \langle\phi_p\phi_q\|\phi^r\phi^s\rangle$ are the one-electron and antisymmetrized two-electron integrals, respectively. Here, we have also introduced the one-particle reduced density matrix (1-RDM) defined as $\gamma_q^p = \langle\Psi_0|\hat{a}_q^p|\Psi_0\rangle$.

2.1. Review of the MR-DSRG Method. The MR-DSRG performs a parametric unitary transformation of the bare Hamiltonian analogous to eq 1, whereby the anti-Hermitian operator $\hat{A}(s)$ depends on the so-called flow parameter (s), a real number defined in the range $[0,\infty)$. The resulting transformed Hamiltonian $[\bar{H}(s)]$ is a function of s defined as

$$\bar{H}(s) = e^{-\hat{A}(s)}\hat{H}e^{\hat{A}(s)} \tag{6}$$

The operator $\hat{A}(s)$ is a sum of many-body operators with rank ranging from one up to the total number of electrons (n)

$$\hat{A}(s) = \sum_{k=1}^{n}\hat{A}_k(s) \tag{7}$$

where $\hat{A}_k(s)$ is the k-body component of $\hat{A}(s)$. In the MR-DSRG, the operators $\hat{A}_k(s)$ are written as $\hat{A}_k(s) = \hat{T}_k(s) - \hat{T}_k^\dagger(s)$, where $\hat{T}_k(s)$ is an s-dependent cluster operator defined as

$$\hat{T}_k(s) = \frac{1}{(k!)^2}\sum_{ij\cdots}^{\mathbf{H}}\sum_{ab\cdots}^{\mathbf{P}} t_{ab\cdots}^{ij\cdots}(s)\{\hat{a}_{ij\cdots}^{ab\cdots}\} \tag{8}$$

Note that the cluster amplitudes $t_{ab\cdots}^{ij\cdots}(s)$ exclude internal excitations, which are labeled only with active orbital indices. The DSRG transformed Hamiltonian has a many-body expansion similar to eq 4,

$$\bar{H}(s) = \bar{E}_0(s) + \sum_{pq}\bar{H}_q^p(s)\{\hat{a}_p^q\} + \frac{1}{4}\sum_{pqrs}\bar{H}_{rs}^{pq}(s)\{\hat{a}_{pq}^{rs}\} + \cdots \tag{9}$$

where,

$$\bar{E}_0(s) = \langle\Psi_0|\bar{H}(s)|\Psi_0\rangle \tag{10}$$

is the DSRG energy and the tensors $\bar{H}_{rs\cdots}^{pq\cdots}(s)$ are analogous to one- and two-electron integrals but dressed with dynamical correlation effects.

Here we briefly summarize the most important features of the DSRG approach, glossing over some of the more technical aspects of this formalism. For more details on the DSRG and its connection to the flow renormalization group, we recommend the reader to consult ref 60. The goal of the DSRG transformation is to decouple the interactions between the reference wave function (Ψ_0) and its excited configurations. Such interactions are the couplings between hole and particle orbitals represented by generalized excitation $[\bar{H}_{ab\cdots}^{ij\cdots}(s)\{\hat{a}_{ij\cdots}^{ab\cdots}\}]$ and de-excitation $[\bar{H}_{ij\cdots}^{ab\cdots}(s)\{\hat{a}_{ab\cdots}^{ij\cdots}\}]$ operators, where $ij\cdots \in \mathbf{H}$ and $ab\cdots \in \mathbf{P}$, excluding cases where all the indices are active orbitals. These terms of $\bar{H}(s)$ that the DSRG transformation aims to suppress are called the off-diagonal components and will be denoted as $\bar{H}_{\mathrm{od}}(s)$. Instead of achieving a full decoupling of the off-diagonal components [i.e., $\bar{H}_{\mathrm{od}}(s) = 0$], we demand that the DSRG transformation achieves a partial decoupling, where most of the electron correlation is recovered, but excitations that may cause numerical instabilities (small denominators) are excluded from the treatment of electron correlation. This goal is achieved by an appropriate parametrization of $\bar{H}_{\mathrm{od}}(s)$ that separates excitations according to the value of their Møller–Plesset energy denominator $[\Delta_{ab\cdots}^{ij\cdots} = \epsilon_i + \epsilon_j + \cdots - \epsilon_a - \epsilon_b - \cdots]$.[55] This parametrization is imposed via the DSRG flow equation, a nonlinear equation of the form

$$\bar{H}_{\mathrm{od}}(s) = \hat{R}(s) \tag{11}$$

where the *source operator* $\hat{R}(s)$ is Hermitian and continuous in s. The source operator in the DSRG flow equation drives the off-diagonal elements of $\bar{H}(s)$ to zero as $s \to \infty$. The operator $\hat{R}(s)$ contains both generalized excitation and de-excitation components, with the former defined in terms of the elements

$$r_{ab\cdots}^{ij\cdots}(s) = [\bar{H}_{ab\cdots}^{ij\cdots}(s) + t_{ab\cdots}^{ij\cdots}(s)\Delta_{ab\cdots}^{ij\cdots}]e^{-s(\Delta_{ab\cdots}^{ij\cdots})^2} \tag{12}$$

and the de-excitation components defined as $r_{ij\cdots}^{ab\cdots}(s) = [r_{ab\cdots}^{ij\cdots}(s)]^*$. Note that eq 11 should be interpreted as a set of many-body conditions,[26] which guarantees equal numbers of amplitudes and conditions, and avoids the multiparentage problem that affects many MRCC methods.[63−69]

In the MR-LDSRG(2) method, we retain up to two-body operators in the cluster operator [i.e., $\hat{A}(s) \approx \hat{A}_1(s) + \hat{A}_2(s)$] and the BCH series for $\bar{H}(s)$ [eq 2] is evaluated using the linearized approximation, keeping up to two-body normal-ordered operators in each commutator. The transformed Hamiltonian can then be evaluated by the following recursive equations

$$\begin{cases} \hat{C}^{(k+1)}(s) &= \dfrac{1}{k+1}[\hat{C}_{1,2}^{(k)}(s), \hat{A}(s)]_{0,1,2}, \\ \bar{H}^{(k+1)}(s) &= \bar{H}^{(k)}(s) + \hat{C}^{(k+1)}(s), \end{cases} \tag{13}$$

starting from $\hat{C}^{(0)}(s) = \bar{H}^{(0)}(s) = \hat{H}$ and iterating until the norm of $\hat{C}_{1,2}^{(k+1)}(s)$ is less than a given convergence threshold.

The solution of eq 11 yields a set of amplitudes $t_{ab\cdots}^{ij\cdots}(s)$ that define the operator $\hat{A}(s)$ and the DSRG transformed Hamiltonian $\bar{H}(s)$. From this latter quantity, the MR-DSRG electronic energy is computed as the expectation value of $\bar{H}$ with respect to the reference [via eq 10]. We refer the energy computed using this approach as the *unrelaxed* energy because the reference coefficients are not optimized. To include reference relaxation effects, we require that Ψ_0 is an eigenstate of $\bar{H}(s)$ within the space of reference determinants, a condition that is equivalent to solving the eigenvalue problem

$$\sum_{\mu=1}^{d}\langle\Phi_\nu|\bar{H}(s)|\Phi^\mu\rangle c_\mu' = E(s)c_\nu' \tag{14}$$

Equation 14 defines a new reference Ψ_0' with expansion coefficients c_μ'. This new reference may be used as a starting point for a subsequent MR-DSRG transformation, and this procedure can be repeated until convergence. The energy

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

obtained at the end of this iterative procedure is referred to as the *fully relaxed* energy. For the nonperturbative MR-DSRG schemes discussed in this work, we use the fully relaxed energy by default, unless otherwise noted.

2.2. Simplifying the MR-DSRG Equations: Sequential Transformation. Our first modification to improve the efficiency of the MR-DSRG consists in transforming the Hamiltonian via a sequence of unitary operators with increasing particle rank

$$\bar{H}(s) = e^{-\hat{A}_n(s)} \cdots e^{-\hat{A}_2(s)} e^{-\hat{A}_1(s)} \hat{H} e^{\hat{A}_1(s)} e^{\hat{A}_2(s)} \cdots e^{\hat{A}_n(s)} \tag{15}$$

We term the MR-DSRG approach based on eq 15 the *sequential* MR-DSRG (sq-MR-DSRG), while we refer to the original formalism based on eq 6 as the *traditional* MR-DSRG. Note that in the limit of $s \to \infty$ and no truncation of $\hat{A}(s)$, both the traditional and sequential MR-DSRG can approach the full configuration interaction limit.[70] However, these schemes are not equivalent when the operator $\hat{A}(s)$ is truncated to a given rank, like in the case of the MR-DSRG(2) approach.

An advantage of the sq-MR-DSRG approach is that $\hat{A}_1(s)$ can be exactly folded into the Hamiltonian via the transformation

$$\tilde{H}(s) = e^{-\hat{A}_1(s)} \hat{H} e^{\hat{A}_1(s)} \tag{16}$$

which preserves the particle rank of the bare Hamiltonian [eq 4]. The resulting $\hat{A}_1(s)$-dressed Hamiltonian [$\tilde{H}(s)$] can be evaluated efficiently [for details, see Section 3.1], especially if the two-electron integrals are approximated with DF or CD. The transformed Hamiltonian for the sq-MR-DSRG truncated to one- and two-body operators [sq-MR-DSRG(2)] is given by

$$\bar{H}(s) = e^{-\hat{A}_2(s)} \tilde{H}(s) e^{\hat{A}_2(s)} \tag{17}$$

In the linear approximation, the evaluation of eq 17 is simpler than in the traditional MR-LDSRG(2) since the total number of tensor contractions is reduced from 39 to 30.[57] As in the traditional MR-LDSRG(2), the sequentially transformed variant still requires the iterative update of both the $\hat{A}_1(s)$ and $\hat{A}_2(s)$ amplitudes (see ref 57 for details). In both the traditional and sequential approaches, the singles and double amplitudes are obtained by solving the same set of equations, except that $\bar{H}(s)$ is computed differently.

Another advantage of the sequential approach is that $\hat{A}_1(s)$ is treated exactly, while in the traditional scheme some contractions involving singles are neglected. To appreciate this point, consider all the contributions to the double-commutator term in the MR-LDSRG(2) that depend on $\hat{A}_1(s)$

$$\begin{aligned}[[\hat{H}, \hat{A}(s)]_{1,2}, \hat{A}(s)]_{0,1,2} \leftarrow\ & [[\hat{H}, \hat{A}_1(s)]_{1,2}, \hat{A}_1(s)]_{0,1,2} \\ & + [[\hat{H}, \hat{A}_1(s)]_{1,2}, \hat{A}_2(s)]_{0,1,2} \\ & + [[\hat{H}, \hat{A}_2(s)]_{1,2}, \hat{A}_1(s)]_{0,1,2}\end{aligned} \tag{18}$$

The first term on the r.h.s. of eq 18 is treated exactly in the MR-LDSRG(2). However, since contractions involving $\hat{A}_2(s)$ generate three-body terms (truncated in the linearized approximation), the contribution of $\hat{A}_1(s)$ in the second and third terms are not included exactly in the MR-LDSRG(2) transformed Hamiltonian. In the sequential approach, all contributions from $\hat{A}_1(s)$ are folded into the operator $\tilde{H}(s)$ and are treated exactly even in the linearized approximation. Therefore, the treatments of singles in the traditional and sequential versions of the MR-LDSRG(2) differ by terms of order three and higher in perturbation theory (assuming that both $\hat{A}_1(s)$ and $\hat{A}_2(s)$ are first-order quantities). Note that neither ansatz is in principle superior to the other as both are exact when the operator $\hat{A}(s)$ is not truncated, so that the different treatment of singles in the sequential theory should not be viewed as an approximation of the traditional scheme.

2.3. Alleviating the Memory Bottleneck: The Non-interacting Virtual Orbital (NIVO) Approximation. The bottleneck that arises in the evaluation of the BCH series for $\bar{H}(s)$ is the cost to store the intermediates generated during the evaluation of the commutator $[\hat{C}_{1,2}^{(k)}(s), \hat{A}(s)]_{0,1,2}$. The two-body components of this commutator have a storage cost of $O(N^4)$, a factor $N^2/N_{\mathbf{H}}^2$ higher than that of CCSD.

To reduce the cost to store $\bar{H}_2(s)$ and $\hat{C}_2^{(k)}(s)$, we propose to discard certain tensor blocks of these operators. By partitioning of orbitals into core (**C**), active (**A**) and virtual (**V**) spaces, each general 4-index tensor may be subdivided into 81 blocks according to the combination of orbital indices, for example **CCCC**, **AAVV**, **CAVA**, etc. In our approach, which we refer to as the noninteracting virtual orbital (NIVO) approximation, we neglect the operator components of $\hat{C}_2^{(k+1)}(s) = [\hat{C}_{1,2}^{(k)}(s), \hat{A}(s)]_2$, $k \geq 0$, with three or more virtual orbital indices (**VVVV**, **VCVV**, **VVVA**, etc.). Removing these blocks, the number of elements in each NIVO-approximated tensor is reduced from $O(N^4)$ to $O(N^2 N_{\mathbf{H}}^2)$, a size comparable to that of the $\hat{A}_2(s)$ tensor. For instance, in the cyclobutadiene computation discussed in Section 4.2, the memory requirements for storing $\bar{H}_2(s)$ or $\hat{C}_2^{(k)}(s)$ using a quintuple-ζ basis set are reduced from 2.7 TB to 6.8 GB by the NIVO approximation.

To justify the NIVO approximation, we analyze its effect on the energy. The first term in the BCH series that is approximated in the sq-MR-LDSRG(2)+NIVO scheme is the commutator $\hat{C}_2^{(1)}(s) = [\tilde{H}(s), \hat{A}_2(s)]_2$. Indicating the terms neglected by the NIVO approximation as $\delta\hat{C}_2^{(1)}(s)$, we see that the lowest-order energy error comes from the expectation value of the triple commutator term $[\delta\hat{C}_0^{(3)}(s)]$

$$\delta\hat{C}_0^{(3)}(s) = \frac{1}{6}[[\delta\hat{C}_2^{(1)}(s), \hat{A}_2(s)]_{1,2}, \hat{A}_2(s)]_0 \tag{19}$$

whose contributions are shown as diagrams in Figure 1. From a perturbation theory perspective, these diagrams are of order four [assuming $\hat{A}_2(s)$ to be of order one] and, therefore, are negligible compared to the leading error (third order) of the linearized commutator approximation. Since beyond the first commutator, the NIVO approximation affects core and virtual

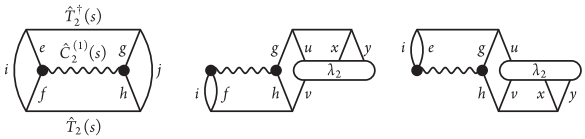


Figure 1. Diagrams that are neglected by the NIVO approximation in the evaluation of the term $\hat{C}_0^{(3)}(s)$ [eq 19]. The wiggly and horizontal solid lines indicate elements of $\hat{C}_2^{(1)}(s)$ and $\hat{A}_2(s)$, respectively. The two-body density cumulant is labeled with the symbol λ_2.

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

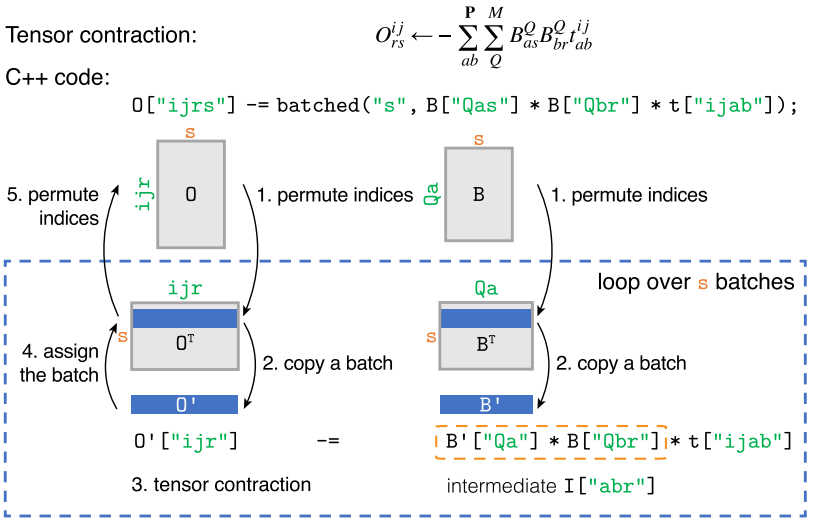


Figure 2. Batched algorithm implemented in the tensor library Ambit to compute the second term of eq 28. (1) The algorithm proceeds by first permuting tensor indices to rearrange the data with the same *s*-index continuously in memory. (2) Small batches of tensors are copied into temporary tensors (`B'["Qa"]` and `O'["ijr"]`). (3) Each batch is contracted, building appropriate intermediates (`I["abr"]`) to minimize the number of floating point operations. (4) The result of the tensor contraction is copied back to a temporary tensor (O^T`["sijr"]`). (5) After looping over all batches, the indices of the temporary tensor O^T are permuted back to the original order.

orbitals differently, it breaks hole-particle symmetry.[104] The most significant consequence of breaking this symmetry is the loss of exactness for two-electron systems, which, however, is already compromised by the linear commutator approximation made in the MR-LDSRG(2) approach. Therefore, in practical applications, the only major consequence of the NIVO scheme is to introduce small energy deviations from the MR-LDSRG(2) results. We also found that restoring hole-particle symmetry in the NIVO approximation by neglecting terms with three or more core orbital indices (**CCCC**, **CCCA**, etc.) leads to numerical instabilities and significantly larger errors in the correlation energy.

Hereafter, we shall append "+NIVO" at the end of the method name to indicate the use of NIVO approximation. For example, the density fitted MR-LDSRG(2) method in the sequential transformation ansatz with NIVO approximation is termed "DF-sq-MR-LDSRG(2)+NIVO".

3. IMPLEMENTATION

The sq-MR-LDSRG(2) method combined with DF and the NIVO approximation was implemented in Forte,[105] an open-source suite of multireference theories for molecular computations. This implementation reuses several components of our previous MR-LDSRG(2) code based on conventional four-index two-electron integrals.[57] The DSRG equations were implemented as tensor contractions using the Ambit tensor library.[106] The DF/CD integrals were obtained from the Psi4 package[107] using the MemDF interface.[108] In the following, we provide the details of our implementation of the sequential ansatz in combination with DF.[109]

3.1. Sequential Transformation. The $\hat{A}_1(s)$-dressed Hamiltonian [$\tilde{H}(s)$, eq 16] can be obtained by a unitary transformation of $\hat{H}$ via the operator $\hat{U}(s) = \exp[\hat{A}_1(s)]$. For clarity, we shall drop the label "(s)" for all *s*-dependent quantities [$\tilde{H}(s)$, $\hat{A}_1(s)$, and $\hat{U}(s)$] in this section. The one- and two-body components of $\tilde{H}$ ($\tilde{h}^{q'}_{p'}$ and $\tilde{v}^{r's'}_{p'q'}$) are given by

$$\tilde{h}^{q'}_{p'} = \sum_{pq} U^{q'}_{q} h^{q}_{p} U^{p}_{p'} \tag{20}$$

$$\tilde{v}^{r's'}_{p'q'} = \sum_{pqrs} U^{r'}_{r} U^{s'}_{s} v^{rs}_{pq} U^{p}_{p'} U^{q}_{q'} \tag{21}$$

Here, the unitary matrix $U^{p'}_{p} = (\mathbb{U})_{p'p}$ and its inverse $U^{p}_{p'} = (\mathbb{U})^{*}_{p'p}$ are given by $\mathbb{U} = e^{\mathbb{A}}$, where the matrix $\mathbb{A}$ is composed of elements of the $\hat{A}_1$ tensor, $(\mathbb{A})_{\mathrm{ia}} = t^{i}_{a}$ and $(\mathbb{A})_{\mathrm{ai}} = -t^{i}_{a}$. Note that we use primed indices only as a way to distinguish labels, yet these indices by no means imply a new set of orbitals.

The $\hat{A}_1$-dressed Hamiltonian written in normal ordered form with respect to Ψ_0 is given by

$$\tilde{H} = \tilde{E}_0 + \sum_{pq} \tilde{f}^{q}_{p} \{\hat{a}^{p}_{q}\} + \frac{1}{4} \sum_{pqrs} \tilde{v}^{rs}_{pq} \{\hat{a}^{pq}_{rs}\} \tag{22}$$

where the transformed energy ($\tilde{E}_0$) is given by

$$\tilde{E}_0 = \sum_{i'j'}^{\mathbf{H}} \tilde{h}^{j'}_{i'} \gamma^{i'}_{j'} + \frac{1}{4} \sum_{i'j'k'l'}^{\mathbf{H}} \tilde{v}^{k'l'}_{i'j'} \gamma^{i'j'}_{k'l'} \tag{23}$$

and the Fock matrix elements ($\tilde{f}^{q'}_{p'}$) are defined as

$$\tilde{f}^{q'}_{p'} = \tilde{h}^{q'}_{p'} + \sum_{i'j'}^{\mathbf{H}} \tilde{v}^{q'j'}_{p'i'} \gamma^{i'}_{j'} \tag{24}$$

Note that the quantities $\gamma^{i'}_{j'}$ and $\gamma^{i'j'}_{k'l'}$ in eqs 23 and 24 are the *untransformed* 1- and 2-RDMs of the reference $|\Psi_0\rangle$ defined as $\gamma^{p}_{q} = \langle\Psi_0|\hat{a}^{p}_{q}|\Psi_0\rangle$ and $\gamma^{pq}_{rs} = \langle\Psi_0|\hat{a}^{pq}_{rs}|\Psi_0\rangle$, respectively.

The two-electron integral transformation [eq 21] has a noticeable cost [$O(N^5)$] and must be repeated each time the $\hat{A}_1$ operator is updated. However, in the implementation based on DF integrals, this transformation may be performed in a significantly more efficient way. In DF, the four-index electron

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

Table 1. Computational Scaling of the Steps Involved in the Evaluation of the DSRG Transformed Hamiltonian ($\bar{H}$)[a]

	time	computation	original	DF	NIVO
traditional	$t_1^{(0)}$	$\hat{H}^{(1)} \leftarrow [\hat{C}^{(0)}, \hat{T}_1]$	$O(N^3N_{\mathbf{H}}N_{\mathbf{P}})$	unchanged	reduced to $O(N^2N_{\mathbf{H}}^2N_{\mathbf{P}})$
	$t_1^{(k)}$, $k \geq 1$	$\hat{O}^{(k+1)} \leftarrow \frac{1}{k+1}[\hat{C}^{(k)}, \hat{T}_1]$	$O(N^3N_{\mathbf{H}}N_{\mathbf{P}})$	unchanged	reduced to $O(\mathrm{NN}_{\mathbf{H}}^3N_{\mathbf{P}})$
sequential	t_1	$\tilde{H} = e^{-\hat{A}_1}\hat{H}e^{\hat{A}_1}$	$O(N^5)$	reduced to $O(N^3M)$	unchanged
traditional/sequential	$t_2^{(0)}$	$\hat{O}^{(1)} \leftarrow [\hat{C}^{(0)}, \hat{T}_2]$	$O(N^2N_{\mathbf{H}}^2N_{\mathbf{P}}^2)$	unchanged	reduced to $O(N_{\mathbf{H}}^2N_{\mathbf{P}}^4)$
	$t_2^{(k)}$, $k \geq 1$	$\hat{O}^{(k+1)} \leftarrow \frac{1}{k+1}[\hat{C}^{(k)}, \hat{T}_2]$	$O(N^2N_{\mathbf{H}}^2N_{\mathbf{P}}^2)$	unchanged	reduced to $O(N_{\mathbf{H}}^4N_{\mathbf{P}}^2)$
	$t_{\mathrm{misc}}^{(k)}$	$\begin{cases}\hat{C}^{(k+1)} = \hat{O}^{(k+1)} + [\hat{O}^{(k+1)}]^\dagger \\ \bar{H}^{(k+1)} = \bar{H}^{(k)} + \hat{C}^{(k+1)}\end{cases}$	$O(N^4)$	unchanged	reduced to $O(N^2N_{\mathbf{H}}^2)$

[a]For each step, we report the computational scaling of the original MR-LDSRG(2) and asymptotic scaling changes when density fitting (DF) or the NIVO approximation are applied.

repulsion integral tensor is implicitly a contraction of a three-index auxiliary tensor (B_{pq}^{Q})

$$\langle pq|rs\rangle \approx \sum_{\mathrm{Q}}^{M} B_{pr}^{\mathrm{Q}}B_{qs}^{\mathrm{Q}} \tag{25}$$

where M is the dimension of the fitting basis in DF. Using this decomposition, the unitary transformation may be directly applied to each auxiliary tensor

$$\tilde{B}_{p'q'}^{\mathrm{Q}} = \sum_{pq} B_{pq}^{\mathrm{Q}}U_{p'}^{p}U_{q'}^{q} \tag{26}$$

reducing the cost to evaluate $\tilde{H}$ down to $O(N^3M)$.

Equations 22 and 23 specify the procedures to obtain $\tilde{H}$ as a unitary transformation of $\hat{H}$. Since $\tilde{H}$ retains the structure of $\hat{H}$, we can reuse most of our previous MR-LDSRG(2) code[57] to implement sq-MR-LDSRG(2) by using $\tilde{h}_{p'}^{q'}$ and $\tilde{v}_{p'q'}^{r's'}$ in place of the bare integrals and removing contractions involving $\hat{A}_1$.

3.2. Batched Tensor Contraction for the DF Algorithm. Despite the storage cost reduction of the DF and NIVO approximations, another potential memory bottleneck is the size of intermediates generated during tensor contractions. For example, consider the following contraction

$$O_{rs}^{ij} \leftarrow \sum_{ab}^{\mathbf{P}} \langle ab\|rs\rangle t_{ab}^{ij} \;\forall\; i, j \in \mathbf{H}, \forall\; r, s \in \mathbf{G} \tag{27}$$

which is also found in the CCSD equations. In the DF case, eq 27 is written as two contractions involving auxiliary tensors

$$O_{rs}^{ij} \leftarrow \sum_{ab}^{\mathbf{P}}\sum_{\mathrm{Q}}^{M} B_{ar}^{\mathrm{Q}}B_{bs}^{\mathrm{Q}}t_{ab}^{ij} - \sum_{ab}^{\mathbf{P}}\sum_{\mathrm{Q}}^{M} B_{as}^{\mathrm{Q}}B_{br}^{\mathrm{Q}}t_{ab}^{ij} \tag{28}$$

The most efficient way to evaluate the second term of eq 28 is to introduce the intermediate tensor $I_{asbr} = \sum_{\mathrm{Q}}^{M} B_{as}^{\mathrm{Q}}B_{br}^{\mathrm{Q}}$ of size $O(N^2N_{\mathbf{P}}^2)$. To avoid storage of this large intermediate, it is common to evaluate eq 28 using a batched algorithm like the one reported in Figure 2, whereby a slice of the tensor I_{asbr} is computed and contracted on the fly with the amplitudes t_{ab}^{ij}. To automate this optimization of the tensor contraction we have coded a generic batching algorithm in the tensor library Ambit.[106] Whereas the Ambit code for the second term in eq 28 is written as

```
O["ijrs"] -= B["Qas"] * B["Qbr"] * t["ijab"];
```

our new implementation allows batching over the index s by simply surrounding the contraction with the `batched()` function decorator

```
O["ijrs"] -= batched("s", B["Qas"] * B["Qbr"] * t["ijab"]);
```

This command is executed by Ambit using the same procedure reported in Figure 2 and only generates intermediate tensors of size $O(NN_{\mathbf{P}}^2)$. Our implementation is general and allows batching over multiple indices (e.g., rs).

3.3. Computational Cost Reduction. Here we discuss timings for all the MR-LDSRG(2) variants introduced in this work. In MR-LDSRG(2) theory, the main computational bottleneck is forming the transformed Hamiltonian $\bar{H}$. The total timing (t_{tot}) for n evaluations of this quantity is partitioned according to

$$t_{\mathrm{tot}} = t_1 + t_2 + t_{\mathrm{misc}} \tag{29}$$

where t_1 and t_2 are the timings to evaluate the commutators involving $\hat{T}_1$ and $\hat{T}_2$, respectively. In the sequential transformation approach, t_1 is instead defined as the time for forming the $\hat{A}_1$-transformed Hamiltonian. The term t_{misc} accounts for the cost to sort and accumulate the results of contractions with $\hat{T}_1$ and $\hat{T}_2$. In Table 1 we report the computational scaling of the various terms that contribute to t_1, t_2, and t_{misc} for the original MR-LDSRG(2) and all the variants considered in this work.

To illustrate the performance of the various schemes developed in this work, we report timings for computing $\bar{H}$ in the case of cyclobutadiene in Figure 3 assuming that all tensors can be stored in random-access memory (see Section 4.2 and the Supporting Information for details). Figure 3 shows that the timing for the traditional MR-LDSRG(2) is dominated by contractions involving $\hat{T}_1$ and $\hat{T}_2$. The cost of the singles contractions can be reduced significantly (3−5 times) by employing the sequentially transformed approach, even though at each iteration of the sq-MR-LDSRG(2) equations it is necessary to build the operator $\tilde{H}$.

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

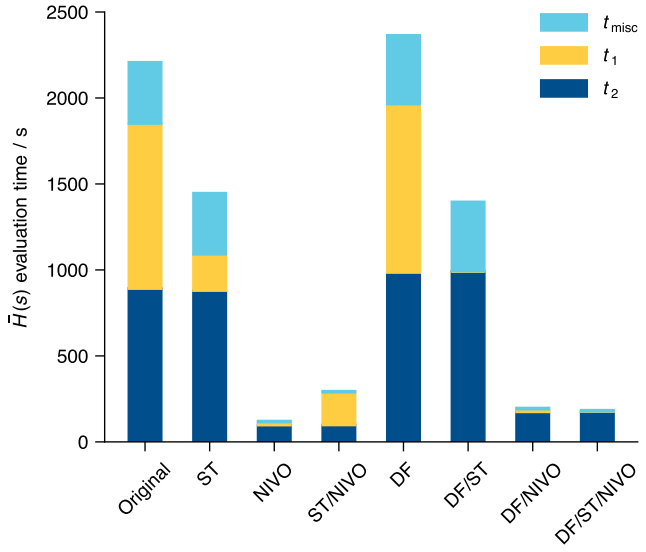


Figure 3. Time to evaluate the DSRG transformed Hamiltonian $[\bar{H}]$ of the ground-state cyclobutadiene when different techniques are introduced in the MR-LDSRG(2) method. These techniques include: density fitting (DF), sequential transformation (ST), and the noninteracting virtual orbital (NIVO) approximation. The total time for computing $[\hat{C}^{(k)}, \hat{T}_1]$ in MR-LDSRG(2) or $\tilde{H} = e^{-\hat{A}_1}\hat{H}e^{\hat{A}_1}$ in sq-MR-LDSRG(2) is labeled as t_1 in this plot. All computations employed the cc-pVTZ basis set and were carried out on an Intel Xeon E5−2650 v2 processor using eight threads.

Applying the NIVO approximation to the original MR-LDSRG(2) leads to a drastic reduction of the total computational time (×18 speedup). This reduction in timing is due to several contributing factors. First, the evaluation of the $\hat{T}_1$ contractions in NIVO is sped up by a factor of $O(nN/N_{\mathbf{H}})$, where n is the number of commutators included in the BCH series. Second, the contributions due to doubles, $[\hat{C}^{(k)}, \hat{T}_2]$, have identical scaling for the first commutator, but for $k \geq 1$, they can be evaluated with a speedup of a factor of $O(N^2/N_{\mathbf{H}}^2)$. Third, the NIVO approximation also reduces t_{misc} significantly because the cost of tensor transpose and accumulation operations are reduced from $O(N^4)$ to $O(N^2N_{\mathbf{H}}^2)$. For large $N/N_{\mathbf{H}}$ ratios, the cost to compute $\bar{H}$ in the NIVO approximation is dominated by the commutator $[\hat{H}, \hat{T}_2]$, with scaling identical to that of CCSD. For comparison, the similarity-transformed Hamiltonian can be evaluated in 24 s with Crawford's CCSD code in Psi4, in 121 s with our MR-LDSRG(2)+NIVO code, and 2208 s with the original MR-LDSRG(2) code (in both cases employing an unrestricted implementation and C_1 symmetry).

For methods combined with DF, we observe an increase in t_2 due to the extra cost to build two-body intermediates from the auxiliary tensors. The traditional and sequential approaches using the DF/NIVO approximations have similar costs, with the latter being slightly faster due to the efficient transformation of the auxiliary tensors $[\tilde{B}$, eq 26] afforded by the DF approximation. For this example, the DF-sq-MR-LDSRG(2)+NIVO computation ran 12 times faster than the one using the original approach. As we will demonstrate in the next section, this method is as accurate as the MR-LDSRG(2), and therefore, it is our recommend choice for large-scale multireference computations.

For our last timing test we study the scaling for a system in which the number of active orbitals is proportional to the number of electrons. To this end we consider a polyacetylene chain containing m carbon atoms (C_mH_{m+2}) with m ranging from 2 to 16 and employ a CAS(me, mo) reference wave function. In the MR-DSRG(2), the term with the highest degree in the active orbitals scales as $O(N_{\mathbf{A}}^6N_{\mathbf{V}})$. This contribution, however, becomes important only when $N_{\mathbf{A}} >$ 30. For smaller active space, the MR-LDSRG(2) equations contain a term proportional to $N^2N_{\mathbf{H}}^2N_{\mathbf{P}}^2$, which in our example is expected to be proportional to m^6. In Figure 4, we show the

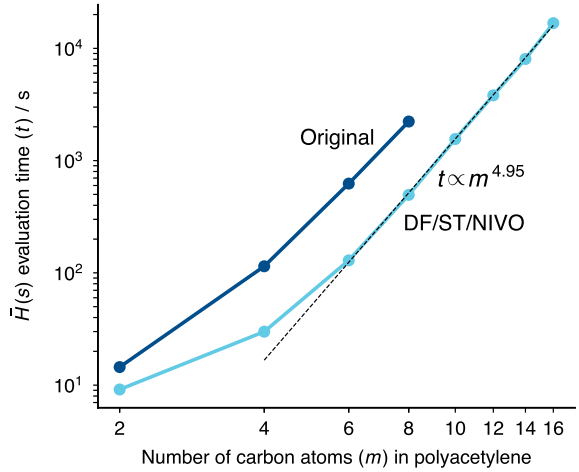


Figure 4. Log−log plot of the time to evaluate the DSRG transformed Hamiltonian $[\bar{H}]$ vs the number of carbon atoms in computations of the ground-state of polyacetylenes (C_mH_{m+2}, m = 2, 4, 6, ···, 16) for the original MR-LDSRG(2) and the DF-sq-MR-LDSRG(2)+NIVO methods. All computations employed the cc-pVDZ basis set, and they were carried out on an Intel Xeon E5−2650 v2 processor using eight threads. The largest computation (m = 16) corresponds to 298 correlated orbitals.

timing for the original and approximate MR-LDSRG(2) and illustrate the asymptotic computational cost, which is found to be close to m^5. Although the techniques introduced in this paper do not change the overall scaling with respect to the original MR-LDSRG(2) method, they significantly reduce the computational time and enable computations on systems with large active spaces. In the polyacetylene system, the original algorithm cannot be applied beyond octatetraene (m = 8) on a single node. In contrast, the DF-sq-MR-LDSRG(2)+NIVO method is four times faster and may be applied to systems twice as large. In addition, we also tested the dependence of the computational time as a function of the number of active orbitals for a fixed number of electrons and orbitals. For example, in computations on the $C_{16}H_{18}$ molecule, increasing the number of active orbitals from 2 to 16 increases the cost by a factor 1.6 (see Figure S1 in the Supporting Information for a plot of the timing for this example). Although restrictions imposed by the use of a CAS reference limit the polyacetylene computations to about m = 16, we estimate that when combined with approximated CAS diagonalization methods, the MR-LDSRG(2) method will be applicable to problems with 25−30 active orbitals. For the large-basis computations on cyclobutadiene reported in Section 4.2 (580 correlated orbitals), we find that active spaces as large as the full valence CAS(20e, 20o) are in principle attainable on a single node

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

since the size of each NIVO-approximated two-body tensor is 18.5 GB. However, in our current implementation, batching of tensor contractions is applied only to intermediate tensors that scale as $O(N^4)$, but not those generating intermediates that scale as $O(N_A N^3)$. Consequently, computations with around 600 correlated orbitals are limited to about 16−18 active orbitals.

4. RESULTS AND DISCUSSION

4.1. First Row Diatomic Molecules. We first benchmark the effect of DF and the NIVO approximation on the traditional and sequential versions of the MR-LDSRG(2). Our test set consists of eight diatomic molecules: BH, HF, LiF, BeO, CO, C_2, N_2, and F_2. Specifically, we computed equilibrium bond lengths (r_e), harmonic vibrational frequencies (ω_e), anharmonicity constants ($\omega_e x_e$), and dissociation energies (D_0) and compare those to experimental data taken from ref 110. The dissociation energy D_0 includes zero-point vibrational energy corrections that account for anharmonicity effects and is computed as $D_0 = D_e - \omega_e/2 + \omega_e x_e/4$ (in a.u.), where D_e is the dissociation energy with respect to the bottom of the potential.[111] Since our current implementation of the MR-DSRG cannot handle half-integer spin states, the energies of the atoms Li, B, C, N, O, and F were computed as half of the energy of the stretched homonuclear diatomic molecule at a distance of 10 Å. All spectroscopic constants were obtained via a polynomial fit of the energy using nine equally spaced points centered around the equilibrium bond length and separated by 0.2 Å. For all eight molecules, we adopted a full-valence active space where the 1s orbital of hydrogen, and the 2s and 2p orbitals of first-row elements are considered as active orbitals. No orbitals were frozen in the CASSCF optimization procedure. The flow parameter for all DSRG computations was set to $s = 0.5\ E_h^{-2}$, as suggested by our previous work.[56] All computations utilized the cc-pVQZ basis set[112] and 1s-like orbitals of the first-row elements were frozen in the CC and MR-DSRG treatments of electron correlation. In DF computations, we employed a combination of different auxiliary basis sets. For CASSCF, the cc-pVQZ-JKFIT auxiliary basis set[83] was used for H, B, C, N, O, and F, and the def2-QZVPP-JKFIT basis set[113] was used for Li and Be. In the DSRG computations, the cc-pVQZ-RI basis set[114] was used for all atoms.

Figure 5 and Table 2 report a comparison of second- and third-order DSRG multireference perturbation theory (DSRG-MRPT2/3),[56,58] the original MR-LDSRG(2), DF-sq-LDSRG(2)+NIVO, CCSD, and CCSD(T). The mean absolute error (MAE) and standard deviation (SD) reported in Table 2 show that MR-LDSRG(2) method is as accurate as CCSD(T) in predicting r_e, $\omega_e x_e$ and D_0, while it predicts ω_e values with accuracy in between CCSD and CCSD(T).

The fact that the MR-LDSRG(2) results are more accurate than those from CCSD suggests that the full-valence treatment employed in the former improves the treatment of electron correlation. It is also rewarding to see that in many instances, the MR-LDSRG(2) has an accuracy similar to that of CCSD(T), despite the fact that the former does not include triples corrections.

To analyze the impact of each modification to the original MR-LDSRG(2) method, in Table 2 we report statistics computed for all variants of the MR-LDSRG(2) considered here. The use of a sequential ansatz has a modest effect on all

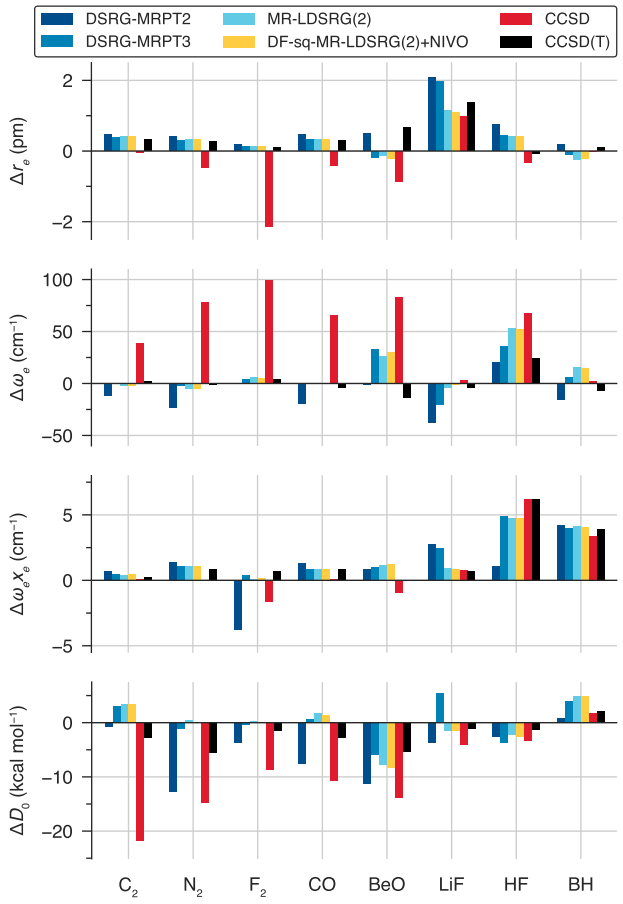


Figure 5. Comparison of second- and third-order MR-DSRG perturbation theory (DSRG-MRPT2, DSRG-MRPT3), MR-LDSRG(2), DF-sq-MR-LDSRG(2)+NIVO, and single reference coupled cluster methods on a test set composed of eight diatomic molecules. Deviations of equilibrium bond lengths (r_e), harmonic vibrational frequencies (ω_e), anharmonicity constants ($\omega_e x_e$), and dissociation energies (D_0) with respect to experimental values.[110] All results were computed with the cc-pVQZ basis. Core orbitals were frozen in the MR-DSRG and coupled cluster computations.

properties and the MAEs with respect experimental results are nearly unchanged, if not slightly improved. The DF and NIVO approximations have an effect on molecular properties that is comparable in magnitude and smaller than the deviation introduced by the sequential ansatz. When these three approximations are combined together, the resulting method shows errors with respect to experimental values that are nearly identical to those from the traditional MR-LDSRG(2). The only noticeable deviation is found for ω_e (MAE 13.9 vs 14.3 cm^{-1}).

In this study, we also investigate the effect of combining the DF-sq-LDSRG(2)+NIVO method with truncation of the BCH expansion, that is, terminating $\bar{H}(s) = \tilde{H}(s) + \sum_{k=1}^{n} \frac{1}{k!}\hat{C}^{(k)}(s)$ at a given integer n. The recursive evaluation of $\bar{H}(s)$ via eq 13 usually requires 10−12 terms to converge. In Table 2 we report statistics computed by approximating the BCH expansion with n up to 2, 3, and 4 commutators. The use of only two commutator introduces noticeable deviations for ω_e and D_0: the MAE with respect to the results obtained using the full BCH series increases by 3.8 cm^{-1} and 0.6 kcal mol^{-1}, respectively. The inclusion of the triply nested commutator significantly reduces these deviations to only 0.7 cm^{-1} and 0.2

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

Table 2. Error Statistics for the Equilibrium Bond Lengths (r_e, in pm), Harmonic Vibrational Frequencies (ω_e, in cm^{-1}), Anharmonicity Constants ($\omega_e x_e$, in cm^{-1}), and Dissociation Energies (D_0, in kcal mol^{-1}) of the Eight Diatomic Molecules Computed Using Various MR-DSRG Schemes[a]

		DSRG-MRPT		MR-LDSRG(2)			sq-MR-LDSRG(2)			DF-sq-MR-LDSRG(2)+NIVO				
		PT2	PT3	conv.[b]	DF	DF+NIVO	conv.[b]	DF	DF+NIVO	comm(2)	comm(3)	comm(4)	CCSD	CCSD(T)
r_e	mean	0.63	0.41	0.30	0.30	0.30	0.28	0.28	0.28	0.28	0.28	0.28	−0.41	0.38
	MAE	0.63	0.48	0.39	0.39	0.38	0.39	0.39	0.39	0.40	0.38	0.39	0.65	0.40
	SD	0.85	0.74	0.49	0.50	0.49	0.48	0.48	0.48	0.49	0.48	0.48	0.92	0.57
	max	2.07	1.96	1.15	1.15	1.15	1.10	1.10	1.10	1.11	1.10	1.09	2.13	1.36
ω_e	mean	−11.2	7.2	11.2	11.1	11.1	11.8	11.9	11.8	13.0	11.9	11.9	54.9	−0.1
	MAE	16.5	13.1	14.3	14.2	14.3	13.8	13.9	13.9	16.0	13.8	13.9	54.9	7.9
	SD	20.0	19.0	22.0	21.9	22.0	21.9	22.1	22.1	26.3	21.6	22.2	64.5	10.7
	max	38.3	36.1	53.2	52.8	53.7	51.2	51.8	51.9	64.5	50.3	52.3	99.3	24.5
$\omega_e x_e$	mean	1.1	1.9	1.7	1.5	1.7	1.7	1.8	1.7	1.9	1.6	1.7	1.0	1.7
	MAE	2.0	1.9	1.7	1.5	1.7	1.7	1.8	1.7	1.9	1.6	1.7	1.6	1.7
	SD	2.4	2.5	2.3	2.1	2.3	2.3	2.6	2.3	2.6	2.2	2.3	2.6	2.6
	max	4.2	4.9	4.7	4.1	4.8	4.8	5.7	4.7	5.9	4.4	4.8	6.2	6.2
D_0	mean	−5.2	0.2	−0.1	−0.1	−0.0	−0.3	−0.3	−0.3	−0.7	−0.3	−0.3	−9.5	−2.4
	MAE	5.4	3.0	2.8	2.8	2.6	2.9	2.9	2.8	3.3	2.9	2.8	9.9	2.9
	SD	6.9	3.6	3.7	3.7	3.5	3.9	3.8	3.8	4.1	3.9	3.8	11.8	3.3
	max	12.7	6.0	7.9	7.8	7.5	8.4	8.4	8.3	8.6	8.4	8.2	21.9	5.6

[a]All results were obtained using the cc-pVQZ basis and core orbitals were frozen in the MR-DSRG and CC computations. The statistical indices are mean signed error (mean), mean absolute error (MAE), standard deviation (SD), and maximum absolute error (max). [b]Computed using conventional four-index two-electron integrals.

kcal mol^{-1}, respectively. The 4-fold commutator term yields r_e, ω_e, and $\omega_e x_e$ that are nearly identical to those from the untruncated BCH series.

In practice, the cost savings of truncating the BCH series when using the NIVO approximations are small because contributions past the first commutator are inexpensive to compute. However, our results suggest that an alternative way to approximate the BCH series could be to keep the first four commutators, evaluating their contribution exactly.

4.2. Cyclobutadiene. Next, we consider the automerization reaction of cyclobutadiene (CBD, C_4H_4). We study the energy difference between the rectangular (D_{2h}) energy minimum and the square transition state (D_{4h}).[115] This reaction is a challenging chemistry problem for both experiment and theory.[115−127] Due to its instability, there are no direct measurements of the reaction barrier, and experiments performed on substituted cyclobutadienes suggest the barrier height falls in the range of 1.6−10 kcal mol^{-1}.[116] In this work, we optimize the equilibrium and transition state geometries using finite differences of energies to compute the barrier height. Specifically, we compare both DF-MR-LDSRG(2)+NIVO and DF-sq-MR-LDSRG(2)+NIVO optimized geometries to those obtained from the second-order complete active space perturbation theory (CASPT2),[4] the partially contracted second-order *n*-electron valence state perturbation theory (pc-NEVPT2),[6] and the internally contracted MRCI with singles and doubles plus the relaxed Davidson correction (ic-MRCISD+Q)[9,128] as implemented in Molpro, and the state-specific MRCC of Mukherjee and coworkers (Mk-MRCC)[21,129] as implemented in Psi4.[130]

To reduce the computational cost, all MR-DSRG calculations performed two steps of the reference relaxation procedure discussed in Section 2.1. A comparison of this procedure with full reference relaxation using the cc-pVDZ basis set shows errors of ca. 0.01 kcal mol^{-1} for absolute energies, 0.0001 Å for bond lengths, and 0.001° for bond angles. The Mk-MRCC implementation used in this work neglects effective Hamiltonian couplings between reference determinants that differ by three or more spin orbitals and therefore yield approximate results when applied to the CAS(4e,4o) reference considered here. To guarantee convergence, we applied a Tikhonov regularization denominator shift[131] of 1 mE_h in all Mk-MRCC calculations. All computations utilized the cc-pV*X*Z (*X* = D, T, Q, 5) basis sets,[112] and the corresponding cc-pV*X*Z-JKFIT[83] and cc-pV*X*Z-RI[114] auxiliary basis sets for DF-CASSCF and DF-DSRG computations, respectively. The 1s core electrons of carbon atoms were frozen in all post-CASSCF methods. All results were computed using semicanonical CASSCF orbitals.

Preliminary computations using the cc-pVDZ basis using the CAS(2e,2o) and CAS(4e,4o) active spaces revealed an interesting aspect of this system. As shown in Figure 6, the

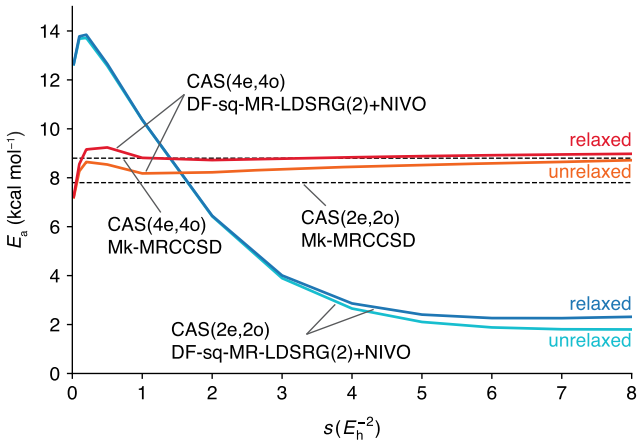


Figure 6. Automerization barrier (E_a) of cyclobutadiene computed using the DF-sq-MR-LDSRG(2)+NIVO theory with varying flow parameters. Results were obtained using the cc-pVDZ basis set. We also applied a Tikhonov regularization denominator shift[131] of 1 mE_h in all Mk-MRCC calculations to guarantee convergence.

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

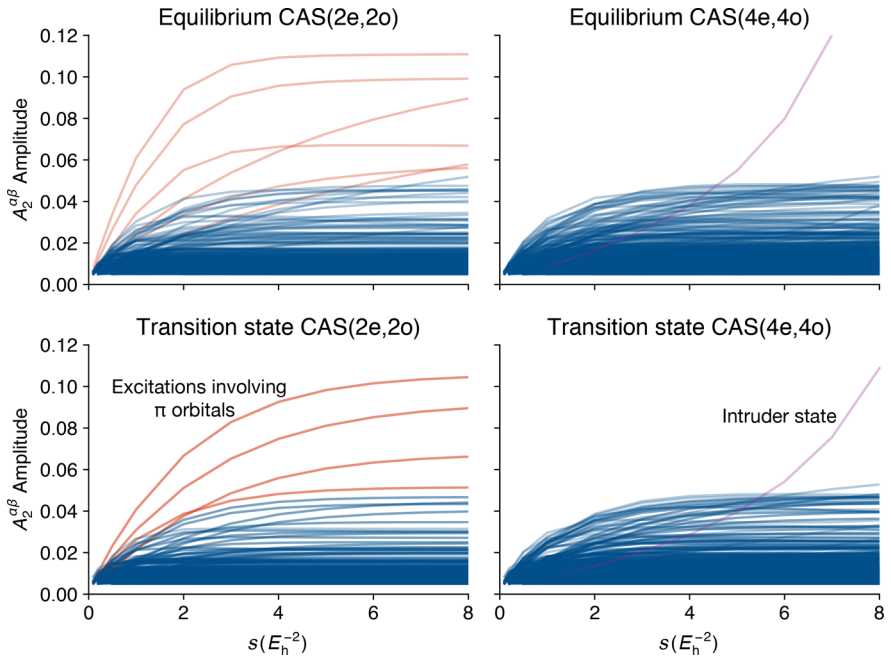


Figure 7. Value of the DF-sq-MR-LDSRG(2)+NIVO double substitution amplitudes involving both alpha and beta electrons ($A_2^{\alpha\beta}$) as a function of the flow parameter s for the rectangular equilibrium and the square transition state of cyclobutadiene. All results were computed with the cc-pVDZ basis set and do not include reference relaxation effects.

s-dependency of the automerization barrier displays a significantly different behavior for these two active spaces. In both cases, the predicted activation energies change significantly for small values of s (<0.2 E_h^{-2}), a normal trend observed for all DSRG computations and due to the increased recovery of dynamical correlation energy. Interestingly, while the CAS(4e,4o) curve flattens out for larger values of s, the CAS(2e,2o) curve shows a significant s-dependence in the range $s \in [0.5, 8]$ E_h^{-2}.

To understand the origin of this difference, we analyze the double substitution amplitudes $[t_{ab}^{ij}(s)]$ as a function of s for both the equilibrium and transition state geometries, as shown in Figure 7. In the CAS(2e,2o) case, we notice some abnormally large amplitudes (indicated in red), some of which are as large as 0.1. These amplitudes correspond to excitations within the four π orbitals of CBD, and their large value suggests that the CAS(2e,2o) space is insufficient to capture all static correlation effects in CBD. The offending amplitudes converge at different rates as s increases and introduce a strong s-dependence in the energy barrier. Note also that in the limit of $s \to \infty$ there is a significant difference in the barrier for the CAS(2e,2o) and CAS(4e,4o) spaces.

In contrast, in the CAS(4e,4o) computations all excitations within the π orbitals are included in the active space and the resulting DSRG amplitudes have absolute values less than 0.05. Diverging amplitudes in computations with CAS(4e,4o) reference wave functions, corresponding to intruder states, can also be seen in Figure 7. Our results reported in Table 3 are all based on the flow parameter value $s = 1.0$ E_h^{-2}, which is significantly far from the region ($s > 5.0$ E_h^{-2}) where amplitudes begin to diverge, and at the same time leads to well-converged absolute energies (see Figure S4 in the Supporting Information). We also performed computations using $s = 0.5$ E_h^{-2} to verify that the automerization energies computed with different values of the flow parameter are consistent. In general, the difference in automerization energies computed with $s =$ 0.5 and 1.0 E_h^{-2} is of the order of 0.6−0.7 kcal mol^{-1}. Note that intruder states are also encountered in Mk-MRCCSD computations based on the CAS(4e,4o) reference and lead to convergence issues that could be avoided only via Tikhonov regularization.

Geometric parameters for the optimized structures and energy barriers of CBD computed with the CAS(4e,4o) reference are reported in Table 3. A comparison of the CASSCF and correlated results shows that dynamical correlation is important in this system as it increases the energy barrier by about 1−3 kcal mol^{-1}. Our best estimate for the automerization barrier of CBD is 10.3 kcal mol^{-1} at the DF-sq-MR-LDSRG(2)+NIVO/cc-pV5Z level of theory. This value is likely to be slightly higher than the exact results since in the Mk-MRCC results perturbative triples corrections contribute to lowering the barrier by ca. 1.5 kcal mol^{-1}. In general, the MR-LDSRG(2) results are between those of Mk-MRCCSD and Mk-MRCCSD(T), reinforcing the same observation we made in the benchmark of diatomic molecules. There is also good agreement between the MR-LDSRG(2) results and the ic-MRCISD+Q and pc-NEVPT2 barriers (ca. 9−9.5 kcal mol^{-1}), while CASPT2 results consistently underestimate the barrier height by about 5 kcal mol^{-1}. For instance, at the D_{4h} geometry the MR-LDSRG(2) C−C bond length is 1.4446 Å, which is almost midway between the Mk-MRCCSD (1.4406 Å) and Mk-MRCCSD(T) (1.4480 Å) values. As expected, the differences between the traditional and sequentially transformed MR-DSRG(2) results are negligible. Our results in cc-pVDZ and cc-pVTZ bases also agree well with other reported theoretical values, especially those computed with multireference methods,[115,118,121−126] and the experimental range reported in ref 116 (see Supporting Information for a list of theoretical predictions). Using our new implementation, we were able to perform, for the first time, nonperturbative multireference computations on cyclo-

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

Table 3. Automerization Reaction Barrier (E_a, in kcal mol^{-1}) and Geometric Parameters (Bond Lengths in Å, Bond Angles in Degrees) of Cyclobutadiene Computed with Various Multireference Methods[a]

method	E_a	D_{2h} C−C[b]	D_{2h} C=C[b]	D_{2h} C−H	D_{2h} ∠C−C−H[c]	D_{4h} C−C	D_{4h} C−H
		cc-pVDZ (n_V = 60)					
CASSCF	6.49	1.5502	1.3567	1.0790	134.87	1.4472	1.0779
CASPT2	3.48	1.5603	1.3712	1.0916	134.94	1.4578	1.0910
pc-NEVPT2	8.33	1.5710	1.3609	1.0920	134.91	1.4566	1.0906
ic-MRCISD+Q	7.66	1.5732	1.3655	1.0924	134.92	1.4610	1.0914
Mk-MRCCSD	8.80	1.5733	1.3623	1.0931	134.91	1.4585	1.0920
Mk-MRCCSD(T)	7.56	1.5772	1.3699	1.0951	134.92	1.4652	1.0941
DF-MR-LDSRG(2)+NIVO (s = 0.5)	9.05	1.5710	1.3636	1.0923	134.90	1.4587	1.0908
DF-sq-MR-LDSRG(2)+NIVO (s = 0.5)	9.09	1.5709	1.3635	1.0922	134.90	1.4587	1.0907
DF-MR-LDSRG(2)+NIVO (s = 1.0)	8.56	1.5769	1.3660	1.0945	134.92	1.4624	1.0932
DF-sq-MR-LDSRG(2)+NIVO (s = 1.0)	8.62	1.5768	1.3659	1.0944	134.92	1.4623	1.0931
		cc-pVTZ (n_V = 160)					
CASSCF	7.44	1.5475	1.3471	1.0694	134.83	1.4409	1.0683
CASPT2	3.87	1.5521	1.3552	1.0771	134.94	1.4451	1.0765
pc-NEVPT2	9.34	1.5617	1.3454	1.0774	134.91	1.4437	1.0760
ic-MRCISD+Q	8.93	1.5624	1.3480	1.0766	134.90	1.4464	1.0755
Mk-MRCCSD	10.09	1.5628	1.3452	1.0775	134.89	1.4442	1.0764
Mk-MRCCSD(T)	8.56	1.5671	1.3535	1.0797	134.90	1.4515	1.0786
DF-MR-LDSRG(2)+NIVO (s = 0.5)	10.57	1.5611	1.3465	1.0767	134.89	1.4447	1.0751
DF-sq-MR-LDSRG(2)+NIVO (s = 0.5)	10.60	1.5608	1.3463	1.0766	134.89	1.4446	1.0750
DF-MR-LDSRG(2)+NIVO (s = 1.0)	9.87	1.5668	1.3488	1.0789	134.91	1.4483	1.0775
DF-sq-MR-LDSRG(2)+NIVO (s = 1.0)	9.93	1.5666	1.3487	1.0788	134.91	1.4481	1.0774
		cc-pVQZ (n_V = 324)					
CASSCF	7.53	1.5467	1.3462	1.0689	134.84	1.4400	1.0678
CASPT2	3.87	1.5488	1.3520	1.0763	134.98	1.4418	1.0757
pc-NEVPT2	9.49	1.5584	1.3421	1.0766	134.95	1.4403	1.0751
ic-MRCISD+Q	9.15	1.5587	1.3443	1.0758	134.94	1.4427	1.0746
Mk-MRCCSD	10.28	1.5591	1.3417	1.0768	134.94	1.4406	1.0756
Mk-MRCCSD(T)	8.69	1.5634	1.3500	1.0791	134.95	1.4480	1.0779
DF-MR-LDSRG(2)+NIVO (s = 0.5)	10.89	1.5576	1.3427	1.0759	134.94	1.4410	1.0743
DF-sq-MR-LDSRG(2)+NIVO (s = 0.5)	10.92	1.5575	1.3426	1.0759	134.94	1.4409	1.0742
DF-MR-LDSRG(2)+NIVO (s = 1.0)	10.16	1.5634	1.3452	1.0782	134.96	1.4447	1.0768
DF-sq-MR-LDSRG(2)+NIVO (s = 1.0)	10.21	1.5631	1.3451	1.0781	134.96	1.4446	1.0766
		cc-pV5Z (n_V = 568)					
DF-MR-LDSRG(2)+NIVO (s = 0.5)[d]	10.94						
DF-sq-MR-LDSRG(2)+NIVO (s = 0.5)[d]	10.96						
DF-MR-LDSRG(2)+NIVO (s = 1.0)[d]	10.26						
DF-sq-MR-LDSRG(2)+NIVO (s = 1.0)[d]	10.30						

[a]All computations employed the CAS(4e,4o) reference and core orbitals constructed from carbon 1s orbitals were frozen for all post-CASSCF computations. The number of virtual orbitals ($n_V = N_V/2$) is reported for each basis set. [b]C−C and C=C refer to the longer and shorter carbon−carbon bonds, respectively. [c]∠C−C−H is the bond angle between the C−H bond and the longer C−C bond. [d]Based on the corresponding cc-pVQZ optimized geometries.

butadiene using the cc-pV5Z basis (580 correlated orbitals) on a single node with 128 GB of memory.

5. CONCLUSIONS

In this work, we describe a strategy to reduce the computational and memory costs of the multireference driven similarity renormalization group (MR-DSRG). We demonstrate that the cost of the linear MR-DSRG with singles and doubles [MR-LDSRG(2)] can be lowered substantially without compromising its accuracy by using a combination of (1) a sequential unitary transformation, (2) density fitting (DF) of the two-electron integrals, and (3) the noninteracting virtual orbital (NIVO) approximation. The sequential MR-DSRG scheme introduced in this work [sq-MR-DSRG] reduces the cost of evaluating single-excitations and allows to treat them exactly. Like in the case of Brueckner coupled cluster values reported,[132−134] this approach reduces the number of algebraic terms in the DSRG equations because there are no terms (diagrams) containing single excitations. The use of DF integrals reduces the memory requirements of the original MR-DSRG(2) from $O(N^4)$ to $O(N^2M)$, where N and M are the number of basis functions and auxiliary basis functions, respectively. Density fitting is particularly convenient when combined with the sequential approach because the contributions of singles can be directly included in the DF auxiliary three-index integrals, reducing the integral transformation cost from $O(N^5)$ to $O(N^3M)$. The NIVO approximation neglects the operator components of a commutator with three or more virtual indices. A formal analysis of this approximation showed that the leading error is of fourth order in perturbation theory.

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

In practice, NIVO is crucial to both avoiding the memory bottleneck of the MR-DSRG(2) and reducing the computational cost to evaluate the transformed Hamiltonian.

To benchmark the MR-LDSRG(2) and sq-MR-LDSRG(2) approaches and assess the impact of the DF and NIVO approximations, we computed the spectroscopic constants of eight diatomic molecules using the full-valence active space and the cc-pVQZ basis set. Compared to experimental data, both MR-DSRG methods yield results that are as accurate as those obtained with CCSD(T). Moreover, the DF-sq-MR-LDSRG(2)+NIVO results are almost identical to those computed without the NIVO approximation: the harmonic vibrational frequencies, anharmonicity constants, and dissociation energies only differ by, on average, 0.1 cm^{-1}, 0.1 cm^{-1}, and 0.2 kcal mol^{-1}, respectively. These results supports our claim that the speedup brought by the NIVO approximation does not sacrifice the accuracy of both variants of the MR-LDSRG(2).

Combining DF and the NIVO approximation, both the traditional and sequential MR-LDSRG(2) can be routinely applied to chemical systems with more than 500 basis function. We demonstrate this point by studying the automerization reaction of cyclobutadiene using a quintuple-ζ basis set (580 basis functions). Our best estimate of the reaction barrier from DF-sq-MR-LDSRG(2)+NIVO/cc-pV5Z is 10.3 kcal mol^{-1}. However, we expect that this result is likely overestimated due to the lack of three-body corrections in the MR-LDSRG(2) theory. Our results agree well with Mk-MRCCSD predictions and other multireference coupled cluster values reported in the literature.

In conclusion, we have shown that it is possible to significantly reduce the cost of MR-LDSRG(2) computations without reducing the accuracy of this approach. The sequential approach and NIVO approximations are general, and can be applied to improve the efficiency of other unitary non-perturbative methods (e.g., unitary coupled cluster theory) and downfolding schemes for classical-quantum hybrid algorithms.[47]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.9b00353.

Detailed timings and formal scaling required to evaluate the transformed Hamiltonian, individual prediction errors for the eight diatomic molecules by different methods, and convergence of DF-sq-MR-LDSRG(2)+NIVO absolute energy of cyclobutadiene molecules with respect to flow parameter (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: tianyuanzhang1993@gmail.com.
*E-mail: cli62@emory.edu.
*E-mail: francesco.evangelista@emory.edu.

ORCID

Tianyuan Zhang: 0000-0001-9655-7927
Chenyang Li: 0000-0002-0059-4174
Francesco A. Evangelista: 0000-0002-7917-6652

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy under Award No. DE-SC0016004, a Alfred P. Sloan Research Fellowship (FG-20166748), and a Camille and Henry Dreyfus Foundation Teacher-Scholar Award (TC-18-045).

■ REFERENCES

(1) (a) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **1990**, *94*, 5483−5488. (b) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218−1226.

(2) Hirao, K. Multireference Møller−Plesset method. *Chem. Phys. Lett.* **1992**, *190*, 374−380.

(3) Nakano, H. Quasidegenerate perturbation theory with multiconfigurational self-consistent-field reference functions. *J. Chem. Phys.* **1993**, *99*, 7983−7992.

(4) Werner, H.-J. Third-order multireference perturbation theory The CASPT3 method. *Mol. Phys.* **1996**, *89*, 645−661.

(5) Mahapatra, U. S.; Datta, B.; Mukherjee, D. Molecular Applications of a Size-Consistent State-Specific Multireference Perturbation Theory with Relaxed Model-Space Coefficients. *J. Phys. Chem. A* **1999**, *103*, 1822−1830.

(6) (a) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J. P. Introduction of *n*-electron valence states for multireference perturbation theory. *J. Chem. Phys.* **2001**, *114*, 10252−10264. (b) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. *n*-electron valence state perturbation theory: A spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *J. Chem. Phys.* **2002**, *117*, 9138−9153. (c) Angeli, C.; Pastore, M.; Cimiraglia, R. New perspectives in multireference perturbation theory: the *n*-electron valence state approach. *Theor. Chem. Acc.* **2007**, *117*, 743−754.

(7) Sokolov, A. Y. Multi-reference algebraic diagrammatic construction theory for excited states: General formulation and first-order implementation. *J. Chem. Phys.* **2018**, *149*, 204113.

(8) Liu, B. *Ab initio* potential energy surface for linear H_3. *J. Chem. Phys.* **1973**, *58*, 1925−1937.

(9) (a) Werner, H.-J.; Knowles, P. J. An efficient internally contracted multiconfiguration−reference configuration interaction method. *J. Chem. Phys.* **1988**, *89*, 5803−5814. (b) Knowles, P. J.; Werner, H.-J. An efficient method for the evaluation of coupling coefficients in configuration interaction calculations. *Chem. Phys. Lett.* **1988**, *145*, 514−522.

(10) Hanrath, M.; Engels, B. New algorithms for an individually selecting MR-CI program. *Chem. Phys.* **1997**, *225*, 197−202.

(11) Szalay, P. G.; Müller, T.; Gidofalvi, G.; Lischka, H.; Shepard, R. Multiconfiguration self-consistent field and multireference configuration interaction methods and applications. *Chem. Rev.* **2012**, *112*, 108−181.

(12) Sivalingam, K.; Krupicka, M.; Auer, A. A.; Neese, F. Comparison of fully internally and strongly contracted multireference configuration interaction procedures. *J. Chem. Phys.* **2016**, *145*, 054104.

(13) Čížek, J. On the Use of the Cluster Expansion and the Technique of Diagrams in Calculations of Correlation Effects in Atoms and Molecules. *Adv. Chem. Phys.* **2007**, *14*, 35−89.

(14) Jeziorski, B.; Monkhorst, H. J. Coupled-cluster method for multideterminantal reference states. *Phys. Rev. A: At., Mol., Opt. Phys.* **1981**, *24*, 1668−1681.

(15) Haque, M. A.; Mukherjee, D. Application of cluster expansion techniques to open shells: Calculation of difference energies. *J. Chem. Phys.* **1984**, *80*, 5058−5069.

(16) Stolarczyk, L. Z.; Monkhorst, H. J. Coupled-cluster method in Fock space. I. General formalism. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *32*, 725−742.

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

(17) Malrieu, J. P.; Durand, P.; Daudey, J. P. Intermediate Hamiltonians as a new class of effective Hamiltonians. *J. Phys. A: Math. Gen.* **1985**, *18*, 809−826.
(18) Lindgren, I.; Mukherjee, D. On the connectivity criteria in the open-shell coupled-cluster theory for general model spaces. *Phys. Rep.* **1987**, *151*, 93−127.
(19) (a) Piecuch, P.; Paldus, J. Orthogonally spin-adapted multi-reference Hilbert space coupled-cluster formalism: diagrammatic formulation. *Theoret. Chim. Acta* **1992**, *83*, 69−103. (b) Paldus, J.; Piecuch, P.; Pylypow, L.; Jeziorski, B. Application of Hilbert-space coupled-cluster theory to simple $(H_2)_2$ model systems: Planar models. *Phys. Rev. A: At., Mol., Opt. Phys.* **1993**, *47*, 2738−2782.
(20) Mášik, J.; Hubač, I. Multireference Brillouin-Wigner Coupled-Cluster Theory. Single-root approach. *Adv. Quantum Chem.* **1998**, *31*, 75−104.
(21) (a) Mahapatra, U. S.; Datta, B.; Bandyopadhyay, B.; Mukherjee, D. State-Specific Multi-Reference Coupled Cluster Formulations: Two Paradigms. *Adv. Quantum Chem.* **1998**, *30*, 163−193. (b) Mahapatra, U. S.; Datta, B.; Mukherjee, D. A state-specific multi-reference coupled cluster formalism with molecular applications. *Mol. Phys.* **1998**, *94*, 157−171. (c) Mahapatra, U. S.; Datta, B.; Mukherjee, D. A size-consistent state-specific multireference coupled cluster theory: Formal developments and molecular applications. *J. Chem. Phys.* **1999**, *110*, 6171−6188.
(22) Li, S. Block-correlated coupled cluster theory: The general formulation and its application to the antiferromagnetic Heisenberg model. *J. Chem. Phys.* **2004**, *120*, 5017−5026.
(23) Hanrath, M. An exponential multireference wave-function Ansatz. *J. Chem. Phys.* **2005**, *123*, 084102.
(24) Evangelista, F. A.; Gauss, J. An orbital-invariant internally contracted multireference coupled cluster approach. *J. Chem. Phys.* **2011**, *134*, 114102.
(25) Hanauer, M.; Köhn, A. Pilot applications of internally contracted multireference coupled cluster theory, and how to choose the cluster operator properly. *J. Chem. Phys.* **2011**, *134*, 204111.
(26) Datta, D.; Kong, L.; Nooijen, M. A state-specific partially internally contracted multireference coupled cluster approach. *J. Chem. Phys.* **2011**, *134*, 214116.
(27) Chen, Z.; Hoffmann, M. R. Orbitally invariant internally contracted multireference unitary coupled cluster theory and its perturbative approximation: Theory and test calculations of second order approximation. *J. Chem. Phys.* **2012**, *137*, 014108.
(28) Hoffmann, M. R.; Simons, J. A unitary multiconfigurational coupled-cluster method: Theory and applications. *J. Chem. Phys.* **1988**, *88*, 993−1002.
(29) Bartlett, R. J.; Kucharski, S. A.; Noga, J. Alternative coupled-cluster ansätze II. The unitary coupled-cluster method. *Chem. Phys. Lett.* **1989**, *155*, 133−140.
(30) Watts, J. D.; Trucks, G. W.; Bartlett, R. J. The unitary coupled-cluster approach and molecular properties. Applications of the UCC(4) method. *Chem. Phys. Lett.* **1989**, *157*, 359−366.
(31) Kutzelnigg, W. Error analysis and improvements of coupled-cluster theory. *Theoret. Chim. Acta* **1991**, *80*, 349−386.
(32) Mertins, F.; Schirmer, J. Algebraic propagator approaches and intermediate-state representations. I. The biorthogonal and unitary coupled-cluster methods. *Phys. Rev. A: At., Mol., Opt. Phys.* **1996**, *53*, 2140−2152.
(33) Taube, A. G.; Bartlett, R. J. New perspectives on unitary coupled-cluster theory. *Int. J. Quantum Chem.* **2006**, *106*, 3393−3401.
(34) Harsha, G.; Shiozaki, T.; Scuseria, G. E. On the difference between variational and unitary coupled cluster theories. *J. Chem. Phys.* **2018**, *148*, 044107.
(35) Yung, M. H.; Casanova, J.; Mezzacapo, A.; McClean, J.; Lamata, L.; Aspuru-Guzik, A.; Solano, E. From transistor to trapped-ion computers for quantum chemistry. *Sci. Rep.* **2015**, *4*, 714.
(36) Peruzzo, A.; McClean, J.; Shadbolt, P.; Yung, M.-H.; Zhou, X.-Q.; Love, P. J.; Aspuru-Guzik, A.; O'Brien, J. L. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **2014**, *5*, 36.
(37) O'Malley, P. J. J.; Babbush, R.; Kivlichan, I. D.; Romero, J.; McClean, J. R.; Barends, R.; Kelly, J.; Roushan, P.; Tranter, A.; Ding, N.; Campbell, B.; Chen, Y.; Chen, Z.; Chiaro, B.; Dunsworth, A.; Fowler, A. G.; Jeffrey, E.; Lucero, E.; Megrant, A.; Mutus, J. Y.; Neeley, M.; Neill, C.; Quintana, C.; Sank, D.; Vainsencher, A.; Wenner, J.; White, T. C.; Coveney, P. V.; Love, P. J.; Neven, H.; Aspuru-Guzik, A.; Martinis, J. M. Scalable Quantum Simulation of Molecular Energies. *Phys. Rev. X* **2016**, *6*, 361.
(38) Shen, Y.; Zhang, X.; Zhang, S.; Zhang, J.-N.; Yung, M.-H.; Kim, K. Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure. *Phys. Rev. A: At., Mol., Opt. Phys.* **2017**, *95*, 020501.
(39) Dumitrescu, E. F.; McCaskey, A. J.; Hagen, G.; Jansen, G. R.; Morris, T. D.; Papenbrock, T.; Pooser, R. C.; Dean, D. J.; Lougovski, P. Cloud Quantum Computing of an Atomic Nucleus. *Phys. Rev. Lett.* **2018**, *120*, 210501.
(40) Ryabinkin, I. G.; Yen, T.-C.; Genin, S. N.; Izmaylov, A. F. Qubit Coupled Cluster Method: ASystematic Approach to Quantum Chemistry on a Quantum Computer. *J. Chem. Theory Comput.* **2018**, *14*, 6317−6326.
(41) Hempel, C.; Maier, C.; Romero, J.; McClean, J.; Monz, T.; Shen, H.; Jurcevic, P.; Lanyon, B. P.; Love, P.; Babbush, R.; Aspuru-Guzik, A.; Blatt, R.; Roos, C. F. Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator. *Phys. Rev. X* **2018**, *8*, 031022.
(42) Romero, J.; Babbush, R.; McClean, J. R.; Hempel, C.; Love, P. J.; Aspuru-Guzik, A. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Sci. Technol.* **2019**, *4*, 014008.
(43) Lee, J.; Huggins, W. J.; Head-Gordon, M.; Whaley, K. B. Generalized Unitary Coupled Cluster Wave functions for Quantum Computation. *J. Chem. Theory Comput.* **2019**, *15*, 311−324.
(44) Li, Y.; Hu, J.; Zhang, X. M.; Song, Z.; Yung, M.-H. Variational Quantum Simulation for Quantum Chemistry. *Adv. Theory Simul.* **2019**, *2*, 1800182.
(45) Kühn, M.; Zanker, S.; Deglmann, P.; Marthaler, M.; Weiß, H. Accuracy and Resource Estimations for Quantum Chemistry on a Near-term Quantum Computer. 2018, arXiv:1812.06814. arXiv.org e-Print archive. https://arxiv.org/abs/1812.06814
(46) Nam, Y.; Chen, J.-S.; Pisenti, N. C.; Wright, K.; Delaney, C.; Maslov, D.; Brown, K. R.; Allen, S.; Amini, J. M.; Apisdorf, J.; Beck, K. M.; Blinov, A.; Chaplin, V.; Chmielewski, M.; Collins, C.; Debnath, S.; Ducore, A. M.; Hudek, K. M.; Keesan, M.; Kreikemeier, S. M.; Mizrahi, J.; Solomon, P.; Williams, M.; Wong-Campos, J. D.; Monroe, C.; Kim, J. Ground-state energy estimation of the water molecule on a trapped ion quantum computer. 2019, arXiv:1902.10171. arXiv.org e-Print archive. https://arxiv.org/abs/1902.10171
(47) Bauman, N. P.; Bylaska, E. J.; Krishnamoorthy, S.; Low, G. H.; Wiebe, N.; Kowalski, K. Downfolding of many-body Hamiltonians using active-space models: extension of the sub-system embedding sub-algebras approach to unitary coupled cluster formalisms. 2019, arXiv:1902.01553. arXiv.org e-Print archive. https://arxiv.org/abs/1902.01553
(48) Evangelista, F. A. Alternative single-reference coupled cluster approaches for multireference problems: the simpler, the better. *J. Chem. Phys.* **2011**, *134*, 224102.
(49) (a) Yanai, T.; Chan, G. K.-L. Canonical transformation theory for multireference problems. *J. Chem. Phys.* **2006**, *124*, 194106. (b) Yanai, T.; Chan, G. K.-L. Canonical transformation theory from extended normal ordering. *J. Chem. Phys.* **2007**, *127*, 104107.
(50) Evangelista, F. A.; Gauss, J. On the approximation of the similarity-transformed Hamiltonian in single-reference and multi-reference coupled cluster theory. *Chem. Phys.* **2012**, *401*, 27−35.
(51) Neuscamman, E.; Yanai, T.; Chan, G. K.-L. Quadratic canonical transformation theory and higher order density matrices. *J. Chem. Phys.* **2009**, *130*, 124102.
(52) Neuscamman, E.; Yanai, T.; Chan, G. K.-L. A review of canonical transformation theory. *Int. Rev. Phys. Chem.* **2010**, *29*, 231−271.

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

(53) Wegner, F. Flow-equations for Hamiltonians. *Ann. Phys.* **1994**, *506*, 77−91.
(54) Kehrein, S. *The Flow Equation Approach to Many-Particle Systems*; Springer: Berlin, 2006.
(55) Evangelista, F. A. A driven similarity renormalization group approach to quantum many-body problems. *J. Chem. Phys.* **2014**, *141*, 054109.
(56) Li, C.; Evangelista, F. A. Multireference Driven Similarity Renormalization Group: A Second-Order Perturbative Analysis. *J. Chem. Theory Comput.* **2015**, *11*, 2097−2108.
(57) (a) Li, C.; Evangelista, F. A. Towards numerically robust multireference theories: The driven similarity renormalization group truncated to one- and two-body operators. *J. Chem. Phys.* **2016**, *144*, 164114. (b) Li, C.; Evangelista, F. A. Erratum: "Towards numerically robust multireference theories: The driven similarity renormalization group truncated to one- and two-body operators" [J. Chem. Phys. 144, 164114 (2016)]. *J. Chem. Phys.* **2018**, *148*, 079903.
(58) (a) Li, C.; Evangelista, F. A. Driven similarity renormalization group: Third-order multireference perturbation theory. *J. Chem. Phys.* **2017**, *146*, 124132. (b) Li, C.; Evangelista, F. A. Erratum: "Driven similarity renormalization group: Third-order multireference perturbation theory" [J. Chem. Phys. 146, 124132 (2017)]. *J. Chem. Phys.* **2018**, *148*, 079902.
(59) Li, C.; Evangelista, F. A. Driven similarity renormalization group for excited states: A state-averaged perturbation theory. *J. Chem. Phys.* **2018**, *148*, 124106.
(60) Li, C.; Evangelista, F. A. Multireference Theories of Electron Correlation Based on the Driven Similarity Renormalization Group. *Annu. Rev. Phys. Chem.* **2019**, *70*, 245.
(61) Meller, J.; Malrieu, J. P.; Caballol, R. State-specific coupled cluster-type dressing of multireference singles and doubles configuration interaction matrix. *J. Chem. Phys.* **1996**, *104*, 4068−4076.
(62) Köhn, A.; Hanauer, M.; Mück, L. A.; Jagau, T.-C.; Gauss, J. State-specific multireference coupled-cluster theory. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, *3*, 176−197.
(63) Lyakh, D. I.; Musiał, M.; Lotrich, V. F.; Bartlett, R. J. Multireference Nature of Chemistry: The Coupled-Cluster View. *Chem. Rev.* **2012**, *112*, 182−243.
(64) Evangelista, F. A. Perspective: Multireference coupled cluster theories of dynamical electron correlation. *J. Chem. Phys.* **2018**, *149*, 030901.
(65) (a) Schucan, T. H.; Weidenmüller, H. A. The effective interaction in nuclei and its perturbation expansion: An algebraic approach. *Ann. Phys.* **1972**, *73*, 108−135. (b) Schucan, T. H.; Weidenmüller, H. A. Perturbation theory for the effective interaction in nuclei. *Ann. Phys.* **1973**, *76*, 483−509.
(66) Salomonson, S.; Lindgren, I.; Mårtensson, A.-M. Numerical Many-Body Perturbation Calculations on Be-like Systems Using a Multi-Configurational Model Space. *Phys. Scr.* **1980**, *21*, 351−356.
(67) Evangelisti, S.; Daudey, J. P.; Malrieu, J. P. *Qualitative intruder-state problems in effective Hamiltonian theory and their solution through intermediate Hamiltonians* **1987**, *35*, 4930−4941.
(68) Zarrabian, S.; Laidig, W. D.; Bartlett, R. J. Convergence properties of multireference many-body perturbation theory. *Phys. Rev. A: At., Mol., Opt. Phys.* **1990**, *41*, 4711−4720.
(69) (a) Kowalski, K.; Piecuch, P. Complete set of solutions of multireference coupled-cluster equations: The state-universal formalism. *Phys. Rev. A: At., Mol., Opt. Phys.* **2000**, *61*, 052506. (b) Kowalski, K.; Piecuch, P. Complete set of solutions of the generalized Bloch equation. *Int. J. Quantum Chem.* **2000**, *80*, 757−781.
(70) Evangelista, F. A.; Hanauer, M.; Köhn, A.; Gauss, J. A sequential transformation approach to the internally contracted multireference coupled cluster method. *J. Chem. Phys.* **2012**, *136*, 204108.
(71) Whitten, J. L. Coulombic potential energy integrals and approximations. *J. Chem. Phys.* **1973**, *58*, 4496−4501.
(72) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. On some approximations in applications of Xα theory. *J. Chem. Phys.* **1979**, *71*, 3396−3402.
(73) Kendall, R. A.; Früchtl, H. A. The impact of the resolution of the identity approximate integral method on modern ab initio algorithm development. *Theor. Chem. Acc.* **1997**, *97*, 158−163.
(74) Beebe, N. H. F.; Linderberg, J. Simplifications in the generation and transformation of two-electron integrals in molecular calculations. *Int. J. Quantum Chem.* **1977**, *12*, 683−705.
(75) Røeggen, I.; Wisløff-Nilssen, E. On the Beebe-Linderberg two-electron integral approximation. *Chem. Phys. Lett.* **1986**, *132*, 154−160.
(76) Koch, H.; Sanchez de Merás, A.; Pedersen, T. B. Reduced scaling in electronic structure calculations using Cholesky decompositions. *J. Chem. Phys.* **2003**, *118*, 9481−9484.
(77) Boman, L.; Koch, H.; Sanchez de Merás, A. Method specific Cholesky decomposition: Coulomb and exchange energies. *J. Chem. Phys.* **2008**, *129*, 134107.
(78) Aquilante, F.; Gagliardi, L.; Pedersen, T. B.; Lindh, R. Atomic Cholesky decompositions: A route to unbiased auxiliary basis sets for density fitting approximation with tunable accuracy and efficiency. *J. Chem. Phys.* **2009**, *130*, 154107.
(79) Parrish, R. M.; Sherrill, C. D.; Hohenstein, E. G.; Kokkila, S. I. L.; Martínez, T. J. Communication: Acceleration of coupled cluster singles and doubles via orbital-weighted least-squares tensor hypercontraction. *J. Chem. Phys.* **2014**, *140*, 181102.
(80) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* **1993**, *213*, 514−518.
(81) Aquilante, F.; Pedersen, T. B.; Lindh, R. Low-cost evaluation of the exchange Fock matrix from Cholesky and density fitting representations of the electron repulsion integrals. *J. Chem. Phys.* **2007**, *126*, 194106.
(82) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.* **1993**, *208*, 359−363.
(83) Weigend, F. A fully direct RI-HF algorithm: Implementation, optimized auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285−4291.
(84) Werner, H.-J.; Manby, F. R.; Knowles, P. J. Fast linear scaling second-order Møller-Plesset perturbation theory (MP2) using local and density fitting approximations. *J. Chem. Phys.* **2003**, *118*, 8149−8160.
(85) Aquilante, F.; Malmqvist, P.-Å.; Pedersen, T. B.; Ghosh, A.; Roos, B. O. Cholesky Decomposition-Based Multiconfiguration Second-Order Perturbation Theory (CD-CASPT2): Application to the Spin-State Energetics of Co III(diiminato)(NPh). *J. Chem. Theory Comput.* **2008**, *4*, 694−702.
(86) Boström, J.; Delcey, M. G.; Aquilante, F.; Serrano-Andrés, L.; Pedersen, T. B.; Lindh, R. Calibration of Cholesky Auxiliary Basis Sets for Multiconfigurational Perturbation Theory Calculations of Excitation Energies. *J. Chem. Theory Comput.* **2010**, *6*, 747−754.
(87) Győrffy, W.; Shiozaki, T.; Knizia, G.; Werner, H.-J. Analytical energy gradients for second-order multireference perturbation theory using density fitting. *J. Chem. Phys.* **2013**, *138*, 104104.
(88) Hannon, K. P.; Li, C.; Evangelista, F. A. An integral-factorized implementation of the driven similarity renormalization group second-order multireference perturbation theory. *J. Chem. Phys.* **2016**, *144*, 204111.
(89) Freitag, L.; Knecht, S.; Angeli, C.; Reiher, M. Multireference Perturbation Theory with Cholesky Decomposition for the Density Matrix Renormalization Group. *J. Chem. Theory Comput.* **2017**, *13*, 451−459.
(90) Hättig, C.; Weigend, F. CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. *J. Chem. Phys.* **2000**, *113*, 5154−5161.
(91) Pedersen, T. B.; Sanchez de Merás, A. M. J.; Koch, H. Polarizability and optical rotation calculated from the approximate coupled cluster singles and doubles CC2 linear response theory using Cholesky decompositions. *J. Chem. Phys.* **2004**, *120*, 8887−8897.

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

(92) Rendell, A. P.; Lee, T. J. Coupled-cluster theory employing approximate integrals: An approach to avoid the input/output and storage bottlenecks. *J. Chem. Phys.* **1994**, *101*, 400−408.
(93) Boström, J.; Pitoňák, M.; Aquilante, F.; Neogrady, P.; Pedersen, T. B.; Lindh, R. Coupled Cluster and Møller-Plesset Perturbation Theory Calculations of Noncovalent Intermolecular Interactions using Density Fitting with Auxiliary Basis Sets from Cholesky Decompositions. *J. Chem. Theory Comput.* **2012**, *8*, 1921−1928.
(94) DePrince, A. E.; Sherrill, C. D. Accuracy and Efficiency of Coupled-Cluster Theory Using Density Fitting/Cholesky Decomposition, Frozen Natural Orbitals, and a t1-Transformed Hamiltonian. *J. Chem. Theory Comput.* **2013**, *9*, 2687−2696.
(95) Epifanovsky, E.; Zuev, D.; Feng, X.; Khistyaev, K.; Shao, Y.; Krylov, A. I. General implementation of the resolution-of-the-identity and Cholesky representations of electron repulsion integrals within coupled-cluster and equation-of-motion methods: Theory and benchmarks. *J. Chem. Phys.* **2013**, *139*, 134105.
(96) Qiu, Y.; Henderson, T. M.; Zhao, J.; Scuseria, G. E. Projected coupled cluster theory. *J. Chem. Phys.* **2017**, *147*, 064111.
(97) Huntington, L. M. J.; Nooijen, M. pCCSD: parameterized coupled-cluster theory with single and double excitations. *J. Chem. Phys.* **2010**, *133*, 184109.
(98) (a) Kats, D.; Manby, F. R. Communication: The distinguishable cluster approximation. *J. Chem. Phys.* **2013**, *139*, 021102. (b) Kats, D. Communication: The distinguishable cluster approximation. II. The role of orbital relaxation. *J. Chem. Phys.* **2014**, *141*, 061101.
(99) Rishi, V.; Perera, A.; Nooijen, M.; Bartlett, R. J. Excited states from modified coupled cluster methods: Are they any better than EOM CCSD? *J. Chem. Phys.* **2017**, *146*, 144104.
(100) Koch, S.; Kutzelnigg, W. Comparison of CEPA and CP-MET methods. *Theor. Chim. Acta* **1981**, *59*, 387−411.
(101) Bartlett, R. J.; Musiał, M. Addition by subtraction in coupled-cluster theory: A reconsideration of the CC and CI interface and the nCC hierarchy. *J. Chem. Phys.* **2006**, *125*, 204105.
(102) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* **1980**, *48*, 157−173.
(103) (a) Mukherjee, D. Normal ordering and a Wick-like reduction theorem for fermions with respect to a multi-determinantal reference state. *Chem. Phys. Lett.* **1997**, *274*, 561−566. (b) Kutzelnigg, W.; Mukherjee, D. Normal order and extended Wick theorem for a multiconfiguration reference wave function. *J. Chem. Phys.* **1997**, *107*, 432−449. (c) Kutzelnigg, W.; Shamasundar, K. R.; Mukherjee, D. Spinfree formulation of reduced density matrices, density cumulants and generalised normal ordering. *Mol. Phys.* **2010**, *108*, 433−451. (d) Kong, L.; Nooijen, M.; Mukherjee, D. An algebraic proof of generalized Wick theorem. *J. Chem. Phys.* **2010**, *132*, 234107. (e) Sinha, D.; Maitra, R.; Mukherjee, D. Generalized antisymmetric ordered products, generalized normal ordered products, ordered and ordinary cumulants and their use in many electron correlation problem. *Comput. Theor. Chem.* **2013**, *1003*, 62−70.
(104) Kats, D.; Usvyat, D.; Manby, F. R. Particle−hole symmetry in many-body theories of electron correlation. *Mol. Phys.* **2018**, *116*, 1496−1503.
(105) *Forte, a suite of quantum chemistry methods for strongly correlated electrons*. For current version see https://github.com/evangelistalab/forte, 2019.
(106) *Ambit, a C++ library for the implementation of tensor product calculations through a clean, concise user interface*. For current version see https://github.com/jturney/ambit, 2018.
(107) (a) Parrish, R. M.; Burns, L. A.; Smith, D. G. A.; Simmonett, A. C.; DePrince, A. E.; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Di Remigio, R.; Richard, R. M.; Gonthier, J. F.; James, A. M.; McAlexander, H. R.; Kumar, A.; Saitow, M.; Wang, X.; Pritchard, B. P.; Verma, P.; Schaefer, H. F.; Patkowski, K.; King, R. A.; Valeev, E. F.; Evangelista, F. A.; Turney, J. M.; Crawford, T. D.; Sherrill, C. D. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theory Comput.* **2017**, *13*, 3185−3197. (b) Smith, D. G. A.; Burns, L. A.; Sirianni, D. A.; Nascimento, D. R.; Kumar, A.; James, A. M.; Schriber, J. B.; Zhang, T.; Zhang, B.; Abbott, A. S.; Berquist, E. J.; Lechner, M. H.; Cunha, L. A.; Heide, A. G.; Waldrop, J. M.; Takeshita, T. Y.; Alenaizan, A.; Neuhauser, D.; King, R. A.; Simmonett, A. C.; Turney, J. M.; Schaefer, H. F.; Evangelista, F. A.; DePrince, A. E.; Crawford, T. D.; Patkowski, K.; Sherrill, C. D. Psi4NumPy: An Interactive Quantum Chemistry Programming Environment for Reference Implementations and Rapid Development. *J. Chem. Theory Comput.* **2018**, *14*, 3504−3511.
(108) Schieber, M.; Smith, D.; Chow, E.; Sherrill, C. Strong parallel scaling for sparsity-screened density-fitted integrals on Knights Landing. *Abstr. Pap.—Am. Chem. Soc.* **2018***255*.
(109) Hohenstein, E. G.; Parrish, R. M.; Martínez, T. J. Tensor hypercontraction density fitting. I. Quartic scaling second- and third-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2012**, *137*, 044103.
(110) Huber, K. P.; Herzberg, G. *Molecular Spectra and Molecular Structure*; Springer: Boston, MA, 1979.
(111) Irikura, K. K. Experimental Vibrational Zero-Point Energies: Diatomic Molecules. *J. Phys. Chem. Ref. Data* **2007**, *36*, 389−397.
(112) (a) Dunning, T. H., Jr Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1018. (b) Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. IV. Calculation of static electrical response properties. *J. Chem. Phys.* **1994**, *100*, 2975.
(113) (a) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297−3305. (b) Weigend, F. Hartree−Fock exchange fitting basis sets for H to Rn. *J. Comput. Chem.* **2008**, *29*, 167−175.
(114) (a) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175−3183. (b) Hättig, C. Optimization of auxiliary basis sets for RI-MP2 and RI-CC2 calculations: Core−valence and quintuple-ζ basis sets for H to Ar and QZVPP basis sets for Li to Kr. *Phys. Chem. Chem. Phys.* **2005**, *7*, 59−66.
(115) Shen, J.; Piecuch, P. Combining active-space coupled-cluster methods with moment energy corrections via the CC(P;Q) methodology, with benchmark calculations for biradical transition states. *J. Chem. Phys.* **2012**, *136*, 144104.
(116) Whitman, D. W.; Carpenter, B. K. Limits on the activation parameters for automerization of cyclobutadiene-1,2-d_2. *J. Am. Chem. Soc.* **1982**, *104*, 6473−6474.
(117) Nakamura, K.; Osamura, Y.; Iwata, S. Second-order Jahn−Teller effect of cyclobutadiene in low-lying states. An MCSCF study. *Chem. Phys.* **1989**, *136*, 67−77.
(118) Balková, A.; Bartlett, R. J. A multireference coupled-cluster study of the ground state and lowest excited states of cyclobutadiene. *J. Chem. Phys.* **1994**, *101*, 8972−8987.
(119) Sancho-García, J. C.; Pittner, J.; Čársky, P.; Hubač, I. Multireference coupled-cluster calculations on the energy of activation in the automerization of cyclobutadiene: Assessment of the state-specific multireference Brillouin−Wigner theory. *J. Chem. Phys.* **2000**, *112*, 8785−8788.
(120) Levchenko, S. V.; Krylov, A. I. Equation-of-motion spin-flip coupled-cluster model with single and double substitutions: Theory and application to cyclobutadiene. *J. Chem. Phys.* **2004**, *120*, 175−185.
(121) Eckert-Maksić, M.; Vazdar, M.; Barbatti, M.; Lischka, H.; Maksić, Z. B. Automerization reaction of cyclobutadiene and its barrier height: An ab initio benchmark multireference average-quadratic coupled cluster study. *J. Chem. Phys.* **2006**, *125*, 064310.
(122) Lyakh, D. I.; Lotrich, V. F.; Bartlett, R. J. The 'tailored' CCSD(T) description of the automerization of cyclobutadiene. *Chem. Phys. Lett.* **2011**, *501*, 166−171.

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414

(123) Shen, J.; Fang, T.; Li, S.; Jiang, Y. Performance of Block Correlated Coupled Cluster Method with the CASSCF Reference Function for the Prediction of Activation Barriers, Spectroscopic Constants in Diatomic Molecules, and Singlet−Triplet Gaps in Diradicals. *J. Phys. Chem. A* **2008**, *112*, 12518−12525.
(124) Li, X.; Paldus, J. Accounting for the exact degeneracy and quasidegeneracy in the automerization of cyclobutadiene via multi-reference coupled-cluster methods. *J. Chem. Phys.* **2009**, *131*, 114103.
(125) Bhaskaran-Nair, K.; Demel, O.; Pittner, J. Multireference state-specific Mukherjee's coupled cluster method with noniterative triexcitations. *J. Chem. Phys.* **2008**, *129*, 184105.
(126) Demel, O.; Shamasundar, K. R.; Kong, L.; Nooijen, M. Application of Double Ionization State-Specific Equation of Motion Coupled Cluster Method to Organic Diradicals. *J. Phys. Chem. A* **2008**, *112*, 11895−11902.
(127) Wu, J. I.-C.; Mo, Y.; Evangelista, F. A.; von Rague Schleyer, P. Is cyclobutadiene really highly destabilized by antiaromaticity? *Chem. Commun.* **2012**, *48*, 8437−8439.
(128) Davidson, E. R.; Silver, D. W. Size consistency in the dilute helium gas electronic structure. *Chem. Phys. Lett.* **1977**, *52*, 403−406.
(129) (a) Mahapatra, U. S.; Datta, B.; Mukherjee, D. Molecular Applications of a Size-Consistent State-Specific Multireference Perturbation Theory with Relaxed Model-Space Coefficients. *J. Phys. Chem. A* **1999**, *103*, 1822−1830. (b) Chattopadhyay, S.; Mahapatra, U. S.; Mukherjee, D. Development of a linear response theory based on a state-specific multireference coupled cluster formalism. *J. Chem. Phys.* **2000**, *112*, 7939−7952. (c) Chattopadhyay, S.; Pahari, D.; Mukherjee, D.; Mahapatra, U. S. A state-specific approach to multireference coupled electron-pair approximation like methods: Development and applications. *J. Chem. Phys.* **2004**, *120*, 5968−5986. (d) Pahari, D.; Chattopadhyay, S.; Deb, A.; Mukherjee, D. An orbital-invariant coupled electron-pair like approximant to a state-specific multi-reference coupled cluster formalism. *Chem. Phys. Lett.* **2004**, *386*, 307−312.
(130) (a) Evangelista, F. A.; Allen, W. D.; Schaefer, H. F. High-order excitations in state-universal and state-specific multireference coupled cluster theories: Model systems. *J. Chem. Phys.* **2006**, *125*, 154113. (b) Evangelista, F. A.; Simmonett, A. C.; Allen, W. D.; Schaefer, H. F.; Gauss, J. Triple excitations in state-specific multireference coupled cluster theory: Application of Mk-MRCCSDT and Mk-MRCCSDT-n methods to model systems. *J. Chem. Phys.* **2008**, *128*, 124104. (c) Evangelista, F. A.; Prochnow, E.; Gauss, J.; Schaefer, H. F. Perturbative triples corrections in state-specific multireference coupled cluster theory. *J. Chem. Phys.* **2010**, *132*, 074107.
(131) Taube, A. G.; Bartlett, R. J. Rethinking linearized coupled-cluster theory. *J. Chem. Phys.* **2009**, *130*, 144112.
(132) Chiles, R. A.; Dykstra, C. E. An electron pair operator approach to coupled cluster wave functions. Application to He_2, Be_2, and Mg_2 and comparison with CEPA methods. *J. Chem. Phys.* **1981**, *74*, 4544−4556.
(133) Handy, N. C.; Pople, J. A.; Head-Gordon, M.; Raghavachari, K.; Trucks, G. W. Size-consistent Brueckner theory limited to double substitutions. *Chem. Phys. Lett.* **1989**, *164*, 185−192.
(134) Stanton, J. F.; Gauss, J.; Bartlett, R. J. On the choice of orbitals for symmetry breaking problems with application to NO_3. *J. Chem. Phys.* **1992**, *97*, 5554−5559.

DOI: 10.1021/acs.jctc.9b00353
J. Chem. Theory Comput. 2019, 15, 4399−4414