Improving the efficiency of the multireference driven similarity renormalization group via sequential transformation, density fitting, and the non-interacting 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, USA

E-mail: tianyuan.zhang@emory.edu; cli62@emory.edu; francesco.evangelista@emory.edu

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)] approach, in which one-body rotations 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 the storage of large four-index intermediates, we propose a non-interacting virtual orbit (NIVO) approximation in which tensor blocks labeled by three and four virtual indices are neglected. The NIVO approximation reduces the computational cost 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 similar to the values 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 perturbation theories (MRPTs)\textsuperscript{1–6} and multireference configuration interaction (MRCI) schemes.\textsuperscript{7–11} Considerable efforts have been dedicated to the development of multireference coupled cluster (MRCC) methods,\textsuperscript{12–26} 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 considerably less attention.\textsuperscript{26–33} Unitary theories have, in principle, two advantages over traditional coupled cluster approaches: 1) the energy satisfies the variational principle, and 2) the transformed Hamiltonian is Hermitian. The latter property is an important advantage in new applications of unitary methods to quantum computing, both in quantum algorithms\textsuperscript{34–45} and downfolding approaches aimed at reducing the number of orbitals in quantum computations.\textsuperscript{46}

One of the main obstacles in the formulation of both single- and multi-reference 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

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

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,\textsuperscript{28,30,32} the transformed Hamiltonian may be computed as the following commutator series

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

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.

Various approximations have been proposed to evaluate the unitarily transformed Hamiltonian. Perhaps the simplest way to approximate the nonterminating unitary series is to truncate the BCH expansion after a certain number of commutators.\textsuperscript{47,48} Proof-of-principle studies on unitary coupled cluster (CC) theory\textsuperscript{47}
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\(^{32} \) have suggested tractable approximations to unitary CC theory based on the Zassenhaus expansion that are exact for a given number of electrons. A common way to truncate the unitary BCH series is to use a recursive approximation of the commutator \([\cdot, \hat{A}]\), as suggested by Yanai and Chan.\(^{49} \) In their linear truncation scheme, these authors proposed to approximate each single commutator \([\cdot, \hat{A}]\) 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. An advantage of this approach is that closed-form expressions for terms like \([\hat{O}, \hat{A}]_{0,1,2}\), where \(\hat{O}\) is an operator containing up to two-body terms can be easily derived. This truncation scheme has been employed in canonical transformation (CT) theory\(^{49} \) and has been used to truncate normal-ordered equations in the flow renormalization group of Wegner.\(^{50,51} \)

We have recently developed a multireference driven similarity renormalization group (MR-DSRG)\(^{52–56} \) approach that avoids the multiple-parentage problem\(^{20,22,57–60} \) and numerical instabilities\(^{59–65} \) encountered in other nonperturbative multireference methods. In the MR-DSRG, we perform a unitary transformation of the Hamiltonian controlled by a flow parameter, which determines to which extend the resulting effective Hamiltonian is band diagonal.\(^{53,56} \) The simplest nonperturbative approximation, the linearized MR-DSRG truncated to two-body operators [MR-LDSRG(2)],\(^{54} \) assumes that \(\hat{A}\) contains up to two-body operators (singles and doubles) and employs the linear commutator approximation of Yanai and Chan. Preliminary benchmarks indicate that the MR-LDSRG(2) method is more accurate than CCSD around equilibrium geometries, and that this accuracy is preserved along potential energy curves, especially for single-bond breaking processes.\(^{66} \)

The cost to evaluate a single commutator in the MR-LDSRG(2) scales as \(O(N_C^2 N_V^2 N^2) = O(N_C^2 N_V^4 + N_C^2 N_V^2 + \ldots)\) 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. 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 ten or more terms are usually required to convergence the energy error in the BCH series to \(10^{-12} E_h\).\(^{54} \) A second reason is that computing the BCH series requires storing large intermediate tensors with memory costs that scale 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, we consider an alternative ansatz for the MR-DSRG based on a sequential similarity transformation. 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) to reduce the memory requirements and the I/O cost by avoiding the storage of two-electron repulsion integrals. Together with Cholesky decomposition (CD) and other tensor decomposition schemes, these techniques have been crucial in enabling computations with 1000 or more basis functions and found application in numerous electronic structure methods, including coupled cluster methods. 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 non-interacting virtual orbital (NIVO) approximation. A perturbative analysis of the NIVO approximation shows that the errors introduced appear at third order. 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 are have modified coefficients or are completely removed.

This paper is organized in the following way. In Sec. 2 we present an overview of the MR-DSRG theory, discuss the sequential MR-DSRG, and introduced the NIVO approximation. Details of the implementation together with a discussion of timings are given in Sec. 3. In Sec. 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, in Sec. 5 we conclude this work with a discussion of the main results.

2. THEORY

We first define the orbital labeling convention employed in this work. The set of molecular spin orbitals \( \mathbf{G} \equiv \{ \phi^p, p = 1, 2, \ldots, N \} \) 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 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 indices \( i, j, k, l \) and those of particle by indices \( a, b, c, d \). General orbitals are labeled by indices.
The MR-DSRG assumes a multideterminantal reference wave function ($\Psi_0$):

$$|\Psi_0\rangle = \sum_{\mu=1}^{d} c_{\mu} |\Phi^{\mu}\rangle.$$  \hspace{1cm} (3)

In this work, the set of determinants \{\Phi^{\mu}; \mu = 1, 2, \ldots, d\} in Eq. (3) is assumed to form a complete active space (CAS) and the corresponding coefficients \{c_{\mu}; \mu = 1, 2, \ldots, d\} are obtained from a CAS self-consistent field (CASSCF) computation.\textsuperscript{100} All operators are then normal ordered with respect to $\Psi_0$ according to the scheme of Mukherjee and Kutzelnigg.\textsuperscript{101} For example, the bare Hamiltonian is expressed as

$$\hat{H} = E_0 + \sum_{pq} f_p^q \{\hat{a}_p^\dagger \hat{a}_q\} + \frac{1}{4} \sum_{pqrs} v_{pq}^{rs} \{\hat{a}_r^p \hat{a}_s^q\},$$  \hspace{1cm} (4)

where $E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle$ is the reference energy and \{\hat{a}_r^p \hat{a}_s^q\} is a product of creation ($\hat{a}_r^p \equiv \hat{a}_{r,s}^p$) and annihilation ($\hat{a}_r^p$) 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

$$f_p^q = h_p^q + \sum_{rs} v_{pq}^{rs} \gamma_{r,s}^p.$$  \hspace{1cm} (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 anti-symmetrized two-electron integrals, respectively. Here, we have also used the one-particle reduced density matrix (1-RDM) defined as $\gamma_{r,s}^p = \langle \Psi_0 | \hat{a}_{r,s}^p | \Psi_0 \rangle$.

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, a real number $s$ defined in the range of $[0, \infty)$. The resulting transformed Hamiltonian [$\hat{H}(s)$] is a function of $s$ defined as

$$\hat{H}(s) = e^{-\hat{A}(s)} \hat{H} e^{\hat{A}(s)}.$$  \hspace{1cm} (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 ($n$),

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

where $\hat{A}_k(s)$ is the $k$-body component of $\hat{A}(s)$. In the 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

$$\hat{T}_k(s) = \frac{1}{(k!)^2} \sum_{ij\cdots ab\cdots} \sum_{p} I_{aq}^p(s)\{\hat{a}_{ij\cdots ab\cdots}\}. $$

Note that the cluster amplitudes $I_{aq}^p(s)$ exclude internal excitations, which are labeled only with active orbital indices. The DSRG transformed Hamiltonian has a many-body expansion similar to Eq. (4),

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

where,

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

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

The goal of the DSRG transformation is to decouple the interactions between the reference wave function ($\Psi_0$) and its excited configurations. Such interactions are the couplings between hole and particle orbitals represented by generalized excitation $[\bar{H}_{pqrs}^{pq}(s)\{\hat{a}_{pq}^q\}]$ and de-excitation $[\bar{H}_{pqrs}^{pq}(s)\{\hat{a}_{pq}^q\}]$ 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 $\bar{H}_{od}(s)$. Instead of achieving a full decoupling of the off-diagonal components [i.e., $\bar{H}_{od}(s) = 0$], we demand that the DSRG transformation achieves a partial decoupling, avoiding the components of $\bar{H}_{od}(s)$ with small or vanishing Møller–Plesset energy denominators. This condition is imposed via the DSRG flow equation, a nonlinear implicit equation of the form

$$\bar{H}_{od}(s) = \bar{R}(s),$$
where the source operator $\hat{R}(s)$ is Hermitian and continuous in $s$. The role of the source operator in the DSRG flow equation is to drive the off-diagonal elements of $\tilde{H}(s)$ to zero, a goal achieved by an appropriate parameterization of $\hat{R}(s)$.

In the linearized MR-DSRG approximation [MR-LDSRG(2)], the BCH series [Eq. (2)] is evaluated by keeping up to two-body normal-ordered operators of each commutator. The transformed Hamiltonian can then be evaluated by the following recursive equations

\[
\begin{align*}
\hat{C}^{(k+1)}(s) &= \frac{1}{k+1}[\hat{C}^{(k)}(s), \hat{A}(s)]_{0,1,2}, \\
\tilde{H}^{(k+1)}(s) &= \tilde{H}^{(k)}(s) + \hat{C}^{(k+1)}(s),
\end{align*}
\]

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

The solution of Eq. (11) yields a set of amplitudes $t_{ij\cdots}^{ab\cdots}(s)$ that define the operator $\hat{A}(s)$ and the DSRG transformed Hamiltonian $\tilde{H}(s)$. From this latter quantity, the MR-DSRG electronic energy is computed as the expectation value with respect to the reference

\[
\tilde{E}_0(s) = \langle \Psi_0 | \tilde{H}(s) | \Psi_0 \rangle.
\]

We refer the energy computed using Eq. (13) as the unrelaxed energy since the reference coefficients are not optimized. To include reference relaxation effects, we require that $\Psi_0$ is an eigenstate of $\tilde{H}(s)$ within the space of reference determinants, a condition that is equivalent to solving the eigenvalue problem

\[
\sum_{\mu=1}^{d} \langle \Phi_\nu | \tilde{H}(s) | \Phi_\mu \rangle c'_\mu = E(s)c'_\nu.
\]

Equation (14) defines a new reference $\Psi'_0$ with expansion coefficients $c'_\mu$. This new reference may be used as a starting point for a subsequent MR-DSRG transformation and this procedure can be repeated until convergence and such converged energy is referred as 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 the MR-DSRG approach is an alternative way to transform the bare Hamiltonian via a sequence of unitary operators with increasing particle rank

\[
\tilde{H}(s) = e^{-\hat{A}_n(s)} \cdots e^{-\hat{A}_2(s)} \hat{H} e^{\hat{A}_1(s)} e^{\hat{A}_2(s)} \cdots e^{\hat{A}_n(s)}. \tag{15}
\]

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. However, these schemes are not equivalent for truncated \( \hat{A}(s) \) [for example, \( n = 2 \) in Eqs. (7) and (15)] due to the fact that operators of different particle rank do not commute, that is, \( [\hat{A}_i(s), \hat{A}_j(s)] \neq 0 \) for \( i \neq j \).

An advantage of the sq-MR-DSRG approach is that \( \hat{A}_1(s) \) can be exactly folded into the Hamiltonian via a unitary transformation. The resulting \( \hat{A}_1(s) \)-dressed Hamiltonian \([\tilde{H}(s)]\)

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

preserves the particle rank of the bare Hamiltonian [Eq. (4)]. The corresponding scalar and tensor components of \( \tilde{H}(s) \) can be obtained by a simple unitary transformation of the one- and two-electron integrals \( f_{po}^q \) and \( v_{pq}^{rs} \) and update of the scalar energy. As will be discussed in Sec. 3.1, the \( \hat{A}_1(s) \)-dressed Hamiltonian can be computed very efficiently when 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 is given by

\[
\tilde{H}(s) = e^{-\hat{A}_2(s)} \tilde{H}(s) e^{\hat{A}_2(s)}. \tag{17}
\]

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. 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 $\hat{A}_1(s)$

$$
[[\hat{H}, \hat{A}(s)]_{1,2}, \hat{A}(s)]_{0,1,2} \leftarrow [[\hat{H}, \hat{A}_1(s)]_{1,2}, \hat{A}_1(s)]_{0,1,2}
+[[\hat{H}, \hat{A}_1(s)]_{1,2}, \hat{A}_2(s)]_{0,1,2}
+[[\hat{H}, \hat{A}_2(s)]_{1,2}, \hat{A}_1(s)]_{0,1,2}.
$$

(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 treated by forming the operator $\hat{\tilde{H}}(s)$, and since the BCH expansion for such transformation does not generate intermediates with rank greater than two, all terms involving $\hat{A}_1(s)$ are treated exactly in the linearized approximation.

### 2.3 Alleviating the memory bottleneck: The non-interacting virtual orbital (NIVO) approximation

In both the traditional and sequential MR-DSRG approaches, the DF approximation reduces the cost to store both the bare and $\hat{A}_1(s)$-dressed Hamiltonian from $O(N^4)$ to $O(N^2M)$. However, in the evaluation of the recursive commutator approximation of $\hat{H}(s)$, two-body operators are generated during the evaluation of each intermediate commutator $[\hat{C}^{(k)}(s)]$ and $\hat{H}(s)$. These quantities have $O(N^4)$ storage cost and, thus, reintroduce the bottleneck avoided with DF.

In order to reduce the cost to store $\hat{H}_2(s)$ and $\hat{C}^{(k)}_2(s)$, we shall neglect 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. We propose a non-interacting virtual orbital (NIVO) approximation, which neglects the operator components of $\hat{C}^{(k+1)}_2(s) = [\hat{C}^{(k)}_{1,2}(s), \hat{A}(s)]_{2, k \geq 0}$, with three or more virtual orbital indices (VVV, VCVV, VVVA, etc.) in the recursive definition of the linearized BCH series. Neglecting these blocks, the number of elements in each NIVO-approximated tensor is reduced from $O(N^4)$ to $O(N^2N^2_H)$, a size comparable to that of the $\hat{A}_2(s)$ tensor. For instance, in the cyclobutadiene computation using a quintuple-$\zeta$ basis set reported in Sec. 4.2, the memory requirements of $\hat{H}_2(s)$ or $\hat{C}^{(k)}_2(s)$ 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 \( \hat{C}_2^{(1)}(s) = [\hat{H}(s), \hat{A}_2(s)]_2 \). Indicating the terms neglected in NIVO as \( \delta \hat{C}_2^{(1)}(s) \), we see that the first energy contribution affected by the NIVO approximation comes from the expectation value of the triple commutator term \( \{ \delta \hat{C}_0^{(3)}(s) \} \)

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

(19)

whose contributions are shown as diagrams in Fig. 1. From a perturbation theory perspective, these diagrams are of order four or higher [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.

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, 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. The DSRG equations were implemented as tensor contractions using the Ambit tensor library, while integrals were generated using the Psi4 package. In the following, we provide the details of our implementation of the sequential ansatz in combination with DF.
3.1 Sequential transformation

The \( \hat{A}_1(s) \)-dressed Hamiltonian \( \hat{H}(s) \), 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 \[ \hat{H}(s), \hat{A}_1(s), \) and \( \hat{U}(s) \] in this section. The one- and two-body components of \( \hat{H} (\hat{h}_p^q, \hat{v}_{pq}^{r's'}) \) are given by

\[ \begin{align*}
\hat{h}_p^q &= \sum_{pq} U^q_p U^p_q h_q^p, \\
\hat{v}_{pq}^{r's'} &= \sum_{pqrs} U^r_p U^s_q v_{rs}^{pq} U^p_q U^q_r.
\end{align*} \tag{20} \tag{21} \]

Here, the unitary matrix \( U_p^p = (U)_{p'p} \) and its inverse \( U_p^p = (U)^*_{p'p} \) are given by \( U = \exp[\hat{A}_1] \), where the matrix \( A \) is composed of elements of the \( \hat{A}_1 \) tensor, \( (A)_{ia} = t_{ia} \) and \( (A)_{ai} = -t_{ia} \). 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 \( \Psi_0 \) is given by

\[ \hat{H} = E_0 + \sum_{pq} \hat{f}_p^q \{ \hat{a}_q^p \} + \frac{1}{4} \sum_{pqrs} \hat{v}_{pq}^{r's'} \{ \hat{a}_q^p \hat{a}_r^s \hat{a}_s^r \hat{a}_r^q \}, \tag{22} \]

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

\[ \tilde{E}_0 = \sum_{ij} \tilde{h}_i^j \gamma_i^j + \frac{1}{4} \sum_{ijkl} \tilde{v}_{ij}^{kl} \gamma_i^j \gamma_k^l, \tag{23} \]

and the Fock matrix elements (\( \tilde{f}_p^q \)) are defined as

\[ \tilde{f}_p^q = \hat{h}_p^q + \sum_{ij} \tilde{v}_{pq}^{i'j'} \gamma_p^{i'j'}. \tag{24} \]

Note that the quantities \( \gamma_i^j \) and \( \gamma_k^l \) in Eqs. (23) and (24) are the untransformed 1- and 2-RDMs of the reference \( |\Psi_0\rangle \) defined as \( \gamma_q^p = \langle \Psi_0 | \hat{a}_q^p | \Psi_0 \rangle \) and \( \gamma_{pq}^{rs} = \langle \Psi_0 | \hat{a}_q^p \hat{a}_r^s | \Psi_0 \rangle \), 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 repulsion
integral tensor as a contraction of a three-index auxiliary tensor \( (B^Q_{pq}) \),

\[
\langle pq|rs \rangle \approx \sum_Q M^Q B^Q_{pq} B^Q_{qs}, \tag{25}
\]

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,

\[
\tilde{B}^Q_{p'q'} = \sum_{pq} B^Q_{pq} U^p_{p'} U^q_{q'}, \tag{26}
\]

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

Equations (22)–(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\(^{54}\) to implement sq-MR-LDSRG(2) by employing \( \tilde{H} \) (instead of \( \hat{H} \)) and removing terms involving \( \hat{A}_1 \).

As described in Ref. 54, we evaluate the commutator \( \hat{C}^{(k+1)} = \frac{1}{k+1} [\hat{C}^{(k)}_{1,2}, \hat{A}]_{0,1,2} \) in Eq. (12) using the following recursive system of equations

\[
\begin{align*}
\hat{O}^{(k+1)} & = \frac{1}{k+1} [\hat{C}^{(k)}_{1,2}, \hat{T}]_{0,1,2}, \\
\hat{C}^{(k+1)} & = \hat{O}^{(k+1)} + [\hat{O}^{(k+1)}]^{\dagger},
\end{align*}
\tag{27}
\]

where \( \hat{O}^{(k+1)} \) is an intermediate containing up to two-body components. The iteration starts from either \( \hat{C}^{(0)} = \hat{H} \), in traditional MR-LDSRG(2), or \( \hat{C}^{(0)} = \tilde{H} \) in the sequential version, optionally applying the NIVO approximation to the two-body intermediate tensors \( \hat{O}_2^{(k)} \), \( \hat{C}_2^{(k)} \) and \( \tilde{H}_2^{(k)} \) for \( k \geq 1 \).

### 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 the intermediate tensors generated during the evaluation of commutators. For example, consider the following contraction,

\[
O^i_{rs} \leftarrow \sum_{ab} \langle rs|ab \rangle t^i_{ab} \quad \forall i, j \in H, \forall r, s \in G, \tag{28}
\]
which is also found in the CCSD equations. In the DF case, Eq. (28) is written as two contractions involving auxiliary tensors,

\[ O_{rs}^{ij} \leftarrow \sum_{ab} \sum_{Q} M_{ar} P_{ab} B_{Q_{bs}}^{ij} - \sum_{ab} \sum_{Q} M_{ar} P_{ab} B_{Q_{bs}}^{ij}. \]  

The most efficient way to evaluate the first term of Eq. (29) is to introduce the intermediate tensor \( I_{arb} = \sum_{Q} B_{Q_{ar}}^{Q_{bs}} \) of size \( O(N^{2}N^{2}) \). To avoid storage of these large intermediates, it is common to evaluate Eq. (29) using a batched algorithm, whereby a slice of the tensor \( I_{arb} \) is computed and contracted on the fly with the amplitudes \( t_{ij}^{ab} \). To automate this optimization of the tensor contraction we have coded a generic batching algorithm in the tensor library Ambit.\(^{103}\) Whereas the Ambit code for the first term in Eq. (29) is written as

\[
0["ijrs"] += B["Qar"] * B["Qbs"] * t["ijab"];
\]

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

\[
0["ijrs"] += \text{batched("r",} B["Qar"] * B["Qbs"] * t["ijab"]);
\]

**Algorithm 1** The batched algorithm to compute \( C_{rs}^{ij} \leftarrow \sum_{ab} \sum_{Q} M_{ar} B_{Q_{bs}}^{ij} \).

1: Permute memory layout of \( C_{rs}^{ij} \) and \( B_{Q_{ar}}^{Q_{bs}} \) so that their \( r \)-subblocks \( C_{rs}^{ij}_{[r]} \) and \( B_{Q_{ar}}^{Q_{bs}}_{[r]} \) are contiguous in memory.
2: for each \( r = 1, 2, \ldots, N_{G} \) do
3: \( I_{arb} := \sum_{Q} B_{Q_{ar}}^{Q_{bs}} \)
4: \( C_{rs}^{ij}_{[r]} \leftarrow \sum_{ab} I_{arb} t_{ij}^{ab} \)  \( i, j \in H, r \in G \)
5: end for
6: Permute \( C_{rs}^{ij} \) back to the original memory layout.

### 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 computational bottleneck is forming the DSRG transformed Hamiltonian \( \bar{H} \). Timings for computing \( \bar{H} \) in the case of cyclobutadiene (see Sec. 4.2 for details) are summarized in Fig. 2. Detailed timings for the evaluation of \( \bar{H} \) using all the combinations of the approximations considered here are reported in the Supplementary Material.
The total timing ($t_{\text{tot}}$) for computing the transformed Hamiltonian in $n$ iterations is partitioned according to

$$t_{\text{tot}} = t_1 + t_2 + t_{\text{misc}},$$

(30)

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 $t_{\text{misc}}$ accounts for the cost to sort and accumulate the results of contractions with $\hat{T}_1$ and $\hat{T}_2$, as shown in the second line of Eq. (27). Figure 2 shows that the timing for the conventional 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 $\bar{H}$.

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

In general, we observe an increase in $t_2$ due to the extra cost to build two-body intermediates from the auxiliary tensors for methods combined with DF. 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 $[\hat{B}, \text{Eq. (26)}]$ 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, the method we recommend for large-scale computations.
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, C₂, N₂, and F₂. Specifically, we computed equilibrium bond lengths (rₑ), harmonic vibrational frequencies (ωₑ), anharmonicity constants (ωₑₓₑ), and dissociation energies (D₀) and compare those to experimental data taken from Ref. 106. The dissociation energy D₀ includes zero-point vibrational energy corrections that account for anharmonicity effects and is computed as D₀ = Dₑ − ωₑ/2 + ωₑₓₑ/4 (in a.u.), where Dₑ is the dissociation energy with respect to the bottom of the potential.¹⁰⁷ 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 15 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 $s = 0.5 \ h^{-2}$, as suggested by our previous work. All computations utilized the cc-pVQZ basis set 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 mixed flavor of the auxiliary basis sets. For CASSCF, the cc-pVQZ-JKFIT auxiliary basis set was used for H, B, C, N, O and F atoms, and the def2-QZVPP-JKFIT basis set was used for Li and Be atoms. The cc-pVQZ-RI basis set was applied to all atoms in DSRG computations.

Figure 3 and Table 1 report a comparison of second- and third-order DSRG multireference perturbation theory (DSRG-MRPT2/3), 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 1 show that MR-LDSRG(2) method is as accurate as CCSD(T) in predicting $r_e$, $\omega_e$, $\omega_{exe}$ and $D_0$, while it predicts $\omega_e$ that are of accuracy.
between that of 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 of static correlation leads to a more accurate treatment of 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 approximation of the MR-LDSRG(2) method, in Table 2 we report the mean absolute difference between properties computed with and without each approximation. The use of a sequential ansatz has a modest effect on all properties, with the largest mean absolute differences observed for \( \omega_e \) (1.6 cm\(^{-1}\)) and \( D_0 \) (0.3 kcal mol\(^{-1}\)). Nevertheless, the MAE with respect experimental results is nearly unchanged, if not slightly improving. 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 conventional MR-LDSRG(2). The only noticeable deviations are found for \( \omega_e \) (MAE 13.9 vs. 14.3 cm\(^{-1}\)) and \( D_0 \) (MAE 2.6 vs. 2.5 kcal mol\(^{-1}\)).

In this study, we also investigate the effect of combining the DF-sq-LDSRG(2)+NIVO method with truncation of the BCH expansion, i.e., terminating \( \bar{H}(s) = \tilde{H}(s) + \sum_{n=1}^{k} \frac{1}{n!} \tilde{C}(n)(s) \) at a given integer \( k \). The recursive evaluation of \( \tilde{H}(s) \) via Eq. (12) usually requires 10–12. Truncation of the BCH series to a few terms

---

### Table 1 Error statistics for the equilibrium bond lengths (\( r_e \), in pm), harmonic vibrational frequencies (\( \omega_e \), in cm\(^{-1}\)), anharmonicity constants \( \omega_{e-xe} \), in cm\(^{-1}\)), and dissociation energies \( (D_0) \), in kcal mol\(^{-1}\)) of the eight diatomic molecules computed using various MR-DSRG schemes. 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).

| \( r_e \) | Mean | MAE | SD | Max | Mean | MAE | SD | Max | Mean | MAE | SD | Max | Mean | MAE | SD | Max |
|-------|------|-----|----|-----|------|-----|----|-----|------|-----|----|-----|------|-----|----|-----|
| PT2   | 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|
| PT3   | 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|
| Com+ \* | 73   | 0.74| 0.49| 0.50| 0.49 | 0.48| 0.48| 0.48 | 0.49 | 0.48| 0.48 | 0.92 | 0.57|
| DF    | 28   | 1.96| 1.15| 1.15| 1.15 | 1.10| 1.10| 1.10 | 1.11 | 1.10| 1.09 | 2.13 | 1.36|

| \( \omega_e \) | Mean | MAE | SD | Max | Mean | MAE | SD | Max | Mean | MAE | SD | Max | Mean | MAE | SD | Max |
|----------------|------|-----|----|-----|------|-----|----|-----|------|-----|----|-----|------|-----|----|-----|
| PT2            | -11.2| 7.2 | 11.1| 11.1| 11.1 | 11.8| 11.9| 11.8 | 13.0 | 11.9| 11.9 | 54.9 | -0.1|
| PT3            | 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 |
| DF             | 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|

| \( \omega_{e-xe} \) | Mean | MAE | SD | Max | Mean | MAE | SD | Max | Mean | MAE | SD | Max | Mean | MAE | SD | Max |
|---------------------|------|-----|----|-----|------|-----|----|-----|------|-----|----|-----|------|-----|----|-----|
| PT2                 | 1.1  | 1.9 | 1.7 | 1.7 | 1.7  | 1.7 | 1.7 | 1.7  | 1.9  | 1.6 | 1.7 | 1.0 | 1.7  |
| PT3                 | 2.0  | 1.9 | 1.7 | 1.7 | 1.7  | 1.7 | 1.7 | 1.7  | 1.9  | 1.6 | 1.7 | 1.6 | 1.7  |
| DF                  | 2.4  | 2.5 | 2.3 | 2.1 | 2.3  | 2.3 | 2.3 | 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 | 4.7 | 4.7  | 5.9  | 4.4 | 4.8 | 6.2 | 6.2  |

| \( D_0 \) | Mean | MAE | SD | Max | Mean | MAE | SD | Max | Mean | MAE | SD | Max | Mean | MAE | SD | Max |
|-----------|------|-----|----|-----|------|-----|----|-----|------|-----|----|-----|------|-----|----|-----|
| PT2       | -5.0 | 0.1 | -0.3| -0.1| -0.0 | -0.6| -0.3| -0.5 | -1.0 | -0.5| -0.3 | -9.5 | -2.4|
| PT3       | 5.3  | 2.8 | 2.6 | 2.8 | 2.6  | 2.6 | 2.9 | 2.5  | 3.1  | 2.7 | 2.8 | 9.9 | 2.9  |
| DF        | 6.8  | 3.5 | 3.5 | 3.7 | 3.5  | 3.7 | 3.8 | 3.6  | 4.0  | 3.7 | 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  |

\* Computed using conventional four-index two-electron integrals.
Table 2 The mean absolute differences in predicting equilibrium bond lengths ($r_e$, in pm), harmonic vibrational frequencies ($\omega_e$, 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 between method pairs that differ by only one technique introduced in this report. Techniques include: sequential transformation (ST), density fitting (DF), NIVO approximation, and commutator truncation of the BCH expansion [comm($k$), $k = 2, 3, 4$]. All results were computed using the cc-pVQZ basis set and core orbitals were frozen in the MR-DSRG computations.

| Technique | MR-LDSRG(2) method pair | Mean absolute difference | $r_e$ | $\omega_e$ | $\omega_{e x e}$ | $D_0$ |
|-----------|--------------------------|--------------------------|------|---------|-------------|------|
| ST        | Original ST              |                          | 0.03 | 1.6     | 0.0         | 0.3  |
| DF        | ST/ST                   |                          | 0.03 | 1.5     | 0.3         | 0.2  |
| DF        | Original DF             |                          | 0.00 | 0.2     | 0.1         | 0.2  |
| ST        | DF/ST                   |                          | 0.00 | 0.2     | 0.1         | 0.3  |
| NIVO      | DF                      | DF/NIVO                  | 0.01 | 0.4     | 0.1         | 0.2  |
|           | DF/ST                   | DF/ST/NIVO               | 0.01 | 0.3     | 0.1         | 0.3  |
| comm(4)   | DF/ST/NIVO/comm(4)      |                          | 0.00 | 0.1     | 0.0         | 0.3  |
| 4th comm  | DF/ST/NIVO/comm(4)      | DF/ST/NIVO/comm(3)       | 0.01 | 0.8     | 0.1         | 0.4  |
| 3rd comm  | DF/ST/NIVO/comm(3)      | DF/ST/NIVO/comm(2)       | 0.04 | 4.3     | 0.3         | 0.6  |

may therefore introduce speedups of up to 3–4 times. In Table 1 and Table 2 we report statistics computed by approximating the BCH expansion up to 2, 3, and 4 commutators. The use of only two commutator introduces noticeable deviations with respect to experiments for $\omega_e$ and $D_0$. Compared to the full BCH series, this truncation level increases the MAE of $\omega_e$ and $D_0$ by 12.6 cm$^{-1}$ and 0.5 kcal mol$^{-1}$, respectively. The inclusion of the triply-nested commutator significantly reduces these deviations to only 0.1 cm$^{-1}$ and 0.2 kcal mol$^{-1}$, respectively. The four-fold commutator term yields $r_e$, $\omega_e$, and $\omega_{e x e}$ that are nearly identical to those from the untruncated BCH series, while the MAE of $D_0$ deviates only by 0.3 kcal mol$^{-1}$.

Since the error introduced by neglecting the four commutator term is smaller or comparable to the other approximations considered here, our results suggest that a BCH series truncated to three commutators may offer a good compromise between accuracy and speed.

4.2 Cyclobutadiene

Next, we consider the automerization reaction of cyclobutadiene (CBD, C$_4$H$_4$). We study the energy difference between the rectangular ($D_{2h}$) energy minimum and the square transition state ($D_{4h}$).\textsuperscript{111} This reaction is a challenging chemistry problem for both experiment and theory.\textsuperscript{111,112,112–117,117–123} 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 1.6–10 kcal mol$^{-1}$.\textsuperscript{112} In this work, we optimized 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 state-specific MRCC of Mukherjee and co-workers (Mk-MRCC)\textsuperscript{20,124} as implemented in Psi4.\textsuperscript{125}

To reduce computational cost, all MR-DSRG calculations performed two steps of the reference relaxation procedure discussed in Sec. 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\textsuperscript{-1} for absolute energies, 0.0001 Å for bond lengths, and 0.001° for bond angles. We applied a Tikhonov regularization denominator shift\textsuperscript{126} of 1 mE\textsubscript{h} in all Mk-MRCC calculations to guarantee convergence. 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. All computations utilized the cc-pV\textsubscript{X}X (X=D, T, Q, 5) basis set,\textsuperscript{108} and the corresponding cc-pV\textsubscript{X}X-JKFIT\textsuperscript{80} and cc-pV\textsubscript{X}X-RI\textsuperscript{110} 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 semi-canonical CASSCF orbitals.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{The 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\textsuperscript{126} of 1 mE\textsubscript{h} in all Mk-MRCC calculations to guarantee convergence.}
\end{figure}

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 Fig. 4, the $s$-dependency of the automerization barrier displays significantly different behavior for these two active spaces. In both cases, the predicted activation energies change significantly for small values of $s$ ($< 0.2 \ E_n^{-2}$), 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 \( [\alpha^{ij}_{ab}(s)] \) as a function of \( s \) for both the equilibrium and transition state geometries, as shown in Fig. 5. 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 \( \pi \) orbitals of CBD, and 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 \( \pi \) 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 Fig. 5. 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 S3 in the Supporting Information). We also performed computations using \( s = 0.5 \ E_h^{-2} \) (reported in Table S4) to verify that the automerization

\[ \text{Figure 5} \] The unrelaxed DF-sq-MR-LDSRG(2)+NIVO/cc-pVDZ double substitution amplitudes involving both alpha and beta electrons \( (\alpha^{ij}_{ab}) \) as a function of the flow parameter \( s \) for the rectangular equilibrium and the square transition state of cyclobutadiene.
Table 3 Automeration reaction barrier ($E_a$, in kcal mol$^{-1}$) and geometry parameters (bond lengths in Å, bond angles in degree) of cyclobutadiene. All DSRG computations used $s = 1.0 E_h^{-2}$. All computations employed the CAS(4e,4o) reference and core orbitals constructed from carbon 1s orbitals were frozen for MR-DSRG and MRCC computations. As such, $n_C = 8$ and $n_A = 4$, where $n_X = N_X/2 (X \in \{C, A, V, H, P, G\})$ is the number molecular spatial orbitals for space $X$.

| Method                        | $E_a$     | $C-C^a$ | $C=C^a$ | $C-H$ | $\angle C=C-H^b$ | $C-C$ | $C-H$ |
|-------------------------------|-----------|---------|---------|-------|-----------------|-------|-------|
|                              |           | cc-pVDZ | cc-pVTZ | cc-pVQZ | cc-pVQZ |        |       |
| CASSCF                        | 6.49      | 1.5502  | 1.3567  | 1.0790 | 134.87         | 1.4472| 1.0779|
| DF-MR-LDSRG(2)+NIVO          | 8.56      | 1.5769  | 1.3660  | 1.0945 | 134.92         | 1.4624| 1.0932|
| DF-sq-MR-LDSRG(2)+NIVO       | 8.62      | 1.5768  | 1.3659  | 1.0944 | 134.92         | 1.4623| 1.0931|
| 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|
|                              |           | cc-pVQZ |         |        |                 |       |       |
|                              | 7.44      | 1.5475  | 1.3471  | 1.0694 | 134.83         | 1.4409| 1.0683|
| DF-MR-LDSRG(2)+NIVO          | 9.87      | 1.5668  | 1.3488  | 1.0789 | 134.91         | 1.4483| 1.0775|
| DF-sq-MR-LDSRG(2)+NIVO       | 9.93      | 1.5666  | 1.3487  | 1.0788 | 134.91         | 1.4481| 1.0774|
| 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|
|                              |           | cc-pVQZ |         |        |                 |       |       |
|                              | 7.53      | 1.5467  | 1.3462  | 1.0689 | 134.84         | 1.4400| 1.0678|
| DF-MR-LDSRG(2)+NIVO          | 10.16     | 1.5634  | 1.3452  | 1.0782 | 134.96         | 1.4447| 1.0768|
| DF-sq-MR-LDSRG(2)+NIVO       | 10.21     | 1.5631  | 1.3451  | 1.0781 | 134.96         | 1.4446| 1.0766|
| 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|

$^a$ $C-C$ and $C=C$ refer to the longer and shorter carbon–carbon bonds, respectively.

$^b$ $\angle C=C–H$ is the bond angle between the $C–H$ bond and the longer $C–C$ bond.

$^c$ Based on the corresponding cc-pVQZ optimized geometries.

energies computed with different values of the flow parameter are consistent. In general, the difference in automeration 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 the energy barrier computed at the CASSCF and correlated levels shows that dynamical correlation is important in this system as it increases by about 1–3 kcal mol$^{-1}$. Our best estimate for the automeration 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 then 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. For instance, the MR-LDSRG(2) predicted C–C bond length at the $D_{4h}$ geometry 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 conventional 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,\textsuperscript{111,114,117,117–122} and the experimental range reported in Ref. 112. Using our new implementation, we were able to perform, for the first time, nonperturbative multireference computations on cyclobutadiene using the cc-pV5Z basis (580 correlated orbitals) on a single node with 128 GB of memory.

5. CONCLUSION

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 non-interacting virtual orbital (NIVO) operator 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 theory,\textsuperscript{127–129} 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 $N$ is the number of basis functions. 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 $O(N^3M)$, where $M$ is the number of auxiliary basis functions. 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. 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 support 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 functions. We demonstrate this point by studying the automerization reaction of cyclobutadiene using a quintuple-$\zeta$ basis set (584 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 multireference coupled cluster 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 nonperturbative methods (e.g., unitary coupled cluster theory) and downfolding schemes for classical-quantum hybrid algorithms.\(^{46}\)

**SUPPORTING INFORMATION**

Detailed timings and formal scaling required to evaluate the transformed Hamiltonian, error statistics for the eight diatomic molecules, cyclobutadiene automerization energies reported in the literature, and convergence of DF-sq-MR-LDSRG(2)+NIVO absolute energy with respect to flow parameter.

**ACKNOWLEDGMENTS**

This work was supported by the U.S. Department of Energy under Award No. DE-SC0016004, a Research Fellowship of the Alfred P. Sloan Foundation, and a Camille Dreyfus Teacher-Scholar Award.
REFERENCES

(1) (a) Andersson, K.; Malmqvist, P.-Å.; 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.; Pastore, M.; Cimiraglia, R. New perspectives in multireference perturbation theory: the n-electron valence state approach. Theor. Chem. Acc. 2007, 117, 743–754.

(7) Liu, B. Ab initio potential energy surface for linear H3. J. Chem. Phys. 1973, 58, 1925–1937.

(8) (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.

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

(10) 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.

(11) 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.

(12) 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. 1969, 14, 35–89.

(13) Jezierski, B.; Monkhorst, H. J. Coupled-cluster method for multideterminantal reference states. Phys. Rev. A 1981, 24, 1668–1681.

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

(15) Stolarczyk, L. Z.; Monkhorst, H. J. Coupled-cluster method in Fock space. I. General formalism. Phys. Rev. A 1985, 32, 725–742.

(16) 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.

(17) 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.
(18) (a) Piecuch, P.; Paldus, J. Orthogonally spin-adapted multi-reference Hilbert space coupled-cluster formalism: diagrammatic formulation. *Theor. Chim. Acta* 1992, 83, 69–103; (b) Paldus, J.; Piecuch, P.; Pylypow, L.; Jeziorski, B. Application of Hilbert-space coupled-cluster theory to simple (H₂)₂ model systems: Planar models. *Phys. Rev. A* 1993, 47, 2738–2782.

(19) Mášik, J.; Hubač, I. Multireference Brillouin-Wigner Coupled-Cluster Theory. Single-root approach. *Adv. Quantum Chem.* 1998, 31, 75–104.

(20) (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.

(21) 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.

(22) Hanrath, M. An exponential multireference wave-function Ansatz. *J. Chem. Phys.* 2005, 123, 084102.

(23) Evangelista, F. A.; Gauss, J. An orbital-invariant internally contracted multireference coupled cluster approach. *J. Chem. Phys.* 2011, 134, 114102.

(24) 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.

(25) Datta, D.; Kong, L.; Nooijen, M. A state-specific partially internally contracted multireference coupled cluster approach. *J. Chem. Phys.* 2011, 134, 214116.

(26) Chen, Z.; Hoffmann, M. R. Orbital-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.

(27) Hoffmann, M. R.; Simons, J. A unitary multiconfigurational coupled-cluster method: Theory and applications. *J. Chem. Phys.* 1988, 88, 993–1002.

(28) 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.

(29) 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.

(30) Kutzelnigg, W. Error analysis and improvements of coupled-cluster theory. *Theor. Chim. Acta* 1991, 80, 349–386.

(31) Mertins, F.; Schirmer, J. Algebraic propagator approaches and intermediate-state representations. I. The biorthogonal and unitary coupled-cluster methods. *Phys. Rev. A* 1996, 53, 2140–2152.

(32) Taube, A. G.; Bartlett, R. J. New perspectives on unitary coupled-cluster theory. *Int. J. Quantum Chem.* 2006, 106, 3393–3401.

(33) Harsha, G.; Shiozaki, T.; Scuseria, G. E. On the difference between variational and unitary coupled cluster theories. *J. Chem. Phys.* 2018, 148, 044107.

(34) 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* 2014, 4, 714.
(35) 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.

(36) 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.; Chiara, B.; Dunssworth, 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.

(37) 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* **2017**, *95*, 020501.

(38) 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.

(39) Ryabinkin, I. G.; Yen, T.-C.; Genin, S. N.; Izmaylov, A. F. Qubit Coupled Cluster Method: A Systematic Approach to Quantum Chemistry on a Quantum Computer. *J. Chem. Theory Comput.* **2018**, *14*, 6317–6326.

(40) 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.

(41) 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.

(42) Lee, J.; Huggins, W. J.; Head-Gordon, M.; Whaley, K. B. Generalized Unitary Coupled Cluster Wave functions for Quantum Computation. *J. Chem. Theory Comput.* **2018**, *15*, 311–324.

(43) Li, Y.; Hu, J.; Zhang, X. M.; Song, Z.; Yung, M.-H. Variational Quantum Simulation for Quantum Chemistry. *Adv. Theory Simul.* **2019**, *309*, 1800182.

(44) Kühn, M.; Zanker, S.; Deglmann, P.; Marthaler, M.; Weiß, H. Accuracy and Resource Estimations for Quantum Chemistry on a Near-term Quantum Computer. arXiv:1812.06814 [quant-ph] 2018.

(45) 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. arXiv:1902.10171v2 [quant-ph] 2019.

(46) 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. arXiv:1902.01553v2 [quant-ph] 2019.

(47) Evangelista, F. A. Alternative single-reference coupled cluster approaches for multireference problems: the simpler, the better. *J. Chem. Phys.* **2011**, *134*, 224102.

(48) Hanauer, M.; Köhn, A. Perturbative treatment of triple excitations in internally contracted multireference coupled cluster theory. *J. Chem. Phys.* **2012**, *136*, 204107.

(49) (a) Yanai, T.; Chan, G. K.-L. Canonical transformation theory for multireference problems. *J. Chem. Phys.* **2006**, *124*, 26.
(b) Yanai, T.; Chan, G. K.-L. Canonical transformation theory from extended normal ordering. *J. Chem. Phys.* 2007, *127*, 104107.

(50) Wegner, F. Flow-equations for Hamiltonians. *Ann. Phys.* 1994, *506*, 77–91.

(51) Kehrein, S. *The Flow Equation Approach to Many-Particle Systems*; Springer Berlin Heidelberg, 2006.

(52) Evangelista, F. A. A driven similarity renormalization group approach to quantum many-body problems. *J. Chem. Phys.* 2014, *141*, 054109.

(53) Li, C.; Evangelista, F. A. Multireference Driven Similarity Renormalization Group: A Second-Order Perturbative Analysis. *J. Chem. Theory Comput.* 2015, *11*, 2097–2108.

(54) (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.

(55) Li, C.; Evangelista, F. A. Driven similarity renormalization group for excited states: A state-averaged perturbation theory. *J. Chem. Phys.* 2018, *148*, 124106.

(56) Li, C.; Evangelista, F. A. Multireference Theories of Electron Correlation Based on the Driven Similarity Renormalization Group. *Annu. Rev. Phys. Chem.* 2019, 70, in press.

(57) 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.

(58) 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.

(59) 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.

(60) Evangelista, F. A. Perspective: Multireference coupled cluster theories of dynamical electron correlation. *J. Chem. Phys.* 2018, *149*, 030901.

(61) (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.

(62) 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.

(63) Evangelisti, S.; Daudèy, J. P.; Malrieu, J. P. Qualitative intruder-state problems in effective Hamiltonian theory and their solution through intermediate Hamiltonians. *Phys. Rev. A* 1987, *35*, 4930–4941.

(64) Zarrabian, S.; Laidig, W. D.; Bartlett, R. J. Convergence properties of multireference many-body perturbation theory. *Phys. Rev. A* 1990, *41*, 4711–4720.

(65) (a) Kowalski, K.; Piecuch, P. Complete set of solutions of multireference coupled-cluster equations: The state-universal formalism. *Phys. Rev. A* 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.
(66) 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.

(67) 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.

(68) Whitten, J. L. Coulombic potential energy integrals and approximations. *J. Chem. Phys.* 1973, 58, 4496–4501.

(69) 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.

(70) Kendall, R. A.; Fruchtl, 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.

(71) 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.

(72) Røeggen, I.; Wisløff-Nilssen, E. On the Beebe-Linderberg two-electron integral approximation. *Chem. Phys. Lett.* 1986, 132, 154–160.

(73) Koch, H.; Sánchez de Merás, A.; Pedersen, T. B. Reduced scaling in electronic structure calculations using Cholesky decompositions. *J. Chem. Phys.* 2003, 118, 9481–9484.

(74) Boman, L.; Koch, H.; de Merás, A. S. Method specific Cholesky decomposition: Coulomb and exchange energies. *J. Chem. Phys.* 2008, 129, 134107.

(75) 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.

(76) 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.

(77) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* 1993, 213, 514–518.

(78) Feyereisen, M.; Fitzgerald, G.; Komormicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.* 1993, 213, 514–518.

(79) Weigend, F. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* 2002, 4, 4285–4291.

(80) 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.

(81) 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.

(82) 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.

(84) 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.

(85) 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.

(86) 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.

(87) 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.

(88) Pedersen, T. B.; de Merás, M. J. S.; 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.

(89) 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.

(90) Boström, J.; Pitoňák, M.; Aquilante, F.; Neogrady, P.; Pedersen, T. B.; Lindh, R. Coupled Cluster and Möller-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.

(91) 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.

(92) 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.

(93) Qiu, Y.; Henderson, T. M.; Zhao, J.; Scuseria, G. E. Projected coupled cluster theory. *J. Chem. Phys.* **2017**, *147*, 064111.

(94) Huntington, L. M. J.; Nooijen, M. pCCSD: parameterized coupled-cluster theory with single and double excitations. *J. Chem. Phys.* **2010**, *133*, 184109.

(95) Evangelista, F. A.; Gauss, J. On the approximation of the similarity-transformed Hamiltonian in single-reference and multireference coupled cluster theory. *Chem. Phys.* **2012**, *401*, 27–35.

(96) (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.

(97) 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.

(98) Koch, S.; Kutzelnigg, W. Comparison of CEPA and CP-MET methods. *Theor. Chim. Acta* **1981**, *59*, 387–411.

(99) Bartlett, R. J.; Musial, 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.
(100) 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.

(101) (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.

(102) Forte, a suite of quantum chemistry methods for strongly correlated electrons. For current version see https://github.com/evangelistalab/forte, 2019.

(103) 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.

(104) (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.

(105) 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.

(106) Huber, K. P.; Herzberg, G. Molecular Spectra and Molecular Structure; Springer: Boston, MA, 1979.

(107) Irikura, K. K. Experimental Vibrational Zero-Point Energies: Diatomic Molecules. J. Phys. Chem. Ref. Data 2007, 36, 389–397.

(108) (a) Dunning Jr, T. H. 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.

(109) (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.

(110) (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.

(111) Shen, J.; Piecuń, 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.

(112) Whitman, D. W.; Carpenter, B. K. Limits on the activation parameters for automerization of cyclobutadiene-1,2-d2. *J. Am. Chem. Soc.* **1982**, *104*, 6473–6474.

(113) 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.

(114) Balková, A.; Bartlett, R. J. A multireference coupled-cluster study of the ground state and lowest excited states of cyclobuta-diene. *J. Chem. Phys.* **1998**, *101*, 8972–8987.

(115) 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.

(116) 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.* **2003**, *120*, 175–185.

(117) Eckert-Maksić, M.; Vazdar, M.; Barbati, 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.

(118) 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.

(119) 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.

(120) Li, X.; Paldus, J. Accounting for the exact degeneracy and quasidegeneracy in the automerization of cyclobutadiene via multireference coupled-cluster methods. *J. Chem. Phys.* **2009**, *131*, 114103.

(121) Bhaskaran-Nair, K.; Demel, O.; Pittner, J. Multireference state-specific Mukherjee’s coupled cluster method with noniterative triexcitations. *J. Chem. Phys.* **2008**, *129*, 184105.

(122) 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.

(123) Wu, J. I.-C.; Mo, Y.; Evangelista, F. A.; Schleyer, P. v. R. Is cyclobutadiene really highly destabilized by antiaromaticity? *Chem. Commun.* **2012**, *48*, 8437–8439.

(124) (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.

(125) (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.

(126) Taube, A. G.; Bartlett, R. J. Rethinking linearized coupled-cluster theory. *J. Chem. Phys.* 2009, 130, 144112.

(127) 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.

(128) 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.

(129) 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.* 1998, 97, 5554–5559.