Annual Review of Physical Chemistry

Multireference Theories of Electron Correlation Based on the Driven Similarity Renormalization Group

Chenyang Li and Francesco A. Evangelista

Department of Chemistry and Cherry L. Emerson Center for Scientific Computation, Emory University, Atlanta, Georgia 30322, USA; emails: cli62@emory.edu, francesco.evangelista@emory.edu

Annu. Rev. Phys. Chem. 2019. 70:275–303

The *Annual Review of Physical Chemistry* is online at physchem.annualreviews.org

https://doi.org/10.1146/annurev-physchem-042018-052416

Keywords

electronic structure theory, electron correlation, many-body perturbation theory, coupled-cluster theory, similarity renormalization group, driven similarity renormalization group

Abstract

The driven similarity renormalization group (DSRG) provides an alternative way to address the intruder state problem in quantum chemistry. In this review, we discuss recent developments of multireference methods based on the DSRG. We provide a pedagogical introduction to the DSRG and its various extensions and discuss its formal properties in great detail. In addition, we report several illustrative applications of the DSRG to molecular systems.

1. INTRODUCTION

Static correlation: strong mixing of electronic configurations

Dynamic correlation: short-range structure of the wave function and dispersion interactions

CAS: complete active space

MRCC: multireference coupled-cluster theory

SRG: similarity renormalization group

DSRG: driven similarity renormalization group

Multireference (MR) theories of electron correlation aim to accurately and efficiently describe molecular systems that are not well described by a single electron configuration. Strong mixing of degenerate or near-degenerate electron configurations (static correlation) occurs in many important areas of chemistry, including bond-breaking processes, open-shell species with orbital degeneracies, electronically excited states, and complexes of transition metals and rare earths in which localized spins are coupled. The theoretical description of near-degenerate states requires a balanced treatment of static and dynamic correlation effects and remains challenging for modern electronic structure methods. The success of conventional single-reference formalisms like Kohn–Sham density functional theory (1) and coupled-cluster (CC) theory (2–5) has motivated the search for MR generalizations that do not break down when static correlation plays an important role (see, e.g., References 6–9). An important class of MR methods are those based on complete active space (CAS) wave functions (10). In these approaches, a CAS wave function is used to account for static correlation effects, and the remaining dynamical correlation is treated at the level of perturbation theory (PT) (11–13), configuration interaction (CI) (14, 15), or CC theory (15–18).

Despite considerable efforts, MRCC methods have not reached the level of sophistication and applicability achieved by other quantum chemistry approaches. There are two major limitations that prevent general applications of MRCC methods: (*a*) numerical instabilities connected to the intruder state problem (19, 20) and (*b*) restrictions on the number of active orbitals in the CAS reference. The intruder state problem occurs naturally in effective Hamiltonian theory (21) and is due to small energy denominators that arise from the interactions between perturber states that are near degenerate with the reference. Intruders lead to characteristic divergences in the correlation energy (i.e., spikes in potential energy surfaces). The question of how to best deal with intruder states in MRCC theories and other nonperturbative schemes still remains open. Conventional solutions to the problem of intruders are successful in regularizing PT (22–26) but cannot be directly generalized to MRCC methods. As all MRCC approaches appear to be intrinsically predisposed to numerical problems, one may wonder if a more general framework with numerically robust equations could be proposed.

Renormalization techniques developed in physics to deal with infinities that arise in quantum field theory offer one solution to the intruder state problem. The SRG (similarity renormalization group) is a many-body formalism developed contemporarily by Wegner (27) and Głazek & Wilson (28) for applications in quantum field theory and condensed matter physics. In the SRG approach, the Hamiltonian is diagonalized via a continuous unitary transformation that depends on an energy cutoff, Λ. The SRG transformation satisfies an important property: It gradually folds correlation effects due to excited configurations into the Hamiltonian, avoiding those states with energy denominators smaller than Λ. Hence, it provides a way to systematically build in correlation effects while avoiding intruder states caused by small denominators. More recently, the SRG formalism was applied in chemistry by White (29), and important developments and applications have been reported in nuclear theory (30–35).

In this review, we summarize our recent development of MR electronic structure methods based on the driven SRG (DSRG) (36), a many-body formalism that combines elements of MRCC theory with the SRG. The renormalization group structure of the DSRG approach leads to MR theories that avoid the intruder state problem. In addition, the many-body framework used to derive the DSRG equations leads to energy and amplitude expressions that are simpler and less expensive than those of other MRCC approaches. The reader is assumed to be familiar with CC theory and many-body techniques (e.g., second quantization, normal-ordered products, Wick's theorem) at the level discussed in References 37–39.

Since the literature on MR theories is a vast subject, we focus exclusively on reviewing the DSRG and other methods related to it. Several excellent reviews exist that cover a wide range of MR methods (15–18). We also note that beside CAS-based methods, several successful CC formalisms have been developed that rely on a single-reference framework (40–48).

2. THEORY

2.1. Preliminaries

In this section, we introduce some basic notation used throughout this review. We start by considering an orthonormal spin orbital basis, $\mathbf{G} \equiv \{\phi^p(x), p = 1, 2, \ldots, N\}$, where $x = (\mathbf{r}, \sigma)$ collects spatial ($\mathbf{r}$) and spin (σ) coordinates. The Hamiltonian written in terms of second quantized creation ($\hat{a}^\dagger$) and annihilation ($\hat{a}$) operators is given by

$$\hat{H} = \sum_{pq}^{\mathbf{G}} h_p^q \hat{a}_q^p + \frac{1}{4} \sum_{pqrs}^{\mathbf{G}} v_{pq}^{rs} \hat{a}_{rs}^{pq}, \tag{1.}$$

where $h_p^q = \int \phi_p^*(x_1) \hat{h} \phi_q(x_1)\, \mathrm{d}x_1$ are one-electron integrals, and $v_{pq}^{rs} = \langle pq \| rs \rangle = \langle pq|rs\rangle - \langle pq|sr\rangle$ are antisymmetrized two-electron integrals with $\langle pq|rs\rangle = \int \phi_p^*(x_1)\phi_q^*(x_2) r_{12}^{-1} \phi_r(x_1)\phi_s(x_2)\, \mathrm{d}x_1 \mathrm{d}x_2$. Here we also employ the compact notation $\hat{a}_{rs\ldots}^{pq\cdots} = \hat{a}_p^\dagger \hat{a}_q^\dagger \cdots \hat{a}_s \hat{a}_r$ to represent products of second quantized operators. A product of operators normal ordered with respect to a reference Fermi vacuum $|\Psi_0\rangle$ is indicated with braces ($\{\hat{a}_{rs\ldots}^{pq\cdots}\}$).

2.2. The Similarity Renormalization Group

To introduce and motivate the DSRG, we start by reviewing the SRG, closely following the formulations of Wegner (27) and Tsukiyama, Bogner, and Schwenk (31). These works assume a single determinant reference Φ with respect to which all operators are normal ordered. Following a common convention (38), we indicate core (**C**, doubly occupied) and virtual (**V**, unoccupied) spin orbitals of Φ with the indices $i, j, k, \ldots$ and $a, b, c, \ldots$, respectively. At the basis of the SRG formalism is a continuous unitary similarity transformation of the Hamiltonian,

$$\hat{H} \to \hat{H}(s) = \hat{U}(s) \hat{H} \hat{U}^\dagger(s), \quad s \in [0, \infty), \tag{2.}$$

that brings the transformed Hamiltonian $\hat{H}(s)$ to a more diagonal (or block-diagonal) form. The quantity s, defined in the range $[0, \infty)$, is a time-like parameter that controls the extent to which $\hat{H}(s)$ is diagonal, and the transformation is chosen to satisfy the condition $\hat{U}^\dagger(0) = \hat{1}$.

The goal of the SRG transformation (Equation 2) is to reduce the magnitude of certain operator components of $\hat{H}(s)$. There is considerable freedom in the way one selects which elements of $\hat{H}(s)$ to suppress, and the choice depends on the intended application of the SRG. For example, to describe ground state correlation effects, we may demand that in the limit of $s \to \infty$ the SRG transformation should zero the components of $\hat{H}(s)$ that couple the reference Φ to all of its excited determinants, $\Phi_{ij\ldots}^{ab\cdots}$; in other words, we may require that

$$\lim_{s\to\infty} \langle \Phi_{ij\ldots}^{ab\cdots} | \hat{H}(s) | \Phi \rangle = \lim_{s\to\infty} \langle \Phi | \hat{H}(s) | \Phi_{ij\ldots}^{ab\cdots} \rangle = 0. \tag{3.}$$

When Equation 3 is satisfied, it follows that Φ is an eigenfunction of $\hat{H}(s)$ with corresponding eigenvalue equal to $E_0(s) = \langle \Phi | \hat{H}(s) | \Phi \rangle$. In discussing the SRG, it is convenient to introduce a notation that allows us to talk about the components of an operator that we want to suppress via the SRG transformation. To this end, we partition $\hat{H}(s)$ into a diagonal [$\hat{H}_\mathrm{d}(s)$] and an off-diagonal

$[\hat{H}_{\rm od}(s)]$ part, $\hat{H}(s) = \hat{H}_{\rm d}(s) + \hat{H}_{\rm od}(s)$. All the operators that couple Φ and $\Phi^{ab\cdots}_{ij\cdots}$ are collected in the off-diagonal part, and the condition expressed by Equation 3 is written as $\lim_{s\to\infty} \hat{H}_{\rm od}(s) = 0$.

Møller–Plesset denominator: $\Delta^{ij\cdots}_{ab\cdots} = \epsilon_i + \epsilon_j + \cdots - \epsilon_a - \epsilon_b - \cdots$

The operator $\hat{U}(s)$ that satisfies Equation 3 is a priori unknown. Taking the derivative with respect to s of the SRG Hamiltonian (Equation 2), it is possible to write an ordinary differential equation that determines $\hat{H}(s)$,

$$\frac{{\rm d}\hat{H}(s)}{{\rm d}s} = [\hat{\eta}(s), \hat{H}(s)], \quad \hat{H}(0) = \hat{H}, \tag{4.}$$

where the quantity $\hat{\eta}(s) = \frac{{\rm d}\hat{U}(s)}{{\rm d}s}\hat{U}^\dagger(s)$ is the anti-Hermitian generator of the transformation. Equation 4 is known as the SRG flow equation or an isospectral flow, as it preserves the eigenvalues of $\hat{H}(s)$. The utility of this differential formulation is that, by a suitable parameterization of $\hat{\eta}(s)$, it is possible to integrate Equation 4 and find a numerical solution $\hat{H}(s)$ that satisfies the condition $\hat{H}_{\rm od}(s) = 0$ without having to explicitly construct $\hat{U}(s)$. As shown by Wegner, the generator $\hat{\eta}^{\rm W}(s) = [\hat{H}_{\rm d}(s), \hat{H}(s)]$ reduces the magnitude of the Hamiltonian couplings in a monotonic way, in the sense that $\frac{\rm d}{{\rm d}s}{\rm Tr}[\hat{H}^2_{\rm od}(s)] \leq 0$ (see Reference 49). The SRG flow equation with the Wegner generator has the structure $\frac{\rm d}{{\rm d}s}\hat{H}(s) = [[\hat{H}_{\rm d}(s), \hat{H}(s)], \hat{H}(s)]$, which is also known in mathematics as a double bracket flow equation and realizes a continuous version of the Jacobi eigenvalue algorithm (50–52).

In application to quantum chemistry, the Hamiltonian at $s = 0$ is a sum of one- and two-body operators (Equation 1). However, as the SRG flow equation is integrated, higher-body operators arise from the commutators that enter into the flow equation, e.g., $[\hat{\eta}(s), \hat{H}(s)]$. When written in normal-ordered form with respect to Φ, the SRG Hamiltonian is a many-body operator with tensor elements that depend on the flow parameter s,

$$\hat{H}(s) = E_0(s) + \sum_{pq}^{\rm G} f^q_p(s)\{\hat{a}^p_q\} + \frac{1}{4}\sum_{pqrs}^{\rm G} v^{rs}_{pq}(s)\{\hat{a}^{pq}_{rs}\} + \frac{1}{36}\sum_{pqrstu}^{\rm G} w^{stu}_{pqr}(s)\{\hat{a}^{pqr}_{stu}\} + \ldots. \tag{5.}$$

To keep the computational cost of integrating the flow equation affordable, we need to truncate the rank of operators and organize the SRG equations to minimize the truncation error. A well established practice involves truncating $\hat{H}(s)$ and all commutators to a given operator rank, and working with operators normal ordered with respect to the reference Φ (31). Normal ordering is essentially a redefinition of products of bare operators such that the new operators describe fluctuations with respect to Φ. In the context of the SRG, this redefinition is important because it brings contributions from higher-rank bare operators into lower-rank operators and reduces truncation errors.

The most important feature of the SRG transformation is the separation of energy scales in the Hamiltonian. To illustrate this point, we consider a perturbative analysis of the SRG, a case in which the flow equation can be integrated analytically (34, 36). Assuming a one-body diagonal zeroth-order Hamiltonian, it is possible to derive closed expressions for the first-order off-diagonal components of the Hamiltonian $[v^{ij,(1)}_{ab}(s)]$,

$$v^{ij,(1)}_{ab}(s) = \langle ij\|ab\rangle e^{-s(\Delta^{ij}_{ab})^2}, \tag{6.}$$

where $\Delta^{ij}_{ab} = \epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b$ is a standard Møller–Plesset denominator and $\epsilon_p = f^p_p(0)$ are canonical orbital energies. Equation 6 shows that the off-diagonal part of $\hat{H}(s)$ decays exponentially as a function of s with an exponent that is determined by the square of Δ^{ij}_{ab}. A dimensional analysis shows that $[s] = ({\rm energy})^{-2}$, and so we can define an energy cutoff $\Lambda = s^{-1/2}$ and express the exponent as $(\Delta^{ab}_{ij}/\Lambda)^2$. For a fixed value of s, if a denominator is large ($|\Delta^{ij}_{ab}| \gg \Lambda$), then the

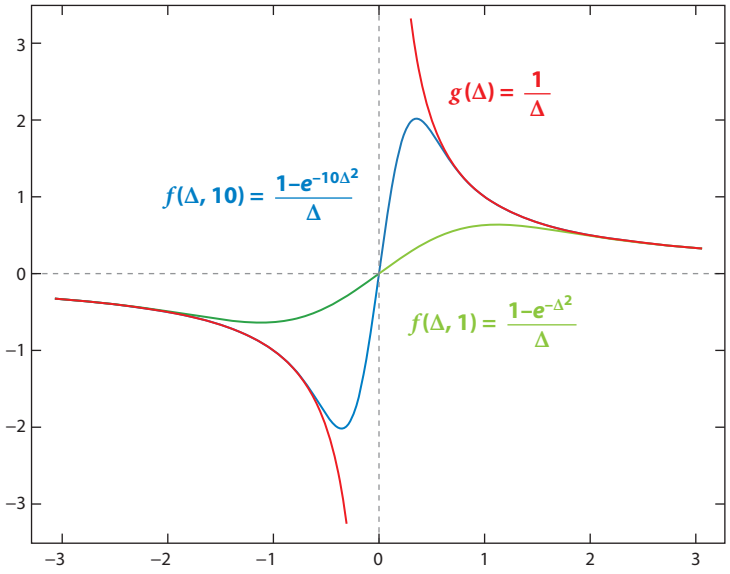

Figure 1

Comparison of the conventional perturbation theory resolvent [$g(\Delta)$] and the the driven similarity renormalization group (DSRG) regularized resolvent [$f(\Delta, s)$] for different values of the flow parameter (s).

corresponding $v_{ab}^{ij,(1)}(s)$ is approximately zero, i.e., an excitation with a large denominator is effectively decoupled from Φ. On the contrary, for small denominators ($|\Delta_{ab}^{ij}| \ll \Lambda$), the off-diagonal part is approximately unchanged, $v_{ab}^{ij,(1)}(s) \approx \langle ij \| ab \rangle$. This result is often summarized by saying that the SRG transformation integrates out the high-energy degrees of freedom (those larger than Λ) of the Hamiltonian.

At second order, the expression of the correlation energy [$E_0^{(2)}(s)$] is analogous to that of the conventional Møller–Plesset PT (MP2), except for a multiplicative factor that regularizes the denominator,

$$E_0^{(2)}(s) = \frac{1}{4} \sum_{ij}^{\mathrm{C}} \sum_{ab}^{\mathrm{V}} \frac{|\langle ij \| ab \rangle|^2}{\Delta_{ab}^{ij}} \left[1 - e^{-2s(\Delta_{ab}^{ij})^2} \right]. \qquad 7.$$

Equation 7 does not diverge when one or more denominators go to zero because the regularized inverse denominator $f(\Delta, s) = (1 - e^{-s\Delta^2})/\Delta$ that enters this expression is always finite. In **Figure 1**, we plot the function $f(\Delta, s)$ for two values of s (1 and 10 E_{h}^{-2}) and compare it to the conventional inverse denominator $g(\Delta) = 1/\Delta$. For any finite value of s, when Δ approaches zero, $f(\Delta, s)$ smoothly goes to zero, while, as expected, $g(\Delta)$ has a singularity at $\Delta = 0$. This observation suggests that one way to avoid the problem of intruders is to integrate the SRG equations up to a finite value of s so that excited configurations with small denominators are not decoupled from the reference Φ. Note, however, that $\lim_{s\to\infty} f(\Delta, s)$ is not equal to $g(\Delta)$.

2.3. The Driven Similarity Renormalization Group

Our development of the DSRG has been motivated by the desire to realize the SRG transformation using a formalism similar to that of CC theory (37–39). In the DSRG, we rewrite the SRG unitary operator $\hat{U}^\dagger(s)$ in terms of an exponential operator $e^{\hat{A}}(s)$, where $\hat{A}(s)$ is an s-dependent

anti-Hermitian operator. The DSRG-transformed Hamiltonian may be written as

$$\bar{H}(s) = e^{-\hat{A}(s)} \hat{H} e^{\hat{A}(s)}, \tag{8.}$$

where we have used the notation $\bar{H}(s)$ to differentiate it from the SRG Hamiltonian, $\hat{H}(s)$. The DSRG transformation is isospectral and canonical, with the latter meaning that the transformed operators $\bar{a}_p^\dagger = e^{-\hat{A}(s)} \hat{a}_p^\dagger e^{\hat{A}(s)}$ and $\bar{a}_q = e^{-\hat{A}(s)} \hat{a}_q e^{\hat{A}(s)}$ satisfy the canonical anticommutator property, i.e., $[\bar{a}_p^\dagger, \bar{a}_q]_+ = [\hat{a}_p^\dagger, \hat{a}_q]_+ = \delta_q^p$ (53). The exponential unitary transformation has a long history (54–56). In quantum chemistry, it is at the basis of unitary CC theory (57–63), canonical diagonalization (29), canonical transformation theory (64–66), and the anti-Hermitian contracted Schrödinger equation method (67, 68). In condensed matter physics, it is often called the Schrieffer–Wolff transformation (69, 70).

Following unitary CC theory (57–63, 71), $\hat{A}(s)$ is written in terms of the standard cluster operator $\hat{T}(s)$ as $\hat{A}(s) = \hat{T}(s) - \hat{T}^\dagger(s)$, where $\hat{T}(s)$ is a sum of operators $\hat{T}_k(s)$, with k ranging from one to the total number of electrons (n),

$$\hat{T}(s) = \hat{T}_1(s) + \hat{T}_2(s) + \cdots + \hat{T}_n(s). \tag{9.}$$

A generic k-body component $\hat{T}_k(s)$ is defined as

$$\hat{T}_k(s) = \frac{1}{(k!)^2} \sum_{ij\cdots}^{\mathbf{C}} \sum_{ab\cdots}^{\mathbf{V}} t_{ab\cdots}^{ij\cdots}(s) \{\hat{a}_{ij\cdots}^{ab\cdots}\}. \tag{10.}$$

The cluster amplitudes $[t_{ab\cdots}^{ij\cdots}(s)]$ that enter the definition of $\hat{T}_k(s)$ are s-dependent and antisymmetric with respect to the separate permutation of upper and lower indices.

In the DSRG approach, we aim to identify the operator $\hat{A}(s)$ such that the off-diagonal parts of the DSRG and SRG Hamiltonians match for all values of s, that is, we require that

$$\left[e^{-\hat{A}(s)} \hat{H} e^{\hat{A}(s)} \right]_{\text{od}} = \left[\hat{U}(s) \hat{H} \hat{U}^\dagger(s) \right]_{\text{od}}, \quad s \in [0, \infty). \tag{11.}$$

Formally, we can identify the right-hand side of Equation 11 with a source operator $\hat{R}(s) = [\hat{U}(s)\hat{H}\hat{U}^\dagger(s)]_{\text{od}}$ and write an equation for $\hat{A}(s)$ of the form

$$\left[e^{-\hat{A}(s)} \hat{H} e^{\hat{A}(s)} \right]_{\text{od}} = \hat{R}(s). \tag{12.}$$

The DSRG equation (Equation 12) is to be interpreted as an operator equation. If we define the residual operator $\hat{\Omega}(s) = [e^{-\hat{A}(s)} \hat{H} e^{\hat{A}(s)}]_{\text{od}} - \hat{R}(s)$ and expand it in its many-body components $\hat{\Omega}(s) = \hat{\Omega}_1(s) + \cdots + \hat{\Omega}_n(s)$ with

$$\hat{\Omega}_k(s) = \frac{1}{(k!)^2} \sum_{ij\cdots}^{\mathbf{C}} \sum_{ab\cdots}^{\mathbf{V}} \omega_{ab\cdots}^{ij\cdots}(s) \{\hat{a}_{ij\cdots}^{ab\cdots}\} + \text{h.c.}, \tag{13.}$$

where "h.c." stands for the Hermitian conjugate of the first term, then the DSRG equations consist of the following set of conditions for all combinations of indices,

$$\omega_{ab\cdots}^{ij\cdots}(s) = 0. \tag{14.}$$

Since $\hat{U}^\dagger(s)$ and $\hat{R}(s)$ are not known, we seek an approximate parameterization of $\hat{R}(s)$ that reproduces the SRG separation of energy scales. To solve this problem, we have proposed to derive $\hat{R}(s)$ by matching the analytical expressions for the first-order SRG and DSRG Hamiltonians (36).

The resulting source operator is expressed as a sum of many-body operators $\hat{R}(s) = \hat{R}_1(s) + \cdots + \hat{R}_n(s)$, with the k-body component defined as

$$\hat{R}_k(s) = \frac{1}{(k!)^2} \sum_{ij\cdots}^{\mathbf{C}} \sum_{ab\cdots}^{\mathbf{V}} r_{ab\cdots}^{ij\cdots}(s)\{\hat{a}_{ij\cdots}^{ab\cdots}\} + \text{h.c.}, \tag{15.}$$

and the coefficients $r_{ab\cdots}^{ij\cdots}(s)$ expressed in terms of $\bar{H}_{\text{od}}(s)$, the cluster amplitudes, and a regularization factor

$$r_{ab\cdots}^{ij\cdots}(s) = \left[\bar{H}_{ab\cdots}^{ij\cdots}(s) + t_{ab\cdots}^{ij\cdots}(s)\Delta_{ab\cdots}^{ij\cdots}\right] e^{-s(\Delta_{ab\cdots}^{ij\cdots})^2}. \tag{16.}$$

The source operator defined by Equation 16 was derived assuming a diagonal Fock operator (canonical orbitals). When working with noncanonical orbitals, a more general formulation is necessary (see Section 3.1).

In **Figure 2*a***, we plot the evolution of the DSRG transformed Hamiltonian driven by the source operator reported in Equation 16. As s increases, the off-diagonal elements of $\bar{H}(s)$, $\bar{H}_{ij}^{ab}(s)$ and $\bar{H}_{ab}^{ij}(s)$, contained in the [**CC**]/[**VV**] blocks are gradually suppressed. **Figure 2*a*** also illustrates that the DSRG transformation performs a separation of energy scales, since it starts by zeroing elements with large Møller–Plesset denominators and proceeds to those with small denominators.

2.4. Multireference Driven Similarity Renormalization Group

The DSRG formalism introduced in Section 2.3 may be easily generalized to the case of a multideterminantal reference. For convenience, we focus on CAS references and, therefore, assume that the spin orbitals are partitioned into core (**C**), active (**A**, partially occupied), and virtual (**V**) subsets of dimension $N_{\mathbf{C}}$, $N_{\mathbf{A}}$, and $N_{\mathbf{V}}$, respectively. We also introduce the hole ($\mathbf{H} = \mathbf{C} \cup \mathbf{A}$) and particle ($\mathbf{P} = \mathbf{A} \cup \mathbf{V}$) composite sets of size $N_{\mathbf{H}} = N_{\mathbf{C}} + N_{\mathbf{A}}$ and $N_{\mathbf{P}} = N_{\mathbf{A}} + N_{\mathbf{V}}$, respectively. These orbital spaces are illustrated in **Figure 3**.

With this partition, one forms the model space of determinants $M_0 = \{\Phi^\mu, \mu = 1, 2, \ldots, d\}$ obtained by distributing a given number of active electrons in the active spin orbitals. The reference wave function $|\Psi_0\rangle$ is assumed to be a linear combination of model space determinants,

$$|\Psi_0\rangle = \sum_{\mu=1}^{d} c_\mu |\Phi^\mu\rangle, \tag{17.}$$

where the coefficients c_μ are initially determined by a CAS CI (CASCI) or a CAS self-consistent field (CASSCF) procedure but may be relaxed in the presence of dynamical correlation. We further assume that Ψ_0 is normalized, i.e., $\langle\Psi_0|\Psi_0\rangle = 1$.

The MR generalization of the DSRG formalism requires a redefinition of normal-ordered products of second-quantized operators. Here, we adopt the generalized normal ordering (GNO) of Mukherjee & Kutzelnigg (72–75). To discuss this formalism, we first introduce the reduced density matrices (RDMs) and density cumulants of Ψ_0. A generic k-particle RDM is defined as the expectation value $\gamma_{rs\cdots}^{pq\cdots} = \langle\Psi_0|\hat{a}_{rs\cdots}^{pq\cdots}|\Psi_0\rangle$, where $\hat{a}_{rs\cdots}^{pq\cdots}$ contains k creation and k annihilation operators. The corresponding density cumulant $\lambda_{rs\cdots}^{pq\cdots}$ is the connected part of $\gamma_{rs\cdots}^{pq\cdots}$ (see References 76–78). For example, the two- and three-body density cumulants are related to the RDMs via the following conditions,

$$\lambda_{rs}^{pq} = \gamma_{rs}^{pq} - \mathcal{A}(\gamma_r^p \gamma_s^q), \qquad \lambda_{stu}^{pqr} = \gamma_{stu}^{pqr} - \mathcal{A}(\gamma_s^p \lambda_{tu}^{qr}) - \mathcal{A}(\gamma_s^p \gamma_t^q \gamma_u^r), \tag{18.}$$

where $\mathcal{A}$ indicates a sum over all unique, separate permutations of upper and lower indices times a parity factor (-1 for each index transposition), e.g., $\mathcal{A}(\gamma_r^p \gamma_s^q) = \gamma_r^p \gamma_s^q - \gamma_s^p \gamma_r^q$.

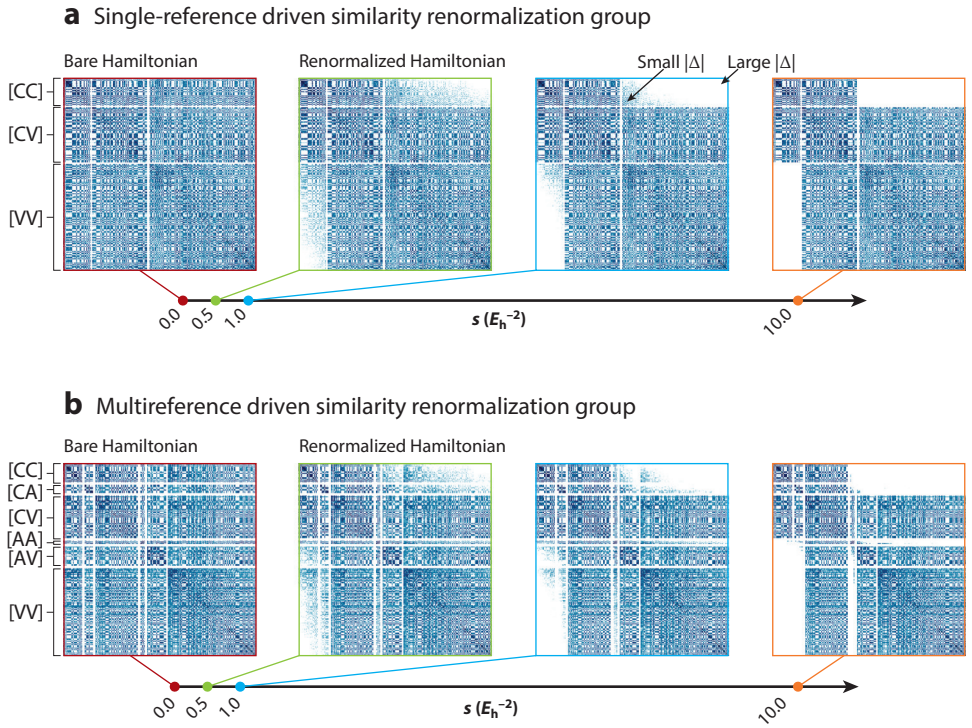

Figure 2

Evolution of the two-body contributions of the transformed Hamiltonian with respect to the flow parameter s for the (*a*) single-reference and (*b*) multireference DSRG formalisms. The two-body Hamiltonian tensor $\bar{H}^{rs}_{pq}(s)$ is plotted as a matrix $\bar{H}_{[pq],[rs]}(s)$ with composite indices $[pq]$ and $[rs]$. The composite indices are blocked into combinations of orbital spaces [**XY**] with $\mathbf{X}, \mathbf{Y} \in \{\mathbf{C}, \mathbf{A}, \mathbf{V}\}$. Within each composite index block, the orbital pairs pq are sorted in increasing order of the sum of orbital energies, $\epsilon_p + \epsilon_q$. The white regions indicate near-zero matrix elements ($<10^{-5}\ E_\mathrm{h}$). Computations were performed on a 90°-twisted butadiene molecule using the 3-21G basis set. RHF orbitals were used for the single-reference LDSRG(2), while the MR-LDSRG(2) computation employed a CAS(4e,4o) reference. The core orbitals and 11 virtual orbitals were excluded from the DSRG treatment of electron correlation. Abbreviations: **A**, active; **C**, core; CAS, complete active space; DSRG, driven similarity renormalization group; LDSRG(2), linearized DSRG with one- and two-body excitations; MR-LDSRG(2), multireference LDSRG(2); RHF, restricted Hartree–Fock; **V**, virtual.

For a generic operator $\hat{a}^{pq\cdots}_{rs\cdots}$, its normal-ordered form $\{\hat{a}^{pq\cdots}_{rs\cdots}\}$ with respect to Ψ_0 is defined as a polynomial in $\hat{a}^\dagger_p, \hat{a}^\dagger_q, \ldots, \hat{a}_s, \hat{a}_r$ with zero expectation value,

$$\langle \Psi_0 | \{\hat{a}^{pq\cdots}_{rs\cdots}\} | \Psi_0 \rangle = 0. \tag{19.}$$

For example, for a one-body operator $\hat{a}^p_q$, the normal-ordered form is given by $\{\hat{a}^p_q\} = \hat{a}^p_q - \gamma^p_q$, while for a two-body operator $\hat{a}^{pq}_{rs}$, we have $\{\hat{a}^{pq}_{rs}\} = \hat{a}^{pq}_{rs} - \mathcal{A}\left(\gamma^p_r \hat{a}^q_s\right) - \lambda^{pq}_{rs} + \gamma^p_r \gamma^q_s - \gamma^p_s \gamma^q_r$. With these definitions, the Hamiltonian operator normal ordered with respect to Ψ_0 is given by

$$\hat{H} = E_0 + \sum_{pq}^{\mathbf{G}} f^q_p \{\hat{a}^p_q\} + \frac{1}{4} \sum_{pqrs}^{\mathbf{G}} v^{rs}_{pq} \{\hat{a}^{pq}_{rs}\}, \tag{20.}$$

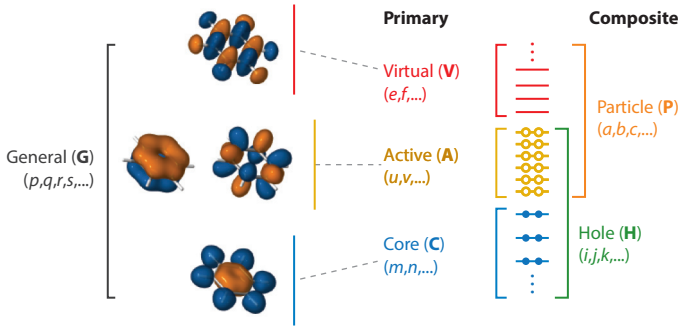

Figure 3

Partitioning of the orbital basis in the multireference driven similarity renormalization group (DSRG) and corresponding orbital indices adopted in this review.

where $E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle$ is the reference energy and the generalized Fock matrix is defined as $f^q_p = h^q_p + \sum_{rs}^{\mathbf{H}} v^{qs}_{pr} \gamma^r_s$.

A generalized version of Wick's theorem simplifies expressions involving products of normal-ordered operators that arise in the MR-DSRG. Two types of contractions appear in the GNO Wick's theorem. The first type involves pairs of second-quantized operators and yields elements of the 1-RDM: $\{\hat{a}^\dagger_p \hat{a}_q\} = \gamma^p_q$, and, $\{\hat{a}_p \hat{a}^\dagger_q\} = \delta^p_q - \gamma^p_q$. These contractions also appear in the single-reference version of Wick's theorem. The second type connects k creation and k annihilation operators ($k \geq 2$) and yields elements of the k-body density cumulants, $\{\hat{a}^\dagger_p \hat{a}^\dagger_q \cdots \hat{a}_s \hat{a}_r\} = \lambda^{pq\cdots}_{rs\cdots}$. These contractions are unique to the GNO formalism and increase the algebraic complexity of the MR-DSRG with respect to its single-reference counterpart. However, for a CAS-type reference, several simplifications can be applied. The 1-RDM is diagonal for the core and virtual blocks, and the density cumulants are nonzero only when all indices correspond to active orbitals.

In the MR-DSRG, the cluster operator $\hat{T}_k(s)$ that enters in the definition of $\hat{A}(s)$ is a generalization of the single-reference version, and it is written explicitly in terms of operators normal ordered with respect to Ψ_0,

$$\hat{T}_k(s) = \frac{1}{(k!)^2} \sum_{ij\cdots}^{\mathbf{H}} \sum_{ab\cdots}^{\mathbf{P}} t^{ij\cdots}_{ab\cdots}(s) \{\hat{a}^{ab\cdots}_{ij\cdots}\}. \qquad 21.$$

Since excitations that involve only active orbitals (internal excitations) have the net effect of changing the coefficients of the determinants that enter the reference (see Equation 17), it is customary to exclude them from the definition of $\hat{T}_k(s)$. This condition corresponds to setting $t^{uv\cdots}_{xy\cdots}(s) = 0$ for all combinations of indices $x, y, \ldots, u, v, \ldots \in \mathbf{A}$.

The source operator used in the MR-DSRG generalizes the single-reference form (Equations 15 and 16) to include excitations that span the hole and particle orbital spaces. As with the cluster operator, internal excitations that enter into $\hat{R}_k(s)$ are removed by setting $r^{uv\cdots}_{xy\cdots}(s) = 0$ for all combinations of indices $x, y, \ldots, u, v, \ldots \in \mathbf{A}$. In the MR-DSRG, we further assume that the orbital energies entering $\hat{R}$ are those from a semicanonical basis.

Internal excitations: involve only active orbitals

Figure 2*b* shows the evolution of the two-body part of $\bar{H}(s)$ in the MR-DSRG using the generalized source operator. This transformation proceeds in a way similar to that of the DSRG, but in the MR case, $\bar{H}(s)$ contains more blocks due to the partitioning of the orbitals in three

spaces. Comparing the two examples, we observe that the decay of the off-diagonal elements in the MR-DSRG transformed Hamiltonian is slower than that of the single-reference DSRG.

FCI: full configuration interaction

To solve the DSRG equations, we must evaluate $\bar{H}(s)$, which may be expressed as a series of commutators of $\hat{H}$ and $\hat{A}(s)$ using the Baker–Campbell–Hausdorff (BCH) formula

$$\bar{H}(s) = e^{-\hat{A}(s)}\hat{H}e^{\hat{A}(s)} = \hat{H} + \sum_{k=1}^{\infty}\frac{1}{k!}\underbrace{[[\ldots[[\hat{H},\hat{A}(s)],\hat{A}(s)],\ldots],\hat{A}(s)]}_{k \text{ nested commutators}}. \quad 22.$$

Like in other theories based on unitary transformations (60, 61, 63), the BCH expansion of the DSRG Hamiltonian does not terminate because $\hat{A}(s)$ contains both excitation and deexcitation operators, $\hat{T}(s)$ and $\hat{T}^\dagger(s)$. Thus, in practical applications it is necessary to use an approximated form of the BCH series.

Once the MR-DSRG equations are solved, the correlation energy is computed as the expectation value of the transformed Hamiltonian $\bar{H}(s)$ as

$$E_0(s) = \langle\Psi_0|\bar{H}(s)|\Psi_0\rangle. \quad 23.$$

Evaluated in this way, the correlation energy neglects any relaxation of the reference due to the coupling of static and dynamical correlation. Relaxation effects can be introduced by requiring Ψ_0 to be an eigenvalue of the transformed Hamiltonian $\bar{H}(s)$, which may achieved by solving the following eigenvalue problem,

$$\sum_{\mu=1}^{d}\langle\Phi_\nu|\bar{H}(s)|\Phi^\mu\rangle c_\mu = E(s)c_\nu, \quad \nu = 1,\ldots,d. \quad 24.$$

The self-consistent iterative solution of the DSRG amplitude and eigenvalue equations (Equations 12 and 24) yields the fully relaxed energy, $E_{\text{fr}}(s)$. In the absence of any approximation, only the fully relaxed MR-DSRG scheme can reproduce the full CI (FCI) energy in the limit $s \to \infty$. Consequently, this is the method of choice for highly accurate computations using nonperturbative approximations. The fully relaxed energy is well approximated by the partially relaxed energy, $E_{\text{pr}}(s)$, obtained by forming $\bar{H}(s)$ using the CASCI reference and diagonalizing it in the basis of reference determinants only once. This second approach is economical and sufficiently accurate for low-order PTs derived from the MR-DSRG.

2.5. Nonperturbative Truncation Schemes

The first hierarchy of MR-DSRG truncation schemes that we consider are nonperturbative approximations in which $\hat{A}(s)$ contains operators of rank up to n [MR-DSRG(n), $n = 2, 3, \ldots$]. The simplest of these methods, MR-DSRG(2), consists in assuming the approximations $\hat{A}(s) \approx \hat{A}_1(s) + \hat{A}_2(s)$ and $\hat{R}(s) \approx \hat{R}_1(s) + \hat{R}_2(s)$. The $\hat{A}_1(s)$ operator, which is given by $\hat{A}_1(s) = \hat{T}_1(s) - \hat{T}_1^\dagger(s)$, with

$$\hat{T}_1(s) = \sum_{m}^{\mathbf{C}}\sum_{e}^{\mathbf{V}} t_e^m(s)\{\hat{a}_m^e\} + \sum_{m}^{\mathbf{C}}\sum_{x}^{\mathbf{A}} t_x^m(s)\{\hat{a}_m^x\} + \sum_{u}^{\mathbf{A}}\sum_{e}^{\mathbf{V}} t_e^u(s)\{\hat{a}_u^e\}, \quad 25.$$

is responsible for mixing core, active, and virtual orbitals, but it excludes active–active rotations, and therefore, it is equivalent to the CASSCF orbital rotation operator (79). The $\hat{A}_2(s)$ operator is responsible for pair correlation effects. Note that in the limit of a single Slater determinant and $\hat{R}(s) = 0$, the MR-DSRG(2) approach is equivalent to unitary CC with singles and doubles (UCCSD). Since CC with singles and doubles (CCSD) and UCCSD are known to differ by fourth-order terms (80) and are found to give nearly identical results in the weakly correlated

regime (81, 82), we expect the MR-DSRG(2) to be of CCSD-like quality in the single-reference limit.

To evaluate the BCH expansion of $\bar{H}(s)$, we follow the linearized commutator approximation of Yanai & Chan (64), which consists in truncating all commutators in Equation 22 up to two-body terms,

$$[\,\cdot\,,\hat{A}(s)] \approx [\,\cdot\,,\hat{A}(s)]_{\{2\}}, \tag{26.}$$

where the subscript $\{2\} \equiv 0, 1, 2$ indicates that only the scalar term and one- and two-body normal-ordered parts of $[\,\cdot\,,\hat{A}(s)]$ are kept. Equations for $[\hat{O},\hat{A}(s)]_{\{2\}}$, where $\hat{O}$ contains both one- and two-body terms, may be easily derived and consist of only 39 terms (83). Note that this approximation does not lead to a closed-form expression for $\bar{H}(s)$, and a recursive evaluation of the the BCH expansion is still necessary. We denote the combination of the MR-DSRG(2) and the linearized commutator approximation as the MR-LDSRG(2). The MR-LDSRG(2) theory gives a transformed Hamiltonian $\bar{H}(s)$ that contains up to two-body operators,

$$\bar{H}_{\{2\}}(s) = E_0(s) + \sum_{pq}^{\mathbf{G}} \bar{H}_p^q(s)\{\hat{a}_q^p\} + \frac{1}{4}\sum_{pqrs}^{\mathbf{G}} \bar{H}_{pq}^{rs}(s)\{\hat{a}_{rs}^{pq}\}. \tag{27.}$$

The cost to evaluate $\bar{H}_{\{2\}}(s)$ scales as $N_{\text{com}} \times \mathcal{O}(N_{\mathbf{H}}^2 N_{\mathbf{P}}^2 N^2)$, where N_{com} is the number of commutator evaluations required to converge the BCH expansion of $\bar{H}_{\{2\}}(s)$. In the limit of $N_{\mathbf{P}} \gg N_{\mathbf{H}}$, this cost is N_{com} times that of single-reference CCSD, which scales as $\mathcal{O}(N_{\mathbf{H}}^2 N_{\mathbf{P}}^4)$. The memory requirements of the MR-LDSRG(2) [$\mathcal{O}(N^4)$] is, however, the same as a vanilla implementation of CCSD.

The error introduced by the linear commutator approximation of Yanai & Chan has been analyzed extensively (64, 66, 84). The lowest-order energy errors introduced by the linearized commutator approximation arise from terms like $\langle\Psi_0|[[\hat{H},\hat{A}(s)]_3,\hat{A}(s)]|\Psi_0\rangle$ since the MR-LDSRG(2) neglects the three-body commutator term $[\hat{H},\hat{A}(s)]_3$. A rigorous evaluation of the double commutator $[[\hat{H},\hat{A}(s)]_3,\hat{A}(s)]$ requires the four-body density cumulant of the reference (85), and this quantity is both expensive to compute and difficult to store in memory for more than 14 spatial active orbitals. However, it has been observed that a quadratic commutator approximation—in which double commutators are computed exactly—can yield larger errors than the linear approximation (85).

For the reference relaxation procedure, we first rewrite $\bar{H}_{\{2\}}(s)$ in normal-ordered form with respect to the true vacuum,

$$\bar{H}_{\{2\}}(s) = \bar{h}_0(s) + \sum_{pq}^{\mathbf{G}} \bar{h}_p^q(s)\hat{a}_q^p + \frac{1}{4}\sum_{pqrs}^{\mathbf{G}} \bar{h}_{pq}^{rs}(s)\hat{a}_{rs}^{pq}, \tag{28.}$$

where the quantities $\bar{h}_0(s)$, $\bar{h}_p^q(s)$, and $\bar{h}_{pq}^{rs}(s)$ may be considered as dressed energy and one- and two-electron integrals, respectively. These quantities are defined in terms of $\bar{H}$ as

$$\bar{h}_0(s) = E_0(s) - \sum_{ij}^{\mathbf{H}} \bar{H}_i^j(s)\gamma_j^i + \frac{1}{2}\sum_{ijkl}^{\mathbf{H}} \bar{H}_{ij}^{kl}(s)\gamma_k^i\gamma_l^j - \frac{1}{4}\sum_{uvxy}^{\mathbf{A}} \bar{H}_{uv}^{xy}(s)\lambda_{xy}^{uv}, \tag{29.}$$

$$\bar{h}_p^q(s) = \bar{H}_p^q(s) - \sum_{ij}^{\mathbf{H}} \bar{H}_{pi}^{qj}(s)\gamma_j^i, \tag{30.}$$

$$\bar{h}_{pq}^{rs}(s) = \bar{H}_{pq}^{rs}(s). \tag{31.}$$

Since this Hamiltonian has the same structure of the bare Hamiltonian $\hat{H}$, it can be straightforwardly interfaced with a standard CASCI code. It may be shown (86) that to compute the matrix elements $\langle\Phi_\mu|\bar{H}_{\{2\}}(s)|\Phi^\nu\rangle$, one only needs to evaluate the components of $\bar{H}_{\{2\}}(s)$ labeled by active indices, i.e., $\bar{H}^x_u(s)$ and $\bar{H}^{xy}_{uv}(s)$ with $u, v, x, y \in \mathbf{A}$.

DSRG-MRPT*n*: *n*th-order DSRG multireference perturbation theory

2.6. Multireference Driven Similarity Renormalization Group Perturbation Theory

In addition to the nonperturbative scheme discussed in Section 2.5, we have also studied low-order perturbative approximations of the MR-DSRG (86–88). We denote these MR-DSRG-based PTs as DSRG-MRPT*n*, where $n = 2, 3, \ldots$ is the order of the perturbative expansion. The main motivation for developing low-order MR PTs is the desire to investigate large systems, for which the MR-LDSRG(2) approximation is too expensive. These approaches are also useful in composite schemes and focal point analyses to approximate the correlation energy in large basis sets or extrapolations to the infinite basis limit (89–91). Perturbative schemes are also valuable because they yield closed-form expressions for the correlation energy, and as such, they provide a way to analyze formal features of the MR-DSRG.

As is customary in developing a perturbative expansion, we begin by partitioning the bare Hamiltonian into a zeroth-order contribution $\hat{H}^{(0)}$ plus a first-order perturbation $\hat{H}^{(1)}$ multiplied by the perturbation parameter ξ,

$$\hat{H} = \hat{H}^{(0)} + \xi\hat{H}^{(1)}. \tag{32.}$$

The transformed Hamiltonian $\bar{H}$ and the operator $\hat{A}(s)$ are expressed as power series in ξ,

$$\bar{H}(s) = \bar{H}^{(0)} + \xi\bar{H}^{(1)}(s) + \xi^2\bar{H}^{(2)}(s) + \xi^3\bar{H}^{(3)}(s) + \cdots, \tag{33.}$$

$$\hat{A}(s) = \xi\hat{A}^{(1)}(s) + \xi^2\hat{A}^{(2)}(s) + \xi^3\hat{A}^{(3)}(s) + \cdots, \tag{34.}$$

where the superscript (n) indicates the nth-order term of a power series. For clarity, in this section, we drop the labels "(s)" for all s-dependent quantities (e.g., $\bar{H}$ and $\hat{A}$). After plugging Equations 32–34 into the definition of the transformed Hamiltonian, applying the BCH formula (Equation 22), and collecting terms of the same power of ξ, we obtain the following zeroth-order through third-order DSRG transformed Hamiltonians:

$$\bar{H}^{(0)} = \hat{H}^{(0)}, \tag{35.}$$

$$\bar{H}^{(1)} = \hat{H}^{(1)} + [\hat{H}^{(0)}, \hat{A}^{(1)}], \tag{36.}$$

$$\bar{H}^{(2)} = \left[\hat{H}^{(0)}, \hat{A}^{(2)}\right] + [\hat{H}^{(1)}, \hat{A}^{(1)}] + \frac{1}{2}[[\hat{H}^{(0)}, \hat{A}^{(1)}], \hat{A}^{(1)}], \tag{37.}$$

$$\begin{aligned}\bar{H}^{(3)} = {} & [\hat{H}^{(0)}, \hat{A}^{(3)}] + [\hat{H}^{(1)}, \hat{A}^{(2)}] + \frac{1}{2}[[\hat{H}^{(1)}, \hat{A}^{(1)}], \hat{A}^{(1)}] \\ & + \frac{1}{2}[[\hat{H}^{(0)}, \hat{A}^{(2)}], \hat{A}^{(1)}] + \frac{1}{2}[[\hat{H}^{(0)}, \hat{A}^{(1)}], \hat{A}^{(2)}] + \frac{1}{6}[[[\hat{H}^{(0)}, \hat{A}^{(1)}], \hat{A}^{(1)}], \hat{A}^{(1)}].\end{aligned} \tag{38.}$$

In this perturbative analysis (87), we assume that the reference wave function is fixed and treat the density matrices and cumulants that result from operator contractions as zeroth-order quantities. Moreover, we employ a one-body diagonal zeroth-order Hamiltonian, $\hat{H}^{(0)} = E_0 + \hat{F}^{(0)}$,

given by the sum of the reference energy E_0 and the diagonal blocks (occupied, active, and virtual) of the generalized Fock matrix $\hat{F}^{(0)}$,

$$\hat{F}^{(0)} = \sum_{mn}^{\mathbf{C}} f_m^n \{\hat{a}_n^m\} + \sum_{uv}^{\mathbf{A}} f_u^v \{\hat{a}_v^u\} + \sum_{ef}^{\mathbf{V}} f_e^f \{\hat{a}_f^e\}. \quad 39.$$

The operator $\hat{F}^{(0)}$ is invariant with respect to separate unitary rotations of core, active, and virtual orbitals. We assume a semicanonical basis ($f_p^q = \delta_q^p \epsilon_p$ for $p, q \in \mathbf{X}$ with $\mathbf{X} \in \{\mathbf{C}, \mathbf{A}, \mathbf{V}\}$) so that $\hat{F}^{(0)}$ takes a particularly simple form, $\hat{F}^{(0)} = \sum_p^{\mathbf{G}} \epsilon_p \{\hat{a}_p^p\}$. With this choice of $\hat{F}^{(0)}$, we find that $[\hat{H}^{(0)}, \hat{A}^{(n)}]_k$ is a k-body off-diagonal operator. This choice has two important consequences: (*a*) the expectation value is $\langle \Psi_0 | [\hat{H}^{(0)}, \hat{A}^{(n)}] | \Psi_0 \rangle = 0$ since internal amplitudes are null, and (*b*) the energy and amplitudes at any order can be solved using a noniterative procedure. It is also easy to show that the zeroth-order energy is equal to the reference energy, $E^{(0)} = E_0$, and the first-order energy is null, $E^{(1)} = 0$.

2.6.1. Second-order energy. The first nontrivial correction to the energy appears at second order,

$$E^{(2)} = \langle \Psi_0 | [\hat{H}^{(1)} + \bar{H}^{(1)}, \hat{T}^{(1)}] | \Psi_0 \rangle, \quad 40.$$

where we have employed $[\hat{H}, \hat{T}]^\dagger = -[\hat{H}, \hat{T}^\dagger]$ to simplify this expression. Solving the first-order DSRG equation $\bar{H}_{\mathrm{od}}^{(1)} = \hat{R}^{(1)}$ leads to the following equations for the first-order amplitudes,

$$t_a^{i,(1)} = \left[f_a^{i,(1)} + \sum_{ux}^{\mathbf{A}} \Delta_u^x t_{ax}^{iu,(1)} \gamma_u^x \right] \frac{1 - e^{-s(\Delta_a^i)^2}}{\Delta_a^i}, \quad 41.$$

$$t_{ab}^{ij,(1)} = v_{ab}^{ij,(1)} \frac{1 - e^{-s(\Delta_{ab}^{ij})^2}}{\Delta_{ab}^{ij}}. \quad 42.$$

Similar to the single-reference SRG-PT2 energy expression (Equation 7), the energy denominators in Equations 41 and 42, are regularized by an exponential function. The equation for singles amplitudes (Equation 41) contains both a first-order Fock-matrix term, $f_a^{i,(1)}$ (like in MP2 for a non-Hartree–Fock reference), and a new contribution from semi-internal doubles amplitudes, $t_{ax}^{iu,(1)}$, contracted with the 1-RDM.

Expressions for the terms that contribute to $E^{(2)}$ (Equation 40) are given in References 87 and 88. The highest body rank of density cumulants appearing in these expressions is three. In the majority of applications where $N_{\mathbf{A}} \ll N_{\mathbf{C}} < N_{\mathbf{V}}$, the time-limiting step in evaluating $E^{(2)}$ is a MP2-like term that scales as $\mathcal{O}(N_{\mathbf{C}}^2 N_{\mathbf{V}}^2)$. When there are 30 or more spatial active orbitals, the evaluation of $E^{(2)}$ can be dominated by a term that involves the three-body density cumulant and that scales as $\mathcal{O}(N_{\mathbf{A}}^6 N_{\mathbf{V}})$. Neglect of the three-body density cumulant in DSRG-MRPT2 is possible (87), and it leads to a numerically stable approach with scalings for large active spaces reduced to only $\mathcal{O}(N_{\mathbf{A}}^4 N_{\mathbf{V}}^2 + N_{\mathbf{A}}^5 N_{\mathbf{V}})$. Due to its simplicity, the DSRG-MRPT2 can be easily combined with density fitting (DF) or Cholesky decomposition of the two-electron integrals, leading to a practical approach to study systems with up to 1,500–2,000 basis functions and 30–40 spatial active orbitals (88).

SA-DSRG: state-averaged DSRG

2.6.2. Third-order energy. The third-order energy contribution is obtained by taking the expectation value of $\bar{H}^{(3)}$ with respect to Ψ_0,

$$E^{(3)} = \langle\Psi_0|[\tilde{H}^{(1)}, \hat{T}^{(2)}]|\Psi_0\rangle + \langle\Psi_0|[\tilde{H}^{(2)}, \hat{T}^{(1)}]|\Psi_0\rangle, \quad 43.$$

where $\tilde{H}^{(1)} = \hat{H}^{(1)} + \bar{H}^{(1)}$ and $\tilde{H}^{(2)} = \bar{H}^{(2)} - \frac{1}{6}[[\hat{H}^{(0)}, \hat{A}^{(1)}], \hat{A}^{(1)}]$ are the modified first- and second-order Hamiltonians, respectively. Note that $\tilde{H}^{(2)}$ contains three-body components $[\tilde{H}_3^{(2)}]$ and they can be fully contacted with $\hat{T}_2^{(1)}$ via four-body density cumulants. To avoid computing this eight-index quantity, we have neglected $\tilde{H}_3^{(2)}$ in Equation 43 (86) and approximated the third-order energy as

$$E^{(3)} \approx \langle\Psi_0|[\tilde{H}^{(1)}, \hat{T}^{(2)}]|\Psi_0\rangle + \langle\Psi_0|[\tilde{H}_{\{2\}}^{(2)}, \hat{T}^{(1)}]|\Psi_0\rangle. \quad 44.$$

This approximation is consistent with the operator truncation scheme used in the MR-LDSRG(2) theory.

Computing $E^{(3)}$ requires the second-order amplitudes, which can be obtained by solving the equation $\bar{H}_{\text{od}}^{(2)} = \hat{R}^{(2)}$. Expressions for the second-order amplitudes are identical to Equations 41 and 42 but with all first-order quantities replaced by the corresponding second-order ones. Specifically, the first-order Hamiltonian elements are replaced by the off-diagonal terms of $\frac{1}{2}[\tilde{H}^{(1)}, \hat{A}^{(1)}]$. With a diagonal one-body zeroth-order Hamiltonian, the second-order amplitude equations can be solved with a noniterative procedure that scales as $\mathcal{O}(N_{\mathbf{H}}^2 N_{\mathbf{P}}^4)$, and this is the time-limiting step of the DSRG-MRPT3 method. Although the asymptotic scaling of DSRG-MRPT3 is identical to that of the MR-LDSRG(2), the former is significantly more efficient and can be applied to larger systems.

2.6.3. Reference relaxation. In the DSRG-MRPT2 and DSRG-MRPT3 methods, the reference CI coefficients are allowed to relax via the partially relaxed scheme described in Section 2.4. As such, the PTs remain simple and efficient, without any iterative update of the reference wave function. Compared to the unrelaxed version, the partially relaxed approach involves two additional steps: (*a*) generating the one- and two-body components of $\bar{H}^{(n)}$ for DSRG-MRPT*n* and (*b*) diagonalizing $\bar{H}^{(n)}$ in the basis of model determinants (Equation 24). Computing the matrix elements $\langle\Phi_\mu|\bar{H}_{\{2\}}^{(n)}(s)|\Phi^\nu\rangle$ required for step *b* has a cost that scales as $\mathcal{O}(N_{\mathbf{A}}^4 N_{\mathbf{V}}^2)$, which is negligible compared to the cost of evaluating the energy and amplitudes (86). However, this estimate does not include the cost of diagonalizing the transformed Hamiltonian, which depends on the specific approach employed and can range from exponential (e.g., CASCI) to polynomial [e.g., density matrix renormalization group (DMRG) (92–94)] in the number of active orbitals.

2.7. Excited States

The MR-DSRG schemes discussed so far focus on one state at a time and may be used to compute the energy of electronic states that are energetically well separated from other states of the same symmetry. However, it is well documented that state-specific theories yield incorrect potential energy surfaces when two or more states become near degenerate (95, 96). To address this problem, we have proposed a state-averaged DSRG (SA-DSRG) framework (97).

In the SA-DSRG ansatz, operators are normal-ordered with respect to an ensemble of n zeroth-order CASCI states, $\mathbb{E}_0 \equiv \{\Psi_0^\alpha, \alpha = 1, 2, \ldots, n\}$. To this ensemble we associate the density operator $\hat{\rho}$,

$$\hat{\rho} = \sum_{\alpha=1}^{n} \omega_\alpha |\Psi_0^\alpha\rangle\langle\Psi_0^\alpha|, \quad 45.$$

where ω_α is the weight of state Ψ_0^α. Here, we assume that all states in $\mathbb{E}_0$ have equal weights ($\omega_\alpha = 1/n$). However, this formalism is also applicable to thermal density matrices where the weights are determined by appropriate Boltzmann factors.

SA-DSRG-PT*n*: *n*th-order perturbation theory derived from state-averaged DSRG

The ensemble average of the Hamiltonian, E_ρ, is then given by

$$E_\rho = \mathrm{Tr}(\hat{\rho}\hat{H}) = \sum_{\alpha=1}^{n} \omega_\alpha \langle \Psi_0^\alpha | \hat{H} | \Psi_0^\alpha \rangle. \quad 46.$$

To discuss the generalization of Wick's theorem to the case of an ensemble, we introduce state-averaged RDMs, $\bar{\gamma}^{pq\cdots}_{rs\cdots} = \sum_{\alpha=1}^{n} \omega_\alpha [\boldsymbol{\gamma}_{\alpha\alpha}]^{pq\cdots}_{rs\cdots}$, where $[\boldsymbol{\gamma}_{\alpha\alpha}]^{pq\cdots}_{rs\cdots} = \langle \Psi_0^\alpha | \hat{a}^{pq\cdots}_{rs\cdots} | \Psi_0^\alpha \rangle$ is the RDM of state Ψ_0^α. The normal-ordered form of an operator $\hat{O}$, denoted $\{\hat{O}\}_\rho$, is obtained by requiring that its ensemble average is zero, i.e., $\mathrm{Tr}(\hat{\rho}\{\hat{O}\}_\rho) = \sum_{\alpha=1}^{n} \omega_\alpha \langle \Psi_0^\alpha | \{\hat{O}\}_\rho | \Psi_0^\alpha \rangle = 0$. This condition implies that contractions in the ensemble GNO yield state-averaged density cumulants (73).

In the SA-DSRG, we perform a single unitary transformation of the bare Hamiltonian imposing the condition $\bar{H}_{\mathrm{od}}(s) = \hat{R}(s)$, where all operators are normal ordered with respect to $\hat{\rho}$. This transformation has the effect of decoupling, on average, the ensemble states from their excited configurations, and may be viewed as a way to identify an average description of dynamical correlation effects optimal for a given ensemble. After forming the DSRG transformed Hamiltonian, it is diagonalized to obtain the energy of n electronic states represented in the ensemble. The basis used to diagonalize $\bar{H}(s)$ may consist of (*a*) the states in $\mathbb{E}_0$ or (*b*) the entire space of CAS determinants. The first option corresponds to a contracted approach as the resulting states can span only $\mathbb{E}_0$. The second option corresponds to an uncontracted scheme, which is similar to the reference relaxation procedure employed in the state-specific MR-DSRG. For either approach, one can use the updated reference states to define a new ensemble and build the transformed Hamiltonian, repeating this procedure until self-consistency is reached.

Since the SA-DSRG framework uses a significant amount of the technologies already developed for MR-DSRG, it can be easily implemented by replacing state-specific density cumulants with their state-averaged counterparts. The the cost of building the SA-DSRG transformed Hamiltonian is independent of the number of ensemble states n. However, computing the ensemble states and their density matrices scales linearly with respect to n. We have currently implemented the SA-DSRG at the level of second- and third-order PT (SA-DSRG-PT2 and SA-DSRG-PT3) by modification of our DSRG-MRPT2 and DSRG-MRPT3 codes, respectively. As done in the state-specific DSRG-MRPT approaches, in the state-averaged variants we employ a partially relaxed scheme where the transformed Hamiltonian is diagonalized only once. By default, the label SA-DSRG-PT corresponds to the approach with a diagonalization space that spans the entire CAS. The alternative approach based on a contracted basis is indicated with the suffix "c," e.g., SA-DSRG-PT2c.

3. FEATURES OF THE DRIVEN SIMILARITY RENORMALIZATION GROUP METHODS

3.1. Orbital Invariance

In this section, we discuss the invariance of the energy and other properties computed with the MR-DSRG methods with respect to unitary rotations of the orbital basis. This property is important because it guarantees that computations performed using different orbital bases (e.g. localized, natural, semicanonical) yield the same result. We restrict our discussion to unitary rotations that separately mix core, active, and virtual orbitals, since the reference wave function is only invariant

Orbital energy: $\epsilon_p = f_p^p(0)$

with respect to these rotations. A unitary transformation from the basis $\{\phi^{p'}\}$ to another basis $\{\phi^q\}$ is written as $\phi^p = \sum_{q'} \phi^{q'} U_{q'}^p$ where $U_{q'}^p = U_{q'p}$ is an element of the unitary matrix $\mathbf{U}$. Restricting the unitary transformation to separate orbital classes implies that the matrix $\mathbf{U}$ is built as a direct sum of unitary matrices that correspond to core $\mathbf{U_C}$, active $\mathbf{U_A}$, and virtual $\mathbf{U_V}$ orbital rotations,

$$\mathbf{U} = \mathbf{U_C} \oplus \mathbf{U_A} \oplus \mathbf{U_V}. \quad 47.$$

The SRG formalism based on the Wegner generator is invariant with respect to unitary orbital transformations, while in the case of the DSRG, the invariance properties are determined by the source operator since the similarity-transformed Hamiltonian is known to be invariant (98, 99). In order to discuss the invariance properties of the DSRG, it is useful to distinguish two categories of orbital invariance. The first one, which we call general invariance, is used to denote invariance with respect to any unitary rotation $\mathbf{U}$ of the form given by Equation 47. Methods such as the internally contracted MRCC (ic-MRCC) approach satisfy general invariance (98, 99). The second category, which we call degenerate invariance, implies invariance only with respect to transformations $\mathbf{U}$ that mix degenerate orbitals (those with same orbital energy ϵ_p). As we see later in this section, this distinction is useful to classify the invariance properties of the DSRG and to compare it to other MR theories. The source operator defined by Equation 16 was derived assuming a semicanonical orbital basis, and it is not invariant with respect to general unitary transformations. However, this source operator is degenerate invariant, which implies that it is possible to formulate a procedure that, given any orbital basis, can reproduce the results in the semicanonical basis (see Reference 83).

To solve the MR-DSRG equations in a general basis, we first evaluate the source operator in the semicanonical basis using Equation 16 and then transform it to the general basis via the unitary matrix $\mathbf{U}$. Equivalently, the MR-DSRG equations are first solved in the semicanonical basis and then the solutions are transformed to the general basis. For example, the semicanonical first-order amplitudes of DSRG-MRPT2 [$t_{cd}^{kl,(1)}$; see Equation 42] are computed via the equation

$$t_{cd}^{kl,(1)} = \left(\sum_{c'd'}^{\mathbf{P}} \sum_{k'l'}^{\mathbf{H}} U_c^{c'} U_d^{d'} v_{c'd'}^{k'l',(1)} U_{k'}^k U_{l'}^l \right) \frac{1 - e^{-s(\Delta_{cd}^{kl})^2}}{\Delta_{cd}^{kl}} \quad 48.$$

and then backtransformed to the general basis via the transformation

$$t_{a'b'}^{i'j',(1)} = \sum_{cd}^{\mathbf{P}} \sum_{kl}^{\mathbf{H}} U_{a'}^c U_{b'}^d t_{cd}^{kl,(1)} U_k^{i'} U_l^{j'}. \quad 49.$$

The evaluation of the energy can then be performed in the general basis by contracting $t_{a'b'}^{i'j',(1)}$ with the electron integrals, the one-body density matrix, and the density cumulants. The above rotation has a cost that scales as $\mathcal{O}(N_{\mathbf{P}}^3 N_{\mathbf{H}}^2)$, which is comparable to the cost of the integral transformation step performed in the DSRG-MRPT2. However, in the density-fitted DSRG-MRPT2 implementation, the integral transformation cost is smaller [$\mathcal{O}(N^4)$], and therefore it is always advantageous to start with semicanonical orbitals. For the DSRG-MRPT3 and MR-LDSRG(2) schemes, the extra cost required to transform all tensors to the semicanonical basis and back to the general basis is negligible compared to that of the tensor contractions.

We conclude this section by discussing the degenerate invariance property of the MR-DSRG. In essence, because degenerate semicanonical orbitals are not uniquely defined, this property guarantees that the procedure outlined above is well defined and yields a unique solution. The same

procedure cannot be applied to MR theories that are not degenerate invariant, like methods based on the Jeziorski–Monkhorst ansatz (100–103). For these methods, even if the equations are solved by transforming to the semicanonical basis, the remaining ambiguity in the definition of degenerate orbitals implies that the energy is not uniquely defined. Consider, for example, a computation of two noninteracting Li atoms using a CAS(2e,2o) reference built from the 2s orbital of each Li atom. Both the fully localized 2s Li basis and the symmetry-adapted delocalized orbitals have degenerate orbital energies and are consistent with the semicanonical condition. In the MR-DSRG methods, degenerate invariance guarantees that computations that use these two bases yield the same energy, but the Jeziorski–Monkhorst methods give different results. In practice, degenerate invariance is necessary for the continuity and uniqueness of the energy obtained by orbital semicanoncalization as a function of the atomic coordinates.

3.2. Size Consistency and Extensivity

In this section, we discuss the size consistency (additivity of energy for noninteracting fragments) and size extensivity (scaling with respect to system size) properties of the MR-DSRG methods. The energy and amplitude equations that define the MR-DSRG method (Equations 8, 12, 16, 23, and 24) only involve quantities that are diagrammatically connected (104, 105). As a consequence, both perturbative and nonperturbative approximations of the MR-DSRG method yield size-extensive energies. Due to the connectivity of the MR-DSRG equations, it also follows that for a multiplicatively separable reference in a basis of localized semicanonical orbitals (these can be constructed from the semicanonical orbitals on each separate fragment), the MR-DSRG correlation energy is additive for separate noninteracting fragments. The orbital invariance properties established in Section 3.1 guarantee that size consistency is also satisfied in any other basis that is unitarily equivalent to the localized semicanonical orbitals. Size consistency and size extensivity are satisfied by the unrelaxed, partially relaxed, and fully relaxed MR-DSRG schemes. Size consistency of the MR-DSRG methods was verified by numerical experiments (83, 106).

3.3. Flow Parameter Dependence

Due to the flow parameter dependence of the MR-DSRG equations, the energy and other properties depend on the value of s used in a computation. This is also the case for other MR theories that employ level shifts (22, 23, 107) or truncation thresholds to eliminate linearly dependent configurations (66, 98, 99, 108–110). On the one hand, the s dependence is undesirable because it requires the user to select a value and introduces an element of arbitrariness in the MR-DSRG results. On the other hand, s may also be viewed as a parameter that can be used to improve the accuracy of the MR-DSRG methods.

A few considerations enter in the choice of the value of s. Since in the limit of $s \to 0$ the DSRG correlation energy goes to zero, the flow parameter should be chosen to be large enough to recover a significant portion of the dynamical correlation energy. At the same time, in the limit of $s \to \infty$, the DSRG is exposed to numerical instabilities that arise from intruders, and therefore, s cannot be chosen to be too large. For the DSRG-MRPT2, it is possible to derive rigorous bounds on the value of s necessary to avoid the appearance of large amplitudes. For example, in Reference 87 we showed that the first-order amplitudes are bound by a constant $t_{\max}$, i.e., $\|t_{ab}^{ij,(1)}\|_{\max} < t_{\max}$, if s satisfies the following inequality,

$$s < \left[\frac{t_{\max}}{0.6382\, \|v_{ab}^{ij}\|_{\max}} \right]^2 . \qquad 50.$$

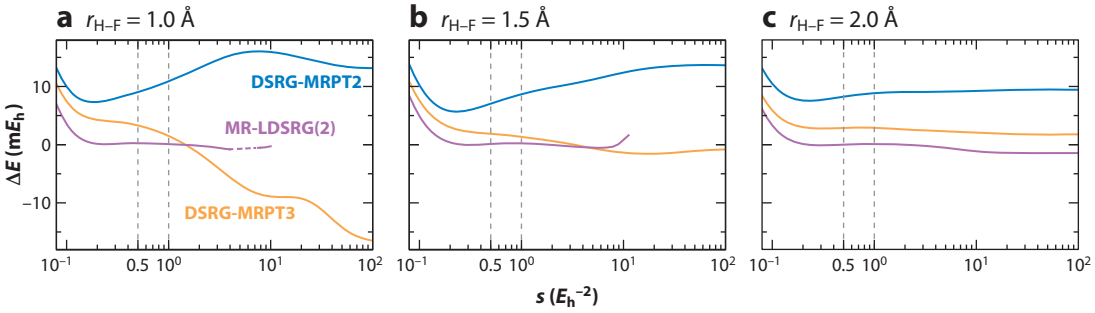

Figure 4

Energy deviations of various MR-DSRG methods with respect to FCI/cc-pVDZ as a function of the flow parameter s computed for HF at different H–F bond lengths: (*a*) 1.0 Å, (*b*) 1.5 Å, and (*c*) 2.0 Å. The purple dashed line indicates the range of s for which the MR-LDSRG(2) equations failed to converge. The vertical dashed lines correspond to values of s equal to 0.5 E_h^{-2}(default) and 1 E_h^{-2}. Abbreviations: cc-pVDZ, correlation-consistent polarized valence double zeta; DSRG-MRPTn, nth-order DSRG multireference perturbation theory; FCI, full configuration interaction; MR-DSRG, multireference driven similarity renormalization group; MR-LDSRG(2), linearized MR-DSRG with one- and two-body excitations.

A study of the s dependence of the unrelaxed DSRG-MRPT2 energy (87) showed that an optimal range for the flow parameter is [0.1, 1.0] E_h^{-2}. In this range, the theory yields the smallest absolute errors and nonparallelism errors (NPEs) for the ground state potential energy curves of HF and N_2. A later test on the potential energy curve of HF showed that the MR-LDSRG(2) energy error remains fairly constant when s falls in the range [0.5, 1.0] E_h^{-2} (83). Therefore, we have adopted a pragmatic approach and performed MR-DSRG computations using the value $s = 0.5\ E_h^{-2}$.

In **Figure 4** we present a comparison of the s dependence of different MR-DSRG theories. We consider the HF molecule at three different bond lengths and compare all results to FCI. For all methods, the correlation energy rapidly decreases from $s = 0$ to $s = 0.2/0.3\ E_h^{-2}$. For s in the range [0.5, 1.0] E_h^{-2}, the energy error of MR-LDSRG(2) remains nearly constant for all three bond lengths, while PTs show a stronger dependence on the flow parameter, especially for shorter bond lengths. **Figure 4*a*** also shows that when $r_{H-F} = 1.0$ Å, it becomes problematic to converge the MR-LDSRG(2) for $s > 4\ E_h^{-2}$, likely due to the presence of an intruder state. In addition, certain semi-internal excitations become significantly larger when going from first to second order in PT, showing that the DSRG-MRPT3 method may be more sensitive to intruders than DSRG-MRPT2.

3.4. Projective Versus Many-Body Formulations and the Residual Conditions

In this section, we discuss the consequences of formulating the DSRG using many-body residual conditions and compare it to methods that employ projective solutions. For simplicity, we assume $\hat{R}(s) = 0$ such that the similarity-transformed Hamiltonian $\bar{H}(s)$ is independent of the flow parameter, and we drop the s dependence in operators and tensors.

Let us first discuss the single-reference DSRG where the Fermi vacuum is a Slater determinant, $\Psi_0 = \Phi$. In this case, the many-body conditions $\omega^{ij\cdots}_{ab\cdots} = 0$ (Equation 14) are equivalent to

$$\bar{H}^{ij\cdots}_{ab\cdots} = 0, \quad \forall i, j \in \mathbf{C}, \quad \forall a, b \in \mathbf{V}, \tag{51.}$$

where $\bar{H}^{ij\cdots}_{ab\cdots}$ are the tensor elements corresponding to the operator $\{\hat{a}^{ab\cdots}_{ij\cdots}\}$ in the many-body expansion of the operator $\bar{H}$ (e.g., Equation 27). We now show that the above condition is

Semi-internal excitations: involve at least one active orbital and one core or virtual orbital

equivalent to satisfying all projections of excited determinants $\Phi_{ij\ldots}^{ab\ldots}$ onto the Schrödinger equation, $\bar{H}|\Phi\rangle = E|\Phi\rangle$. Introducing the projective residual $S_{ab\ldots}^{ij\ldots} = \langle \Phi_{ij\ldots}^{ab\ldots}|\bar{H} - E|\Phi\rangle$, it is easy to show that

$$S_{ab\ldots}^{ij\ldots} = \bar{H}_{ab\ldots}^{ij\ldots}, \tag{52.}$$

where we have used the fact that $\langle \Phi_{ij\ldots}^{ab\ldots}|\Phi\rangle = 0$ and applied Wick's theorem. Therefore, in the single-reference case, the many-body conditions (Equation 51) are equivalent to satisfying $S_{ab\ldots}^{ij\ldots} = 0$, even for truncated theories. We call the latter the projective conditions because they are based on projections of the similarity-transformed Hamiltonian.

For the MR case, the many-body conditions are just a generalization of Equation 51, but the equivalence between the projective and many-body approaches is lost. First, if we compute the projection of the Schrödinger equation residual with respect to an active space determinant Φ_v (S_v),

$$S_v = \langle \Phi_v|\bar{H} - E|\Psi_0\rangle = \sum_{\mu=1}^{d} \langle \Phi_v|\bar{H}|\Phi^\mu\rangle c_\mu - E c_v, \tag{53.}$$

we find that $S_v = 0$ is satisfied for all Φ_v only if the wave function is obtained via the fully relaxed procedure, since this condition is equivalent to satisfying the MR-DSRG eigenvalue equation (Equation 24).

Second, projections of the residual onto excited configurations,

$$S_{ab\ldots}^{ij\ldots} \equiv \langle \Psi_0|\{\hat{a}_{ab\ldots}^{ij\ldots}\}(\bar{H} - E)|\Psi_0\rangle = \langle \Psi_0|\{\hat{a}_{ab\ldots}^{ij\ldots}\}\bar{H}|\Psi_0\rangle, \tag{54.}$$

are only in part equal to zero. To illustrate this point, consider the residual for single excitations, S_a^i, which takes the form

$$S_a^i = \sum_{j}^{\mathbf{H}} \sum_{b}^{\mathbf{P}} \gamma_j^i \eta_a^b \bar{H}_b^j - \frac{1}{2} \sum_{j}^{\mathbf{H}} \sum_{uxy}^{\mathbf{A}} \gamma_j^i \lambda_{au}^{xy} \bar{H}_{xy}^{ju} + \cdots - \frac{1}{6} \sum_{j}^{\mathbf{H}} \sum_{uvxyz}^{\mathbf{A}} \gamma_j^i \lambda_{auv}^{xyz} \bar{H}_{xyz}^{juv} + \cdots. \tag{55.}$$

In analyzing this equation it is convenient to distinguish two cases. If all orbital indices in the residual S_a^i are external, meaning that they involve only core or virtual orbitals (S_e^m), then all terms that involve density cumulants are zero for a CASCI reference and the above expression simplifies to

$$S_e^m = \bar{H}_e^m, \quad \forall m \in \mathbf{C}, \quad \forall e \in \mathbf{V}, \tag{56.}$$

showing that the projective and many-body conditions are equivalent. In the case of excitations involving at least one active index (semi-internal), the expression for S_a^i depends on the 1-, 2-, and higher-body components of the tensors $\bar{H}_{ab\ldots}^{ij\ldots}$, and the residual condition $S_a^i = 0$ can be satisfied only if all of these components are null. It follows then that if the many-body equations are truncated to a given order, then the residual conditions $S_a^i = 0$ cannot be satisfied for semi-internal excitations. This result may be generalized to a generic residual $S_{ab\ldots}^{ij\ldots}$ and summarized by saying that for truncated MR theories, the projective and many-body conditions are equivalent only for external excitations (those not involving active indices). In most practical situations where $N_{\mathbf{C}} \gg N_{\mathbf{A}}$, we find that semi-internal excitations are significantly outnumbered by externals, and the projective and many-body conditions are for the most part equivalent.

Density cumulant $\lambda^{pq\cdots}_{rs\cdots}$: the connected (or irreducible) part of $\gamma^{pq\cdots}_{rs\cdots}$

Reduce density matrix (RDM) of Ψ_0: $\gamma^{pq\cdots}_{rs\cdots} = \langle\Psi_0|\hat{a}^{pq\cdots}_{rs\cdots}|\Psi_0\rangle$

The main advantage of the many-body conditions is that they require density cumulants of smaller rank when compared to the projective equations. However, the amplitude equations acquire Fock space character due to the reduced dependence on the one-body density matrix and the density cumulants. In the many-body approach, the intruder state problem is more severe than in the projective methods, and therefore it is essential to employ an approach like the DSRG to regularize the amplitude equations. For example, depending on the Hamiltonian partitioning, a second-order perturbative expansion of the projective equations may lead to the CASPT2 (CAS second-order PT) (111) or NEVPT2 (*n*-electron valence second-order PT) (12) approaches. The former is subject to intruders, while the later avoids them thanks to the two-body terms in the zeroth-order Hamiltonian. Instead, the corresponding many-body formalisms with or without two-body terms in the zeroth-order Hamiltonian can be shown to suffer from the intruder state problem (86).

The projective conditions are used in several internally contracted MR theories, such as CASPT2 (111), canonical transformation theory (64–66), and ic-MRCC (99, 112). However, projective approaches have several drawbacks: (*a*) It is necessary to perform an expensive orthogonalization of the excitation configurations, $\{\hat{a}^{ab\cdots}_{ij\cdots}\}|\Psi_0\rangle$; (*b*) the energy and amplitude equations are more complex than the corresponding many-body ones; and (*c*) the elimination of linearly dependent excited configurations introduces discontinuities on the potential energy surface. The use of many-body conditions has been proposed by Nooijen and coworkers (108, 113, 114) as a way to avoid the orthogonalization of the excited configurations, $\{\hat{a}^{ab\cdots}_{ij\cdots}\}|\Psi_0\rangle$, and the problems associated with it.

4. IMPLEMENTATION

All the DSRG schemes discussed above are implemented in Forte (115), a suite of MR quantum chemistry methods based on the open source quantum chemistry package Psi4 (116). The MR-DSRG module can be broken down into three components: (*a*) an active space solver that defines the reference wave function, (*b*) an orbital semicanonicalization module, and (*c*) a dynamic correlation solver that determines the DSRG-transformed Hamiltonian.

The first component obtains the reference by diagonalizing the Hamiltonian within the active space using exact or approximate methods [CASSCF, DMRG, selected CI (117, 118), or variational 2-RDM (119) methods]. From the reference wave function, the one-body density matrix and the two- and three-body density cumulants are computed. Storage of these quantities in memory for large active spaces with 30–40 spatial active orbitals requires less than 128 GB of memory for a spin-integrated implementation, and it is generally feasible on a single-node machine. Note that for SA-DSRG computations on n electronic states, the memory requirement doubles and the computational time is n multiplied by the time for evaluating the density cumulants.

The second component of our MR-DSRG code performs a semicanonicalization of the orbitals, a step required in the definition of the source operator. The orbital rotation matrix $\mathbf{U}$ is obtained as described in Section 3.1 and is used to transform the original orbital coefficient matrix $\mathbf{C}'$ into the semicanonical basis $\mathbf{C}$ via $\mathbf{C} = \mathbf{C}'\mathbf{U}$. The active part of the one-body density matrix and the density cumulants in the semicanonical basis are obtained by rotating each index using only the active part of the unitary transformation. This procedure is preferred over recomputing the densities via a second Hamiltonian diagonalization in the semicanonical basis, because the CI coefficients can have arbitrary phases and they are not uniquely defined for degenerate configurations.

The third component solves the MR-DSRG equations using a truncation level requested by the user. Details of the implementation of each approximate scheme can be found in References 88, 86, and 83 for DSRG-MRPT2, DSRG-MRPT3, and MR-LDSRG(2), respectively. Here we briefly

discuss the limitations of the current implementation. First, all DSRG equations are spin integrated such that only unique nonzero spin components of the amplitudes and integrals are computed and stored. Second, sparsity due to point group symmetry is not exploited in the storage and contractions of integrals, densities, and amplitudes. This optimization is desirable especially for computing small molecules with high symmetry where it can bring 4–64 times speedup, depending on the number of irreducible representations.

CASPT*n*: complete active space *n*th-order perturbation theory

With these two limitations in mind, we now analyze the number of basis functions that can be handled for each level of theory. The DF version of DSRG PTs are available in the current version of Forte. In the DF-DSRG-MRPT2 (88), the largest storage demands come from holding six tensors of size $\frac{1}{16}N_{\mathbf{A}}^2N_{\mathbf{P}}^2$. In the DF-DSRG-MRPT3 (86), it is convenient to store all three-index DF integrals in memory, as well as nine tensors of size $\frac{1}{16}N_{\mathbf{H}}^2N_{\mathbf{P}}^2$. As such, practical applications on a single computer node are restricted to 2,000 and 700 basis functions for DF-DSRG-MRPT2 and DF-DSRG-MRPT3, respectively.

Our current implementation of the MR-LDSRG(2) does not use DF and requires storage of nine intermediates (three quantities, each with three spin cases) of size $\frac{1}{16}N^4$. This memory requirement limits practical applications of the MR-LDSRG(2) to systems with about 200 basis functions. Work is underway in our group to implement an alternative DSRG ansatz that reduces the computational and storage costs of the MR-LDSRG(2) and takes advantage of DF.

5. APPLICATIONS

The MR-DSRG computations reported in the literature so far focus largely on benchmarking different approximation schemes and a few selected applications. We have tested the DSRG-MRPT2, DSRG-MRPT3, and MR-LDSRG(2) schemes on the ground state dissociation curves of HF, F_2, H_2O_2, C_2H_6, and N_2 (83, 86, 87). Results for these systems show that the accuracy of MR-LDSRG(2) is higher than that of DSRG-MRPT3, which is significantly higher than that of DSRG-MRPT2 (83, 86). For excited states, we have studied 134 low-lying singlets of 28 molecules using the state-specific DSRG-MRPT2 combined with valence CI singles and doubles (VCISD) references. The resulting VCISD-DSRG-PT2 theory yields excitation energies with an accuracy comparable to that of other MRPT2 approaches (106). In the same study, we also reported computations on formaldehyde that use the DSRG-MRPT2 with references containing up to 34 active orbitals. More recently, SA-DSRG-PT2 and SA-DSRG-PT3 have been benchmarked on the avoided crossing of LiF, and the conical intersection of NH_3 and the penta-2,4-dieniminium cation (97). It is found that the accuracy of SA-DSRG-PT2 is comparable to other multistate second-order PTs, while the SA-DSRG-PT3 results are in excellent agreements with those from MR CI with singles and doubles (MRCISD) (14, 120). Besides these benchmark systems, we have computed the low-lying singlet excited states of octatetraene (97) and the singlet–triplet splittings of (*a*) *p*-benzyne (83, 87), (*b*) all isomers of naphthyne (88), and (*c*) 9,10-anthracyne (86). DF-DSRG-MRPT2 computations on a transition metal complex using 1,700 basis functions have also been reported (88). Here, we briefly review the results for the N_2 dissociation curve and the vertical excitation energies of octatetraene.

5.1. Ground State Dissociation Curve of N_2

Accurately predicting the dissociation curve of N_2 has long been a challenging problem for MR methods. As shown in **Figure 5**, the NPE of CASPT2 is 9.5 mE_h, significantly larger than that observed for single-bond breaking (see **Supplemental Figure 1** and **Supplemental Table 1**). This deterioration is also observed for the DSRG-MRPT2, yet more pronounced (NPE = 18.5 mE_h). The NPE is significantly reduced when going from second- to third-order PTs. For example,

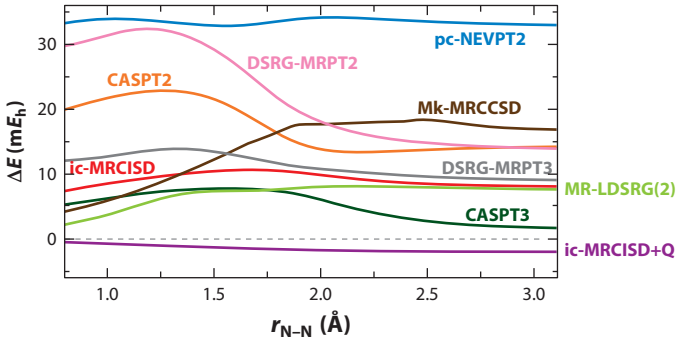

Figure 5

Energy deviations of various multireference methods [based on a CASSCF(6e,6o) reference] with respect to FCI for the ground state of N_2. All computations used the cc-pVDZ basis set, and the molecular orbitals mainly constructed from atomic 1s orbitals were excluded from post-CASSCF treatments. All DSRG results assumed $s = 0.5\ E_h^{-2}$. Abbreviations: CAS, complete active space; CASPT*n*, CAS *n*th-order perturbation theory; CASSCF, CAS self-consistent field; cc-pVDZ, correlation-consistent polarized valence double zeta; DSRG, driven similarity renormalization group; DSRG-MRPT*n*, *n*th-order DSRG multireference perturbation theory; FCI, full configuration interaction; ic-MRCISD, internally contracted multireference configuration interaction with singles and doubles; ic-MRCISD+Q, ic-MRCISD with Davidson correction; Mk-MRCCSD, Mukherjee state-specific multireference coupled cluster with singles and doubles; MR-LDSRG(2), second-order linearized MR-DSRG; pc-NEVPT2, partially contracted *n*-electron valence second-order perturbation theory.

the NPE of DSRG-MRPT3 is only 4.80 mE_h and comparable to that of CASPT3 (6.0 mE_h). Surprisingly, among all PTs, the one that best parallels the FCI curve is pc-NEVPT2 (partially contracted NEVPT2), with an NPE of only 1.3 mE_h. Going from DSRG-MRPT3 to MR-LDSRG(2), more dynamic correlation is gained, but the NPE is worsened by 1.1 mE_h, yet it is still significantly smaller than the Mk-MRCCSD (Mukherjee state-specific MR CCSD) result (14.2 mE_h).

The above observations shed light on how to improve the MR-DSRG methods. Comparing DSRG-MRPT2 to DSRG-MRPT3 and MR-LDSRG(2) suggests that the double commutator $[[\hat{H}, \hat{A}(s)], \hat{A}(s)]$ in the BCH expansion of $\bar{H}(s)$ is essential in describing triple bond breaking in N_2 at a level of accuracy comparable to that of ic-MRCISD (internally contracted MRCISD). Furthermore, contrary to CASPT2 and pc-NEVPT2, the MR-DSRG methods considered here neglect the four-body density cumulant. In the case of N_2, it is known that the three-body density cumulant vanishes at the dissociation limit, while the four-body density cumulant has a large norm (78). This observation suggests that commutator approximations that include the four-body density cumulant may be important to reduce the NPE for the N_2 curve.

pc-NEVPT2: partially contracted *n*-electron valence second-order perturbation theory

Mk-MRCCSD: Mukherjee state-specific multireference coupled cluster with singles and doubles

ic-MRCISD: internally contracted multireference configuration interaction with singles and doubles

5.2. Excitation Energies of (E,E)-1,3,5,7-Octatetraene

We conclude this section by highlighting the application of the SA-DSRG PTs to the vertical excitation energies of (E,E)-1,3,5,7-octatetraene (OTE). In particular, we focus on the ordering of the lowest excited states, $2\ ^1A_g^-$ and $1\ ^1B_u^+$ (97). For the dipole-forbidden $2\ ^1A_g^-$ state, the dominant configuration is a double excitation from the Hartree–Fock determinant, whose description generally necessitates MR methods. On the other hand, the $1\ ^1B_u^+$ state primarily consists of a HOMO/LUMO (highest occupied/lowest unoccupied molecular orbital) single excitation and it is dipole allowed. Predicting the excitation energy of $1\ ^1B_u^+$ requires a sophisticated treatment of

dynamic correlation effects. As such, various theoretical studies have given contrasting predictions of the relative ordering of these two states.

In **Supplemental Table 2**, we summarize the vertical excitation energies of the $2\,^1A_g^-$ and $1\,^1B_u^+$ states of OTE obtained using various theories. Comparing these excitation energies directly to experiment is not straightforward because these computations employed moderate-size basis sets that do not contain diffuse functions. Thus, we restrict our comparison to other computational methods. The $1\,^1B_u^+$ state is dominated by single excitations and is well described by single-reference equation-of-motion CC theories. As such, we consider the CC3 (approximate CC with singles, doubles, and triples) method to provide the best theoretical estimate for the excitation energy (4.94 eV) of the $1\,^1B_u^+$ state. The $2\,^1A_g^-$ state requires a MR method, and we take the excitation energy provided by the ic-MRCCSD-LR (internally contracted MR CCSD linear response) approach (4.65 eV) as the most accurate theoretical estimate.

Among all single-reference methods, only ADC(3) (third-order algebraic diagrammatic construction) predicts the $2\,^1A_g^-$ state lying below the $1\,^1B_u^+$ state. However, the ADC(3) excitation energies are generally too low, especially for the $2\,^1A_g^-$ state, which is 0.92 eV below the ic-MRCCSD-LR result. In the case of MR theories, the excitation energy of $2\,^1A_g^-$ is largely insensitive to the different treatments of dynamic electron correlation. On the contrary, the energy of the $1\,^1B_u^+$ state strongly depends on the level of theory. For example, excitation energies from XMS-CASPT2 (extended multistate CASPT2) and QD-sc-NEVPT2 (quasi-degenerate strongly contracted NEVPT2)/CAS(8e,8o) are more than 0.7 eV too low, leading to an incorrect ordering of the two states. This underestimation has been attributed to the lack of dynamic σ polarization ($\pi \rightarrow \pi^*$ excitations coupled with $\sigma \rightarrow \sigma^*$ excitations) in the CAS(8e,8o) wave function (121). If we enlarge the active space to CAS(8e,16o), the QD-sc-NEVPT2 excitation energy for the $1\,^1B_u^+$ state is found to be in better agreement with the CC3 value. Instead, for all SA-DSRG-PT2/3 methods, enlarging the active space has little effect on the excitation energy of the $1\,^1B_u^+$ state.

CC3: approximate coupled cluster with singles, doubles, and triples model

ic-MRCCSD-LR: internally contracted multireference coupled cluster with singles and doubles linear response

ADC(*n*): *n*th-order algebraic diagrammatic construction

XMS-CASPT2: extended multistate complete active space second-order perturbation theory

QD-sc-NEVPT2: quasi-degenerate strongly contracted *n*-electron valence second-order perturbation theory

6. CONCLUSIONS AND FUTURE WORK

The DSRG is a formalism for creating numerically robust MR electron correlation theories. In the past several years, we have developed a series of computationally feasible MR-DSRG methods, ranging from an efficient second-order PT to a more accurate, CCSD-like nonperturbative scheme. These MR-DSRG methods address some of the most critical limitations inherent to internally contracted MR theories, including the intruder state problem, the unfavorable scaling with respect to the number of active orbitals, and energy discontinuities in potential energy surfaces due to the elimination of redundant excitations. However, the use of a renormalization group approach introduces some drawbacks, such as the dependence of the results on the flow parameter.

Under the current MR-DSRG formalism, several important directions deserve future explorations. Since MR-DSRG perturbative schemes are sufficiently accurate and avoid expensive terms involving four-body density matrices, it would be useful to develop analytic gradient theories of the SA-DSRG-PT2 and SA-DSRG-PT3 approaches for applications to geometry optimizations and on-the-fly molecular dynamics on ground and excited potential energy surfaces. Another important development is to go beyond a treatment of dynamical correlation that includes only single and double substitutions. The importance of perturbative triple corrections in determining the success of the CCSD(T) (CCSD with perturbative triples) method (122, 123) suggests that attaining quantitative agreement with experiment requires an analogous extension of the MR-LDSRG(2) approach. Another route to improving the accuracy of MR-DSRG

theory is to go beyond the linearized commutator approximation employed in MR-LDSRG(2). As discussed previously, this approximation is known to introduce errors at the third order of perturbation.

An important open challenge is to reduce or eliminate the dependence of the MR-DSRG results on the flow parameter. In this respect, the MR-DSRG exposes a problem inherent to all internally contracted theories that use numerical thresholds to eliminate linear dependencies in the basis of contracted configurations. The elimination of the discontinuities in the potential energy surfaces seen in other internally contracted theories represents a major breakthrough for MR-DSRG, but the presence of an external parameter like s appears to be unavoidable. There are at least two ways to improve upon our current way of performing MR-DSRG computations. One is to identify a value of s that minimizes error metrics of a benchmark set. In a recent study (124) on MP2 with a modified denominator of the form $\max(\Delta_{ab}^{ij}, \tau)$, the threshold value $\tau \approx 2.4\, E_\mathrm{h}$ was used, which is consistent with a flow parameter $s \approx 0.17\, E_\mathrm{h}^{-2}$ and suggests that our default value $s = 0.5\, E_\mathrm{h}^{-2}$ is close to the optimal value. A direct path to reducing the s dependence is to account for the off-diagonal components of the Hamiltonian that are not driven to zero by the DSRG transformation. This can be done by diagonalizing $\bar{H}$ in a space of configurations that extends beyond the active space. This strategy has been successfully applied in the MR equation-of-motion CC of Nooijen and coworkers (113, 114). We expect that this solution would also alleviate the dependence of MR-DSRG results on the choice of the active space. Both solutions are under investigation in our group.

SUMMARY POINTS

1. The multireference driven similarity renormalization group (MR-DSRG) is a general framework to add dynamical correlation effects on top of any given zeroth-order reference wave function.
2. The MR-DSRG methods avoid the intruder state problem by performing a gradual diagonalization of the Hamiltonian controlled by the flow parameter s. An s-dependent Hermitian source operator drives the Hamiltonian to a band-diagonal structure.
3. The MR-DSRG scheme avoids the problem of linear dependencies in the excitation manifold inherent to internally contracted theories by using many-body conditions.
4. Under the linearized commutator approximation, the MR-DSRG methods are currently applicable to active spaces with up to 30–40 active orbitals, because at most the three-body density cumulants of the reference are required.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work was supported by the US Department of Energy under award number DE-SC0016004, a Research Fellowship of the Alfred P. Sloan Foundation, and a Camille Dreyfus Teacher–Scholar Award. The authors would like to thank Scott Bogner, Heiko Hergert, Marcel Nooijen, Piotr

Piecuch, Garnet Chan, and Andreas Köhn for insightful discussions. The authors would also like to thank Jeffrey B. Schriber for comments on the manuscript.

LITERATURE CITED

1. Kohn W, Sham LJ. 1965. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* 140:A1133–38
2. Coester F, Kümmel H. 1960. Short-range correlations in nuclear wave functions. *Nucl. Phys.* 17:477–85
3. Coester F. 1958. Bound states of a many-particle system. *Nucl. Phys.* 7:421–24
4. Čížek J. 1969. On the use of the cluster expansion and the technique of diagrams in calculations of correlation effects in atoms and molecules. *Adv. Chem. Phys.* 14:35–89
5. Čížek J. 1966. On correlation problem in atomic and molecular systems: calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods. *J. Chem. Phys.* 45:4256
6. Cohen AJ, Mori-Sanchez P, Yang W. 2008. Insights into current limitations of density functional theory. *Science* 321:792–94
7. Kulik HJ. 2015. Perspective: treating electron over-delocalization with the DFT+U method. *J. Chem. Phys.* 142:240901
8. Furche F, Perdew JP. 2006. The performance of semilocal and hybrid density functionals in 3*d* transition-metal chemistry. *J. Chem. Phys.* 124:044103
9. Cramer CJ, Truhlar DG. 2009. Density functional theory for transition metals and transition metal chemistry. *Phys. Chem. Chem. Phys.* 11:10757–816
10. Roos BO, Taylor PR, Siegbahn PEM. 1980. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* 48:157–73
11. Andersson K, Malmqvist PÅ, Roos BO, Sadlej AJ, Wolinski K. 1990. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* 94:5483–88
12. Angeli C, Cimiraglia R, Evangelisti S, Leininger T, Malrieu JP. 2001. Introduction of *n*-electron valence states for multireference perturbation theory. *J. Chem. Phys.* 114:10252–64
13. Chaudhuri RK, Freed KF, Hose G, Piecuch P, Kowalski K, et al. 2005. Comparison of low-order multireference many-body perturbation theories. *J. Chem. Phys.* 122:134105
14. Werner HJ, Knowles PJ. 1988. An efficient internally contracted multiconfiguration–reference configuration interaction method. *J. Chem. Phys.* 89:5803–14
15. Szalay PG, Müller T, Gidofalvi G, Lischka H, Shepard R. 2012. Multiconfiguration self-consistent field and multireference configuration interaction methods and applications. *Chem. Rev.* 112:108–81
16. Lyakh DI, Musiał M, Lotrich VF, Bartlett RJ. 2012. Multireference nature of chemistry: the coupled-cluster view. *Chem. Rev.* 112:182–243
17. Köhn A, Hanauer M, Mück LA, Jagau TC, Gauss J. 2013. State-specific multireference coupled-cluster theory. *WIREs Comput. Mol. Sci.* 3:176–97
18. Evangelista FA. 2018. Perspective: multireference coupled cluster theories of dynamical electron correlation. *J. Chem. Phys.* 149:030901
19. Jankowski K, Paldus J, Grabowski I, Kowalski K. 1992. Applicability of valence-universal multireference coupled-cluster theories to quasidegenerate electronic states. I. Models involving at most two-body amplitudes. *J. Chem. Phys.* 97:7600–12
20. Paldus J, Piecuch P, Pylypow L, Jeziorski B. 1993. Application of Hilbert-space coupled-cluster theory to simple $(H_2)_2$ model systems: planar models. *Phys. Rev. A* 47:2738–82
21. Evangelisti S, Daudey JP, Malrieu JP. 1987. Qualitative intruder-state problems in effective Hamiltonian theory and their solution through intermediate Hamiltonians. *Phys. Rev. A* 35:4930–41
22. Roos BO, Andersson K. 1995. Multiconfigurational perturbation theory with level shift—the Cr_2 potential revisited. *Chem. Phys. Lett.* 245:215–23
23. Roos BO, Andersson K, Fülscher MP, Serrano-Andrés L, Pierloot K, et al. 1996. Applications of level shift corrected perturbation theory in electronic spectroscopy. *J. Mol. Struct. THEOCHEM* 388:257–76

24. Malrieu JP, Durand P, Daudey JP. 1985. Intermediate Hamiltonians as a new class of effective Hamiltonians. *J. Phys. A* 18:809–26
25. Dyall KG. 1995. The choice of a zeroth-order Hamiltonian for second-order perturbation theory with a complete active space self-consistent-field reference function. *J. Chem. Phys.* 102:4909–18
26. Chaudhuri RK, Finley JP, Freed KF. 1997. Comparison of the perturbative convergence with multireference Möller–Plesset, Epstein–Nesbet, forced degenerate and optimized zeroth order partitionings: the excited BeH_2 surface. *J. Chem. Phys.* 106:4067–81
27. Wegner F. 1994. Flow-equations for Hamiltonians. *Ann. Phys.* 506:77–91
28. Głazek SD, Wilson KG. 1994. Perturbative renormalization group for Hamiltonians. *Phys. Rev. D* 49:4214–18
29. White SR. 2002. Numerical canonical transformation approach to quantum many-body problems. *J. Chem. Phys.* 117:7472–82
30. Bogner SK, Furnstahl RJ, Perry RJ. 2007. Similarity renormalization group for nucleon-nucleon interactions. *Phys. Rev. C* 75:061001
31. Tsukiyama K, Bogner SK, Schwenk A. 2011. In-medium similarity renormalization group for nuclei. *Phys. Rev. Lett.* 106:222502
32. Tsukiyama K, Bogner SK, Schwenk A. 2012. In-medium similarity renormalization group for open-shell nuclei. *Phys. Rev. C* 85:061304
33. Hergert H, Bogner SK, Binder S, Calci A, Langhammer J, et al. 2013. In-medium similarity renormalization group with chiral two- plus three-nucleon interactions. *Phys. Rev. C* 87:034307
34. Hergert H, Bogner SK, Morris TD, Schwenk A, Tsukiyama K. 2016. The in-medium similarity renormalization group: a novel ab initio method for nuclei. *Phys. Rep.* 621:165–222
35. Hergert H. 2017. In-medium similarity renormalization group for closed and open-shell nuclei. *Phys. Scr.* 92:023002
36. Evangelista FA. 2014. A driven similarity renormalization group approach to quantum many-body problems. *J. Chem. Phys.* 141:054109
37. Crawford TD, Schaefer HF. 2000. An introduction to coupled cluster theory for computational chemists. *Rev. Comp. Chem.* 14:33–136
38. Bartlett RJ, Musiał M. 2007. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* 79:291–352
39. Shavitt I, Bartlett RJ. 2009. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*. Cambridge, UK: Cambridge Univ. Press
40. Piecuch P, Kowalski K, Pimienta ISO, McGuire MJ. 2002. Recent advances in electronic structure theory: method of moments of coupled-cluster equations and renormalized coupled-cluster approaches. *Int. Rev. Phys. Chem.* 21:527–655
41. Krylov AI. 2008. Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: the Hitchhiker's guide to Fock space. *Annu. Rev. Phys. Chem.* 59:433–62
42. Piecuch P. 2010. Active-space coupled-cluster methods. *Mol. Phys.* 108:2987–3015
43. Musiał M, Perera A, Bartlett RJ. 2011. Multireference coupled-cluster theory: the easy way. *J. Chem. Phys.* 134:114108
44. Shen J, Piecuch P. 2012. Merging active-space and renormalized coupled-cluster methods via the CC(P; Q) formalism, with benchmark calculations for singlet-triplet gaps in biradical systems. *J. Chem. Theory Comput.* 8:4968–88
45. Small DW, Head-Gordon M. 2009. Tractable spin-pure methods for bond breaking: local many-electron spin-vector sets and an approximate valence bond model. *J. Chem. Phys.* 130:084103
46. Parkhill JA, Head-Gordon M. 2010. A tractable and accurate electronic structure method for static correlations: the perfect hextuples model. *J. Chem. Phys.* 133:024103
47. Bulik IW, Henderson TM, Scuseria GE. 2015. Can single-reference coupled cluster theory describe static correlation? *J. Chem. Theory Comput.* 11:3171–79
48. Qiu Y, Henderson TM, Zhao J, Scuseria GE. 2017. Projected coupled cluster theory. *J. Chem. Phys.* 147:064111
49. Kehrein S. 2006. *The Flow Equation Approach to Many-Particle Systems*. Berlin: Springer-Verlag

50. Chu MT, Driessel KR. 1990. The projected gradient method for least squares matrix approximations with spectral constraints. *SIAM J. Numer. Anal.* 27:1050–60
51. Brockett RW. 1991. Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems. *Linear Algebra Appl.* 146:79–91
52. Bach V, Bru JB. 2010. Rigorous foundations of the Brockett–Wegner flow for operators. *J. Evol. Equ.* 10:425–42
53. Anderson A. 1993. Quantum canonical transformations: physical equivalence of quantum theories. *Phys. Lett. B* 305:67–70
54. Van Vleck J. 1929. On σ-type doubling and electron spin in the spectra of diatomic molecules. *Phys. Rev.* 33:467–506
55. Kemble EC. 2005. *The Fundamental Principles of Quantum Mechanics: With Elementary Applications*. New York: Dover
56. Primas H. 1963. Generalized perturbation theory in operator form. *Rev. Mod. Phys.* 35:710–11
57. Kutzelnigg W. 1982. Quantum chemistry in Fock space. I. The universal wave and energy operators. *J. Chem. Phys.* 77:3081–97
58. Kutzelnigg W, Koch S. 1983. Quantum chemistry in Fock space. II. Effective Hamiltonians in Fock space. *J. Chem. Phys.* 79:4315–35
59. Kutzelnigg W. 1984. Quantum chemistry in Fock space. III. Particle-hole formalism. *J. Chem. Phys.* 80:822–30
60. Bartlett RJ, Kucharski SA, Noga J. 1989. Alternative coupled-cluster Ansätze II: the unitary coupled-cluster method. *Chem. Phys. Lett.* 155:133–40
61. Watts JD, Trucks GW, Bartlett RJ. 1989. The unitary coupled-cluster approach and molecular properties. Applications of the UCC(4) method. *Chem. Phys. Lett.* 157:359–66
62. Kutzelnigg W. 1991. Error analysis and improvements of coupled-cluster theory. *Theor. Chim. Acta* 80:349–86
63. Taube AG, Bartlett RJ. 2006. New perspectives on unitary coupled-cluster theory. *Int. J. Quantum Chem.* 106:3393
64. Yanai T, Chan GKL. 2006. Canonical transformation theory for multireference problems. *J. Chem. Phys.* 124:194106
65. Yanai T, Chan GKL. 2007. Canonical transformation theory from extended normal ordering. *J. Chem. Phys.* 127:104107
66. Neuscamman E, Yanai T, Chan GKL. 2010. A review of canonical transformation theory. *Int. Rev. Phys. Chem.* 29:231–71
67. Mazziotti DA. 2007. Anti-Hermitian part of the contracted Schrödinger equation for the direct calculation of two-electron reduced density matrices. *Phys. Rev. A* 75:022505
68. Mazziotti DA. 2012. Two-electron reduced density matrix as the basic variable in many-electron quantum chemistry and physics. *Chem. Rev.* 112:244–62
69. Schrieffer JR, Wolff PA. 1966. Relation between the Anderson and Kondo Hamiltonians. *Phys. Rev.* 149:491–92
70. Bravyi S, DiVincenzo DP, Loss D. 2011. Schrieffer–Wolff transformation for quantum many-body systems. *Ann. Phys.* 326:2793–826
71. Kutzelnigg W. 2010. Unconventional aspects of coupled-cluster theory. In *Recent Progress in Coupled Cluster Methods*, ed. P Čársky, J Paldus, J Pittner, pp. 299–356. Chall. Adv. Comput. Chem. Phys. 11. Dordrecht, Neth.: Springer
72. Mukherjee D. 1995. A coupled cluster approach to the electron correlation problem using a correlated reference state. In *Recent Progress in Many-Body Theories*, ed. E Schachinger, H Mitter, H Sormann, pp. 127–33. Boston: Springer
73. Kutzelnigg W, Mukherjee D. 1997. Normal order and extended Wick theorem for a multiconfiguration reference wave function. *J. Chem. Phys.* 107:432–49
74. Mukherjee D. 1997. Normal ordering and a Wick-like reduction theorem for fermions with respect to a multi-determinantal reference state. *Chem. Phys. Lett.* 274:561–66

75. Sinha D, Maitra R, Mukherjee D. 2013. Generalized antisymmetric ordered products, generalized normal ordered products, ordered and ordinary cumulants and their use in many electron correlation problem. *Comput. Theor. Chem.* 1003:62–70
76. Mazziotti DA. 1998. Approximate solution for electron correlation through the use of Schwinger probes. *Chem. Phys. Lett.* 289:419–27
77. Kutzelnigg W, Mukherjee D. 1999. Cumulant expansion of the reduced density matrices. *J. Chem. Phys.* 110:2800–9
78. Hanauer M, Köhn A. 2012. Meaning and magnitude of the reduced density matrix cumulants. *Chem. Phys.* 401:50–61
79. Werner HJ, Knowles PJ. 1985. A second order multiconfiguration SCF procedure with optimum convergence. *J. Chem. Phys.* 82:5053
80. Kutzelnigg W. 1998. Almost variational coupled cluster theory. *Mol. Phys.* 94:65–71
81. Cooper B, Knowles PJ. 2010. Benchmark studies of variational, unitary and extended coupled cluster methods. *J. Chem. Phys.* 133:234102
82. Evangelista FA. 2011. Alternative single-reference coupled cluster approaches for multireference problems: The simpler, the better. *J. Chem. Phys.* 134:224102
83. Li C, Evangelista FA. 2016. Towards numerically robust multireference theories: the driven similarity renormalization group truncated to one- and two-body operators. *J. Chem. Phys.* 144:164114. Erratum. 2018. *J. Chem. Phys.* 148:079903
84. Evangelista FA, Gauss J. 2012. On the approximation of the similarity-transformed Hamiltonian in single-reference and multireference coupled cluster theory. *Chem. Phys.* 401:27–35
85. Neuscamman E, Yanai T, Chan GKL. 2009. Quadratic canonical transformation theory and higher order density matrices. *J. Chem. Phys.* 130:124102
86. Li C, Evangelista FA. 2017. Driven similarity renormalization group: third-order multireference perturbation theory. *J. Chem. Phys.* 146:124132. Erratum. 2018. *J. Chem. Phys.* 148:079902
87. Li C, Evangelista FA. 2015. Multireference driven similarity renormalization group: a second-order perturbative analysis. *J. Chem. Theory Comput.* 11:2097–108
88. Hannon KP, Li C, Evangelista FA. 2016. An integral-factorized implementation of the driven similarity renormalization group second-order multireference perturbation theory. *J. Chem. Phys.* 144:204111
89. East ALL, Allen WD. 1993. The heat of formation of NCO. *J. Chem. Phys.* 99:4638–50
90. Császár AG, Allen WD, Schaefer HF. 1998. In pursuit of the *ab initio* limit for conformational energy prototypes. *J. Chem. Phys.* 108:9751–64
91. Tajti A, Szalay PG, Császár AG, Kállay M, Gauss J, et al. 2004. HEAT: high accuracy extrapolated *ab initio* thermochemistry. *J. Chem. Phys.* 121:11599
92. White SR. 1992. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* 69:2863–66
93. Schollwöck U. 2011. The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.* 326:96–192
94. Chan GKL, Sharma S. 2011. The density matrix renormalization group in quantum chemistry. *Annu. Rev. Phys. Chem.* 62:465–81
95. Nakano H. 1993. Quasidegenerate perturbation theory with multiconfigurational self-consistent-field reference functions. *J. Chem. Phys.* 99:7983–92
96. Malrieu JP, Heully JL, Zaitsevskii A. 1995. Multiconfigurational second-order perturbative methods: overview and comparison of basic properties. *Theor. Chim. Acta* 90:167–87
97. Li C, Evangelista FA. 2018. Driven similarity renormalization group for excited states: a state-averaged perturbation theory. *J. Chem. Phys.* 148:124106
98. Evangelista FA, Gauss J. 2011. An orbital-invariant internally contracted multireference coupled cluster approach. *J. Chem. Phys.* 134:114102
99. Hanauer M, Köhn A. 2011. Pilot applications of internally contracted multireference coupled cluster theory, and how to choose the cluster operator properly. *J. Chem. Phys.* 134:204111
100. Jeziorski B, Monkhorst HJ. 1981. Coupled-cluster method for multideterminantal reference states. *Phys. Rev. A* 24:1668–81

101. Evangelista FA, Allen WD, Schaefer HF. 2007. Coupling term derivation and general implementation of state-specific multireference coupled cluster theories. *J. Chem. Phys.* 127:024102
102. Kong L. 2010. Orbital invariance issue in multireference methods. *Int. J. Quantum Chem.* 110:2603–13
103. Evangelista FA, Gauss J. 2010. Insights into the orbital invariance problem in state-specific multireference coupled cluster theory. *J. Chem. Phys.* 133:044101
104. Bartlett RJ. 1981. Many-body perturbation theory and coupled cluster theory for electron correlation in molecules. *Annu. Rev. Phys. Chem.* 32:359–401
105. Hanrath M. 2009. On the concepts of connectivity, separability, and consistency: an illustration by partitioned diagrams and numerical probing. *Chem. Phys.* 356:31–38
106. Li C, Verma P, Hannon KP, Evangelista FA. 2017. A low-cost approach to electronic excitation energies based on the driven similarity renormalization group. *J. Chem. Phys.* 147:074107
107. Ghigo G, Roos BO, Malmqvist PÅ. 2004. A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chem. Phys. Lett.* 396:142–49
108. Datta D, Kong L, Nooijen M. 2011. A state-specific partially internally contracted multireference coupled cluster approach. *J. Chem. Phys.* 134:214116
109. Yanai T, Kurashige Y, Neuscamman E, Chan GKL. 2012. Extended implementation of canonical transformation theory: parallelization and a new level-shifted condition. *Phys. Chem. Chem. Phys.* 14:7809–20
110. Sivalingam K, Krupicka M, Auer AA, Neese F. 2016. Comparison of fully internally and strongly contracted multireference configuration interaction procedures. *J. Chem. Phys.* 145:054104
111. Andersson K, Malmqvist PÅ, Roos BO. 1992. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* 96:1218–26
112. Hanauer M, Köhn A. 2012. Communication: restoring full size extensivity in internally contracted multireference coupled cluster theory. *J. Chem. Phys.* 137:131103
113. Datta D, Nooijen M. 2012. Multireference equation-of-motion coupled cluster theory. *J. Chem. Phys.* 137:204107
114. Demel O, Datta D, Nooijen M. 2013. Additional global internal contraction in variations of multireference equation of motion coupled cluster theory. *J. Chem. Phys.* 138:134108
115. Evangelista Lab. 2019. *Forte: an open-source suite of quantum chemistry methods for strongly correlated electrons*. Psi4 Plugin Software. **https://github.com/evangelistalab/forte**
116. Parrish RM, Burns LA, Smith DGA, Simmonett AC, DePrince AE, et al. 2017. Psi4 1.1: an open-source electronic structure program emphasizing automation, advanced libraries, and interoperability. *J. Chem. Theory Comput.* 13:3185–97
117. Evangelista FA. 2014. Adaptive multiconfigurational wave functions. *J. Chem. Phys.* 140:124114
118. Schriber JB, Evangelista FA. 2016. Communication: an adaptive configuration interaction approach for strongly correlated electrons with tunable accuracy. *J. Chem. Phys.* 144:161106
119. Fosso-Tande J, Nguyen TS, Gidofalvi G, DePrince AE III. 2016. Large-scale variational two-electron reduced-density-matrix-driven complete active space self-consistent field methods. *J. Chem. Theory Comput.* 12:2260–71
120. Knowles PJ, Werner HJ. 1988. An efficient method for the evaluation of coupling coefficients in configuration interaction calculations. *Chem. Phys. Lett.* 145:514–22
121. Angeli C, Pastore M. 2011. The lowest singlet states of octatetraene revisited. *J. Chem. Phys.* 134:184302
122. Urban M, Noga J, Cole S, Bartlett RJ. 1985. Towards a full CCSDT model for electron correlation. *J. Chem. Phys.* 83:4041–46
123. Raghavachari K, Trucks GW, Pople JA, Head-Gordon M. 1989. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* 157:479–83
124. Ohnishi Yy, Ishimura K, Ten-no S. 2014. Interaction energy of large molecules from restrained denominator MP2-F12. *J. Chem. Theory Comput.* 10:4857–61