The Journal of Chemical Physics

Driven similarity renormalization group for excited states: A state-averaged perturbation theory

Cite as: J. Chem. Phys. **148**, 124106 (2018); https://doi.org/10.1063/1.5019793
Submitted: 16 December 2017 . Accepted: 21 February 2018 . Published Online: 23 March 2018

Chenyang Li, and Francesco A. Evangelista

COLLECTIONS

Paper published as part of the special topic on JCP Editors' Choice 2018

This paper was selected as an Editor's Pick

View Online Export Citation CrossMark

ARTICLES YOU MAY BE INTERESTED IN

Perspective: Multireference coupled cluster theories of dynamical electron correlation
The Journal of Chemical Physics **149**, 030901 (2018); https://doi.org/10.1063/1.5039496

Novel strategy to implement active-space coupled-cluster methods
The Journal of Chemical Physics **148**, 124108 (2018); https://doi.org/10.1063/1.5004971

Communication: Density functional theory model for multi-reference systems based on the exact-exchange hole normalization
The Journal of Chemical Physics **148**, 121101 (2018); https://doi.org/10.1063/1.5025334

© 2018 Author(s).

Driven similarity renormalization group for excited states: A state-averaged perturbation theory

Chenyang Li[a] and Francesco A. Evangelista[b]
Department of Chemistry and Cherry Emerson Center for Scientific Computation, Emory University, Atlanta, Georgia 30322, USA

(Received 16 December 2017; accepted 21 February 2018; published online 23 March 2018)

The multireference driven similarity renormalization group (MRDSRG) approach [C. Li and F. A. Evangelista, J. Chem. Theory Comput. **11**, 2097 (2015)] is generalized to treat quasi-degenerate electronic excited states. The new scheme, termed state-averaged (SA) MRDSRG, is a state-universal approach that considers an ensemble of quasi-degenerate states on an equal footing. Using the SA-MRDSRG framework, we implement second- (SA-DSRG-PT2) and third-order (SA-DSRG-PT3) perturbation theories. These perturbation theories can treat a manifold of near-degenerate states at the cost of a single state-specific computation. At the same time, they have several desirable properties: (1) they are intruder-free and size-extensive, (2) their energy expressions can be evaluated non-iteratively and require at most the three-body density cumulant of the reference states, and (3) the reference states are allowed to relax in the presence of dynamical correlation effects. Numerical benchmarks on the potential energy surfaces of lithium fluoride, ammonia, and the penta-2,4-dieniminium cation reveal that the SA-DSRG-PT2 method yields results with accuracy similar to that of other second-order quasi-degenerate perturbation theories. The SA-DSRG-PT3 results are instead consistent with those from multireference configuration interaction with singles and doubles (MRCISD). Finally, we compute the vertical excitation energies of (E,E)-1,3,5,7-octatetraene. The ordering of the lowest three states is predicted to be $2\,^1A_g^- < 1\,^1B_u^+ < 1\,^1B_u^-$ by both SA-DSRG-PT2 and SA-DSRG-PT3, in accordance with MRCISD plus Davidson correction. *Published by AIP Publishing.* https://doi.org/10.1063/1.5019793

I. INTRODUCTION

Multireference (MR) theories are perhaps the most general methods to study electronic excited states. These approaches can be classified into three broad categories: (1) MR perturbation theory (MRPT),[1–7] (2) MR configuration interaction (MRCI),[8–10] and (3) MR coupled cluster (MRCC) methods.[11–22] Among these schemes, second-order MRPT (MRPT2) approaches are the most widely applied because of their low cost. Multireference perturbation theories are also sufficiently accurate to treat many excited states that are incorrectly described by single-reference methods.[23,24]

State-specific formulations of MRPT2 target a collection of excited states independently. However, it has long been realized that they yield qualitatively incorrect potential energy surfaces (PESs) in the vicinity of near-degenerate electronic states.[25] A solution to this problem is offered by quasi-degenerate (QD) perturbation theories (PTs) or multi-state (MS) approaches. These methods are based—to various degrees—on the formalism of effective Hamiltonian theory.[25–30] Numerous realizations of MS and QD theories have been suggested.[31–36] Some of the most popular second-order schemes include multi-configurational QDPT (MC-QDPT2)[31] and its extended version (XMCQDPT2),[37] MS complete-active-space (CAS) PT (MS-CASPT2),[32] and QD *n*-electron valence state PT (QD-NEVPT2).[33]

These methods differ in several aspects. In Nakano's MC-QDPT2,[31] the perturbers are uncontracted singly and doubly excited configuration state functions (CSFs). The energies of a manifold of states are obtained by diagonalizing a state-universal effective Hamiltonian built using second-order perturbation theory. On the contrary, MS-CASPT2 and QD-NEVPT2 first perform a series of state-specific computations using contracted perturbers. Subsequently, a second-order effective Hamiltonian is built using the basis of state-specific wave functions and diagonalized to obtain the energies of all states. Because the MC-QDPT2 uses uncontracted perturbers, its computational cost is proportional to the number of model space CSFs, that is, exponential scaling with respect to the number of active orbitals. By contrast, MS-CASPT2 and QD-NEVPT2 scale polynomially with respect to the number of active orbitals although this estimate does not account for the cost of computing complete-active-space references and their density matrices.

Another distinction lies in the choice of the zeroth-order Hamiltonian [$\hat{H}^{(0)}$]. This operator is normally defined by projecting a simple model Hamiltonian onto individual CAS configuration interaction (CASCI) eigenvectors and excited configurations. In MS-CASPT2, the model Hamiltonian is the generalized Fock operator ($\hat{F}$), while MC-QDPT2 only exploits the diagonal blocks of $\hat{F}$. QD-NEVPT2 uses a more

[a]Electronic mail: cli62@emory.edu
[b]Electronic mail: francesco.evangelista@emory.edu

sophisticated model Hamiltonian suggested by Dyall,[38] where diagonal blocks of $\hat{F}$ are augmented by two-electron interactions within the active orbitals. Choosing an appropriate set of projectors requires extra caution since some choices have been shown to violate important formal properties of the Hamiltonian like size consistency.[37,39] Moreover, Granovsky[37] points out that a zeroth-order Hamiltonian with a diagonal projector is not invariant with respect to unitary rotations of the model space. As a result, the PES may be erratic in the degenerate region where the CASCI vectors are not uniquely defined. In XMCQDPT2,[37] this problem is solved by using a more general projector that is invariant to unitary rotations of the model space. An extended version of MS-CASPT2 has also been developed.[40]

Both MS-CASPT2[32] and QD-NEVPT2[33] computations proceed via a two-step approach. In the first step, state-specific perturbed wave functions are obtained for all states of interest. These state-specific computations use a zeroth-order Hamiltonian partitioned differently for each state (multipartitioning scheme).[41] In the second step, the second-order effective Hamiltonian is formed and diagonalized.[32,33,42] As a consequence, MS-CASPT2 and QD-NEVPT2 inherit some of the difficulties of their parent state-specific methods. Such limitations include, for example, the intruder-state problem[43–45] in CASPT2,[46,47] the presence of the four-particle density matrix,[48,49] and the need to remove linear dependencies in the excitation manifold.[2,5] The latter two aspects are important because they restrict the number of active orbitals that can be used in CASPT2 and NEVPT2 applications.

In our previous work,[50–53] we have shown that the MRPT2 based on driven similarity renormalization group[54,55] (DSRG) provides a viable alternative to state-specific CASPT2 and NEVPT2. The DSRG-MRPT2 method addresses both the intruder-state problem and the limitations resulting from high-order density matrices without sacrificing much of the accuracy. In analogy to the *in-medium* similarity renormalization group,[56–62] the DSRG performs a continuous unitary transformation of the Hamiltonian that depends on a time-like quantity—the flow parameter s. This parameter controls the extent to which dynamic correlation effects are folded into the transformed Hamiltonian. Specifically, for a given value of s, excitations that correspond to energy denominators larger than a cutoff energy $\Lambda = s^{-1/2}$ are folded into an effective Hamiltonian. Configurations that have energy denominators smaller than Λ are not decoupled from the reference. Excitations that cause the intruder-state problem have small energy denominators and are avoided by the DSRG transformation. As a consequence, the DSRG effective Hamiltonian generated with a finite value of s is free from intruders.

Expensive contractions that involve higher-order density matrices are removed by formulating the DSRG as a set of Fock-space many-body conditions[30,63–65] in combination with Mukherjee and Kutzelnigg's generalized normal ordering (MK-GNO).[66–72] This choice has two important consequences. On the one hand, the use of many-body conditions avoids the need to orthogonalize the internally contracted excitation manifold.[19,73] On the other hand, the DSRG-MRPT2 equations require at most three-body density cumulants, which directly arise from contractions of second quantized operators in the MK-GNO algebra. The combination of these two advantages makes the DSRG-MRPT2 applicable to systems with more than 30 active orbitals.[53] This aspect is also true in the case of higher-level MR-DSRG theories, such as third-order MRPT (DSRG-MRPT3),[52] which still need at most the three-body cumulant in the linear recursive commutator approximation.[74,75]

In this work, we adapt the MR-DSRG framework to treat degenerate electronic states. To this end, we propose a novel approach in which a single state-averaged (SA) DSRG transformation is performed on the Hamiltonian. The equations that determine this transformation are many-body equations normal ordered with respect to a vacuum chosen to be a statistical ensemble of model space solutions. One consequence of redefining the vacuum in such a way is that contractions of operators depend on SA density matrices and cumulants.[67] Like in the state-specific version, the state-averaged MR-DSRG transformation uses one unitary transformation. However, the latter decouples the ensemble of reference states from their corresponding internally contracted excited configurations in an averaged way. Once the DSRG equations are solved, an effective Hamiltonian is formed and diagonalized to obtain the model states dressed with correlation effects. This procedure can be performed once or it can be iterated until the solutions reach stationarity. Due to the use of state-averaged density in this QD formulation of MR-DSRG, we term it state-averaged MRDSRG (SA-MRDSRG).

Herein, we focus on second- and third-order perturbation theories derived from the SA-MRDSRG, which are denoted as SA-DSRG-PT2 and SA-DSRG-PT3, respectively. The zeroth-order Hamiltonian contains only the diagonal blocks of $\hat{F}$, defined using the SA one-particle density matrix of the model space. As a consequence, the transformed Hamiltonian can be obtained in a single non-iterative procedure that does not depend on the number of model states. Another major advantage of this formalism is that the dominant cost to compute SA-DSRG-PT2/3 energy corrections is identical to that of a *single* state-specific DSRG-MRPT2/3 calculation. The SA theories also share several nice properties of the state-specific counterparts as they are intruder-free, size extensive,[76] and they require only the one-, two-, and three-body density cumulants of the reference.

This article is organized as follows. In Sec. II, we first review the MK-GNO formalism for an ensemble of states (Sec. II A) and then introduce the SA-DSRG-PT2 and SA-DSRG-PT3 methods (Sec. II B). Numerical applications of the SA-DSRG-PT2 and SA-DSRG-PT3 schemes to lithium fluoride, ammonia, the penta-2,4-dieniminium cation, and (E,E)-1,3,5,7-octatetraene are presented in Sec. III. Finally, we end this work by considering some future extensions of the current state-averaged perturbation theories (Sec. IV).

II. THEORY

A. Generalized normal ordering for an ensemble of states

We first review the theory of generalized normal ordering of Mukherjee and Kutzelnigg (MK-GNO)[67] for a statistical

ensemble ($\mathbb{E}$) of n electronic states $\mathbb{E} \equiv \{\Psi^\alpha, \alpha = 1, 2, \ldots, n\}$. The ensemble is characterized by a density operator ($\hat{\rho}$),

$$\hat{\rho} = \sum_{\alpha=1}^{n} \omega_\alpha |\Psi^\alpha\rangle\langle\Psi^\alpha|, \tag{1}$$

where $\omega_\alpha \geq 0$ is the weight of state Ψ^α in the ensemble, and the sum of the weights is equal to one, $\sum_{\alpha=1}^{n} \omega_\alpha = 1$. In this work, we will assume that states have equal weights ($\omega_\alpha = 1/n$), but the formalism presented is applicable to more general situations. For brevity, the following discussion concentrates only on particle-number-conserving operators.

If we use the notation $\{\hat{A}\}$ to indicate the operator $\hat{A}$ normal ordered with respect to a single electronic state Ψ, the normal order condition is $\langle\Psi|\{\hat{A}\}|\Psi\rangle = 0$. In the case of an ensemble, this condition is replaced by an analogous one, where the average of the ensemble normal-ordered $\hat{A}$ ($\{\hat{A}\}_\rho$) vanishes

$$\langle\{\hat{A}\}_\rho\rangle_\rho = \mathrm{Tr}(\hat{\rho}\,\{\hat{A}\}_\rho) = \sum_{\alpha=1}^{n} \omega_\alpha \langle\Psi^\alpha|\{\hat{A}\}_\rho|\Psi^\alpha\rangle = 0. \tag{2}$$

Using the MK Wick theorem, we may identify the relevant contractions by imposing the normal-ordered condition [Eq. (2)] on a hierarchy of operators with increasing rank. Starting from a general one-body operator, $\hat{a}^p_q = \hat{a}^\dagger_p \hat{a}_q$, defined in terms of bare creation ($\hat{a}^\dagger_p$) and annihilation ($\hat{a}_q$) operators, the MK Wick theorem allows us to rewrite this quantity as the sum of a normal-ordered pair plus a single contraction

$$\hat{a}^p_q = \{\hat{a}^p_q\}_\rho + \hat{a}^\dagger_p \hat{a}_q. \tag{3}$$

Taking the ensemble average of this expression and employing the definition of normal ordering [Eq. (2)], we can derive the value of the contraction in Eq. (3),

$$\hat{a}^\dagger_p \hat{a}_q = \langle\hat{a}^p_q\rangle_\rho = \sum_{\alpha=1}^{n} \omega_\alpha \langle\Psi^\alpha|\hat{a}^p_q|\Psi^\alpha\rangle = \sum_{\alpha=1}^{n} \omega_\alpha \, [\boldsymbol{\gamma}_{\alpha\alpha}]^p_q = \bar{\gamma}^p_q, \tag{4}$$

where $\bar{\gamma}^p_q = [\bar{\boldsymbol{\gamma}}]^q_p$ is the state-averaged one-particle density matrix and $[\boldsymbol{\gamma}_{\alpha\alpha}]^p_q = \langle\Psi^\alpha|\hat{a}^p_q|\Psi^\alpha\rangle$ is the density matrix for state Ψ^α in the ensemble. In general, the expectation value of an arbitrary second-quantized operator $\hat{a}^{pq\cdots}_{rs\cdots} = \hat{a}^\dagger_p \hat{a}^\dagger_q \ldots \hat{a}_s \hat{a}_r$ is given by the corresponding SA density matrix element

$$\langle\hat{a}^{pq\cdots}_{rs\cdots}\rangle_\rho = \bar{\gamma}^{pq\cdots}_{rs\cdots} = \sum_{\alpha=1}^{n} \omega_\alpha [\boldsymbol{\gamma}_{\alpha\alpha}]^{pq\cdots}_{rs\cdots}. \tag{5}$$

In the case of a two-body operator ($\hat{a}^{pq}_{rs}$), the MK Wick theorem yields

$$\begin{aligned}
\hat{a}^{pq}_{rs} &= \{\hat{a}^{pq}_{rs}\}_\rho + \hat{a}^\dagger_q \hat{a}_s \{\hat{a}^p_r\}_\rho \\
&\quad - \hat{a}^\dagger_q \hat{a}_r \{\hat{a}^p_s\}_\rho + \hat{a}^\dagger_p \hat{a}_r \{\hat{a}^q_s\}_\rho - \hat{a}^\dagger_p \hat{a}_s \{\hat{a}^q_r\}_\rho \\
&\quad + \hat{a}^\dagger_p \hat{a}^\dagger_q \hat{a}_s \hat{a}_r + \hat{a}^\dagger_p \hat{a}^\dagger_q \hat{a}_s \hat{a}_r + \hat{a}^\dagger_p \hat{a}^\dagger_q \hat{a}_s \hat{a}_r \\
&= \{\hat{a}^{pq}_{rs}\}_\rho + \sum_{\mathcal{P}} (-1)^{\mathcal{P}} \bar{\gamma}^q_s \{\hat{a}^p_r\}_\rho \\
&\quad + \hat{a}^\dagger_p \hat{a}^\dagger_q \hat{a}_s \hat{a}_r + \bar{\gamma}^p_r \bar{\gamma}^q_s - \bar{\gamma}^p_s \bar{\gamma}^q_r,
\end{aligned} \tag{6}$$

where $\sum_{\mathcal{P}} (-1)^{\mathcal{P}} \bar{\gamma}^q_s \{\hat{a}^p_r\}_\rho$ implies the sum of all partitions of upper and lower indices of the SA density and the normal-ordered operator with an appropriate sign factor. Note that for a general reference, we also introduce four- and higher-leg contractions that involve more than two operators. Taking the ensemble average of Eq. (6), we can identify the four-leg contraction with a state-averaged two-body density cumulant ($\bar{\lambda}^{pq}_{rs}$)

$$\hat{a}^\dagger_p \hat{a}^\dagger_q \hat{a}_s \hat{a}_r = \bar{\lambda}^{pq}_{rs} \equiv \bar{\gamma}^{pq}_{rs} - \bar{\gamma}^p_r \bar{\gamma}^q_s + \bar{\gamma}^p_s \bar{\gamma}^q_r. \tag{7}$$

This procedure may be repeated for higher body operators. In general, one finds that multi-leg contractions correspond to density cumulants defined in terms of state averaged density matrices.

From the ensemble normal ordering condition [Eq. (2)] also follows a generalization of Wick's theorem for products of normal-ordered operators of the form $\{\hat{A}\}_\rho\{\hat{B}\}_\rho$ where multi-leg contractions appear (e.g., see Refs. 67 and 55 and the supplementary material in Ref. 50).

B. The SA-DSRG-PT2 and SA-DSRG-PT3 methods

1. Reference CAS ensemble

The vacuum used in SA-MRDSRG is an ensemble of n zeroth-order electronic states, $\mathbb{E}_0 \equiv \{\Psi^\alpha_0, \alpha = 1, 2, \ldots, n\}$, obtained from a CASCI or SA-CASSCF procedure.[9,77–79] Each state in the ensemble is expanded using a common set of determinants that forms a CAS $\{\Phi^I, I = 1, 2, \ldots, N_{\mathrm{CAS}}\}$

$$|\Psi^\alpha_0\rangle = \sum_{I=1}^{N_{\mathrm{CAS}}} c^\alpha_I |\Phi^I\rangle. \tag{8}$$

To define an active space ensemble, we partition the set of spin orbitals $\mathbf{G} \equiv \{\phi^p, p = 1, 2, \ldots, N\}$ into core ($\mathbf{C}$), active ($\mathbf{A}$), and virtual ($\mathbf{V}$) subsets of dimension $N_{\mathbf{C}}, N_{\mathbf{A}}$, and $N_{\mathbf{V}}$, respectively. As usual, we introduce the hole ($\mathbf{H} = \mathbf{C} \cup \mathbf{A}$) and particle ($\mathbf{P} = \mathbf{A} \cup \mathbf{V}$) composite sets with size $N_{\mathbf{H}} = N_{\mathbf{C}} + N_{\mathbf{A}}$ and $N_{\mathbf{P}} = N_{\mathbf{A}} + N_{\mathbf{V}}$, respectively. An important consequence of this orbital partitioning is that the density cumulants (for example, $\bar{\lambda}^{pq}_{rs}$) can be nonzero if and only if all the indices belong to the set of active orbitals ($\forall\, p, q, r, s \in \mathbf{A}$).

Using the MK-GNO, we may write the bare Hamiltonian ($\hat{H}$) in a normal-ordered form with respect to the ensemble

$$\hat{H} = E_0 + \sum_{pq}^{\mathbf{G}} \bar{f}^q_p \{\hat{a}^p_q\}_\rho + \frac{1}{4} \sum_{pqrs}^{\mathbf{G}} v^{rs}_{pq} \{\hat{a}^{pq}_{rs}\}_\rho, \tag{9}$$

where $E_0 = \langle\hat{H}\rangle_\rho = \sum_{\alpha=1}^{n} \omega_\alpha \langle\Psi^\alpha_0|\hat{H}|\Psi^\alpha_0\rangle$ is the SA reference energy and $\bar{f}^q_p$ is the SA Fock matrix

$$\bar{f}^q_p = h^q_p + \sum_{ij}^{\mathbf{H}} v^{qj}_{pi} \bar{\gamma}^i_j, \tag{10}$$

defined by the one-electron (h^q_p) and antisymmetrized two-electron ($v^{rs}_{pq} = \langle\phi_p\phi_q||\phi^r\phi^s\rangle$) integrals and the SA density [$\bar{\gamma}^i_j$, see Eq. (4)].

2. State-averaged transformation of the Hamiltonian

Our previous formulations of the DSRG[50,54,55] are based on a state-specific transformation of the Hamiltonian via a

unitary operator $[\hat{U}(s)]$

$$\hat{H} \to \bar{H}(s) = \hat{U}^\dagger(s)\hat{H}\hat{U}(s), \quad s \geq 0, \tag{11}$$

where $\hat{U}(s)$ is a continuous function of the time-like flow parameter s. To extend the DSRG to a state-averaged formalism, we propose to fold dynamic correlation energy for all the states in the ensemble via one unitary transformation of the Hamiltonian, as in Eq. (11). The resulting DSRG transformed Hamiltonian $[\bar{H}(s)]$ written in a normal-ordered form with respect to the ensemble density is an operator that includes three- and higher-body terms

$$\bar{H}(s) = \bar{E}_0(s) + \sum_{pq}^{\mathbf{G}} \bar{H}^q_p(s)\{\hat{a}^p_q\}_\rho + \frac{1}{4}\sum_{pqrs}^{\mathbf{G}} \bar{H}^{rs}_{pq}(s)\{\hat{a}^{pq}_{rs}\}_\rho + \ldots, \tag{12}$$

where $\bar{E}_0(s) = \langle \bar{H}(s)\rangle_\rho$, the quantities $\bar{H}^{rs\cdots}_{pq\cdots}(s)$ are tensors, and the second quantized operators are normal ordered with respect to the ensemble of states. It is convenient to write $\hat{U}(s)$ in terms of an s-dependent cluster operator $\hat{T}(s)$,

$$\hat{U}(s) = \exp[\hat{T}(s) - \hat{T}^\dagger(s)] = \exp[\hat{A}(s)], \tag{13}$$

where $\hat{A}(s) \equiv \hat{T}(s) - \hat{T}^\dagger(s)$ is an anti-Hermitian operator. The cluster operator is a sum of many-body operators, $\hat{T}(s) = \hat{T}_1(s) + \hat{T}_2(s) + \ldots + \hat{T}_n(s)$, truncated up to n-body terms, where the generic k-body term $\hat{T}_k(s)$ is written in terms of s-dependent cluster amplitudes $t^{ij\cdots}_{ab\cdots}(s)$,

$$\hat{T}_k(s) = \frac{1}{(k!)^2}\sum_{ij\cdots}^{\mathbf{H}}\sum_{ab\cdots}^{\mathbf{P}} t^{ij\cdots}_{ab\cdots}(s)\{\hat{a}^{ab\cdots}_{ij\cdots}\}_\rho. \tag{14}$$

We further require that $\hat{T}(s)$ does not include internal excitations (involving only active indices). This condition is enforced by imposing $t^{xy\cdots}_{uv\cdots}(s) = 0$, $\forall u, v, x, y, \cdots \in \mathbf{A}$.

In the state-specific DSRG, the unitary transformation folds dynamic correlation effects into $\bar{H}(s)$ by achieving a partial decoupling of a reference wave function Ψ_0 from its excited configurations. This goal is attained by solving a nonlinear equation in which the non-diagonal part of $\bar{H}(s)$ (indicated with the superscript "N" and responsible for the coupling) is driven to zero by an Hermitian operator, $\hat{R}(s)$, termed the source operator

$$[\bar{H}(s)]^{\mathrm{N}} = \hat{R}(s). \tag{15}$$

The source operator is chosen in such a way that for $s = 0$ then $\bar{H}(0) = \hat{H}$, and as s increases, the transformed Hamiltonian smoothly transitions from the original Hamiltonian to one with no coupling between the reference and its excited configurations, that is, $\lim_{s\to\infty}[\bar{H}(s)]^{\mathrm{N}} = 0$. In the state-averaged extension of the DSRG, the operator $\hat{U}(s)$ is determined by the flow equation [Eq. (15)], but contrary to the state-specific approach, we impose many-body conditions on operators normal ordered with respect to the ensemble of states. Therefore, the source operator $\hat{R}(s)$ in Eq. (15) is expressed as a sum of many-body operators, with the k-body ($k = 1, 2, \ldots, n$) component given by

$$\hat{R}_k(s) = \frac{1}{(k!)^2}\sum_{ij\ldots}^{\mathbf{H}}\sum_{ab\ldots}^{\mathbf{P}} r^{ij\cdots}_{ab\ldots}(s)(\{\hat{a}^{ab\cdots}_{ij\ldots}\}_\rho + \{\hat{a}^{ij\cdots}_{ab\ldots}\}_\rho). \tag{16}$$

Here, $r^{ij\cdots}_{ab\ldots}$ is a rank $2k$ tensor (assumed to be real) whose explicit expression will be given later in Sec. II B 3. Note also that the source operator is a nondiagonal operator, that is, $[\hat{R}(s)]^{\mathrm{N}} = \hat{R}(s)$ with $r^{xy\cdots}_{uv\ldots}(s) = 0, \forall u, v, x, y, \cdots \in \mathbf{A}$.

The DSRG equation [Eq. (15)] should be interpreted as an *operator equation* that leads to a set of many-body conditions[19,30,54,65] that implicitly determine $\hat{U}(s)$. This means that if we define the operator $\hat{O}(s) = [\bar{H}(s)]^{\mathrm{N}} - \hat{R}(s)$, whose many-body expansion is

$$\hat{O}(s) = \sum_{i}^{\mathbf{H}}\sum_{a}^{\mathbf{P}} o^i_a(s)(\{\hat{a}^a_i\}_\rho + \{\hat{a}^i_a\}_\rho) + \frac{1}{4}\sum_{ij}^{\mathbf{H}}\sum_{ab}^{\mathbf{P}} o^{ij}_{ab}(s)(\{\hat{a}^{ab}_{ij}\}_\rho + \{\hat{a}^{ij}_{ab}\}_\rho) + \ldots, \tag{17}$$

we interpret the DSRG equation, $\hat{O}(s) = 0$, as the set of conditions

$$o^{ij\cdots}_{ab\ldots}(s) = o^{ab\cdots}_{ij\ldots}(s) = 0, \tag{18}$$

for all combinations of indices but excluding internal excitations.

Note that the SA-DSRG formalism is in many ways significantly different from that of traditional MRPTs based on a projective scheme. In the latter, one introduces a basis of perturber functions and defines corresponding subspaces and projectors. The many-body formalism presented here instead emphasizes the transformation of the Hamiltonian and therefore does not require defining a set of perturbers. In place of subspaces and their corresponding projectors, the many-body version introduces a partitioning of operators into diagonal and non-diagonal components.

3. SA-MRDSRG perturbative analysis

In this work, we consider second- and third-order perturbation theories based on the SA-MRDSRG framework, namely, SA-DSRG-PT2 and SA-DSRG-PT3. The bare Hamiltonian is partitioned into a zeroth-order contribution $\hat{H}^{(0)}$ and a first-order perturbation $\hat{H}^{(1)}$. The zeroth-order term is chosen to contain the SA reference energy and the diagonal blocks of the SA Fock matrix

$$\hat{H}^{(0)} = E_0 + \hat{F}^{(0)}, \tag{19}$$

$$\hat{F}^{(0)} = \sum_{mn}^{\mathbf{C}} \bar{f}^{n,(0)}_m\{\hat{a}^m_n\}_\rho + \sum_{uv}^{\mathbf{A}} \bar{f}^{v,(0)}_u\{\hat{a}^u_v\}_\rho + \sum_{ef}^{\mathbf{V}} \bar{f}^{f,(0)}_e\{\hat{a}^e_f\}_\rho = \sum_{p}^{\mathbf{G}} \epsilon_p\{\hat{a}^p_p\}_\rho, \tag{20}$$

where $\epsilon_p = \bar{f}^{p,(0)}_p$. The last line of Eq. (20) implies the use of semicanonical molecular orbitals.[50,80] This basis is obtained by separately rotating the core, active, and virtual orbitals so that the corresponding blocks of the SA Fock matrix are diagonal.

We then perform an order-by-order expansion of the cluster amplitudes, the source operator, the DSRG transformed Hamiltonian, and DSRG equations.[50,52] The zeroth- till third-order transformed Hamiltonian are given by

$$\bar{H}^{(0)}(s) = \hat{H}^{(0)}, \tag{21}$$

$$\bar{H}^{(1)}(s) = \hat{H}^{(1)} + [\hat{H}^{(0)}, \hat{A}^{(1)}(s)], \tag{22}$$

$$\bar{H}^{(2)}(s) = [\hat{H}^{(0)}, \hat{A}^{(2)}(s)] + \frac{1}{2}[\tilde{H}^{(1)}(s), \hat{A}^{(1)}(s)], \tag{23}$$

$$\bar{H}^{(3)}(s) \approx [\hat{H}^{(0)}, \hat{A}^{(3)}(s)] + \frac{1}{2}[\tilde{H}^{(1)}(s), \hat{A}^{(2)}(s)] + \frac{1}{2}[\tilde{H}^{(2)}_{1,2}(s), \hat{A}^{(1)}(s)], \tag{24}$$

where we introduced modified first-order $\tilde{H}^{(1)}(s) = \hat{H}^{(1)} + \bar{H}^{(1)}(s)$ and second-order $\tilde{H}^{(2)}(s) = \bar{H}^{(2)}(s) - \frac{1}{6}[[\hat{H}^{(0)}, \hat{A}^{(1)}(s)], \hat{A}^{(1)}(s)]$ operators. For convenience, we approximate the equation for the third-order transformed Hamiltonian by neglecting three-body terms in the operator $\tilde{H}^{(2)}(s)$.[52] This approximation is indicated with the subscript "1,2."[55,74]

For a source operator [Eq. (16)] with matrix elements defined as[54]

$$r^i_a(s) = [f^i_a + \sum_{uv}^{\mathbf{A}} \Delta^u_v \bar{\gamma}^u_v t^{iv}_{au}(s)] \exp[-s(\Delta^i_a)^2], \tag{25}$$

$$r^{ij}_{ab}(s) = v^{ij}_{ab} \exp[-s(\Delta^{ij}_{ab})^2], \tag{26}$$

the kth-order (k = 1, 2) DSRG flow equation, $[\bar{H}^{(k)}(s)]^{\mathrm{N}} = \hat{R}^{(k)}(s)$, yields the following amplitude equations:

$$t^{i,(k)}_a(s) = [\bar{f}^{i,(k)}_a(s) + \sum_{ux}^{\mathbf{A}} \Delta^x_u \bar{\gamma}^x_u t^{iu,(k)}_{ax}(s)] \frac{1 - e^{-s(\Delta^i_a)^2}}{\Delta^i_a}, \tag{27}$$

$$t^{ij,(k)}_{ab}(s) = v^{ij,(k)}_{ab}(s) \frac{1 - e^{-s(\Delta^{ij}_{ab})^2}}{\Delta^{ij}_{ab}}. \tag{28}$$

The quantities $\bar{f}^{i,(k)}_a$ and $v^{ij,(k)}_{ab}$ are the non-diagonal components of $\bar{H}^{(k)}(s) - [\hat{H}^{(0)}, \hat{A}^{(k)}(s)]$, and $\Delta^{ij}_{ab} = \epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b$ is a Møller–Plesset denominator. Equations (27) and (28) show that, for finite value of s, when a denominator approaches zero, then the cluster amplitude smoothly goes to zero (rather than diverge). We emphasize that computing the scalar term of the second- and third-order energy [Eqs. (23) and (24)] requires one-, two-, and three-body cumulants of the reference.[50,51] Consequently, if we exclude the cost to evaluate the state-averaged one-, two-, and three-body density cumulants, the scaling with respect to the number of active orbitals of the SA-DSRG-PT2/3 is at worst $\mathcal{O}(N_{\mathbf{V}} N^6_{\mathbf{A}})$.

4. Ground- and excited-state energies

The SA-DSRG-PTX (X = 2, 3) energies are obtained by diagonalizing the DSRG Hamiltonian summed up to Xth-order

$$\bar{H}^{[X]}(s) = \bar{H}^{(0)} + \sum_{i=1}^{X} \bar{H}^{(i)}(s). \tag{29}$$

Here we consider two diagonalization methods. The first one is a contracted scheme, which we indicate with the name SA-DSRG-PTXc. In this approach, we diagonalize $\bar{H}^{[X]}(s)$ within the space of basis states in the ensemble $\mathbb{E}_0$,

$$\sum_{\beta}^{n} \langle \Psi^\alpha_0 | \bar{H}^{[X]}(s) | \Psi^\beta_0 \rangle \, C^\xi_\beta = C^\xi_\alpha E_\xi(s), \quad \xi = 1, \ldots, n, \tag{30}$$

which yields the energy of state ξ [$E_\xi(s)$] and the corresponding zeroth-order wave function

$$|\Psi^\xi_{0'}\rangle = \sum_{\alpha}^{n} C^\xi_\alpha |\Psi^\alpha_0\rangle = \sum_{I=1}^{N_{\mathrm{CAS}}} \Big(\sum_{\alpha}^{n} C^\xi_\alpha c^\alpha_I \Big) |\Phi^I\rangle. \tag{31}$$

This may be considered an unrelaxed state averaged approach since the weights of determinants in the final wave function are partially constrained by the coefficients in the configuration-interaction expansion of the reference states.

The second option is an uncontracted scheme in which we diagonalize $\bar{H}(s)$ in the space of CAS determinants $\{\Phi^I, I = 1, 2, \ldots, N_{\mathrm{CAS}}\}$,

$$\sum_{J}^{N_{\mathrm{CAS}}} \langle \Phi_I | \bar{H}^{[X]}(s) | \Phi^J \rangle \tilde{C}^\xi_J = \tilde{C}^\xi_I \tilde{E}_\xi(s), \quad \xi = 1, \ldots, n. \tag{32}$$

The corresponding wave function is

$$|\tilde{\Psi}^\xi_{0'}\rangle = \sum_{I}^{N_{\mathrm{CAS}}} \tilde{C}^\xi_I |\Phi^I\rangle, \tag{33}$$

which is more flexible since each determinant has an associated coefficient that is optimized.

In this work, we truncate the $\bar{H}^{[X]}(s)$ operator after two-body terms, but more accurate schemes that include three- and higher-body terms are possible. The $\bar{H}^{[X]}(s)$ operator is then completely specified by a scalar term and quantities analogous to the one- and two-electron integrals of the bare Hamiltonian [$\bar{H}^{q,[X]}_p(s)$ and $\bar{H}^{rs,[X]}_{pq}(s)$]. In practice, to diagonalize the operator $\bar{H}^{[X]}(s)$ in the basis of CASCI determinants or CASCI solutions, we first rewrite it in a normal-ordered form with respect to the true vacuum. In this form, $\bar{H}^{[X]}(s)$ may be diagonalized using a standard full CI code.

When the ensemble contains only one state, the uncontracted diagonalization approach [Eq. (32)] is equivalent to the partially relaxed state-specific DSRG-MRPT2 method discussed in Ref. 52. Both the contracted and uncontracted SA-DSRG-MRPTX schemes are computational advantageous because they have a cost that is nearly identical to that of one state-specific unrelaxed DSRG-MRPTX computation. The additional cost of the SA-DSRG-MRPTX comes from computing the active one- and two-body components of $\bar{H}^{[X]}(s)$ [$\bar{H}^{v,[X]}_u(s)$ and $\bar{H}^{xy,[X]}_{uv}(s)$ with $u, v, x, y \in \mathbf{A}$] and diagonalizing this operator in the basis of the contracted or uncontracted determinant basis. Generating $\bar{H}^{[X]}(s)$ has a small computational cost that scales as $\mathcal{O}(N^2_{\mathbf{V}} N^4_{\mathbf{A}})$. The costs to diagonalize $\bar{H}^{[X]}(s)$ depends on the type of reference used. In the case of CASCI, the computational complexity for diagonalization is a factorial of the number of active orbitals. However, the SA-DSRG-PT methods may also use approximate references that have polynomial scaling with respect to the number of active orbitals. Moreover, it is found that computing the Xth-order energy does not require evaluating the Xth-order amplitudes.[52] More complex schemes that treat relaxation effects via the iterative solution of the DSRG amplitude and eigenvalue equations are not considered here.

It is easy to see that both the contracted and uncontracted SA-DSRG-PT approaches avoid the intruder-state problem. In these methods, the energy is obtained by diagonalizing the

DSRG Hamiltonian $\bar{H}^{[X]}(s)$, computed by contracting the one- and two-electron integrals, the cluster amplitudes, and the SA density cumulants. As mentioned in Sec. II B 3 (see Sec. 2.4 in Ref. 50 for more details), the SA-DSRG-PT amplitudes are bounded [see Eqs. (27) and (28)] for any finite value of s. Moreover, the one- and two-electron integrals and the SA density cumulants are also bounded and independent of s. As a consequence, for finite values of s, the DSRG Hamiltonian $\bar{H}^{[X]}(s)$ is bounded and its eigenvalues cannot diverge.

The SA-DSRG-PTX amplitude equations and the matrix elements of the $\bar{H}^{[X]}(s)$ Hamiltonian are diagrammatically connected. This is a consequence of the fact that the DSRG Hamiltonian can be expressed as a series of commutators of $\hat{H}$ and the operator $\hat{A}(s)$.[53] If the reference wave function is chosen appropriately, then the SA-DSRG-PTX energy is size extensive, and the energy of noninteracting fragments is the sum of fragment energies. We tested size consistency for a pair of noninteracting ethylene molecules, studying the $\pi \rightarrow \pi^*$ transition. Both size consistency and size intensivity of the excitation energies were verified within numerical accuracy.[81]

The SA-DSRG-PTX (X = 2, 3) approach can be easily implemented by modifying a state-specific DSRG-MRPTX code. Specifically, the Fock matrix and cumulants of DSRG-MRPTX for individual reference states are replaced by the corresponding state-averaged quantities. As a result, SA-DSRG-PTX is as computationally efficient as the state-specific version of DSRG-MRPTX.[50–52] The SA-DSRG-PT2 method is also simpler than other internally contracted multi-state MRPT2 approaches. For instance, both MS-CASPT2 and QD-NEVPT2 need to solve for each root either separately[32,33] or in a coupled way.[40] In SA-DSRG-PT2 (and SA-DSRG-PT3), however, cluster amplitudes are solved only once for all states and this procedure requires only the one-, two-, and three-body SA density cumulants. As such, SA-DSRG-PTX is easily applicable to systems with more than 30 active orbitals.

5. Static properties

Static properties can also be conveniently evaluated under the state-averaged MRDSRG framework. Taking the dipole operator ($\hat{\boldsymbol{\mu}}$) as an example, we start by writing it in a normal-ordered form as

$$\hat{\boldsymbol{\mu}} = \boldsymbol{\mu}_0 + \sum_{pq}^{\mathbf{G}} \boldsymbol{\mu}_p^q \{\hat{a}_q^p\}_\rho. \tag{34}$$

Here $\boldsymbol{\mu}_p^q = -\langle \phi_p | \mathbf{r} | \phi^q \rangle$ are the dipole integrals in the molecular orbital basis and the vector $\boldsymbol{\mu}_0 = \sum_{ij}^{\mathbf{H}} \boldsymbol{\mu}_i^j \bar{\gamma}_j^i$ is the ensemble expectation value of the dipole operator. Dynamic correlation effects are folded in the dipole operator like in the case of the Hamiltonian. The resulting effective dipole operator $[\bar{\boldsymbol{\mu}}(s)]$ is given by

$$\bar{\boldsymbol{\mu}}(s) = \hat{U}^\dagger(s) \hat{\boldsymbol{\mu}} \hat{U}(s). \tag{35}$$

Considering the bare dipole operator as a zeroth-order quantity, we can derive an order-by-order expansion of the transformed dipole operator

$$\bar{\boldsymbol{\mu}}^{(0)}(s) = \hat{\boldsymbol{\mu}}, \tag{36}$$

$$\bar{\boldsymbol{\mu}}^{(1)}(s) = [\hat{\boldsymbol{\mu}}, \hat{A}^{(1)}(s)], \tag{37}$$

$$\bar{\boldsymbol{\mu}}^{(2)}(s) = [\hat{\boldsymbol{\mu}}, \hat{A}^{(2)}(s)] + \frac{1}{2}[\bar{\boldsymbol{\mu}}^{(1)}(s), \hat{A}^{(1)}(s)], \tag{38}$$

$$\bar{\boldsymbol{\mu}}^{(3)}(s) \approx [\hat{\boldsymbol{\mu}}, \hat{A}^{(3)}(s)] + \frac{1}{2}[\bar{\boldsymbol{\mu}}^{(1)}(s), \hat{A}^{(2)}(s)] + \frac{1}{2}[\tilde{\boldsymbol{\mu}}_{1,2}^{(2)}(s), \hat{A}^{(1)}(s)], \tag{39}$$

where $\tilde{\boldsymbol{\mu}}_{1,2}^{(2)}(s) = \bar{\boldsymbol{\mu}}_{1,2}^{(2)}(s) - \frac{1}{6}[\bar{\boldsymbol{\mu}}^{(1)}(s), \hat{A}^{(1)}(s)]_{1,2}$ and, consistently with the treatment of the energy, we discard three- and higher-body operators.

Accordingly, the SA-DSRG-PTX (X = 2, 3) transformed dipole operator is given by

$$\bar{\boldsymbol{\mu}}^{[2]}(s) = \boldsymbol{\mu} + \bar{\boldsymbol{\mu}}^{(1)}(s) + \frac{1}{2}[\bar{\boldsymbol{\mu}}^{(1)}(s), \hat{A}^{(1)}(s)], \tag{40}$$

$$\bar{\boldsymbol{\mu}}^{[3]}(s) = \boldsymbol{\mu} + \bar{\boldsymbol{\mu}}^{(1)}(s) + \bar{\boldsymbol{\mu}}^{(2)}(s) + \frac{1}{2}[\bar{\boldsymbol{\mu}}^{(1)}(s), \hat{A}^{(2)}(s)] + \frac{1}{2}[\tilde{\boldsymbol{\mu}}_{1,2}^{(2)}(s), \hat{A}^{(1)}(s)]. \tag{41}$$

Notice that we ignore the $[\hat{\boldsymbol{\mu}}, \hat{A}^{(X)}(s)]$ term in $\bar{\boldsymbol{\mu}}^{[X]}(s)$ because $\boldsymbol{\mu}$ is partitioned differently than $\hat{H}$. In the case of the Hamiltonian, contributions from the commutator $[\hat{H}^{(0)}, \hat{A}^{(X)}(s)]$ will not enter the SA-DSRG-PTX energy expression.[52] Hence, we discard the analogous contribution from the dipole operator. Moreover, we do not consider the orbital response with respect to the external field. This approach to evaluate the dipole moment operator requires the transformation of each component of the dipole operator. Generally, it is more efficient to generate the SA-DSRG density and then contract it with the MO dipole integrals. However, to evaluate the density, explicit expressions are required that involve doubly and triply nested commutators. As such, the explicit evaluation of the density is not explored in this work.

Permanent and transition dipole moments can be computed as matrix elements of the effective dipole operator with respect to the basis of eigenstates of $\bar{H}^{[X]}$. For example, if the energies are obtained by Eq. (32), the dipole moment will be given by $\bar{\boldsymbol{\mu}}_{\alpha\beta}^{[X]}(s) = \langle \tilde{\Psi}_{0'}^{\alpha} | \bar{\boldsymbol{\mu}}^{[X]}(s) | \tilde{\Psi}_{0'}^{\beta} \rangle$. This expression can be explicitly written as

$$\begin{aligned}
\bar{\boldsymbol{\mu}}_{\alpha\beta}^{[X]}(s) = \delta_\alpha^\beta \Big[& \bar{\boldsymbol{\mu}}_0^{[X]}(s) - \sum_{uv}^{\mathbf{A}} (\bar{\boldsymbol{\mu}}^{[X]})_u^v(s) \bar{\gamma}_v^u \\
& - \sum_{uvxy}^{\mathbf{A}} (\bar{\boldsymbol{\mu}}^{[X]})_{uv}^{xy}(s) \Big(\frac{1}{4} \bar{\lambda}_{xy}^{uv} - \frac{1}{2} \bar{\gamma}_x^u \bar{\gamma}_y^v \Big) \Big] \\
& + \sum_{uv}^{\mathbf{A}} \Big[(\bar{\boldsymbol{\mu}}^{[X]})_u^v(s) - \sum_{xy}^{\mathbf{A}} (\bar{\boldsymbol{\mu}}^{[X]})_{ux}^{vy}(s) \bar{\gamma}_x^y \Big] [\boldsymbol{\gamma}'_{\alpha\beta}]_v^u \\
& + \frac{1}{4} \sum_{uvxy}^{\mathbf{A}} (\bar{\boldsymbol{\mu}}^{[X]})_{uv}^{xy}(s) [\boldsymbol{\gamma}'_{\alpha\beta}]_{xy}^{uv},
\end{aligned} \tag{42}$$

where we have introduced the Kronecker delta δ_α^β and the one- and two-body transition density matrices $[\boldsymbol{\gamma}'_{\alpha\beta}]_v^u = \langle \tilde{\Psi}_{0'}^{\alpha} | \hat{a}_v^u | \tilde{\Psi}_{0'}^{\beta} \rangle$ and $[\boldsymbol{\gamma}'_{\alpha\beta}]_{xy}^{uv} = \langle \tilde{\Psi}_{0'}^{\alpha} | \hat{a}_{xy}^{uv} | \tilde{\Psi}_{0'}^{\beta} \rangle$. Note that the averaged densities in Eq. (42) result from writing $\bar{\boldsymbol{\mu}}$ in a normal-ordered form with respect to the true vacuum and are

computed from the density matrices of the CASCI/CASSCF zeroth-order states [Eq. (8)].

III. RESULTS

We tested the SA-DSRG-PT2 and SA-DSRG-PT3 methods on several benchmark systems: lithium fluoride, ammonia, the penta-2,4-dieniminium cation, and (E,E)-1,3,5,7-octatetraene. The results are compared to those obtained with extended MS-CASPT2 (XMS-CASPT2),[2,32,40,82] quasi-degenerate strongly contracted NEVPT2 (QD-sc-NEVPT2),[5,33] extended MC-QDPT2 (XMCQDPT2),[31,37] complete active space third-order perturbation theory (CASPT3),[82] internally contracted MRCI with singles and doubles (ic-MRCISD) as formulated by Knowles and Werner,[8,83] and ic-MRCISD with Davidson correction (ic-MRCISD+Q).[84,85] The DSRG flow parameter was set to 0.5 E_h^{-2}, a value that previous studies[50,55] showed to yield good agreement with benchmark results. Unless noted, the XMS-CASPT2 level shift[46] was set to 0.1 E_h. We also used a 0.02 E_h^2 intruder-state-avoidance shift[86] in all XMCQDPT2 computations. The 1*s*-like orbitals of second-period elements were excluded from all post-CASSCF treatment of electron correlation.

The SA-DSRG perturbation theories were implemented in Forte,[87] an open-source plugin to Psi4[88] for developing novel multireference methods. The QD-sc-NEVPT2 and XMCQDPT2 energies were obtained using the Orca 4.0[89] and Gamess 2016[90] packages, respectively. The XMS-CASPT2, CASPT3, and ic-MRCISD computations were carried out using the Molpro 2015.1 package.[91]

A. Lithium fluoride

We first consider the lowest two $^1\Sigma^+$ states of lithium fluoride. This system is a classic example of ionic-neutral avoided crossing and a testbed for many quasi-degenerate and multi-state perturbation theories.[31–33,35,37,40,92] The computed distance of the avoided crossing is sensitive to many factors, including the active space,[93] the basis set, the frozen-core approximation,[94] and the relative energy of the first two states.[95] Going from the equilibrium geometry to the dissociation limit, the ground state wave function reduces its ionic character (due to the $|1\sigma^2 2\sigma^2 3\sigma^2 1\pi^4 4\sigma^2\rangle$ configuration) and becomes dominated by the covalent configuration $|1\sigma^2 2\sigma^2 3\sigma^2 1\pi^4 4\sigma^1 5\sigma^1\rangle$. Thus, the minimal active space required to describe the dissociation of LiF consists of the 4σ and 5σ orbitals. However, a state-averaged CASSCF(2,2) wave function underestimates the electron affinity of fluorine[93] and it is an inadequate starting point for perturbation theories.[33] In accordance with Varandas,[94] we employ a larger active space consisting of six electrons in seven orbitals (three a_1, two b_1, and two b_2 orbitals in C_{2v} symmetry). A mixed aug(F)-cc-pVTZ basis set was used, which was built from the cc-pVTZ and aug-cc-pVTZ basis sets for Li[96] and F,[97] respectively. The reference wavefunctions and molecular orbitals were obtained by a SA-CASSCF procedure, averaging over the lowest two $^1\Sigma^+$ states, which we denote as SA2-CASSCF(6e,7o).

The potential energy curves (PECs) for the lowest two $^1\Sigma^+$ states of LiF near the avoid crossing region are shown in Fig. 1. It is well-known that the state-specific MRPTs cannot correctly describe the avoided crossing of LiF.[25,31,32] Indeed, an unphysical double crossing is observed between the lowest two $^1\Sigma^+$ states of the state-specific unrelaxed DSRG-MRPT2 (u-DSRG-MRPT2).[50,52] Contrarily, all SA-DSRG perturbation theories correctly predict an avoided crossing. As expected from our discussion in Sec. II B 4, in the weak-mixing region, we see good correspondence between the state-specific (u-DSRG-MRPT2) and state-averaged theory with constrained coefficients (SA-DSRG-PT2c). We also note that the larger diagonalization space used in the uncontracted approach (SA-DSRG-PT2) helps recover more electron correlation but shifts the avoided crossing point towards a shorter distance.

Results for a number of MRPT2 methods are shown in the right panel of Fig. 1. When compared to SA-DSRG-PT2, other multi-state second-order perturbation theories recover more correlation energy. The energy gap and the position of the avoided crossing (minimum energy gap) predicted by SA-DSRG-PT2 are similar to those of QD-sc-NEVPT2. In comparison, based on these criteria, XMS-CASPT2 and

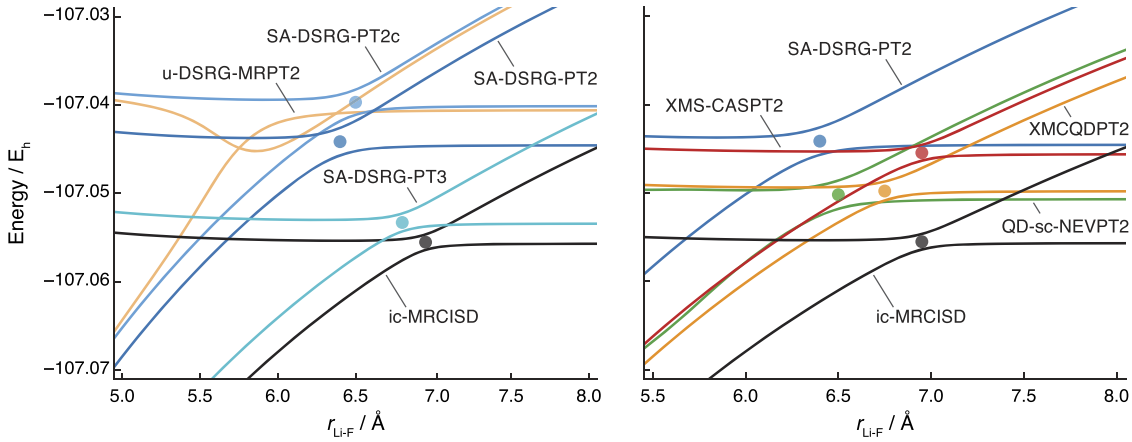

FIG. 1. Potential energy curves for the lowest two $^1\Sigma^+$ states of lithium fluoride computed using various methods and the aug(F)-cc-pVTZ basis set. For each method, a circle marks the average energy of the two states at the bond length with the smallest energy gap.

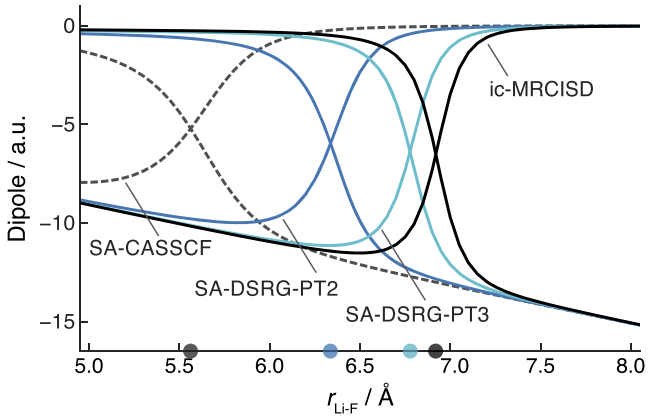

FIG. 2. Dipole moment for the lowest two $^1\Sigma^+$ states of lithium fluoride computed using various methods and the aug(F)-cc-pVTZ basis set. The circles on the r_{Li-F} axis indicate the crossing points.

XMCQDPT2 show better agreement with ic-MRCISD. Note that on the PECs of QD-sc-NEVPT2, there are small humps around 5.5 Å due to the model-space invariance problem of the zeroth-order Hamiltonian.[37,40] These artificial humps are cured by the extended multi-state methods like XMCQDPT2 and XMS-CASPT2.

Going from second- to third-order, we observe a significant improvement of the agreement between DSRG perturbation theory and ic-MRCISD. To highlight the avoided crossing point, in Fig. 2 we report the SA-DSRG-PT2 and -PT3 dipole moment of the two lowest $^1\Sigma^+$ states of LiF. The SA2-CASSCF(6e,7o) reference yields a flawed crossing point, at a bond distance that is 1.5 Å too short with respect to the ic-MRCISD result. This defect is corrected by the inclusion of dynamic correlation; however, a third-order SA-MRDSRG treatment is required to achieve very good agreement with ic-MRCISD.

B. Ammonia

Next, we study the ground ($\widetilde{X}$, $^1A_1$) and first singlet excited ($\widetilde{A}$, $^1A_2''$) states of ammonia, which are important in understanding the photodissociation process $NH_3 \rightarrow NH_2 + H$.[99–104] This is a classical example of systems with a conical intersection between the ground and first excited state, which are notoriously difficult for excited state methods like time-dependent density functional theory (TD-DFT)[105] and equation-of-motion coupled cluster theory.[106–108] On the ground-state PES, two C_{3v} minima can be interconverted via a planar D_{3h} transition state (TS). The PES of the $\widetilde{A}$ state has a D_{3h} symmetry minimum and a C_{2v} transition state for hydrogen dissociation. Several conical intersections (CIs) that connect these surfaces at C_{2v} or C_s geometries have been located.[103] Here we consider the C_{2v} CI using the coordinate system introduced by Truhlar and co-workers.[104,108,109] As shown in Fig. 3, the geometry of NH_3 is specified by two coordinates: the bond length of N—H_1 (r_{N-H_1}) and the pyramidalization angle β. The other two N—H bonds are set equal to 1.039 Å. The geometry is such that for each bond N—H_i (i = 1, 2, 3), the corresponding angle $\angle H_i NN'$ is equal to β. When $\beta = 90°$, the geometry is planar, while other values of β give pyramidal geometries. Moreover, the angle between any

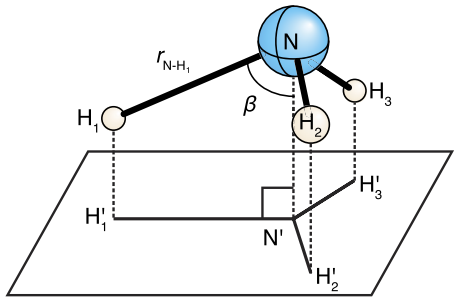

FIG. 3. The coordinate system of NH_3. The bond length of N—H_1 (r_{N-H_1}) and the pyramidalization angle β are the varying coordinates. The bond lengths of N—H_2 and N—H_3 are fixed at 1.039 Å, and all projected bond angles ($\angle H_i' N' H_j'$, $i, j \in [1, 2, 3]$, $i \neq j$) are equal to 120°. The NH_3 molecule is drawn using CHEMVP.[98]

two N—H_i bonds projected onto the plane is constrained to 120°. For more details, we refer the readers to Refs. 104, 108, and 109.

To assess the accuracy of SA-MRDSRG perturbation theories, we employed a full-valence SA2-CASSCF(8e,7o) reference and the aug-cc-pVTZ basis set.[97] We also carried out TD-DFT computations (Tamm-Dancoff approximation) with the Becke's three-parameter exchange[110] and Lee-Yang-Parr correlation[111] (B3LYP) functional as implemented in ORCA 4.0.[89]

Figure 4 displays adiabatic PESs of the lowest two singlet states of NH_3 near the CI region. In accordance to previous work,[104,108,109] all MR theories predict a double-cone CI except for TD-B3LYP. Among these MR theories, those accounting for dynamic electron correlation yield a noticeably lower barrier for the ground state interconversion of the two equivalent pyramidalized geometries at shorter N—H_1 bond lengths. For example, when r_{N-H_1} = 1.8 Å, the SA2-CASSCF(8e,7o) energy difference between $\beta = 90°$ and $\beta = 60°$ is 1.45 eV, which is 0.30 and 0.26 eV higher than the SA-DSRG-PT2 and ic-MRCISD values, respectively. The inclusion of dynamic correlation also moves the location of the CI to larger N—H_1 bond distances. In comparison to ic-MRCISD, both SA-DSRG-PT2 and XMS-CASPT2 slightly underestimate the energetics in the vicinity of the CI, as visible in the ground-state contour plots shown in Fig. 4. Quantitative agreement with ic-MRCISD for the energetics and position of the CI is instead achieved by the SA-DSRG-PT3 method.

C. Penta-2,4-dieniminium cation

The third system examined here is the penta-2,4-dieniminium cation [PSB3, $^+NH_2(CH)_4CH_2$], a prototype of the retinal protonated Schiff base (PSB) chromophore that has been studied extensively.[112–116] In particular, we test the SA-DSRG-PT2 and SA-DSRG-PT3 methods on the three pathways introduced by Olivucci and co-workers[112] using ground-state SA2-CASSCF(6,6)/6-31G* optimized geometries. As shown in Fig. 5, two pathways directly connect the *cis* and *trans* isomers of PSB3 and correspond to (i) the charge-transfer minimum-energy path (MEP_{CT}) via a TS (TS_{CT}) with enhanced electron density on the N=C_1 bond and (ii) the diradical MEP (MEP_{DIR}) with a TS (TS_{DIR}) that has two unpaired electrons on the π orbitals of the C_5–C_4–C_3 and

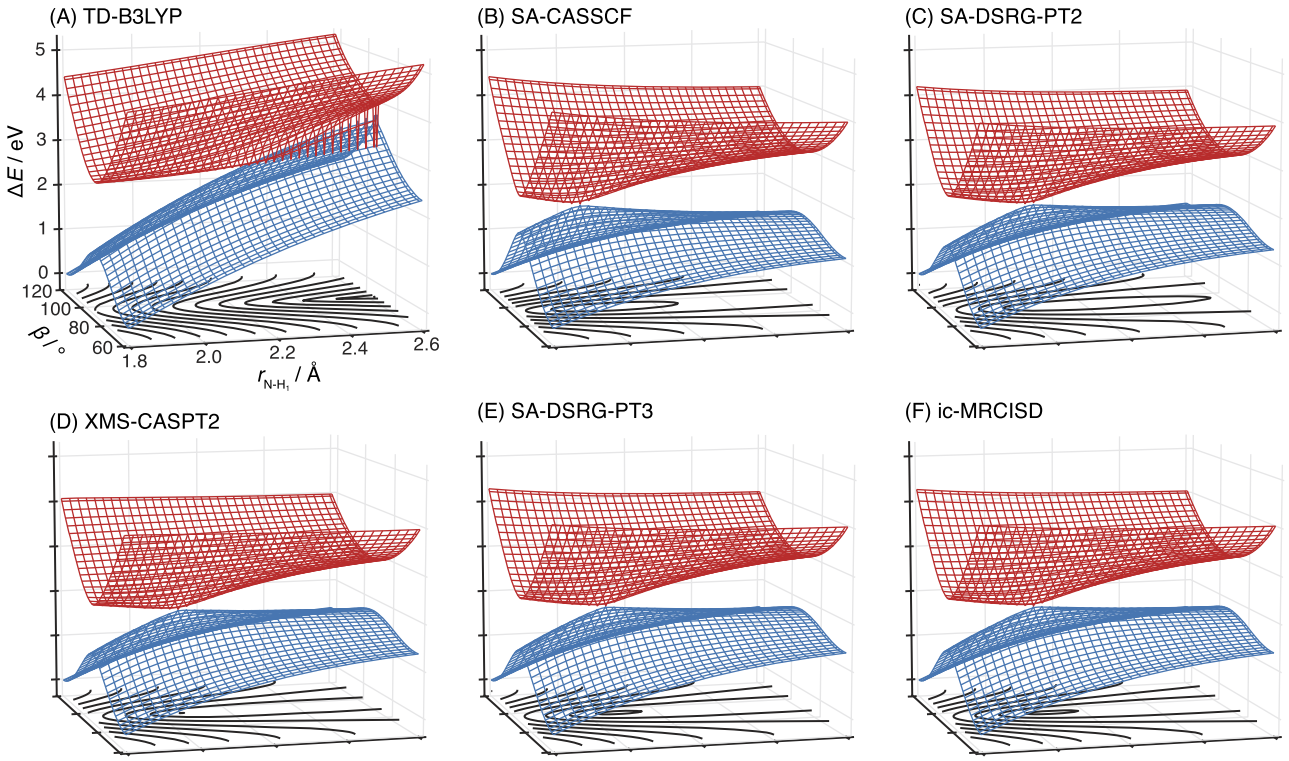

FIG. 4. Potential energy surfaces (relative to the ground-state NH_3 with $r_{N-H_1} = 1.80$ Å and $\beta = 60.0°$) for the lowest two singlet states of ammonia computed using various methods and the aug-cc-pVTZ basis set. Contour lines (0.25 eV interval) are also used to represent the ground-state potential energy surface.

C_2–C_1–N backbones. The third pathway connects the charge-transfer TS_{CT} and diradical TS_{DIR} via a conical intersection.

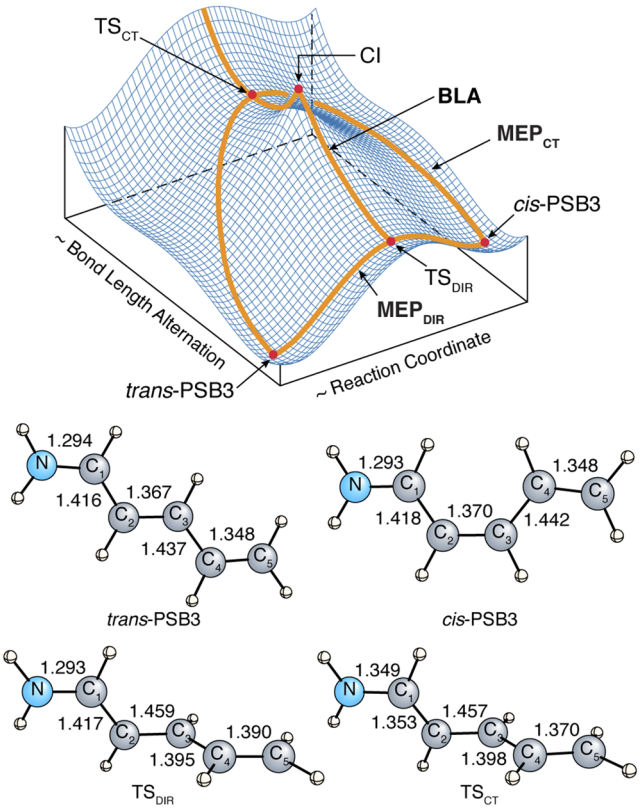

FIG. 5. Schematic potential energy surface and ground-state CASSCF(6,6)/6-31G* equilibrium geometries of PSB3 from Ref. 112. The bond lengths are in Ångstrom. Molecules are drawn using CHEMVP.[98]

This pathway is characterized by the bond length alternation coordinate [BLA, defined as $\frac{1}{2}(r_{C_1-C_2} + r_{C_3-C_4}) - \frac{1}{3}(r_{N=C_1} + r_{C_2=C_3} + r_{C_4=C_5})$]. To allow direct comparison with the results of Ref. 112, we use a SA2-CASSCF(6,6) reference and the 6-31G* basis set.[117]

Previous work[112] showed that SA2-CASSCF(6,6) overestimates the energy of PSB3 charge-transfer states and yields an incorrect location of the conical intersection. Including dynamic correlation improves the description of these states and moves the CI closer to TS_{DIR}. As such, TS_{DIR} becomes a minimum on the excited potential energy surface. As shown in the potential energy profiles in Fig. 6, both the SA-DSRG-PT2 and SA-DSRG-PT3 correctly reproduce these features. For example, along the BLA coordinate [Fig. 6(a)], both SA-DSRG perturbation theories predict the CI at large BLA values, in agreement with results from XMS-CASPT2, ic-MRCISD, and ic-MRCISD+Q. In Fig. 6(b), we see a drastic decrease of the ground-state barrier height along the MEP_{CT} coordinate when including dynamic correlation. For instance, the SA2-CASSCF(6,6) barrier (58.7 kcal mol^{-1}) is reduced by 9.8, 12.1, 7.2, 6.3, and 10.0 kcal mol^{-1} for SA-DSRG-PT2, XMS-CASPT2, SA-DSRG-PT3, ic-MRCISD, and ic-MRCISD+Q, respectively. We also notice that the SA-DSRG-PT3 results for PSB3 are generally consistent with those from ic-MRCISD, while, fortuitously, the SA-DSRG-PT2 curves closely follow those computed with ic-MRCISD+Q.

D. (E,E)-1,3,5,7-octatetraene

Our final example is the (E,E)-1,3,5,7-octatetraene (OTE) molecule. An accurate description for the low-lying excited states of OTE has long been a difficult problem for

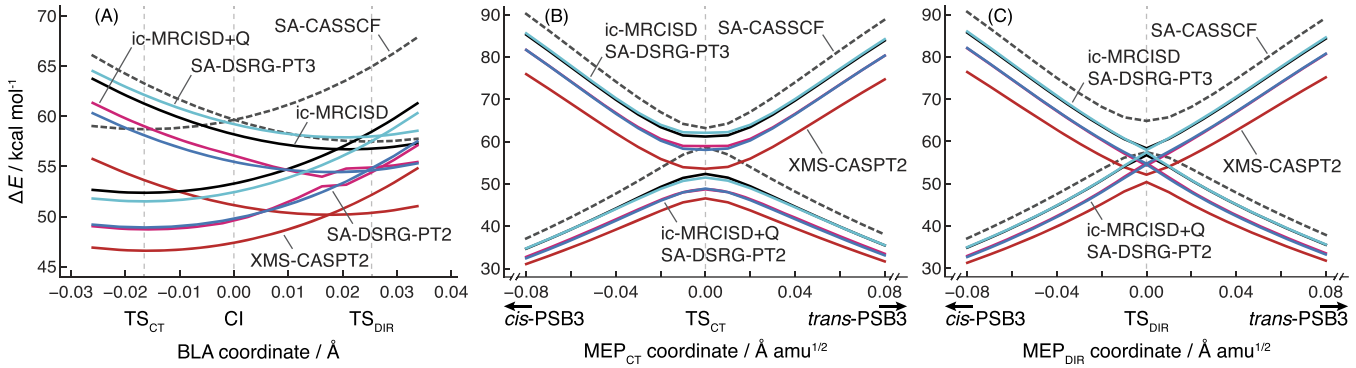

FIG. 6. Energy (relative to *cis*-PSB3) profiles for the lowest two singlet states of PSB3 along the (a) bond length alternation (BLA), (b) charge-transfer minimum-energy path ($\mathrm{MEP_{CT}}$), and (c) diradical minimum-energy path ($\mathrm{MEP_{DIR}}$) coordinates computed using various methods with the 6-31G* basis set. The CASSCF(6,6)/6-31G* optimized geometries (taken from Ref. 112) of $\mathrm{TS_{CT}}$, $\mathrm{TS_{DIR}}$, and CI are indicated by vertical dashed lines.

theory.[23,24,118–126] For instance, the $2\,^1\mathrm{A}_g^-$ state possesses strong double-excitation character and it is poorly described by single-reference linear response methods.[123] Another example is the $1\,^1\mathrm{B}_u^+$ state, which can be considered as an intramolecular charge-transfer state, and, as such, it requires careful considerations of dynamic correlation effects.[120,123,125] Here, states are labeled with Mulliken symbols for the irreducible representations of the C_{2h} point group plus a superscript "+/−" to indicate the particle-hole symmetry (i.e., +/− state shows mainly ionic/covalent character, respectively).[118,119,121,127]

We computed the vertical excitation energies of several low-lying singlet valence excited states of OTE. A series of SA-CASSCF references were considered and they are generally denoted as SA8-CASSCF(8e,*n*o) with n = 8, 10, 12, 14, 16. The smallest active space, CASSCF(8e,8o), involves all π and π^* (four b_g and four a_u) orbitals. Then, more active orbitals were included in a step-by-step manner where two more orbitals (one b_g and one a_u) were considered in each step until the "double-π" active space (eight b_g and eight a_u) is reached.[123] In all cases, the CASSCF orbitals were optimized by averaging over eight states (five $^1\mathrm{A_g}$ and three $^1\mathrm{B_u}$). Since only valence excited states were of interests, the basis sets were chosen not to include any diffuse functions. Specifically, we employed the def2-TZVP[128] and cc-pVQZ[129] basis sets. In all DSRG computations, the two-electron integrals were factorized using the pivoted incomplete Cholesky decomposition[51,130–133] with a 10^{-6} a.u. threshold for def2-TZVP and 10^{-4} a.u. for cc-pVQZ. The geometry of OTE, optimized at the MP2/6-31G* level of theory, was taken from Ref. 23.

To help identify the "+/−" character of a state, we localized the SA-CASSCF active orbitals using the Pipek–Mezey scheme based on Mulliken charges.[135] The resulting localized molecular orbitals (LMOs) are shown in Fig. 7. A state is then labeled as "+" (ionic) if two electrons share a common LMO in its dominant configurations. Conversely, for a "−" (covalent) state, there is at most one electron occupying each of the LMOs in those important configurations.

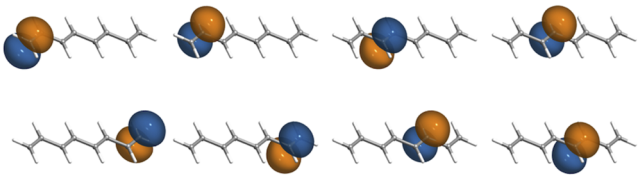

FIG. 7. The set of Pipek–Mezey localized active orbitals computed at the SA8-CASSCF(8e,8o)/def2-TZVP level of theory. Orbitals are drawn using VMD.[134]

The excitation energy ($\omega_{\alpha\leftarrow 0}$) is obtained by subtracting the ground-state energy (E_0) from the energy of an excited state α (E_α),

$$\omega_{\alpha\leftarrow 0} = E_\alpha - E_0, \tag{43}$$

and the corresponding oscillator strength $f_{\alpha\leftarrow 0}$ is computed as

$$f_{\alpha\leftarrow 0} = \frac{2}{3}|\boldsymbol{\mu}_{\alpha\leftarrow 0}|^2 \omega_{\alpha\leftarrow 0}, \tag{44}$$

where $\boldsymbol{\mu}_{\alpha\leftarrow 0} = \bar{\boldsymbol{\mu}}^{[X]}_{\alpha 0}(s)$ is the transition dipole moment. In CASPT2, it is often the case that E_0 is chosen to come from a multistate computation if the excited state has the same symmetry as the ground state, otherwise, the value from state-specific computation is used. In practice, the difference that arises from these two choices is small (0.04 eV for OTE).[24] Since the ensemble formalism at the basis of the SA-MRDSRG approach allows us to consider states with different symmetries, we always compute excitation energies from the SA-DSRG-PTX (X = 2, 3) eigenvalues of the DSRG Hamiltonian, a treatment that is consistent with the definition of transition dipole moment [Eq. (42)].

Table I presents the vertical excitation energies of OTE along with the oscillator strengths of several valence excited states computed with SA8-CASSCF(8e,8o) reference using the def2-TZVP basis set. It is well-known that dynamic correlation effects are essential to account for σ polarization of ionic states.[136,137] Indeed, all post-CASSCF methods yield significantly lower excitation energies than the corresponding CASSCF values for the ionic $1\,^1\mathrm{B}_u^+$ and $1\,^1\mathrm{A}_g^+$ states. For example, the SA-DSRG-PT2 excitation energy for the $1\,^1\mathrm{B}_u^+$ state (4.85 eV) is 1.8 eV lower than that of CASSCF (6.65 eV). Both SA-DSRG-PT2 and SA-DSRG-PT3 predict the following ordering:

$$2\,^1\mathrm{A}_g^- < 1\,^1\mathrm{B}_u^+ < 1\,^1\mathrm{B}_u^- < 3\,^1\mathrm{A}_g^- < 1\,^1\mathrm{A}_g^+ < 4\,^1\mathrm{A}_g^-,$$

which agrees with that of ic-MRCISD+Q but differs slightly from the ic-MRCISD ordering ($4\,^1\mathrm{A}_g^- < 1\,^1\mathrm{A}_g^+$). Compared

TABLE I. Vertical excitation energies (in eV) of several low-lying valence excited states of octatetraene based on the SA8-CASSCF(8e,8o)/def2-TZVP reference. The oscillator strengths (in a.u.) are given in parentheses. All excitation energies are from this work. The geometry was optimized using MP2/6-31G*, taken from Ref. 23.

							SA-DSRG				
State	CASSCF[a]	NEVPT2[b]	QD-NEVPT2[b]	XMCQDPT2	XMS-CASPT2[c]	CASPT3[d]	PT2c	PT2	PT3	ic-MRCISD[e]	ic-MRCISD+Q[f]
$2\,^1A_g^-$	4.68	4.75	4.74	4.48	4.42	4.62	4.60	4.57	4.75	4.72	4.71
$1\,^1B_u^+$	6.65 (1.861)	4.02	4.02 (1.127)	4.26	4.18	4.85	4.69	4.85(1.742)	5.41(1.789)	5.76(1.605)	5.29
$1\,^1B_u^-$	5.86 (0.000)	5.99	5.99 (0.001)	5.69	5.40	5.82	5.80	5.77(0.001)	5.98(0.002)	5.92(0.015)	5.92
$3\,^1A_g^-$	6.59	6.73	6.73	6.48			6.55	6.52	6.73	6.64	6.62
$1\,^1A_g^+$	8.94	6.08	6.08	5.77			6.83	6.92	7.58	7.91	7.27
$4\,^1A_g^-$	7.49	7.75	7.76	7.38			7.48	7.44	7.70	7.59	7.61
$2\,^1B_u^-$	8.26 (0.000)	8.54	8.54 (0.000)	8.21			8.27	8.25 (0.002)	8.50 (0.000)	8.37 (0.000)	8.38

[a]Computed using Molpro and exact two-electron integrals.
[b]Strongly contracted variant computed in this work. Oscillator strengths are approximated using CASSCF transition dipole moments but QD-sc-NEVPT2 energies.
[c]Only two ^1A_g and two ^1B_u states are considered in XMS-CASPT2 with a 0.2 level shift.
[d]Obtained using a 0.2 E_h level shift.
[e]Excitation energies are computed in C_{2h} symmetry. Oscillator strengths are obtained by a separate computation in C_s symmetry.
[f]Rotated reference coefficients and energies are used to compute cluster corrections. Coefficients of rotated reference functions are all ≥0.893.

to ic-MRCISD+Q, SA-DSRG-PT2 consistently underestimates the excitation energies on average by 0.21 eV, while the SA-DSRG-PT3 overestimates excitation energies on average by 0.12 eV. Interestingly, the excitation energy for the $1\,^1B_u^+$ state obtained at the SA-DSRG-PT3 (5.41 eV) or ic-MRCISD+Q (5.29 eV) level is notably higher than other reported estimates, including CC3/TZVP (4.94 eV),[23] MS-CASPT2/TZVP (4.70 eV),[23] and QD-sc-NEVPT2/ANO1 (4.22 eV),[123] yet consistent with the results of state-specific strongly contracted third-order NEVPT (sc-NEVPT3)/ANO1 (5.14 eV), naïve QD-sc-NEVPT3/ANO1 (5.14 eV),[123,138] and linear response internally contracted MRCC with singles and doubles (ic-MRCCSD-LR)/TZVP (5.05 eV).[124] Note that we used rotated reference wave functions to compute the cluster corrections of ic-MRCISD,[85] and the rotated reference energy deviates from the corresponding fixed reference energy by at most 0.01 eV.[139]

In Table II, we show the vertical excitation energies computed using different SA8-CASSCF references and the cc-pVQZ basis set. As more orbitals are added to the CAS(8e,8o) space, it is possible to find SA-CASSCF solutions that correspond to local minima, especially for the enlarged active spaces CAS(8e,*n*o) with n = 10, 12, 14, 16. However, since the additional orbitals introduced are weakly occupied, these local minima are expected to be close in energy to the minimum energy solution. Indeed, for the CAS(8e,12o) space, we were able to locate a second solution with average CASSCF energy ca. 0.4 mE_h above the optimal solution. Semicanonical orbitals for all active spaces considered here computed using the cc-pVQZ basis set are reported in the supplementary material.

Comparing the CAS(8e,8o) results in Tables I and II, we see almost no changes in the excitation energies when increasing the size of the basis set. This observation also agrees with previous studies.[120,123] As expected, increasing the size of the active-space leads to better agreement between SA-DSRG-PT3 and ic-MRCISD+Q excitation energies. We also find that the ionic states benefit most from doubling the size of the active space. Focusing on the first three states of SA-DSRG-PT3, the excitation energy of the ionic state $1\,^1B_u^+$ is lowered by 0.23 eV, while the $2\,^1A_g^-$ and $1\,^1B_u^-$ states are stabilized by less than 0.05 eV. The incomplete active spaces CAS(8e,10o)–CAS(8e,12o) give an unbalanced description of the eight electronic states. For example, going from a CAS(8e,12o) to a CAS(8e,14o) reference, the additional b_g and a_u orbitals stabilize the $3\,^1A_g^-$ and $2\,^1B_u^-$ states, causing their excitation energy at the SA-DSRG-PT3 level to drop by 0.25 and 0.22 eV, respectively. The CAS(8e,14o) reference instead yields excitation

TABLE II. Vertical excitation energies (in eV) of several low-lying valence excited states of octatetraene obtained using SA-DSRG-PT2 and SA-DSRG-PT3 methods with various SA8-CASSCF references and the cc-pVQZ basis set.

	CAS(8e,8o)				CAS(8e,10o)				CAS(8e,12o)				CAS(8e,14o)				CAS(8e,16o)			
		SA-DSRG				SA-DSRG				SA-DSRG				SA-DSRG				SA-DSRG		
State	CASSCF	PT2c	PT2	PT3	CASSCF	PT2c	PT2	PT3	CASSCF	PT2c	PT2	PT3	CASSCF	PT2c	PT2	PT3	CASSCF	PT2c	PT2	PT3
$2\,^1A_g^-$	4.68	4.60	4.56	4.75	4.70	4.56	4.52	4.74	4.71	4.57	4.54	4.74	4.72	4.55	4.54	4.71	4.71	4.55	4.54	4.71
$1\,^1B_u^+$	6.65	4.67	4.83	5.40	6.37	4.80	4.82	5.30	6.19	4.87	4.85	5.24	6.13	4.89	4.86	5.19	6.00	4.93	4.89	5.17
$1\,^1B_u^-$	5.86	5.80	5.76	5.97	5.86	5.75	5.71	5.96	5.87	5.75	5.73	5.96	5.88	5.74	5.73	5.94	5.86	5.73	5.72	5.94
$3\,^1A_g^-$	6.59	6.54	6.51	6.73	6.57	6.47	6.44	6.69	6.57	6.48	6.47	6.70	6.36	6.25	6.26	6.45	6.36	6.24	6.25	6.45
$1\,^1A_g^+$	8.94	6.79	6.89	7.57	8.29	7.03	6.83	7.34	8.21	6.97	6.83	7.29	8.21	6.94	6.83	7.26	8.12	6.94	6.85	7.22
$4\,^1A_g^-$	7.49	7.48	7.43	7.70	7.51	7.41	7.36	7.68	7.51	7.41	7.38	7.68	7.53	7.39	7.38	7.66	7.52	7.37	7.36	7.65
$2\,^1B_u^-$	8.25	8.26	8.24	8.50	8.24	8.18	8.16	8.45	8.25	8.19	8.18	8.47	8.07	7.98	7.99	8.25	8.07	7.97	7.98	8.25

energies that are consistent with those from the full double-π space.

Finally, we comment on the vertical excitation energy of the $2\,^1A_g^-$, $1^1B_u^+$, and $1^1B_u^-$ states. Recent MS-CASPT2,[123] QD strongly contracted (sc) and partially contracted (pc) NEVPT2,[123] and state-specific time-dependent NEVPT2 (t-NEVPT2)[125] studies suggest the ordering $1\,^1B_u^+ < 2\,^1A_g^- < 1^1B_u^-$. The energy difference between $1^1B_u^+$ and $2^1A_g^-$ states is predicted to be around 0.7 eV by QD-sc-NEVPT2, QD-pc-NEVPT2, and t-NEVPT2 based on the CAS(8e,8o) reference.[123,125] This energy gap is perfectly reproduced by our sc-NEVPT2 and QD-sc-NEVPT2 computations reported in Table I. The same ordering is also obtained by XMS-CASPT2 and XMCQDPT2, but both methods predict that the $2\,^1A_g^-$ state lies only 0.2 eV above the $1^1B_u^+$ state. It is well-documented that NEVPT2 overestimates dynamic correlation for ionic states when the zeroth-order wave function does not capture enough dynamic σ polarization.[123,140] Indeed, the energy gap between the $2\,^1A_g^-$ and $1\,^1B_u^+$ states is reduced by more than 0.6 eV for QD-sc-NEVPT2 and QD-pc-NEVPT2 if the CAS(8e,16o) is employed.[123] By contrast, both second- and third-order SA-DSRG-PT yield the ordering of $2\,^1A_g^- < 1\,^1B_u^+ < 1\,^1B_u^-$ for all the SA-CASSCF references considered here. This ordering was also found in previous studies that used CASPT2,[120] multireference Møller-Plesset perturbation theory,[121] and recently third-order algebraic diagrammatic construction method.[141] Moreover, it is also consistent with the ordering of state-specific CASPT3, ic-MRCISD, and ic-MRCISD+Q presented in Table I. Our best estimates for the vertical excitation energies of the $2\,^1A_g^-$, $1\,^1B_u^+$, and $1\,^1B_u^-$ states are 4.7, 5.2, and 5.9 eV, respectively, at the SA-DSRG-PT3/SA8-CASSCF(8e,16o)/cc-pVQZ level of theory. These values are also in good agreement with those from ic-MRCCSD-LR/CASSCF(8e,8o)/TZVP (4.65 eV for $2\,^1A_g^-$ and 5.05 eV for $1\,^1B_u^+$)[124] although it is expected that the vertical excitation energy of the $1^1B_u^+$ state will be lowered if the basis set includes diffuse functions.

IV. CONCLUSIONS

In this work, we introduced a state-average approach to electronic excited states based on the multireference driven similarity renormalization group (MRDSRG) framework. Both MRDSRG and SA-MRDSRG are built upon the operator algebra of Mukherjee and Kutzelnigg's (MK) generalized normal ordering.[66–72] Like in our previous state-specific scheme based on a *single* multi-configurational state, we perform one unitary transformation of the Hamiltonian.[50,52,55] However, in the present work, we average dynamic correlation effects over a manifold of excited states by generalizing the MRDSRG equation to an *ensemble* of reference states. This subtle but profound difference leads to a state-universal scheme in which dynamic correlations is folded into a transformed Hamiltonian. This Hamiltonian is then diagonalized to obtain excited state energies and properties. This effectively amounts to mixing and relaxing the model space functions in the presence of effective interactions that include dynamic correlation effects. As such, the SA-MRDSRG is capable to handle quasi-degenerate states, such as avoided crossings and conical intersections.

The SA-MRDSRG scheme is closely related to the multireference equation-of-motion coupled cluster (MR-EOM-CC) method of Nooijen and co-workers,[19,73] and the density-averaged MS internally contracted MRCC ($\bar{\gamma}$-MS-icMRCC) theory of Aoto and Köhn.[142] All methods share the "transform-then-diagonalize" philosophy, where the transformed Hamiltonian is obtained by solving a set of state-averaged cluster amplitudes. Note that only the SA-MRDSRG Hamiltonian remains Hermitian after the transformation. The major difference between the SA-MRDSRG and other methods resides in the definition of the amplitude equations. The SA-MRDSRG uses many-body conditions for all amplitudes, while the $\bar{\gamma}$-MS-icMRCC equations are obtained by left-projecting the Schrödinger equation onto excited configurations. Instead, MR-EOM-CC is a hybrid method that combines a projective scheme for singles and many-body conditions for higher excitations.

We have implemented the second- and third-order perturbation theories based on the SA-MRDSRG (SA-DSRG-PT2 and SA-DSRG-PT3). Due to the state-universal character of SA-MRDSRG, the cost of the SA-DSRG-PT2 and SA-DSRG-PT3 methods does not depend on the number of targeted states. This is a major advantage of our formalism because it means that the cost to compute the energy and properties of n excited states is comparable to that of a single state-specific DSRG-PT computation. In comparison, other multi-state approaches require computing n energies and $n(n-1)/2$ couplings between states to evaluate the energy of n electronic states.

To simplify the SA-DSRG-PT equations, we adopted a zeroth-order Hamiltonian that contains only the diagonal blocks of the state-averaged Fock operator and applied the linear recursive commutator approximation[74,75] to the Hamiltonian. Consequently, only one-, two-, and three-body state-averaged density cumulants appear in the energy expressions of SA-DSRG-PT2 and SA-DSRG-PT3, evaluating which require non-iterative procedures that scale as $\mathcal{O}[(N_\mathbf{C}^2 + N_\mathbf{C}N_\mathbf{A}^2 + N_\mathbf{A}^4)N_\mathbf{V}^2]$ and $\mathcal{O}(N_\mathbf{C}^2N_\mathbf{V}^4)$, respectively (these estimates do not include the integral transformation, which scales as $N_\mathbf{C}N^4$, and the cost to diagonalize the CASCI Hamiltonian and build the density cumulants).

To demonstrate the ability of SA-MRDSRG perturbation theories to accurately treat near-degenerate electronic states, we studied the avoided crossing of lithium fluoride, the C_{2v} conical intersection of ammonia, the *cis–trans* isomerization pathways of the penta-2,4-dieniminium cation, and eight valence excited states of (E,E)-1,3,5,7-octatetraene. The LiF and NH_3 examples demonstrate that the accuracy of SA-DSRG-PT2 is similar to that of other quasi-degenerate second-order perturbation theories. Moreover, the SA-DSRG-PT3 results are found to be in very good agreement with those from ic-MRCISD. In the other two test cases, dynamic correlation effects are found critical to correctly predict the energetics of ionic states, a point already noticed in previous studies.[112,120,123] For the challenging octatetraene molecule, the SA-DSRG-PT3 computations based on the SA8-CASSCF(8e,16o)/cc-pVQZ reference predict the relative ordering of lowest two excited states

to be $2\,^1A_g^-(4.7\,\text{eV}) < 1\,^1B_u^+(5.2\,\text{eV})$, in accordance with ic-MRCISD+Q and ic-MRCCSD-LR results.

In summary, we have proposed a general and economical state-average approach for computing excited states based on the MRDSRG framework. The ensemble-based formalism presented here has been explored only in the context of valence excitation energies and several other areas of application are possible. In this work, states that enter the ensemble average were equally weighted; however, by modification of the weights is it possible to extend this treatment to other situations, like finite-temperature computations. While economical and general, the inherent averaging of dynamic correlation effects assumed by the SA-MRDSRG scheme may lead to a loss of accuracy when computing numerous electronic states. This deterioration of the accuracy should be expected when treating states with different character (e.g., ionic/diradical), larger molecules, and a significantly larger number of states. Work is in progress to create a method that combines some of the features of the state-averaged formalism (robustness, ability to treat conical intersections) with the accuracy of state-specific methods.

SUPPLEMENTARY MATERIAL

See supplementary material for the energies of LiF, NH_3, PSB3, and octatetraene, a Python function for generating the *xyz* coordinates of NH_3, and the active orbitals of octatetraene optimized using several CAS(8e,*n*o) references ($n = 8$, 10, 12, 14, 16).

ACKNOWLEDGMENTS

C.L. would like to thank Dr. Yang Guo for helpful discussions on performing QD-NEVPT2 computations using Orca. This work was supported by the U.S. Department of Energy under Award No. DE-SC0016004 and by a Research Fellowship of the Alfred P. Sloan Foundation.

[1]I. Shavitt and L. T. Redmon, J. Chem. Phys. **73**, 5711 (1980); B. Kirtman, *ibid.* **75**, 798 (1981).
[2]K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990); K. Andersson, P.-Å. Malmqvist, and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).
[3]K. Hirao, Chem. Phys. Lett. **190**, 374 (1992).
[4]U. S. Mahapatra, B. Datta, and D. Mukherjee, J. Phys. Chem. A **103**, 1822 (1999); Chem. Phys. Lett. **299**, 42 (1999); F. A. Evangelista, A. C. Simmonett, H. F. Schaefer, D. Mukherjee, and W. D. Allen, Phys. Chem. Chem. Phys. **11**, 4728 (2009); S. Chattopadhyay, R. K. Chaudhuri, U. S. Mahapatra, A. Ghosh, and S. S. Ray, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **6**, 266 (2016).
[5]C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J. P. Malrieu, J. Chem. Phys. **114**, 10252 (2001); C. Angeli, R. Cimiraglia, and J. P. Malrieu, *ibid.* **117**, 9138 (2002); C. Angeli, M. Pastore, and R. Cimiraglia, Theor. Chem. Acc. **117**, 743 (2007).
[6]Z. Rolik, Á. Szabados, and P. R. Surján, J. Chem. Phys. **119**, 1922 (2003).
[7]R. K. Chaudhuri, K. F. Freed, G. Hose, P. Piecuch, K. Kowalski, M. Włoch, S. Chattopadhyay, D. Mukherjee, Z. Rolik, Á. Szabados, G. Tóth, and P. R. Surján, J. Chem. Phys. **122**, 134105 (2005); M. R. Hoffmann, D. Datta, S. Das, D. Mukherjee, Á. Szabados, Z. Rolik, and P. R. Surján, *ibid.* **131**, 204104 (2009).
[8]P. J. Knowles and H.-J. Werner, Chem. Phys. Lett. **145**, 514 (1988); H.-J. Werner and P. J. Knowles, J. Chem. Phys. **89**, 5803 (1988).
[9]P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, and R. Shepard, Chem. Rev. **112**, 108 (2012).
[10]M. Saitow, Y. Kurashige, and T. Yanai, J. Chem. Theory Comput. **11**, 5120 (2015).
[11]B. Jeziorski and H. J. Monkhorst, Phys. Rev. A **24**, 1668 (1981).
[12]A. Banerjee and J. Simons, Int. J. Quantum Chem. **19**, 207 (1981).
[13]J. Mášik and I. Hubač, Adv. Quantum Chem. **31**, 75 (1998); J. Pittner, P. Nachtigall, P. Čársky, J. Mášik, and I. Hubač, J. Chem. Phys. **110**, 10275 (1999).
[14]U. S. Mahapatra, B. Datta, and D. Mukherjee, Mol. Phys. **94**, 157 (1998); J. Chem. Phys. **110**, 6171 (1999); S. Das, D. Mukherjee, and M. Kállay, *ibid.* **132**, 074103 (2010).
[15]F. A. Evangelista, W. D. Allen, and H. F. Schaefer, J. Chem. Phys. **125**, 154113 (2006).
[16]X. Li and J. Paldus, J. Chem. Phys. **119**, 5320 (2003).
[17]S. Li, J. Chem. Phys. **120**, 5017 (2004); T. Fang, J. Shen, and S. Li, *ibid.* **129**, 234106 (2008).
[18]F. A. Evangelista and J. Gauss, J. Chem. Phys. **134**, 114102 (2011).
[19]D. Datta, L. Kong, and M. Nooijen, J. Chem. Phys. **134**, 214116 (2011).
[20]M. Hanauer and A. Köhn, J. Chem. Phys. **134**, 204111 (2011); **137**, 131103 (2012); **136**, 204107 (2012).
[21]D. I. Lyakh, M. Musiał, V. F. Lotrich, and R. J. Bartlett, Chem. Rev. **112**, 182 (2012).
[22]A. Köhn, M. Hanauer, L. A. Mück, T.-C. Jagau, and J. Gauss, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **3**, 176 (2013).
[23]M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, and W. Thiel, J. Chem. Phys. **128**, 134110 (2008).
[24]M. R. Silva-Junior, M. Schreiber, S. P. A. Sauer, and W. Thiel, J. Chem. Phys. **133**, 174318 (2010).
[25]J.-P. Malrieu, J.-L. Heully, and A. Zaitsevskii, Theor. Chim. Acta **90**, 167 (1995).
[26]C. Bloch, Nucl. Phys. **6**, 329 (1958).
[27]J. des Cloizeaux, Nucl. Phys. **20**, 321 (1960).
[28]B. Brandow, Rev. Mod. Phys. **39**, 771 (1967).
[29]K. F. Freed, J. Chem. Phys. **60**, 1765 (1974).
[30]I. Lindgren, J. Phys. B: At., Mol. Phys. **7**, 2441 (1974); Int. J. Quantum Chem. **14**(S12), 33 (1978); I. Lindgren and J. Morrison, *Atomic Many-Body Theory* (Springer-Verlag, Berlin, Heidelberg, 1986).
[31]H. Nakano, J. Chem. Phys. **99**, 7983 (1993); H. Nakano, R. Uchiyama, and K. Hirao, J. Comput. Chem. **23**, 1166 (2002).
[32]J. Finley, P.-Å. Malmqvist, B. O. Roos, and L. Serrano-Andrés, Chem. Phys. Lett. **288**, 299 (1998).
[33]C. Angeli, S. Borini, M. Cestari, and R. Cimiraglia, J. Chem. Phys. **121**, 4043 (2004).
[34]W. Liu and M. R. Hoffmann, Theor. Chem. Acc. **133**, 1481 (2014); Y. Lei, W. Liu, and M. R. Hoffmann, Mol. Phys. **115**, 2696 (2017).
[35]S. Sharma, G. Jeanmairet, and A. Alavi, J. Chem. Phys. **144**, 034103 (2016).
[36]S. Sinha Ray, P. Ghosh, R. K. Chaudhuri, and S. Chattopadhyay, J. Chem. Phys. **146**, 064111 (2017).
[37]A. A. Granovsky, J. Chem. Phys. **134**, 214113 (2011).
[38]K. G. Dyall, J. Chem. Phys. **102**, 4909 (1995).
[39]H. J. J. van Dam, J. H. van Lenthe, and P. Pulay, Mol. Phys. **93**, 431 (1998); H. van Dam, J. H. van Lenthe, and P. Ruttink, Int. J. Quantum Chem. **72**, 549 (1999).
[40]T. Shiozaki, W. Győrffy, P. Celani, and H.-J. Werner, J. Chem. Phys. **135**, 081106 (2011).
[41]A. Zaitsevskii and J.-P. Malrieu, Chem. Phys. Lett. **233**, 597 (1995).
[42]T. Yanai, M. Saitow, X.-G. Xiong, J. Chalupsky, Y. Kurashige, S. Guo, and S. Sharma, J. Chem. Theory Comput. **13**, 4829 (2017).
[43]S. Evangelisti, J. P. Daudey, and J. P. Malrieu, Phys. Rev. A **35**, 4930 (1987).
[44]J. Paldus, P. Piecuch, L. Pylypow, and B. Jeziorski, Phys. Rev. A **47**, 2738 (1993).
[45]K. Kowalski and P. Piecuch, Phys. Rev. A **61**, 052506 (2000).
[46]B. O. Roos and K. Andersson, Chem. Phys. Lett. **245**, 215 (1995); B. O. Roos, K. Andersson, M. P. Fülscher, L. Serrano-Andrés, K. Pierloot, M. Merchán, and V. Molina, J. Mol. Struct.: THEOCHEM **388**, 257 (1996).
[47]N. Forsberg and P.-Å. Malmqvist, Chem. Phys. Lett. **274**, 196 (1997).
[48]D. Zgid, D. Ghosh, E. Neuscamman, and G. K.-L. Chan, J. Chem. Phys. **130**, 194107 (2009).
[49]Y. Kurashige and T. Yanai, J. Chem. Phys. **135**, 094104 (2011).
[50]C. Li and F. A. Evangelista, J. Chem. Theory Comput. **11**, 2097 (2015).
[51]K. P. Hannon, C. Li, and F. A. Evangelista, J. Chem. Phys. **144**, 204111 (2016).

[52]C. Li and F. A. Evangelista, J. Chem. Phys. **146**, 124132 (2017); **148**, 079902 (2018).
[53]C. Li, P. Verma, K. P. Hannon, and F. A. Evangelista, J. Chem. Phys. **147**, 074107 (2017).
[54]F. A. Evangelista, J. Chem. Phys. **141**, 054109 (2014).
[55]C. Li and F. A. Evangelista, J. Chem. Phys. **144**, 164114 (2016); **148**, 079903 (2018).
[56]S. D. Głazek and K. G. Wilson, Phys. Rev. D **48**, 5863 (1993).
[57]F. Wegner, Ann. Phys. **506**, 77 (1994).
[58]S. Kehrein, *The Flow Equation Approach to Many-Particle Systems* (Springer Berlin Heidelberg, 2006).
[59]K. Tsukiyama, S. K. Bogner, and A. Schwenk, Phys. Rev. Lett. **106**, 222502 (2011).
[60]H. Hergert, S. Binder, A. Calci, J. Langhammer, and R. Roth, Phys. Rev. Lett. **110**, 242501 (2013).
[61]H. Hergert, S. K. Bogner, T. D. Morris, A. Schwenk, and K. Tsukiyama, Phys. Rep. **621**, 165 (2016); H. Hergert, Phys. Scr. **92**, 023002 (2017).
[62]E. Gebrerufael, K. Vobig, H. Hergert, and R. Roth, Phys. Rev. Lett. **118**, 152503 (2017).
[63]W. Kutzelnigg, J. Chem. Phys. **77**, 3081 (1982); W. Kutzelnigg and S. Koch, *ibid.* **79**, 4315 (1983); W. Kutzelnigg, *ibid.* **80**, 822 (1984); **82**, 4166 (1985).
[64]L. Z. Stolarczyk and H. J. Monkhorst, Phys. Rev. A **32**, 725 (1985); **32**, 743 (1985); **37**, 1908 (1988); **37**, 1926 (1988).
[65]M. Nooijen and R. J. Bartlett, J. Chem. Phys. **104**, 2652 (1996).
[66]D. Mukherjee, Chem. Phys. Lett. **274**, 561 (1997).
[67]W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **107**, 432 (1997).
[68]U. S. Mahapatra, B. Datta, B. Bandyopadhyay, and D. Mukherjee, Adv. Quantum Chem. **30**, 163 (1998).
[69]K. R. Shamasundar, J. Chem. Phys. **131**, 174109 (2009).
[70]L. Kong, M. Nooijen, and D. Mukherjee, J. Chem. Phys. **132**, 234107 (2010).
[71]W. Kutzelnigg, K. R. Shamasundar, and D. Mukherjee, Mol. Phys. **108**, 433 (2010).
[72]D. Sinha, R. Maitra, and D. Mukherjee, Comput. Theor. Chem. **1003**, 62 (2013).
[73]D. Datta and M. Nooijen, J. Chem. Phys. **137**, 204107 (2012); O. Demel, D. Datta, and M. Nooijen, *ibid.* **138**, 134108 (2013); L. M. J. Huntington and M. Nooijen, *ibid.* **142**, 194111 (2015); L. M. J. Huntington, O. Demel, and M. Nooijen, J. Chem. Theory Comput. **12**, 114 (2016).
[74]T. Yanai and G. K.-L. Chan, J. Chem. Phys. **124**, 194106 (2006); **127**, 104107 (2007); E. Neuscamman, T. Yanai, and G. K.-L. Chan, *ibid.* **130**, 124102 (2009).
[75]F. A. Evangelista and J. Gauss, Chem. Phys. **401**, 27 (2012).
[76]R. J. Bartlett and G. D. Purvis, Int. J. Quantum Chem. **14**, 561 (1978); R. J. Bartlett, Annu. Rev. Phys. Chem. **32**, 359 (1981).
[77]B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, Chem. Phys. **48**, 157 (1980).
[78]K. Ruedenberg, M. W. Schmidt, M. M. Gilbert, and S. T. Elbert, Chem. Phys. **71**, 41 (1982).
[79]H.-J. Werner and P. J. Knowles, J. Chem. Phys. **82**, 5053 (1985).
[80]N. C. Handy, J. A. Pople, M. Head-Gordon, K. Raghavachari, and G. W. Trucks, Chem. Phys. Lett. **164**, 185 (1989).
[81]We tested size consistency and size intensivity of the excitation energies using a SA2-CASCI(2,2) reference for a monomer and a SA4-CASCI(4,4) reference for the dimer. To test size consistency, we verified that the following condition $E_{M\cdots M} = E_M + E_M$, where E_M and $E_{M\cdots M}$ indicate the energies of a monomer and two noninteracting monomers, respectively. In addition, we also confirmed the size intensivity of the excitation energy by testing the condition $E_{M\cdots M^*} = E_M + E_{M^*}$, where E_{M^*} is the energy of an excited monomer. Both tests were satisfied to within 6×10^{-10} E_h when calculations were converged to 10^{-10} E_h.
[82]H.-J. Werner, Mol. Phys. **89**, 645 (1996).
[83]P. J. Knowles and H.-J. Werner, Theor. Chim. Acta **84**, 95 (1992).
[84]S. R. Langhoff and E. R. Davidson, Int. J. Quantum Chem. **8**, 61 (1974).
[85]H.-J. Werner, M. Kállay, and J. Gauss, J. Chem. Phys. **128**, 034305 (2008).
[86]H. A. Witek, Y.-K. Choe, J. P. Finley, and K. Hirao, J. Comput. Chem. **23**, 957 (2002).
[87]Forte, a suite of quantum chemistry methods for strongly correlated electrons. For the current version, see https://github.com/evangelistalab/forte, 2017.
[88]J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, N. J. Russ, M. L. Leininger, C. L. Janssen, E. T. Seidl, W. D. Allen, H. F. Schaefer, R. A. King, E. F. Valeev, C. D. Sherrill, and T. D. Crawford, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 556 (2012); R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, J. Chem. Theory Comput. **13**, 3185 (2017).
[89]F. Neese, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 73 (2011); **8**, e1327 (2017).
[90]M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993); M. S. Gordon and M. W. Schmidt, in *Theory and Applications of Computational Chemistry*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, Amsterdam, 2005), pp. 1167–1189.
[91]H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schuetz, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 242 (2012); H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and M. Wang, MOLPRO, version 2015.1, a package of *ab initio* programs, 2015, see http://www.molpro.net.
[92]S. Chattopadhyay, R. K. Chaudhuri, and U. S. Mahapatra, J. Chem. Phys. **129**, 244108 (2008).
[93]C. W. Bauschlicher and S. R. Langhoff, J. Chem. Phys. **89**, 4246 (1988).
[94]A. J. C. Varandas, J. Chem. Phys. **131**, 124128 (2009).
[95]S. L. Li, D. G. Truhlar, M. W. Schmidt, and M. S. Gordon, J. Chem. Phys. **142**, 064106 (2015).
[96]B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning, and A. K. Wilson, Theor. Chem. Acc. **128**, 69 (2010).
[97]R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
[98]cheMVP is free, open-source software designed to make clean molecule drawings suitable for publications and presentations, written by A. Simmonett, J. M. Turney, and H. P. Shelton. For the current version, see https://github.com/CCQC/cheMVP, 2010.
[99]P. Rosmus, P. Botschwina, H.-J. Werner, V. Vaida, P. C. Engelking, and M. I. McCarthy, J. Chem. Phys. **86**, 6677 (1987).
[100]M. I. McCarthy, P. Rosmus, H.-J. Werner, P. Botschwina, and V. Vaida, J. Chem. Phys. **86**, 6693 (1987).
[101]J. Biesner, L. Schnieder, J. Schmeer, G. Ahlers, X. Xie, K. H. Welge, M. N. R. Ashfold, and R. N. Dixon, J. Chem. Phys. **88**, 3607 (1988).
[102]R. N. Dixon, Mol. Phys. **88**, 949 (1996).
[103]D. R. Yarkony, J. Chem. Phys. **121**, 628 (2004).
[104]S. Nangia and D. G. Truhlar, J. Chem. Phys. **124**, 124309 (2006).
[105]B. G. Levine, C. Ko, J. Quenneville, and T. J. Martínez, Mol. Phys. **104**, 1039 (2006).
[106]J. F. Stanton, J. Chem. Phys. **115**, 10382 (2001).
[107]A. Köhn and A. Tajti, J. Chem. Phys. **127**, 044105 (2007).
[108]E. F. Kjønstad, R. H. Myhre, T. J. Martínez, and H. Koch, J. Chem. Phys. **147**, 164105 (2017); Z. H. Li, R. Valero, and D. G. Truhlar, Theor. Chem. Acc. **118**, 9 (2007).
[109]S. L. Li, A. V. Marenich, X. Xu, and D. G. Truhlar, J. Phys. Chem. Lett. **5**, 322 (2014).
[110]A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[111]C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[112]S. Gozem, M. Huntress, I. Schapiro, R. Lindh, A. A. Granovsky, C. Angeli, and M. Olivucci, J. Chem. Theory Comput. **8**, 4069 (2012).
[113]T. Mori, K. Nakano, and S. Kato, J. Chem. Phys. **133**, 064107 (2010).
[114]O. Valsson and C. Filippi, J. Chem. Theory Comput. **6**, 1275 (2010).
[115]S. Gozem, F. Melaccio, R. Lindh, A. I. Krylov, A. A. Granovsky, C. Angeli, and M. Olivucci, J. Chem. Theory Comput. **9**, 4495 (2013); M. Huix-Rotllant, M. Filatov, S. Gozem, I. Schapiro, M. Olivucci, and N. Ferré, *ibid.* **9**, 3917 (2013); D. Tuna, D. Lefrancois, Ł. Wolański, S. Gozem, I. Schapiro, T. Andruniów, A. Dreuw, and M. Olivucci, *ibid.* **11**, 5758 (2015).
[116]J. W. Park and T. Shiozaki, J. Chem. Theory Comput. **13**, 2561 (2017).
[117]M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, and J. A. Pople, J. Chem. Phys. **77**, 3654 (1982).

[118]R. Pariser, J. Chem. Phys. **24**, 250 (1956).
[119]P. Tavan and K. Schulten, J. Chem. Phys. **85**, 6602 (1986).
[120]L. Serrano-Andrés, R. Lindh, B. O. Roos, and M. Merchán, J. Phys. Chem. **97**, 9360 (1993).
[121]K. Nakayama, H. Nakano, and K. Hirao, Int. J. Quantum Chem. **66**, 157 (1998).
[122]Y. Kurashige, H. Nakano, Y. Nakao, and K. Hirao, Chem. Phys. Lett. **400**, 425 (2004).
[123]C. Angeli and M. Pastore, J. Chem. Phys. **134**, 184302 (2011).
[124]P. K. Samanta, D. Mukherjee, M. Hanauer, and A. Köhn, J. Chem. Phys. **140**, 134108 (2014).
[125]A. Y. Sokolov, S. Guo, E. Ronca, and G. K.-L. Chan, J. Chem. Phys. **146**, 244102 (2017).
[126]J. B. Schriber and F. A. Evangelista, J. Chem. Theory Comput. **13**, 5354 (2017).
[127]K. Hirao, H. Nakano, K. Nakayama, and M. Dupuis, J. Chem. Phys. **105**, 9227 (1996).
[128]F. Weigend and R. Ahlrichs, Phys. Chem. Chem. Phys. **7**, 3297 (2005).
[129]T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).
[130]N. H. F. Beebe and J. Linderberg, Int. J. Quantum Chem. **12**, 683 (1977).
[131]H. Koch, A. Sánchez de Merás, and T. B. Pedersen, J. Chem. Phys. **118**, 9481 (2003).
[132]F. Aquilante, T. B. Pedersen, and R. Lindh, J. Chem. Phys. **126**, 194106 (2007); F. Aquilante, L. Boman, J. Boström, H. Koch, R. Lindh, A. S. de Merás, and T. B. Pedersen, in *Linear-Scaling Techniques in Computational Chemistry and Physics*, Challenges and Advances in Computational Chemistry and Physics, edited by R. Zalesny, M. G. Papadopoulos, P. G. Mezey, and J. Leszczynski (Springer Netherlands, 2011), Vol. 13, pp. 301–343.
[133]B. Peng and K. Kowalski, J. Chem. Theory Comput. **13**, 4179 (2017).
[134]W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graphics **14**, 33 (1996).
[135]J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).
[136]W. T. Borden and E. R. Davidson, Acc. Chem. Res. **29**, 67 (1996).
[137]C. Angeli, J. Comput. Chem. **30**, 1319 (2009).
[138]C. Angeli, R. Cimiraglia, and J.-P. Malrieu, Theor. Chem. Acc. **116**, 434 (2006).
[139]In addition, we note that although ic-MRCISD excitation energies were computed using the C_{2h} point group, to evaluate the transition dipole moment between ^1A_g and ^1B_u states, we had to perform ic-MRCISD computations using the C_s point group. Due to the procedure used to generate the CI basis functions implemented in MOLPRO, these two computations yield slightly different results. However, the ic-MRCISD energies computed with C_s and C_{2h} symmetry differ at most by 0.6 mE_h and excitation energies are nearly unaffected (≤ 0.01 eV). Excitation energies computed using both C_s and C_{2h} symmetry are reported in the supplementary material.
[140]I. Schapiro, K. Sivalingam, and F. Neese, J. Chem. Theory Comput. **9**, 3567 (2013).
[141]S. A. Mewes, F. Plasser, A. Krylov, and A. Dreuw, J. Chem. Theory Comput. **14**, 710 (2018).
[142]Y. A. Aoto and A. Köhn, J. Chem. Phys. **144**, 074103 (2016).