Active-space symmetry-adapted-cluster configuration-interaction and equation-of-motion coupled-cluster methods for high accuracy calculations of potential energy surfaces of radicals

Yuhki Ohtsuka[a)]
Fukui Institute for Fundamental Chemistry, Kyoto University, Kyoto 606-8103, Japan

Piotr Piecuch[b)]
Fukui Institute for Fundamental Chemistry, Kyoto University, Kyoto 606-8103, Japan and Department of Chemistry, Michigan State University, East Lansing, Michigan 48824

Jeffrey R. Gour
Department of Chemistry, Michigan State University, East Lansing, Michigan 48824

Masahiro Ehara
Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Kyoto 615-8510, Japan

Hiroshi Nakatsuji[c)]
Fukui Institute for Fundamental Chemistry, Kyoto University, Kyoto 606-8103, Japan and Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Kyoto 615-8510, Japan

(Received 16 October 2006; accepted 14 March 2007; published online 25 April 2007)

The electron-attached (EA) and ionized (IP) symmetry-adapted-cluster configuration-interaction (SAC-CI) methods and their equation-of-motion coupled-cluster (EOMCC) analogs provide an elegant framework for studying open-shell systems. As shown in this study, these schemes require the presence of higher-order excitations, such as the four-particle-three-hole ($4p$-$3h$) or four-hole–three-particle ($4h$-$3p$) terms, in the electron attaching or ionizing operator R in order to produce accurate ground- and excited-state potential energy surfaces of radicals along bond breaking coordinates. The full inclusion of the $4p$-$3h$/$4h$-$3p$ excitations in the EA/IP SAC-CI and EOMCC methods leads to schemes which are far too expensive for calculations involving larger radicals and realistic basis sets. In order to reduce the large costs of such schemes without sacrificing accuracy, the active-space EA/IP EOMCC methodology [J. R. Gour *et al.*, J. Chem. Phys. **123**, 134113 (2005)] is extended to the EA/IP SAC-CI approaches with $4p$-$3h$/$4h$-$3p$ excitations. The resulting methods, which use a physically motivated set of active orbitals to pick out the most important $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ excitations, represent practical computational approaches for high-accuracy calculations of potential energy surfaces of radicals. To illustrate the potential offered by the active-space EA/IP SAC-CI approaches with up to $4p$-$3h$/$4h$-$3p$ excitations, the results of benchmark calculations for the potential energy surfaces of the low-lying doublet states of CH and OH are presented and compared with other SAC-CI and EOMCC methods, and full CI results. [DOI: 10.1063/1.2723121]

I. INTRODUCTION

Radicals represent an important class of molecular systems due to their high reactivity and significance as reactants, products, and intermediates in elementary chemical processes. It is, therefore, essential to develop practical *ab initio* approaches that would enable one to provide a highly accurate description of ground and excited electronic states of radical and other open-shell species, particularly since it is very difficult to obtain reliable information about the relevant potential energy surfaces by experimental techniques due to the short lifetimes of radical intermediates. Unfortunately, the development of accurate and, at the same time, affordable quantum chemistry methods for the ground- and excited-state potential energy surfaces of radicals poses a challenging problem for the existing *ab initio* methodologies. This, in particular, applies to methods based on the coupled-cluster (CC) theory,[1–5] including the equation-of-motion CC (EOMCC) approaches[6–10] and their closely related symmetry-adapted-cluster (SAC)/symmetry-adapted-cluster configuration-interaction[11–21] (SAC-CI) and response or time-dependent CC (Refs. 22–25) predecessors. Indeed, the low-lying electronic states of radicals often display a manifestly multideterminantal character, which cannot be captured by the basic CC singles and doubles[26–28] (CCSD) and EOMCC singles and doubles (EOMCCSD),[7–9] SAC singles

[a)]Present address: Department of Theoretical Studies, Institute for Molecular Science, Myodaiji, Okazaki, Aichi 444-8585, Japan.
[b)]Author to whom correspondence should be addressed. Electronic mail: piecuch@chemistry.msu.edu
[c)]Electronic mail: hiroshi@sbchem.kyoto-u.ac.jp

and doubles (SAC-SD) and SAC-CI-SD-*R* (SAC-CI singles and doubles),[11–21] and linear response CCSD[29,30] approximations, exploiting the unrestricted Hartree-Fock (UHF) or restricted open-shell Hartree-Fock (ROHF) references (see, e.g., Refs. 31–35). The basic electron-attached (EA) and ionized (IP) EOMCC methods with up to two-particle-one-hole ($2p$-$1h$) and two-hole-one-particle ($2h$-$1p$) excitations, which are referred to as the EA-EOMCCSD (Refs. 36 and 37) or EA-EOMCCSD($2p$-$1h$) (Refs. 34 and 35) and IP-EOMCCSD (Refs. 38–41) or IP-EOMCCSD($2h$-$1p$) (Refs. 34 and 35) schemes, and the analogous and historically older EA and IP SAC-CI approximations truncated at the $2p$-$1h$ and $2h$-$1p$ excitations,[17–19,42–44] which belong to the family of the SAC-CI-SD-*R* approximations (cf., e.g., Ref. 44) and which are referred to here and elsewhere in this article as the SAC-CI($2p$-$1h$) and SAC-CI($2h$-$1p$) schemes, can be useful in calculations of electron affinities and ionization potentials, but they are, in most cases, insufficient for reliable calculations of excited states of radicals (see, e.g., Refs. 34–36 and 45–47). In these methods, one obtains the electronic states of radicals by applying the electron attaching or ionizing operator *R* with the $1p$ and $2p$-$1h$ or $1h$ and $2h$-$1p$ components to the ground state of the related closed-shell reference system provided by the CCSD or SAC-SD approaches, and one needs higher-than-$2p$-$1h$/$2h$-$1p$ excitations in *R* to model the electronic structure of radical species. Breaking bonds in ground and excited states of radicals is particularly difficult to describe by the basic EOMCC and SAC-CI models.[34,35]

One might attempt to address the above difficulties by turning to the sophisticated multireference CC (MRCC) methodologies of either the Fock-space/valence-universal[48,49] or Hilbert-space/state-universal[50] type (cf., e.g., Refs. 51–53 for selected reviews), but none of the existing genuine MRCC methods and computer codes are ready for routine chemical applications. In fact, the use of the true multistate MRCC methods based on complete active spaces may become prohibitively expensive when the radical systems of interest become large and when one has to rely on larger multidimensional reference spaces. Moreover, all genuine MRCC methods exploiting the generalized Bloch equation and the effective Hamiltonian formalism continue to face their own nontrivial challenges, including the need to deal with unphysical multiple[54–56] and singular[54,55,57–60] solutions, intruder states,[54,55,58] and the existence of the so-called intruder solutions that may tremendously complicate the MRCC calculations and the subsequent analysis of the resulting wave functions and energies.[54,56] This is not to say that genuine MRCC methods should not continue to be developed. On the contrary, recent years have witnessed renewed interest and significant progress in the area of valence-universal and state-universal MRCC calculations (cf., e.g., Refs. 53, 54, and 61–72). We are rather reflecting on the need for the development of relatively inexpensive, intuitive, and yet highly accurate CC or response CC, EOMCC, and SAC-CI models that would be free from the mathematical and numerical problems plaguing genuine multistate MRCC methods, that would not exceed the relatively low computer costs of the basic EOMCC and SAC-CI approaches of the CCSD type by a large factor, and that would enable one to avoid the prohibitive costs of the accurate, but very expensive conventional EOMCC and SAC-CI methods with a full treatment of triple and higher-than-triple excitations. Indeed, the complete inclusion of triple or triple and quadruple excitations in the EOMCC theory via the full EOMCC singles, doubles, and triples[73,74] (EOMCCSDT) (cf., also, Ref. 75) and EOMCC singles, doubles, triples, and quadruples[31,76] (EOMCCSDTQ) approaches exploiting the UHF or ROHF references greatly improves the results when the excited states of radicals of interest become significantly multideterminantal.[31] However, the large computer costs of the EOMCCSDT, EOMCCSDTQ, and the analogous high-order SAC-CI schemes, referred to as the SAC-CI-general-*R* methods (cf., e.g., Refs. 42–45), which are characterized by the iterative steps that scale as $n_o^3 n_u^5$ or $\mathcal{N}^8$ in the triples case and $n_o^4 n_u^6$ or $\mathcal{N}^{10}$ in the case of quadruple excitations, where n_o and n_u are the numbers of occupied and unoccupied orbitals in a basis set, respectively, and $\mathcal{N}$ is a measure of the system size, limit the applicability of the EOMCC and SAC-CI methods with a full account of triple or triple and quadruple excitations to very small molecular problems. The high cost of the open-shell EOMCCSDT calculations employing the UHF or ROHF references can be considerably reduced by exploiting the recently developed extensions of the iterative CC3 (Refs. 77–80) and noniterative CR-EOMCCSD(T) (Refs. 81–83) approaches to open shells;[32,33] but there may always be cases in which the CC3, CR-EOMCCSD(T), and similar approximate triples methods do not provide a sufficiently accurate description of the electronic states of radicals of interest. The situations examined in the present article, where one often has to go beyond triple excitations, are examples of such cases. Moreover, the use of the spin-orbital formalism in the existing open-shell implementations of the CC3 and CR-EOMCCSD(T) methods leads to problems such as spin contamination of the resulting electronic states. The same is true when one uses the open-shell EOMCCSDT,[31] EOMCCSDTQ,[31,76] and other high-order EOMCC schemes[76] employing the ROHF or UHF references.

The recently implemented EA-EOMCCSDT (Ref. 84) and IP-EOMCCSDT (Refs. 85 and 86) methods, which incorporate up to $3p$-$2h$ and $3h$-$2p$ excitations in the linear excitation operator *R*, and the analogous and historically older EA and IP SAC-CI approaches with up to $3p$-$2h$ and $3h$-$2p$ excitations in the *R* operator,[42–44] which belong to the SAC-CI-general-*R*(SDT) family[44] and which are referred to here and elsewhere in this article, for consistency reasons, as the SAC-CI($3p$-$2h$) and SAC-CI($3h$-$2p$) approaches, lead to the orthogonally spin-adapted wave functions of radicals and should be accurate enough in at least some cases of radical excited states, but there are two problems with these kinds of methods. The first problem is the computer cost. For example, the EA-EOMCCSDT and IP-EOMCCSDT methods, as implemented in Refs. 84–86, are characterized by the expensive iterative $n_o^2 n_u^5$ and $n_o^3 n_u^4$ ($\mathcal{N}^7$-like) steps in the diagonalization of the similarity-transformed Hamiltonian, combined with even more expensive $n_o^3 n_u^5$ ($\mathcal{N}^8$-like) steps of the underlying ground-state CCSDT calculations. Iterative steps of these types (particularly, the $n_o^2 n_u^5$ steps of EA-

EOMCCSDT and the $n_o^3n_u^5$ steps of the underlying CCSDT) are far too expensive for routine chemical applications. The original SAC-CI-general-*R*(SDT) [i.e., SAC-CI(3*p*-2*h*) and SAC-CI(3*h*-2*p*)] approaches[42–44] and the recently developed IP-EOMCCSDT-3,[87] IP-CC3,[87] IP-EOMCCSD(3*h*-2*p*),[34,35,88] and EA-EOMCCSD(3*p*-2*h*)[34,35,88] methods eliminate the most expensive $n_o^3n_u^5$ steps of CCSDT from the calculations, but the iterative $n_o^3n_u^4$ steps defining the diagonalization of the Hamiltonian in the IP triples case and the iterative $n_o^2n_u^5$ steps characterizing the diagonalization of the Hamiltonian in the EA triples case remain.

The second, more fundamental, problem, which is particularly relevant to the studies reported in this paper, is that even the full account of 3*p*-2*h* and 3*h*-2*p* excitations in the EA and IP EOMCC and SAC-CI calculations is, in many cases, insufficient to provide the correct description of ground- and excited-state potential energy surfaces of radical species along the relevant bond breaking coordinates. This has been pointed out in Ref. 35 and the present paper provides a clear demonstration that one has to consider at least the 4*p*-3*h* and 4*h*-3*p* terms in the electron attaching or ionizing operator *R* defining the SAC-CI and EOMCC wave function Ansätze to obtain the high-quality potential energy surfaces of radicals, particularly when the excited states of interest gain significant 3*p*-2*h* and 3*h*-2*p* components, as is often the case when bonds are stretched or broken. As shown in this paper, the EA-EOMCCSD(3*p*-2*h*) method[34,35,88] and its SAC-CI(3*p*-2*h*) counterpart,[42–44] which use up to 3*p*-2*h* excitations in the electron attaching operator *R*, are accurate enough for the electronic states of radicals dominated by the 1*p* and 2*p*-1*h* components. Similarly, the IP-EOMCCSD(3*h*-2*p*) (Refs. 34, 35, and 88) and SAC-CI(3*h*-2*p*) (Refs. 42–44) approaches provide a reasonably accurate description of the electronic states dominated by the 1*h* and 2*h*-1*p* components. Unfortunately, the ground- and excited-state wave functions of radicals gain significant 3*p*-2*h* and 3*h*-2*p* components relative to the corresponding closed-shell reference molecule at larger internuclear separations, which, as shown in this work, cause major problems to the EA-EOMCCSD(3*p*-2*h*)/SAC-CI(3*p*-2*h*) and IP-EOMCCSD(3*h*-2*p*)/SAC-CI(3*h*-2*p*) methods.

The fact that one needs the higher-order EA and IP SAC-CI or EOMCC schemes with at least the 4*p*-3*h* or 4*h*-3*p* components in the electron attaching or ionizing operator *R* to study ground- and excited-state potential energy surfaces of radicals is actually not a surprise. A radical can be viewed as a system obtained by attaching an electron to or removing an electron from the related closed-shell molecule. Assuming for a moment that a reference closed-shell molecule of interest is a singly bonded species, one needs at least triple (i.e., 3*p*-3*h*) excitations to describe bond breaking in it in a quantitative manner (cf., e.g., Refs. 81, 83, and 89–92). The superposition of the 3*p*-3*h* excitations in the closed-shell reference molecule with the 1*p* or 1*h* excitations that one needs to apply in order to form a radical from it results in the 4*p*-3*h* or 4*h*-3*p* excitations that should, therefore, be present in the EA or IP SAC-CI and EOMCC calculations to obtain a quantitative description of the ground- and excited-state potential energy surfaces of radical species along the relevant bond breaking coordinates. A similar argument can also be used to explain why it is often enough to use the EA-EOMCCSD(3*p*-2*h*)/SAC-CI(3*p*-2*h*) and IP-EOMCCSD(3*h*-2*p*)/SAC-CI(3*h*-2*p*) methods, which ignore the 4*p*-3*h* or 4*h*-3*p* components, to describe ground and excited states of radicals near the equilibrium geometry (i.e., in the Franck-Condon region).[34,35,88] At shorter internuclear distances, double (i.e., 2*p*-2*h*) excitations are often good enough to describe the closed-shell reference molecule in a reasonable manner. The superposition of the 2*p*-2*h* excitations in the closed-shell reference molecule with the 1*p* or 1*h* excitations that one needs to apply in order to form a radical species from it results in the need for 3*p*-2*h* or 3*h*-2*p* excitations in the EA or IP SAC-CI and EOMCC calculations, but higher-order 4*p*-3*h* or 4*h*-3*p* excitations can often be neglected.

The above discussion, supported by the ample numerical evidence provided in this paper, implies that it is essential to consider the 4*p*-3*h* and 4*h*-3*p* excitations in the electron attaching and ionizing operators defining the EA and IP SAC-CI and analogous EOMCC schemes in order to obtain the accurate ground- and excited-state potential energy surfaces of radicals. The problem is that the excitation amplitudes defining the 4*p*-3*h* and 4*h*-3*p* operators, whose numbers scale as $n_o^3n_u^4$ and $n_o^4n_u^3$, respectively, are far too numerous for routine calculations for larger radicals and larger basis sets. Moreover, the full treatment of the 4*p*-3*h* and 4*h*-3*p* excitations in the electron attaching and ionizing operators, as in the recently implemented EA- and IP-EOMCCSDTQ approaches,[47] leads to schemes which use the prohibitively expensive CPU steps that scale as $n_o^3n_u^6$ and $n_o^4n_u^5$ or $\mathcal{N}^9$ in the Hamiltonian diagonalization part, and $n_o^3n_u^5$ ($\mathcal{N}^8$) or $n_o^4n_u^6$ ($\mathcal{N}^{10}$) in the part that deals with the underlying ground-state CCSDT or CCSDTQ calculations for the reference closed-shell system (although this is not absolutely necessary,[10] particularly when radicals are examined,[34,35] the authors of Ref. 47 chose the more expensive CCSDTQ option in the underlying CC calculations for the closed-shell reference system that precede the EA- and IP-EOMCCSDTQ calculations). The complete treatment of all many-body components of the electron attaching and ionizing operators *R* that define the SAC-CI-general-*R* methods with up to 4*p*-3*h* and 4*h*-3*p* excitations[42,44] [the EA and IP SAC-CI-general-*R* (SDTQ) schemes in Ref. 44 and the SAC-CI(4*p*-3*h*) and SAC-CI(4*h*-3*p*) approaches in this work] leads to similar computer costs. Clearly, methods of this kind can only be used in benchmark or small-molecule calculations. In fact, as mentioned earlier, a full treatment of the less demanding 3*p*-2*h* and 3*h*-2*p* components of the electron attaching and ionizing operators *R* of the EA and IP SAC-CI and EOMCC methods leads to schemes that are already very expensive. It is important to develop the EA and IP SAC-CI and EOMCC methods that can reduce the large costs of handling the 3*p*-2*h*/3*h*-2*p* and 4*p*-3*h*/4*h*-3*p* terms without a substantial loss of accuracy.

The need for an approximate treatment of the 3*p*-2*h*/3*h*-2*p* and 4*p*-3*h*/4*h*-3*p* excitations has already been pointed out in the original papers on the EA and IP SAC-CI-general-*R* methods[42–44] that precede the initial papers on the

EA and IP EOMCC approaches,[36–41] which ignore these terms altogether. This has prompted the development of the highly efficient EA and IP SAC-CI methods with the perturbative selection (PS) of the $3p$-$2h$/$3h$-$2p$, $4p$-$3h$/$4h$-$3p$, and other higher-order excitations, which also neglect or approximate higher-order nonlinear terms in the cluster operator defining the ground electronic state of the reference closed-shell molecule[42,44] (see Ref. 93 for the original idea of the SAC-CI approaches with the PS, abbreviated here and elsewhere in this article as the SAC-CI/PS schemes). There are several advantages of the SAC-CI/PS methods, the most important one being the fact that in these approaches we do not have to limit ourselves to a particular excitation order and, at least in principle, can select the $3p$-$2h$/$3h$-$2p$, $4p$-$3h$/$4h$-$3p$, and other appropriate higher-order excitations purely numerically, based on their significance in the perturbation theory analysis rather than on their excitation rank. As a result, the SAC-CI/PS approaches can be successfully applied to a wide range of molecular problems, including large systems such as biological molecules, without losing much accuracy. There also are disadvantages. The perturbative selection of the $3p$-$2h$/$3h$-$2p$, $4p$-$3h$/$4h$-$3p$, and other higher-order excitations may lead to substantial numerical noise if we try to apply the SAC-CI/PS methods to potential energy surfaces, since different sets of higher-order excitations will be selected at different geometries. This problem can be alleviated by merging the $3p$-$2h$/$3h$-$2p$, $4p$-$3h$/$4h$-$3p$, and other higher-order excitations selected in different regions of potential energy surfaces. In this regard, the Group SUM method,[94] which gives a common set of the suitably selected excitation operators, is particularly useful, producing smooth potential energy surfaces within the SAC-CI/PS framework. However, there may be cases where it is not immediately obvious which regions of potential energy surfaces should initially be probed to obtain the adequate selection of the $3p$-$2h$/$3h$-$2p$, $4p$-$3h$/$4h$-$3p$, and other higher-order excitations that would work in all regions. In fact, the same remarks apply to other methods based on numerical selections of higher-order excitations, including, for example, the popular multireference CI approach with a selection of single and double excitations from a multideterminantal reference wave function of Refs. 95–97. It is also quite difficult to combine the PS scheme with the EA, IP, and other EOMCC methods, which traditionally rely on the many-body diagrammatic formulation and factorization of nonlinear terms that enter the EOMCC wave functions and the similarity-transformed Hamiltonian of the underlying CC theory.

The above discussion implies that it is important to consider alternative formulations of the EA and IP SAC-CI and EOMCC methods with $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ excitations, in which selection of the dominant $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ terms is done mathematically rather than via numerical thresholds. An idea, which can help us to design the relatively inexpensive EA and IP SAC-CI and EOMCC methods that could provide highly accurate results for ground- and excited-state potential energy surfaces of radicals, is that of the active-space CC or single-reference-like, state-selective MRCC approaches pioneered by Adamowicz, Piecuch, and co-workers[89–92,98–106] and active-space EOMCC theories introduced and fully developed by Kowalski and Piecuch.[73,74,107] In these methods and their various subsequent implementations, by Piecuch and co-workers and Adamowicz and co-workers (cf., e.g., Refs. 108–112) and others (cf., e.g., Refs. 76 and 113–118), one combines the single-reference CC/EOMCC formalism with a multireference concept of active orbitals, which are used to select a relatively small subset of the dominant triply and other higher-than-doubly excited clusters that reflect the nature of the electronic quasidegeneracy or excited states of interest, reducing the computer costs of the parent CC/EOMCC calculations with a full treatment of higher-than-double excitations by a large factor. In particular, Gour *et al.* have recently proposed the active-space EA and IP EOMCC methods in which one uses small sets of active orbitals to select the most important $3p$-$2h$/$3h$-$2p$, $4p$-$3h$/$4h$-$3p$, and other higher-order contributions to the electron attaching and ionizing operators R that define the EA and IP EOMCC theories and their multiply attached and ionized variants.[34,35,88] Just like the parent EA and IP EOMCC approaches with the full treatment of the $3p$-$2h$/$3h$-$2p$, $4p$-$3h$/$4h$-$3p$, and other higher components of R, the active-space EA and IP EOMCC methods and their multiply attached and ionized analogs discussed in Ref. 34 provide an orthogonally spin-adapted description of ground and excited states of radical and other open-shell species. Among the active-space EA and IP EOMCC approaches discussed in Refs. 34, 35, and 88 are the EA-EOMCCSDt and IP-EOMCCSDt approximations, in which the dominant $3p$-$2h$ and $3h$-$2p$ components of the electron attaching and ionizing operators R defining the aforementioned EA-EOMCCSD($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$) schemes are selected via active orbitals. The EA-EOMCCSDt/IP-EOMCCSDt methods based on diagonalizing the similarity-transformed Hamiltonian of CCSD in the space of all $1p/1h$ and $2p-1h/2h-1p$ excitations and a small subset of $3p$-$2h$/$3h$-$2p$ excitations defined through active orbitals have been tested in Refs. 34, 35, and 88, demonstrating a considerable promise in applications to excitation energies of radicals calculated in the vicinity of the equilibrium geometry. As shown in Refs. 34, 35, and 88, the EA-EOMCCSDt/IP-EOMCCSDt approaches give excellent vertical and adiabatic excitation energies of radical species at a small fraction of the cost associated with the parent EA-EOMCCSD($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$) calculations, which use all $3p$-$2h$ and $3h$-$2p$ excitations. However, as indicated in Ref. 35, the EA-EOMCCSDt/IP-EOMCCSDt methods are not sufficiently accurate to describe potential energy surfaces along bond breaking coordinates, since ground and excited states of radicals often gain significant $3p$-$2h$ and $3h$-$2p$ components when larger internuclear separations are examined and these, as already pointed out above, cannot be well described by the EA-EOMCCSD($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$) levels of theory. This prompts the need for the implementation of the active-space EA-EOMCCSDtq and IP-EOMCCSDtq methods, as defined in Ref. 34, and their SAC-CI counterparts, in which one uses active orbitals to select the most important $4p$-$3h$ and $4h$-$3p$ excitations, in addition to the $3p$-$2h$ and $3h$-$2p$ excitations present in the EA-EOMCCSDt and IP-EOMCCSDt methods.

The basic idea of EA-EOMCCSDtq, IP-EOMCCSDtq, and the analogous active-space SAC-CI methods developed and tested in this work is to reduce the enormous computer costs of the EA and IP EOMCC and SAC-CI calculations with the full treatment of $4p$-$3h$ and $4h$-$3p$ excitations by systematically selecting the dominant $4p$-$3h$ and $4h$-$3p$ components and their $3p$-$2h$ and $3h$-$2p$ counterparts via a small subset of active orbitals.

In this paper, we focus on the development, implementation, and testing of the active-space SAC-CI methods with the $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ terms. SAC-CI offers an advantage over the current implementations of the EOMCC theory in that it is easier to impose the active-space logic on the $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ components within the language of electron configurations exploited in SAC-CI, as compared with the diagrammatic many-body language used in EOMCC. We will eventually pursue the more complete active-space EA-EOMCCSDtq and IP-EOMCCSDtq methods suggested in Ref. 34, which use the CCSDt (Refs. 73, 74, 89–92, and 98–107) or CCSDtq (Refs. 90, 99–102, and 104) wave functions to describe the ground state of the reference closed-shell system and which do not neglect any nonlinear terms in the cluster operator, but it is important to investigate first what are the potential benefits of incorporating the high-order $4p$-$3h$ and $4h$-$3p$ terms via active orbitals and what types of improvements in the full and active-space EA-EOMCCSD($3p$-$2h$)/SAC-CI($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$)/SAC-CI($3h$-$2p$) results can be obtained when the ground- and excited-state potential energy surfaces of radicals along bond breaking coordinates are examined with the active-space methods truncated at the $4p$-$3h$ and $4h$-$3p$ excitations. The SAC-CI methodology is very useful in this regard.

The primary objectives of the present work are the following: (i) to show that one needs the genuine $4p$-$3h$ and $4h$-$3p$ excitations in the electron attaching and ionizing operators R in order to obtain an accurate representation of the ground- and excited-state potential energy surfaces of radicals along the relevant bond breaking coordinates, demonstrating that the $3p$-$2h$ and $3h$-$2p$ components (and, of course, the lower-order $1p$ and $2p$-$1h$ or $2h$ and $2h$-$1p$ terms) are generally insufficient; (ii) to demonstrate that one can obtain accurate ground- and excited-state potential energy surfaces, which can compete with the results of SAC-CI calculations with a full treatment of $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ components and be close to the exact, full CI potentials, using small subsets of active orbitals to select the dominant $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ excitations; and (iii) to show that the active-space SAC-CI methods with the $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ excitations selected mathematically by redefining the relevant R operators are at least as effective as the SAC-CI/PS approaches, in which the $4p$-$3h$/$4h$-$3p$ components are selected numerically based on perturbative arguments and thresholds for neglecting small contributions. We also demonstrate that it is not sufficient to use the CI methods with up to $4p$-$3h$ and $4h$-$3p$ components to obtain accurate ground- and excited-state potential energy surfaces of the CH and OH radicals. One needs to apply the electron attaching and ionizing operators R to the correlated CC or SAC ground state to obtain high-quality potential energy surfaces of radical species along the relevant bond breaking coordinates. The relevant test calculations are reported for the ground- and excited-state potential energy curves representing a few low-lying doublet states of the CH (the EA case) and OH (the IP case) radicals, which are small enough to allow for the exact, full CI, calculations. The results reported in this article are encouraging enough to pursue the implementation of the active-space EA-EOMCCSDtq and IP-EOMCCSDtq methods of Ref. 34 in the future, since one can regard the EA-EOMCCSDtq and IP-EOMCCSDtq approaches as nothing else than the EA and IP EOMCC analogs of the active-space SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods considered here. This statement reflects on our belief that the historically older SAC-CI methodology and the more recent EOMCC formalism, being so closely related to each other, may both benefit from being developed side by side. The present study may serve as an illustration of how much one can learn by working with both methodologies at the same time. Since there have been no prior studies that use the SAC-CI and EOMCC methods together, we use this paper as an opportunity to emphasize that up to rather unimportant details, which can be dealt with if necessary, the EA and IP SAC-CI methods and their EOMCC counterparts represent essentially the same methodology.

II. THEORY

The active-space EA and IP SAC-CI and EOMCC methods are based on the idea of selecting the most important $3p$-$2h$/$3h$-$2p$, $4p$-$3h$/$4h$-$3p$, and other higher-order excitations in the corresponding full schemes. Thus, we begin this section with a review of the basic elements of the EA and IP SAC-CI and EOMCC methodologies. This enables us to make an important point that there are many similarities between the EA and IP SAC-CI methods and their EOMCC analogs, which have not always been appreciated in the literature and which may benefit the development of both methodologies, as is the case in this work.

A. An overview of the EA and IP SAC-CI and EOMCC methods

The EA and IP SAC-CI and EOMCC methods use the following wave function *Ansatz* to represent the electronic states of the $(N+1)$- or $(N-1)$-electron system:

$$|\Psi_\mu^{(N\pm1)}\rangle = R_\mu^{(N\pm1)}|\Psi_0\rangle, \tag{1}$$

where

$$|\Psi_0\rangle = e^S|\Phi\rangle \tag{2}$$

is the SAC (in the SAC-CI case) or CC (in the EOMCC case) ground state of an N-electron closed-shell system, and $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$ are the electron attaching and ionizing operators that generate the $(N+1)$- or $(N-1)$-electron states from the N-electron SAC/CC wave function $|\Psi_0\rangle$. Here, $|\Phi\rangle$ is a closed-shell, N-electron reference determinant [in the applications presented in this work, the restricted Hartree-Fock (RHF) configuration] and S is the cluster operator,

$$S = \sum_{n=1}^{M_S} S_n, \quad S_n = \left(\frac{1}{n!}\right)^2 s_{a_1\cdots a_n}^{i_1\cdots i_n} a^{a_1}\cdots a^{a_n} a_{i_n}\cdots a_{i_1}, \tag{3}$$

with $s_{a_1\cdots a_n}^{i_1\cdots i_n}$ representing the corresponding cluster amplitudes. We use the usual notation where $i,j,\ldots$ or $i_1,i_2,\ldots$ $(a,b,\ldots$ or $a_1,a_2,\ldots)$ are the spin-orbitals occupied (unoccupied) in the reference determinant $|\Phi\rangle$ and a^p (a_p) are the creation (annihilation) operators associated with the spin-orbital basis set $\{|p\rangle\}$. Whenever appropriate, we also use the Einstein summation convention over the repeated upper and lower indices. In general, the excitation level M_S in Eq. (3) satisfies $M_S \leqslant N$. In the basic SAC-SD and CCSD calculations, which provide the ground-state wave function of the reference N-electron closed-shell system for all of the EA and IP SAC-CI and EOMCC calculations reported in this work, $M_S=2$.

The electron attaching and ionizing operators, $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$, respectively, entering Eq. (1), are defined as

$$R_\mu^{(N+1)} = \sum_{n=0}^{M_R} R_{\mu,(n+1)p-nh} \tag{4}$$

and

$$R_\mu^{(N-1)} = \sum_{n=0}^{M_R} R_{\mu,(n+1)h-np}, \tag{5}$$

where

$$R_{\mu,(n+1)p-nh} = \frac{1}{n!(n+1)!} r_{aa_1\cdots a_n}^{i_1\cdots i_n} a^a a^{a_1}\cdots a^{a_n} a_{i_n}\cdots a_{i_1} \tag{6}$$

and

$$R_{\mu,(n+1)h-np} = \frac{1}{n!(n+1)!} r_{a_1\cdots a_n}^{ii_1\cdots i_n} a^{a_1}\cdots a^{a_n} a_{i_n}\cdots a_{i_1} a_i, \tag{7}$$

with $M_R=N$ in the exact theory and $M_R<N$ in the approximate approaches. In this paper, we use the following EA and IP SAC-CI and EOMCC truncation schemes:

(i) EA-EOMCCSD($2p$-$1h$) and IP-EOMCCSD($2h$-$1p$). These are the basic EA and IP EOMCC approaches, in which $M_S=2$ and $M_R=1$. The EA-EOMCCSD($2p$-$1h$) and IP-EOMCCSD($2h$-$1p$) methods are also known in the literature as the EA-EOMCCSD and IP-EOMCCSD approaches.[36–41] They use the CCSD approximation for S and the $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$ operators are truncated at the $2p$-$1h$ and $2h$-$1p$ components, so that

$$R_\mu^{(N+1)} = R_{\mu,1p} + R_{\mu,2p\text{-}1h} = r_a a^a + \tfrac{1}{2} r_{ab}^j a^a a^b a_j \tag{8}$$

and

$$R_\mu^{(N-1)} = R_{\mu,1h} + R_{\mu,2h\text{-}1p} = r^i a_i + \tfrac{1}{2} r_b^{ij} a^b a_j a_i, \tag{9}$$

respectively. The EA-EOMCCSD($2p$-$1h$) and IP-EOMCCSD($2h$-$1p$) methods are the EOMCC analogs of the SAC-CI($2p$-$1h$) and SAC-CI($2h$-$1p$) schemes, also known as the EA and IP SAC-CI-SD-R methods.[44] We do not report the SAC-CI($2p$-$1h$) and SAC-CI($2h$-$1p$) numerical results in this work, since they are very similar to those obtained with the EA-EOMCCSD($2p$-$1h$) and IP-EOMCCSD($2h$-$1p$) approaches. Moreover, neither SAC-CI nor EOMCC schemes truncated at the $2p$-$1h$ and $2h$-$1p$ excitations provide accurate potential energy surfaces of radical species. The EA-EOMCCSD($2p$-$1h$) and IP-EOMCCSD($2h$-$1p$) results are sufficient to illustrate this point.

(ii) EA-EOMCCSD($3p$-$2h$), IP-EOMCCSD($3h$-$2p$), SAC-CI($3p$-$2h$), and SAC-CI($3h$-$2p$). In these schemes, $M_S=2$ and $M_R=2$. In other words, we continue to use the CCSD (the EOMCC case) or SAC-SD (the SAC-CI case) ground state of an N-electron closed-shell system as a correlated reference state $|\Psi_0\rangle$, while truncating the $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$ operators at the $3p$-$2h$ and $3h$-$2p$ components to obtain

$$R_\mu^{(N+1)} = R_{\mu,1p} + R_{\mu,2p\text{-}1h} + R_{\mu,3p\text{-}2h} \tag{10}$$

and

$$R_\mu^{(N-1)} = R_{\mu,1h} + R_{\mu,2h\text{-}1p} + R_{\mu,3h\text{-}2p}, \tag{11}$$

respectively, where the $R_{\mu,1p}$, $R_{\mu,2p\text{-}1h}$, $R_{\mu,1h}$, and $R_{\mu,2h\text{-}1p}$ components are defined in the same manner as in Eqs. (8) and (9) and $R_{\mu,3p\text{-}2h}$ and $R_{\mu,3h\text{-}2p}$ are given by

$$R_{\mu,3p-2h} = \tfrac{1}{12} r_{abc}^{jk} a^a a^b a^c a_k a_j \tag{12}$$

and

$$R_{\mu,3h-2p} = \tfrac{1}{12} r_{bc}^{ijk} a^b a^c a_k a_j a_i, \tag{13}$$

respectively. We use the EA-EOMCCSD($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$) methods and the analogous SAC-CI($3p$-$2h$) and SAC-CI($3h$-$2p$) schemes to show that the more complete treatment of the higher-order nonlinear terms in S in the EA-EOMCCSD($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$) approaches is not sufficient to eliminate large errors in the SAC-CI($3p$-$2h$) and SAC-CI($3h$-$2p$) results at larger internuclear separations. One needs the genuine $4p$-$3h$ and $4h$-$3p$ excitations in $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$, respectively, to improve the EA-EOMCCSD($3p$-$2h$)/SAC-CI($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$)/SAC-CI($3h$-$2p$) results at significantly stretched geometries. The active-space variants of the EA-EOMCCSD($3p$-$2h$)/SAC-CI($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$)/SAC-CI($3h$-$2p$) methods are described in Sec. II B.

(iii) SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$). These are the key truncation schemes for the considerations reported in this paper. In these approximations, $M_S=2$ and $M_R=3$. Thus, we use the SAC-SD ground state of an N-electron closed-shell system as the correlated reference state $|\Psi_0\rangle$, while truncating the $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$ operators at the $4p$-$3h$ and $4h$-$3p$ components to obtain

$$R_\mu^{(N+1)} = R_{\mu,1p} + R_{\mu,2p\text{-}1h} + R_{\mu,3p\text{-}2h} + R_{\mu,4p\text{-}3h} \tag{14}$$

and

$$R_\mu^{(N-1)} = R_{\mu,1h} + R_{\mu,2h\text{-}1p} + R_{\mu,3h\text{-}2p} + R_{\mu,4h\text{-}3p}, \tag{15}$$

respectively, where the $R_{\mu,1p}$, $R_{\mu,2p\text{-}1h}$, $R_{\mu,1h}$, $R_{\mu,2h\text{-}1p}$, $R_{\mu,3p\text{-}2h}$, and $R_{\mu,3h\text{-}2p}$ components are the same as in Eqs. (8)–(11), and $R_{\mu,4p\text{-}3h}$ and $R_{\mu,4h\text{-}3p}$ are defined as

$$R_{\mu,4p\text{-}3h} = \frac{1}{144} r^{jkl}_{abcd} a^a a^b a^c a^d a_l a_k a_j \tag{16}$$

and

$$R_{\mu,4h\text{-}3p} = \frac{1}{144} r^{ijkl}_{bcd} a^b a^c a^d a_l a_k a_j a_i, \tag{17}$$

respectively. One could define the EA-EOMCCSD($4p$-$3h$) and IP-EOMCCSD($4h$-$3p$) truncation schemes in a similar manner. Although the EA-EOMCCSD($4p$-$3h$) and IP-EOMCCSD($4h$-$3p$) approaches would offer a complete treatment of the nonlinear terms in S_1 and S_2 resulting from the use of the CC Ansatz for the N-electron ground state $|\Psi_0\rangle$ [Eq. (2)], we focus on the corresponding and simpler SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods and their relatively inexpensive active-space variants, which are sufficient to eliminate the large errors in describing the ground- and excited-state potential energy curves of the CH and OH radicals at larger internuclear separations produced by the EA-EOMCCSD($3p$-$2h$)/SAC-CI($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$)/SAC-CI($3h$-$2p$) schemes. The active-space SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) approaches are described in Sec. II B.

In addition to the above EA and IP EOMCC and SAC-CI methods, we use the CI($4p$-$3h$) and CI($4h$-$3p$) approaches, in which one obtains the electronic wave functions $|\Psi^{(N\pm1)}_\mu\rangle$ by applying the electron attaching and ionizing operators $R^{(N+1)}_\mu$ and $R^{(N-1)}_\mu$ with up to $4p$-$3h$ and $4h$-$3p$ excitations, defined by Eqs. (14) and (15), respectively, directly to the N-electron reference determinant $|\Phi\rangle$ rather than to the correlated CC or SAC ground state $|\Psi_0\rangle$ [Eq. (2)]. The purpose of comparing the CI($4p$-$3h$)/CI($4h$-$3p$) results with the corresponding results of the SAC-CI($4p$-$3h$)/SAC-CI($4h$-$3p$) calculations is to show that the mere inclusion of the $4p$-$3h$ and $4h$-$3p$ excitations in $R^{(N+1)}_\mu$ and $R^{(N-1)}_\mu$, respectively, is not sufficient to obtain the accurate potential energy curves of the CH and OH radicals. One has to apply the $R^{(N+1)}_\mu$ and $R^{(N-1)}_\mu$ operators with a full or active-space treatment of the $4p$-$3h$ and $4h$-$3p$ components to the correlated ground state of the N-electron reference system.

Before discussing the active-space EA and IP SAC-CI and EOMCC methods developed and/or tested in this paper, let us emphasize that although there are technical differences between the SAC-CI and EOMCC approaches, which are related to different algorithms that are used to determine the S and R operators and the additional approximations in the SAC-CI models that are not exploited in the standard implementations of the EOMCC theories, the similarities between the SAC-CI and EOMCC methodologies are so great that the main conclusions drawn from the SAC-CI calculations apply to the EOMCC calculations and vice versa. Indeed, in the EA and IP SAC-CI approaches which, historically speaking, precede the analogous EA and IP EOMCC methods by a decade or so, one determines the cluster operator S by solving the system of equations

$$\langle\Phi|(H - E^{(N)}_0)e^S|\Phi\rangle = 0, \tag{18}$$

$$\langle\Phi^{a_1\cdots a_n}_{i_1\cdots i_n}|(H - E^{(N)}_0)e^S|\Phi\rangle = 0,$$

$$i_1 < \cdots < i_n, \quad a_1 < \cdots < a_n, \quad n = 1, \ldots, M_S, \tag{19}$$

where H is the electronic Hamiltonian, $E^{(N)}_0$ is the ground-state energy of the N-electron closed-shell system, and $|\Phi^{a_1\cdots a_n}_{i_1\cdots i_n}\rangle \equiv a^{a_1}\cdots a^{a_n}a_{i_n}\cdots a_{i_1}|\Phi\rangle$ are the n-tuply excited determinants relative to $|\Phi\rangle$ [in practical implementations of SAC and SAC-CI, the spin- and symmetry-adapted configuration state functions (CSFs) corresponding to determinants $|\Phi^{a_1\cdots a_n}_{i_1\cdots i_n}\rangle$]. Equations (18) and (19) are obtained by left-projecting the electronic Schrödinger equation for the CC/SAC wave function $|\Psi_0\rangle$ [Eq. (2)] on the reference determinant $|\Phi\rangle$ and the excited determinants $|\Phi^{a_1\cdots a_n}_{i_1\cdots i_n}\rangle$ corresponding to the excitations included in S. If we do not neglect any nonlinear terms in S, the SAC equations [Eqs. (18) and (19)] become equivalent to the standard CC equations of Čížek,

$$E^{(N)}_0 = \langle\Phi|\bar{H}|\Phi\rangle, \tag{20}$$

$$\langle\Phi^{a_1\cdots a_n}_{i_1\cdots i_n}|\bar{H}|\Phi\rangle = 0,$$

$$i_1 < \cdots < i_n, \quad a_1 < \cdots < a_n, \quad n = 1, \ldots, M_S, \tag{21}$$

where

$$\bar{H} = e^{-S}He^S = (He^S)_C \tag{22}$$

is the similarity-transformed Hamiltonian of the CC theory and subscript C refers to the connected part of a given operator expression. This is because the energy-dependent terms in Eq. (19) cancel out the disconnected terms contributing to the product of H and e^S, leaving us with the energy-independent system of nonlinear CC equations defined by Eq. (21). In particular, the full SAC-SD approximation, in which all nonlinear terms in S_1 and S_2 are retained, is equivalent to the standard CCSD approximation. Although full SAC-SD calculations are, in principle, possible using the routines that form part of GAUSSIAN 03,[119] the present computational algorithm used to determine all nonlinear terms of the full SAC-SD approach is not particularly efficient; so in the conventional SAC-SD calculations, which are used in this work to generate the S_1 and S_2 clusters for the subsequent SAC-CI($2p$-$1h$), SAC-CI($2h$-$1p$), SAC-CI($3p$-$2h$), SAC-CI($3h$-$2p$), SAC-CI($4p$-$3h$), and SAC-CI($4h$-$3p$) calculations, all nonlinear terms in Eqs. (18) and (19) other than $\frac{1}{2}S^2_2$ are neglected. Thus, the SAC-SD and CCSD results differ. However, since the nonlinear terms neglected in SAC-SD calculations are usually rather unimportant, the differences between the SAC-SD and CCSD results are often small. They may become larger when the S_1 and S_2 clusters become large, as is the case at stretched nuclear geometries, but in those cases neither CCSD nor SAC-SD are accurate. This is shown in Sec. III, where we compare the SAC-SD and CCSD potential energy curves of the CH^+ and OH^- ions, which serve as reference closed-shell systems for the EA and IP SAC-CI and EOMCC calculations for the CH and OH

radicals. The fact that the cluster operator S of SAC is traditionally defined via the spin- and symmetry-adapted excited CSFs and that SAC equations are programed and solved in a CI-like fashion, using the energy-dependent equations [Eqs. (18) and (19)] rather than the energy-independent equations defining all standard CC methods [Eq. (21)], which are usually derived using diagrammatic techniques, does not play any significant role in discussing the differences between the SAC-SD and CCSD results, particularly when the SAC-SD and CCSD approaches are applied to closed-shell systems and spin- and symmetry-adapted RHF reference determinants, as is the case in this work.

The above similarities between SAC/SAC-CI and CC/EOMCC methods persist when we analyze the corresponding equations defining the EA and IP formalisms. Indeed, once S is determined by solving Eqs. (18) and (19), the amplitudes r_a and $r^{i_1\cdots i_n}_{aa_1\cdots a_n}$ defining the EA SAC-CI wave functions $|\Psi^{(N+1)}_\mu\rangle$ or the amplitudes r^i and $r^{ii_1\cdots i_n}_{a_1\cdots a_n}$ defining the IP SAC-CI wave functions $|\Psi^{(N-1)}_\mu\rangle (n=1,\ldots,M_R)$ are obtained by solving the eigenvalue problem

$$HR^{(N\pm1)}_\mu e^S|\Phi\rangle = E^{(N\pm1)}_\mu R^{(N\pm1)}_\mu e^S|\Phi\rangle, \tag{23}$$

in the relevant subspace of $\mathcal{H}^{(N+1)}$ (the EA case) or $\mathcal{H}^{(N-1)}$ (the IP case). Here and elsewhere in this paper, $\mathcal{H}^{(N+1)}$ and $\mathcal{H}^{(N-1)}$ designate the appropriate $(N+1)$- and $(N-1)$-electron subspaces of the Fock space. Thus, the subspace of $\mathcal{H}^{(N+1)}$ used to solve the EA SAC-CI eigenvalue problem [the $(N+1)$-electron variant of Eq. (23)] is spanned by the $|\Phi^a\rangle = a^a|\Phi\rangle$ and $|\Phi^{aa_1\cdots a_n}_{i_1\cdots i_n}\rangle = a^a a^{a_1}\cdots a^{a_n}a_{i_n}\cdots a_{i_1}|\Phi\rangle$ $(n=1,\ldots,M_R)$ determinants, while the subspace of $\mathcal{H}^{(N-1)}$ used to solve the IP SAC-CI eigenvalue problem [the $(N-1)$-electron variant of Eq. (23)] is spanned by the $|\Phi_i\rangle = a_i|\Phi\rangle$ and $|\Phi^{a_1\cdots a_n}_{ii_1\cdots i_n}\rangle = a^{a_1}\cdots a^{a_n}a_{i_n}\cdots a_{i_1}a_i|\Phi\rangle$ $(n=1,\ldots,M_R)$ determinants [in analogy to the ground-state SAC calculations for the N-electron closed-shell reference system, one uses the spin- and symmetry-adapted CSFs corresponding to $|\Phi^a\rangle$ and $|\Phi^{aa_1\cdots a_n}_{i_1\cdots i_n}\rangle$ in the EA case or $|\Phi_i\rangle$ and $|\Phi^{a_1\cdots a_n}_{ii_1\cdots i_n}\rangle$ in the IP case when solving Eq. (23) in the $\mathcal{H}^{(N+1)}$ or $\mathcal{H}^{(N-1)}$ subspace]. Again, if we do not neglect any nonlinear terms in S, the EA/IP SAC-CI equations [Eqs. (18), (19), and (23)] are equivalent to the corresponding EA/IP EOMCC equations, which are defined by Eqs. (20) and (21), and the non-Hermitian eigenvalue problem

$$\bar{H}R^{(N\pm1)}_\mu|\Phi\rangle = E^{(N\pm1)}_\mu R^{(N\pm1)}_\mu|\Phi\rangle, \tag{24}$$

in which we diagonalize the similarity-transformed Hamiltonian $\bar{H}$ [Eq. (22)] in the relevant $(N+1)$- or $(N-1)$-electron subspace, $\mathcal{H}^{(N+1)}$ or $\mathcal{H}^{(N-1)}$, respectively, as defined above, to obtain the r_a and $r^{i_1\cdots i_n}_{aa_1\cdots a_n}$ amplitudes of $R^{(N+1)}_\mu$ in the EA case or the r^i and $r^{ii_1\cdots i_n}_{a_1\cdots a_n}$ amplitudes of $R^{(N-1)}_\mu$ in the IP case, where $n=1,\ldots,M_R$.

It is only when we start neglecting nonlinear terms in S in the EA and IP SAC-CI eigenvalue problem [Eq. (23)] that the mathematical equivalence between the EA/IP SAC-CI and EOMCC equations is lost. Moreover, the *ad hoc* truncations in the many-body expansions defining S and $R^{(N\pm1)}_\mu$ may affect the size intensivity of results (discussed in detail, in the EOMCC context, in Refs. 30 and 120; cf. also Refs. 33, 116, and 121–123) by introducing the disconnected components of the operator product $\bar{H}R^{(N\pm1)}_\mu$ into Eq. (24) that do not cancel out if $M_R > M_S$.[10] Indeed, as shown in Ref. 10, the eigenvalue problem defined by Eq. (24), understood as a diagonalization of $\bar{H}$ [Eq. (22)] with S truncated at the M_S-body clusters as in Eq. (3), in the $\mathcal{H}^{(N+1)}$ subspace spanned by $|\Phi^a\rangle$ and $|\Phi^{aa_1\cdots a_n}_{i_1\cdots i_n}\rangle$ or in the $\mathcal{H}^{(N-1)}$ subspace spanned by $|\Phi_i\rangle$ and $|\Phi^{a_1\cdots a_n}_{ii_1\cdots i_n}\rangle$ $(n=1,\ldots,M_R)$, becomes equivalent to the explicitly connected eigenvalue problem

$$(\bar{H}_{N,\mathrm{open}}R^{(N\pm1)}_\mu)_C|\Phi\rangle = \omega^{(N\pm1)}_\mu R^{(N\pm1)}_\mu|\Phi\rangle, \tag{25}$$

only when $M_R \leqslant M_S$ and only when the cluster operator S satisfies Eq. (21), meaning that no terms nonlinear in S are neglected. Here, we define

$$\omega^{(N\pm1)}_\mu = E^{(N\pm1)}_\mu - E^{(N)}_0, \tag{26}$$

where $E^{(N)}_0$ is the ground-state CC energy of the N-electron reference system [Eq. (20)] and

$$\bar{H}_{N,\mathrm{open}} \equiv (H_N e^S)_{C,\mathrm{open}} = e^{-S}H_N e^S - (H_N e^S)_{C,\mathrm{closed}} = \bar{H} - E^{(N)}_0, \tag{27}$$

where $\bar{H}$ is the similarity-transformed Hamiltonian defined by Eq. (22), with S defined by Eq. (3), $H_N = H - \langle\Phi|H|\Phi\rangle$ is the Hamiltonian in the normal-ordered form relative to the Fermi vacuum $|\Phi\rangle$, and the subscripts "open," "closed," and C refer to open (i.e., having external lines), closed (i.e., having no external lines), and connected parts of a given operator expression. If the condition $M_R \leqslant M_S$ is not satisfied, we can only write

$$\bar{H}_{N,\mathrm{open}}R^{(N\pm1)}_\mu|\Phi\rangle = \omega^{(N\pm1)}_\mu R^{(N\pm1)}_\mu|\Phi\rangle, \tag{28}$$

even when all relevant nonlinear terms in S are included in the calculations. This introduces size-intensivity errors into the resulting "excitation" energies $\omega^{(N\pm1)}_\mu$, Eq. (26), when $M_R > M_S$. In fact, one might even argue, based on the analogy between the EA and IP EOMCC methods and the valence-universal MRCC approach in the (1,0) and (0,1) sectors of the Fock space,[124] that it may be more appropriate to use the $M_R = M_S - 1$ condition for the truncations in S and $R^{(N\pm1)}_\mu$ to obtain the formally correct EA and IP EOMCC approximations [the simplest EA-EOMCCSD($2p$-$1h$) =EA-EOMCCSD (Refs. 36 and 37) and IP-EOMCCSD($2h$-$1p$)=IP-EOMCCSD (Refs. 38–41) approximations and their higher-order EA-EOMCCSDT,[84] IP-EOMCCSDT,[85,86] EA-EOMCCSDTQ,[47] and IP-EOMCCSDTQ[47] analogs are in this category], but the analysis presented in Ref. 10, supported by the numerical evidence presented in Refs. 34, 35, and 88 and in Sec. III, shows that the $M_R = M_S$ condition that leads to the explicitly connected EA and IP EOMCC schemes, such as EA-EOMCCSD($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$) (Refs. 34, 35, and 88), is perfectly acceptable in calculations of ground and excited states of radical species.

The issue that is far more important from the point of view of the accuracies in applications involving potential

energy surfaces of radicals is the inclusion of the $4p$-$3h$/$4h$-$3p$ excitations in $R_\mu^{(N\pm1)}$. Indeed, the $M_R \leqslant M_S$ condition, which is needed to convert Eqs. (23) and (28) into the explicitly connected Eq. (25), is satisfied by the EA-EOMCCSD($2p$-$1h$), IP-EOMCCSD($2h$-$1p$), EA-EOMCCSD($3p$-$2h$), and IP-EOMCCSD($3h$-$2p$) approaches, for which M_S=2 and M_R equals 1, 1, 2, and 2, respectively. All of these approaches lead to a rigorously size intensive description of the "excitation" energies $\omega_\mu^{(N\pm1)}$ defined by Eq. (26), but the corresponding results at larger internuclear separations of radicals are generally poor. The $M_R \leqslant M_S$ condition is also satisfied by the SAC-CI analogs of the above four approaches, including the SAC-CI($3p$-$2h$) and SAC-CI($3h$-$2p$) methods and their active-space variants, although the SAC-CI($3p$-$2h$) and SAC-CI($3h$-$2p$) approaches tested in this work neglect the nonlinear terms in S in Eq. (23), i.e., the corresponding wave functions

$$|\Psi_\mu^{(N+1)}\rangle = (R_{\mu,1p} + R_{\mu,2p\text{-}1h} + R_{\mu,3p\text{-}2h})e^{S_1+S_2}|\Phi\rangle \tag{29}$$

and

$$|\Psi_\mu^{(N-1)}\rangle = (R_{\mu,1h} + R_{\mu,2h\text{-}1p} + R_{\mu,3h\text{-}2p})e^{S_1+S_2}|\Phi\rangle \tag{30}$$

are approximated by

$$|\Psi_\mu^{(N+1)}\rangle = (R_{\mu,1p} + R_{\mu,2p\text{-}1h} + R_{\mu,3p\text{-}2h})(1+S_2)|\Phi\rangle \tag{31}$$

and

$$|\Psi_\mu^{(N-1)}\rangle = (R_{\mu,1h} + R_{\mu,2h\text{-}1p} + R_{\mu,3h\text{-}2p})(1+S_2)|\Phi\rangle, \tag{32}$$

respectively, and this leads to a departure from a strictly size intensive behavior. This is not a problem for the analysis presented in this work, where all of the calculations are performed for small many-electron systems, for which the exact, full CI results can be generated for comparison purposes. In fact, the full treatment of nonlinear terms in S offered by the corresponding EA-EOMCCSD($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$) approaches, which rely on Eqs. (29) and (30) rather than Eqs. (31) and (32) and on the explicitly connected form of the eigenvalue problem [Eq. (25)], does not improve the relatively poor SAC-CI($3p$-$2h$) and SAC-CI($3h$-$2p$) results at larger internuclear separations of the CH and OH radicals. As shown in Sec. III, one has to include the genuine $4p$-$3h$ and $4h$-$3p$ components of the $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$ operators, as in Eqs. (14) and (15), and their active-space analogs discussed in Sec. II B to obtain an accurate description of the ground- and excited-state potential energy curves of CH and OH. This is done in the present paper via the SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods and their active-space variants in which, in analogy to the aforementioned linearized forms of the SAC-CI($3p$-$2h$) and SAC-CI($3h$-$2p$) approaches, instead of the complete wave functions

$$|\Psi_\mu^{(N+1)}\rangle = (R_{\mu,1p} + R_{\mu,2p\text{-}1h} + R_{\mu,3p\text{-}2h} + R_{\mu,4p\text{-}3h})e^{S_1+S_2}|\Phi\rangle \tag{33}$$

and

$$|\Psi_\mu^{(N-1)}\rangle = (R_{\mu,1h} + R_{\mu,2h\text{-}1p} + R_{\mu,3h\text{-}2p} + R_{\mu,4h\text{-}3p})e^{S_1+S_2}|\Phi\rangle, \tag{34}$$

we use the simplified forms

$$|\Psi_\mu^{(N+1)}\rangle = (R_{\mu,1p} + R_{\mu,2p\text{-}1h} + R_{\mu,3p\text{-}2h} + R_{\mu,4p\text{-}3h}) \times (1+S_2)|\Phi\rangle \tag{35}$$

and

$$|\Psi_\mu^{(N-1)}\rangle = (R_{\mu,1h} + R_{\mu,2h\text{-}1p} + R_{\mu,3h\text{-}2p} + R_{\mu,4h\text{-}3p}) \times (1+S_2)|\Phi\rangle, \tag{36}$$

respectively, in which only the lead terms linear in S_2 are retained. The SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods are not rigorously size intensive, since they do not satisfy the condition $M_R \leqslant M_S$ (M_R=3 and M_S=2 in this case) and neglect nonlinear terms in S but, as shown in Sec. III, they are sufficiently accurate to provide an excellent description of the entire potential energy curves of the low-lying states of the CH and OH radicals. Clearly, it would eventually be more proper to use the more complete and size intensive EA-EOMCCSDT($4p$-$3h$) and IP-EOMCCSDT($4h$-$3p$) approximations, in which all nonlinear terms resulting from e^S are retained and M_R=M_S=3, or the EA-EOMCCSDTQ($4p$-$3h$)=EA-EOMCCSDTQ and IP-EOMCCSDTQ($4h$-$3p$)=IP-EOMCCSDTQ approximations, as recently implemented in Ref. 47, in which M_R=3 and M_S=4, but this is not necessary to prove one of the main points of the present paper, which has not received much attention in the literature, that one needs the $4p$-$3h$ and $4h$-$3p$ components of the $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$ operators to obtain accurate ground- and excited-state potential energy curves of radical species along the relevant bond breaking coordinates. Moreover, the development of the EA-EOMCCSDT($4p$-$3h$), IP-EOMCCSDT($4h$-$3p$), EA-EOMCCSDTQ($4p$-$3h$), and IP-EOMCCSDTQ($4h$-$3p$) codes, albeit in principle possible,[47] leads to computational schemes that are prohibitively expensive anyway. Thus, for now, we will rely on the simplified SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods, as described by Eqs. (35) and (36), respectively, in our analyses. Even these severely truncated approaches are usually much too expensive, particularly when larger systems/larger basis sets are employed. Thus, it is essential to simplify the SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods even further. The idea advocated in this paper is that of the active-space SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods in which the large numbers of r_{abc}^{jk} and r_{abcd}^{jkl} amplitudes defining the $3p$-$2h$ and $4p$-$3h$ components of $R_\mu^{(N+1)}$ and the large numbers of r_{bc}^{ijk} and r_{bcd}^{ijkl} amplitudes defining the $3h$-$2p$ and $4h$-$3p$ components of $R_\mu^{(N-1)}$ are significantly reduced through the use of active orbitals.

B. The active-space EA and IP SAC-CI and EOMCC approaches

As shown in this paper, the EA and IP SAC-CI approaches truncated at $4p$-$3h$/$4h$-$3p$ excitations are capable of providing a highly accurate description of the entire potential energy curves of the ground and low-lying excited states of

radical species, but, as pointed out in the Introduction, the resulting SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods and their EA and IP EOMCC analogs are much too expensive to be widely applicable. In fact, even the less expensive EA and IP SAC-CI and EOMCC approaches truncated at $3p$-$2h$/$3h$-$2p$ excitations, which work well in the Frank-Condon region,[34,35,43,88] are not very practical, particularly when the number of electrons in a radical system of interest is larger or a larger basis set is employed. This is a consequence of the fact that the numbers of $r^{jk}_{abc}/r^{ijk}_{bc}$ and $r^{jkl}_{abcd}/r^{ijkl}_{bcd}$ amplitudes defining the relevant $R^{(N\pm1)}_{\mu}$ operators, particularly the numbers of the $4p$-$3h$ and $4h$-$3p$ amplitudes r^{jkl}_{abcd} and r^{ijkl}_{bcd}, which are $\sim n_o^3 n_u^4$ and $\sim n_o^4 n_u^3$, respectively, are far too large for the majority of applications. The SAC-CI($4p$-$3h$)/PS and SAC-CI($4h$-$3p$)/PS methods mentioned in the Introduction, in which the small r^{jkl}_{abcd} and r^{ijkl}_{bcd} amplitudes that do not significantly perturb the suitably chosen zero-order wave functions, obtained with the CI($3p$-$2h$) and CI($3h$-$2p$) approaches, are eliminated from the calculations with the help of numerical thresholds, following the recipes described in Refs. 42, 44, and 93, substantially reduce the costs of the SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) calculations, but there are good reasons, also discussed in the Introduction, for seeking alternative ways of reducing these costs. The active-space SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods developed in this work and their recently proposed EA and IP EOMCC analogs, such as the EA-EOMCCSDt, EA-EOMCCSDtq, IP-EOMCCSDt, and IP-EOMCCSDtq approaches defined in Ref. 34, represent one of the most promising solutions to this problem.

The active-space EA and IP SAC-CI methods examined in this paper and their EOMCC counterparts discussed in Refs. 34, 35, and 88 are based on the observation that not all $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ amplitudes $r^{jk}_{abc}/r^{ijk}_{bc}$ and $r^{jkl}_{abcd}/r^{ijkl}_{bcd}$ are equally important. In many cases, one can *a priori* select the dominant $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ terms based on the model in which a radical is interpreted as a species obtained by either attaching an electron to one of the lowest-energy unoccupied orbitals or removing an electron from one of the highest-energy occupied orbitals of the related closed-shell system. For example, the CH and OH radicals examined in Sec. III are obtained, at least at the zero-order level, by attaching an electron to the lowest-energy unoccupied 1π orbitals of the CH^+ closed-shell ion (the CH case) or by removing an electron from the highest-energy occupied 1π orbitals of the OH^- closed-shell ion (the OH case), respectively. Thus, in analogy to the active-space CC and EOMCC methods, pioneered by Adamowicz, Piecuch, and co-workers,[73,74,90,91,98–100,107] and multireference theories, one can treat the 1π shells of CH^+ and OH^- as active orbitals for selecting the most important $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ amplitudes in the EA and IP SAC-CI and EOMCC calculations for CH and OH. The choice of 1π shells of CH^+ and OH^- as active orbitals may not be sufficient for the calculations at larger C-H and O-H distances that interest us in this work, since the 1π orbitals of the CH^+ and OH^- systems become asymptotically degenerate with the corresponding 3σ and 4σ orbitals, but this can be easily addressed by selecting all valence orbitals of CH^+ and OH^- that correlate with the $2p$ shells of C and O and $1s$ shell of H (the 3σ, $1\pi_x$, $1\pi_y$, and 4σ orbitals of CH^+ and OH^-) as active orbitals for the EA and IP SAC-CI and EOMCC calculations with $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ excitations (see Fig. 1). This is equivalent to stating that at larger C-H distances, the low-lying electronic states of the CH radical are formed by the process of adding an electron to the 1π or 4σ unoccupied valence orbitals of CH^+, which can be combined with the electronic excitations from the highest occupied 3σ orbital that belongs to the same valence shell as the 1π or 4σ orbitals. In the case of OH, we might say that at larger O-H separations, the low-lying electronic states of the OH radical are formed by the process of removing an electron from the 1π shell of OH^-, which can be coupled with the electronic excitations from the 3σ and 1π occupied orbitals to the 4σ orbital that again belongs to the same valence shell as the 3σ and 1π orbitals.

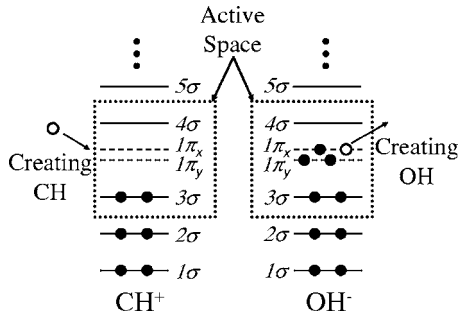

FIG. 1. Orbital levels of the CH^+ and OH^- ions and a schematic representation of the electron attachment and ionization processes that lead to the formation of the CH and OH radicals from the CH^+ and OH^- reference closed-shell systems. Valence shells of CH^+ and OH^- that play a dominant role in the relevant electron attachment and ionization processes and that are used in the active-space EA and IP EOMCC and SAC-CI calculations discussed in the text are emphasized with the help of dotted frames.

The above situation, where the process of forming radical species from the related closed-shell species is defined by a relatively small subset of orbitals which are involved in the relevant electron attachment or electron removal process, occurs in a large number of molecular systems. It is, therefore, reasonable, in a manner similar to that of multireference approaches, to use these orbitals as active orbitals in the EA-EOMCC and IP-EOMCC calculations. Formally, in order to define the active-space EA and IP EOMCC and SAC-CI methods, we first divide the available spin-orbitals of a closed-shell N-electron system into four disjoint groups of core or inactive occupied spin-orbitals (labeled by lowercase bold letters $\mathbf{i},\mathbf{j},\ldots$), active spin-orbitals occupied in the closed-shell reference determinant $|\Phi\rangle$ (labeled by uppercase bold letters $\mathbf{I},\mathbf{J},\ldots$), active spin-orbitals unoccupied in $|\Phi\rangle$ (labeled by uppercase bold letters $\mathbf{A},\mathbf{B},\ldots$), and virtual or inactive unoccupied spin-orbitals (labeled by lowercase bold letters $\mathbf{a},\mathbf{b},\ldots$). We continue to designate the occupied and unoccupied spin-orbitals in $|\Phi\rangle$ by the italic characters $i,j,\ldots$ and $a,b,\ldots$, respectively, if the active/inactive character of the spin-orbitals is not specified. Once the above orbital classification scheme is established, we use it to define the electron attaching and electron removing operators $R^{(N\pm1)}_{\mu}$ of the active-space EA and IP EOMCC and SAC-CI methods, following the general strategy described in Ref. 34.

For example, the active-space EA-EOMCCSD(3p-2h) and SAC-CI(3p-2h) methods are obtained by replacing the 3p-2h component $R_{\mu,3p\text{-}2h}$ of the electron attaching operator $R_\mu^{(N+1)}$, Eq. (10), by[34]

$$r_{\mu,3p\text{-}2h} = \sum_{j>k,\mathbf{A}<b<c} r_{\mathbf{A}bc}^{jk} a^{\mathbf{A}} a^b a^c a_k a_j. \tag{37}$$

Thus, the active-space EA-EOMCCSD(3p-2h) and SAC-CI(3p-2h) methods reduce the relatively large number of all 3p-2h amplitudes r_{abc}^{jk} that enter $R_{\mu,3p\text{-}2h}$, which is (if we ignore symmetries) $\sim n_o^2 n_u^3$, to $\sim N_u n_o^2 n_u^2$, where N_u is the number of active orbitals unoccupied in the reference determinant $|\Phi\rangle$. Assuming that $N_u \ll n_u$, this is a major reduction in the computational effort compared to full EA-EOMCCSD(3p-2h) and SAC-CI(3p-2h) calculations. The active-space EA-EOMCCSD(3p-2h) and SAC-CI(3p-2h) approaches using N_u unoccupied active orbitals are designated by EA-EOMCCSD(3p-2h)$\{N_u\}$ and SAC-CI(3p-2h)$\{N_u\}$, respectively. Clearly, the EA-EOMCCSD(3p-2h)$\{n_u\}$ and SAC-CI(3p-2h)$\{n_u\}$ methods, in which all orbitals are chosen to be active, are equivalent to the full EA-EOMCCSD(3p-2h) and SAC-CI(3p-2h) approaches. The relatively small set of the unknown amplitudes $r_{\mathbf{A}bc}^{jk}$ defining $r_{\mu,3p\text{-}2h}$ [Eq. (37)], in which at least one of the three unoccupied spin-orbital indices is active, and the remaining 1p and 2p-1h amplitudes r_a and r_{ab}^j that enter the (N+1)-electron wave functions of the active-space EA-EOMCCSD(3p-2h)$\{N_u\}$ approach,

$$|\Psi_\mu^{(N+1)}\rangle = (R_{\mu,1p} + R_{\mu,2p\text{-}1h} + r_{\mu,3p\text{-}2h}) e^{S_1+S_2} |\Phi\rangle, \tag{38}$$

where S_1 and S_2 are the singly and doubly excited clusters obtained in the preceding CCSD calculations for the closed-shell, N-electron, reference system, are obtained by diagonalizing the similarity-transformed Hamiltonian of CCSD,

$$\bar{H}_{N,\text{open}}^{(\text{CCSD})} \equiv (H_N e^{S_1+S_2})_{C,\text{open}}, \tag{39}$$

in the subspace of $\mathcal{H}^{(N+1)}$ spanned by the $|\Phi^a\rangle$, $|\Phi_j^{ab}\rangle$, and $|\Phi_{jk}^{\mathbf{A}bc}\rangle$ determinants [cf. Eq. (25)]. In the case of the active-space SAC-CI(3p-2h)$\{N_u\}$ approach, as implemented in this paper, where the wave functions $|\Psi_\mu^{(N+1)}\rangle$, [Eq. (38)], are simplified to [see Eq. (31)]

$$|\Psi_\mu^{(N+1)}\rangle = (R_{\mu,1p} + R_{\mu,2p\text{-}1h} + r_{\mu,3p\text{-}2h})(1 + S_2) |\Phi\rangle, \tag{40}$$

where S_2 is obtained in the ground-state SAC-SD calculations for the N-electron reference system, the relevant r_a, r_{ab}^j, and $r_{\mathbf{A}bc}^{jk}$ amplitudes defining $R_{\mu,1p}$, $R_{\mu,2p\text{-}1h}$, and $r_{\mu,3p\text{-}2h}$ are obtained by solving the system of equations obtained by projecting the eigenvalue problem given by Eq. (23), in which e^S is replaced by $(1+S_2)$ and $R_\mu^{(N+1)}$ by $(R_{\mu,1p}+R_{\mu,2p\text{-}1h}+r_{\mu,3p\text{-}2h})$, onto the CSFs that correspond to the $|\Phi^a\rangle$, $|\Phi_j^{ab}\rangle$, and $|\Phi_{jk}^{\mathbf{A}bc}\rangle$ determinants.

Similarly, the active-space IP-EOMCCSD(3h-2p) and SAC-CI(3h-2p) methods are obtained by replacing the 3h-2p component $R_{\mu,3h\text{-}2p}$ of the $R_\mu^{(N-1)}$ operator [Eq. (11)], by[34]

$$r_{\mu,3h\text{-}2p} = \sum_{\mathbf{I}>j>k,b<c} r_{bc}^{\mathbf{I}jk} a^b a^c a_k a_j a_{\mathbf{I}}. \tag{41}$$

Again, the active-space IP-EOMCCSD(3h-2p) and SAC-CI(3h-2p) methods offer considerable savings in the computer effort, compared to the full IP-EOMCCSD(3h-2p) and SAC-CI(3h-2p) approaches, reducing the relatively large number of $\sim n_o^3 n_u^2$ of all r_{bc}^{ijk} amplitudes to a smaller number of $\sim N_o n_o^2 n_u^2$ amplitudes $r_{bc}^{\mathbf{I}jk}$ that enter the $r_{\mu,3h\text{-}2p}$ operator [Eq. (41)], where $N_o < n_o$ is the number of active orbitals occupied in $|\Phi\rangle$. The active-space IP-EOMCCSD(3h-2p) and SAC-CI(3h-2p) approaches using N_o occupied active orbitals are designated by IP-EOMCCSD(3h-2p)$\{N_o\}$ and SAC-CI(3h-2p)$\{N_o\}$, respectively. The IP-EOMCCSD(3h-2p)$\{n_o\}$ and SAC-CI(3h-2p)$\{n_o\}$ methods, in which all orbitals are active, are equivalent to the parent IP-EOMCCSD(3h-2p) and SAC-CI(3h-2p) approximations. The $r_{bc}^{\mathbf{I}jk}$ amplitudes defining $r_{\mu,3h\text{-}2p}$ [Eq. (41)], in which at least one of the three occupied spin-orbital indices is active, and the remaining 1h and 2h-1p amplitudes r^i and r_b^{ij} that define the (N−1)-electron wave functions of the active-space IP-EOMCCSD(3h-2p)$\{N_o\}$ approach,

$$|\Psi_\mu^{(N-1)}\rangle = (R_{\mu,1h} + R_{\mu,2h\text{-}1p} + r_{\mu,3h\text{-}2p}) e^{S_1+S_2} |\Phi\rangle, \tag{42}$$

are obtained by diagonalizing the similarity-transformed Hamiltonian obtained in the ground-state CCSD calculations for the closed-shell, N-electron, reference system [Eq. (39)] in the subspace of $\mathcal{H}^{(N-1)}$ spanned by the $|\Phi_i\rangle$, $|\Phi_{ij}^b\rangle$, and $|\Phi_{\mathbf{I}jk}^{bc}\rangle$ determinants [cf. Eq. (25)]. In the active-space SAC-CI(3h-2p)$\{N_o\}$ approach, as implemented in this work, where the wave functions $|\Psi_\mu^{(N-1)}\rangle$ given by Eq. (42) are simplified to [see Eq. (32)]

$$|\Psi_\mu^{(N-1)}\rangle = (R_{\mu,1h} + R_{\mu,2h\text{-}1p} + r_{\mu,3h\text{-}2p})(1 + S_2) |\Phi\rangle, \tag{43}$$

the r^i, r_b^{ij}, and $r_{bc}^{\mathbf{I}jk}$ coefficients defining $R_{\mu,1h}$, $R_{\mu,2h\text{-}1p}$, and $r_{\mu,3h\text{-}2p}$ are determined by solving the system of equations obtained by projecting the eigenvalue problem given by Eq. (23), in which e^S and $R_\mu^{(N-1)}$ are replaced by $(1+S_2)$ and $(R_{\mu,1h}+R_{\mu,2h\text{-}1p}+r_{\mu,3h\text{-}2p})$, respectively, onto the CSFs corresponding to the $|\Phi_i\rangle$, $|\Phi_{ij}^b\rangle$, and $|\Phi_{\mathbf{I}jk}^{bc}\rangle$ determinants.

The savings in the computer effort offered by the active-space EA and IP SAC-CI and EOMCC approaches truncated at the 3p-2h and 3h-2p excitations can be quite significant compared to the parent EA-EOMCCSD(3p-2h)/SAC-CI(3p-2h) and IP-EOMCCSD(3h-2p)/SAC-CI(3h-2p) approximations, particularly in the EA case where the N_u/n_u ratio which defines these savings is usually a small number. These savings become even larger when the active-space schemes with 4p-3h and 4h-3p excitations are considered, since in that case the active-space restrictions on the indices defining the relevant r_{abcd}^{jkl} and r_{bcd}^{ijkl} amplitudes apply to both occupied and unoccupied spin-orbitals. For example, following the general prescription described in Sec. II C of Ref. 34 and using Eqs. (35) and (36), we define the wave functions $|\Psi_\mu^{(N+1)}\rangle$ and $|\Psi_\mu^{(N-1)}\rangle$ exploited in the active-space SAC-CI(4p-3h) and SAC-CI(4h-3p) methods employing N_o active occupied and N_u active unoccupied spin-orbitals,

designated as the SAC-CI(4*p*-3*h*)$\{N_o,N_u\}$ and SAC-CI(4*h*-3*p*)$\{N_o,N_u\}$ schemes, respectively, in the following manner:

$$|\Psi_\mu^{(N+1)}\rangle = (R_{\mu,1p} + R_{\mu,2p\text{-}1h} + r_{\mu,3p\text{-}2h} + r_{\mu,4p\text{-}3h}) \times (1+S_2)|\Phi\rangle \tag{44}$$

and

$$|\Psi_\mu^{(N-1)}\rangle = (R_{\mu,1h} + R_{\mu,2h\text{-}1p} + r_{\mu,3h\text{-}2p} + r_{\mu,4h\text{-}3p}) \times (1+S_2)|\Phi\rangle, \tag{45}$$

where S_2 is a doubly excited cluster operator obtained in the ground-state SAC-SD calculations for the N-electron reference system. The $r_{\mu,3p\text{-}2h}$ and $r_{\mu,3h\text{-}2p}$ operators entering Eqs. (44) and (45) are exactly the same as those used in the active-space EA-EOMCCSD(3*p*-2*h*)/SAC-CI(3*p*-2*h*) and IP-EOMCCSD(3*h*-2*p*)/SAC-CI(3*h*-2*p*) calculations [see Eqs. (37) and (41)], whereas the $r_{\mu,4p\text{-}3h}$ and $r_{\mu,4h\text{-}3p}$ components are given by[34]

$$r_{\mu,4p\text{-}3h} = \sum_{\mathbf{J}>k>l,\mathbf{A}<\mathbf{B}<c<d} r_{\mathbf{AB}cd}^{\mathbf{J}kl} a^{\mathbf{A}} a^{\mathbf{B}} a^c a^d a_l a_k a_{\mathbf{J}} \tag{46}$$

and

$$r_{\mu,4h\text{-}3p} = \sum_{\mathbf{I}>\mathbf{J}>k>l,\mathbf{B}<c<d} r_{\mathbf{B}cd}^{\mathbf{IJ}kl} a^{\mathbf{B}} a^c a^d a_l a_k a_{\mathbf{J}} a_{\mathbf{I}}, \tag{47}$$

respectively. The explicit equations for the unknown r_a, r_{ab}^i, $r_{\mathbf{A}bc}^{jk}$, and $r_{\mathbf{AB}cd}^{\mathbf{J}kl}$ amplitudes, which define the $R_{\mu,1p}$, $R_{\mu,2p\text{-}1h}$, $r_{\mu,3p\text{-}2h}$, and $r_{\mu,4p\text{-}3h}$ operators entering the wave functions $|\Psi_\mu^{(N+1)}\rangle$ of the SAC-CI(4*p*-3*h*)$\{N_o,N_u\}$ approach [Eq. (44)], are obtained by projecting the SAC-CI eigenvalue problem given by Eq. (23), in which e^S is replaced by $(1+S_2)$ and $R_\mu^{(N+1)}$ by $(R_{\mu,1p}+R_{\mu,2p\text{-}1h}+r_{\mu,3p\text{-}2h}+r_{\mu,4p\text{-}3h})$, onto the CSFs corresponding to the $|\Phi^a\rangle$, $|\Phi_j^{ab}\rangle$, $|\Phi_{jk}^{\mathbf{A}bc}\rangle$, and $|\Phi_{\mathbf{J}kl}^{\mathbf{AB}cd}\rangle$ determinants. Similarly, the explicit equations for the r^i, r_b^{ij}, $r_{bc}^{\mathbf{I}jk}$, and $r_{\mathbf{B}cd}^{\mathbf{IJ}kl}$ amplitudes, which define the $R_{\mu,1h}$, $R_{\mu,2h\text{-}1p}$, $r_{\mu,3h\text{-}2p}$, and $r_{\mu,4h\text{-}3p}$ operators entering the wave functions $|\Psi_\mu^{(N-1)}\rangle$ of the SAC-CI(4*h*-3*p*)$\{N_o,N_u\}$ approach [Eq. (45)], are obtained by projecting the SAC-CI eigenvalue problem [Eq. (23)], where e^S is replaced by $(1+S_2)$ and $R_\mu^{(N-1)}$ by $(R_{\mu,1h}+R_{\mu,2h\text{-}1p}+r_{\mu,3h\text{-}2p}+r_{\mu,4h\text{-}3p})$, onto the CSFs corresponding to determinants $|\Phi_i\rangle$, $|\Phi_{ij}^b\rangle$, $|\Phi_{\mathbf{I}jk}^{bc}\rangle$, and $|\Phi_{\mathbf{IJ}kl}^{\mathbf{B}cd}\rangle$. Clearly, we could easily write the analogous expressions for the active-space EA-EOMCCSD(4*p*-3*h*)$\{N_o,N_u\}$ and IP-EOMCCSD(4*h*-3*p*)$\{N_o,N_u\}$ methods. In the case of EA-EOMCCSD(4*p*-3*h*)$\{N_o,N_u\}$, the explicit equations for the relevant r_a, r_{ab}^i, $r_{\mathbf{A}bc}^{jk}$, and $r_{\mathbf{AB}cd}^{\mathbf{J}kl}$ amplitudes, which define the EA-EOMCCSD(4*p*-3*h*)$\{N_o,N_u\}$ wave functions $|\Psi_\mu^{(N+1)}\rangle$ via the formula

$$|\Psi_\mu^{(N+1)}\rangle = (R_{\mu,1p} + R_{\mu,2p\text{-}1h} + r_{\mu,3p\text{-}2h} + r_{\mu,4p\text{-}3h}) e^{S_1+S_2}|\Phi\rangle, \tag{48}$$

where S_1 and S_2 are determined by solving the CCSD equations for the N-electron reference system, would be obtained by projecting the eigenvalue problem involving the similarity-transformed Hamiltonian of CCSD [Eq. (39)], as given by Eq. (28), on $|\Phi^a\rangle$, $|\Phi_j^{ab}\rangle$, $|\Phi_{jk}^{\mathbf{A}bc}\rangle$, and $|\Phi_{\mathbf{J}kl}^{\mathbf{AB}cd}\rangle$. Similarly, the explicit equations for the r^i, r_b^{ij}, $r_{bc}^{\mathbf{I}jk}$, and $r_{\mathbf{B}cd}^{\mathbf{IJ}kl}$ amplitudes, which define the IP-EOMCCSD(4*h*-3*p*)$\{N_o,N_u\}$ wave functions $|\Psi_\mu^{(N-1)}\rangle$ via

$$|\Psi_\mu^{(N-1)}\rangle = (R_{\mu,1h} + R_{\mu,2h\text{-}1p} + r_{\mu,3h\text{-}2p} + r_{\mu,4h\text{-}3p}) e^{S_1+S_2}|\Phi\rangle, \tag{49}$$

would be obtained by projecting Eq. (28), with the similarity-transformed Hamiltonian $\bar{H}_{N,\text{open}}$ replaced by $\bar{H}_{N,\text{open}}^{(\text{CCSD})}$ [Eq. (39)], and $R_\mu^{(N-1)}$ approximated by $(R_{\mu,1h}+R_{\mu,2h\text{-}1p}+r_{\mu,3h\text{-}2p}+r_{\mu,4h\text{-}3p})$, on the $|\Phi_i\rangle$, $|\Phi_{ij}^b\rangle$, $|\Phi_{\mathbf{I}jk}^{bc}\rangle$, and $|\Phi_{\mathbf{IJ}kl}^{\mathbf{B}cd}\rangle$ determinants. Based on the discussion in Sec II A, ultimately one would like to pursue the active-space variants of the more complete and rigorously size intensive EA-EOMCCSDT(4*p*-3*h*) and IP-EOMCCSDT(4*h*-3*p*) approximations or their recently implemented[47] EA-EOMCCSDTQ(4*p*-3*h*)=EA-EOMCCSDTQ and IP-EOMCCSDTQ(4*h*-3*p*)=IP-EOMCCSDTQ analogs, designated in Ref. 34 as the EA-EOMCCSDtq and IP-EOMCCSDtq approaches, in which the electron attaching and electron removing operators, $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$, respectively, are defined in exactly the same manner as in the active-space SAC-CI(4*p*-3*h*)$\{N_o,N_u\}$ and SAC-CI(4*h*-3*p*)$\{N_o,N_u\}$ schemes and in which the ground state of the N-electron reference system is obtained in the active-space CCSDt (Ref. 73, 74, 89–92, and 98–107) or CCSDtq (90, 99–102, and 104) calculations. We plan to work on such approaches in the future, following the efficient computer coding strategy described in Ref. 88. The primary objective of this study is to show that one can accurately and efficiently handle the most expensive 4*p*-3*h* and 4*h*-3*p* components of the $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$ operators, which are crucial for obtaining high quality potential energy curves of radical systems, via small active orbital spaces. The active-space SAC-CI(4*p*-3*h*)$\{N_o,N_u\}$ and SAC-CI(4*h*-3*p*)$\{N_o,N_u\}$ approaches are sufficient to prove this point.

Clearly, the SAC-CI(4*p*-3*h*)$\{n_o,n_u\}$ and SAC-CI(4*h*-3*p*)$\{n_o,n_u\}$ approaches, in which all orbitals are active, are equivalent to the parent SAC-CI(4*p*-3*h*) and SAC-CI(4*h*-3*p*) theories. This simple relationship between the active-space and parent methods is one of the advantages of all active-space approaches, since we always know that by increasing the active orbital space we can systematically approach the parent approximations. Of course, the main advantage of the active-space SAC-CI(4*p*-3*h*)$\{N_o,N_u\}$ and SAC-CI(4*h*-3*p*)$\{N_o,N_u\}$ approaches and their EA and IP EOMCC analogs with an active-space treatment of the 3*p*-2*h*/4*p*-3*h* and 3*h*-2*p*/4*h*-3*p* excitations, which we plan to develop in the future, is that they offer substantial savings in the computer effort compared to the parent SAC-CI(4*p*-3*h*) and SAC-CI(4*h*-3*p*) theories, and similar EOMCC approximations. In particular, the SAC-CI(4*p*-3*h*)$\{N_o,N_u\}$ method reduces the large number of 4*p*-3*h* amplitudes used in the regular, all-orbital SAC-CI(4*p*-3*h*) approach, which can be estimated at $\sim n_o^3 n_u^4$ to $\sim N_o N_u^2 n_o^2 n_u^2$. Since $N_o < n_o$ and, in the vast majority of cases, $N_u \ll n_u$, the number of 4*p*-3*h* amplitudes used in the

TABLE I. The dimensions of the full (3*p*-2*h*), active-space (3*p*-2*h*){3}, full (4*p*-3*h*), active-space (4*p*-3*h*){1,3}, and (4*p*-3*h*)/PS doublet eigenvalue problems, using the C_{2v} Abelian subgroup of $C_{\infty v}$, for the CH radical, as described by the the [5*s*3*p*1*d*/3*s*1*p*] basis set of Ref. 125.

State symmetry	(3*p*-2*h*)	(3*p*-2*h*){3}[a]	(4*p*-3*h*)	(4*p*-3*h*){1,3}[b]	(4*p*-3*h*)/PS[c]
$^2A_1$ ($^2\Sigma^+$,$^2\Delta$)	5823	2170	35 803	3688	2268–3030
$^2A_2$ ($^2\Sigma^-$,$^2\Delta$)	3711	1448	28 919	3344	1626–4071
$^2B_1$, $^2B_2$ ($^2\Pi$)	4755	1823	32 315	3495	1381–3481

[a]The active space consisted of the $1\pi_x$, $1\pi_y$, and 4σ orbitals of CH^+.
[b]The active space consisted of the 3σ occupied and $1\pi_x$, $1\pi_y$, and 4σ unoccupied orbitals of CH^+.
[c]The threshold λ_e for selection of the 4*p*-3*h* components of $R_\mu^{(N+1)}$ was set at 10^{-7} hartree.

active-space SAC-CI(4*p*-3*h*)$\{N_o,N_u\}$ calculations is a tiny fraction of all amplitudes r_{abcd}^{jkl}. This is true even for smaller radicals and smaller basis sets, such as the CH radical described by the [5*s*3*p*1*d*/3*s*1*p*] basis set of Ref. 125, tested in this work. As shown in Table I, the number of the spin- and symmetry-adapted CSFs corresponding to the largest block of 4*p*-3*h* excitations entering the SAC-CI(4*p*-3*h*){1,3} calculations for the CH/[5*s*3*p*1*d*/3*s*1*p*] system, using the 3σ, $1\pi_x$, $1\pi_y$, and 4σ orbitals of the CH^+ reference ion as active orbitals, is approximately 10% of all 4*p*-3*h* CSFs needed to define the full SAC-CI(4*p*-3*h*) eigenvalue problem. There is also a significant reduction in the number of 3*p*-2*h* excitations (by a factor of ~2.6–2.7), which enter the SAC-CI(3*p*-2*h*) and SAC-CI(4*p*-3*h*) calculations, when we use the above active orbitals. All of this translates into a reduction of the CPU time by almost two orders of magnitude, when we compare the active-space SAC-CI(4*p*-3*h*){1,3} and regular SAC-CI(4*p*-3*h*) calculations for the CH/[5*s*3*p*1*d*/3*s*1*p*] system. As shown in Table II, the savings offered by the SAC-CI(4*h*-3*p*){3,1} approach, when applied to the OH radical, as described by the 6–31G(*d*,*p*) basis set,[126,127] are less impressive, partly due to the fact that in this case N_o (=3) is almost identical to n_o (=4) and partly due to the small dimension of the employed basis set, but they are still substantial. In fact, in this case the number of spin- and symmetry-adapted 3*h*-2*p* amplitudes $r_{bc}^{\mathbf{I}jk}$ is the same as the number of all 3*h*-2*p* amplitudes r_{bc}^{ijk}, since there is only one correlated occupied orbital outside the active space. Thus, all of the observed savings in the computer effort are due to the active-space treatment of 4*h*-3*p* excitations. Clearly, the savings offered by the SAC-CI(4*h*-3*p*)$\{N_o,N_u\}$ would be greater if larger systems/larger basis sets were considered. In general, the SAC-CI(4*h*-3*p*)$\{N_o,N_u\}$ method and its IP-EOMCC analogs, including the IP-EOMCCSD(4*h*-3*p*)$\{N_o,N_u\}$ and IP-EOMCCSDtq approximations discussed above and in Ref. 34, reduce the large number of all 4*h*-3*p* amplitudes used, for example, in the regular SAC-CI(4*h*-3*p*) approach, which can be estimated at $\sim n_o^4 n_u^3$ to $\sim N_o^2 N_u n_o^2 n_u^2$. Since $N_o < n_o$ and for larger basis sets $N_u \ll n_u$, the number of 4*h*-3*p* amplitudes used in the active-space SAC-CI(4*h*-3*p*)$\{N_o,N_u\}$ calculations is much smaller than a number of all amplitudes r_{bcd}^{ijkl}.

TABLE II. The dimensions of the full (3*h*-2*p*), full (4*h*-3*p*), active-space (4*h*-3*p*){3,1}, and (4*h*-3*p*)/PS doublet eigenvalue problems, using the C_{2v} Abelian subgroup of $C_{\infty v}$, for the OH radical, as described by the 6-31G(*d*,*p*) basis set (Refs. 126 and 127).

State symmetry	(3*h*-2*p*)	(4*h*-3*p*)	(4*h*-3*p*){3,1}[a]	(4*h*-3*p*)/PS[b]
$^2A_1$ ($^2\Sigma^+$,$^2\Delta$)	1166	10 629	3075	1869–4896
$^2A_2$ ($^2\Sigma^-$,$^2\Delta$)	1090	10 467	2993	1263–3525
$^2B_1$, $^2B_2$ ($^2\Pi$)	1128	10 545	3031	1765–3708

[a]The active space consisted of the 3σ, $1\pi_x$, and $1\pi_y$ occupied and 4σ unoccupied orbitals of OH^-.
[b]The threshold λ_e for selection of the 4*h*-3*p* components of $R_\mu^{(N-1)}$ was set at 10^{-7} hartree.

Tables I and II also illustrate the point that one can achieve significant reductions in the numbers of 4*p*-3*h* and 4*h*-3*p* amplitudes by exploiting the SAC-CI/PS approaches. The main difference between the active-space EA and IP SAC-CI and EOMCC methods pursued in this work and the corresponding SAC-CI/PS approaches, including the SAC-CI(4*p*-3*h*)/PS and SAC-CI(4*h*-3*p*)/PS methods used in our calculations for CH and OH, in which the 4*p*-3*h*/4*h*-3*p* components are selected numerically based on the suitable perturbative analysis exploiting the CI(3*p*-2*h*)/CI(3*h*-2*p*) reference states, is that in the active-space methods we use the *a priori* selected active orbitals which reflect on the electron attachment or ionization process when going from the *N*-electron reference system to the $(N\pm1)$-electron radical of interest, whereas in the SAC-CI/PS methods we let the numerical thresholds decide which 4*p*-3*h*/4*h*-3*p* amplitudes are kept in the calculations and which are removed.

The active-space EA and IP EOMCC and SAC-CI methods, including the active-space SAC-CI(4*p*-3*h*) and SAC-CI(4*h*-3*p*) methods developed in this work, offer considerable savings in the computer effort when compared to the corresponding parent approximations. The key question is if the significant cost reduction offered by the active-space methods is not done at the expense of sacrificing the accuracy of the parent EA and IP EOMCC and SAC-CI calculations. This and related questions are addressed in the next section.

III. NUMERICAL EXAMPLES AND DISCUSSION: POTENTIAL ENERGY CURVES OF THE LOW-LYING ELECTRONIC STATES OF CH AND OH

In order to examine the significance of the 4*p*-3*h*/4*h*-3*p* excitations in the EA/IP SAC-CI and, potentially, EOMCC calculations for ground- and excited-state potential energy surfaces of radical species along bond breaking coordinates and to demonstrate the effectiveness of the active-space EA/IP SAC-CI and EOMCC methods in recovering the dominant 3*p*-2*h*/3*h*-2*p* and 4*p*-3*h*/4*h*-3*p* contributions that

are needed to achieve high accuracies, we present the results of various SAC-CI and EOMCC calculations for the low-lying doublet states of the CH radical, as obtained with the $[5s3p1d/3s1p]$ basis set of Ref. 125, and the low-lying doublet states of the OH radical, as computed with the 6–31G(d,p) basis set.[126,127] In each case, we considered a wide range of nuclear geometries, including larger internuclear separations from the stretched-bond and asymptotic regions and several geometries from the region of the minimum on the corresponding ground-state potential energy curve.

The ground-state RHF orbitals of the closed-shell reference systems (CH^+ for CH and OH^- for OH) were employed throughout and the spherical components of the d functions were used in the case of the 6–31G(d,p) basis set of the OH/OH^- system. In all correlated (SAC, CC, SAC-CI, EOMCC, and CI) calculations, the lowest-energy orbital of CH^+ (the CH case) and the lowest-energy orbital of OH^- (the OH case) were kept frozen. The $[5s3p1d/3s1p]$ basis set of Ref. 125 used in the calculations for CH and the 6–31G(d,p) basis set used in the calculations for OH are small enough to enable the exact, full CI calculations, which we performed using the GAMESS package.[128] The availability of the full CI results, which are provided in the supplementary material,[129] enables us to assess the accuracy of various EA and IP SAC-CI and EOMCC approaches. By having access to full CI wave functions of the CH and OH radicals, we can also analyze the relationship between the performance of various EA/IP SAC-CI and EOMCC methods and the significance of the $1p/1h$, $2p$-$1h/2h$-$1p$, $3p$-$2h/3h$-$2p$, etc., contributions to these wave functions. All calculations for CH and OH reported in this work were based on exploiting the C_{2v} symmetry.

All of the EA and IP SAC-CI calculations, including the SAC-CI$(3p$-$2h)$, SAC-CI$(3h$-$2p)$, SAC-CI$(4p$-$3h)$, and SAC-CI$(4h$-$3p)$ calculations based on the wave function *Ansätze* defined by Eqs. (31), (32), (35), and (36), respectively, the corresponding SAC-CI$(4p$-$3h)$/PS and SAC-CI$(4h$-$3p)$/PS calculations, in which the smallest $4p$-$3h$ and $4h$-$3p$ contributions in the $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$ operators that do not significantly perturb the zero-order states obtained in the CI$(3p$-$2h)$ and CI$(3h$-$2p)$ calculations are ignored, and the underlying SAC-SD calculations for the closed-shell reference systems (CH^+ for CH and OH^- for OH) were performed using the routines developed at Kyoto University that form part of GAUSSIAN 03.[119] The threshold λ_e for selection of the $4p$-$3h$ and $4h$-$3p$ components of the $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$ operators in the SAC-CI$(4p$-$3h)$/PS and SAC-CI$(4h$-$3p)$/PS calculations was set at 10^{-7} hartree. The active-space SAC-CI$(3p$-$2h)$, SAC-CI$(3h$-$2p)$, SAC-CI$(4p$-$3h)$, and SAC-CI$(4h$-$3p)$ calculations were carried out using the new computer programs developed in this work, which were obtained by properly imposing the selection schemes discussed in detail in Sec. II B on the lists of spin- and symmetry-adapted CSFs corresponding to the $|\Phi_{jk}^{abc}\rangle$ and $|\Phi_{jkl}^{abcd}\rangle$ (the EA case) and $|\Phi_{ijk}^{bc}\rangle$ and $|\Phi_{ijkl}^{bcd}\rangle$ (the IP case) determinants that define the corresponding Hamiltonian eigenvalue problems. These programs are very efficient and fully utilize the advantages of the active-space approaches, primarily by reducing the dimensions of the Hamiltonian matrices to be evaluated in the EA and IP SAC-CI calculations, as discussed in detail in Sec. II B (cf. Tables I and II). Once the final lists of spin- and symmetry-adapted CSFs that correspond to the $|\Phi^a\rangle$, $|\Phi_j^{ab}\rangle$, and $|\Phi_{jk}^{\mathbf{A}bc}\rangle$ or $|\Phi_i\rangle$, $|\Phi_{ij}^b\rangle$, and $|\Phi_{\mathbf{I}jk}^{bc}\rangle$ or $|\Phi^a\rangle$, $|\Phi_j^{ab}\rangle$, $|\Phi_{jk}^{\mathbf{A}bc}\rangle$, and $|\Phi_{\mathbf{J}kl}^{\mathbf{AB}cd}\rangle$, or $|\Phi_i\rangle$, $|\Phi_{ij}^b\rangle$, $|\Phi_{\mathbf{I}jk}^{bc}\rangle$, and $|\Phi_{\mathbf{IJ}kl}^{\mathbf{B}cd}\rangle$ determinants are constructed, our programs proceed to the calculation of the nonzero Hamiltonian matrix elements that enter the eigenvalue problem defining the active-space SAC-CI$(3p$-$2h)\{N_u\}$, SAC-CI$(3h$-$2p)\{N_o\}$, SAC-CI$(4p$-$3h)\{N_o,N_u\}$, or SAC-CI$(4h$-$3p)\{N_o,N_u\}$ approach of interest. The full and active-space EA-EOMCCSD$(3p$-$2h)$ and IP-EOMCCSD$(3h$-$2p)$ calculations and the corresponding EA-EOMCCSD$(2p$-$1h)$ and IP-EOMCCSD$(2h$-$1p)$ calculations were performed using the highly efficient, vectorized computer codes employing the suitably defined recursively generated intermediates and fast matrix multiplication routines, which were described elsewhere.[34,35,88] In particular, for the details of the EA-EOMCCSD$(3p$-$2h)\{N_u\}$ and IP-EOMCCSD$(3h$-$2p)\{N_o\}$ computer programs used in this work and the examination of their efficiency versus parent EA-EOMCCSD$(3p$-$2h)$ and IP-EOMCCSD$(3h$-$2p)$ calculations, we refer the reader to Ref. 88. The EA and IP EOMCC codes and the CCSD program used in this work have been interfaced with the RHF and integral routines in GAMESS.[128]

A. The CH radical

Based on the model of the CH radical discussed in Sec. II B (cf. Fig. 1), in which the low-lying electronic states of CH are formed by the process of adding an electron to the 1π or 4σ unoccupied valence orbitals of CH^+, which can be combined with the electronic excitations from the highest occupied 3σ orbital that belongs to the same valence shell as the 1π or 4σ orbitals, we have chosen the 3σ, $1\pi_x$, $1\pi_y$, and 4σ valence orbitals of CH^+ that correlate with the $2p$ shell of C and $1s$ shell of H as active orbitals for the SAC-CI$(3p$-$2h)\{N_u\}$, EA-EOMCCSD$(3p$-$2h)\{N_u\}$, and SAC-CI$(4p$-$3h)\{N_o,N_u\}$ calculations discussed in this subsection. Thus, our choice of N_o and N_u for the active-space EA SAC-CI and EOMCC calculations reported in this work is $N_o=1$ and $N_u=3$. Before discussing the results of the active-space SAC-CI$(3p$-$2h)\{3\}$, EA-EOMCCSD$(3p$-$2h)\{3\}$, and SAC-CI$(4p$-$3h)\{1,3\}$ calculations, which are compared with the exact, full CI results, and the results obtained with the EA-EOMCCSD$(2p$-$1h)$, SAC-CI$(3p$-$2h)$, EA-EOMCCSD$(3p$-$2h)$, SAC-CI$(4p$-$3h)$, SAC-CI$(4p$-$3h)$/PS, and CI$(4p$-$3h)$ approaches at a few C–H distances in Table III, and at a larger number of C–H distances in the supplementary material[129] and Fig. 2, we analyze the relationship between the significance of the $1p$, $2p$-$1h$, and $3p$-$2h$ contributions to the full CI wave functions representing the low-lying doublet states of CH (shown in Table IV) and the performance of the regular EA SAC-CI and EOMCC schemes in which all orbitals are active. We focus on the $X\,^2\Pi$ ground state and the $A\,^2\Delta$, $B\,^2\Sigma^-$, and $C\,^2\Sigma^+$ excited states. The $X\,^2\Pi$ and $B\,^2\Sigma^-$ states correlate with the lowest-energy $C(^3P)+H(^2S)$ asymptote, whereas the $A\,^2\Delta$ and

TABLE III. The differences between various EA EOMCC, SAC-CI, and CI(4*p*-3*h*) energies and the corresponding full CI energies (Ref. 129) (in mhartree) at representative internuclear separations R_{C-H} (in Å), and the associated MUE and NPE values (also in mhartree) for the low-lying doublet states of the CH radical, as described by the [5*s*3*p*1*d*/3*s*1*p*] basis set of Ref. 125.

		R_{C-H} (Å)										
State	Method	0.75	1.119 786[a]	1.30	1.50	1.75	2.00	2.50	3.00	4.00	MUE	NPE
$X\ ^2\Pi$	EA-EOMCCSD(2*p*-1*h*)	0.561	1.094	1.905	3.576	7.160	12.539	25.450	32.192	27.295	32.192	31.631
	EA-EOMCCSD(3*p*-2*h*)	0.818	1.012	1.195	1.520	2.184	3.186	5.471	6.539	6.526	6.539	5.721
	EA-EOMCCSD(3*p*-2*h*){3}[b]	1.769	2.395	2.641	2.594	2.935	3.763	5.834	6.822	6.798	6.822	5.053
	SAC-CI(3*p*-2*h*)	−0.236	−0.221	−0.212	−0.198	−0.152	0.009	0.785	1.526	2.227	2.227	2.463
	SAC-CI(3*p*-2*h*){3}[b]	0.679	1.123	1.199	0.848	0.567	0.552	1.123	1.787	2.464	2.464	1.912
	SAC-CI(4*p*-3*h*)	−1.030	−1.155	−1.276	−1.502	−1.969	−2.615	−3.834	−4.462	−4.946	4.946	3.916
	SAC-CI(4*p*-3*h*){1,3}[c]	0.011	0.425	0.417	−0.201	−1.012	−1.768	−2.896	−3.381	−3.815	3.815	4.293
	SAC-CI(4*p*-3*h*)/PS	0.170	0.408	0.222	0.023	−0.668	−0.913	−1.891	−2.251	−1.975	2.251	2.659
	CI(4*p*-3*h*)	1.093	1.237	1.373	1.630	2.175	2.977	4.749	5.778	6.474	6.474	5.381
$A\ ^2\Delta$	EA-EOMCCSD(2*p*-1*h*)	57.722	68.792	79.721	98.442	131.490	168.105	224.670	254.739	278.327	278.327	220.605
	EA-EOMCCSD(3*p*-2*h*)	1.260	1.676	2.591	5.061	10.780	17.315	25.699	28.706	29.483	29.483	28.223
	EA-EOMCCSD(3*p*-2*h*){3}[b]	1.733	2.244	3.230	5.646	11.213	17.625	25.891	28.895	29.772	29.772	28.039
	SAC-CI(3*p*-2*h*)	1.292	1.866	2.925	5.596	11.571	18.367	27.572	31.623	33.800	33.800	32.508
	SAC-CI(3*p*-2*h*){3}[b]	1.764	2.431	3.564	6.185	12.009	18.682	27.763	31.782	33.956	33.956	32.192
	SAC-CI(4*p*-3*h*)	0.038	0.110	0.180	0.309	0.525	0.727	1.024	1.215	1.381	1.381	1.343
	SAC-CI(4*p*-3*h*){1,3}[c]	0.990	1.376	1.904	2.026	1.771	1.670	1.874	2.197	2.672	2.672	1.682
	SAC-CI(4*p*-3*h*)/PS	1.100	1.185	1.498	1.518	2.481	2.282	2.446	1.913	1.874	2.481	1.381
	CI(4*p*-3*h*)	0.469	0.713	0.990	1.577	2.804	4.186	6.233	7.364	8.234	8.234	7.765
$B\ ^2\Sigma^-$	EA-EOMCCSD(2*p*-1*h*)	82.720	99.558	116.599	143.887	185.074	224.236	278.423	304.581	323.029	323.029	240.309
	EA-EOMCCSD(3*p*-2*h*)	2.679	4.225	6.399	10.506	17.021	22.849	29.843	32.502	33.358	33.358	30.679
	EA-EOMCCSD(3*p*-2*h*){3}[b]	3.284	4.873	7.093	11.097	17.433	23.163	30.105	32.803	33.821	33.821	30.537
	SAC-CI(3*p*-2*h*)	2.740	4.485	6.827	11.141	17.881	23.928	31.574	35.013	36.910	36.910	34.170
	SAC-CI(3*p*-2*h*){3}[b]	3.337	5.130	7.523	11.738	18.297	24.240	31.802	35.216	37.105	37.105	33.768
	SAC-CI(4*p*-3*h*)	−0.003	0.094	0.207	0.392	0.622	0.788	1.006	1.146	1.286	1.286	1.289
	SAC-CI(4*p*-3*h*){1,3}[c]	1.999	2.964	3.578	3.185	2.507	2.268	2.321	2.576	3.087	3.578	1.579
	SAC-CI(4*p*-3*h*)/PS	1.152	1.160	1.574	1.797	2.064	2.885	2.834	3.069	3.713	3.713	2.562
	CI(4*p*-3*h*)	0.841	1.129	1.503	2.227	3.406	4.504	6.054	6.943	7.683	7.683	6.842
$C\ ^2\Sigma^+$	EA-EOMCCSD(2*p*-1*h*)	46.283	57.533	67.438	89.126	20.725	18.885	21.429	23.682	26.010	89.126	70.241
	EA-EOMCCSD(3*p*-2*h*)	2.480	2.771	3.436	5.185	5.691	4.706	3.411	1.744	−1.175	5.691	6.866
	EA-EOMCCSD(3*p*-2*h*){3}[b]	3.186	3.737	4.584	6.189	6.203	5.054	3.656	1.964	−0.930	6.203	7.133
	SAC-CI(3*p*-2*h*)	2.489	2.951	3.753	5.488	4.796	3.592	2.976	2.768	2.856	5.488	2.999
	SAC-CI(3*p*-2*h*){3}[b]	3.207	3.916	4.901	6.496	5.301	3.940	3.208	2.948	3.002	6.496	3.548
	SAC-CI(4*p*-3*h*)	0.218	0.442	0.671	0.866	−0.100	−0.586	−0.636	−0.500	−0.223	0.866	1.527
	SAC-CI(4*p*-3*h*){1,3}[c]	1.146	1.819	2.622	2.788	0.980	0.180	−0.039	0.080	0.416	2.909	2.948
	SAC-CI(4*p*-3*h*)/PS	1.829	2.376	3.430	3.465	1.347	0.606	0.156	0.157	0.199	3.601	3.445
	CI(4*p*-3*h*)	0.948	1.887	3.202	6.176	8.313	7.902	8.567	9.076	9.497	9.497	8.549

[a]Equilibrium geometry taken from Ref. 130.
[b]The active space consisted of the $1\pi_x$, $1\pi_y$, and 4σ orbitals of CH^+.
[c]The active space consisted of the 3σ, $1\pi_x$, $1\pi_y$, and 4σ orbitals of CH^+.

$C\ ^2\Sigma^+$ states correlate with the $C(^1D)+H(^2S)$ asymptote [see Fig. 2(a), which shows the full CI curves]. One of the interesting features of the low-lying states of CH is the crossing between the potential energy curves representing the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ states in the Franck-Condon region. It is interesting to examine what are the minimum levels of the EA SAC-CI and EOMCC theories that can properly describe various features of the full CI curves shown in Fig. 2(a), including, among others, the crossing of the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ curves in the Franck-Condon region and the asymptotic degeneracies of the $X\ ^2\Pi$ and $B\ ^2\Sigma^-$ as well as $A\ ^2\Delta$ and $C\ ^2\Sigma^+$ states.

The results in Table III indicate that the basic level of the EA EOMCC or SAC-CI theory, in which the electron attaching operator $R_\mu^{(N+1)}$ is truncated at the 2*p*-1*h* excitations, is practically useless. In the case of excited states, the differences between the EA-EOMCCSD(2*p*-1*h*) and full CI energies in the R_{C-H}=0.75–4.0 Å region (R_{C-H} is the C–H separation and R_{C-H}=1.119 786 Å is the experimental equilibrium bond length in CH taken from Ref. 130) are 57.722–278.327 mhartree for the $A\ ^2\Delta$ state, 82.720–323.029 mhartree for the $B\ ^2\Sigma^-$ state, and 18.885–89.126 mhartree for the $C\ ^2\Sigma^+$ state. The analogous SAC-CI(2*p*-1*h*) results (not shown) are equally poor. As shown in Fig. 2(b), the EA-EOMCCSD(2*p*-1*h*) curves for the $A\ ^2\Delta$, $B\ ^2\Sigma^-$, and $C\ ^2\Sigma^+$ states do not resemble the corre-

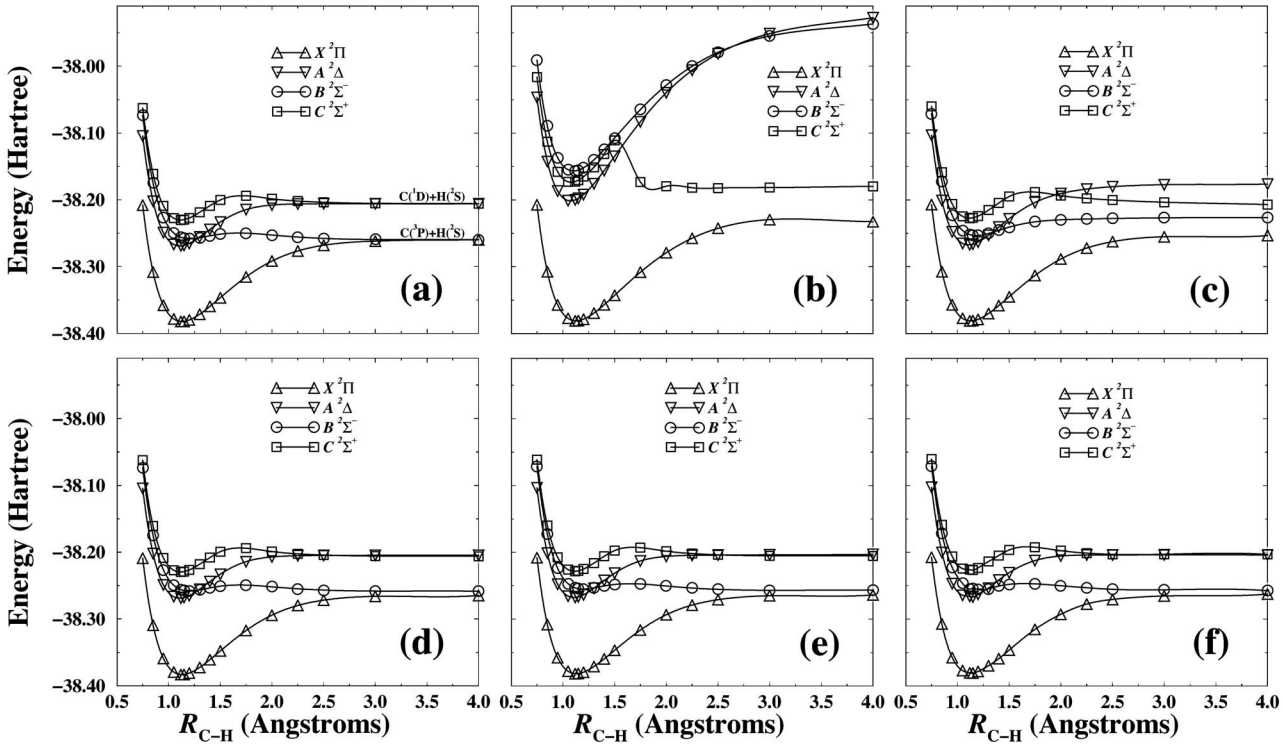

FIG. 2. Potential energy curves for the ground and low-lying doublet excited states of the CH radical, as described by the [5*s*3*p*1*d*/3*s*1*p*] basis set of Ref. 125. Energies are in hartree and the C–H distance R_{C-H} is in Å. (a) The full CI results, (b) the EA-EOMCCSD(2*p*-1*h*) results, (c) the EA-EOMCCSD(3*p*-2*h*) results, (d) the SAC-CI(4*p*-3*h*) results, (e) the SAC-CI(4*p*-3*h*){1,3} results, and (f) the SAC-CI(4*p*-3*h*)/PS results.

sponding full CI curves [shown in Fig. 2(a)]. The EA-EOMCCSD(2*p*-1*h*) curves are characterized by large maximum unsigned errors (MUEs) and large nonparallelity errors (NPE values; NPE is defined as the difference between the most positive and most negative signed errors relative to full CI along a given potential energy curve). They are also qualitatively wrong. In particular, the EA-EOMCCSD(2*p*-1*h*) approach produces an incorrect ordering of excited states and cannot describe the crossing of the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ states at $R_{C-H} \approx 1.3$ Å. The situation only worsens as R_{C-H} approaches the asymptotic region.

The poor performance of the EA-EOMCCSD(2*p*-1*h*) [or SAC-CI(2*p*-1*h*)] approach in the calculations of excited states of CH can be understood if we realize that the $A\ ^2\Delta$, $B\ ^2\Sigma^-$, and $C\ ^2\Sigma^+$ states have significant 2*p*-1*h* components at all C–H distances (see Table IV). They also gain substan-

TABLE IV. An analysis of the major full CI configurations (all configurations with a coefficient ⩾0.15 for at least one of the selected values of R_{C-H} are included) for the low-lying doublet states of the CH radical, as described by the [5*s*3*p*1*d*/3*s*1*p*] basis set of Ref. 125.

		Coefficients for various values of R_{C-H}				
State	Configuration orbital occupancy	1.119 786 Å[a]	1.50 Å	2.00 Å	3.00 Å	Excitation type[b]
$X\ ^2\Pi$	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^2(1\pi_x)^1\|$	0.942	0.913	0.828	0.573	1*p*
	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1(1\pi_x)^1(4\sigma)^1\|$[c]	<0.01, 0.029	−0.042, 0.212	−0.171, 0.370	−0.439, 0.479	2*p*-1*h*
	$\|(1\sigma)^2(2\sigma)^2(1\pi_x)^1(4\sigma)^2\|$	<0.01	−0.088	−0.245	−0.408	3*p*-2*h*
$A\ ^2\Delta$	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1[(1\pi_x)^2-(1\pi_y)^2]\|$	−0.681	−0.648	−0.534	0.362	2*p*-1*h*
	$\|(1\sigma)^2(2\sigma)^2[(1\pi_x)^2-(1\pi_y)^2](4\sigma)^1\|$	0.014	0.163	0.415	−0.569	3*p*-2*h*
$B\ ^2\Sigma^-$	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1(1\pi_x)^1(1\pi_y)^1\|$[c]	0.472, 0.818	0.440, 0.763	0.357, 0.619	−0.250, −0.432	2*p*-1*h*
	$\|(1\sigma)^2(2\sigma)^2(1\pi_x)^1(1\pi_y)^1(4\sigma)^1\|$[c]	<0.01, −0.013	0.144, −0.249	0.310, −0.536	−0.401, 0.694	3*p*-2*h*
$C\ ^2\Sigma^+$	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^2(4\sigma)^1\|$	−0.033	0.166	0.710	0.730	1*p*
	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1(4\sigma)^2\|$	0.015	0.053	0.161	0.295	2*p*-1*h*
	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1[(1\pi_x)^2+(1\pi_y)^2]\|$	0.675	0.628	0.334	0.209	2*p*-1*h*
	$\|(1\sigma)^2(2\sigma)^2[(1\pi_x)^2+(1\pi_y)^2](4\sigma)^1\|$	−0.011	−0.174	−0.288	−0.334	3*p*-2*h*

[a]Equilibrium bond length taken from Ref. 130.
[b]Relative to the ground-state reference configuration of CH^+, $|1\sigma^2 2\sigma^2 3\sigma^2|$.
[c]The two different coefficients shown are for the two doublet configuration state functions, each corresponding to a different intermediate spin state, that result from coupling the spins of the three unpaired electrons in this orbital occupation scheme.

tial $3p$-$2h$ components, when $R_{C-H} > 1.5$ Å. This is particularly true for the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ states (for the $C\ ^2\Sigma^+$ state, the increase of the $3p$-$2h$ components is less pronounced). The EA-EOMCCSD($2p$-$1h$) and SAC-CI($2p$-$1h$) approaches can provide reasonable results only when the electronic states of interest are of an almost pure $1p$ character. As shown in Table IV, the $A\ ^2\Delta$, $B\ ^2\Sigma^-$, and $C\ ^2\Sigma^+$ states do not satisfy this requirement and this leads to huge errors in the EA-EOMCCSD($2p$-$1h$) results for these states. The only state for which the EA-EOMCCSD($2p$-$1h$) approach can find some limited use is the ground state. The ground state of CH is dominated by the $1p$ excitations in the $R_{C-H} < 2.0$ Å region and this correlates quite well with the relatively small errors in the EA-EOMCCSD($2p$-$1h$) results for the $X\ ^2\Pi$ state in this region.

As shown in Table III and Fig. 2(c), the explicit inclusion of the $3p$-$2h$ excitations in the electron attaching operator $R_\mu^{(N+1)}$ leads to substantial improvements in the poor EA-EOMCCSD($2p$-$1h$) and SAC-CI($2p$-$1h$) results. These improvements are particularly impressive when $R_{C-H} \leqslant 1.5$ Å. Indeed, the 57.722–98.442, 82.720–143.887, and 46.283–89.126 mhartree errors in the EA-EOMCCSD($2p$-$1h$) results for the $A\ ^2\Delta$, $B\ ^2\Sigma^-$, and $C\ ^2\Sigma^+$ states, respectively, in the $R_{C-H} \leqslant 1.5$ Å region reduce to 1.260-5.061, 2.679-10.506, and 2.480–5.185 mhartree, respectively, when the EA-EOMCCSD($3p$-$2h$) approach is employed, and 1.292-5.596, 2.740-11.141, and 2.489–5.488 mhartree, respectively, when the SAC-CI($3p$-$2h$) method is used. The EA-EOMCCSD($3p$-$2h$) and SAC-CI($3p$-$2h$) results for the ground state are virtually perfect in the $R_{C-H} \leqslant 1.5$ Å region as well. For example, there is practically no difference between the EA-EOMCCSD($3p$-$2h$) and full CI potential energy curves in the $R_{C-H} \leqslant 1.5$ Å region and the EA-EOMCCSD($3p$-$2h$) approach restores the crossing of the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ states at $R_{C-H} \approx 1.3$ Å, which the EA-EOMCCSD($2p$-$1h$) approach could not describe [cf. Figs. 2(a) and 2(c)].

The fact that the EA-EOMCCSD($3p$-$2h$) and SAC-CI($3p$-$2h$) approaches work so well in the $R_{C-H} \leqslant 1.5$ Å region is a consequence of the absence of the significant higher-than-$2p$-$1h$ contributions in the $X\ ^2\Pi$, $A\ ^2\Delta$, $B\ ^2\Sigma^-$, and $C\ ^2\Sigma^+$ states in this region. This is particularly true in the vicinity of the minimum on the ground-state potential energy curve. For the $X\ ^2\Pi$, $A\ ^2\Delta$, and $C\ ^2\Sigma^+$ states, the $3p$-$2h$ and higher-than-$3p$-$2h$ contributions remain relatively small in the entire $R_{C-H} \leqslant 1.5$ Å region. For the $B\ ^2\Sigma^-$ state, they begin to grow as R_{C-H} approaches 1.5 Å and this results in the increase of the errors characterizing the EA-EOMCCSD($3p$-$2h$) and SAC-CI($3p$-$2h$) results for this state at $R_{C-H} \approx 1.5$ Å.

The EA-EOMCCSD($3p$-$2h$) and SAC-CI($3p$-$2h$) approaches provide considerable improvements in the poor EA-EOMCCSD($2p$-$1h$) [and SAC-CI($2p$-$1h$)] results in the $R_{C-H} > 1.5$ Å region as well. For example, the EA-EOMCCSD($3p$-$2h$) approach reduces the 131.490–278.327 and 185.074–323.029 mhartree errors in the EA-EOMCCSD($2p$-$1h$) results for the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ states in the $R_{C-H} = 1.75$–4.0 Å region down to 10.780–29.483 and 17.021–33.358 mhartree, respectively. SAC-CI($3p$-$2h$) provides very similar improvements. In fact, since the $3p$-$2h$ contributions to the $X\ ^2\Pi$ and $C\ ^2\Sigma^+$ states do not grow as fast as in the case of the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ states and since the $X\ ^2\Pi$ and $C\ ^2\Sigma^+$ states remain largely dominated by the $1p$ and $2p$-$1h$ excitations for almost all values of R_{C-H}, the EA-EOMCCSD($3p$-$2h$) and SAC-CI($3p$-$2h$) approaches provide a very good description of the entire $X\ ^2\Pi$ and $C\ ^2\Sigma^+$ potential energy curves, with the MUE and NPE values ranging between 2 and 7 mhartree in the entire 0.75 Å $\leqslant R_{C-H} \leqslant 4.0$ Å region. This can also be seen by comparing the $X\ ^2\Pi$ and $C\ ^2\Sigma^+$ potentials obtained with EA-EOMCCSD($3p$-$2h$) [shown in Fig. 2(c)] with the corresponding full CI curves [shown in Fig. 2(a)].

In spite of the considerable improvements in the results offered by the EA-EOMCCSD($3p$-$2h$) and SAC-CI($3p$-$2h$) approaches, one cannot rely on these approaches in quantitative calculations of the entire potential energy curves. The 17.315–29.483 and 18.367–33.800 mhartree errors resulting from the EA-EOMCCSD($3p$-$2h$) and SAC-CI($3p$-$2h$) calculations for the $A\ ^2\Delta$ state in the $R_{C-H} \geqslant 2.0$ Å region and the 22.849–33.358 and 23.928–36.910 mhartree errors obtained with the EA-EOMCCSD($3p$-$2h$) and SAC-CI($3p$-$2h$) methods for the $B\ ^2\Sigma^-$ state in the same region are too large to be acceptable in high accuracy calculations. The relatively good description of the entire $X\ ^2\Pi$ and $C\ ^2\Sigma^+$ curves and the relatively poor description of the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ curves at larger C–H distances by the EA-EOMCCSD($3p$-$2h$) and SAC-CI($3p$-$2h$) approaches result in a significant breakdown of the asymptotic degeneracies of the electronic states of CH shown in Fig. 2(a). For example, the $X\ ^2\Pi$ and $B\ ^2\Sigma^-$ states should become degenerate as $R_{C-H} \to \infty$, but they are far from being degenerate at larger C–H distances when the EA-EOMCCSD($3p$-$2h$) method is employed, since the $B\ ^2\Sigma^-$ state dissociates incorrectly in the EA-EOMCCSD($3p$-$2h$) calculations [see Fig. 2(c)]. A similar remark is true for the $A\ ^2\Delta$ and $C\ ^2\Sigma^+$ states, which should become degenerate as $R_{C-H} \to \infty$, but remain separated by a rather large energy gap at larger C–H distances in the EA-EOMCCSD($3p$-$2h$) calculations, since only one of the two states ($C\ ^2\Sigma^+$) is reasonably well described by the EA-EOMCCSD($3p$-$2h$) approach.

The above analysis demonstrates that one has to go beyond the EA-EOMCCSD($3p$-$2h$) and SAC-CI($3p$-$2h$) approximations to obtain a uniformly accurate description of the entire potential energy curves representing the low-lying states of CH. This is a consequence of the presence of the relatively large $3p$-$2h$ contributions to the electronic states of CH (particularly, the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ states) at larger internuclear separations. If we did not have access to the EA-EOMCCSD($3p$-$2h$) data and had only to rely on the results obtained with the quasilinearized forms of the SAC-CI($3p$-$2h$) wave functions used in this paper, we might speculate that larger errors in the results for the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ states at larger C–H distances are a consequence of ignoring the nonlinear terms in the cluster operator S in Eq. (31) defining the SAC-CI($3p$-$2h$) states rather than due to the neglect of higher-than-$3p$-$2h$ components in the electron attaching operator $R_\mu^{(N+1)}$. We might even speculate that perhaps the nonlinear terms in S neglected in the SAC-SD cal-

TABLE V. The differences between the CCSD, CCSDT, CR-CC(2,3), and SAC-SD ground-state energies of the CH^+ ion, as described by the [5*s*3*p*1*d*/3*s*1*p*] basis set of Ref. 125, and the corresponding full CI ground-state energies (Ref. 129) (in mhartree) at representative internuclear separations R_{C-H} (in Å), and the associated MUE and NPE values relative to full CI (also in mhartree).

	R_{C-H} (Å)										
Method	0.75	1.119 786[a]	1.30	1.50	1.75	2.00	2.50	3.00	4.00	MUE	NPE
CCSD	1.622	1.813	2.023	2.390	3.107	4.123	6.639	8.602	9.703	9.703	8.081
CCSDT	0.110	0.116	0.136	0.179	0.275	0.414	0.706	0.719	0.172	0.719	0.610
CR-CC(2,3)	0.156	0.170	0.211	0.249	0.322	0.441	0.707	0.711	0.168	0.711	0.558
SAC-SD	1.648	1.857	2.099	2.551	3.543	5.221	11.138	18.981	29.896	29.896	28.248

[a]The equilibrium bond length of CH taken from Ref. 130.

culations for the CH^+ reference system are responsible for the inaccuracies observed in the SAC-CI(3*p*-2*h*) calculations at larger C–H distances. Fortunately, we have access to the full CCSD and EA-EOMCCSD(3*p*-2*h*) results, in addition to the SAC-SD and SAC-CI(3*p*-2*h*) data, so that we do not have to speculate about these issues. As shown in Table III, there is virtually no difference between the results obtained with the quasilinearized form of SAC-CI(3*p*-2*h*) used in this paper and the results of the EA-EOMCCSD(3*p*-2*h*) calculations, which do not neglect any nonlinear terms in S that enter the EA-EOMCCSD(3*p*-2*h*)/SAC-CI(3*p*-2*h*) eigenvalue problem. Moreover, as shown in Table V, there is almost no difference in the results of the SAC-SD calculations for the ground-state potential energy curve of CH^+, in which all nonlinear terms other than $\frac{1}{2}S_2^2$ are neglected, and the analogous results obtained with the full CCSD approach, in which all nonlinear terms in S_1 and S_2 are kept. The SAC-SD and CCSD curves begin to deviate only in the $R_{C-H} \geqslant 2.5$ Å region. In any case, since we use the full CCSD approach to produce the correlated reference states for the EA-EOMCCSD(3*p*-2*h*) calculations and since we do not neglect any relevant nonlinear terms in S in the EA-EOMCCSD(3*p*-2*h*) approach, and yet the EA-EOMCCSD(3*p*-2*h*) results for the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ states at larger C–H separations are rather poor, we believe that the primary reason for the observed failures of the SAC-CI(3*p*-2*h*) and EA-EOMCCSD(3*p*-2*h*) approximations in the region of larger R_{C-H} values is the absence of the 4*p*-3*h* excitations in the electron attaching operator $R_\mu^{(N+1)}$ defining the EA SAC-CI and EOMCC theories.

The importance of the 4*p*-3*h* components of $R_\mu^{(N+1)}$ at larger C–H distances can be seen if we analyze the CI(4*p*-3*h*) data in Table III. The description of the $X\ ^2\Pi$, $A\ ^2\Delta$, $B\ ^2\Sigma^-$, and $C\ ^2\Sigma^+$ potential energy curves of CH by the CI(4*p*-3*h*) approach, in which the electron attaching operator $R_\mu^{(N+1)}$ truncated at 4*p*-3*h* excitations is directly applied to the reference determinant $|\Phi\rangle$, is not perfect and the maximum errors in the CI(4*p*-3*h*) results range from 6.474 to 9.497 mhartree, but we already observe substantial improvements in the results for the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ states, compared to the SAC-CI(3*p*-2*h*) and EA-EOMCCSD(3*p*-2*h*) calculations (MUE reduction by a factor of 4), when the CI(4*p*-3*h*) method is employed.

The most accurate results are obtained when we apply the electron attaching operator $R_\mu^{(N+1)}$ truncated at 4*p*-3*h* excitations to the properly correlated ground state of CH^+ rather than to the uncorrelated, single-determinantal state $|\Phi\rangle$. This is clearly seen when we look at the SAC-CI(4*p*-3*h*) results in Table III and when we compare the potential energy curves resulting from the SAC-CI(4*p*-3*h*) calculations shown in Fig. 2(d) with the corresponding full CI curves shown in Fig. 2(a). The 29.483, 33.800, and 8.234 mhartree maximum errors in the EA-EOMCCSD(3*p*-2*h*), SAC-CI(3*p*-2*h*), and CI(4*p*-3*h*) results for the $A\ ^2\Delta$ state, the maximum errors of 33.358, 36.910, and 7.683 mhartree characterizing the EA-EOMCCSD(3*p*-2*h*), SAC-CI(3*p*-2*h*), and CI(4*p*-3*h*) calculations for the $B\ ^2\Sigma^-$ state, and the 5.691, 5.488, and 9.497 mhartree maximum errors characterizing the EA-EOMCCSD(3*p*-2*h*), SAC-CI(3*p*-2*h*), and CI(4*p*-3*h*) results for the $C\ ^2\Sigma^+$ state reduce to 1.381, 1.286, and 0.866 mhartree, respectively, when the SAC-CI(4*p*-3*h*) approach exploiting the quasilinearized form of the wave function [Eq. (35)] is employed. Based on the above discussion of the relative performance of EA SAC-CI, EOMCC, and CI(4*p*-3*h*) methods and based on the remarks made in Sec. II A, one of the best solutions that would, most likely, enable us to obtain the virtually exact description of the entire potential energy curves of the ground and excited states of CH should originate from the EA-EOMCCSDT(4*p*-3*h*) calculations, which would employ the CCSDT ground state of CH^+ and the $R_\mu^{(N+1)}$ operator truncated at the 4*p*-3*h* excitations. Unfortunately, the EA-EOMCCSDT(4*p*-3*h*) approach has not been implemented yet [the automated implementation of the related EA-EOMCCSDTQ(4*p*-3*h*)=EA-EOMCCSDTQ approximation that uses the CCSDTQ rather than the CCSDT wave function as a reference ground state in the EA EOMCC calculations truncated at 4*p*-3*h* excitations has recently been reported,[47] but we have no access to this implementation]. As pointed out in Sec. II A, the EA-EOMCCSDT(4*p*-3*h*) method would provide a rigorously size intensive description of the electronic excitations in CH. Moreover, as implied by the excellent full CCSDT results for CH^+ and the results obtained by the recently formulated CR-CC(2,3) approach,[131,132] which offers an approximate and yet highly accurate treatment of S_3 clusters, both shown in Table V, the explicit inclusion of S_3 clusters in the underlying calculations for CH^+ would lead to the virtually exact description of the ground state of CH^+, which serves as a reference for the EA EOMCC and SAC-CI calculations. These observations, combined with the great im-

provements in the description of the ground and excited states of CH offered by the presence of the $4p$-$3h$ excitations in the electron attaching operator $R_\mu^{(N+1)}$, make us believe that the use of the EA-EOMCCSDT($4p$-$3h$) approach would result in the virtually exact description of the entire potential energy curves of CH. On the other hand, it is encouraging to observe that the SAC-CI($4p$-$3h$) method, in which most of the nonlinear terms in S and the S_3 components are ignored and which is less expensive than EA-EOMCCSDT($4p$-$3h$), provides a highly accurate description of the entire ground- and excited-state potential energy curves of CH. We might wonder if one could obtain similarly good results with the EA-EOMCCSDT($3p$-$2h$) approach, in which S is truncated at S_3 and the $4p$-$3h$ excitations in $R_\mu^{(N+1)}$ are ignored, since CCSDT provides the perfect description of CH^+; we will examine this issue in the future, but right now we believe that one needs the $4p$-$3h$ excitations in $R_\mu^{(N+1)}$, since the CI($4p$-$3h$) calculations greatly improve the EA-EOMCCSD($3p$-$2h$) results for the most difficult $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ states.

Now that we have established that the SAC-CI($4p$-$3h$) approach provides high-quality potential curves for the low-lying states of CH, it is important to examine if one can retain the accuracy of the SAC-CI($4p$-$3h$) calculations without an inclusion of all $3p$-$2h$ amplitudes r_{abc}^{jk} and all $4p$-$3h$ amplitudes r_{abcd}^{jkl}, which are far too numerous for the majority of practical applications. One possible way to reduce large costs of the SAC-CI($4p$-$3h$) calculations without a substantial loss of accuracy is offered by the SAC-CI($4p$-$3h$)/PS approach, in which the $4p$-$3h$ components of the $R_\mu^{(N+1)}$ operators that do not significantly perturb the CI($3p$-$2h$) wave functions and whose energy contributions are smaller than the threshold λ_e (set in our calculations at 10^{-7} hartree) are eliminated from the SAC-CI($4p$-$3h$) diagonalization. As shown in Table III, the SAC-CI($4p$-$3h$)/PS results are practically as good as those provided by the SAC-CI($4p$-$3h$) approach, in spite of the fact that the total number of spin- and symmetry-adapted amplitudes defining the $R_\mu^{(N+1)}$ operator used in the SAC-CI($4p$-$3h$)/PS calculations represents only 4%-14% of all amplitudes r used in the SAC-CI($4p$-$3h$) calculations (see Table I). In addition to the high accuracy and the substantial reduction in the computer effort offered by the SAC-CI($4p$-$3h$)/PS approach, the SAC-CI($4p$-$3h$)/PS method has an advantage of being relatively easy to use, since the elimination of the small r_{abcd}^{jkl} amplitudes is completely automatic once the suitable selection threshold λ_e is chosen. There is, however, one problem with the SAC-CI($4p$-$3h$)/PS approach, namely, the numerical noise that the SAC-CI($4p$-$3h$)/PS calculations may produce due to the fact that different sets of the $4p$-$3h$ contributions to the $R_\mu^{(N+1)}$ operators are selected at different nuclear geometries. The active-space SAC-CI($4p$-$3h$) approach developed in this work has an advantage over the SAC-CI($4p$-$3h$)/PS method in the fact that it uses the same set of r amplitudes defining the $R_\mu^{(N+1)}$ operator at all nuclear geometries, producing smooth potential energy curves while offering a substantial reduction in the dimensionality of the corresponding eigenvalue problem through the selection of the $3p$-$2h$ amplitudes r_{abc}^{jk} and $4p$-$3h$ amplitudes r_{abcd}^{jkl} via active orbitals, as discussed in Sec. II B (see Table I). Yet, as shown in Table III, the active-space SAC-CI($4p$-$3h$) approach employing only one occupied and three unoccupied active orbitals, and using approximately 10% of all amplitudes r defining the $R_\mu^{(N+1)}$ operator truncated at $4p$-$3h$ excitations, is at least as effective in producing highly accurate potential energy curves of CH as the SAC-CI($4p$-$3h$)/PS method. The small maximum errors of 3.815, 2.672, 3.578, and 2.909 mhartree characterizing the SAC-CI($4p$-$3h$)$\{1,3\}$ results for the entire potential energy curves of the $X\ ^2\Pi$, $A\ ^2\Delta$, $B\ ^2\Sigma^-$, and $C\ ^2\Sigma^+$ states of CH are either similar to or only slightly larger than the 4.946, 1.381, 1.286, and 0.866 mhartree maximum errors obtained with the SAC-CI($4p$-$3h$) approach and similar to the 2.251, 2.481, 3.713, and 3.601 mhartree maximum errors characterizing the corresponding SAC-CI($4p$-$3h$)/PS results. As a result, the SAC-CI($4p$-$3h$)$\{1,3\}$ potential energy curves shown in Fig. 2(e) can hardly be distinguished from the highly accurate SAC-CI($4p$-$3h$) and SAC-CI($4p$-$3h$)/PS curves shown in Figs. 2(d) and 2(f), respectively. There is also almost no difference between the SAC-CI($4p$-$3h$)$\{1,3\}$ curves [shown in Fig. 2(e)], and the full CI curves [shown in Fig. 2(a)]. All essential features of the full CI curves, including the crossing of the $A\ ^2\Delta$ and $B\ ^2\Sigma^-$ curves in the Franck-Condon region and the asymptotic degeneracies of the $X\ ^2\Pi$ and $B\ ^2\Sigma^-$ as well as $A\ ^2\Delta$ and $C\ ^2\Sigma^+$ states of CH are accurately reproduced by the SAC-CI($4p$-$3h$)$\{1,3\}$ calculations.

As one might expect, the SAC-CI($3p$-$2h$)$\{3\}$ and EA-EOMCCSD ($3p$-$2h$)$\{3\}$ schemes are less accurate than the SAC-CI($4p$-$3h$)$\{1,3\}$ method, particularly in the region of larger C–H separations, but it is encouraging to observe that the SAC-CI($3p$-$2h$)$\{3\}$ and EA-EOMCCSD($3p$-$2h$)$\{3\}$ approaches, which reduce the dimensionality of the SAC-CI($3p$-$2h$) and EA-EOMCCSD($3p$-$2h$) eigenvalue problems by a factor of 2.6-2.7, faithfully reproduce the results of the parent SAC-CI($3p$-$2h$) and EA-EOMCCSD($3p$-$2h$) calculations at all C–H separations. The relatively small 0.679–1.200, 1.764–6.185, 3.337–11.738, and 3.207–6.496 mhartree errors in the SAC-CI($3p$-$2h$)$\{3\}$ results for the $X\ ^2\Pi$, $A\ ^2\Delta$, $B\ ^2\Sigma^-$, and $C\ ^2\Sigma^+$ states, respectively, in the $R_{\mathrm{C-H}} \leqslant 1.5$ Å region and the very similar errors of 1.769–2.641, 1.733–5.646, 3.284–11.097, and 3.186–6.189 mhartree characterizing the EA-EOMCCSD($3p$-$2h$)$\{3\}$ results for the $X\ ^2\Pi$, $A\ ^2\Delta$, $B\ ^2\Sigma^-$, and $C\ ^2\Sigma^+$ states in the same region are not much different than the 0.198–0.236, 1.292–5.596, 2.740–11.141, and 2.489–5.488 mhartree errors in the corresponding SAC-CI($3p$-$2h$) results and the 0.818–1.520, 1.260–5.061, 2.679–10.506, and 2.480–5.185 mhartree errors in the results of the EA-EOMCCSD($3p$-$2h$) calculations (see Table III and Ref. 129). As in the case of the parent SAC-CI($3p$-$2h$) and EA-EOMCCSD($3p$-$2h$) approximations, the active-space SAC-CI($3p$-$2h$)$\{3\}$ and EA-EOMCCSD($3p$-$2h$)$\{3\}$ approaches fail to provide an accurate description of the asymptotic region, but the differences between the SAC-CI($3p$-$2h$)$\{3\}$/EA-EOMCCSD($3p$-$2h$)$\{3\}$ energies and their SAC-CI($3p$-$2h$)/EA-EOMCCSD($3p$-$2h$) counterparts remain small even when $R_{\mathrm{C-H}} > 1.5$ Å. For example, the relatively large errors

in the EA-EOMCCSD($3p$-$2h$){3} results for the $A\,^2\Delta$ and $B\,^2\Sigma^-$ states in the $R_{C-H}=1.75-4.0$ Å region, where these states gain significant $3p$-$2h$ components, of 11.213–29.772 and 17.433–33.821 mhartree, respectively, are similar to the 10.780–29.483 and 17.021–33.358 mhartree errors obtained with the all-orbital EA-EOMCCSD($3p$-$2h$) approach (cf. Table III). It is virtually impossible to distinguish between the ground- and excited-state potential energy curves of the low-lying states of CH obtained in the active-space SAC-CI($3p$-$2h$){3} and EA-EOMCCSD($3p$-$2h$){3} calculations and the analogous potential energy curves resulting from the parent SAC-CI($3p$-$2h$) and EA-EOMCCSD($3p$-$2h$) calculations.

We can conclude this subsection by stating that the active-space SAC-CI($3p$-$2h$)$\{N_u\}$, EA-EOMCCSD($3p$-$2h$)$\{N_u\}$, and SAC-CI($4p$-$3h$)$\{N_o,N_u\}$ methods are as accurate as their considerably more expensive SAC-CI($3p$-$2h$), EA-EOMCCSD($3p$-$2h$), and SAC-CI($4p$-$3h$) counterparts. In particular, the active-space SAC-CI($4p$-$3h$){1,3} approach, in which one uses only four active orbitals to select the most numerous $3p$-$2h$ and $4p$-$3h$ excitations, provides an excellent description of the entire potential energy curves of the ground and low-lying excited states of the CH radical, including the Franck-Condon and asymptotic regions, at a small fraction of the computer effort involved in the SAC-CI($4p$-$3h$) calculations.

B. The OH radical

The zero-order description of the OH radical discussed in Sec. II B (see Fig. 1), in which the low-lying electronic states of OH are formed by removing an electron from the occupied 1π shell of OH^-, which can be further coupled with the electronic excitations from the 3σ and 1π occupied orbitals to the 4σ unoccupied orbital that belongs to the same valence shell as the 3σ and 1π orbitals, suggests that the most natural choice of active orbitals for the SAC-CI($3h$-$2p$)$\{N_o\}$, IP-EOMCCSD($3h$-$2p$)$\{N_o\}$, and SAC-CI($4h$-$3p$)$\{N_o,N_u\}$ calculations for OH is represented by the 3σ, $1\pi_x$, $1\pi_y$, and 4σ valence orbitals of OH^-. These orbitals correlate with the $2p$ shell of O and $1s$ shell of H. This is the choice made in the active-space SAC-CI($3h$-$2p$)$\{N_o\}$, IP-EOMCCSD($3h$-$2p$)$\{N_o\}$, and SAC-CI($4h$-$3p$)$\{N_o,N_u\}$ calculations for OH discussed in this subsection. In other words, in the active-space IP SAC-CI and EOMCC calculations performed in this study, the N_o and N_u values are set at 3 and 1, respectively. As explained in Sec. II B, with this particular choice of the active space, the number of spin- and symmetry-adapted $3h$-$2p$ amplitudes $r_{bc}^{\mathbf{I}jk}$ is the same as the number of all $3h$-$2p$ amplitudes r_{bc}^{ijk}, since there is only one correlated occupied orbital outside the active space. This immediately implies that the SAC-CI($3h$-$2p$){3} and IP-EOMCCSD($3h$-$2p$){3} calculations for OH are equivalent to the regular, all-orbital SAC-CI($3h$-$2p$) and IP-EOMCCSD($3h$-$2p$) calculations, respectively. Thus, in Table VI and Fig. 3, and in the more detailed additional tables in the supplementary material,[129] where the results of the IP SAC-CI and EOMCC calculations for OH are collected, we focus on the energies obtained in the active-space SAC-CI($4h$-$3p$){3,1} calculations, which we compare with the corresponding full CI, IP-EOMCCSD($2h$-$1p$), IP-EOMCCSD($3h$-$2p$), SAC-CI($3h$-$2p$), SAC-CI($4h$-$3p$), and SAC-CI($4h$-$3p$)/PS data.

In analogy to the CH radical, we first analyze the relationship between the importance of the $1h$, $2h$-$1p$, and $3h$-$2p$ contributions to the full CI wave functions representing the low-lying doublet states of OH (shown in Table VII) and the performance of the normal IP SAC-CI and EOMCC schemes in which all orbitals are active. We focus on the $X\,^2\Pi$ ground state and the $1\,^2\Delta$, $2\,^2\Pi$, $A\,^2\Sigma^+$, $1\,^2\Sigma^-$, and $B\,^2\Sigma^+$ excited states. As shown in Fig. 3(a), which displays the exact potential energy curves obtained in full CI calculations, the $X\,^2\Pi$ and $1\,^2\Sigma^-$ states correlate with the lowest-energy $O(^3P)+H(^2S)$ asymptote, the $1\,^2\Delta$, $2\,^2\Pi$, and $A\,^2\Sigma^+$ states correlate with the next $O(^1D)+H(^2S)$ asymptote, and the $B\,^2\Sigma^+$ state dissociates into $O(^1S)+H(^2S)$. We begin our analysis by examining the effectiveness of various IP SAC-CI and EOMCC approximations in describing the most important features of the full CI curves shown in Fig. 3(a), including, for example, the state ordering, the two interesting crossings of the $A\,^2\Sigma^+$ and $1\,^2\Sigma^-$ states at the O-H distance $R_{O-H}\approx 1.5$ Å and of the $2\,^2\Pi$ and $B\,^2\Sigma^+$ states at $R_{O-H}\approx 1.3$ Å, and the asymptotic degeneracies of the $X\,^2\Pi$ and $1\,^2\Sigma^-$ states and the $1\,^2\Delta$, $2\,^2\Delta$, and $A\,^2\Sigma^+$ states.

The performance of the basic IP EOMCC and SAC-CI approximations, in which the electron removing operator $R_\mu^{(N-1)}$ is truncated at the $2h$-$1p$ excitations, is generally very poor. In the following, we only discuss the results of the IP-EOMCCSD($2h$-$1p$) calculations, since the analogous SAC-CI($2h$-$1p$) results are essentially identical. As shown in Table VI, the only two states that are reasonably well described by the IP-EOMCCSD($2h$-$1p$) method are the $X\,^2\Pi$ and $A\,^2\Sigma^+$ states, and even in this case the applicability of the IP-EOMCCSD($2h$-$1p$) and SAC-CI($2h$-$1p$) approaches is limited to the region of the minimum on the $X\,^2\Pi$ curve and to the region of the relatively small stretches of the O-H bond which do not exceed $R_{O-H}=1.5$ Å (R_{O-H} is the O-H separation and $R_{O-H}=0.969\,66$ Å is the experimental equilibrium bond length in OH taken from Ref. 133). This can be understood by analyzing the leading contributions to the full CI wave functions of the $X\,^2\Pi$ and $A\,^2\Sigma^+$ states shown in Table VII. In the $R_{O-H}\leqslant 1.5$ Å region, the $X\,^2\Pi$ and $A\,^2\Sigma^+$ states are dominated by $1h$ excitations. The IP EOMCC and SAC-CI approximations truncated at $2h$-$1p$ excitations work well for such states. Unfortunately, once we enter the $R_{O-H}>1.5$ Å region, the $X\,^2\Pi$ and $A\,^2\Sigma^+$ states gain significant $2h$-$1p$ components and, in the $R_{O-H}\approx 3.0$ Å region, the relatively large $3h$–$2p$ contributions. The presence of the relatively large higher-than-$1h$ contributions in the wave functions representing the $X\,^2\Pi$ and $A\,^2\Sigma^+$ states in the $R_{O-H}>1.5$ Å region results in a rapid deterioration of the quality of the IP-EOMCCSD($2h$-$1p$) results. Indeed, the relatively small errors in the IP-EOMCCSD($2h$-$1p$) results relative to full CI in the $R_{O-H}\leqslant 1.5$ Å region, which do not exceed 6.231 mhartree for the $X\,^2\Pi$ state and 8.921 mhartree

TABLE VI. The differences between various IP EOMCC, SAC-CI, and CI($4h$-$3p$) energies and the corresponding full CI energies (Ref. 129) (in mhartree) at representative internuclear separations R_{O-H} (in Å), and the associated MUE and NPE values (also in mhartree) for the low-lying doublet states of the OH radical, as described by the 6-31G(d,p) basis set (Refs. 126 and 127).

		R_{O-H} (Å)										
State	Method	0.77	0.969 66[a]	1.07	1.27	1.50	1.75	2.00	2.50	3.00	MUE	NPE
$X\ ^2\Pi$	IP-EOMCCSD($2h$-$1p$)	−4.293	−3.081	−2.165	0.592	6.231	16.711	31.180	59.437	73.473	73.473	77.766
	IP-EOMCCSD($3h$-$2p$)	1.483	1.433	1.381	1.323	1.621	2.842	4.955	9.188	11.777	11.777	10.454
	SAC-CI($3h$-$2p$)	−1.066	−1.542	−1.817	−2.437	−3.134	−3.350	−2.520	0.463	2.508	3.350	5.858
	SAC-CI($4h$-$3p$)	−2.217	−2.557	−2.724	−3.117	−3.732	−4.430	−4.776	−4.421	−3.841	4.776	2.559
	SAC-CI($4h$-$3p$)\{3,1\}[b]	−1.039	−1.363	−1.555	−2.003	−2.583	−3.159	−3.381	−2.895	−2.257	3.381	2.342
	SAC-CI($4h$-$3p$)/PS	−0.428	−0.957	−1.291	−1.732	−2.437	−3.261	−3.773	−3.461	−3.024	3.773	3.345
	CI($4h$-$3p$)	2.172	2.676	2.947	3.556	4.502	5.996	7.698	9.808	10.142	10.142	7.970
$1\ ^2\Delta$	IP-EOMCCSD($2h$-$1p$)	230.128	225.597	225.822	236.405	258.298	284.966	310.557	352.144	378.026	378.026	152.747
	IP-EOMCCSD($3h$-$2p$)	3.033	7.191	9.759	15.539	22.729	30.069	35.936	42.334	42.747	42.747	39.714
	SAC-CI($3h$-$2p$)	2.801	6.879	9.511	15.802	24.361	33.982	42.585	54.844	61.313	61.313	58.512
	SAC-CI($4h$-$3p$)	−0.781	−0.647	−0.554	−0.429	−0.388	−0.390	−0.385	−0.372	−0.497	0.781	0.409
	SAC-CI($4h$-$3p$)\{3,1\}[b]	0.440	1.485	2.021	2.577	2.559	2.381	2.283	2.246	2.129	2.577	2.137
	SAC-CI($4h$-$3p$)/PS	−0.006	0.201	0.271	0.404	0.557	0.730	0.519	0.638	0.625	0.803	0.809
	CI($4h$-$3p$)	12.209	12.359	12.461	13.532	16.098	19.456	22.462	26.366	28.181	28.181	15.972
$2\ ^2\Pi$	IP-EOMCCSD($2h$-$1p$)	256.555	171.011	141.150	107.797	89.486	78.830	72.216	67.703	68.382	256.555	188.852
	IP-EOMCCSD($3h$-$2p$)	7.793	4.281	2.400	1.544	1.363	1.135	1.136	2.621	3.894	7.793	6.658
	SAC-CI($3h$-$2p$)	7.405	4.259	2.536	1.833	1.698	1.289	0.706	0.425	0.803	7.405	7.030
	SAC-CI($4h$-$3p$)	−0.857	−0.318	−0.157	−0.089	−0.144	−0.335	−0.727	−1.421	−1.531	1.531	1.442
	SAC-CI($4h$-$3p$)\{3,1\}[b]	0.536	1.699	2.107	2.270	1.996	1.563	1.028	0.258	0.158	2.270	2.112
	SAC-CI($4h$-$3p$)/PS	1.591	0.893	0.938	1.020	0.815	0.417	−0.113	−0.795	−0.965	1.591	2.556
	CI($4h$-$3p$)	13.126	10.889	9.027	6.845	5.597	4.883	4.707	5.594	6.128	13.126	8.419
$A\ ^2\Sigma^+$	IP-EOMCCSD($2h$-$1p$)	−4.256	−1.200	0.599	4.263	8.921	15.291	21.221	27.904	31.552	31.552	35.808
	IP-EOMCCSD($3h$-$2p$)	1.295	1.387	1.460	1.682	2.312	3.872	5.613	7.065	6.272	7.065	5.770
	SAC-CI($3h$-$2p$)	−1.359	−1.655	−1.758	−1.852	−1.409	0.438	2.874	5.895	6.905	6.905	8.757
	SAC-CI($4h$-$3p$)	−2.317	−2.702	−2.878	−3.188	−3.439	−3.506	−3.460	−3.492	−3.651	3.651	1.334
	SAC-CI($4h$-$3p$)\{3,1\}[b]	−1.330	−1.680	−1.900	−2.311	−2.505	−2.373	−2.183	−2.113	−2.237	2.505	1.175
	SAC-CI($4h$-$3p$)/PS	−1.843	−2.152	−2.280	−2.621	−2.843	−2.981	1.502	−3.042	−3.193	3.193	4.695
	CI($4h$-$3p$)	2.248	2.840	3.159	3.840	4.942	6.942	9.093	11.516	12.298	12.298	10.050
$1\ ^2\Sigma^-$	IP-EOMCCSD($2h$-$1p$)	262.848	262.749	265.988	283.913	315.124	350.078	380.988	426.231	450.090	450.090	187.712
	IP-EOMCCSD($3h$-$2p$)	7.410	12.392	15.535	22.612	31.064	38.914	44.518	49.731	49.263	49.731	42.321
	SAC-CI($3h$-$2p$)	7.043	11.957	15.187	22.856	32.788	42.968	51.235	61.994	67.298	67.298	60.255
	SAC-CI($4h$-$3p$)	−0.698	−0.564	−0.467	−0.318	−0.247	−0.237	−0.253	−0.340	−0.597	0.698	0.461
	SAC-CI($4h$-$3p$)\{3,1\}[b]	0.821	1.939	2.510	3.107	3.093	2.899	2.763	2.603	2.338	3.107	2.286
	SAC-CI($4h$-$3p$)/PS	0.391	0.903	0.841	1.160	1.086	0.928	1.491	1.804	1.587	1.942	1.551
	CI($4h$-$3p$)	12.495	12.775	12.945	14.087	16.484	19.273	21.544	24.282	25.486	25.486	12.991
$B\ ^2\Sigma^+$	IP-EOMCCSD($2h$-$1p$)	196.170	188.632	190.368	197.080	245.827	257.391	274.583	310.450	335.385	335.385	146.753
	IP-EOMCCSD($3h$-$2p$)	0.765	4.203	6.234	10.556	15.433	19.675	22.722	26.164	26.161	26.164	25.399
	SAC-CI($3h$-$2p$)	0.515	3.851	5.945	10.732	16.702	22.512	27.370	34.916	39.324	39.324	38.809
	SAC-CI($4h$-$3p$)	−0.946	−0.797	−0.668	−0.411	−0.159	0.075	0.352	1.021	1.464	1.464	2.410
	SAC-CI($4h$-$3p$)\{3,1\}[b]	0.217	1.244	1.790	2.435	2.561	2.501	2.567	3.053	3.452	3.452	3.235
	SAC-CI($4h$-$3p$)/PS	−0.196	−0.074	0.105	0.299	0.484	0.772	0.931	1.688	2.118	2.118	2.505
	CI($4h$-$3p$)	12.418	12.821	13.172	14.836	17.759	20.523	22.439	25.122	26.759	26.759	14.341

[a]Equilibrium bond length taken from Ref. 133.
[b]The active space consisted of the 3σ, $1\pi_x$, $1\pi_y$, and 4σ orbitals of OH^-.

for the $A\ ^2\Sigma^+$ state, increase to 16.711-73.473 and 15.291–31.552 mhartree, respectively, when the 1.75A $\leqslant R_{O-H} \leqslant$ 3.0 Å region is examined.

The situation for the remaining states of OH considered in this work, which have large $2h$-$1p$ and, in some cases, large $3h$-$2p$ components independent of the value of R_{O-H} (cf. Table VII), is even more dramatic. Indeed, the errors in the IP-EOMCCSD($2h$-$1p$) results for the $1\ ^2\Delta$, $2\ ^2\Pi$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states are 225.279–378.026, 67.703–256.555, 262.378–450.090, and 188.632–335.385 mhartree, respectively, when the entire 0.77 A $\leqslant R_{O-H} \leqslant$ 3.0 Å region is considered (see Table VI and Ref. 129). In consequence, the IP-EOMCCSD($2h$-$1p$) potential energy curves for the $1\ ^2\Delta$, $2\ ^2\Pi$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states are completely pathological

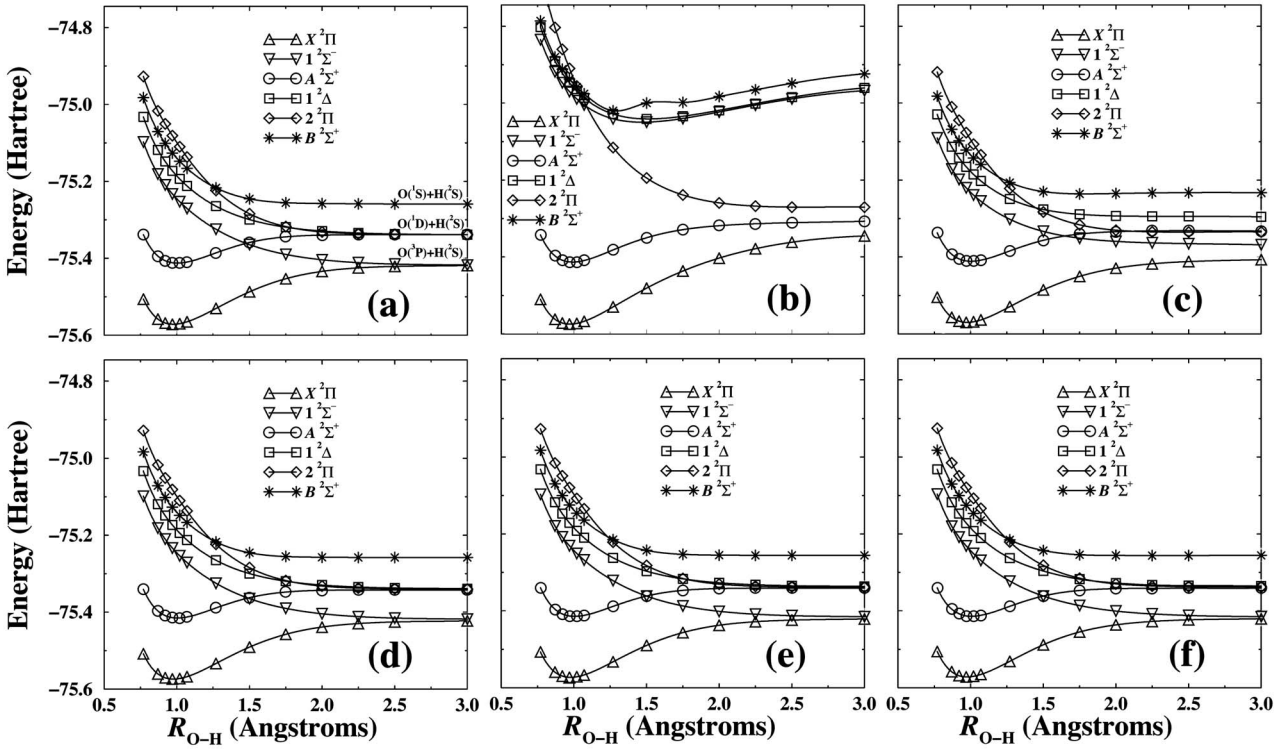

FIG. 3. Potential energy curves for the ground and low-lying doublet excited states of the OH radical, as described by the 6-31G(d,p) basis set (Refs. 126 and 127). Energies are in hartree and the O–H distance R_{O-H} is in Å. (a) the full CI results, (b) the IP-EOMCCSD($2h$-$1p$) results, (c) the IP-EOMCCSD($3h$-$2p$) results, (d) the SAC-CI($4h$-$3p$) results, (e) the SAC-CI($4h$-$3p$)$\{3,1\}$ results, and (f) the SAC-CI($4h$-$3p$)/PS results.

[cf. Figs. 3(a) and 3(b)]. The IP-EOMCCSD($2h$-$1p$) curves are characterized by the very large MUE and NPE values. The energy gaps between electronic states and their relative ordering resulting from the IP-EOMCCSD($2h$-$1p$) calculations are incorrect, even in the vicinity of the minimum on the ground-state potential energy curve. The $A\ ^2\Sigma^+$ and $1\ ^2\Sigma^-$ curves, which should cross at $R_{O-H}\approx 1.5$ Å, do not cross in the IP-EOMCCSD($2h$-$1p$) calculations, since the $1\ ^2\Sigma^-$ state is significantly shifted up in energy. The $2\ ^2\Pi$ and $B\ ^2\Sigma^+$ curves obtained in the IP-EOMCCSD($2h$-$1p$) calculations cross, but not at the right geometry. The asymptotic degeneracies of the $X\ ^2\Pi$ and $1\ ^2\Sigma^-$ states and the $1\ ^2\Delta$, $2\ ^2\Pi$, and $A\ ^2\Sigma^+$ states are completely broken by the IP-EOMCCSD($2h$-$1p$) method.

As shown in Table VI and Fig. 3(c), the inclusion of the $3h$-$2p$ excitations in the $R_\mu^{(N-1)}$ operator leads to large improvements in the poor IP-EOMCCSD($2h$-$1p$) [and SAC-CI($2h$-$1p$)] results. These improvements are particularly substantial in the region of the minimum on the ground-state potential energy curve and for the $1\ ^2\Delta$, $2\ ^2\Pi$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states, which have significant $2h$-$1p$ components. For example, the 225.597, 171.011, 262.749, and 188.632 mhartree errors in the IP-EOMCCSD($2h$-$1p$) results for the $1\ ^2\Delta$, $2\ ^2\Pi$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states at $R_{O-H}=0.969\ 66$ Å (the experimental equilibrium geometry of the ground-state OH radical) reduce to 7.191, 4.281, 12.392, and 4.203 mhartree, respectively, when the IP-EOMCCSD($3h$-$2p$) approach is used, and 6.879, 4.259, 11.957, and 3.851 mhartree, respectively, when the SAC-CI($3h$-$2p$) method is employed (see Table VI). For the remaining $X\ ^2\Pi$ and $A\ ^2\Sigma^+$ states, which in the Franck-Condon region are dominated by the $1h$ excitations and which are adequately described by the IP-EOMCCSD($2h$-$1p$) approach when $R_{O-H}\approx 0.969\ 66$ Å, we observe further error reduction in the already good IP-EOMCCSD($2h$-$1p$) results (the $X\ ^2\Pi$ case) or no essential changes in the quality of the IP-EOMCCSD($2h$-$1p$) energies (the $A\ ^2\Sigma^+$ case). If we limit ourselves to the Franck-Condon region, there is almost no difference between the IP-EOMCCSD($3h$-$2p$) potential energy curves representing the $X\ ^2\Pi$, $1\ ^2\Delta$, $2\ ^2\Pi$, $A\ ^2\Sigma^+$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states of OH [shown in Fig. 3(c)] and the corresponding full CI curves [shown in Fig. 3(a)]. In particular, the IP-EOMCCSD($3h$-$2p$) method restores the correct state ordering in the region of the minimum on the ground-state curve of OH, destroyed by the IP-EOMCCSD($2h$-$1p$) model. It also provides a reasonable description of the crossing of the $A\ ^2\Sigma^+$ and $1\ ^2\Sigma^-$ states, which in the full CI calculations appears at $R_{O-H}\approx 1.5$ Å and in the IP-EOMCCSD($3h$-$2p$) calculations at $R_{O-H}\approx 1.6$ Å, and of the crossing of the $2\ ^2\Pi$ and $B\ ^2\Sigma^+$ states, which both in the full CI and in the IP-EOMCCSD($3h$-$2p$) calculations appears at $R_{O-H}\approx 1.3$ Å. The IP-EOMCCSD($3h$-$2p$) method provides an overall very good description of the ground and excited states of OH in the Franck-Condon region. The same is true for SAC-CI($3h$-$2p$).

The IP-EOMCCSD($3h$-$2p$) and SAC-CI($3h$-$2p$) approaches are also quite effective in describing the entire potential energy curves of the $X\ ^2\Pi$, $2\ ^2\Pi$, and $A\ ^2\Sigma^+$ states, which (with an exception of the $X\ ^2\Pi$ state in the $R_{O-H}\approx 3.0$ Å region) are dominated by the $1h$ and $2h$-$1p$ excita-

TABLE VII. An analysis of the major full CI configurations (all configurations with a coefficient $\geqslant 0.15$ for at least one of the selected values of R_{O-H} are included) for the low-lying doublet states of the OH radical, as described by the 6-31G(d,p) basis set (Refs. 126 and 127).

		Coefficients for various values of R_{O-H}				
State	Configuration orbital occupancy	0.77 Å	0.969 66 Å[a]	1.50 Å	3.00 Å	Excitation type[b]
$X\ ^2\Pi$	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^2(1\pi_x)^2(1\pi_y)^1\|$	0.961	0.951	0.873	−0.431	1h
	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1(1\pi_x)^2(1\pi_y)^1(4\sigma)^1\|$[c]	<0.01, 0.070	0.017, 0.145	0.120, 0.356	0.472, 0.621	2h-1p
	$\|(1\sigma)^2(2\sigma)^2(1\pi_x)^2(1\pi_y)^1(4\sigma)^2\|$	<0.01	−0.024	−0.144	0.356	3h-2p
$1\ ^2\Delta$	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^2[(1\pi_x)^2-(1\pi_y)^2](4\sigma)^1\|$	−0.651	0.633	0.555	0.268	2h-1p
	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1[(1\pi_x)^2-(1\pi_y)^2](4\sigma)^2\|$	0.075	0.144	−0.367	−0.620	3h-2p
$2\ ^2\Pi$	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^2(1\pi_x)^2(1\pi_y)^1\|$	−0.040	0.027	−0.114	0.250	1h
	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1(1\pi_x)^2(1\pi_y)^1(4\sigma)^1\|$[c]	−0.528, 0.746	−0.407, 0.812	−0.086, 0.934	−0.352, 0.839	2h-1p
	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1(1\pi_x)^2(1\pi_y)^1(5\sigma)^1\|$	−0.027	0.178	−0.180	0.031	2h-1p
	$\|(1\sigma)^2(2\sigma)^2(1\pi_x)^2(1\pi_y)^1(4\sigma)^2\|$	0.124	−0.138	−0.047	−0.196	3h-2p
$A\ ^2\Sigma^+$	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1(1\pi_x)^2(1\pi_y)^2\|$	0.960	0.958	0.941	0.755	1h
	$\|(1\sigma)^2(2\sigma)^2(1\pi_x)^2(1\pi_y)^2(4\sigma)^1\|$	0.060	0.101	<0.01	0.254	2h-1p
	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^2[(1\pi_x)^2+(1\pi_y)^2](4\sigma)^1\|$	<0.01	<0.01	0.115	−0.150	2h-1p
	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1[(1\pi_x)^2+(1\pi_y)^2](4\sigma)^2\|$	<0.01	0.012	−0.098	−0.356	3h-2p
$1\ ^2\Sigma^-$	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^2(1\pi_x)^1(1\pi_y)^1(4\sigma)^1\|$[c]	−0.459, 0.795	<0.01, 0.892	−0.380, 0.658	0.185, −0.321	2h-1p
	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1(1\pi_x)^1(1\pi_y)^1(4\sigma)^2\|$[c]	−0.052, 0.091	−0.108, 0.186	−0.277, 0.479	−0.440, 0.763	3h-2p
$B\ ^2\Sigma^+$	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1(1\pi_x)^2(1\pi_y)^2\|$	0.019	0.012	−0.207	0.524	1h
	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^2[(1\pi_x)^2+(1\pi_y)^2](4\sigma)^1\|$	0.636	0.613	0.527	0.217	2h-1p
	$\|(1\sigma)^2(2\sigma)^2(1\pi_x)^2(1\pi_y)^2(4\sigma)^1\|$	−0.148	−0.158	−0.169	0.166	2h-1p
	$\|(1\sigma)^2(3\sigma)^1(1\pi_x)^2(1\pi_y)^2(4\sigma)^2\|$	0.013	0.025	0.081	−0.173	3h-2p
	$\|(1\sigma)^2(2\sigma)^2(3\sigma)^1[(1\pi_x)^2+(1\pi_y)^2](4\sigma)^2\|$	−0.072	0.139	−0.345	0.499	3h-2p

[a]Equilibrium bond length taken from Ref. 133.
[b]Relative to the ground-state reference configuration of OH^-, $|1\sigma^2 2\sigma^2 3\sigma^2 1\pi_x^2 1\pi_y^2|$.
[c]The two different coefficients shown are for the two doublet configuration state functions, each corresponding to a different intermediate spin state, that result from coupling the spins of the three unpaired electrons in this orbital occupation scheme.

tions, with the errors relative to full CI that do not exceed a few millihartrees, but they fail to accurately describe the $1\ ^2\Delta$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states at larger O–H separations (see Table VI and Ref. 129). Indeed, although the IP-EOMCCSD(3h-2p) and SAC-CI(3h-2p) methods improve the pathological description of the $1\ ^2\Delta$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states by the IP-EOMCCSD(2h-1p) and SAC-CI(2h-1p) approaches, reducing the 225.279–378.026, 262.378–450.090, and 188.632–335.385 mhartree errors in the IP-EOMCCSD(2h-1p) results for these three states in the entire $0.77\ \text{Å} \leqslant R_{O-H} \leqslant 3.0$ Å region to 3.033–42.747, 7.410–49.731, and 0.765–26.164 mhartree, respectively, in the IP-EOMCCSD(3h-2p) case and 2.801–61.313, 7.043–67.298, and 0.515–39.324 mhartree, respectively, in the SAC-CI(3h-2p) case, the errors in the IP-EOMCCSD(3h-2p) and SAC-CI(3h-2p) results for the $1\ ^2\Delta$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states at larger O–H distances are too large for high accuracy calculations. The IP-EOMCCSD(3h-2p) and SAC-CI(3h-2p) approaches break the asymptotic degeneracies of the $X\ ^2\Pi$ and $1\ ^2\Sigma^-$ states, which should dissociate into $O(^3P)+H(^2S)$, and of the $1\ ^2\Delta$, $2\ ^2\Pi$, and $A\ ^2\Sigma^+$ states, which should dissociate into $O(^1D)+H(^2S)$, since they fail to correctly describe the $1\ ^2\Sigma^-$ and $1\ ^2\Delta$ states at larger O–H separations. For example, the IP-EOMCCSD(3h-2p) method restores the asymptotic degeneracy of the $2\ ^2\Pi$ and $A\ ^2\Sigma^+$ states, broken by IP-EOMCCSD(2h-1p), but this is not sufficient to obtain a correct description of the asymptotic region, since the $2\ ^2\Pi$ and $A\ ^2\Sigma^+$ states should also become degenerate with the $1\ ^2\Delta$ state as $R_{O-H} \to \infty$ [see Fig. 3(c)].

The IP-EOMCCSD(3h-2p) and SAC-CI(3h-2p) approximations cannot provide an accurate description of the asymptotic region, since the $1\ ^2\Delta$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states gain large 3h-2p contributions at stretched O–H distances and, as argued in the Introduction, one has to include the 4h-3p excitations in the IP EOMCC and SAC-CI calculations in order to obtain an accurate description of the electronic states characterized by larger 3h-2p contributions. The significance of the 4h-3p components in the electron removing operator $R_\mu^{(N-1)}$ at larger O–H separations, particularly when the $1\ ^2\Delta$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states are examined, can be seen by comparing the IP-EOMCCSD(3h-2p) and SAC-CI(3h-2p) results, which ignore the 4h-3p contributions altogether, with the results of the CI(4h-3p) calculations, in which the $R_\mu^{(N-1)}$ operator truncated at 4h-3p excitations is directly applied to the reference determinant $|\Phi\rangle$. Although, as shown in Table VI, the description of the potential energy curves representing the $1\ ^2\Delta$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states of OH by the CI(4h-3p) approach is not fully quantitative and the maximum errors in the CI(4h-3p) results are still as large as 28.181 mhartree for the $1\ ^2\Delta$ state, 25.486 mhartree for the $1\ ^2\Sigma^-$ state, and 26.759 mhartree for the $B\ ^2\Sigma^+$ state, we already observe major improvements in

TABLE VIII. The differences between the CCSD, CCSDT, CR-CC(2,3), and SAC-SD ground-state energies of the OH^- ion, as described by the 6-31G(d,p) basis set (Refs. 126 and 127), and the corresponding full CI ground-state energies (Ref. 129) (in mhartree) at representative internuclear separations R_{O-H} (inÅ), and the associated MUE and NPE values relative to full CI (also in mhartree).

	R_{O-H} (Å)										
Method	0.77	0.969 66[a]	1.07	1.27	1.50	1.75	2.00	2.50	3.00	MUE	NPE
CCSD	2.087	2.714	3.172	4.424	6.508	9.487	12.958	20.410	26.865	26.865	24.778
CCSDT	0.464	0.594	0.652	0.735	0.779	0.826	0.927	1.341	1.840	1.840	1.376
CR-CC(2,3)	0.338	0.446	0.479	0.501	0.458	0.435	0.618	1.179	0.701	1.179	0.841
SAC-SD	2.032	2.657	3.100	4.322	6.541	10.202	15.111	27.648	41.660	41.660	39.628

[a]The equilibrium bond length of OH taken from Ref. 133.

the results, compared to the SAC-CI($3h$-$2p$) and IP-EOMCCSD($3h$-$2p$) calculations, both in the maximum errors and in the corresponding NPE values, when the CI($4h$-$3p$) method is employed. This implies that the inclusion of the genuine $4h$-$3p$ components in the $R_\mu^{(N-1)}$ operator is the key step toward improving the overall description of the potential energy curves of the low-lying states of OH. Other factors, such as the use of the quasilinearized form of the SAC-CI($3h$-$2p$) wave functions [Eq. (32)] and the neglect of several nonlinear terms in the cluster operator S in the SAC-SD calculations for the OH^- reference system that precede the SAC-CI($3h$-$2p$) calculations for OH are of lesser importance, since the IP-EOMCCSD($3h$-$2p$) approach, which uses a more complete form of the wave function compared to SAC-CI($3h$-$2p$), provides the results which are essentially as inaccurate in the region of larger O–H distances as those obtained with SAC-CI($3h$-$2p$). As shown in Table VI, there are relatively small differences between the results obtained with the quasilinearized form of SAC-CI($3h$-$2p$) used in this paper and the results of the IP-EOMCCSD($3h$-$2p$) calculations, which do not neglect any nonlinear terms in S that enter the full IP-EOMCCSD($3h$-$2p$)/SAC-CI($3h$-$2p$) eigenvalue problem. Moreover, as shown in Table VIII, with an exception of the $R_{O-H} \geqslant 2.5$ Å region, the differences between the results of the SAC-SD calculations for the ground-state potential energy curve of OH^-, in which all nonlinear terms other than $\frac{1}{2}S_2^2$ are neglected, and the analogous results obtained with the full CCSD approach, in which all nonlinear terms in S_1 and S_2 are retained, are on the order of a fraction of a millihartree or a few millihartrees, at most. The SAC-SD and CCSD curves for OH^- begin to differ more substantially in the $R_{C-H} \geqslant 2.5$ Å region, but this is not too important in the context of the present discussion, since neither SAC-SD nor CCSD works well in this region (see Table VIII). A similar remark applies to the SAC-CI($3h$-$2p$) versus IP-EOMCCSD($3h$-$2p$) results for the most difficult $1\,^2\Delta$, $1\,^2\Sigma^-$, and $B\,^2\Sigma^+$ states, which begin to deviate more substantially only in the $R_{C-H} \geqslant 2.5$ Å region where neither SAC-CI($3h$-$2p$) nor IP-EOMCCSD($3h$-$2p$) works well. We can conclude that since we do not neglect any relevant nonlinear terms in S in the IP-EOMCCSD($3h$-$2p$) calculations and yet the IP-EOMCCSD($3h$-$2p$) results for the $1\,^2\Delta$, $1\,^2\Sigma^-$, and $B\,^2\Sigma^+$ states of OH at larger O–H distances remain quite poor, the failures of the SAC-CI($3h$-$2p$) and IP-EOMCCSD($3h$-$2p$) approximations in the region of larger R_{O-H} values are primarily due to the absence of the $4h$-$3p$ excitations in the electron removing operators $R_\mu^{(N-1)}$ defining the SAC-CI($3h$-$2p$) and IP-EOMCCSD($3h$-$2p$) schemes.

The CI($4h$-$3p$) approach improves the poor description of the $1\,^2\Delta$, $1\,^2\Sigma^-$, and $B\,^2\Sigma^+$ states at larger O–H separations by the IP-EOMCCSD($3h$-$2p$), SAC-CI($3h$-$2p$), and other lower-order IP EOMCC and SAC-CI methods, but, as shown in Table VI, the most accurate results for the electronic states of OH discussed in this work are obtained when we apply the electron removing operator $R_\mu^{(N-1)}$ truncated at $4p$-$3h$ excitations to the correlated ground state of OH^-, as is done in the SAC-CI($4h$-$3p$) calculations, rather than to the uncorrelated, single-determinantal state $|\Phi\rangle$ used in the CI($4h$-$3p$) model. This becomes clear when we examine the SAC-CI($4h$-$3p$) results in Table VI and when we compare the SAC-CI($4h$-$3p$) potential energy curves shown in Fig. 3(d) with the corresponding full CI curves shown in Fig. 3(a). The maximum errors of 42.747, 61.313, and 28.181 mhartree characterizing the IP-EOMCCSD($3h$-$2p$), SAC-CI($3h$-$2p$), and CI($4h$-$3p$) calculations for the $1\,^2\Delta$ state, the 49.731, 67.298, and 25.486 mhartree maximum errors characterizing the IP-EOMCCSD($3h$-$2p$), SAC-CI($3h$-$2p$), and CI($4h$-$3p$) results for the $1\,^2\Sigma^-$ state, and the maximum errors of 26.164, 39.324, and 26.759 mhartree characterizing the IP-EOMCCSD($3h$-$2p$), SAC-CI($3h$-$2p$), and CI($4h$-$3p$) calculations for the $B\,^2\Sigma^+$ state reduce to 0.781, 0.698, and 1.464 mhartree, respectively, when the SAC-CI($4h$-$3p$) approach is employed (see Table VI). For the remaining states of OH listed in Table VI, the results of the SAC-CI($4h$-$3p$) calculations are equally good. We might be able to obtain additional small improvements if we used the rigorously size intensive IP-EOMCCSDT($4h$-$3p$) approach, in which the $R_\mu^{(N-1)}$ operator truncated at the $4h$-$3p$ excitations is applied to the full CCSDT ground state of OH^-, but the IP-EOMCCSDT($4h$-$3p$) method has not been implemented yet [again, the related IP-EOMCCSDTQ($4h$-$3p$)=IP-EOMCCSDTQ approach that uses the CCSDTQ rather than the CCSDT reference ground state in the IP EOMCC calculations truncated at $4h$-$3p$ excitations has recently been implemented,[47] but this program is not available to us]. As implied by the excellent performance of the full CCSDT approach in the calculations for OH^- and the analogous results obtained with the CR-CC(2,3) approach, which offers

an approximate and yet highly accurate treatment of S_3 clusters in all cases involving single bond breaking,[131,132] both shown in Table VIII, the explicit inclusion of S_3 clusters would lead to the virtually exact description of the ground state of OH^-, which serves as a reference for the IP EOMCC and SAC-CI calculations for OH. This observation, combined with the great improvements in accuracy offered by the presence of the $4h$-$3p$ excitations in the electron removing operator $R_\mu^{(N-1)}$, makes us believe that the IP-EOMCCSDT($4p$-$3h$) approach would give the virtually exact description of the entire potential energy curves of OH. Having said all this, it is most encouraging to observe that the simpler SAC-CI($4h$-$3p$) method, in which nonlinear terms in S and the S_3 components are ignored and which is less expensive than the IP-EOMCCSDT($4h$-$3p$) approach, is capable of providing excellent results for the entire ground- and excited-state potential energy curves of OH, including the asymptotic region. In addition to the small maximum errors and NPE values characterizing the SAC-CI($4h$-$3p$) results for OH, the SAC-CI($4h$-$3p$) approach restores the asymptotic degeneracies of the $X\ ^2\Pi$ and $1\ ^2\Sigma^-$ states, which dissociate into $O(^3P)+H(^2S)$, and of the $1\ ^2\Delta$, $2\ ^2\Pi$, and $A\ ^2\Sigma^+$ states, which dissociate into $O(^1D)+H(^2S)$, which are not properly described by the lower-order IP SAC-CI and EOMCC models [cf. Figs. 3(d) and 3(a)]. Moreover, as shown in Fig. 3(d), the SAC-CI($4h$-$3p$) approach provides the virtually perfect description of the crossings of the $A\ ^2\Sigma^+$ and $1\ ^2\Sigma^-$ states at $R_{O-H}\approx 1.5$ Å and of the $2\ ^2\Pi$ and $B\ ^2\Sigma^+$ states at $R_{O-H}\approx 1.3$ Å, further improving the IP-EOMCCSD($3h$-$2p$) and SAC-CI($3h$-$2p$) results in this regard. In analogy to the CH case, we might wonder if one could obtain similarly accurate results with the IP-EOMCCSDT($3h$-$2p$) method, in which S is truncated at S_3 and the $4h$-$3p$ components of $R_\mu^{(N-1)}$ are ignored, since CCSDT works perfectly for OH^-; we plan to study this issue in the future, but at this time we believe that one needs the $4h$-$3p$ excitations in $R_\mu^{(N-1)}$, since the CI($4h$-$3p$) calculations greatly improve the IP-EOMCCSD($3h$-$2p$) results for the most difficult $1\ ^2\Delta$, $2\ ^2\Pi$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states.

Thus far, we have demonstrated that the SAC-CI($4h$-$3p$) approach provides high-quality potential energy curves for the ground and excited states of OH. The final question remains if one can retain the high accuracy of the SAC-CI($4h$-$3p$) calculations without a full inclusion of all $4h$-$3p$ amplitudes r_{bcd}^{ijkl}. In analogy to the CH case, one can reduce large costs of the SAC-CI($4h$-$3p$) calculations by using the SAC-CI($4h$-$3p$)/PS approach, in which the $4h$-$3p$ components of the $R_\mu^{(N-1)}$ operators that do not significantly perturb the CI($3h$-$2p$) reference wave function and whose energy contributions are smaller than the threshold λ_e (set in our calculations at 10^{-7} hartree) are neglected and eliminated from the SAC-CI($4h$-$3p$) diagonalization. As shown in Table VI, the SAC-CI($4h$-$3p$)/PS results are essentially as good as those obtained with the complete SAC-CI($4h$-$3p$) approach, in spite of the fact that the total number of spin- and symmetry-adapted amplitudes defining the $R_\mu^{(N-1)}$ operator used in the SAC-CI($4h$-$3p$)/PS calculations represents 12%–46% of all r amplitudes used in the SAC-CI($4h$-$3p$) calculations (see Table II). There is, however, a problem, mentioned in the Introduction, of the numerical noise which the SAC-CI($4h$-$3p$)/PS calculations may produce due to the fact that different sets of the $4h$-$3p$ contributions to the $R_\mu^{(N-1)}$ operators are selected at different nuclear geometries. The active-space SAC-CI($4h$-$3p$) approach developed in this work uses the same set of r amplitudes defining the $R_\mu^{(N-1)}$ operator at all nuclear geometries, producing smooth potential energy curves. At the same time, the active-space SAC-CI($4h$-$3p$) method offers a substantial reduction in the dimensionality of the corresponding eigenvalue problem through the selection of the dominant $3h$-$2p$ and $4h$-$3p$ amplitudes via active orbitals, as discussed in Sec. II B (see Table II). As shown in Table VI, the active-space SAC-CI($4h$-$3p$)$\{3,1\}$ approach employing only three occupied and one unoccupied active orbitals, and using approximately 29% of all amplitudes r defining the $R_\mu^{(N-1)}$ operator truncated at $4h$-$3p$ excitations, is as effective in producing highly accurate potential energy curves of OH as the SAC-CI($4h$-$3p$)/PS and complete SAC-CI($4h$-$3p$) methods. The small maximum errors relative to full CI of 3.381, 2.577, 2.270, 2.505, 3.107, and 3.452 mhartree characterizing the SAC-CI($4p$-$3h$)$\{3,1\}$ results for the entire potential energy curves of the $X\ ^2\Pi$, $1\ ^2\Delta$, $2\ ^2\Pi$, $A\ ^2\Sigma^+$, $1\ ^2\Sigma^-$, and $B\ ^2\Sigma^+$ states of OH are similar to or only slightly larger than the 4.776, 0.781, 1.531, 3.651, 0.698, and 1.464 mhartree maximum errors obtained with the SAC-CI($4h$-$3p$) approach and the 3.773, 0.803, 1.591, 3.193, 1.942, and 2.118 mhartree maximum errors resulting from the SAC-CI($4h$-$3p$)/PS calculations. Similar remarks apply to the NPE values characterizing the potential energy curves of OH obtained in the SAC-CI($4h$-$3p$)$\{3,1\}$ calculations. In consequence, the SAC-CI($4h$-$3p$)$\{3,1\}$ potential energy curves shown in Fig. 3(e) can hardly be distinguished from the highly accurate SAC-CI($4h$-$3p$) and SAC-CI($4h$–$3p$)/PS curves shown in Figs. 3(d) and 3(f), respectively. There is practically no difference between the ground- and excited-state potential energy curves obtained in the relatively inexpensive active-space SAC-CI($4h$-$3p$)$\{3,1\}$ calculations shown in Fig. 3(e) and the full CI curves shown in Fig. 3(a). All essential features of the full CI curves, including the asymptotic degeneracies of the $X\ ^2\Pi$ and $1\ ^2\Sigma^-$ states, which dissociate into $O(^3P)+H(^2S)$, and of the $1\ ^2\Delta$, $2\ ^2\Pi$, and $A\ ^2\Sigma^+$ states, which dissociate into $O(^1D)+H(^2S)$ and the aforementioned crossings of the $A\ ^2\Sigma^+$ and $1\ ^2\Sigma^-$ states at $R_{O-H}\approx 1.5$ Å and of the $2\ ^2\Pi$ and $B\ ^2\Sigma^+$ states at $R_{O-H}\approx 1.3$ Å, are accurately described by the active-space SAC-CI($4h$-$3p$)$\{3,1\}$ approach.

We conclude by stating that the active-space SAC-CI($4h$-$3p$)$\{N_o,N_u\}$ approach exploiting small numbers of occupied and unoccupied active orbitals is practically as accurate as its considerably more expensive SAC-CI($4h$-$3p$) counterpart. The active-space SAC-CI($4h$-$3p$)$\{3,1\}$ approach employing four active orbitals to select the dominant $3h$-$2p$ and $4h$-$3p$ excitations provides an excellent description of the potential energy curves of the ground and excited states of the OH radical, including the Franck-Condon and asymptotic regions, at the relatively small fraction of the effort involved in the parent SAC-CI($4h$-$3p$) calculations.

IV. SUMMARY AND CONCLUDING REMARKS

In the present paper, we have focused on one of the major challenges of modern electronic structure theory, which is the development of practical methods that can accurately describe ground- and excited-state potential energy surfaces of radical species along bond breaking coordinates. Specifically, we have shown that one can develop relatively inexpensive and easy-to-use *ab initio* schemes for high accuracy calculations of ground- and excited-state potential energy surfaces of radicals by combining the EA/IP SAC-CI and EOMCC methods with $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ excitations with the idea of using the physically motivated sets of active orbitals to select the most important $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ excitations. The active-space variants of the SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) approaches turned out to be particularly successful, enabling us to obtain highly accurate potential energy curves of the ground and excited states of the CH and OH radicals along the relevant bond breaking coordinates, including the Franck-Condon and asymptotic regions, at a small fraction of the computer cost associated with the regular SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) calculations. We have demonstrated that it is sufficient to use small numbers of active orbitals that correlate with the valence shells of the radical species of interest and small fractions of all $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ excitations selected via active orbitals to obtain excellent results of the full SAC-CI($4p$-$3h$)/SAC-CI($4h$-$3p$) quality. We have also shown that in general the active-space EA and IP EOMCC and SAC-CI methods with $3p$-$2h$/$3h$-$2p$ or $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ excitations faithfully reproduce the results obtained with the parent EA and IP EOMCC and SAC-CI approaches including these high-order excitations.

By comparing the potential energy curves of the ground and excited states of the CH and OH systems obtained in the full and active-space SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) calculations with the corresponding full CI curves, full and active-space EA-EOMCCSD($3p$-$2h$)/SAC-CI($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$)/SAC-CI($3h$-$2p$) results, and potential energy curves obtained with the basic EA-EOMCCSD($2p$-$1h$) and IP-EOMCCSD($2h$-$1p$) approximations, we have shown that one needs the $4p$-$3h$ and $4h$-$3p$ excitations in the electron attaching and ionizing operators $R_\mu^{(N\pm1)}$ in order to obtain an accurate representation of the entire ground- and excited-state potential energy surfaces of radicals along bond breaking coordinates. The full and active-space EA-EOMCCSD($3p$-$2h$)/SAC-CI($3p$-$2h$) and IP-EOMCCSD($3h$-$2p$)/SAC-CI($3h$-$2p$) methods, in which the $4p$-$3h$ and $4h$-$3p$ components of $R_\mu^{(N\pm1)}$ are ignored, provide accurate excitation energies in the Franck-Condon region, but they fail at larger internuclear separations, where electronic states of radicals often gain significant $3p$-$2h$ and $3h$-$2p$ contributions. A rationale for this behavior of the EA/IP SAC-CI and EOMCC methods has been provided in the Introduction. The EA/IP SAC-CI and EOMCC methods with $4p$-$3h$/$4h$-$3p$ excitations are expected to work well in the Franck-Condon and asymptotic regions of many radical species, since radicals can be viewed as systems obtained by attaching an electron to or removing an electron from the related closed-shell molecule and it is well known that the triply excited (i.e., $3p$-$3h$) clusters are usually sufficient to obtain a quantitative description of single bond breaking in closed-shell systems. Our numerical results for the CH and OH radicals obtained with the full and active-space SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods confirm this expectation.

As shown in this study, the active-space variants of the SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods are at least as effective in describing ground and excited states of radicals as the SAC-CI($4p$-$3h$)/PS and SAC-CI($4h$-$3p$)/PS approaches, in which one selects the $4p$-$3h$ and $4h$-$3p$ excitations numerically, based on their significance in perturbing the CI($3p$-$2h$) and CI($3h$-$2p$) wave functions. The active-space SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods and their parent SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) approximations, in which all orbitals are active, have also been shown to be significantly more accurate than the analogous CI methods with up to $4p$-$3h$ and $4h$-$3p$ excitations. One has to apply the electron attaching and ionizing operators $R_\mu^{(N\pm1)}$ truncated at the $4p$-$3h$ and $4h$-$3p$ components to the correlated CC or SAC ground state of the N-electron reference system to obtain high-quality potential energy surfaces of the $(N\pm1)$-electron radical species along bond breaking coordinates.

Based on the numerical results and theoretical analysis presented in this paper, we expect that the active-space variants of the EA-EOMCCSD($4p$-$3h$) and IP-EOMCCSD($4h$-$3p$) approaches, which would represent the EOMCC analogs of the active-space SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods examined in this work, and, particularly, the active-space variants of the rigorously size intensive EA-EOMCCSDT($4p$-$3h$) and IP-EOMCCSDT($4h$-$3p$) methods, referred to in Ref. 34 as the EA-EOMCCSDtq and IP-EOMCCSDtq schemes, should be at least as effective as the active-space SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) approximations discussed here. The efficient computer implementation of the EA-EOMCCSDtq and IP-EOMCCSDtq methods proposed in Ref. 34, which would use active orbitals to select the most important S_3 components in the cluster operator S defining the underlying CCSDT/SAC-SDT calculations for the N-electron reference systems and the dominant $3p$-$2h$/$3h$-$2p$ and $4p$-$3h$/$4h$-$3p$ excitations in the corresponding operators $R_\mu^{(N\pm1)}$, would be a useful next step toward the development of high accuracy methods for radicals and other open-shell systems. On the other hand, it is encouraging to observe that the active-space SAC-CI($4p$-$3h$) and SAC-CI($4h$-$3p$) methods, which are less expensive than the EA- and IP-EOMCCSDtq schemes, provide highly accurate potential energy surfaces of radical species.

ACKNOWLEDGMENTS

This study was also supported by the Grant for Creative Scientific Research from the Ministry of Education, Science, Culture, and Sports of Japan to one of the authors (H.N.) and

the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy [Grant No. DE-FG02-01ER15228, to another author (P.P.)].

[1] F. Coester, Nucl. Phys. **7**, 421 (1958).
[2] F. Coester and H. Kümmel, Nucl. Phys. **17**, 477 (1960).
[3] J. Čížek, J. Chem. Phys. **45**, 4256 (1966).
[4] J. Čížek, Adv. Chem. Phys. **14**, 35 (1969).
[5] J. Čížek and J. Paldus, Int. J. Quantum Chem. **5**, 359 (1971).
[6] K. Emrich, Nucl. Phys. A **351**, 379 (1981).
[7] J. Geertsen, M. Rittby, and R. J. Bartlett, Chem. Phys. Lett. **164**, 57 (1989).
[8] D. C. Comeau and R. J. Bartlett, Chem. Phys. Lett. **207**, 414 (1993).
[9] J. F. Stanton and R. J. Bartlett, J. Chem. Phys. **98**, 7029 (1993).
[10] P. Piecuch and R. J. Bartlett, Adv. Quantum Chem. **34**, 295 (1999).
[11] H. Nakatsuji and K. Hirao, Chem. Phys. Lett. **47**, 569 (1977).
[12] H. Nakatsuji and K. Hirao, J. Chem. Phys. **68**, 2053 (1978).
[13] H. Nakatsuji and K. Hirao, J. Chem. Phys. **68**, 4279 (1978).
[14] H. Nakatsuji, Chem. Phys. Lett. **59**, 362 (1978).
[15] H. Nakatsuji, Chem. Phys. Lett. **67**, 329 (1979).
[16] H. Nakatsuji, Chem. Phys. Lett. **67**, 334 (1979).
[17] H. Nakatsuji and K. Hirao, Int. J. Quantum Chem. **20**, 1301 (1981).
[18] H. Nakatsuji, K. Ohta, and K. Hirao, J. Chem. Phys. **75**, 2952 (1981).
[19] H. Nakatsuji, K. Ohta, and T. Yonezawa, J. Phys. Chem. **87**, 3068 (1983).
[20] H. Nakatsuji, in *Computational Chemistry: Reviews of Current Trends*, edited by J. Leszczyński (World Scientific, Singapore, 1997), Vol. 2, pp. 62–124, and references therein.
[21] H. Nakatsuji, Bull. Chem. Soc. Jpn. **78**, 1705 (2005), and references therein.
[22] H. Monkhorst, Int. J. Quantum Chem., Quantum Chem. Symp. **11**, 421 (1977).
[23] E. Dalgaard and H. Monkhorst, Phys. Rev. A **28**, 1217 (1983).
[24] D. Mukherjee and P. K. Mukherjee, Chem. Phys. **39**, 325 (1979).
[25] M. Takahashi and J. Paldus, J. Chem. Phys. **85**, 1486 (1986).
[26] G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).
[27] G. E. Scuseria, A. C. Scheiner, T. J. Lee, J. E. Rice, and H. F. Schaefer III, J. Chem. Phys. **86**, 2881 (1987).
[28] P. Piecuch and J. Paldus, Int. J. Quantum Chem. **36**, 429 (1989).
[29] H. Koch and P. Jørgensen, J. Chem. Phys. **93**, 3333 (1990).
[30] H. Koch, H. J. Aa. Jensen, P. Jørgensen, and T. Helgaker, J. Chem. Phys. **93**, 3345 (1990).
[31] S. Hirata, J. Chem. Phys. **121**, 51 (2004).
[32] C. E. Smith, R. A. King, and T. D. Crawford, J. Chem. Phys. **122**, 054110 (2005).
[33] M. Włoch, J. R. Gour, K. Kowalski, and P. Piecuch, J. Chem. Phys. **122**, 214107 (2005).
[34] J. R. Gour, P. Piecuch, and M. Włoch, J. Chem. Phys. **123**, 134113 (2005).
[35] J. R. Gour, P. Piecuch, and M. Włoch, Int. J. Quantum Chem. **106**, 2854 (2006).
[36] M. Nooijen and R. J. Bartlett, J. Chem. Phys. **102**, 3629 (1995).
[37] M. Nooijen and R. J. Bartlett, J. Chem. Phys. **102**, 6735 (1995).
[38] R. J. Bartlett and J. F. Stanton, in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd (VCH, New York, 1994), Vol. 5, pp. 65–169.
[39] M. Nooijen and J. G. Snijders, Int. J. Quantum Chem., Quantum Chem. Symp. **26**, 55 (1992).
[40] M. Nooijen and J. G. Snijders, Int. J. Quantum Chem. **48**, 15 (1993).
[41] J. F. Stanton and J. Gauss, J. Chem. Phys. **101**, 8938 (1994).
[42] H. Nakatsuji, Chem. Phys. Lett. **177**, 331 (1991).
[43] H. Nakatsuji and M. Ehara, J. Chem. Phys. **98**, 7179 (1993).
[44] H. Nakatsuji, M. Ehara, and T. Momose, J. Chem. Phys. **100**, 5821 (1994).
[45] M. Ishida, K. Toyota, M. Ehara, M. J. Frisch, and H. Nakatsuji, J. Chem. Phys. **120**, 2593 (2004).
[46] S. Hirata, M. Nooijen, and R. J. Bartlett, Chem. Phys. Lett. **328**, 459 (2000).
[47] M. Kamiya and S. Hirata, J. Chem. Phys. **125**, 074111 (2006).
[48] I. Lindgren and D. Mukherjee, Phys. Rep. **151**, 93 (1987).
[49] D. Mukherjee and S. Pal, Adv. Quantum Chem. **20**, 291 (1989).
[50] B. Jeziorski and H. J. Monkhorst, Phys. Rev. A **24**, 1668 (1981).
[51] J. Paldus, in *Methods in Computational Molecular Physics*, NATO Advanced Study Institute, Series B: Physics, edited by S. Wilson and G. H. F. Diercksen (Plenum, New York, 1992), Vol. 293, pp. 99–194.
[52] J. Paldus and X. Li, Adv. Chem. Phys. **110**, 1 (1999).
[53] P. Piecuch and K. Kowalski, Int. J. Mol. Sci. **3**, 676 (2002).
[54] K. Kowalski and P. Piecuch, Phys. Rev. A **61**, 052506 (2000).
[55] J. Paldus, P. Piecuch, L. Pylypow, and B. Jeziorski, Phys. Rev. A **47**, 2738 (1993).
[56] K. Kowalski and P. Piecuch, Int. J. Quantum Chem. **80**, 757 (2000).
[57] P. Piecuch, R. Toboła, and J. Paldus, Chem. Phys. Lett. **210**, 243 (1993).
[58] P. Piecuch and J. Paldus, Phys. Rev. A **49**, 3479 (1994).
[59] K. Jankowski, J. Paldus, I. Grabowski, and K. Kowalski, J. Chem. Phys. **97**, 7600 (1992); **101**, 1759(E) (1994).
[60] K. Jankowski, J. Paldus, I. Grabowski, and K. Kowalski, J. Chem. Phys. **101**, 3085 (1994).
[61] K. R. Shamasundar and S. Pal, J. Chem. Phys. **114**, 1981 (2001); **115**, 1979(E) (2001).
[62] P. Piecuch and J. I. Landman, Parallel Comput. **26**, 913 (2000).
[63] K. Kowalski and P. Piecuch, Chem. Phys. Lett. **334**, 89 (2001).
[64] K. Kowalski and P. Piecuch, J. Mol. Struct.: THEOCHEM **547**, 191 (2001).
[65] K. Kowalski and P. Piecuch, Mol. Phys. **102**, 2425 (2004).
[66] X. Li and J. Paldus, J. Chem. Phys. **119**, 5320 (2003).
[67] X. Li and J. Paldus, J. Chem. Phys. **119**, 5334 (2003).
[68] X. Li and J. Paldus, J. Chem. Phys. **119**, 5346 (2003).
[69] X. Li and J. Paldus, J. Chem. Phys. **120**, 5890 (2004).
[70] M. Musiał and R. J. Bartlett, J. Chem. Phys. **121**, 1670 (2004).
[71] M. Musiał, L. Meissner, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **122**, 224110 (2005).
[72] M. Hanrath, J. Chem. Phys. **123**, 084102 (2005).
[73] K. Kowalski and P. Piecuch, J. Chem. Phys. **115**, 643 (2001).
[74] K. Kowalski and P. Piecuch, Chem. Phys. Lett. **347**, 237 (2001).
[75] S. A. Kucharski, M. Włoch, M. Musiał, and R. J. Bartlett, J. Chem. Phys. **115**, 8263 (2001).
[76] M. Kállay and J. Gauss, J. Chem. Phys. **121**, 9257 (2004).
[77] H. Koch, O. Christiansen, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. **244**, 75 (1995).
[78] O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. **103**, 7429 (1995).
[79] O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. **105**, 1451 (1996).
[80] O. Christiansen, H. Koch, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. **256**, 185 (1996).
[81] P. Piecuch, K. Kowalski, I. S. O. Pimienta, and M. J. McGuire, Int. Rev. Phys. Chem. **21**, 527 (2002).
[82] K. Kowalski and P. Piecuch, J. Chem. Phys. **120**, 1715 (2004).
[83] P. Piecuch, K. Kowalski, I. S. O. Pimienta, P.-D. Fan, M. Lodriguito, M. J. McGuire, S. A. Kucharski, T. Kuś, and M. Musiał, Theor. Chem. Acc. **112**, 349 (2004).
[84] M. Musiał and R. J. Bartlett, J. Chem. Phys. **119**, 1901 (2003).
[85] M. Musiał, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **118**, 1128 (2003).
[86] M. Musiał and R. J. Bartlett, Chem. Phys. Lett. **384**, 210 (2004).
[87] Y. J. Bomble, J. C. Saeh, J. F. Stanton, P. G. Szalay, M. Kállay, and J. Gauss, J. Chem. Phys. **122**, 154107 (2005).
[88] J. R. Gour and P. Piecuch, J. Chem. Phys. **125**, 234107 (2006).
[89] K. B. Ghose, P. Piecuch, and L. Adamowicz, J. Chem. Phys. **103**, 9331 (1995).
[90] P. Piecuch, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **110**, 6103 (1999).
[91] P. Piecuch, S. A. Kucharski, and V. Špirko, J. Chem. Phys. **111**, 6679 (1999).
[92] K. Kowalski and P. Piecuch, Chem. Phys. Lett. **344**, 165 (2001).
[93] H. Nakatsuji, Chem. Phys. **75**, 425 (1983).
[94] H. Nakatsuji, J. Ushio, and T. Yonezawa, Can. J. Chem. **63**, 1857 (1985).
[95] R. J. Buenker and S. D. Peyerimhoff, Theor. Chim. Acta **35**, 33 (1974).
[96] R. J. Buenker and S. D. Peyerimhoff, Theor. Chim. Acta **39**, 217 (1975).
[97] P. J. Bruna and S. D. Peyerimhoff, Adv. Chem. Phys. **67**, 1 (1998).
[98] N. Oliphant and L. Adamowicz, Int. Rev. Phys. Chem. **12**, 339 (1993).
[99] P. Piecuch, N. Oliphant, and L. Adamowicz, J. Chem. Phys. **99**, 1875 (1993).
[100] P. Piecuch and L. Adamowicz, J. Chem. Phys. **100**, 5792 (1994).
[101] P. Piecuch and L. Adamowicz, Chem. Phys. Lett. **221**, 121 (1994).
[102] P. Piecuch and L. Adamowicz, J. Chem. Phys. **102**, 898 (1995).

[103] K. B. Ghose and L. Adamowicz, J. Chem. Phys. **103**, 9324 (1995).
[104] V. Alexandrov, P. Piecuch, and L. Adamowicz, J. Chem. Phys. **102**, 3301 (1995).
[105] K. B. Ghose, P. Piecuch, S. Pal, and L. Adamowicz, J. Chem. Phys. **104**, 6582 (1996).
[106] L. Adamowicz, P. Piecuch, and K. B. Ghose, Mol. Phys. **94**, 225 (1998).
[107] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 8490 (2000).
[108] K. Kowalski, S. Hirata, M. Włoch, P. Piecuch, and T. L. Windus, J. Chem. Phys. **123**, 074319 (2005).
[109] P. Piecuch, S. Hirata, K. Kowalski, P.-D. Fan, and T. L. Windus, Int. J. Quantum Chem. **106**, 79 (2006).
[110] L. Adamowicz, J.-P. Malrieu, and V. V. Ivanov, J. Chem. Phys. **112**, 10075 (2000).
[111] V. V. Ivanov and L. Adamowicz, J. Chem. Phys. **112**, 9258 (2000).
[112] D. I. Lyakh, V. V. Ivanov, and L. Adamowicz, J. Chem. Phys. **122**, 024108 (2005).
[113] J. Olsen, J. Chem. Phys. **113**, 7140 (2000).
[114] J. W. Krogh and J. Olsen, Chem. Phys. Lett. **344**, 578 (2001).
[115] M. Kállay, P. G. Szalay, and P. G. Surján, J. Chem. Phys. **117**, 980 (2002).
[116] L. V. Slipchenko and A. I. Krylov, J. Chem. Phys. **123**, 084107 (2005).
[117] P.-D. Fan and S. Hirata, J. Chem. Phys. **124**, 104108 (2006).
[118] A. Kohn and J. Olsen, J. Chem. Phys. **125**, 174110 (2006).
[119] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.
[120] L. Meissner and R. J. Bartlett, J. Chem. Phys. **102**, 7490 (1995).
[121] D. Mukhopadhyay, S. Mukhopadhyay, R. Chaudhuri, and D. Mukherjee, Theor. Chim. Acta **80**, 441 (1991).
[122] L. Meissner and R. J. Bartlett, J. Chem. Phys. **94**, 6670 (1991).
[123] J. F. Stanton, J. Chem. Phys. **101**, 8928 (1994).
[124] L. Meissner, J. Chem. Phys. **108**, 9227 (1998).
[125] J. Olsen, A. M. Sánchez de Merás, H. J. Aa. Jensen, and P. Jørgensen, Chem. Phys. Lett. **154**, 380 (1989).
[126] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).
[127] P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**, 213 (1973).
[128] M. W. Schmidt, K. K. Baldridge, J. A. Boatz *et al.*, J. Comput. Chem. **14**, 1347 (1993).
[129] See EPAPS Document No. E-JCPSA6-126-311716 for the ground- and excited-state energies of CH and OH obtained with the EA/IP EOMCC, SAC-CI, truncated CI, and full CI methods, and the corresponding CC, SAC, and full CI ground-state energies of CH^+ and OH^- at several internuclear separations. This document can be reached via a direct link in the online article's HTML reference section or via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html).
[130] M. Zachwieja, J. Mol. Spectrosc. **170**, 285 (1995).
[131] P. Piecuch, M. Włoch, J. R. Gour, and A. Kinal, Chem. Phys. Lett. **418**, 463 (2005).
[132] P. Piecuch and M. Włoch, J. Chem. Phys. **123**, 224105 (2005).
[133] K. P. Huber and G. Herzberg, Constants of Diatomic Molecules, data prepared by J. W. Gallagher and R. D. Johnson III, in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, June 2005, edited by P. J. Linstrom and W. G. Mallard, National Institute of Standards and Technology, Gaithersburg, MD, 20899 (http://webbook.nist.gov).