JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

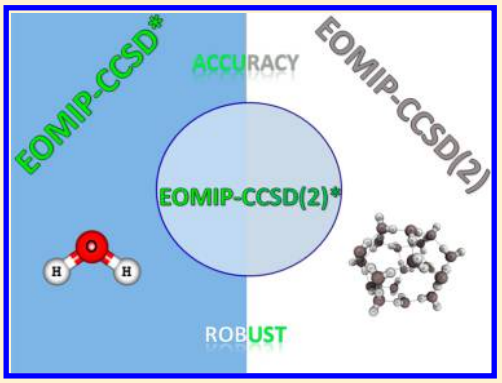
EOMIP-CCSD(2)*: An Efficient Method for the Calculation of Ionization Potentials

Achintya Kumar Dutta, Nayana Vaval, and Sourav Pal*

Physical Chemistry Division, CSIR−National Chemical Laboratory, Pune-411008, India

Supporting Information

ABSTRACT: A new approximation within the domain of EOMIP-CC method is proposed. The proposed scheme is based on the perturbative truncation of the similarity transformed effective Hamiltonian matrix. We call it the EOMIP-CCSD(2)* method, which scales as noniterative N^6 and its storage requirement is very less, compared to the conventional EOMIP−CCSD method. The existing EOMIP-CCSD(2) method has a tendency to overestimate the ionization potential (IP) values. On the other hand, our new strategy corrects for the problem of such an overestimation, which is evident from the excellent agreement achieved with the experimental values. Furthermore, not only the ionization potential but also geometry and IR frequencies of problematic double radicals are estimated correctly, and the results are comparable to the CCSD(T) method, obviously at lesser computational cost. The EOMIP-CCSD(2)* method works even for the core ionization and satellite IP, where the earlier EOMIP-CCSD(2) approximation dramatically fails.



1. INTRODUCTION

The equation-of-motion coupled cluster method has been very successful over the years for the calculation of excitation energies (EE),[1] ionization potential (IP),[2] and electron affinities (EA)[3] as a direct difference of energy. It addresses both the dynamic and nondynamic parts of the electron correlation in a balanced manner and provides a "black box" methodology to describe different Fock space sectors of interest, without going into the complications of the corresponding multireference coupled cluster theory.[4]

The method is generally used in singles and doubles excitation approximation (EOM-CCSD),[1] which scales as N^6 power of the basis set. On the other hand, the storage requirement is similar to that of the standard single-reference coupled cluster method in the same level of truncation (CCSD), which restricts the applicability of the EOMCC method only to small atomic and molecular systems in a reasonable basis. Therefore, it is necessary to have a scheme in the EOMCC framework that not only scales lower with the number of basis function but also requires less storage space to make it applicable to large systems. The coupled cluster method and many-body perturbation theory (MBPT)[5] has an innate relationship to each other. Therefore, the most obvious way of approximating the coupled cluster method is on the perturbation order. Nooijen and Snijders[6] were the first to propose the idea of replacing coupled cluster T amplitudes by MBPT(2) amplitudes. They applied the scheme in the context of the ionization problem, which makes the calculations of ionization potential an N^5 scaling method. It requires less storage space, since four particle intermediates are absent from the algebraic expressions. However, the method is very successful for the calculation of the IPs but it does not provide a straightforward definition of the total energy. As a result, the method is not suitable for the calculation of property. Stanton and Gauss[7] latter generalized this approach to provide a hierarchical approximation to the standard EOM-CCSD method. They used the term EOM-CCSD(n), where n denotes the order in perturbation of the ground state and at large values of n, the EOM-CCSD(n) method converges to standard EOM-CCSD method. It is capable of calculating the difference of energy with reasonable accuracy. Furthermore, the method has the advantage of a clear definition of total energy, which makes it suitable for the final state property calculations. The lowest order approximation to EOM-CCSD(n) is EOM-CCSD(2), where the ground state wave function is the first-order perturbed wave function, which corresponds to the MBPT(2) energy as the ground-state energy. The EOM-CCSD(2) approximation was originally implemented by Stanton and Gauss[7] for the ionization problem (EOMIP-CCSD(2)) and the excitation energies (EOMEE-CCSD(2)). Similar developments were pursued by Dutta et al. in the context of electron affinity[8] and spin flip variants[9] of the EOMCC method. Recently, Pal and co-workers[10] have shown that the EOMIP-CCSD(2) method can be used to calculate geometry and IR frequency of large doublet radicals with accuracy comparable to that of the standard EOMIP-CCSD method. However, Dutta et al.[11] have shown that, although the EOMIP-CCSD(2) method is very

Received: October 20, 2014
Published: May 14, 2015

DOI: 10.1021/ct500927h
J. Chem. Theory Comput. 2015, 11, 2461−2472

good for the study of final state properties, it is not particularly suitable for the calculation of IP itself. The missing relaxation effect due to the missing T_1 amplitudes leads to a systematic overestimation of IP in EOMIP-CCSD(2) method. However, the advantage of low computational expenses inherent to the idea of truncating the similarity transformed Hamiltonian on perturbation order makes the method too attractive to be discarded on the above ground. However, at the same time, it is necessary to pursue new theoretical developments within the framework of EOMIP-CCSD(2) method to correct for the problem of the overestimation of IP values. The objective of this paper is to suggest a modification of the standard EOMIP-CCSD(2) method that can account for the missing relaxation effect without significantly increasing the computational cost.

The paper is organized as follows. The next section (section 2) gives the theory and computational details of the new method. The numerical performance of the new method is discussed in section 3. The concluding remarks are provides in section 4.

2. THEORY

In the EOMIP-CC framework, the final states are obtained by diagonalizing the similarity transformed Hamiltonian in $(N-1)$ electron space.

$$\bar{H} = e^{-T} H e^{T} = (H e^{T})_c \tag{1}$$

In the eigenvalue equation form, this can be written as

$$[\bar{H}, \hat{\Omega}]|\Phi_0\rangle = (\bar{H}\Omega)_c|\Phi_0\rangle = \omega_k \hat{\Omega}_k|\Phi_0\rangle \tag{2}$$

where ω_k is the IP value corresponding to the kth state. $\hat{\Omega}_k$ is the corresponding EOM operator and, for the IP problem, it has the following form:

$$\begin{aligned}\Omega_k^{\mathrm{IP}} &= R_1 + R_2 + \cdots \\ &= \sum_i R_i(k) i + \sum_{i>j,a} R_{ij}^{a}(k) \hat{a}^{\dagger}\hat{j}\hat{i} + \cdots\end{aligned} \tag{3}$$

Now eq 2 gives an exact expression for the IP, when the ground state is exact, i.e.,

$$\bar{H}|\Phi_0\rangle = E_0|\Phi_0\rangle \tag{4}$$

Since $\bar{H}$ is non-Hermitian, there exist different right (R) and left (L) eigenvectors, which are biorthogonal and can be normalized to satisfy

$$L_k R_l = \delta_{kl} \tag{5}$$

The resulting method is equivalent to the (0,1) sector of the Fock space multireference coupled cluster (FSMRCC) method for the principal ionizations.[12]

The main source of error in the EOMIP-CCSD(2) approximation, as pointed out by Dutta et al.,[11] is the missing relaxation effect due to truncated T amplitudes, which cannot be compensated by the R_1 and R_2 operators. A straightforward way to account for the missing relaxation effect is to include higher-order terms in the EOM matrix. The full inclusion of R_3 operator would increase the scaling to iterative N^7, which is not feasible, except for very small molecules. However, it is possible to perform a selective inclusion of R_3 in a noniterative way with a N^6 scaling. Among the various possible schemes available for noniterative inclusion of R_3 operator, we have followed the scheme described by Stanton and co-workers.[13−15]

(a). EOMIP-CCSD*. Following Löwdin's partitioning technique,[16] eq 2 can be partitioned into P and Q space, where P represents the principal configuration space, and Q represents its orthogonal complement.

$$\begin{bmatrix} \bar{H}_{pp} & \bar{H}_{pq} \\ \bar{H}_{qp} & \bar{H}_{qq} \end{bmatrix} \begin{bmatrix} R_p \\ R_q \end{bmatrix} = \omega \begin{bmatrix} R_p \\ R_q \end{bmatrix} \tag{6}$$

and

$$[L_p \quad L_q] \begin{bmatrix} \bar{H}_{pp} & \bar{H}_{pq} \\ \bar{H}_{qp} & \bar{H}_{qq} \end{bmatrix} = [L_p \quad L_q]\omega \tag{7}$$

where $R_p(L_p)$ and $R_q(L_q)$ represent the projection of the right (left) eigenvector on P and Q spaces.

Expanding eq 6, we get

$$\bar{H}_{pp}R_p + \bar{H}_{pq}R_q = \omega R_p \tag{8}$$

$$\bar{H}_{qp}R_p + \bar{H}_{qq}R_q = \omega R_q \tag{9}$$

Rearranging eq 9 yields

$$R_q = [\omega - \bar{H}_{qq}]^{-1}\bar{H}_{qp}R_p \tag{10}$$

Inserting R_q back into eq 8, we get

$$\bar{H}_{\mathrm{eff}}R_p \equiv (\bar{H}_{pp} + \bar{H}_{pq}[\omega - \bar{H}_{qq}]^{-1}\bar{H}_{qp})R_p = \omega R_p \tag{11}$$

Projecting eq 11 with L_p gives

$$\begin{aligned}\langle L_p|\bar{H}_{\mathrm{eff}}|R_p\rangle &\equiv \langle L_p|[\bar{H}_{pp} + \bar{H}_{pq}[\omega - \bar{H}_{qq}]^{-1}\bar{H}_{qp}]|R_p\rangle \\ &= \omega\langle L_p|R_p\rangle\end{aligned} \tag{12}$$

The eigenvalues of H_{eff} are solely defined in the P space, for the first several eigenvalues. The exact eigenvalue ω can be expressed as the sum of zeroth-order energy ω_0, as of yet undetermined, and an energy correction $\Delta\omega$.

The operator inverse in eq 12 can be expressed as

$$\begin{aligned}[\omega - \bar{H}_{qq}]^{-1} &= [\omega_0 + \Delta\omega - \bar{H}_{qq}^{[0]} - \bar{H}_{qq}^{[1]} - \bar{H}_{qq}^{[2]}\cdots]^{-1} \\ &= [(\omega_0 - \bar{H}_{qq}^{[0]})(1 - [\omega_0 - \bar{H}_{qq}^{[0]}]^{-1}[\bar{H}_{qq}^{[1]} + \bar{H}_{qq}^{[2]}\cdots - \Delta\omega])]^{-1} \\ &\equiv [(\omega_0 - \bar{H}_{qq}^{[0]})(1 - [\omega_0 - \bar{H}_{qq}^{[0]}]^{-1}[V_{qq} - \Delta\omega])]^{-1}\end{aligned} \tag{13}$$

where

$$V_{qq} = \bar{H}_{qq}^{[1]} + \bar{H}_{qq}^{[2]} + \cdots$$

Equation 13 can be expanded in an inverse series:

$$\begin{aligned}[\omega - \bar{H}_{qq}]^{-1} &= [\omega_0 - \bar{H}_{qq}^{[0]}]^{-1} \\ &+ [\omega_0 - \bar{H}_{qq}^{[0]}]^{-1}(V_{qq} - \Delta\omega)[\omega_0 - \bar{H}_{qq}^{[0]}]^{-1} \\ &+ [\omega_0 - \bar{H}_{qq}^{[0]}]^{-1}(V_{qq} - \Delta\omega)[\omega_0 - \bar{H}_{qq}^{[0]}]^{-1}(V_{qq} - \Delta\omega)[\omega_0 - \bar{H}_{qq}^{[0]}]^{-1} \\ &+ \cdots\end{aligned} \tag{14}$$

Now, the energy correction to EOMIP-CCSD can be derived by defining p as $p \equiv h \cup 2hp$, $\bar{H}_{pp}$ is taken as the zeroth-order Hamiltonian, and ω_0 can be taken as the EOMIP-CCSD energy. Equation 12 can be written as

$$\langle L_p|\bar{H}_{\mathrm{eff}}|R_p\rangle = E_{\mathrm{EOMIP}} + \Delta\omega = E_{\mathrm{EOMIP}} + \langle \delta L|D|\delta R\rangle \tag{15}$$

and

DOI: 10.1021/ct500927h
J. Chem. Theory Comput. 2015, 11, 2461−2472

$$\langle \delta L| = \langle L|\bar{H}_{pq}[\omega_0 - H^{[0]}_{qq}]^{-1} \tag{16}$$

$$|\delta R\rangle = [\omega_0 - H^{[0]}_{qq}]^{-1} H_{qp}|R\rangle \tag{17}$$

The similarity transformed Hamiltonian can be expressed in perturbational orders:

$$\bar{H} = (He^T)_c = \bar{H}^{[1]} + \bar{H}^{[2]} + \bar{H}^{[3]} + \bar{H}^{[4]} + \cdots \tag{18}$$

In the above expression, the hole−hole and the particle−particle block of the Fock matrix is treated as zeroth-order and the rest of the H is treated as first-order. The T_1 and T_2 amplitudes for the reference state are taken as second order and first order in correlation, respectively. The projection of L and R on $1h$ determinants (L_h and R_h) are taken as zeroth order and the projection on $2h1p$ determinants (L_{2h1p} and R_{2h1p}) are taken as first order in correlation. With this definition, eq 12 can be written as

$$\langle L_p|\bar{H}_{\text{eff}}|R_p\rangle = \langle L_p|\bar{H}_{pq}[\omega_0 - E_0 - \bar{H}^0_{qq}]^{-1}\bar{H}_{qq}[\omega_0 - E_0 - \bar{H}^0_{qq}]^{-1}\bar{H}_{qp}|R_p\rangle \tag{19}$$

Equation 19 contains terms that are fourth order and higher in correlation. Because of their negligible contribution and high computational cost associated with their evaluation, eq 19 has not been considered. Instead, eq 15, which has terms only up to third order in perturbation, has been used to evaluate the energy correction. The elements of $\bar{H}_{pq}$ and $\bar{H}_{qp}$ are divided according to hole−particle contribution and only the terms with the lowest nonvanishing contributions have been considered.

Following the above guide line, Stanton and Gauss[13,14] have shown that the only surviving contributions are those which connect the reference determinant ($|0\rangle$) to determinants generated by $3h2p$ operators (i.e., those obtained by excitation of two electrons and the removal of the third).

The spin−orbital notation of eqs 16 and 17, as described in ref 14, are given below:

$$D^{ijk}_{ab} l^{ijk}_{ab} = P(ijk) l^k \langle ab||ij\rangle - P(ijk) \sum_e l^{ij}_e \langle ek||ab\rangle - P(ab)P(ijk) \sum_m l^{mk}_a \langle ij||mb\rangle \tag{20}$$

$$\begin{aligned} D^{ab}_{ijk} r^{ab}_{ijk} = {} & -P(ijk) \sum_e r^e_{ij} \langle ab||ek\rangle - P(ab)P(ijk) \sum_m r^a_{mk} \langle mb||ij\rangle \\ & - P(ab)P(ijk) \sum_{me} r_m t^{ae}_{ij} \langle mb||ke\rangle + P(kji) \sum_{mn} r_m t^{ab}_{in} \langle mn||kj\rangle \\ & - P(ijk) \sum_m r^e_{mk} [P(ij) \sum_{en} t^{ab}_{nj} \langle mn||ei\rangle] \\ & + P(ijk)P(ab) \sum_{mef} r^e_{mk} t^{fb}_{ij} \langle ma||ef\rangle \\ & + \frac{1}{2} P(ijk)P(ab) \sum_{mne} r^a_{mn} t^{eb}_{ij} \langle mn||ek\rangle \end{aligned} \tag{21}$$

$$D^{ijk}_{ab} = D^{ab}_{ijk} = \omega_0 - E_0 + f_{ii} + f_{jj} + f_{kk} - f_{aa} - f_{bb} \tag{22}$$

However, the method has some ambiguity,[15] because of the inclusion of second-order three-body contributions, the final three terms in eq 21, in the Hamiltonian. At the same time, the following second-order two-body contributions are excluded from the model.

$$\begin{aligned} D^{ab}_{ijk} r^{ab}_{ijk} \Leftarrow {} & P(ijk) \sum_m t^{eb}_{mk} [P(ij) \sum_{en} r^a_{nj} \langle mn||ei\rangle] \\ & - P(ijk)P(ab) \sum_{mef} t^{eb}_{mk} r^e_{mk} \langle ma||ef\rangle \\ & - \frac{1}{2} P(ijk)P(ab) \sum_{mne} t^{ab}_{mn} r^e_{ij} \langle mn||ek\rangle \end{aligned} \tag{23}$$

Both sets of contractions, the last three terms of eq 21 and the three appearing in eq 23 are of overall fourth-order contribution to energy.

Saeh and Stanton[15] have shown, from limited numerical analysis, the energy corrections obtained when both classes of terms are included to be very close to those obtained when both classes are excluded. So they redefined the EOMIP-CCSD* model to express eq 21 as the following:

$$\begin{aligned} D^{ab}_{ijk} r^{ab}_{ijk} = {} & -P(ijk) \sum_e r^e_{ij} \langle ab||ek\rangle - P(ab)P(ijk) \sum_m r^a_{mk} \langle mb||ij\rangle \\ & - P(ab)P(ijk) \sum_{me} r_m t^{ae}_{ij} \langle mb||ke\rangle + P(kji) \sum_{mn} r_m t^{ab}_{in} \langle mn||kj\rangle \end{aligned} \tag{24}$$

The resulting method gives IP values that are correct up to at least third-order. A similar philosophy has been recently employed by Dutta et al.[17] to derive perturbative triples correction to the Fock space multireference coupled cluster method, which gives results that are almost identical to those ontained using the EOMIP-CCSD* method.

For a more detailed discussion on the theory of EOMIP-CCSD* method, the readers are advised to consult the original implementation papers.[13−15]

The fT and dT scheme for including noniterative triples in EOMIP-CCSD by Krylov and co-workers[18] were derived following a similar philosophy.

(b). EOMIP-CCSD(2)*. Following the original EOMIP-CCSD(2) scheme reported by Stanton and Gauss,[7] the CC amplitudes can be approximated by the MBPT(2) amplitudes for a second-order truncated similarity transformed Hamiltonian ($H^{[2]}$)

$$\bar{H} = (He^T)_c \approx (He^{T'})_c \tag{25}$$

where the perturbative approximation to the T amplitudes can be written as

$$T_1' = \frac{f_{ia}}{\varepsilon_i - \varepsilon_a}$$

$$T_2' = \frac{\langle ab||ij\rangle}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} \tag{26}$$

T_1 is zero for restricted closed-shell and unrestricted MBPT(2) reference. Using these T' amplitudes, one can generate a modified similarity transformed Hamiltonian $\bar{H}'$, which can be used as the reference for subsequent EOMIP calculations. After solving for the right vector, one must solve for the left vector, while the correction to the right and left vectors is constructed in a noniterative fashion using eqs 20, 22, and 24. Finally, the energy correction is calculated using eq 15. We call this new approximation EOMIP-CCSD(2)*.

The original EOMIP-CCSD(2) method scales as iterative N^5. Now, the energy correction as described in eqs 20 and 24 scales as noniterative N^6. Thus, overall the EOMIP-CCSD(2)* method is noniterative N^6 scaling, compared to the iterative

DOI: 10.1021/ct500927h
J. Chem. Theory Comput. 2015, 11, 2461−2472

N^6 scaling of standard the EOMIP-CCSD method. However, the situation is more favorable than the above statement indicates. The most expensive terms arising due to the $3h2p$ corrections in our EOMIP-CCSD(2)* method scale as $n_h^3 n_p^3$ and $n_h^4 n_p^2$, where n_p and n_h represent the number of unoccupied and occupied orbitals, respectively, in the reference Hartree−Fock determinant. The reference state CCSD calculation involves an iterative step that scales with $n_h^2 n_p^4$. Now, for typical applications in a reasonable basis set, n_p is much greater than n_h. Therefore, the cost of a single EOMIP-CCSD(2)* calculation should be substantially less than a single iteration of the reference state CCSD equations.

Now, let us consider the storage requirements. In the EOMIP-CCSD(2)* method, there are no four particle intermediates contributing in the energy correction. The four particle integrals can indeed contract with T_1 amplitudes to contribute to the 3-particle−1-hole intermediate. However, a first-order perturbed ground-state reference for the RHF or UHF reference case will lead to zero T_1 amplitudes. Therefore, the EOMIP-CCSD(2)* method does not contain any four particle terms, which leads to a significant decrease in the storage requirements. However, the storage requirement has increased from that in the original EOMIP-CCSD(2) approximation, because of the presence of 3-particle−1-hole intermediates in the energy correction part, which were absent in the original EOMIP-CCSD(2) approximation. Thus, both in terms of CPU scaling and storage requirements, the EOMIP-CCSD(2)* method lies between the EOMIP-CCSD(2) method and the standard EOMIP-CCSD method.

The method can be trivially implemented by using $\bar{H}$ modified according to eq 25 in any EOMIP-CCSD* code. In the present work, the implementation of EOMIP-CCSD(2) and EOMIP-CCSD* in the software CFOUR[19] has been used in a handy way.

(c). Computational Details. All the EOMIP-CCSD(2)* calculations, as well as other EOMCC and single-reference coupled cluster calculations, were performed using the quantum chemistry package CFOUR.[19] The valence IP values for test molecules are calculated using a hierarchy of Dunning's correlation consistent cc-pVXZ (X = D, T, and Q) basis sets.[20] The core−valence correlation cc-pCVXZ (X = D, T, and Q) basis sets[21] were used for the calculation of the core IP. The calculated IP values in the EOMIP-CCSD(2) method were extrapolated using the formula presented by Pal and co-workers[11] in their benchmark study. Experimental geometries were used in all the cases, except for thymine, where the B3LYP/6-311++G** optimization geometry was used. All the geometries are provided in the Supporting Information. The structure optimization and frequency calculations of doublet radicals were performed using the numerical gradient technique.

3. RESULTS AND DISCUSSION

Table 1 presents the wall times for the EOMIP-CCSD(2) method, along with the EOMIP-CCSD(2) and the standard EOMIP-CCSD methods, for a series of water clusters ($(H_2O)_n$, where n = 1−8). All the calculations were performed using a single core in an i7 desktop computer with 3.40 GHz CPU speed and 16 GB of RAM. It can be seen that the CPU timing of the EOMIP-CCSD(2)* method lies between that of the standard EOMIP-CCSD and that of EOMIP-CCSD(2) method, as expected from the theoretical foundation of the method described in the previous section. It can be seen that as the system size increases, the EOMIP-CCSD(2)* starts to become progressively less expensive, compared to the standard EOMIP-CCSD method.

Table 1. Wall Timings[a,b] for the EOMIP-CCSD, EOMIP-CCSD(2) and EOMIP-CCSD(2)* Method in the cc-pVDZ Basis Set

number of H_2O units	Wall Timing (s)		
	EOMIP-CCSD	EOMIP-CCSD(2)	EOMIP-CCSD(2)*
1	1.16	1.13	1.33
2	11.54	2.89	4.84
3	108.88	9.88	29.42
4	490.52	30.67	70.61
5	1516.96	119.20	255.79
6	3795.23	289.19	1620.46
7	15129.76	673.03	4030.58
8	42946.41	1682.45	9436.54

[a]All calculations were performed using an i7 desktop computer with 3.40 GHz CPU speed and 16 GB of RAM. Calculations were performed on a single core. [b]Calculations were performed assuming C_1 symmetry.

(a). Valence Ionization Spectra. The performance of the EOMIP-CCSD(2)* method for valence ionization energies is benchmarked for small molecules such as N_2, H_2O, H_2CO, C_2H_2, and O_3 in a hierarchy of Dunning's correlation consistent cc-pVXZ (X = D, T, Q) basis sets (see Tables 2−6), and the results are compared with experimental results, wherever available. For the sake of comparison, we also quote the corresponding EOMIP-CCSD(2) and extrapolated EOMIP-CCSD(2) results.

Table 2 presents the valence ionization energies of the first three states of N_2. It can be seen that the EOMIP-CCSD(2) method overestimates the IP values, compared to the standard EOMIP-CCSD method in the cc-pVDZ basis. The EOMIP-CCSD(2)* method corrects for this overestimation and gives IP values, which are in superior agreement with the highly accurate EOMIP-CCSD* method. The agreement is even better than the standard EOMIP-CCSD method. The IP values in all the methods increase from cc-pVDZ to cc-pVTZ basis set. The extrapolated EOMIP-CCSD(2) method shows much improvement over the original EOMIP-CCSD(2) approximation. However, the values are inferior, compared to the EOMIP-CCSD(2)* method, which are in very good agreement with the EOMIP-CCSDT method. The IP values in all the methods increase slightly, as we go from cc-pVTZ to cc-pVQZ basis set. The EOMIP-CCSD(2)* method gives IP values which are within 0.15 eV of the experimental value,[19] and the results are in even better agreement than that in the standard EOMIP-CCSD method. EOMIP-CCSD(2)* values are also in better agreement with EOMIP-CCSDT than the EOMIP-CCSD method.

Table 3 presents the vertical ionization energies corresponding to the first three occupied orbitals of water. It can be seen that, in the cc-pVDZ basis set, all the EOMIP methods lead to very similar results. The IP values for all three states increase from cc-pVDZ to cc-pVTZ basis. Here, it should be noted that the inclusion of triples has a negligible effect on the valence IPs of water. The IP values further increase as we go from the cc-pVTZ basic set to the cc-pVQZ basis set. The experimental IP values corresponding to $1b_2$ and $3a_1$ are well reproduced at the EOMIP-CCSD(2)*/cc-pVQZ level of theory. However, the IP

DOI: 10.1021/ct500927h
J. Chem. Theory Comput. 2015, 11, 2461−2472

Table 2. Ionization Energies of N_2

	Ionization Energy						
state	EOMIP-CCSD	EOMIP-CCSD(2)	extrapolated EOMIP-CCSD(2)	EOMIP-CCSD(2)*	EOMIP-CCSD*	EOMIP-CCSDT	Exp[a]
			cc-pVDZ Basis Set				
$3\sigma_g$	15.19	15.39		15.19	15.02	15.10	15.60
$1\pi_u$	16.96	17.14		16.59	16.45	16.67	16.98
$2\sigma_u$	18.45	18.56		18.44	18.35	18.34	18.78
			cc-pVTZ Basis Set				
$3\sigma_g$	15.59	15.85	15.65	15.54	15.33	15.46	15.60
$1\pi_u$	17.22	17.48	17.30	16.87	16.66	16.95	16.98
$2\sigma_u$	18.81	18.98	18.87	18.75	18.61	18.65	18.78
			cc-pVQZ Basis Set				
$3\sigma_g$	15.72	16.02	15.82	15.68	15.43	15.57	15.60
$1\pi_u$	17.34	17.64	17.46	17.00	16.75	17.05	16.98
$2\sigma_u$	18.93	19.15	19.04	18.88	18.69	18.75	18.78

[a]Data taken from ref 34.

Table 3. Ionization Energies of H_2O

	Ionization Energy (eV)						
state	EOMIP-CCSD	EOMIP-CCSD(2)	extrapolated EOMIP-CCSD(2)	EOMIP-CCSD(2)*	EOMIP-CCSD*	EOMIP-CCSDT	Exp[a]
			cc-pVDZ Basis Set				
$1b_2$	11.80	11.74		11.85	11.91	11.94	12.62
$3a_1$	14.11	14.04		14.11	14.23	14.25	14.73
$1b_1$	18.47	18.37		18.40	18.50	18.53	18.55
			cc-pVTZ Basis Set				
$1b_2$	12.40	12.43	12.43	12.39	12.39	12.46	12.62
$3a_1$	14.63	14.63	14.63	14.60	14.62	14.69	14.73
$1b_1$	18.83	18.81	18.81	18.75	18.79	18.85	18.55
			cc-pVQZ Basis Set				
$1b_2$	12.62	12.69	12.69	12.60	12.55	12.63	12.62
$3a_1$	14.82	14.87	14.87	14.79	14.76	14.84	14.73
$1b_1$	19.00	19.01	19.01	18.91	18.92	18.99	18.55

[a]Data taken from ref 34.

Table 4. Ionization Energies of H_2CO

	Ionization Energy (eV)						
state	EOMIP-CCSD	EOMIP-CCSD(2)	extrapolated EOMIP-CCSD(2)	EOMIP-CCSD(2)*	EOMIP-CCSD*	EOMIP-CCSDT	Exp[a]
			cc-pVDZ Basis Set				
$2b_2$	10.34	10.25		10.34	10.34	10.40	10.88
$1b_1$	14.29	14.18		14.15	14.15	14.27	14.5
$5a_1$	15.71	15.63		15.66	15.66	15.71	16.0
$1b_2$	17.08	17.03		16.78	16.78	16.77	16.5
			cc-pVTZ Basis Set				
$2b_2$	10.75	10.77	10.77	10.63	10.64	10.74	10.88
$1b_1$	14.57	14.58	14.58	14.32	14.33	14.52	14.5
$5a_1$	16.05	16.07	16.07	15.87	15.87	15.97	16.0
$1b_2$	17.37	17.39	17.39	17.00	17.00	17.03	16.5
			cc-pVQZ Basis Set				
$2b_2$	10.90	10.97	10.97	10.79	10.75	10.86	10.88
$1b_1$	14.69	14.77	14.77	14.47	14.42	14.62	14.5
$5a_1$	16.19	16.28	16.28	16.03	15.97	16.07	16.0
$1b_2$	17.48	17.54	17.54	17.13	17.08	17.13	16.5

[a]Data taken from ref 34.

value corresponding to the $1b_1$ state in the EOMIP-CCSD(2)* method is overestimated by 0.36 eV, compared to the experiments. Similarly, the benchmark EOMIP-CCSDT method also overestimates the IP value for the $1b_1$ state. It should be noted that the EOMIP-CCSD(2)* method gives slightly better agreement with the experiment than the EOMIP-CCSD method for all three states, although the values in both methods are very similar to each other.

The first four valence ionized states of formaldehyde are reported in Table 4. The EOMIP-CCSD(2) method agrees well with the EOMIP-CCSD values using the cc-pVDZ basis set. The EOMIP-CCSD(2)* and EOMIP-CCSD* methods

DOI: 10.1021/ct500927h
J. Chem. Theory Comput. 2015, 11, 2461−2472

Table 5. Ionization Energies of C_2H_2

	Ionization Energy (eV)						
state	EOMIP-CCSD	EOMIP-CCSD(2)	Extrapolated EOMIP-CCSD(2)	EOMIP-CCSD(2)*	EOMIP-CCSD*	EOMIP-CCSDT	Exp[a]
			cc-pVDZ Basis Set				
$^2\Pi_u$	11.33	11.35		11.07	11.08	11.23	11.49
$^2\Sigma^+_g$	16.98	17.01		16.84	16.81	16.89	16.7
$^2\Sigma^+_u$	18.88	18.86		18.72	18.75	18.78	18.7
			cc-pVTZ Basis Set				
$^2\Pi_u$	11.55	11.66	11.64	11.31	11.22	11.43	11.49
$^2\Sigma^+_g$	17.21	17.32	17.29	17.07	16.98	17.08	16.7
$^2\Sigma^+_u$	19.09	19.15	19.15	18.93	18.89	18.96	18.7
			cc-pVQZ Basis Set				
$^2\Pi_u$	11.63	11.80	11.78	11.41	11.27	11.50	11.49
$^2\Sigma^+_g$	17.30	17.46	17.43	17.18	17.04	17.15	16.7
$^2\Sigma^+_u$	19.17	19.27	19.27	19.03	18.94	19.02	18.7

[a]Data taken from ref 35.

Table 6. Ionization Energies of O_3

	Ionization Energy (eV)						
state	EOMIP-CCSD	EOMIP-CCSD(2)	extrapolated EOMIP-CCSD(2)	EOMIP-CCSD(2)*	EOMIP-CCSD*	EOMIP-CCSDT	Exp[a]
			cc-pVDZ Basis Set				
$1a_2$	12.35	12.76		12.26	11.93	12.20	12.73
$6a_1$	12.45	12.84		12.39	12.07	12.33	13.00
$3b_1$	13.11	13.52		12.88	12.61	13.12	13.54
			cc-pVTZ Basis Set				
$1a_2$	12.77	13.24	12.83	12.61	12.24	12.56	12.73
$6a_1$	12.85	13.30	12.91	12.72	12.36	12.67	13.00
$3b_1$	13.41	13.93	13.52	13.21	12.75	13.44	13.54
			cc-pVQZ Basis Set				
$1a_2$	12.97	13.49	13.08	12.81	12.40	12.74	12.73
$6a_1$	13.05	13.54	13.15	12.91	12.51	12.84	13.00
$3b_1$	13.58	14.15	13.74	13.39	12.88	13.60	13.54

[a]Data taken from ref 35.

give slightly lower IP values for the $1b_1$ and $1b_2$ states; however, the values are similar to the EOMIP-CCSD method for the other two states. Upon increasing the basis set from cc-pVDZ to cc-pVTZ, the IP values in all the methods increase considerably. The EOMIP-CCSD(2)* method gives almost the same values as that in the EOMIP-CCSD* method, and the results are slightly lower than those of the EOMIP-CCSD method using the cc-pVTZ basis set. The IP values further increase from the cc-pVTZ basis set to the cc-pVQZ basis set, and the results are within 0.1 eV of experimental values at the EOMIP-CCSD(2)*/cc-pVQZ level of theory, except for the $1b_2$ state, which is overestimated by 0.58 eV. However, the EOMIP-CCSDT method also shows similar value for the $1b_2$ state. It should be noted that the EOMIP-CCSD(2)* results are in better agreement with the experimental result[19] than both the EOMIP-CCSD method and the EOMIP-CCSD(2) method for the $1b_2$ state.

Table 5 presents the IP values corresponding to $^2\Pi_u$, $^2\Sigma^+_g$, and $^2\Sigma^+_u$ states of acetylene at different level of truncation of EOM methods. In the cc-pVDZ basis set, the IP values in the EOMIP-CCSD(2)* method give very good agreement with that in the EOMIP-CCSD* method, and the values are considerably lower than the corresponding EOMIP-CCSD and EOMIP-CCSD(2) results. The IP values in all the methods undergo a blue shift as we go from the cc-pVDZ basis set to the cc-pVTZ basis set. The EOMIP-CCSD(2)* method continues to give lower values, compared to the EOMIP-CCSD method. However, the former yields better agreement with the highly accurate EOMIP-CCSD* method. The IP values undergo a further blue shift from the cc-pVTZ basis set to the cc-pVQZ basis set. The EOMIP-CCSD(2)* method in the cc-pVQZ basis set gives very good agreement with the experimental value for the $^2\Pi_u$ state, but it overestimates the IP values for the other two states. However, the IP values are in better agreement with the experiment, compared to the EOMIP-CCSD method, which overestimates them by values as high as 0.6 eV with the cc-pVQZ basis set. The EOMIP-CCSD(2) method leads to further overestimation, and its extrapolated version does not offer any significant improvement.

Table 6 presents the ionization potentials corresponding to the first three states of ozone. The ozone ground state has considerable multireference character and is known to offer a significant challenge for all of the EOMCC methods based on a MBPT(2) reference.[8,9,11] In the cc-pVDZ basis set, the EOMIP-CCSD(2) values are much overestimated, compared to the EOMIP-CCSD values. All of the IP values undergo a blue shift from the cc-pVDZ basis set to the cc-pVTZ basis set. However, the qualitative trend remains the same. The IP values further increase from the cc-pVTZ basis set to the cc-pVQZ basis set. The EOMIP-CCSD(2) method significantly overestimates the IP values compared to the experimental values. The extrapolated version shows significant improvement over the original EOMIP-CCSD(2) approximation. The EOMIP-CCSD(2)* method and the EOMIP-CCSD method with the

DOI: 10.1021/ct500927h
J. Chem. Theory Comput. 2015, 11, 2461−2472

Table 7. Core-Ionized Energies in EOMCC Methods

	Core-Ionized Energy (eV)					
molecule	EOMIP-CCSD	EOMIP-CCSD(2)	extrapolated EOMIP-CCSD(2)	EOMIP-CCSD(2)*	EOMIP-CCSD*	Exp
			cc-pCVDZ Basis Set			
H_2O	542.69	542.81		541.17	541.06	539.75[a]
CH_4	293.18	293.31		292.22	292.22	290.86[b]
N_2	412.61	413.27		412.19	411.59	409.9[b]
HF	697.24	697.19		696.36	695.38	693.80[b]
NH_3	408.17	408.36		407.06	406.89	405.52[b]
			cc-pCVTZ Basis Set			
H_2O	541.13	541.65	541.53	540.03	539.54	539.75
CH_4	291.99	292.40	292.27	291.33	290.96	290.86
N_2	411.13	412.05	411.39	410.91	410.06	409.9
HF	695.41	695.81	695.81	694.04	693.66	693.80
NH_3	406.84	407.36	407.17	406.00	405.51	405.52
			cc-pCVQZ Basis Set			
H_2O	541.35	541.92	541.80	540.13	539.62	539.75
CH_4	291.99	292.49	292.36	290.84	290.78	290.86
N_2	411.33	412.28	411.62	410.84	409.70	409.9
HF	695.74	696.19	696.19	694.27	693.83	693.80
NH_3	406.99	407.59	407.40	406.13	405.55	405.52

[a]Values taken from ref 36. [b]Values taken from ref 37.

Table 8. Satellite IP Values in EOMCC Methods

	Satellite IP Value (eV)					
molecule	EOMIP-CCSD	EOMIP-CCSD(2)	extrapolated EOMIP-CCSD(2)	EOMIP-CCSD(2)*	EOMIP-CCSD*	Exp[a]
			cc-pVDZ Basis Set			
$N_2(^2\Sigma_u^+)$	28.80	28.45		25.10	25.11	25.51
$CO(^2\Pi)$	26.24	26.18		23.07	23.16	23.4
			cc-pVTZ Basis Set			
$N_2(^2\Sigma_u^+)$	29.57	30.17	30.17	25.76	25.26	25.51
$CO(^2\Pi)$	26.76	26.80	26.80	23.26	23.25	23.4
			cc-pVQZ Basis Set			
$N_2(^2\Sigma_u^+)$	29.83	30.43	30.43	25.87	25.31	25.51
$CO(^2\Pi)$	26.96	27.06	27.06	23.36	23.28	23.4

[a]Data taken from ref 34.

cc-pVQZ basis set show good agreement with the experimental value.

(b). Core-Ionization Spectra. The ionization of electrons from the core orbitals by high-energy radiation leads to a variety of interesting physical and chemical phenomena and it often presents a significant challenge for the conventional *ab initio* methods. Dutta et al.[11] have shown that the EOMIP-CCSD(2) method and its extrapolated version fail to model the core-ionized states. Table 7 presents core ionization energies of H_2O, CH_4, N_2, HF, and NH_3 in a hierarchy of Dunning's core valence correlation consistent cc-pCVXZ (X = D, T, and Q) basis sets. The EOMIP-CCSD(2) method, in the cc-pCVDZ basis set, significantly overestimates the IP values compared to the standard EOMIP-CCSD method. The EOMIP-CCSD(2)* method, on the other hand, gives much lower values, even compared to the EOMIP-CCSD method. The core ionization energies for all the molecules undergo a blue shift from the cc-pCVDZ basis set to the cc-pCVTZ basis set. However, the qualitative trend remains the same. The core ionization energies in all of the EOM methods further increase as we go to the cc-pCVQZ basis set. The EOMIP-CCSD* method gives the best agreement with the experiment, using the cc-pCVQZ basis set, and the results are within 0.2 eV of the experiments. The EOMIP-CCSD(2)* method gives slightly overestimated values, especially for the N_2, where the core ionization energies are overestimated by ∼1 eV. However, the error looks significantly less if percentage errors are consider, which are as small as 0.27%. At the same time, the EOMIP-CCSD method itself shows a higher error bar (∼1.5 eV) in the cc-pCVQZ basis set for the core-ionized state of N_2. Generally, the EOMIP-CCSD(2)* method gives much better agreement with the experiment than the EOMIP-CCSD method for core-ionized states for all five of the molecules studied. The original EOMIP-CCSD(2) method, on the other hand, grossly overestimates the IP values and its extrapolated version does not show any significant improvement.

(c). Satellite Peaks. The satellite IP peaks are characterized by ionization of an electron, along with simultaneous excitation of another electron from an occupied orbital to a virtual orbital. The satellite peaks are generally associated with large relaxation effects and even the standard EOMIP-CCSD method fails to offer a reasonable description. Table 8 provides satellite IP values for CO and N_2. It can be seen that the EOMIP-CCSD(2)* method in the cc-pVDZ basis set gives much lower values, compared to the EOMIP-CCSD method, although the former predicts IP values, which are in good agreement with the highly accurate EOMIP-CCSD* method for both CO and N_2.

DOI: 10.1021/ct500927h
J. Chem. Theory Comput. 2015, 11, 2461−2472

The IP values in all of the methods undergo a blue shift from the cc-pVDZ basis set to the cc-pVTZ basis set. However, the qualitative trend remains the same, except that the EOMIP-CCSD(2) method yields grossly overestimated values for the $^2\Sigma_u^+$ state of N_2 and the extrapolated version does not provide any improvement. The IP values further increase in the cc-pVQZ basis set. The EOMIP-CCSD(2)* method gives very good agreement with the experimental results for the satellite IP values of CO and N_2, and the agreement is even better than that in the standard EOMIP-CCSD method.

Here, it should be mentioned that it is not justified to come to any firm conclusion about the relative accuracy of the different EOMCC methods for satellite IP values, from the study of only two states. However, a detailed study of the satellite IP values using different truncations of the EOMIP-CC method is outside the scope of the present manuscript.

(d). Error Analysis. Table 9 provides the statistical analysis of the absolute error of valence IP in the cc-pVQZ basis set, compared to the experimental results in the different approximation of the EOMIP-CC method. It can be seen that the average and RMS error in the EOMIP-CCSD(2)* method is 0.15 and 0.24 eV, respectively. These are nearly half of the values observed in EOMIP-CCSD(2) and its extrapolated version. The errors are even less than that observed in the standard EOMIP-CCSD and EOMIP-CCSD* method.

Table 9. Analysis of Errors in IP in Different Approximation to EOMIP-CC in the cc-pVQZ Basis Set, Compared to Experimental Results

method	abs max deviation	abs min deviation	abs avg deviation	abs RMS deviation
EOMIP-CCSD	0.98	0.0	0.26	0.36
EOMIP-CCSD(2)	1.04	0.07	0.46	0.53
extrapolated EOMIP-CCSD(2)	1.04	0.07	0.35	0.43
EOMIP-CCSD(2)*	0.63	0.02	0.15	0.24
EOMIP-CCSD*	0.66	0.03	0.25	0.32

It may not be a wise idea to compare the calculated vertical IP values in the EOMIP-CCSD(2)* method with the experimental results, which are often the band maxima in the spectra. Therefore, we have also compared our results with benchmark EOMIP-CCSDT results in the same basis set to remove any scope of ambiguity. Table 10 presents statistical analysis of the errors in different hierarchical approximations of EOMIP-CC in the cc-pVQZ basis set from the EOMIP-CCSDT results. It can be seen that the absolute average and RMS error in the EOMIP-CCSD(2)* method, compared to benchmark EOMIP-CCSDT values, is 0.07 and 0.09 eV, respectively. The absolute average and RMS error are nearly one-third of that observed in the EOMIP-CCSD(2) method and the extrapolated version of the EOMIP-CCSD(2) method. The errors are even less than those observed in the standard EOMIP-CCSD method or its partial triples corrected version EOMIP-CCSD* method.

Table 10. Analysis of Errors in IP in Different Approximation to EOMIP-CC in the cc-pVQZ Basis Set, Compared to Benchmark EOMIP-CCSDT Results

method	abs max dev	abs min deviation	abs avg deviation	abs RMS deviation
EOMIP-CCSD	0.35	0.01	0.13	0.17
EOMIP-CCSD(2)	0.75	0.02	0.33	0.40
extrapolated EOMIP-CCSD(2)	0.41	0.02	0.22	0.25
EOMIP-CCSD(2)*	0.21	0.0	0.07	0.09
EOMIP-CCSD*	0.72	0.06	0.19	0.25

The EOMIP-CCSD(2)* method shows significant improvement over the original EOMIP-CCSD(2) approximation for valence, core, and satellite IP values. In the benchmark cc-pVQZ basis set (cc-pCVQZ for the core IP), the IP values are in very good agreement with the experimental values, and the results are even better than the standard EOMIP-CCSD method for most of the cases. It has been shown that a MBPT(2) truncated reference state leads to systematic underestimation of IP values in EOM-CC,[11] when R-vectors have been truncated at R_1 and R_2. On the other hand, the inclusion of perturbative triples in the target state over the CCSD ground state for IP problems in both EOMCC and FSMRCC leads to a systematic underestimation of the IP values.[17] Therefore, in a combination of both, i.e., when perturbative triples are included in the target state and the ground state kept at MBPT(2) truncation, the errors should cancel out. This is presumably the reason for the good results that are obtained in the EOMIP-CCSD(2)* method. However, it is not very straightforward to draw a quantitative explanation. The EOMIP-CCSD(2)* method also works for molecules such as ozone, where the reference state has significant multireference character (reflected by high T1 diagnosis values, as seen from Table 11), where the original EOMIP-CCSD(2) approximation fails drastically.

Table 11. T1 Diagnosis Values in the cc-pVTZ Basis Set

molecule	T1 value
N_2	0.013
H_2O	0.007
H_2CO	0.015
C_2H_4	0.011
ozone	0.028

Figure 1 provides a schematic depiction of the relative change in the position of the target state caused by the inclusion of the R_3 operator in the EOMIP-CCSD(2)* method. It can be seen that the increase in the target state energy, which is caused by the truncated T amplitudes in the EOMIP-CCSD(2) approximation, is corrected by the R_3 operator in the EOMIP-CCSD(2)* method. This leads to a better balance

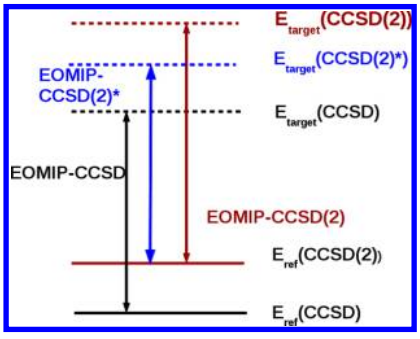


Figure 1. Relative ordering of reference and target state in different variants of the EOM approach to the IP problem.

DOI: 10.1021/ct500927h
J. Chem. Theory Comput. 2015, 11, 2461−2472

between the errors in the reference state and the target state, while the systematic error cancellation results in improved IP values.

(e). Vertical Ionization Potential of Thymine. To show the robustness of the new method, we have calculated the valence IP values of thymine. The EOMCC investigation of IP values of thymine were extensively persuaded by Krylov and co-workers.[22,23] Table 12 presents the computed and experimental vertical IP values for the first five ionized states of thymine.

Table 12. Vertical Ionization Energies of Thymine

molecule	Ionization Energy (eV)				
	$1\ ^2A''$	$1\ ^2A'$	$2\ ^2A''$	$2\ ^2A'$	$3\ ^2A''$
cc-pVDZ Basis Set					
EOMIP-CCSD(2)	8.98	9.77	10.26	10.67	12.47
EOMIP-CCSD	8.79	9.72	10.10	10.63	12.33
EOMIP-CCSD(2*)	8.60	9.46	9.90	10.34	12.03
cc-pVTZ Basis Set					
EOMIP-CCSD(2)	9.44	10.31	10.78	11.20	12.88
extrapolated EOMIP-CCSD(2)	9.25	10.26	10.62	11.16	12.74
EOMIP-CCSD	9.14	10.16	10.52	11.06	12.66
EOMIP-CCSD(2*)	8.98	9.88	10.33	10.74	12.37
Exp[38]	9.02	9.95	10.40	10.88	12.10

It can be seen that the EOMIP-CCSD(2) method overestimates the IP values for all five states, compared to the standard EOMIP-CCSD method, in both the cc-pVDZ and cc-pVTZ basis sets. The extrapolated version shows some improvement over the original EOMIP-CCSD(2) method; however, the values are still overestimated, compared to the EOMIP-CCSD approximation. In the cc-pVTZ basis set, the EOMIP-CCSD method itself overestimates the IP values compared to experimental IP values. The EOMIP-CCSD(2)* method, on the other hand, gives a very good agreement, compared to the experimental result, except the ($3\ ^2A''$) state, where the experimental value is overestimated by 0.27 eV. Generally, the EOMIP-CCSD(2)* method gives better agreement with the experimental results with the cc-pVTZ basis set than the standard EOMIP-CCSD method.

(f). Geometry and IR Frequency. The relaxation effect introduced by the R_3 operator improves the description of the total energy in the EOMIP-CCSD(2)* method over the original EOMIP-CCSD(2) approximation. Therefore, the EOMIP-CCSD(2)* method may offer an improved description of the final state properties. We have investigated the geometry and IR frequencies of NO_2, NO_3, and a test set of six diatomic doublet radicals, some of which have been used for the benchmarking of the original EOMIP-CCSD(2) approximation.[10]

NO_2. NO_2 provides significant challenge for the standard single-reference *ab initio* methods. The UHF- and ROHF-based MP2 and even the CCSD method all fail to provide reasonable agreement with the experiment for either the geometry or the IR frequency of NO_2.

Table 13 provides the geometry and IR frequencies of NO_2 computed using different variants of single-reference and equation of motion coupled cluster methods in the aug-cc-pVTZ basis set. The EOMIP-CCSD(2)* method shows a deviation of 0.002 Å from the experiment for the bond lengths. The results are better than the original EOMIP-CCSD(2) approximation or the standard EOMIP-CCSD method.

Table 13. Geometry and Harmonic Vibrational Frequency of Nitrogen Dioxide (NO_2) in the aug-cc-pVTZ Basis Set

method	bond length (Å)	bond angle (θ)	ω_1	ω_2	ω_3
UCCSD(T)	1.294	123.9	344	577	1134
ROCCSD(T)	1.194	134.8	761	1359	1694
EOM-IP-CCSD	1.186	133.7	795	1443	1745
EOMIP-CCSD(2)	1.186	134.9	769	1388	1784
EOM-IP-CCSD(2)*	1.192	134.7	760	1347	1560
EOM-IP-CCSD*	1.191	133.7	784	1404	1717
experiment	1.194[a]	133.9[a]	750[b]	1325[b]	1634[b]

[a]Values taken from ref 39. [b]Values taken from ref 40.

However, the bond angle is slightly overestimated in the EOMIP-CCSD(2)* method.

The UHF-based CCSD(T) method (UCCSD(T)) fails drastically for both bond lengths and bond angles. The ROHF-based CCSD(T) method (ROCCSD(T)), on the other hand, reproduces the bond length exactly. However, the bond angle is slightly overestimated, in a manner similar to that observed in the case of EOMIP-CCSD(2)* method. For IR frequencies, the ROCCSD(T) method gives the best agreement with experiments, and the largest deviation is observed for the asymmetric stretching mode(ω_3), which is overestimated, compared to the experimental value, by 64 cm^{-1}. The EOMIP-CCSD(2)* method gives a similar performance, with the difference being that the asymmetric stretching mode(ω_3) gets underestimated by 74 cm^{-1}. The performance is superior than that of the original EOMIP-CCSD(2) approximation or even the standard EOMIP-CCSD method. The UCCSD(T) method fails for all the three modes of vibration. Surprisingly, the EOMIP-CCSD* method gives slightly inferior performance, compared to the EOMIP-CCSD(2)* approximation.

NO_3. The equilibrium geometry of NO_3 has been a matter of long-standing debate. The experimental geometry[24] of NO_3 is D_{3h} and most of the single-reference methods, even the coupled cluster method,[25−27] predict a C_{2V} geometry. Multireference methods such as FSMRCCSD, MRCI, and the EOMIP-CCSD(2) method, on the other hand, predict a D_{3h} geometry.[10,28,29]

Table 14 provides the geometry and IR frequencies of NO_3 computed using the aug-cc-pVTZ basis set. Both the UCCSD-(T) and ROCCSD(T) method lead to a C_{2V} geometry with two long (L_1) bonds and one short (L_2) bond. In the ROCCSD(T) method, the long bond is in reasonable agreement with the experimental value. However, the short bond is underestimated by 0.039 Å. The UCCSD(T) method gives inferior performance for both long and short bonds. All the EOM methods lead to a D_{3h} geometry. The EOMIP-CCSD(2)* method gives the best agreement with the experimental results ($|\Delta r_e|$ = 0.008 Å), among all the methods used in this study.

The UCCSD(T) and ROCCSD(T) methods give very poor agreement for all modes of vibrations of NO_3, except for the umbrella and symmetric stretching modes. The EOMIP-CCSD(2)* method, on the other hand, gives very good agreement with the experimental values, except for the two asymmetric stretching modes. Especially, the asymmetric bending modes in EOMIP-CCSD(2)* method show significant improvement over the original EOMIP-CCSD(2) approxima-

DOI: 10.1021/ct500927h
J. Chem. Theory Comput. 2015, 11, 2461−2472

Table 14. Geometry and Harmonic Vibrational Frequency of Nitrogen Trioxide (NO_3) in aug-cc-pVTZ Basis Set

	Bond Length (Å)		Vibrational Frequency (cm^{-1})					
method	L_1	L_2	ω_1 (asym bend)	ω_2 (asym bend)	ω_3 (umbrella)	ω_4 (sym strech)	ω_5 (asym stretch)	ω_6 (asym stretch)
UCCSD(T)	1.291	1.198	664	683	732	1031	1063	1615
ROCCSD(T)	1.252	1.201	414	506	779	896	1082	1499
EOM-IP-CCSD(2)	1.228	1.228	66	66	800	1140	1176	1176
EOMIP-CCSD	1.221	1.221	305	305	836	1170	1191	1191
EOM-IP-CCSD(2)*	1.232	1.232	172	172	785	1114	1179	1179
EOM-IP-CCSD*	1.226	122.6	349	349	822	1146	1188	1188
experiment	1.240[a]	1.240[a]	250[b]	250[b]	762[c]	1060[c]	1480[c]	1480[c]

[a]Values taken from ref 24. [b]Values taken from ref 41. [c]Values taken from ref 42.

Table 15. Geometry of Doublet Diatomic Molecules in the aug-cc-pVQZ Basis Set

	Geometry (Å)						
molecule	UCCSD(T)	ROCCSD(T)	EOM-IP-CCSD	EOMIP-CCSD(2)	EOM-IP-CCSD(2)*	EOM-IP-CCSD*	Exp[a]
OH	0.970	0.969	0.966	0.966	0.968	0.968	0.969
O_2^+	1.115	1.115	1.107	1.112	1.117	1.109	1.116
CN	1.167	1.173	1.161	1.164	1.173	1.167	1.172
F_2^+	1.306	1.305	1.295	1.295	1.303	1.302	1.322
CO^+	1.112	1.116	1.104	1.108	1.117	1.112	1.115
NO	1.148	1.151	1.150	1.145	1.147	1.151	1.151

[a]Data taken from ref 33.

Table 16. IR Frequency of Doublet Diatomic Molecules in the aug-cc-pVQZ Basis Set

	IR Frequency (cm^{-1})						
molecule	UCCSD(T)	ROCCSD(T)	EOM-IP-CCSD	EOMIP-CCSD(2)	EOM-IP-CCSD(2)*	EOM-IP-CCSD*	Exp[a]
OH	3746	3749	3802	3892	3841	3751	3738
O_2^+	1940	1942	2022	1942	1897	2002	1905
CN	2137	2068	2174	2134	2055	2119	2069
F_2^+	1126	1128	1175	1178	1137	1138	1073
CO^+	2303	2223	2331	2288	2200	2247	2212
NO	2104	1918	2004	2022	1967	1952	1904

[a]Data taken from ref 33.

tion. On the other hand, the two asymmetric stretching modes show considerable deviation from the experimental values in all the EOM methods. However, Stanton[30] has shown that the assignment of the experimental peak at 1480 cm^{-1} is not unambiguous, and detailed investigations are required for the assignment of these modes, which is outside the scope of the present study. For an elaborate discussion on the *ab initio* determination of the structure and properties of NO_3, readers are referred to refs 31 and 32.

Diatomics. The diatomic doublet radical suffers from a high degree of symmetry breaking and other typical problems associated with the theoretical treatment of open-shell molecules. They are often used as test cases for benchmarking the accuracy of multireference methods.

The bond length and IR frequencies of six doublet radicals OH, O_2^+, CN, F_2^+, CO^+ and NO are given in Tables 15 and 16, respectively. The ROCCSD(T) method gives the best agreement with the experimental bond length and IR frequency.[33] The EOMIP-CCSD(2)* method gives a comparable performance, and it shows significant improvement over the original EOMIP-CCSD(2) approximation. The EOMIP-CCSD(2)* results are generally in better agreement with the experiment than those obtained from the standard EOMIP-CCSD method, except for the case of NO. The UCCSD(T) and EOMIP-CCSD* methods show a mixed performance: while they are in good agreement in some cases, they also lead to inferior performance for others.

g. Adiabatic Ionization Energy. Table 17 presents adiabatic ionization energies of six small molecules in the cc-pVQZ basis set, using the different truncation schemes of EOMCC. It can be seen that the EOMIP-CCSD(2) method significantly overestimates the IP values, compared to benchmark EOMIP-CCSDT results. On the other hand, the EOMIP-CCSD* method underestimates the IP values. The EOMIP-CCSD(2)* method provides a better error balance and gives IP

Table 17. Adiabatic Ionization Energy in the cc-pVQZ Basis Set[a]

	Ionization Energy (eV)				
molecule	EOM-IP-CCSD	EOMIP-CCSD(2)	EOM-IP-CCSD(2)*	EOM-IP-CCSD*	EOMIP-CCSDT
NO_3[b]	3.78	4.07	3.64	3.39	3.62
OH	1.06	1.18	1.07	0.98	1.06
O_2^+	10.62	11.01	10.81	10.45	10.84
CN	5.05	5.27	4.76	4.61	4.82
F_2^+	15.46	15.57	15.48	15.33	15.55
CO^+	14.21	14.36	13.63	13.91	13.98

[a]Geometry and zero point energy calculated in the aug-cc-pVDZ basis set. [b]Here, *f* functions were removed to keep it computationally viable.

DOI: 10.1021/ct500927h
J. Chem. Theory Comput. 2015, 11, 2461−2472

values that are in excellent agreement with the EOMIP-CCSDT result. The trend is consistent with that observed in the case of vertical IP values.

4. CONCLUSIONS

In this paper, we propose a new approximation within the hierarchy of EOMIP-CC methods for the calculation of ionization potential. Our EOMIP-CCSD(2)* method corrects for the missing relaxation effect caused by the truncated T amplitudes in the original EOMIP-CCSD(2) approximation by the partial inclusion of the R_3 operator in the EOM part. The EOMIP-CCSD(2)* method scales as noniterative N^6 and has a much smaller storage requirement, compared to the standard EOMIP-CCSD method, and can be applied to large systems. The resulting EOMIP-CCSD(2)* method is free from the problem of overestimation of IP values shown by the original EOMIP-CCSD(2) method and its extrapolated versions. The superiority of the method is especially prominent for the ionization of core electrons and satellite peaks, where the relaxation effect plays an important role and the new method outperforms the standard EOMIP-CCSD method for the above-mentioned cases.

The EOMIP-CCSD(2)* method is capable of accurate simulation of geometry and IR frequencies of problematic doublet radicals and gives excellent agreement with the experimental results. The results using the EOMIP-CCSD(2)* method are comparable to those obtained using the single-reference CCSD(T) method, and they are even better than the standard EOMIP-CCSD method for most of the cases.

Similar development is also possible in the context of Fock space multireference coupled cluster method. It has additional advantages in calculating the IP values corresponding to multiple states in a single calculation, as it does not require the calculation of a left vector, as in EOMCC. Work in that direction is currently underway and will be reported in a future publication.

■ ASSOCIATED CONTENT

Supporting Information

Cartesian coordinates of the studied molecules are given in the Supporting Information. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/ct500927h.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: s.pal@ncl.res.in.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge a grant from the CSIR XIIth Five-Year Plan Project on Multiscale Simulations of Material (MSM) and facilities of the Centre of Excellence in Scientific Computing at NCL. A.K.D. thanks the Council of Scientific and Industrial Research (CSIR) for a Senior Research Fellowship. S.P. acknowledges the DST J. C. Bose Fellowship project and CSIR SSB grant towards completion of the work. A.K.D. acknowledges Dr. Robert Izsak and Himadri Pathak for a critical reading of the manuscript.

■ REFERENCES

(1) Stanton, J. F.; Bartlett, R. J. The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys.* **1993**, *98*, 7029−7039.
(2) Stanton, J. F.; Gauss, J. Analytic energy derivatives for ionized states described by the equation-of-motion coupled cluster method. *J. Chem. Phys.* **1994**, *101*, 8938−8944.
(3) Nooijen, M.; Bartlett, R. J. Equation of motion coupled cluster method for electron attachment. *J. Chem. Phys.* **1995**, *102*, 3629−3647.
(4) Mukherjee, D.; Pal, S., Use of Cluster Expansion Methods in the Open-Shell Correlation Problem. In *Advances in Quantum Chemistry*, Vol. *20*; Academic Press: San Diego, CA, 1989; pp 291−373.
(5) Bartlett, R. J. Many-Body Perturbation Theory and Coupled Cluster Theory for Electron Correlation in Molecules. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359−401.
(6) Nooijen, M.; Snijders, J. G. Second order many-body perturbation approximations to the coupled cluster Green's function. *J. Chem. Phys.* **1995**, *102*, 1681−1688.
(7) Stanton, J. F.; Gauss, J. Perturbative treatment of the similarity transformed Hamiltonian in equation-of-motion coupled-cluster approximations. *J. Chem. Phys.* **1995**, *103*, 1064−1076.
(8) Dutta, A. K.; Gupta, J.; Pathak, H.; Vaval, N.; Pal, S. Partitioned EOMEA-MBPT(2): An Efficient N5 Scaling Method for Calculation of Electron Affinities. *J. Chem. Theory Comput.* **2014**, *10*, 1923−1933.
(9) Dutta, A. K.; Pal, S.; Ghosh, D. Perturbative approximations to single and double spin flip equation of motion coupled cluster singles doubles methods. *J. Chem. Phys.* **2013**, *139*, 124116.
(10) Dutta, A. K.; Vaval, N.; Pal, S. Performance of the EOMIP-CCSD(2) Method for Determining the Structure and Properties of Doublet Radicals: A Benchmark Investigation. *J. Chem. Theory Comput.* **2013**, *9*, 4313−4331.
(11) Dutta, A. K.; Vaval, N.; Pal, S., Assessment of Low Scaling Approximations to EOM-CCSD Method for Ionization Potential. Manuscript under review.
(12) Musial, M.; Bartlett, R. J. Multireference Fock-space coupled-cluster and equation-of-motion coupled-cluster theories: The detailed interconnections. *J. Chem. Phys.* **2008**, *129*, 134105−134112.
(13) Stanton, J. F.; Gauss, J. A simple correction to final state energies of doublet radicals described by equation-of-motion coupled cluster theory in the singles and doubles approximation. *Theor. Chem. Acc.* **1996**, *93*, 303−313.
(14) Stanton, J. F.; Gauss, J. A simple correction to final state energies of doublet radicals described by equation-of-motion coupled cluster theory in the singles and doubles approximation (Erratum). *Theor. Chem. Acc.* **1997**, *95*, 97−98.
(15) Saeh, J. C.; Stanton, J. F. Application of an equation-of-motion coupled cluster method including higher-order corrections to potential energy surfaces of radicals. *J. Chem. Phys.* **1999**, *111*, 8275−8285.
(16) Löwdin, P. O. Studies in perturbation theory: Part I. An elementary iteration-variation procedure for solving the Schrödinger equation by partitioning technique. *J. .Mol. Spectrosc.* **1963**, *10*, 12−33.
(17) Dutta, A. K.; Vaval, N.; Pal, S. A new scheme for perturbative triples correction to (0, 1) sector of Fock space multi-reference coupled cluster method: Theory, implementation, and examples. *J. Chem. Phys.* **2015**, *142*, 044113.
(18) Manohar, P. U.; Stanton, J. F.; Krylov, A. I. Perturbative triples correction for the equation-of-motion coupled-cluster wave functions with single and double substitutions for ionized states: Theory, implementation, and examples. *J. Chem. Phys.* **2009**, *131*, 14112.
(19) Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G.; Auer, A. A.; Bartlett, R. J.; Benedikt, U.; Berger, C.; Bernholdt, D. E.; Bomble, Y. J., CFOUR, a quantum chemical program package. For the current version, see http://www.cfour.de 2009. (accessed Oct. 1, 2014).
(20) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(21) Peterson, K. A.; Dunning, T. H. Accurate correlation consistent basis sets for molecular core-valence correlation effects: The second

row atoms Al-Ar, and the first row atoms Ba-Ne revisited. *J. Chem. Phys.* **2002**, *117*, 10548−10560.
(22) Bravaya, K. B.; Kostko, O.; Ahmed, M.; Krylov, A. I. The effect of stacking, H-bonding, and electrostatic interactions on the ionization energies of nucleic acid bases: Adenine−adenine, thymine−thymine and adenine−thymine dimers. *Phys. Chem. Chem. Phys.* **2010**, *12*, 2292−2307.
(23) Bravaya, K. B.; Kostko, O.; Dolgikh, S.; Landau, A.; Ahmed, M.; Krylov, A. I. Electronic Structure and Spectroscopy of Nucleic Acid Bases: Ionization Energies, Ionization-Induced Structural Changes, and Photoelectron Spectra. *J. Phys. Chem. A* **2009**, *114*, 12305−12317.
(24) Ishiwata, T.; Tanaka, I.; Kawaguchi, K.; Hirota, E. Infrared diode laser spectroscopy of the NO_3 ν_3 band. *J. Chem. Phys.* **1985**, *82*, 2196−2205.
(25) Crawford, T. D.; Stanton, J. F. Some surprising failures of Brueckner coupled cluster theory. *J. Chem. Phys.* **2000**, *112*, 7873−7879.
(26) Gauss, J.; Stanton, J. F.; Bartlett, R. J. Analytic evaluation of energy gradients at the coupled-cluster singles and doubles level using quasi-restricted Hartree−Fock open-shell reference functions. *J. Chem. Phys.* **1991**, 952639−952645.
(27) Stanton, J. F.; Gauss, J.; Bartlett, R. J. Potential nonrigidity of the NO_3 radical. *J. Chem. Phys.* **1991**, *94*, 4084−4087.
(28) Kaldor, U. The ground state geometry of the NO_3 radical. *Chem. Phys. Lett.* **1990**, *166*, 599−601.
(29) Eisfeld, W.; Morokuma, K. A detailed study on the symmetry breaking and its effect on the potential surface of NO_3. *J. Chem. Phys.* **2000**, *113*, 5587−5597.
(30) Stanton, J. F. On the vibronic level structure in the NO_3 radical. I. The ground electronic state. *J. Chem. Phys.* **2007**, *126*, 134309−134320.
(31) Simmons, C. S.; Ichino, T.; Stanton, J. F., The ν_3 Fundamental in NO_3 Has Been Seen Near 1060 cm^{-1}, Albeit Some Time Ago. *J. Phys. Chem. Lett. 3*, 1946−1950.
(32) Stanton, J. F. Note: Is it symmetric or not? *J. Chem. Phys.* **2013**, *139*, 046102.
(33) Huber, K. P.; Herzberg, G. *Molecular Structure and Molecular Spectra. IV. Constants of Diatomic Molecules*; Van Rostrand−Reinhold: New York, 1979.
(34) Kimura, K.; Katsumata, S.; Achiba, Y.; Yamazaki, T.; Iwata, S. *Handbook of HeI Photoelectron Spectra of Fundamental Organic Molecules*; Japan Scientific Societies Press: Tokyo, 1981; pp 27−221.
(35) Musial, M.; Bartlett, R. J. EOM-CCSDT study of the low-lying ionization potentials of ethylene, acetylene and formaldehyde. *Chem. Phys. Lett.* **2004**, *384*, 210−214.
(36) Ohtsuka, Y.; Nakatsuji, H. Inner-shell ionizations and satellites studied by the open-shell reference symmetry-adapted cluster/symmetry-adapted cluster configuration-interaction method. *J. Chem. Phys.* **2006**, *124*, 054110.
(37) Ehara, M.; Nakatsuji, H. Geometry Relaxations After Inner-Shell Excitations and Ionizations. *Collect. Czek. Chem. Commun.* **2008**, *73*, 771−785.
(38) Lauer, G.; Schafer, W.; Schweig, A. Functional subunits in the nucleic acid bases uracil and thymine. *Tetrahedron Lett.* **1975**, *16*, 3939−3942.
(39) Lafferty, W. J.; Sams, R. L. The high resolution infrared spectrum of the $2\nu_2 + \nu_3$ and $\nu_1 + \nu_2 + \nu_3$ bands of $^{14}N^{16}O_2$: Vibration and vibration−rotation constants of the electronic ground state of $^{14}N^{16}O_2$. *J. Mol. Spectrosc.* **1977**, *66*, 478−492.
(40) Morino, Y.; Tanimoto, M.; Saito, S.; Hirota, E.; Awata, R.; Tanaka, T. Microwave spectrum of nitrogen dioxide in excited vibrational states—Equilibrium structure. *J. Mol. Spectrosc.* **1983**, *98*, 331−348.
(41) Weaver, A.; Arnold, D. W.; Bradforth, S. E.; Neumark, D. M. Examination of the $^2A'_2$ and $^2E''$ states of NO_3 by ultraviolet photoelectron spectroscopy of NO_3^-. *J. Chem. Phys.* **1991**, *94*, 1740−1751.
(42) Friedl, R. R.; Sander, S. P. Fourier transform infrared spectroscopy of the nitrate radical ν_2 and ν_3 bands: Absolute line strength measurements. *J. Phys. Chem.* **1987**, *91*, 2721−2726.

DOI: 10.1021/ct500927h
J. Chem. Theory Comput. 2015, 11, 2461−2472