



**HAL**  
open science

## How accurate are EOM-CC4 vertical excitation energies?

Pierre-Francois Loos, Devin A. Matthews, Filippo Lipparini, Denis Jacquemin

► **To cite this version:**

Pierre-Francois Loos, Devin A. Matthews, Filippo Lipparini, Denis Jacquemin. How accurate are EOM-CC4 vertical excitation energies?. *The Journal of Chemical Physics*, 2021, 154 (22), pp.221103. 10.1063/5.0055994 . hal-03218908v2

**HAL Id: hal-03218908**

**<https://hal.science/hal-03218908v2>**

Submitted on 22 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# How accurate are EOM-CC4 vertical excitation energies?

Pierre-François Loos,<sup>1, a)</sup> Devin A. Matthews,<sup>2</sup> Filippo Lipparini,<sup>3</sup> and Denis Jacquemin<sup>4, b)</sup>

<sup>1)</sup>Laboratoire de Chimie et Physique Quantiques (UMR 5626), Université de Toulouse, CNRS, UPS, France

<sup>2)</sup>Department of Chemistry, Southern Methodist University, Dallas, Texas 75275, USA

<sup>3)</sup>Dipartimento di Chimica e Chimica Industriale, University of Pisa, Via Moruzzi 3, 56124 Pisa, Italy

<sup>4)</sup>Université de Nantes, CNRS, CEISAM UMR 6230, F-44000 Nantes, France

We report the first investigation of the performance of EOM-CC4 — an approximate equation-of-motion coupled-cluster model which includes iterative quadruple excitations — for vertical excitation energies in molecular systems. By considering a set of 28 excited states in 10 small molecules for which we have computed CCSDTQP and FCI reference energies, we show that, in the case of excited states with a dominant contribution from the single excitations, CC4 yields excitation energies with sub-kJ mol<sup>-1</sup> accuracy (*i.e.*, error below 0.01 eV), in very close agreement with its more expensive CCSDTQ parent. Therefore, if one aims at high accuracy, CC4 stands as a highly competitive approximate method to model molecular excited states, with a significant improvement over both CC3 and CCSDT. Our results also evidence that, although the same qualitative conclusions hold, one cannot reach the same level of accuracy for transitions with a dominant contribution from the double excitations.

## I. INTRODUCTION

Single-reference coupled-cluster (CC) theory provides a hierarchy of size-extensive methods delivering increasingly accurate energies and properties via the systematic increase of the maximum excitation degree of the cluster operator  $\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots + \hat{T}_k$  (where  $k \leq n$  and  $n$  is the number of electrons).<sup>1-5</sup> Without any truncation (*i.e.*,  $k = n$ ), the so-called full CC (FCC) method is equivalent to full configuration interaction (FCI), hence providing the exact energy and wave function of the system for a given atomic basis set. However, it is not computationally viable due to its exponential scaling with system size, and one has to resort to truncated CC methods (*i.e.*,  $k \ll n$ ) for computational convenience. Popular choices are CC with singles and doubles (CCSD),<sup>1,6</sup> CC with singles, doubles, and triples (CCSDT),<sup>7,8</sup> CC with singles, doubles, triples, and quadruples (CCSDTQ),<sup>9,10</sup> and CC with singles, doubles, triples, quadruples, and pentuples (CCSDTQP)<sup>11,12</sup> with corresponding computational scalings of  $O(N^6)$ ,  $O(N^8)$ ,  $O(N^{10})$ , and  $O(N^{12})$ , respectively (where  $N$  denotes the number of orbitals). An alternative, systematically-improvable family of methods is defined by the CC2,<sup>13</sup> CC3,<sup>14,15</sup> and CC4<sup>16</sup> series of models which have been introduced by the Aarhus group in the context of CC response theory.<sup>17</sup> These iterative methods scale as  $O(N^5)$ ,  $O(N^7)$ , and  $O(N^9)$ , respectively, and can be seen as cheaper approximations of CCSD, CCSDT, and CCSDTQ, by skipping the most expensive terms and avoiding the storage of the higher-excitation amplitudes. A somewhat similar strategy has been applied to define the CCSDT-3<sup>18,19</sup> and CCSDTQ-3<sup>16</sup> models based on arguments stemming from perturbation theory. Of course, a large number of other approximate CC models have been developed over the years and we refer the interested reader to specialized reviews for more details.<sup>3-5,20</sup>

Coupled-cluster methods have been particularly successful for small- and medium-sized molecules in the field of thermodynamics, kinetics, and spectroscopy, thanks to

the computations of accurate equilibrium geometries,<sup>21</sup> potential energy surfaces, vibrational frequencies,<sup>22</sup> Born-Oppenheimer corrections,<sup>23</sup> and a vast panel of properties such as dipoles (and higher moments),<sup>21</sup> NMR chemical shifts,<sup>22</sup> magnetizabilities,<sup>24</sup> polarizabilities,<sup>25</sup> etc. Although originally developed for ground-state energies and properties, CC has been successfully extended to excited states<sup>26</sup> thanks to the equation-of-motion (EOM)<sup>27-30</sup> and linear-response (LR)<sup>31-35</sup> formalisms which are known to produce identical excitation energies but different properties. In EOM-CC, one determines vertical excitation energies via the diagonalization of the similarity-transformed Hamiltonian  $\bar{H} = e^{-\hat{T}}\hat{H}e^{\hat{T}}$ . A more general procedure to compute excitation energies that can be applied to any approximate CC model (as the ones mentioned above) consists in diagonalizing the so-called CC Jacobian obtained via the differentiation of the CC amplitude equations. Again, by increasing the maximum excitation degree of  $\hat{T}$ , one can systematically produce increasingly accurate EOM/LR excitation energies with the “complete” series CCSD,<sup>29</sup> CCSDT,<sup>36,37</sup> CCSDTQ,<sup>38</sup> CCSDTQP<sup>39</sup> ... or the “approximate” series CC2,<sup>13</sup> CC3,<sup>14</sup> CC4<sup>39</sup> ... while retaining the formal scaling of their ground-state analog. From here on, we drop the EOM prefix as all our calculations are performed within this approach.

Recently, a large collection of deterministic, stochastic or hybrid selected CI (SCI) methods<sup>40-42</sup> has (re)appeared<sup>43-57</sup> in the electronic structure landscape providing an alternative route to highly-accurate ground- and excited-state energies<sup>58-70</sup> (see Refs. 71 and 72 for recent reviews). The general idea behind SCI methods is simple: rather than exploring the entire FCI space by systematically increasing the maximum excitation degree of the determinants taken into account (leading to the slowly convergent and size-inconsistent series of methods CISD, CISDT, CISDTQ, ...), one performs a sparse exploration of the FCI space by selecting only the most energetically important determinants thanks to a suitable criterion usually based on perturbation theory.<sup>42,67,73,74</sup> By iteratively increasing the number of determinants of the variational space and supplementing it with a second-order perturbative correction (PT2), the SCI+PT2 family of methods has been recently shown to produce near-FCI correlation and excitation energies for small-

<sup>a)</sup>Electronic mail: loos@irsamc.ups-tlse.fr

<sup>b)</sup>Electronic mail: denis.jacquemin@univ-nantes.fr

and medium-size molecules in compact basis sets.<sup>58,61–65,67,70</sup> Although the formal scaling of such algorithms remain exponential, the prefactor is greatly reduced which explains their current attractiveness in the electronic structure community and much wider applicability than their standard FCI parent.

Taking advantage of the high accuracy of CC and SCI+PT2 methods, we have very recently created a large dataset gathering more than 500 highly-accurate vertical excitation energies for electronic transitions of various natures (valence, Rydberg,  $n \rightarrow \pi^*$ ,  $\pi \rightarrow \pi^*$ , singlet, doublet, triplet, charge-transfer, and double excitations) in small- and medium-sized molecules ranging from diatomics to molecules as large as naphthalene.<sup>61,63–65,70,71</sup> The main purpose of this so-called QUEST database is to provide reference excitation energies in order to perform fair and reliable benchmarks between electronic structure methods and assess their strengths and weakness for a large panel of chemical scenarios. Most of these reference transition energies, which rely exclusively on high-level *ab initio* calculations, can be reasonably considered as chemically accurate, *i.e.*, within 1 kcal mol<sup>-1</sup> or 0.043 eV of the FCI limit. However, their accuracy may rapidly deteriorate, in particular, as the system size grows. Indeed, it is usually challenging to compute reliable SCI+PT2 estimates or CCSDTQ excitation energies for molecules with more than four non-hydrogen atoms. Therefore, for the larger molecules of the QUEST database, we mostly relied on CCSDT to define reference excitation energies. Of course, it would be highly valuable to have access to more accurate methods (including, at least, quadruple excitations) in order to refine these theoretical best estimates.

In this context, the main purpose of the present study is to assess the relative accuracy of the approximate iterative CC4 model against the more expensive CCSDTQ and CCSDTQP methods in the case of vertical excitation energies, as well as their absolute accuracy with respect to FCI. To do so, we consider a set of 10 small molecules (NH<sub>3</sub>, C<sub>2</sub>, BH, BF, CO, N<sub>2</sub>, HCl, H<sub>2</sub>S, HNO and H<sub>2</sub>O) and we compare the excitation energies associated with 28 singlet excited states of various natures ( $n \rightarrow \pi^*$ ,  $\pi \rightarrow \pi^*$ , Rydberg, valence, charge-transfer, and double excitations) and spatial symmetries obtained with various high-level CC methods. Although a small number of studies have been published on the performance of CC4 for ground-state energies and properties,<sup>16,75</sup> this work stands, to the best of our knowledge, as the first to consider CC4 for the computation of excited-states energies. As we shall see below, CC4 is an excellent approximation to its CCSDTQ parent, and produces, in the case of excited states with a dominant contribution from the single excitations, excitation energies with sub-kJ mol<sup>-1</sup> accuracy (*i.e.*, error below 0.01 eV) for this set of small molecules, well below the chemical accuracy threshold.

## II. COMPUTATIONAL DETAILS

All the methods considered in the present study are listed in Table I alongside their formal computational scaling and the electronic structure software employed to compute the excitation energies. In a nutshell, we have used by default

TABLE I. Methods considered in the present study, their formal computational scaling and the electronic structure software employed to compute excitation energies. Here,  $N$  is the number of orbitals.

Method	Scaling	Code	Ref.
CC2	$O(N^5)$	CFOUR	76
CCSD	$O(N^6)$	CFOUR	76
CC3	$O(N^7)$	CFOUR	76
CCSDT	$O(N^8)$	CFOUR	76
CC4	$O(N^9)$	CFOUR	76
CCSDTQ	$O(N^{10})$	CFOUR	76
CCSDTQP	$O(N^{12})$	MRCC	77
CIPSI	$O(e^N)$	QUANTUM PACKAGE	67

CFOUR<sup>76</sup> to compute the CC energies at the notable exception of CCSDTQP for which we have employed MRCC and its automated implementation of high-order CC methods.<sup>77</sup> The present CFOUR calculations have been performed with the new and fast CC module (xncc) written by one of authors (DAM) which couples a general algebraic and graphical interpretation of the non-orthogonal spin-adaptation approach with highly efficient storage format and implementation techniques designed to minimize data movement and to avoid costly tensor transposes.<sup>78</sup> The FCI estimates were obtained with the SCI algorithm known as “*configuration interaction using a perturbative selection made iteratively*” (CIPSI) implemented in QUANTUM PACKAGE.<sup>67</sup> The error bars associated with the extrapolation step of the CIPSI calculations (see Ref. 67) have been computed using our recently developed protocol presented in Ref. 70.

Because this is, to our knowledge, the first implementation of EOM-CC4, we have verified its accuracy by comparing the excitation energies obtained from solving for the right- and left-hand wave functions. Coupled cluster is a non-hermitian theory and thus the right- and left-hand eigenfunctions of the Jacobian are distinct, albeit with the same eigenvalue. For the left-hand solution, we have reused the already-verified code for the ground-state  $\Lambda$  equations which describe the amplitude relaxation contribution in the analytic gradient theory.<sup>75</sup> The structure of the left-hand EOM-CC and  $\Lambda$  equations are identical, and so simply interfacing this existing code with a Davidson solver<sup>79</sup> provides left-hand EOM-CC solutions; this procedure has been checked for other known-good methods such as CCSDT and CC3.

All calculations have been performed in the frozen-core approximation and the CC3/aug-cc-pVTZ geometries of the systems considered here have been extracted from previous studies.<sup>61,80</sup> In the following, we consider diffuse-containing Dunning’s double- and triple- $\zeta$  basis sets (aug-cc-pVDZ and aug-cc-pVTZ). Note that CCSDTQP energies could only be computed for the smaller basis (aug-cc-pVDZ).

## III. RESULTS AND DISCUSSION

Table II gathers, for the two considered basis sets, 25 vertical excitation energies with a (strongly) dominant contribution

TABLE II. Vertical excitation energies (in eV) of a selection of singly-excited states obtained at various levels of theory with the aug-cc-pVDZ and aug-cc-pVTZ basis sets.  $\%T_1$  is the percentage of single excitations involved in the transition computed at the CC3/aug-cc-pVTZ level. For the aug-cc-pVDZ basis, the mean absolute error (MAE), mean signed error (MSE), and maximum error (Max) with respect to CCSDTQP is reported. For the FCI data, the error bars reported in parenthesis correspond to one standard deviation.

Mol.	State	$\%T_1$	aug-cc-pVDZ							aug-cc-pVTZ							
			CC2	CCSD	CC3	CCSDT	CC4	CCSDTQ	CCSDTQP	FCI	CC2	CCSD	CC3	CCSDT	CC4	CCSDTQ	FCI
NH <sub>3</sub>	<sup>1</sup> A <sub>2</sub>	93	6.249	6.455	6.464	6.462	6.479	6.480	6.482	6.483(1)	6.387	6.600	6.573	6.571	6.585	6.586	6.593(22)
	<sup>1</sup> E	93	7.733	8.024	8.061	8.057	8.078	8.079	8.081	8.082(1)	7.847	8.148	8.146	8.143	8.161	8.161	8.171(20)
	<sup>1</sup> A <sub>1</sub>	94	9.400	9.649	9.664	9.659	9.677	9.677	9.680	9.681(8)	9.051	9.334	9.318	9.314	9.331	9.331	9.340(19)
	<sup>1</sup> A <sub>2</sub>	93	10.148	10.376	10.396	10.391	10.409	10.409	10.411	10.412(1)	9.654	9.953	9.945	9.939	9.957	9.957	9.967(19)
BH	<sup>1</sup> Π	95	2.862	2.970	2.955	2.946	2.947	2.947	2.947	2.947(0)	2.831	2.928	2.910	2.900	2.901	2.901	2.901(0)
BF	<sup>1</sup> Π	94	6.509	6.534	6.478	6.491	6.484	6.485	6.485	6.485(1)	6.445	6.464	6.410	6.423	6.416	6.417	6.418(2)
CO	<sup>1</sup> Π	93	8.724	8.671	8.572	8.574	8.562	8.563	8.561	8.563(4)	8.638	8.587	8.486	8.492	8.479	8.480	
	<sup>1</sup> Σ <sup>-</sup>	93	10.381	10.096	10.122	10.062	10.055	10.057	10.057	10.056(1)	10.297	9.986	9.992	9.940	9.930	9.932	
	<sup>1</sup> Δ	91	10.685	10.210	10.225	10.178	10.167	10.169	10.168	10.168(1)	10.604	10.123	10.119	10.076	10.064	10.066	
	<sup>1</sup> Σ <sup>+</sup>	91	11.089	11.171	10.917	10.944	10.925	10.926	10.919		11.106	11.222	10.943	10.987	10.961	10.963	
	<sup>1</sup> Σ <sup>+</sup>	92	11.628	11.710	11.483	11.518	11.510	11.510	11.506		11.626	11.751	11.489	11.540	11.521	11.523	
	<sup>1</sup> Π	92	11.878	11.973	11.737	11.767	11.757	11.758	11.753		11.825	11.960	11.690	11.737	11.719	11.720	
	<sup>1</sup> Π <sub>g</sub>	92	9.528	9.495	9.442	9.417	9.409	9.411	9.409	9.411(3)	9.439	9.408	9.344	9.326	9.317	9.319	
N <sub>2</sub>	<sup>1</sup> Σ <sub>u</sub> <sup>-</sup>	97	10.428	10.197	10.059	10.060	10.063	10.055	10.054	10.054(0)	10.320	9.996	9.885	9.890	9.883	9.878	9.879(4)
	<sup>1</sup> Δ <sub>u</sub>	95	10.961	10.607	10.433	10.436	10.439	10.430	10.428	10.429(0)	10.863	10.443	10.293	10.302	10.294	10.287	10.289(12)
	<sup>1</sup> Σ <sub>g</sub> <sup>+</sup>	92	13.077	13.326	13.229	13.202	13.171	13.182	13.181	13.180(1)	12.833	13.151	13.013	12.999	12.962	12.974	
	<sup>1</sup> Π <sub>u</sub>	82	13.309	13.451	13.279	13.174	13.128	13.131	13.127		13.152	13.422	13.223	13.140	13.091	13.095	
	<sup>1</sup> Σ <sub>u</sub> <sup>+</sup>	92	12.937	13.250	13.146	13.130	13.099	13.109	13.107		12.888	13.263	13.120	13.118	13.078	13.090	
	<sup>1</sup> Π <sub>u</sub>	87	14.091	13.765	13.635	13.591	13.551	13.560	13.558		13.963	13.674	13.494	13.455	13.409	13.419	
	<sup>1</sup> Π	94	7.895	7.862	7.819	7.815	7.822	7.822	7.823	7.823(0)	7.959	7.906	7.840	7.834	7.837	7.837	7.838(1)
HCl	<sup>1</sup> B <sub>1</sub>	94	6.157	6.141	6.098	6.098	6.102	6.103	6.103	6.103(1)	6.304	6.294	6.240	6.237	6.238	6.238	6.240(7)
H <sub>2</sub> S	<sup>1</sup> A <sub>2</sub>	94	6.431	6.343	6.293	6.286	6.286	6.286	6.286	6.286(0)	6.345	6.246	6.192	6.185	6.181	6.181	6.181(6)
	<sup>1</sup> B <sub>1</sub>	93	7.089	7.447	7.511	7.497	7.531	7.528	7.532	7.533(0)	7.234	7.597	7.605	7.591	7.623	7.620	7.626(3)
H <sub>2</sub> O	<sup>1</sup> A <sub>2</sub>	93	8.743	9.213	9.293	9.279	9.317	9.313	9.318	9.318(0)	8.889	9.361	9.382	9.368	9.405	9.400	9.407(7)
	<sup>1</sup> A <sub>1</sub>	93	9.486	9.861	9.921	9.903	9.940	9.937	9.941	9.941(0)	9.580	9.957	9.966	9.949	9.986	9.981	9.986(2)
MAE			0.257	0.109	0.028	0.018	0.003	0.002									
MSE			0.020	0.075	0.013	0.000	0.000	0.000									
Max			0.575	0.324	0.152	0.047	0.011	0.007									

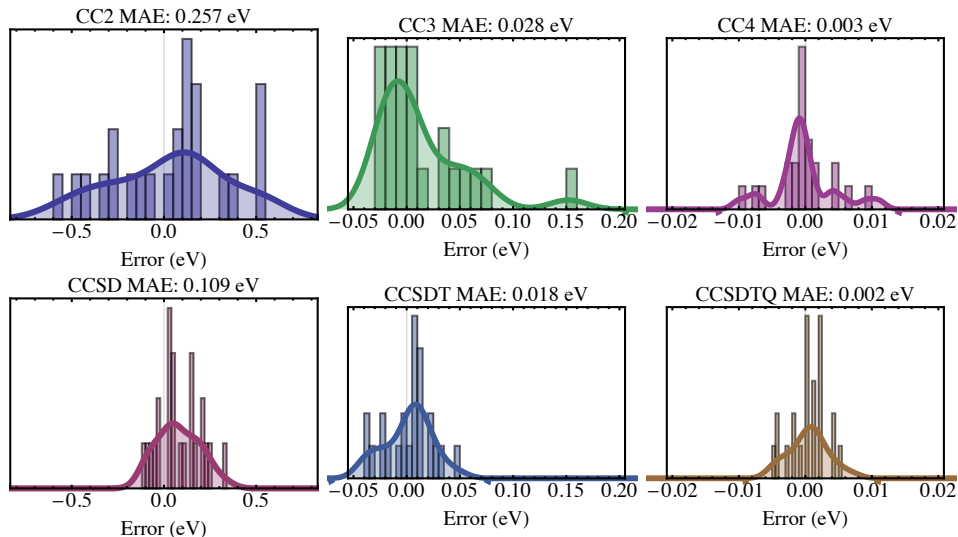


FIG. 1. Distribution of the error (in eV) in excitation energies (with respect to CCSDTQP) for CC2, CC3, and CC4 (top) and CCSD, CCSDT, and CCSDTQ (bottom) obtained with the aug-cc-pVDZ basis. Note the various error ranges ( $X$  axis) for the various methods. See Table II for the raw data.

from the single excitations computed for a set of 8 molecules with various CC models as well as the FCI estimates computed with CIPSI. First, we underline that the FCI estimates show how accurate the CCSDTQP reference data are, with a maximum deviation of 0.002 eV when one considers the aug-cc-pVDZ basis. For the larger aug-cc-pVTZ basis, the CCSDTQ and FCI remain in excellent agreement although the

error bars associated with the extrapolated FCI values prevent us from any quantitative comparisons.

The mean absolute errors (MAEs) and mean signed errors (MSEs) with respect to CCSDTQP computed in the aug-cc-pVDZ basis are reported in the bottom of Table II for the CC2-to-CCSDTQ models. The distribution of the errors are reported in Fig. 1 for each level of theory. These statistical

TABLE III. Vertical excitation energies (in eV) for a selection of transitions with a dominant contribution from the double excitations obtained at various levels of theory with the aug-cc-pVDZ basis set.  $\%T_1$  is the percentage of single excitations involved in the transition (computed at the CC3/aug-cc-pVTZ level).

Mol.	State	$\%T_1$	CC3	CCSDT	CC4	CCSDTQ	CCSDTQP	FCI
C <sub>2</sub>	<sup>1</sup> Δ <sub>g</sub>	1	3.107	2.632	2.341	2.241	2.214	2.213(0)
	<sup>1</sup> Σ <sub>g</sub> <sup>+</sup>	1	3.283	2.874	2.602	2.521	2.505	2.503(1)
HNO	<sup>1</sup> A'	0	5.247	4.756	4.454	4.424		4.397(1)

quantities nicely illustrate the systematic improvement of the transition energies when one ramps up the computational effort following the series CC2, CCSD, CC3, CCSDT, CC4, and CCSDTQ. We note that the errors decrease by roughly one order of magnitude when switching from CCSD to CC3 and from CCSDT to CC4, whilst improvements of approximately 50% “only” are noted when going from the “approximate” model to the “complete” method (*i.e.*, from CC2 to CCSD, from CC3 to CCSDT, and from CC4 to CCSDTQ). In other words, CC4 brings significant improvements in terms of MAE and MSE as compared to the third-order methods, CC3 and CCSDT, which demonstrates the importance of quadruple excitations when one aims at very high accuracy. Besides, for the two basis sets, there is an outstanding similarity between the CC4 and CCSDTQ excitation energies with mean absolute and signed deviations below (equal to) 0.001 eV and a maximum deviation of 0.011 eV (0.007 eV) between the two sets of data obtained with the aug-cc-pVDZ (aug-cc-pVTZ) basis set. Therefore, including quadruples allows to reach sub-kJ mol<sup>-1</sup> accuracy (*i.e.*, average error below 0.01 eV) for singly-excited states with only a rather minor improvement in going from CC4 to CCSDTQ.

A closer inspection at Table II shows that the largest deviations appear for the transitions with the smallest  $\%T_1$  values (where  $\%T_1$  is the percentage of single excitations involved in the transition which is computed at the CC3/aug-cc-pVTZ level in the present case). This is particularly noticeable for the two <sup>1</sup>Π<sub>u</sub> transitions of N<sub>2</sub>. This suggests that the performance of the various CC models discussed above might be highly dependent on the nature of the transitions, as discussed in Ref. 63. To investigate further this point, we report in Table III vertical excitation energies for transitions with a dominant contribution from the doubly-excited determinants in the carbon dimer, C<sub>2</sub>, and nitroxyl, HNO, computed with the aug-cc-pVDZ basis for the methods including at least triple excitations; the second-order methods, CC2 and CCSD, being unable to faithfully locate these states with a large contribution of double excitations. These transitions can be labeled as “pure” double excitations as they involve an insignificant amount of single excitations ( $\%T_1 \approx 0$ ), hence providing a very stringent test for the EOM-CC formalism. In this case, as shown in Fig. 2, the differences between methods are magnified, but the conclusions drawn in the previous paragraph hold: CC4 is an excellent approximation to CCSDTQ (with a maximum deviation of 0.1 eV for the <sup>1</sup>Δ<sub>g</sub> transition of C<sub>2</sub>) and a massive improvement over CC3 (where the error can be as large as 0.9 eV) and, to a lesser extent, over CCSDT. However, for these

transitions with a dominant double excitation character, CC4 does not permit to reach chemical accuracy with errors of the order of 0.1 eV compared to CCSDTQ and FCI. The outcome might differ for transitions of mixed characters ( $\%T_1 \approx 70$ ) such as the well-known <sup>1</sup>A<sub>g</sub> excited state of butadiene.<sup>63,81–85</sup>

## IV. CONCLUSION

Thanks to the results gathered in the present study, we can conclude, for this set of small molecules at least, that CC4 is a rather competitive approximation to its more expensive CCSDTQ parent as well as a very significant improvement over both its third-order version, CC3, and the “complete” CCSDT method. This is particularly true in the case of transitions with a dominant contribution from the single excitations (Table II) when one reaches sub-kJ mol<sup>-1</sup> accuracy. For states with a dominant contribution from the double excitations, we have seen (Table III) that the same qualitative conclusions hold but one cannot reach chemical accuracy for the set of “pure” double excitations that we have considered.

These findings are promising, though we are well aware that the conclusions obtained for small and larger molecules might differ significantly. For example, CCSD outperforms CC2 for compact molecules (as here), but the opposite trend is often found for larger compounds.<sup>61,70,71</sup> Therefore, although further investigations on larger compounds are definitely required, the present results are very encouraging as CC4, with its  $O(N^9)$  scaling, can be applied to significantly larger molecules than CCSDTQ [which scales as  $O(N^{10})$ ]. This will likely allow us to revisit, in the future, some of the theoretical best estimates defined in the QUEST database.<sup>70,71</sup>

## ACKNOWLEDGMENTS

PFL has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement No. 863481). DAM has received funding from the National Science Foundation (grant No. OAC-2003931). This work was performed using HPC resources from CALMIP (Toulouse) under allocation 2021-18005 and of the CCIPL computational center installed in Nantes.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Zenodo at <http://doi.org/10.5281/zenodo.4739288>.

## REFERENCES

- <sup>1</sup>J. Čížek, *J. Chem. Phys.* **45**, 4256 (1966).
- <sup>2</sup>J. Paldus, J. Čížek, and I. Shavitt, *Phys. Rev. A* **5**, 50 (1972).
- <sup>3</sup>T. D. Crawford and H. F. Schaefer, in *Reviews in Computational Chemistry* (John Wiley & Sons, Ltd, 2000) pp. 33–136.

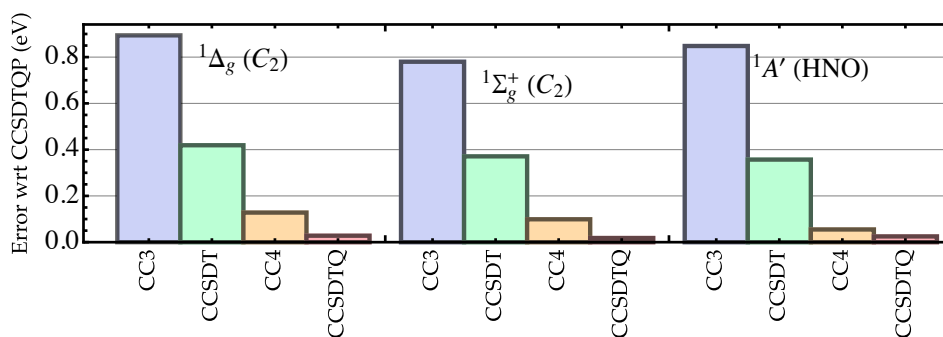


FIG. 2. Error in excitation energies (with respect to FCI) computed with CC3, CCSDT, CC4, and CCSDTQ in the aug-cc-pVDZ basis for transitions with a dominant contribution from the double excitations gathered in Table III.

- <sup>4</sup>R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**, 291 (2007).
- <sup>5</sup>I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*, Cambridge Molecular Science (Cambridge University Press, Cambridge, 2009).
- <sup>6</sup>G. D. Purvis and R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).
- <sup>7</sup>J. Noga and R. J. Bartlett, *J. Chem. Phys.* **86**, 7041 (1987).
- <sup>8</sup>G. E. Scuseria and H. F. Schaefer, *Chem. Phys. Lett.* **152**, 382 (1988).
- <sup>9</sup>N. Oliphant and L. Adamowicz, *J. Chem. Phys.* **95**, 6645 (1991).
- <sup>10</sup>S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **97**, 4282 (1992).
- <sup>11</sup>S. Hirata and R. J. Bartlett, *Chem. Phys. Lett.* **321**, 216 (2000).
- <sup>12</sup>M. Kállay and P. R. Surján, *J. Chem. Phys.* **115**, 2945 (2001).
- <sup>13</sup>O. Christiansen, H. Koch, and P. Jørgensen, *Chem. Phys. Lett.* **243**, 409 (1995).
- <sup>14</sup>O. Christiansen, H. Koch, and P. Jørgensen, *J. Chem. Phys.* **103**, 7429 (1995).
- <sup>15</sup>H. Koch, O. Christiansen, P. Jørgensen, A. M. Sanchez de Merás, and T. Helgaker, *J. Chem. Phys.* **106**, 1808 (1997).
- <sup>16</sup>M. Kállay and J. Gauss, *J. Chem. Phys.* **123**, 214105 (2005).
- <sup>17</sup>O. Christiansen, P. Jørgensen, and C. Hättig, *Int. J. Quantum Chem.* **68**, 1 (1998).
- <sup>18</sup>M. Urban, J. Noga, S. J. Cole, and R. J. Bartlett, *J. Chem. Phys.* **83**, 4041 (1985).
- <sup>19</sup>J. Noga, R. J. Bartlett, and M. Urban, *Chem. Phys. Lett.* **134**, 126 (1987).
- <sup>20</sup>P. Piecuch, K. Kowalski, I. S. O. Pimienta, and M. J. Mcguire, *Int. Rev. Phys. Chem.* **21**, 527 (2002).
- <sup>21</sup>M. Kállay, J. Gauss, and P. G. Szalay, *J. Chem. Phys.* **119**, 2991 (2003).
- <sup>22</sup>M. Kállay and J. Gauss, *J. Chem. Phys.* **120**, 6841 (2004).
- <sup>23</sup>J. Gauss, A. Tajti, M. Kállay, J. F. Stanton, and P. G. Szalay, *J. Chem. Phys.* **125**, 144111 (2006).
- <sup>24</sup>J. Gauss, M. Kállay, and F. Neese, *J. Phys. Chem. A* **113**, 11541 (2009).
- <sup>25</sup>M. Kállay and J. Gauss, *J. Mol. Struct. THEOCHEM* **768**, 71 (2006).
- <sup>26</sup>H. J. Monkhorst, *Int. J. Quantum Chem.* **12**, 421 (1977).
- <sup>27</sup>K. Emrich, *Nuc. Phys. A* **351**, 379 (1981).
- <sup>28</sup>D. C. Comeau and R. J. Bartlett, *Chem. Phys. Lett.* **207**, 414 (1993).
- <sup>29</sup>J. F. Stanton and R. J. Bartlett, *J. Chem. Phys.* **98**, 7029 (1993).
- <sup>30</sup>A. I. Krylov, *Annu. Rev. Phys. Chem.* **59**, 433 (2008).
- <sup>31</sup>S. Ghosh, D. Mukherjee, and S. Bhattacharyya, *Mol. Phys.* **43**, 173 (1981).
- <sup>32</sup>H. Sekino and R. J. Bartlett, *Int. J. Quantum Chem.* **26**, 255 (1984).
- <sup>33</sup>H. Koch and P. Jørgensen, *J. Chem. Phys.* **93**, 3333 (1990).
- <sup>34</sup>H. Koch, H. J. A. Jensen, P. Jørgensen, and T. Helgaker, *J. Chem. Phys.* **93**, 3345 (1990).
- <sup>35</sup>R. J. Rico and M. Head-Gordon, *Chem. Phys. Lett.* **213**, 224 (1993).
- <sup>36</sup>K. Kowalski and P. Piecuch, *J. Chem. Phys.* **115**, 643 (2001).
- <sup>37</sup>S. A. Kucharski, M. Włoch, M. Musiał, and R. J. Bartlett, *J. Chem. Phys.* **115**, 8263 (2001).
- <sup>38</sup>S. Hirata, *J. Chem. Phys.* **121**, 51 (2004).
- <sup>39</sup>M. Kállay and J. Gauss, *J. Chem. Phys.* **121**, 9257 (2004).
- <sup>40</sup>C. F. Bender and E. R. Davidson, *Phys. Rev.* **183**, 23 (1969).
- <sup>41</sup>J. L. Whitten and M. Hackmeyer, *J. Chem. Phys.* **51**, 5584 (1969).
- <sup>42</sup>B. Huron, J. P. Malrieu, and P. Rancurel, *J. Chem. Phys.* **58**, 5745 (1973).
- <sup>43</sup>G. H. Booth, A. J. W. Thom, and A. Alavi, *J. Chem. Phys.* **131**, 054106 (2009).
- <sup>44</sup>J. E. Deustua, J. Shen, and P. Piecuch, *Phys. Rev. Lett.* **119**, 223003 (2017).
- <sup>45</sup>A. A. Holmes, N. M. Tubman, and C. J. Umrigar, *J. Chem. Theory Comput.* **12**, 3674 (2016).
- <sup>46</sup>J. J. Eriksen, F. Lipparini, and J. Gauss, *J. Phys. Chem. Lett.* **8**, 4633 (2017).
- <sup>47</sup>E. Xu, M. Uejima, and S. L. Ten-no, *Phys. Rev. Lett.* **121**, 113001 (2018).
- <sup>48</sup>E. Giner, A. Scemama, and M. Caffarel, *Can. J. Chem.* **91**, 879 (2013).
- <sup>49</sup>E. Giner, A. Scemama, and M. Caffarel, *J. Chem. Phys.* **142**, 044115 (2015).
- <sup>50</sup>F. A. Evangelista, *J. Chem. Phys.* **140**, 124114 (2014).
- <sup>51</sup>N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, and K. B. Whaley, *J. Chem. Phys.* **145**, 044112 (2016).
- <sup>52</sup>W. Liu and M. R. Hoffmann, *J. Chem. Theory Comput.* **12**, 1169 (2016).
- <sup>53</sup>Y. Ohtsuka and J.-y. Hasegawa, *J. Chem. Phys.* **147**, 034102 (2017).
- <sup>54</sup>P. M. Zimmerman, *J. Chem. Phys.* **146**, 104102 (2017).
- <sup>55</sup>J. P. Coe, *J. Chem. Theory Comput.* **14**, 5739 (2018).
- <sup>56</sup>Y. Garniron, A. Scemama, E. Giner, M. Caffarel, and P. F. Loos, *J. Chem. Phys.* **149**, 064103 (2018).
- <sup>57</sup>J. J. Eriksen and J. Gauss, *J. Phys. Chem. Lett.* **27**, 7910 (2019).
- <sup>58</sup>A. A. Holmes, C. J. Umrigar, and S. Sharma, *J. Chem. Phys.* **147**, 164111 (2017).
- <sup>59</sup>J. Li, M. Otten, A. A. Holmes, S. Sharma, and C. J. Umrigar, *J. Chem. Phys.* **149**, 214110 (2018).
- <sup>60</sup>J. Li, Y. Yao, A. A. Holmes, M. Otten, Q. Sun, S. Sharma, and C. J. Umrigar, *Phys. Rev. Research* **2**, 012015 (2020).
- <sup>61</sup>P. F. Loos, A. Scemama, A. Blondel, Y. Garniron, M. Caffarel, and D. Jacquemin, *J. Chem. Theory Comput.* **14**, 4360 (2018).
- <sup>62</sup>A. D. Chien, A. A. Holmes, M. Otten, C. J. Umrigar, S. Sharma, and P. M. Zimmerman, *J. Phys. Chem. A* **122**, 2714 (2018).
- <sup>63</sup>P.-F. Loos, M. Boggio-Pasqua, A. Scemama, M. Caffarel, and D. Jacquemin, *J. Chem. Theory Comput.* **15**, 1939 (2019).
- <sup>64</sup>P. F. Loos, F. Lipparini, M. Boggio-Pasqua, A. Scemama, and D. Jacquemin, *J. Chem. Theory Comput.* **16**, 1711 (2020).
- <sup>65</sup>P. F. Loos, A. Scemama, M. Boggio-Pasqua, and D. Jacquemin, *J. Chem. Theory Comput.* **16**, 3720 (2020).
- <sup>66</sup>P.-F. Loos, Y. Damour, and A. Scemama, *J. Chem. Phys.* **153**, 176101 (2020).
- <sup>67</sup>Y. Garniron, K. Gasperich, T. Applencourt, A. Benali, A. Ferté, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, P. Barbaresco, N. Renon, G. David, J. P. Malrieu, M. Vénil, M. Caffarel, P. F. Loos, E. Giner, and A. Scemama, *J. Chem. Theory Comput.* **15**, 3591 (2019).
- <sup>68</sup>J. J. Eriksen, T. A. Anderson, J. E. Deustua, K. Ghanem, D. Hait, M. R. Hoffmann, S. Lee, D. S. Levine, I. Magoulas, J. Shen, N. M. Tubman, K. B. Whaley, E. Xu, Y. Yao, N. Zhang, A. Alavi, G. K.-L. Chan, M. Head-Gordon, W. Liu, P. Piecuch, S. Sharma, S. L. Ten-no, C. J. Umrigar, and J. Gauss, *J. Phys. Chem. Lett.* **11**, 8922 (2020).
- <sup>69</sup>Y. Yao, E. Giner, J. Li, J. Toulouse, and C. J. Umrigar, "Almost exact energies for the Gaussian-2 set with the semistochastic heat-bath configuration interaction method." (2020), [arXiv:2004.10059](https://arxiv.org/abs/2004.10059) [physics.chem-ph].
- <sup>70</sup>M. Vénil, A. Scemama, M. Caffarel, F. Lipparini, M. Boggio-Pasqua, D. Jacquemin, and P.-F. Loos, *WIREs Comput. Mol. Sci.* **n/a**, e1517.

- <sup>71</sup>P. F. Loos, A. Scemama, and D. Jacquemin, *J. Phys. Chem. Lett.* **11**, 2374 (2020).
- <sup>72</sup>J. J. Eriksen, *J. Phys. Chem. Lett.* **12**, 418 (2021).
- <sup>73</sup>Y. Garniron, A. Scemama, P.-F. Loos, and M. Caffarel, *J. Chem. Phys.* **147**, 034101 (2017).
- <sup>74</sup>S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar, *J. Chem. Theory Comput.* **13**, 1595 (2017).
- <sup>75</sup>D. A. Matthews, *J. Chem. Theory Comput.* **16**, 6195 (2020).
- <sup>76</sup>D. A. Matthews, L. Cheng, M. E. Harding, F. Lipparini, S. Stopkowitz, T.-C. Jagau, P. G. Szalay, J. Gauss, and J. F. Stanton, *J. Chem. Phys.* **152**, 214108 (2020).
- <sup>77</sup>M. Kállay, P. R. Nagy, D. Mester, Z. Rolik, G. Samu, J. Csontos, J. Csóka, P. B. Szabó, L. Gyevi-Nagy, B. Hégyely, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P. D. Mezei, and Á. Ganyecz, *J. Chem. Phys.* **152**, 074107 (2020).
- <sup>78</sup>D. A. Matthews and J. F. Stanton, *J. Chem. Phys.* **142**, 064108 (2015).
- <sup>79</sup>E. R. Davidson, *J. Comput. Phys.* **17**, 87 (1975).
- <sup>80</sup>A. Chrayteh, A. Blondel, P. F. Loos, and D. Jacquemin, *J. Chem. Theory Comput.* **17**, 416 (2021).
- <sup>81</sup>N. T. Maitra, R. J. C. F. Zhang, and K. Burke, *J. Chem. Phys.* **120**, 5932 (2004).
- <sup>82</sup>B. Saha, M. Ehara, and H. Nakatsuji, *J. Chem. Phys.* **125**, 014316 (2006).
- <sup>83</sup>M. A. Watson and G. K.-L. Chan, *J. Chem. Theory Comput.* **8**, 4013 (2012).
- <sup>84</sup>Y. Shu and D. G. Truhlar, *J. Am. Chem. Soc.* **139**, 13770 (2017).
- <sup>85</sup>G. M. J. Barca, A. T. B. Gilbert, and P. M. W. Gill, *J. Chem. Theory Comput.* **14**, 1501 (2018).