

# Coupled-cluster calculations of structure and vibrational frequencies of ozone: Are triple excitations enough?

John D. Watts and Rodney J. Bartlett

*Quantum Theory Project, P.O. Box 118435, University of Florida, Gainesville, Florida 32611*

(Received 17 July 1997; accepted 6 November 1997)

Coupled-cluster calculations with full inclusion of singles, doubles, and triples (CCSDT) with a double-zeta plus polarization and correlation-consistent polarized valence triple-zeta basis sets have been used to calculate the structure and harmonic vibrational frequencies of the ozone molecule. These results have been compared with those of more approximate CC methods, and the effects of the different terms in the triple excitation equation are analyzed. The effect of basis set extension on the CCSDT results has been estimated. In the limit of a large basis set, it appears that the complete CCSDT method will give smaller bond lengths than experiment and an asymmetric stretching frequency about  $50\text{ cm}^{-1}$  above the experimental value. That is, it would appear that connected quadruple excitations are needed for quantitative calculations of the structure and frequencies of ozone. © 1998 American Institute of Physics. [S0021-9606(98)03806-9]

## I. INTRODUCTION

The ozone molecule ( $\text{O}_3$ ) has long been considered to be a demanding test case for quantum chemical methods. As it is a textbook multireference example, several single reference coupled-cluster (CC) methods which include triple excitation effects have had mixed results in describing ozone's harmonic force field.<sup>1-5</sup> Lee and Scuseria<sup>4</sup> performed a series of CC singles-and-doubles (CCSD) and CCSD with a perturbative treatment of triples [CCSD(T)] calculations with atomic natural orbital (ANO) basis sets. With a  $4s3p2d1f$  basis set, the CCSD(T) frequencies agreed well with experiment. The largest error was  $32\text{ cm}^{-1}$ , which was for the asymmetric stretching mode ( $\omega_3$ ). Of the three modes, this one is the most sensitive to the correlation treatment.<sup>1</sup> In previous work,<sup>5</sup> we reported the structure and frequencies with a double-zeta plus polarization (DZP) basis set with the complete CC singles, doubles, and triples (CCSDT) method. The frequencies were within  $12\text{ cm}^{-1}$  of experiment, but basis-set extension will reduce the calculated bond lengths and increase the stretching frequencies significantly, yielding poorer agreement with experiment. This implies that with a large basis set, it appears the CCSDT method cannot give a quantitative description of ozone, suggesting the critical importance of connected quadruple excitations. Naively one would expect this, as double excitations from the second important configuration in  $\text{O}_3$  are quadruple excitations from the first. However, we have to distinguish  $T_4$  and  $T_2^2$  contributions. In a study with multireference configuration interaction (MRCI) and complete active space (CAS) methods, Borowski *et al.*<sup>6</sup> obtain good frequencies when their MRCI data are corrected for size inextensivity, and they agree with our suggestion that quadruple excitations may be important. Good results have also been obtained from a multireference averaged quadratic CC (MRAQCC) study, using a two-configuration reference.<sup>7</sup>

In this work, we have performed a complete CCSDT calculation using a significantly larger basis set [the

correlation-consistent polarized valence triple-zeta (cc-pVTZ) set], with the intention of approaching more closely the limit of the CCSDT method. Along with the CCSDT calculation, we have also performed calculations with several iterative approximations to CCSDT, and larger basis set calculations with the CCSD(T) method, which provides a non-iterative treatment of triple excitations.

## II. COMPUTATIONAL METHODS

All of the calculations were performed with the ACES II program system.<sup>8,9</sup> Along with the DZP basis set used previously,<sup>10,11</sup> three types of correlation-consistent (cc) sets were used. The DZP set comprises Dunning's  $4s2p$  contraction, augmented with a  $d$  function with an exponent of 1.211. The first type of cc sets comprised the basic polarized valence sets,<sup>12</sup> namely cc-pVDZ, cc-pVTZ, cc-pVQZ, and cc-pV5Z. The latter was used with and without  $h$  functions (cc-pV5Z' denotes no  $h$  functions). The second type of cc sets comprised those augmented with diffuse functions,<sup>13</sup> namely aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ. The final cc set used was the cc-pCVTZ set,<sup>14</sup> which is triple-zeta in the valence space and contains additional functions for describing core correlation effects. The DZP set was used with Cartesian  $d$  functions, while spherical harmonic polar-

TABLE I. Summary of  $T_3$  equations for iterative CC methods.  $H_N = F_N + W_N$  is the normal-ordered Hamiltonian, the  $T_n$  are cluster operators,  $|0\rangle$  is the reference determinant, and  $|_{ijk}^{abc}\rangle$  is a triply excited determinant. All methods except CCSDT-1a include  $W_N T_1 T_3$  in the  $T_2$  equation.

Method	$T_3$ equation
CCSDT-1a, b	$\langle_{ijk}^{abc}   (F_N T_3 + H_N T_2)_c   0 \rangle = 0$
CCSDT-1c	$\langle_{ijk}^{abc}   (F_N T_3 + H_N T_2 + W_N T_1 T_2)_c   0 \rangle = 0$
CC3	$\langle_{ijk}^{abc}   (F_N T_3 + H_N e^{T_1} T_2)_c   0 \rangle = 0$
CCSDT-2	$\langle_{ijk}^{abc}   (F_N T_3 + H_N T_2^2/2)_c   0 \rangle = 0$
CCSDT-3	$\langle_{ijk}^{abc}   (F_N T_3 + H_N e^{T_1 + T_2})_c   0 \rangle = 0$
CCSDT	$\langle_{ijk}^{abc}   (H_N e^{T_1 + T_2 + T_3})_c   0 \rangle = 0$

TABLE II. Geometry and frequencies of  $X^1A_1$  O<sub>3</sub> (DZP and cc-pVTZ basis sets).

Method	$r_e/\text{\AA}$	$\theta_e/^\circ$	$\omega_1(a_1)/\text{cm}^{-1}$	$\omega_2(a_1)/\text{cm}^{-1}$	$\omega_3(b_2)/\text{cm}^{-1}$
DZP					
CCSD	1.263	117.4	1256	748	1240
CCSDT-1a	1.295	116.6	1076	674	680
CCSDT-1b	1.292	116.6	1098	694	1093
CCSDT-1c	1.294	116.8	1092	687	994
CC3	1.293	116.8	1096	688	1041
CCSDT-2	1.283	116.8	1158	712	1182
CCSDT-3	1.285	116.9	1150	707	1117
CCSDT	1.286	116.7	1141	705	1077
cc-pVTZ					
CCSD	1.250	117.6	1278	763	1266
CCSDT-1a	1.284	116.7	1098	700	845
CCSDT-1b	1.282	116.7	1119	705	1135
CCSDT-1c	1.283	116.9	1110	697	1031
CC3	1.283	116.9	1114	698	1068
CCSDT-2	1.273	116.9	1174	723	1216
CCSDT-3	1.274	117.0	1165	717	1142
CCSDT	1.274	116.8	1163	717	1117
Expt.	1.272	116.8	1135	716	1089

ization functions were used for all other sets. In all calculations except those with the cc-pCVTZ set, the core electrons were not correlated. In the DZP calculations the highest three virtual orbitals were discarded.

Several iterative CC methods were used, CCSD,<sup>15</sup> CCSDT-1a,<sup>16</sup> CCSDT-1b,<sup>17</sup> CCSDT-1c,<sup>18</sup> CC3,<sup>18</sup> CCSDT-2,<sup>19</sup> CCSDT-3,<sup>19</sup> and CCSDT.<sup>20</sup> The CCSD method is a complete treatment of single and double excitation clusters and is exact for two electron systems. The CCSDT method is a complete treatment of single, double, and triple excitation clusters and is exact for three electron systems. The other methods are iterative approximations to CCSDT, in which varying numbers of terms in the  $T_3$  equation are neglected. They are summarized in Table I.

The noniterative CCSD(T) method<sup>2</sup> was used with the various cc basis sets.

### III. RESULTS

Table II contains structures and harmonic vibrational frequencies obtained from the iterative CC methods with the DZP and cc-pVTZ basis sets. Table III contains CCSD(T)

results with the various cc basis sets. The three vibrational modes are symmetric stretching ( $\omega_1$ ), bending ( $\omega_2$ ), and asymmetric stretching ( $\omega_3$ ).

We consider the DZP basis set results in Table II first. The DZP basis set results for the CCSD, CCSDT-1a, CCSDT-2, CCSDT-3, and CCSDT methods were reported previously (see Ref. 5 and references therein). The CCSDT-1a method contains the lowest-order contribution to the  $T_3$  equation and its value for  $\omega_3$  is much lower than those given by the other methods. The  $T_3$  equations for CCSDT-1a and CCSDT-1b are the same, but the latter method includes the  $T_1T_3$  term in the  $T_2$  equation. This term is normally of minor significance, but in this case it leads to smaller bond lengths than CCSDT-1a and a significantly improved value of  $\omega_3$ . The CCSDT-1c method extends CCSDT-1b by including the lowest-order  $T_1$  contribution to the  $T_3$  equation.<sup>21</sup> This term increases the bond length marginally, but lowers  $\omega_3$  by  $99\text{ cm}^{-1}$ . The CC3 method includes all terms containing  $T_1$  and linear  $T_2$ . The CC3 structure is almost the same as that of CCSDT-1c, and the values of  $\omega_1$  and  $\omega_2$  given by the two methods are virtually identical, but

TABLE III. Geometry and frequencies of  $X^1A_1$  O<sub>3</sub> with the CCSD(T) method and several basis sets.

Method	$r_e/\text{\AA}$	$\theta_e/^\circ$	$\omega_1(a_1)/\text{cm}^{-1}$	$\omega_2(a_1)/\text{cm}^{-1}$	$\omega_3(b_2)/\text{cm}^{-1}$
cc-pVDZ	1.284	116.6	1118	704	977
cc-pVTZ	1.275	116.9	1153	716	1054
cc-pVQZ	1.269	117.1	1169	725	1081
cc-pV5Z'	1.268	117.1	1169	725	1079
cc-pV5Z	1.267	117.1			
aug-cc-pVDZ	1.285	116.6	1115	703	970
aug-cc-pVTZ	1.276	117.1	1147	714	1038
aug-cc-pVQZ	1.269	117.2	1168	724	1077
cc-pCVTZ	1.273	117.0	1158	720	1062
Expt.	1.272	116.8	1135	716	1089

the CC3 value of  $\omega_3$  is  $47\text{ cm}^{-1}$  closer to the CCSDT result. The CCSDT-2 method adds the  $H_N T_2^2/2$  term to the CCSDT-1b method and one can see that this significantly reduces the bond length and increases the frequencies, particularly  $\omega_3$ . For a system like ozone, with a large contribution to the wave function from a doubly excited configuration, one might expect the  $T_2^2$  term to be important, as first demonstrated by Magers *et al.*<sup>1</sup> Compared with CCSDT, the CCSDT-2 method overcorrects the bond length, but it is apparent that the  $T_2^2$  term is the single term most responsible for reducing the bond length. Without this term, it is not possible to approach the CCSDT structure. Going from CCSDT-2 to CCSDT-3, the bond length is increased slightly and the frequencies are reduced and are closer to CCSDT. Of all methods considered, the CCSDT-3 structure,  $\omega_1$ , and  $\omega_2$  are closest to those of CCSDT, and the improvements relative to CCSDT-1a, CCSDT-1c, CC3, and CCSDT-2 are significant. However, the CCSDT-3 value of  $\omega_3$  is  $40\text{ cm}^{-1}$  greater than the CCSDT value, while the CC3 result is  $36\text{ cm}^{-1}$  below the CCSDT result. As noted previously,<sup>5</sup> the CCSDT/DZP frequencies are in excellent agreement with experiment, but the bond length is not, which illustrates the basis set inadequacy.

Now we consider the cc-pVTZ basis set results in Table II. Many of the general trends shown by the DZP basis set results are repeated for the cc-pVTZ basis set. The CCSDT-1a, CCSDT-1b, CCSDT-1c, and CC3 structures are all similar, with the CC3 and CCSDT-1b frequencies, especially  $\omega_3$ , closer to the CCSDT results. Once again, the CCSDT-3 structure,  $\omega_1$ , and  $\omega_2$  agree very well with CCSDT results. In contrast with the results for the DZP basis set, the CCSDT-3 value for  $\omega_3$  is closer to the CCSDT result than is the CC3 value.

The CCSDT/cc-pVTZ structure is much closer to experiment than is the CCSDT/DZP structure, as expected. In accord with the decrease in bond lengths, the CCSDT/cc-pVTZ stretching frequencies,  $\omega_1$  and  $\omega_3$ , are higher than the DZP values, and the agreement with experiment is not as good, as previously anticipated.<sup>5</sup> This points to the importance of connected quadruple excitations ( $T_4$ ).

There are several sources of error in the CCSDT/cc-pVTZ results. The first of these is the one-particle basis set. An estimate of this can be made by comparing results of CCSD(T) calculations with a series of cc basis sets (Table III). Going from cc-pVTZ to cc-pVQZ leads to an appreciable bond length reduction ( $0.006\text{ \AA}$ ), so that the CCSD(T)/cc-pVQZ bond length is below experiment, and increases in  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$  by 16, 9, and  $27\text{ cm}^{-1}$ , respectively. Going beyond cc-pVQZ (to cc-pV5Z or aug-cc-pVQZ) leads to much smaller changes in geometry and frequencies. Assuming the same basis set effects for CCSD(T) and CCSDT, we estimate the CCSDT/cc-pVQZ structure to be  $r_e = 1.268\text{ \AA}$ ,  $\theta_e = 117.0^\circ$  and the frequencies to be  $\omega_1 = 1179$ ,  $\omega_2 = 726$ , and  $\omega_3 = 1144\text{ cm}^{-1}$ . These should be very close to the complete valence basis set results for CCSDT. Based on our CCSDT/DZP results<sup>5</sup> and basis set effects for other methods, Borowski *et al.*<sup>6</sup> estimated large basis set CCSDT results of  $r_e = 1.268\text{ \AA}$ ,  $\theta_e = 117.0^\circ$ ,  $\omega_1 = 1177\text{ cm}^{-1}$ ,  $\omega_2 = 731\text{ cm}^{-1}$ ,

and  $\omega_3 = 1140\text{ cm}^{-1}$ , which are close to our estimated CCSDT/cc-pVQZ results.

The second source of error is the neglect of core correlation. Effects of core correlation should be qualitatively the same as the basis set error, again decreasing bond lengths and increasing stretching frequencies. To investigate, we performed CCSD(T) calculations with the cc-pCVTZ basis set, and the results are shown in Table III. Comparing the cc-pVTZ and cc-pCVTZ results, we see a small reduction in bond length ( $0.002\text{ \AA}$ ) and small increases in frequencies.

It seems evident that complete basis set CCSDT results will differ significantly from experiment. It appears that  $r_e$  will be underestimated by about  $0.006\text{ \AA}$ , and  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$  will be overestimated by about 40, 15, and  $50\text{ cm}^{-1}$ , respectively. Neglecting relativistic effects and errors in the Born–Oppenheimer approximation, the remaining errors are the correlation treatment. Apparently, quadruple and higher excitations must increase the bond lengths and decrease the frequencies appreciably.

#### IV. CONCLUSIONS

A series of iterative approximations to the CCSDT method have been used to calculate the structure and harmonic vibrational frequencies of  $\text{O}_3$ . The calculations have been performed with DZP and cc-pVTZ basis sets. The DZP calculations extend previous studies of this type,<sup>5</sup> while the cc-pVTZ calculations provide a set of complete CCSDT results with a reasonably large valence basis set, developed for correlated methods.

The effects of different types of terms in the  $T_3$  equation are elucidated. It is seen that as one progresses from the simplest approximations to CCSDT, different terms have opposing effects. The  $T_2^2$  term reduces the bond length and increases frequencies; the lowest-order term containing  $T_1$ , namely  $W_N T_1 T_2$ , increases the bond length and reduces frequencies; further terms containing  $T_1$  and  $T_2$  reduce the bond length and increase the frequencies; and the terms in the  $T_3$  equation which are not in the CCSDT-3 method slightly increase the bond length and reduce the frequencies. Of the different approximate methods employed, the CCSDT-3 method achieves closest agreement with CCSDT.

An analysis of basis set effects gives an improved estimate of CCSDT results with a complete basis set. This indicates that it will be necessary to include connected quadruple excitations to describe the structure and frequencies quantitatively.

#### ACKNOWLEDGMENTS

This work has been supported in part by the United States Air Force Office of Scientific Research. The Ohio Supercomputer Center is thanked for provision of facilities for the development of the ACES II program system. We thank the High Performance Computing Center at the USAE Waterways Experiment Station, Vicksburg, MS for facilities to perform the larger basis set CCSD(T) calculations.

- <sup>1</sup>D. H. Magers, W. N. Lipscomb, R. J. Bartlett, and J. F. Stanton, *J. Chem. Phys.* **91**, 1945 (1989); J. F. Stanton, W. W. Lipscomb, D. H. Magers, and R. J. Bartlett, *ibid.* **90**, 1077 (1989).
- <sup>2</sup>K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
- <sup>3</sup>K. Raghavachari, G. W. Trucks, J. A. Pople, and E. Replogle, *Chem. Phys. Lett.* **158**, 207 (1989).
- <sup>4</sup>T. J. Lee and G. E. Scuseria, *J. Chem. Phys.* **93**, 489 (1990).
- <sup>5</sup>J. D. Watts, J. F. Stanton, and R. J. Bartlett, *Chem. Phys. Lett.* **178**, 471 (1991).
- <sup>6</sup>P. Borowski, K. Andersson, P.-Å Malmqvist, and B. O. Roos, *J. Chem. Phys.* **97**, 5568 (1992).
- <sup>7</sup>P. G. Szalay and R. J. Bartlett, *Chem. Phys. Lett.* **214**, 481 (1993).
- <sup>8</sup>ACES II is a program product of the Quantum Theory Project, University of Florida, J. F. Stanton, J. Gauss, J. D. Watts, M. Nooijen, N. Oliphant, S. A. Perera, P. G. Szalay, W. J. Lauderdale, S. R. Gwaltney, S. Beck, A. Balková, D. E. Bernholdt, K.-K. Baeck, P. Rozyczko, H. Sekino, C. Huber, and R. J. Bartlett. Integral packages included are VMOL (J. Almlöf and P. R. Taylor); VPROPS (P. R. Taylor); and ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, J. Olsen, and P. R. Taylor).
- <sup>9</sup>J. F. Stanton, J. Gauss, J. D. Watts, W. J. Lauderdale, and R. J. Bartlett, *Int. J. Quantum Chem. Symp.* **26**, 879 (1992).
- <sup>10</sup>T. H. Dunning, Jr., *J. Chem. Phys.* **53**, 2823 (1970).
- <sup>11</sup>L. T. Redmon, G. D. Purvis III, and R. J. Bartlett, *J. Am. Chem. Soc.* **101**, 2856 (1979).
- <sup>12</sup>T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- <sup>13</sup>R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6769 (1992).
- <sup>14</sup>D. E. Woon and T. H. Dunning, Jr., *J. Chem. Phys.* **103**, 4572 (1995).
- <sup>15</sup>G. D. Purvis III and R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).
- <sup>16</sup>Y. S. Lee, S. A. Kucharski, and R. J. Bartlett, *J. Chem. Phys.* **81**, 5096 (1984).
- <sup>17</sup>M. Urban, J. Noga, S. J. Cole, and R. J. Bartlett, *J. Chem. Phys.* **83**, 4041 (1985).
- <sup>18</sup>H. Koch, O. Christiansen, P. Jørgensen, A. Sanchez de Merás, and T. Helgaker, *J. Chem. Phys.* **106**, 1808 (1997).
- <sup>19</sup>J. Noga, R. J. Bartlett, and M. Urban, *Chem. Phys. Lett.* **134**, 126 (1987).
- <sup>20</sup>J. Noga and R. J. Bartlett, *J. Chem. Phys.* **86**, 7041 (1987); G. E. Scuseria and H. F. Schaefer III, *Chem. Phys. Lett.* **152**, 382 (1988); J. D. Watts and R. J. Bartlett, *Int. J. Quantum Chem. Symp.* **27**, 51 (1993).
- <sup>21</sup>This is the ground state reference of a method introduced for excited states, EOM-CCSDT-1c and EOM-CCSD(T') J. D. Watts and R. J. Bartlett, *Chem. Phys. Lett.* **258**, 581 (1996).