

Toward the limits of predictive electronic structure theory: Connected quadruple excitations for large basis set calculations

Stanisław A. Kucharski^{a)}

Quantum Theory Project, Department of Chemistry, University of Florida, Gainesville, Florida 32611

Maciej Kolaski

Institute of Chemistry, Silesian University, Szkolna 9, 40-006 Katowice, Poland

Rodney J. Bartlett

Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, Florida 32611

(Received 27 January 2000; accepted 21 June 2000)

The general inclusion of the T_4 operator into the coupled cluster equations requires an n^{10} computational procedure, and n^9 in the lowest order, as in the CCSDTQ-1 (coupled cluster singles, doubles, triples, and lowest order quadruples) method. Coupled cluster methods with full inclusion of singles, doubles, triples, and an efficient noniterative inclusion of connected quadruples (CCSDT(Q_f)) have been introduced in [J. Chem. Phys. **108**, 9221 (1998)]. Since the connected quadruple part in the latter method scales as n^7 (CCSDT itself is n^8) it offers an attractive route to assess the connected quadruple contribution for larger basis sets. We present a detailed description of the Q_f algorithm with explicit algebraic formulas for the spin-orbital formalism as well as for a nonorthogonal spin adapted approach. The method has been applied to obtain the equilibrium geometry and harmonic frequencies for the C_2 molecule for a sequence of correlation consistent polarized (core) valence (cc-p(C)VXZ, X=D,T,Q,5) basis sets. For the largest basis sets, cc-pCVQZ and cc-pV5Z, the connected quadruple excitations lower the harmonic frequency by 10 cm^{-1} and raise the bond length by 0.0014 \AA , providing results that agree with experiment to 3 cm^{-1} and 0.0003 \AA . © 2001 American Institute of Physics. [DOI: 10.1063/1.1288917]

I. INTRODUCTION

It is commonly believed that the coupled cluster (CC) (Refs. 1–14) method is an excellent theoretical tool for the high accuracy prediction of properties of molecules. There are many examples of accurate calculations with an inclusion of single and double excitations only (CCSD),^{2–6} but a substantial improvement of the results has been achieved upon inclusion into the CC equations the connected triple excitations.^{7–11} When this is done in an iterative way with a complete incorporation of the T_3 operator we have CCSDT.^{7–9} Its numerous applications demonstrate that for the majority of molecules studied, the correlation energy is very close to the full CI value (i.e., the theoretical limit), usually within a few tenths of a mHartree. Very good results are also obtained with approximate iterative variants of the CCSDT method: CCSDT-1,¹⁰ CCSDT-2, and CCSDT-3.¹¹ The advantage of using CCSDT- n iterative approaches is that they scale with the size of the system as n^7 instead of the n^8 scaling of the full CCSDT. However, most applications are made with the noniterative approximation to the CCSDT (or CCSDT-1) known as CCSD(T).^{12,13} The latter method, along with its earlier incarnation, CCSD[T],¹² has only one step scaling as n^7 , making it significantly more efficient than the iterative formulations. All these methods work very well

for geometries close to equilibrium, or, more generally, when nondynamical correlation effects are not dominant. However, when the reference function is of multiconfigurational character, as in bond breaking and certain other situations,⁵ nondynamical effects play a significant role, and multireference approaches^{14–20} would appear to be the better solution. However, the true multireference formulations are more complicated, are more likely to suffer from convergence and intruder state problems, and are not as applicable in a “black box” form. The alternative is to work within a single reference formalism and to include into CC theory sufficiently high cluster operators to account for important nondynamical correlation along with the dynamical correlation.

To illustrate, the simplest multireference problem is when the reference function is composed of two determinants differing by a double excitation. For a reference function of this type, the nondynamical effects that are not addressed by T_2^2 or other disconnected products of lower rank clusters, can be assessed by inclusion into the CC expansion of the connected T_4 cluster operator. Unfortunately, the construction of the T_4 operator within a CC theory is computationally very demanding, as it requires an n^{10} algorithm for the full account and n^9 for the approximate (CCSDTQ-1) account of the T_4 operator. One way proposed to circumvent this difficulty is to extract information about the T_4 cluster from an external source such as the projected unrestricted Hartree–Fock method^{21,22} and various CI methods.^{23–29} A similar procedure is followed by Jankowski *et al.*^{30,31} in his

^{a)}Permanent address: Institute of Chemistry, Silesian University, Szkolna 9, Katowice, Poland.

split amplitude strategy when one part of the amplitude is imported from the external source while the other is obtained by solving simplified CC equations.

The approach adopted in our laboratory is different. It is based on the computation of the high-order contribution solely within the CC theory, i.e., without importing the T_4 amplitudes from external sources. In addition, we require that the scheme would be capable of recovering a substantial portion of the T_4 correlation, i.e., at least at the CCSDTQ-1 level. The proposed method has been described in Ref. 32 and later successfully applied in very accurate calculations of the N_2 equilibrium geometry and harmonic frequency³³ and to the interpretation of the harmonic force field in the ozone molecule.³⁴ This approach is also complementary to active-orbital variants discussed elsewhere.³⁵

The methodology developed in Ref. 32 is to construct a connected quadruple analog of the CCSD(T) method which could be denoted as CCSDT(Q). The main concept in this approach is to solve the CC equations for the T_1 , T_2 , and T_3 amplitudes (not necessary at the full CCSDT level) and then to obtain a T_4 contribution in a noniterative way, evaluated at the lowest order. An important advantage of this method is the possibility to exploit a factorization procedure³⁶ to remove the ‘‘long’’ T_4 denominator. Subsequently, we apply the procedure classified in Ref. 37 as vertical factorization which is effected by splitting a diagram (containing only ‘‘short’’ denominators) into parts and by doing calculations for each part independently. This procedure has already been exploited in the expectation value coupled cluster theory (XCC),^{38,39} where XCC(5), i.e., XCC correct through the fifth-order MBPT energy, was programmed.³⁹ A similar scheme was also employed in MBPT(5) (Ref. 40) and MBPT(6) (Ref. 41) calculations as well as in the development of a method correct through fifth order based on the iterative solution of the CCSD equations.⁴² In the next section we present a derivation of the coupled cluster equations involving the connected quadruple excitation operator.

II. COMPUTATIONAL METHODS

In order to present the hierarchy of approximations used in the current work transparently, we begin with the presentation of the equations defining the full CCSDTQ method.^{43,44} To do so we first construct the CCSDT equations which can be written as

$$\langle \Phi_{ij}^a | (H_N(1 + T_1 + T_2 + T_3 + T_1^2/2 + T_1T_2 + T_1^3/6))_c | 0 \rangle = 0, \quad (1)$$

$$\langle \Phi_{ij}^{ab} | (H_N(1 + T_1 + T_2 + T_3 + T_1^2/2 + T_1T_2 + T_1T_3 + T_2^2/2 + T_1^3/6 + T_1^2T_2/2 + T_1^4/24))_c | 0 \rangle = 0, \quad (2)$$

$$\langle \Phi_{ijk}^{abc} | (H_N(T_2 + T_3 + T_1T_2 + T_1T_3 + T_2^2/2 + T_2T_3 + T_1^2T_2/2 + T_1^2T_3/2 + T_1T_2^2/2 + T_1^3T_2/6))_c | 0 \rangle = 0, \quad (3)$$

where $H_N = W_N + F_N$ is the normal-ordered Hamiltonian, the T_n are cluster operators, $T_n = (n!)^{-2} \sum t_{ij\dots}^{ab\dots} \{a^\dagger b^\dagger \dots j i\}$; $|0\rangle$ is the reference determinant, and $\langle \dots |$ is an n -tuply excited determinant. On the basis of Eqs. (1)–(3) we formulate the CCSDTQ approach,

$$T_1(\text{CCSDTQ}) = T_1(\text{CCSDT}), \quad (4)$$

$$T_2(\text{CCSDTQ}) = T_2(\text{CCSDT}) + R_2(W_N T_4), \quad (5)$$

$$T_3(\text{CCSDTQ}) = T_3(\text{CCSDT}) + R_3(H_N T_4) + R_3(W_N T_1 T_4)_c, \quad (6)$$

$$T_4(\text{CCSDTQ}) = R_4(H_N(T_3 + T_4 + T_1T_3 + T_1T_4 + T_2^2/2 + T_2T_3 + T_2T_4 + T_3^2/2 + T_1^2T_3/2 + T_1^2T_4/2 + T_1T_2^2/2 + T_1T_2T_3 + T_2^3/6 + T_1^2T_2^2/4 + T_1^3T_3/6))_c, \quad (7)$$

where $T_n(\text{CCSDT})$ represents the set of terms occurring in the T_n equation of the CCSDT method, in Eqs. (1)–(3), and the R_n operator is defined as

$$R_n(X) = (n!)^{-2} \sum \frac{\langle \Phi_{ij\dots}^{ab\dots} | X | 0 \rangle}{e_i + e_j + \dots - e_b - e_a} \{a^\dagger b^\dagger \dots j i\} \quad (8)$$

to ensure the presence of the required denominator and the proper projection subspace for the sequence of operators represented by X .

The most natural approximation to the CCSDTQ scheme would be to retain in Eqs. (4)–(7) those terms which would make the resulting method correct through fifth order. To achieve that we need to introduce two major additions to the CCSDT equations; add a new term into the T_2 equations and to construct the simplified T_4 equation, leaving the T_3 equation unaltered. (The T_1 equation is complete at the CCSDT level and cannot be altered for higher rank methods.) In such a way we obtain the equations defining the CCSDTQ-1 approximation,⁴⁵

$$T_1(\text{CCSDTQ-1}) = T_1(\text{CCSDT}), \quad (9)$$

$$T_2(\text{CCSDTQ-1}) = T_2(\text{CCSDT}) + R_2(W_N T_4), \quad (10)$$

$$T_3(\text{CCSDTQ-1}) = T_3(\text{CCSDT}), \quad (11)$$

$$T_4(\text{CCSDTQ-1}) = R_4(W_N(T_2^2/2 + T_3))_c. \quad (12)$$

In Ref. 32 we introduced an approximation which allows us to use factorized formulas in the expressions containing the T_4 operator. That crucial modification enables us to reduce the rank of the computational procedure from n^9 to n^7 . The equations defining the CCSDTQ_{f-1} approximation then take the following form:

$$T_1(\text{CCSDTQ}_f-1) = T_1(\text{CCSDT}), \quad (13)$$

$$T_2(\text{CCSDTQ}_f-1) = T_2(\text{CCSDT}) + \frac{1}{2} R_2 \{ T_2^{(1)\dagger} [W_N (T_2^2/2 + T_3)] \}, \quad (14)$$

$$T_3(\text{CCSDTQ}_f-1) = T_3(\text{CCSDT}). \quad (15)$$

$T_2^{(1)}$ means the first-order contribution to T_2 , $R_2 W_N$. Hence we do not even explicitly consider the T_4 equation; instead we have a new term in the T_2 equation depending on the T_2 and T_3 amplitudes; $T_2^{(1)}$ indicates the first order T_2 amplitudes.

In analogy to the CCSD(T) method, we introduced in Ref. 32 a noniterative version of the CCSDTQ-1 method, termed CCSDT(Q), where we solve the CCSDT equations and use the converged \bar{T}_2 and \bar{T}_3 amplitudes to evaluate the T_4 contribution according to the following formulas:

$$E_{Q_f}^5 = \langle 0 | \bar{T}_2^\dagger W_N T_4 | 0 \rangle, \quad (16)$$

$$T_4 = R_4(W_N(\bar{T}_2^2/2 + \bar{T}_3))_c. \quad (17)$$

Now the total energy within the CCSDT(Q) approximation is given as

$$E_{\text{CCSDT(Q)}} = E_{\text{CCSDT}} + E_{Q_f}^5. \quad (18)$$

Finally, we may employ a factorized formula for the quadruple contribution as given by Eqs. (16), (17) to obtain

$$E_{Q_f}^5 = \frac{1}{2} \langle 0 | \bar{T}_2^\dagger T_2^{(1)\dagger} [W_N(\bar{T}_2^2/2 + \bar{T}_3)]_c | 0 \rangle. \quad (19)$$

The total energy within the CCSDT(Q_f) approximation is given analogously as in Eq. (18),

$$E_{\text{CCSDT(Q}_f)} = E_{\text{CCSDT}} + E_{Q_f}^5. \quad (20)$$

Out of the four approximations introduced for the CCSDTQ model, the most efficient computationally is CCSDT(Q_f) defined by Eqs. (19) and (20). Here we may summarize the most important characteristics of the latter approach:

- (i) The method is size-extensive as it includes only connected diagrams;
- (ii) it is correct through fifth order in the MBPT energy;
- (iii) the T_4 part scales as n^7 ;
- (iv) the correlation corrections for the energy and properties studied are very close to those obtained with the standard CCSDTQ-1 method,⁴⁵ and within about 1 mHartree of the full CCSDTQ, provided the latter result is available. The CCSDTQ, however, requires an n^{10} computational procedure while CCSDTQ-1 is n^9 which places both beyond reach for larger basis set calculations.

III. DEVELOPMENT OF THE FACTORIZED CONTRIBUTION TO THE T_2 EQUATION

The crucial step in the development of the factorized quadruple approach is a construction of the diagrammatic terms corresponding to the expression,

$$D_{ij}^{ab} t_{ij}^{ab} = \frac{1}{2} \langle \Phi_{ij}^{ab} | T_2^{(1)\dagger} (W_N T_2^2/2)_c | 0 \rangle \quad (21a)$$

$$+ \frac{1}{2} \langle \Phi_{ij}^{ab} | T_2^{(1)\dagger} (W_N T_3)_c | 0 \rangle \quad (21b)$$

representing the last term in Eq. (14); D_{ij}^{ab} denotes a standard denominator, i.e., $D_{ij}^{ab} = e_i + e_j - e_a - e_b$. Now the $E_{Q_f}^5$ expression, Eq. (20), takes the form

$$E_{Q_f}^5 = \sum_{abij} (2t_{ij}^{ab} - t_{ji}^{ab}) D_{ij}^{ab} t_{ij}^{ab}. \quad (22)$$

Note that t_{ij}^{ab} refers to the T_2 amplitude obtained from the factorized T_4 contribution, while t_{ij}^{ab} represents the T_2 amplitude obtained from the solution of the CC equations at the CCSDT level. In order to simplify the notation we will sup-

TABLE I. Definitions of symbols.

Symbol	Definition
$\langle rs tu \rangle$	$\langle rs tu \rangle - \langle rs ut \rangle$
v_{tu}^{rs}	$\langle rs tu \rangle$
w_{ab}^{ij}	$\frac{1}{2} \langle ij ab \rangle / (e_i + e_j - e_a - e_b)$
\bar{w}_{ab}^{ij}	$w_{ab}^{ij} - w_{ba}^{ij}$
\bar{v}_{ut}^{rs}	$2v_{tu}^{rs} - v_{ut}^{rs}$
\bar{w}_{tu}^{rs}	$2w_{tu}^{rs} - w_{ut}^{rs}$
t_{ij}^{ab}	$2t_{ij}^{ab} - t_{ji}^{ab}$
t_{ij}^{abc}	$2t_{ijk}^{abc} - t_{ikj}^{abc}$
\bar{t}_{ijk}^{abc}	$2t_{ijk}^{abc} - t_{ikj}^{abc} - t_{jik}^{abc}$

press the overbar on T_2 and T_3 . Out of the total number of 42 antisymmetrized diagrammatic terms 30 are obtained from the $\frac{1}{2} \langle \Phi_{ij}^{ab} | T_2^{(1)\dagger} (W_N T_2^2/2)_c | 0 \rangle$ component, Eq. (21a), and 12 from the $\frac{1}{2} \langle \Phi_{ij}^{ab} | T_2^{(1)\dagger} (W_N T_3)_c | 0 \rangle$ one, Eq. (21b). In Table II we divide all the algebraic terms corresponding to Eq. (21a) into three groups based on the form of the W_N operator: 7 diagrams for the four-particle component of W_N , i.e., $\langle ab || cd \rangle$, indicated in the first column of Table II by the symbol *pppp*; the next group of 7 obtained from the four-hole component of W_N , *hhhh*, and the last 16 terms corresponding to the two-particle-two-hole component of W_N , *hphp*. All the symbols used in the last column of Table II represent the algebraic quantities defined in Table I; t_{ij}^{ab} being the usual antisymmetrized amplitude. The third column of Table II gives the number of Goldstone terms corresponding to a particular antisymmetrized contribution.

Similarly, in Table III we list the algebraic expressions corresponding to the term $\frac{1}{2} \langle \Phi_{ij}^{ab} | T_2^{(1)\dagger} (W_N T_3)_c | 0 \rangle$, Eq. (21b). Here we have two groups corresponding to the *ppph* term (three-particle-one-hole indices) and to the *phhh* one (one-particle-three-hole indices). Each group comprises 6 antisymmetrized diagrams corresponding to 38 Goldstone ones. The total number of Goldstone terms corresponding to both terms in Eq. (21) is 300. The number of Goldstone terms is decisive here since the actual expressions which have been coded are derived on that basis and only apply to closed shell systems. The general strategy in the derivation of the working formulas for the term in Eq. (21) is as follows: (i) derive all the antisymmetrized diagrams corresponding to the term in question; (ii) for each diagram construct all Goldstone terms; (iii) apply ‘‘vertical’’ factorization. To illustrate the last step let us first consider the $T_2^\dagger (W_N T_2^2)_c/4$ term (appropriate projections assumed), Eq. (21a). Here we have four vertices and owing to the fact that we do not have any denominator present in the diagram as a whole (denominators are included into T_2 vertices) we may evaluate this diagram by combining two vertices at a time, i.e., adopt a three step procedure. In addition it can be done in any desired sequence of the operators. Let us assume that the contribution in question can be represented symbolically as a product of four operators, $\hat{T} = \hat{A} \hat{B} \hat{C} \hat{D}$. By combining any two of them, e.g., \hat{A} and \hat{B} we have $\hat{X} = \hat{A} \hat{B}$, and $\hat{T} = \hat{X} \hat{C} \hat{D}$. We call the resulting vertex (i.e., operator) \hat{X} an intermediate of the first rank. Next we combine the latter intermediate, \hat{X} , with one of the remaining vertices, e.g., \hat{C} ,

TABLE II. Algebraic expressions corresponding to antisymmetrized diagrams originating from the $\frac{1}{2}\langle\Phi_{ij}^{ab}||[T_2^{(1)\dagger}(W_N T_2^2/2)_c]|0\rangle$ term.

Form of W_N	Term No.	No. of Goldstone terms	Expression ^a
pppp	1	4	$+\frac{1}{2}P(ab)\tilde{w}_{cd}^{kl}t_{ij}^{ce}t_{ij}^{fb}\langle da ef\rangle$
	2	2	$-\frac{1}{2}\tilde{w}_{cd}^{kl}t_{ij}^{ce}t_{ij}^{fd}\langle ba ef\rangle$
	3	16	$+P(ij)(ab)\tilde{w}_{cd}^{kl}t_{ki}^{ce}t_{lj}^{fb}\langle ad ef\rangle$
	4	6	$+\tilde{w}_{cd}^{kl}t_{ij}^{ce}t_{ij}^{fd}\langle ab ef\rangle$
	5	2	$-\frac{1}{2}P(ab)\tilde{w}_{cd}^{kl}t_{ij}^{ce}t_{ij}^{fb}\langle cd ef\rangle$
	6	4	$-\frac{1}{2}P(ab)\tilde{w}_{cd}^{kl}t_{ij}^{ce}t_{ij}^{fd}\langle cb ef\rangle$
	7	6	$+\frac{1}{2}P(ab)\tilde{w}_{cd}^{kl}t_{ik}^{ce}t_{lj}^{fb}\langle cd ef\rangle$
hhhh	8	4	$+\frac{1}{2}P(ij)\tilde{w}_{cd}^{kl}t_{km}^{cd}t_{nj}^{ab}\langle mn li\rangle$
	9	2	$-\frac{1}{2}\tilde{w}_{cd}^{kl}t_{km}^{cd}t_{nl}^{ba}\langle mn ij\rangle$
	10	16	$+P(ij)(ab)\tilde{w}_{cd}^{kl}t_{km}^{ca}t_{nj}^{db}\langle mn il\rangle$
	11	6	$+\tilde{w}_{cd}^{kl}t_{ca}^{bd}t_{nl}^{mn}\langle mn ij\rangle$
	12	2	$-\frac{1}{2}P(ij)\tilde{w}_{cd}^{kl}t_{mi}^{cd}t_{nj}^{ab}\langle mn kl\rangle$
	13	4	$-\frac{1}{2}P(ij)\tilde{w}_{cd}^{kl}t_{mi}^{cd}t_{nl}^{ba}\langle mn kj\rangle$
	14	6	$\frac{1}{2}P(ij)\tilde{w}_{cd}^{kl}t_{im}^{ca}t_{nj}^{db}\langle mn kl\rangle$
hphp	15	4	$-\frac{1}{2}P(ab)\tilde{w}_{cd}^{kl}t_{cd}^{eb}t_{ij}^{ma}\langle ma le\rangle$
	16	8	$+\frac{1}{2}P(ij)(ab)\tilde{w}_{cd}^{kl}t_{cd}^{eb}t_{ij}^{ma}\langle ma ie\rangle$
	17	8	$+\frac{1}{2}P(ij)(ab)\tilde{w}_{cd}^{kl}t_{cd}^{eb}t_{ij}^{ma}\langle ma ke\rangle$
	18	4	$-\frac{1}{4}P(ij)(ab)\tilde{w}_{cd}^{kl}t_{cd}^{ea}t_{ij}^{mb}\langle mb je\rangle$
	19	8	$-P(ab)\tilde{w}_{cd}^{kl}t_{ca}^{eb}t_{ij}^{md}\langle md el\rangle$
	20	16	$-P(ij)(ab)\tilde{w}_{cd}^{kl}t_{ca}^{ed}t_{ij}^{mb}\langle mb ie\rangle$
	21	8	$+P(ab)\tilde{w}_{cd}^{kl}t_{ca}^{ed}t_{ij}^{mb}\langle mb le\rangle$
	22	16	$-P(ij)(ab)\tilde{w}_{cd}^{kl}t_{ca}^{ed}t_{ij}^{mb}\langle mb ie\rangle$
	23	16	$-P(ij)(ab)\tilde{w}_{cd}^{kl}t_{ca}^{ed}t_{ij}^{mb}\langle mb ke\rangle$
	24	8	$+\frac{1}{2}P(ij)(ab)\tilde{w}_{cd}^{kl}t_{ca}^{be}t_{ij}^{md}\langle md je\rangle$
	25	16	$-P(ij)(ab)\tilde{w}_{cd}^{kl}t_{ca}^{ed}t_{ij}^{mb}\langle mb ke\rangle$
	26	8	$+\frac{1}{2}P(ij)(ab)\tilde{w}_{cd}^{kl}t_{ca}^{ed}t_{ij}^{mb}\langle mb je\rangle$
27	4	$-\tilde{w}_{cd}^{kl}t_{ca}^{ed}t_{ij}^{mb}\langle mc le\rangle$	
28	8	$+P(ij)\tilde{w}_{cd}^{kl}t_{ca}^{ab}t_{ij}^{ed}\langle mc je\rangle$	
29	8	$-P(ij)\tilde{w}_{cd}^{kl}t_{ca}^{ab}t_{ij}^{ed}\langle mc ek\rangle$	
30	4	$-\frac{1}{2}P(ij)\tilde{w}_{cd}^{kl}t_{ca}^{ab}t_{ij}^{ed}\langle mc je\rangle$	

^aSummation over repeated indices assumed; i, j, \dots , run over occupied one-particle states; a, b, \dots , run over virtual one-particle states; $P(ij)$ or $P(ab)$ implies sum of two components differing by permutation of i, j or a, b indices, respectively; $P(ij)(ab)$ implies sum of four terms differing by permutation of i, j and a, b indices; $t_{i_1 i_2}^{a_1 a_2}$ denotes antisymmetrized double excitation amplitude.

to obtain an intermediate of the second rank, $\hat{Y}: \hat{Y} = \hat{X}\hat{C}$. The last step combines the intermediate of the second rank, \hat{Y} , and the operator \hat{D} , to generate the contribution to the T_2 equation, $\hat{T} = \hat{Y}\hat{D}$.

In the case of the $T_2^\dagger(W_N T_3)_c/2$ term we have only three vertices, hence after a creation of the intermediate of the first rank we obtain, directly, the contribution to the T_2 equation. The crucial point in this step-by-step construction of the final T_2 contribution is a choice of the appropriate sequence of vertices. A general hint is that the vertices to be combined at each step should have the maximum possible number of contraction lines, or, in other words, the maximum number of common indices. This would ensure adopting the lowest rank computational procedure in the evaluation of the desired contribution to the considered CC equation.

The working formulas for the factorized contribution to the T_2 equation, Eq. (21), are presented in Tables IV, V, and VI. Table IV contains the expressions corresponding to the

TABLE III. Algebraic expressions corresponding to antisymmetrized diagrams originating from the $\frac{1}{2}\langle\Phi_{ij}^{ab}||[T_2^{(1)\dagger}(W_N T_3)]|0\rangle$ term.

Form of W_N w operator	Term No.	No. of Goldstone terms	Expression ^a
ppph	31	4	$+\frac{1}{4}P(ij)\tilde{w}_{cd}^{kl}t_{cd}^{cde}\langle ab ej\rangle$
	32	12	$+\frac{1}{4}P(ij)(ab)\tilde{w}_{cd}^{kl}t_{cd}^{cea}\langle db ej\rangle$
	33	3	$-\frac{1}{2}P(ij)\tilde{w}_{cd}^{kl}t_{cd}^{cde}\langle ab le\rangle$
	34	12	$+P(ab)\tilde{w}_{cd}^{kl}t_{cd}^{cde}\langle ad el\rangle$
	35	4	$-\frac{1}{2}P(ij)\tilde{w}_{cd}^{kl}t_{cd}^{cde}\langle cd ei\rangle$
	36	3	$+\frac{1}{2}\tilde{w}_{cd}^{kl}t_{cd}^{cab}\langle cd el\rangle$
hhhp	37	4	$-\frac{1}{4}P(ab)\tilde{w}_{cd}^{kl}t_{cd}^{cda}\langle mb ij\rangle$
	38	12	$-\frac{1}{2}P(ij)(ab)\tilde{w}_{cd}^{kl}t_{cd}^{cda}\langle mb lj\rangle$
	39	3	$\frac{1}{2}P(ab)\tilde{w}_{cd}^{kl}t_{cd}^{cab}\langle ma kl\rangle$
	40	12	$-P(ij)\tilde{w}_{cd}^{kl}t_{cd}^{cab}\langle md il\rangle$
	41	4	$+\frac{1}{2}P(ab)\tilde{w}_{cd}^{kl}t_{cd}^{cab}\langle dm ij\rangle$
	42	3	$-\frac{1}{2}\tilde{w}_{cd}^{kl}t_{cd}^{cab}\langle md kl\rangle$

^aSee footnote to Table II.

first rank intermediates; the upper half (separated by the horizontal line) contains those corresponding to Eq. (21a), the lower one, to those originating from the term in Eq. (21b). Since the latter term involves only three operators the first rank intermediates enter directly the amplitude equation, i.e., they appear in Table VI. The first rank intermediates connected with Eq. (21a) are next exploited in the construction of the intermediates of the second rank listed in Table V and the latter are then used to obtain the contribution to t_{ij}^{ab} as shown in Table VI. Note that all quantities collected in Tables IV, V, and VI are symmetric with respect to the simultaneous permutation of upper and lower indices, i.e., $x_{i_1 i_2 \dots}^{a_1 a_2 \dots} = x_{i_2 i_1 \dots}^{a_2 a_1 \dots} = \text{etc}$. The indices are summed over doubly occupied spatial orbitals. Contrary to the quantities in Tables

TABLE IV. First rank intermediates.

Intermediate	Expression ^a	Scaling
ξ_b^a	$-t_{mn}^{ea} \tilde{w}_{eb}^{mn}$	$n_0^2 n_v^3$
ξ_j^i	$t_{mj}^{ef} \tilde{w}_{ef}^{mi}$	$n_0^3 n_v^2$
ξ_{kl}^{ij}	$t_{kl}^{ef} \tilde{w}_{ef}^{ij}$	$n_0^2 n_v^4$
ξ_{cd}^{ab}	$t_{mn}^{ab} \tilde{w}_{cd}^{mn}$	$n_0^2 n_v^4$
ξ_{ib}^{aj}	$t_{im}^{ae} \tilde{w}_{eb}^{mj} - t_{mi}^{ae} \tilde{w}_{eb}^{mj}$	$n_0^3 n_v^3$
ξ_{bj}^{ai}	$t_{mj}^{ae} \tilde{w}_{be}^{mi}$	$n_0^3 n_v^3$
ξ_{ab}^{ij}	$\tilde{w}_{ae}^{im} t_{mb}^{ej} - \tilde{w}_{ae}^{im} t_{mb}^{ej}$	$n_0^3 n_v^3$
ξ_{ab}^{ij}	$w_{ae}^{mj} t_{mb}^{ie}$	$n_0^3 n_v^3$
ξ_{ab}^{ij}	$w_{ej}^{ij} t_{ab}^{mb} + w_{ab}^{mn} t_{ij}^{mn}$	$n_0^2 n_v^4$
ξ_a^i	$\tilde{t}_{imn}^{ef} \tilde{w}_{ef}^{mi}$	$n_0^3 n_v^3$
ξ_a^i	$u_{am}^{ef} \tilde{w}_{ef}^{im} - v_{mn}^{ie} \tilde{w}_{mn}^{ae}$	$n_0^3 n_v^2$
ξ_{ik}^{aj}	$\tilde{t}_{imk}^{ef} \tilde{w}_{ef}^{mj}$	$n_0^4 n_v^3$
ξ_{ic}^{ab}	$\tilde{t}_{imn}^{ae} \tilde{w}_{eb}^{mn}$	$n_0^3 n_v^4$
ξ_{ak}^{ij}	$u_{ka}^{ef} w_{ef}^{ij} + v_{mk}^{ie} w_{ae}^{im}$	$n_0^3 n_v^3, n_0^4 n_v^2$
ξ_{ac}^{ib}	$u_{mn}^{ib} w_{ac}^{mn} + v_{mc}^{ie} w_{ae}^{im}$	$n_0^3 n_v^3, n_0^2 n_v^4$

^aSymbols used are defined in Table I; summation over repeated indices assumed; m, n run over doubly occupied hole states; e, f run over doubly occupied virtual states; $t_{i_1 i_2 \dots}^{a_1 a_2 \dots}$ represents the (nonorthogonally) spin adapted CC amplitude.

^bEach of the two components has a different scaling as indicated in the last column.

TABLE V. Second rank intermediates.^a

Intermediate	Expression ^b	Scaling
χ_b^a	$-\xi_{ij}^e \bar{v}_{eb}^{fa}$ $-\xi_{mb}^e \bar{v}_{mb}^{na}$ $-\xi_{cd}^e \bar{v}_{eb}^{cd}$ $-\xi_{me}^a \bar{v}_{bn}^{me} - \xi_{em}^a \bar{v}_{mb}^{ne}$	n_v^4 $n_0^2 n_v^2$ n_v^5 $n_0^2 n_v^3$
χ_j^i	$-\xi_{ni}^e \bar{v}_{mj}^{ni}$ $-\xi_{ij}^e \bar{v}_{ej}^{fi}$ $+\xi_{jk}^m \bar{v}_{nm}^{ki}$ $+\xi_{jf}^e \bar{v}_{em}^{if} + \xi_{jj}^e \bar{v}_{em}^{fi}$	n_v^4 $n_0^2 n_v^2$ n_0^5 $n_0^3 n_v^2$
χ_{cd}^{ab}	$+\xi_{ed}^e \bar{v}_{ed}^{ab}$ $-\xi_{ef}^e \bar{v}_{ed}^{fb} + \xi_{fe}^e \bar{v}_{ed}^{fb} + \xi_{fd}^e \bar{v}_{ce}^{fb}$ $+\xi_{na}^e \bar{v}_{mb}^{na} + \xi_{dm}^e \bar{v}_{cn}^{mb} - \xi_{cm}^e \bar{v}_{nd}^{mb}$	n_v^4 n_v^6 $n_0^2 n_v^4$
χ_{kl}^{ij}	$-\xi_{m}^i \bar{v}_{kl}^{mj}$ $-\xi_{km}^i \bar{v}_{nl}^{mj} + \xi_{ni}^e \bar{v}_{nl}^{mj} + \xi_{km}^e \bar{v}_{nl}^{im}$ $+\xi_{fk}^e \bar{v}_{jl}^{ij} + \xi_{ij}^e \bar{v}_{ek}^{ij} - \xi_{kf}^e \bar{v}_{el}^{ij} + \xi_{ef}^i \bar{v}_{kl}^{ij}$	n_0^5 n_0^6 $n_0^2 n_v^2$
χ_{ib}^{aj}	$-\xi_{ib}^e \bar{v}_{ib}^{am}$ $+\xi_{ie}^e \bar{v}_{ie}^{aj}$ $+\xi_{in}^e \bar{v}_{mb}^{an} + \xi_{nb}^e \bar{v}_{im}^{nj}$ $+\xi_{eb}^e \bar{v}_{ej}^{aj} + \xi_{ij}^e \bar{v}_{eb}^{aj}$ $-\xi_{ib}^e \bar{v}_{em}^{aj} - \xi_{bi}^e \bar{v}_{me}^{aj} - \xi_{em}^e \bar{v}_{ib}^{em} - \xi_{me}^e \bar{v}_{bi}^{em}$ $-\xi_{eb}^e \bar{v}_{mi}^{ea} - \xi_{be}^e \bar{v}_{mi}^{ea}$ $+\frac{1}{2} \xi_{ab}^e \bar{v}_{im}^{ae} + \frac{1}{2} \xi_{be}^e \bar{v}_{im}^{ae}$	$n_0^3 n_v^2$ $n_0^2 n_v^3$ $n_0^2 n_v^2$ $n_0^2 n_v^4$ $n_0^3 n_v^3$ $n_0^3 n_v^3$ $n_0^3 n_v^3$
χ_{bj}^{ai}	$-\xi_{ib}^e \bar{v}_{bj}^{am}$ $+\xi_{ie}^e \bar{v}_{ie}^{ai}$ $-\xi_{in}^e \bar{v}_{bm}^{an} + \xi_{nb}^e \bar{v}_{im}^{nj} - \xi_{bn}^e \bar{v}_{mj}^{ni} + \xi_{nb}^e \bar{v}_{mj}^{ni}$ $-\xi_{eb}^e \bar{v}_{ej}^{ai} + \xi_{ij}^e \bar{v}_{eb}^{ai} - \xi_{ij}^e \bar{v}_{eb}^{fa} + \xi_{ji}^e \bar{v}_{eb}^{fa}$ $-\xi_{bj}^e \bar{v}_{em}^{ai} - \xi_{bj}^e \bar{v}_{me}^{ai} - \xi_{em}^e \bar{v}_{bj}^{em} - \xi_{me}^e \bar{v}_{bj}^{em}$ $-\xi_{eb}^e \bar{v}_{jm}^{ea} - \xi_{be}^e \bar{v}_{jm}^{ea}$ $-\frac{1}{2} \xi_{be}^e \bar{v}_{mj}^{ae}$	$n_0^3 n_v^2$ $n_0^2 n_v^3$ $n_0^2 n_v^2$ $n_0^2 n_v^4$ $n_0^3 n_v^3$ $n_0^3 n_v^3$ $n_0^3 n_v^3$
θ_{ij}^{ab}	$\xi_{ie}^e \bar{v}_{mj}^{ab} - \xi_{ei}^e \bar{v}_{mj}^{ab}$	$n_0^3 n_v^3$
$\theta_{ij}'^{ab}$	$\xi_{ej}^e \bar{v}_{im}^{ab}$	$n_0^3 n_v^3$
$\theta_{ij}''^{ab}$	$\xi_{ij}^e \bar{v}_{mn}^{ab}$	$n_0^4 n_v^2$

^aEach intermediate is constructed from the first rank intermediate, defined in Table IV, and two-electron integral or amplitude.

^bSee footnote a to Table IV.

II and III they are not antisymmetric, i.e., the permutation of a pair of upper (or lower) indices creates a new amplitude.

As we may notice from the last column of Table V the bottleneck of this scheme is a construction of the χ_{cd}^{ab} intermediate which requires an n_v^6 procedure. In fact there is another possible factorization scheme replacing the n_v^6 by $n_0^2 n_v^5$. We coded both ways, however, since in the current code only the n_v^6 route exploits spatial symmetry, the latter was used in the actual calculations.

IV. RESULTS AND DISCUSSION

The equilibrium properties of the C_2 molecule have been investigated in several studies. The coupled cluster methods have been employed in the study by Watts and Bartlett⁴⁶ where a number of CCSDT- n variants including full CCSDT were applied at the DZP level. For larger basis sets the CCSD(T) approximation was used. A large basis set study has been reported by Dunning and co-workers⁴⁷ where both coupled cluster methods [at the CCSD and CCSD(T) level]

TABLE VI. Contributions to T_2 equation.^a

	Expression ^b	Scaling
$D_{ij}^{ab} t_{ij}^{ab}$	$\chi_{ij}^e \bar{v}_{ij}^{be} + \xi_{ij}^e \bar{v}_{ij}^{ab}$ $-\chi_{ij}^e \bar{v}_{mj}^{ab} - \xi_{mj}^e \bar{v}_{ij}^{mb}$ $\chi_{ij}^e \bar{v}_{ij}^{ef} + \frac{1}{2} \bar{v}_{ij}^{ab} (\theta_{ij}^{ef} + \theta_{ij}'^{ef})$ $+\xi_{ij}^e \bar{v}_{ej}^{fb} - \xi_{ij}^e \bar{v}_{ej}^{fa} - \xi_{ij}^e \bar{v}_{ej}^{fb}$ $\chi_{ij}^e \bar{v}_{mn}^{ab} + \frac{1}{2} \bar{v}_{ij}^{mn} (\theta_{mn}^{ab} + \theta_{mn}'^{ab})$ $-\xi_{in}^e \bar{v}_{mj}^{nb} + \xi_{ni}^e \bar{v}_{mj}^{nb} + \xi_{nj}^e \bar{v}_{im}^{nb}$ $\chi_{ij}^e \bar{v}_{mj}^{eb} + \chi_{ei}^e \bar{v}_{mj}^{eb} + \chi_{ej}^e \bar{v}_{im}^{eb}$ $+\bar{v}_{ej}^{am} \theta_{ij}^{eb} + \bar{v}_{ie}^{am} \theta_{jm}^{eb} + \bar{v}_{ej}^{am} \theta_{mi}^{eb} + \bar{v}_{ie}^{am} \theta_{mj}^{eb}$ $+\bar{v}_{ie}^{am} \theta_{mj}^{eb} + \bar{v}_{ie}^{am} \theta_{jm}^{eb} + \bar{v}_{ej}^{am} \theta_{im}^{eb}$ $-\xi_{ij}^e \bar{v}_{em}^{ab} + \xi_{me}^e \bar{v}_{ij}^{ab} - \bar{v}_{ij}^{ab} \xi_{em}^e$ $-\bar{v}_{ab}^{em} \xi_{mn}$ $-\bar{v}_{imn} \xi_{ej}$ $-\bar{v}_{aj}^{em} \xi_{fb}$	$n_0^2 n_v^3$ $n_0^3 n_v^2$ $n_0^2 n_v^4$ $n_0^2 n_v^4$ $n_0^4 n_v^2$ $n_0^4 n_v^2$ $n_0^3 n_v^3$ $n_0^3 n_v^3$ $n_0^3 n_v^3$ $n_0^3 n_v^3$ $n_0^3 n_v^3$ $n_0^3 n_v^3$ $n_0^3 n_v^3$ $n_0^4 n_v^3$ $n_0^4 n_v^3$ $n_0^4 n_v^3$

^aEach contribution is obtained by combining the two-electron integral or CC amplitude and second rank intermediate [in connection with Eq. (21a)] or first rank intermediate [in connection with Eq. (21b)].

^bSee footnote a to Table IV.

have been employed as well as several multireference CI approaches. In other papers^{48–51} CI methods in either single- and multireference formulations were used.

In the current calculations, we report results from the three standard coupled cluster approaches which include the T_3 operator: CCSD(T), CCSDT-3, and full CCSDT. On top of the converged T_2 and T_3 amplitudes, the lowest order connected T_4 contribution has been evaluated according to the CCSDX(Q_f) scheme, where X refers to the T operator at various levels of approximation. The calculations have been performed for the series of correlation consistent basis sets of Dunning,⁵² beginning with the small pVDZ containing 28 functions, and ending at the pV5Z set with 182 functions. For the basis sets designated by a pVXZ acronym (X=D, T, Q, etc.) only the valence electrons were correlated; the pCVXZ sets indicate that all electron correlated calculations were performed. In both cases all virtual orbitals were included in the correlation. The calculated values of the equilibrium bond length and harmonic frequency have been collected in Table VII and Table VIII, respectively. In addition Table IX contains the bond lengths and vibrational frequencies obtained for small basis sets with other methods treating the T_4 operator in more rigorous ways.

V. EQUILIBRIUM GEOMETRY

In Table VII we list the computed bond distances for the C_2 molecule. For each method applied we observe that with an improved quality of basis, the computed bond lengths get smaller. For the smallest basis sets, i.e., pVDZ and aug-pVDZ, R_e remains in the range of 1.270–1.273 Å, for the largest basis set used in the present work, pV5Z, it is shorter by ≈ 0.03 Å assuming values 1.241–1.246, depending on the method. Comparing the computed bond lengths for various levels of the T_3 operator, we observe that by upgrading CCSD(T) to the iterative CCSDT-3, the bond length increases by 0.001 Å, whereas going from CCSDT-3 to the full CCSDT has the opposite effect of nearly the same magnitude. As a result, the CCSDT bond length for the C_2 mol-

TABLE VII. Geometry of the C₂ molecule with coupled cluster methods and correlation consistent basis sets. Å. Experiment, $R_{\text{exp}}=1.242$.^a

Basis set ^b	No. of b.f.	CC								
		SD(T)	ΔQ^c	SD(TQ _f)	SDT-3	ΔQ^c	SDT-3(Q _f)	SDT	ΔQ^c	SDT(Q _f)
pVDZ	28	1.2705	-11	1.2694	1.2714	3	1.2717	1.2707	12	1.2719
pVDZ+	46	1.2720	-10	1.2710	1.2729	3	1.2732	1.2721	13	1.2734
pCVDZ	36	1.2680	-11	1.2669	1.2690	3	1.2693	1.2683	11	1.2694
pCVDZ+	54	1.2694	-11	1.2683	1.2703	3	1.2706	1.2695	12	1.2707
pVTZ	60	1.2507	-10	1.2497	1.2518	4	1.2522	1.2506	14	1.2520
pVTZ+	92	1.2508	-10	1.2498	1.2518	5	1.2523	1.2507	13	1.2520
pCVTZ	86	1.2465	-11	1.2454	1.2476	4	1.2480	1.2463	14	1.2477
pCVTZ+	118	1.2471	-11	1.2460	1.2482	4	1.2486	1.2468	15	1.2483
pVQZ	110	1.2458	-10	1.2448	1.2469	5	1.2474	1.2455	14	1.2469
pVQZ+	160	1.2460	-10	1.2450	1.2472	4	1.2476	1.2457	15	1.2472
pCVQZ	168	1.2425	-11	1.2414	1.2437	4	1.2441	1.2422	14	1.2436
pV5Z	182	1.2447	-10	1.2437	1.2459	5	1.2464	1.2444	14	1.2458
pV6Z		1.2444		1.2434 ^d	1.2456 ^d		1.2461 ^d	1.2441 ^d		1.2455 ^d
pCV6Z		1.2411 ^e		1.2401 ^f	1.2423 ^f		1.2428 ^f	1.2408 ^f		1.2422 ^f

^aReference 56.^bCorrelation consistent basis sets from Ref. 51, “+” denotes augmented set; pVXZ basis sets indicate frozen core calculations, pCVXZ, all electron calculations.^cThe difference between values in adjacent columns, i.e., the net T_4 effect (10^{-4} Å).^dExtrapolated, based on additivity of the computed CCSD(T) value.^eReference 47.^fExtrapolated, based on additivity of the CCSD(T) value from Ref. 47.

ecule is very close to that obtained from the CCSD(T) method, the difference being 0.0001–0.0003 Å. However, when the T_4 operator is added, its effect strongly depends on the quality of the T_3 approximation. An inclusion of the T_4 operator on top of the CCSD(T) method decreases the bond length by 0.001 Å, see the third column in Table VII, while for the full CCSDT, T_4 increases the bond length by ≈ 0.0015 Å (last column of Table VII). As a result the CCSDT(Q_f) bond lengths are larger than those given by CCSD(TQ_f) by 0.0022–0.0025 Å. For the pV5Z basis, the crudest approach, CCSD(T), gives a value of 1.2447 Å while

the most accurate one for the same basis gives 1.2458 Å. For the former method we also ran calculations for the pV6Z basis and obtained R_e equal to 1.2444 Å, which is lower by 0.0003 Å compared to the corresponding value for the pV5Z basis set, see the penultimate row in Table VII. Based on that value we evaluated, by assuming additive approximations based on CCSD(T) the expected R_e values for the remaining methods, i.e., they were reduced by 0.0003 Å, e.g. For CCSDT(Q_f) the extrapolated bond length is 1.2455 Å, see the same row in Table VII, which could be considered as

TABLE VIII. Harmonic frequency of the C₂ molecule with coupled cluster methods and correlation consistent basis sets (cm⁻¹). Experiment, $\omega=1855.0$.^a

Basis set ^b	No. of b.f.	CC								
		SD(T)	ΔQ^c	SD(TQ _f)	SDT-3	ΔQ^c	SDT3(Q _f)	SDT	ΔQ^c	SDT(Q _f)
pVDZ	28	1828	6	1834	1830	-2	1828	1829	-8	1821
pVDZ+	46	1815	6	1821	1817	-2	1815	1816	-9	1807
pCVDZ	36	1827	7	1834	1829	-2	1827	1828	-8	1820
pCVDZ+	54	1816	6	1822	1819	-2	1817	1818	-8	1810
pVTZ	60	1845	7	1852	1846	-2	1844	1847	-9	1838
pVTZ+	92	1841	6	1847	1841	-2	1839	1844	-10	1834
pCVTZ	86	1856	7	1863	1855	-1	1853	1858	-9	1849
pCVTZ+	118	1853	7	1860	1851	-2	1849	1854	-10	1844
pVQZ	110	1856	7	1863	1856	-2	1854	1859	-10	1849
pVQZ+	160	1855	5	1860	1854	-2	1852	1857	-9	1848
pCVQZ	168	1867	7	1874	1865	-2	1863	1869	-10	1859
pV5Z	182	1859	7	1866	1858	-2	1856	1861	-9	1852
pV6Z		1860		1867 ^d	1859 ^d		1857 ^d	1862 ^d		1853 ^d
pCV6Z		1871 ^e		1878 ^f	1870 ^f		1868 ^f	1873 ^f		1864 ^f

^aReference 56.^bSee footnote b to Table VII.^cThe difference between values in adjacent columns, i.e., the net T_4 effect.^dExtrapolated, assuming the same 1 cm⁻¹ change as occurs in CCSD(T).^eReference 47.^fExtrapolated, assuming the same change as occurs for CCSD(T) in Ref. 47.

TABLE IX. Performance of the coupled cluster methods including connected quadruple excitations at various level of approximation in the calculations of the equilibrium geometry and harmonic frequency for the C₂ molecule. Bond length in Å, frequencies in cm⁻¹.

Basis set ^a	CC					
	SDT	SDT(Q _f)	SDT(Q)	SDTQ _f -1	SDTQ-1	SDTQ
	<i>R_e</i>					
pVDZ	1.2707	1.2719	1.2717	1.2721	1.2720	1.2723
pVDZ+	1.2721	1.2734	1.2732	1.2736	1.2735	1.2738
pCVDZ	1.2683	1.2694	1.2693	1.2696	1.2695	1.2698
pVTZ	1.2506	1.2520	1.2519	1.2522	1.2521	1.2523
	<i>ω</i>					
pVDZ	1829	1821	1821	1820	1820	1816
pVDZ+	1816	1807	1807	1807	1807	1802
pCVDZ	1828	1820	1820	1820	1820	1815
pVTZ	1847	1838	1838	1838	1837	1833

^aSee footnote b to Table VII.

close to the basis set limit for CCSDT(Q_f) and for the valence correlation.

In order to relate the computed value to experiment we need to include the correlation effects also for the core electrons; otherwise the error of the computed value would reach ≈0.003 Å. As has been found on many occasions, inclusion of core electrons significantly reduces the bond length in diatomics, e.g., for the N₂ molecule the bond contraction amounts to ≈0.002 Å.^{33,47,53–55} For the C₂ molecule the core correlation lowers the bond length by ≈0.003–0.004 Å depending on the method and basis set; compare the values in rows corresponding to pVXZ and pCVXZ basis sets. Our results are in agreement with other findings.⁴⁷ The largest basis used here for which the calculations were made with and without core electron correlation is pVQZ; the core effect in that case is equal to 0.0033 Å for the majority of the methods [CCSD(T), CCSDT-3(Q_f), CCSDT, and CCSDT(Q_f)]. For the remaining two methods it is equal to 0.0034 Å [CCSD(TQ_f)] and to 0.0032 Å (CCSDT-3); the differences, however, come mostly from the roundup errors. Applying the same bond length shift to the extrapolated pV6Z values we were able to evaluate the equilibrium bond length close to the basis set limit with highly sophisticated correlation calculations, see the last row of Table VII. We observe that for the CCSD(T) method we have *R_e* equal to 1.2411 Å,⁴⁷ i.e., 0.0013 Å shorter than experiment, *R_{exp}* = 1.2424.⁵⁶ Adding to this value the *T₄* correction would move *R_e* in the wrong direction, i.e., the error would increase to 0.0025 Å. On the other hand, the CCSDT-3 approach would give a value much closer to experiment, the extrapolated *R_e* is 1.2423 and the *T₄* correction built on the CCSDT-3 amplitudes is small; it increases the bond length by 0.0004 Å so the final value of 1.2427 is still close to the experimental one. We believe that the most reliable results obtained here are due to the CCSDT method corrected with the *T₄* contribution. The CCSDT bond length of 1.2408 Å is shorter by 0.0016 Å than experiment, however, the *T₄* works here in the proper direction increasing the value by 0.0014 Å to 1.2422 Å. We note that all the methods considered here give very accurate estimates of the equilibrium bond length

for the C₂ molecule. However, if we insist on accuracy below 0.001 Å we need a reliable estimate of the connected quadruple excitation effect, which amounts to 0.0014 Å.

VI. THE HARMONIC FREQUENCY

The harmonic frequencies for the C₂ molecule are listed in Table VIII. The first observation which should be made is an apparent insensitivity of the frequency values with respect to the form of the *T₃* equation considered. For example, the differences between the CCSD(T) and CCSDT-3 models as well as those between CCSD(T) and CCSDT oscillate between 0 and 2 cm⁻¹, in one or two cases only assuming 3 cm⁻¹. In addition the CCSDT and CCSDT-3 values change in opposite directions with respect to the CCSD(T) frequencies. These irregularities mirror those observed in the case of the equilibrium bond length calculations. However, the important observation is that the *T₄* corrections are practically independent of the basis sets. They do depend strongly, though, on the quality of the *T₃* operator. Thus, *T₄* corrections built from CCSD(T) are equal in the majority of cases to +7 cm⁻¹, where the plus sign means that the CCSD(TQ_f) frequency is larger than that of CCSD(T), whereas the same correction built on the full CCSDT method is, for larger basis sets, equal to -9 or -10 cm⁻¹, see respective columns in Table VIII with Δ*Q* headings. This behavior of the harmonic frequency is consistent with the *T₄* corrections to the equilibrium bond length discussed in the previous section. A similar stability is observed with respect to the core electron effect on condition that the basis sets of triple zeta quality or better are considered. For those cases an inclusion of core electrons into calculations increases the harmonic frequency by 10–11 cm⁻¹. The stability of the latter two effects makes it possible to extrapolate the results of more demanding methods to the larger basis sets, the procedure applied successfully above in case of the equilibrium bond length.

Thus considering the valence correlation only, we obtain for the pV6Z basis set and CCSD(T) the value of 1860 cm⁻¹. Based on that we may evaluate the full CCSDT result

to be 1862 cm^{-1} . However, the T_4 correction when added to the CCSD(T) value gives 1867 cm^{-1} ($+7\text{ cm}^{-1}$), whereas when added to the CCSDT reduces it to 1853 cm^{-1} (-9 cm^{-1}). The latter numbers are expected to be reasonably close to the basis set limit for the CCSD(T_{Q_f}) and CCSDT(Q_f) methods for the valence correlation.

In order to get an estimate of the basis set limit for the harmonic frequency of the C_2 molecule with all electrons correlated, we evaluate its value for the pCV6Z basis set. We may do it in two ways, either by introducing the core electron shift to the values obtained for the pV6Z basis set (see penultimate row of Table VIII) or based upon the CCSD(T) value obtained for the pCV6Z basis in Ref. 47. Both ways lead to the same values within 1 cm^{-1} , see the last row of Table VIII. For example, the CCSDT-3 value obtained in that way is 1869 cm^{-1} , whereas the full CCSDT is somewhat higher being equal to 1873 cm^{-1} . We observe that all three values computed with the inclusion of the T_3 operator in various approximations give a C_2 frequency that is slightly too high comparing to the experimental value of 1855.0 cm^{-1} .⁵⁶ The deviations are 16, 14, and 18 cm^{-1} (rounding up the experimental value to full wave number) for the CCSD(T), CCSDT-3, and CCSDT methods, respectively. Now in order to account for the connected quadruple effect, we add the T_4 correction on top of those obtained with the T_3 operator. Thus for the CCSD(T_{Q_f}) method we obtain 1878 cm^{-1} ($+7\text{ cm}^{-1}$), for the CCSDT-3(Q_f) 1867 cm^{-1} (-2 cm^{-1}) and finally for the CCSDT(Q_f) approach 1863 cm^{-1} (-10 cm^{-1}). We notice that also in this case the T_4 correction works in the wrong direction for the CCSD(T) method. In the case of the most rigorous T_3 approach, i.e., the full CCSDT scheme, the T_4 correction works properly reducing the deviation from experiment from 18 cm^{-1} for the CCSDT method to 8 for CCSDT(Q_f). Although the results obtained here for the most rigorous scheme, CCSDT(Q_f) are very close to the experiment, both in case of the C_2 equilibrium bond length as for the harmonic frequency, there still exists a small discrepancy of 0.0003 \AA and 8 cm^{-1} , respectively, the source of which is not clear. One of the possibilities is that the errors are connected with the approximations introduced at the T_4 operator level. In order to clarify this issue we made an attempt to evaluate T_4 in a more rigorous way.

VII. METHODS WITH MORE RIGOROUS INCLUSION OF THE T_4 OPERATOR

The most efficient approach to the T_4 operator within CC theory is the method outlined above where the vertical factorization technique is applied to all T_4 diagrams which creates the CCSDX(Q_f) method with X representing the T operator at various levels of approximation. In Ref. 32 we have shown for small basis set examples that the above method gives practically identical results to those from other approximate models of the CCSDTQ method, i.e., CCSDTQ-1, CCSDT Q_f -1, and CCSDT(Q). Here we make another test of the validity of the factorized method by performing calculations for a broader range of basis sets. In addition we managed to perform the calculation with the full CCSDTQ method for the same basis sets. In Table IX we collected the results of the equilibrium geometry and har-

monic frequency calculations for the four basis sets: pVDZ, pCVDZ, augmented pVDZ, and pVTZ. The methods used are CCSDT(Q), CCSDT Q_f -1, CCSDTQ-1, and the full CCSDTQ. For comparison we also list the results for CCSDT(Q_f) taken from Tables VII and VIII. We want to point out that the present calculations fully support what has been said before on the performance of the (Q_f) method. Namely all methods correct through fifth order give the equilibrium bond length within 0.0001 \AA for all the basis sets considered. Similarly, for the harmonic frequency the results stay within 1 cm^{-1} . This means that the observation made in Ref. 32 with respect to the modest basis sets can be extended to larger ones like the pVTZ set used in the current calculations.

In addition we list in the last column of Table IX the results obtained with the full CCSDTQ method. Taking into account that the sizes of the basis sets are 28, 36, 46, and 60, respectively, we want to emphasize that these are the largest calculations ever made with the full CCSDTQ method. The results obtained are consistent for all basis sets. We observe that an inclusion of the T_4 operator in a rigorous manner increases the bond length by 0.0004 \AA for the DZ type basis sets and by 0.0003 for the TZ set [we compare the CCSDT(Q_f) and CCSDTQ values in Table IX]. Consequently, the change in the harmonic frequency is of opposite sign, being equal to -5 cm^{-1} for all the basis sets considered. Assuming similar behavior of the CCSDTQ method for the remaining larger basis sets, we may estimate the pCV6Z harmonic frequency at the CCSDTQ level to be equal to $\approx 1858\text{ cm}^{-1}$. This compares well with the harmonic experimental value of 1855.0 cm^{-1} .⁵⁶

VIII. CONCLUSIONS

The detailed derivation of all algebraic terms occurring in the factorized expressions for the fifth-order noniterative quadruple contribution is presented. The principal advantage of the current formulation relies on the exploitation of the vertical factorization introduced in Ref. 32 which allows us to combine two vertices at a time within a complex diagram. The first two vertices combined create an intermediate of the first rank and the latter combined with another T vertex or integral results in an intermediate of the second rank which when combined with the remaining vertex, gives rise to the contribution to the T_2 amplitude. At each step the rank of the computational procedure was no worse than n^6 for the ($W_N T_2^2/2$) part and n^7 for the ($W_N T_3$) part. All the terms are derived on the basis of Goldstone diagrams. Although the total number of Goldstone diagrams is 300, the intermediates introduced reduce this significantly since each intermediate can be used for constructing more than one diagram.

The CCSDT(Q_f) method was applied to the calculation of the equilibrium geometry and harmonic frequency for the C_2 molecule using systematic basis sets: the largest of which were pCVQZ and pV5Z with 168 and 182 basis functions, respectively. The T_4 correction was built upon the amplitudes obtained with three methods: CCSD(T), CCSDT-3, and the full CCSDT. For each of the methods the T_4 correction to the equilibrium geometry or to the harmonic fre-

quency is very stable and almost independent of the basis set employed. This permits an extrapolation of the computed frequencies to very large basis sets that approach the basis set limit.

To test the validity of the CCSDT(Q_r) approximation we performed calculations with a more rigorous incorporation of the connected quadruples into the CC theory including CCSDTQ. Here, however, due to the high rank of the computation the calculations were possible only up to the pVTZ basis. Extrapolating the results obtained for the CCSDTQ method to the pCV6Z basis set, we obtain 1.2425 Å for the equilibrium bond length [experimental R_e is equal to 1.24244 Å (Ref. 56)] and 1858 cm^{-1} for the harmonic frequency [experimental value, 1854.7 cm^{-1} (Ref. 56)]. Although our results are obtained with exceptionally high quality wave functions there is still some minor disagreement between the theory and experiment ($\approx 3 \text{ cm}^{-1}$). Possible sources of this inconsistency could be several including inaccuracies in extracting the harmonic frequency experimentally (unlikely), the Born–Oppenheimer approximation, relativistic effects and spin–orbit coupling. However, the most likely source are higher than T_4 excitations. See Ref. 57 for the connected T_5 results.

ACKNOWLEDGMENTS

This work has been supported by the United States Air Force Office of Scientific Research under Grant No. AFOSR-F49620-95-1-0130 and also by KBN, Poland, under Grant No. 2P30310607. Some of the calculations were done on the server of the Silesian Academic Computer Network located in Silesian University of Technology, Gliwice, Poland.

- ¹J. Cizek, *J. Chem. Phys.* **45**, 4256 (1966); *Adv. Chem. Phys.* **14**, 15 (1969); J. Paldus, J. Cizek, and I. Shavitt, *Phys. Rev. A* **5**, 50 (1974).
- ²R. J. Bartlett and G. D. Purvis III, *Int. J. Quantum Chem., Quantum Chem. Symp.* **14**, 561 (1978); *Phys. Scr.* **21**, 255 (1980).
- ³J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, *Int. J. Quantum Chem., Quantum Chem. Symp.* **14**, 545 (1978).
- ⁴P. C. Taylor, G. Bacskay, N. S. Hush, and A. C. Hurley, *J. Chem. Phys.* **69**, 1971 (1978); **69**, 4669 (1978).
- ⁵For reviews, see R. J. Bartlett, *J. Phys. Chem.* **93**, 1697 (1989); in *Modern Electronic Structure Theory, Part 2*, edited by D. R. Yarkony (World Scientific, New York, 1995), pp. 1047–1131; R. J. Bartlett and J. F. Stanton, in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd (VCH, New York, 1994), Vol. 5, pp. 65–169; M. Urban, I. Cernusak, V. Kello, and J. Noga, in *Methods in Computational Chemistry*, edited by S. Wilson (Plenum, New York, 1987), Vol. 1, p. 117.
- ⁶G. D. Purvis III and R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).
- ⁷J. Noga and R. J. Bartlett, *J. Chem. Phys.* **86**, 7041 (1987); **89**, 3401 (1988).
- ⁸G. Scuseria and H. F. Schaefer III, *Chem. Phys. Lett.* **152**, 382 (1988).
- ⁹J. D. Watts and R. J. Bartlett, *J. Chem. Phys.* **93**, 6104 (1989).
- ¹⁰Y. S. Lee, S. A. Kucharski, and R. J. Bartlett, *J. Chem. Phys.* **81**, 5906 (1984).
- ¹¹J. Noga, R. J. Bartlett, and M. Urban, *Chem. Phys. Lett.* **134**, 126 (1987).
- ¹²M. Urban, J. Noga, S. J. Cole, and R. J. Bartlett, *J. Chem. Phys.* **83**, 4041 (1986); S. J. Cole and R. J. Bartlett, *ibid.* **86**, 873 (1987).
- ¹³K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
- ¹⁴J. Paldus, in *Methods in Computational Molecular Physics*, Vol. 293 in NATO Advanced Study Institute, Series B: Physics, edited by S. Wilson and G. H. F. Diercksen (Plenum, New York, 1992), pp. 99–194.
- ¹⁵B. Jeziorski and H. J. Monkhorst, *Phys. Rev. A* **24**, 1668 (1981).
- ¹⁶I. Lindgren and D. Mukherjee, *Phys. Rep.* **151**, 93 (1987).
- ¹⁷B. Jeziorski and J. Paldus, *J. Chem. Phys.* **90**, 2714 (1989).
- ¹⁸P. Piecuch and J. Paldus, *J. Chem. Phys.* **101**, 5875 (1994).
- ¹⁹S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **95**, 8227 (1991).
- ²⁰S. A. Kucharski and R. J. Bartlett, *Int. J. Quantum Chem., Quantum Chem. Symp.* **22**, 383 (1988); **26**, 107 (1992).
- ²¹J. Paldus, J. Cizek, and M. Takahashi, *Phys. Rev. A* **30**, 2193 (1984).
- ²²P. Piecuch, R. Tobola, and J. Paldus, *Phys. Rev. A* **54**, 1210 (1996).
- ²³L. Stolarczyk, *Chem. Phys. Lett.* **217**, 1 (1994).
- ²⁴G. Peris, J. Planelles, and J. Paldus, *Int. J. Quantum Chem.* **62**, 137 (1997).
- ²⁵X. Li, G. Peris, J. Planelles, F. Rajadell, and J. Paldus, *J. Chem. Phys.* **107**, 90 (1997).
- ²⁶G. Peris, F. Rajadell, X. Li, J. Planelles, and J. Paldus, *Mol. Phys.* **94**, 235 (1998).
- ²⁷X. Li and J. Paldus, *J. Chem. Phys.* **107**, 6257 (1997); **108**, 637 (1998).
- ²⁸X. Li and J. Paldus, *Chem. Phys. Lett.* **286**, 145 (1998).
- ²⁹X. Li and J. Paldus, *J. Chem. Phys.* **110**, 2844 (1999).
- ³⁰K. Jankowski and K. Kowalski, *Chem. Phys. Lett.* **256**, 141 (1996).
- ³¹K. Jankowski, I. Grabowski, and K. Kowalski, *J. Chem. Phys.* **109**, 6255 (1998).
- ³²S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **108**, 9221 (1998).
- ³³S. A. Kucharski, J. D. Watts, and R. J. Bartlett, *Chem. Phys. Lett.* **302**, 295 (1999).
- ³⁴S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **110**, 8223 (1999).
- ³⁵P. Piecuch, S. A. Kucharski, and R. J. Bartlett, *J. Chem. Phys.* **110**, 6103 (1999).
- ³⁶L. M. Frantz and R. L. Mills, *Nucl. Phys.* **15**, 16 (1960).
- ³⁷S. A. Kucharski and R. J. Bartlett, *Adv. Quantum Chem.* **18**, 281 (1986).
- ³⁸R. J. Bartlett and J. Noga, *Chem. Phys. Lett.* **150**, 29 (1988); R. J. Bartlett, S. A. Kucharski, J. Noga, J. D. Watts, and G. W. Trucks, *Lecture Notes in Chemistry* (Springer, Berlin, 1989), Vol. 52, p. 125; R. J. Bartlett, S. A. Kucharski, and J. Noga, *Chem. Phys. Lett.* **155**, 133 (1989).
- ³⁹J. Noga, S. A. Kucharski, and R. J. Bartlett, *J. Chem. Phys.* **90**, 3399 (1989).
- ⁴⁰S. A. Kucharski, J. Noga, and R. J. Bartlett, *J. Chem. Phys.* **90**, 7282 (1989).
- ⁴¹S. A. Kucharski and R. J. Bartlett, *Chem. Phys. Lett.* **237**, 264 (1995).
- ⁴²R. J. Bartlett, J. D. Watts, S. A. Kucharski, and J. Noga, *Chem. Phys. Lett.* **165**, 513 (1990).
- ⁴³S. A. Kucharski and R. J. Bartlett, *Theor. Chim. Acta* **80**, 387 (1991).
- ⁴⁴S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **97**, 4282 (1992).
- ⁴⁵S. A. Kucharski and R. J. Bartlett, *Chem. Phys. Lett.* **158**, 550 (1989).
- ⁴⁶J. D. Watts and R. J. Bartlett, *J. Chem. Phys.* **96**, 6073 (1992).
- ⁴⁷K. A. Peterson, A. K. Wilson, D. E. Woon, and T. H. Dunning, Jr., *Theor. Chem. Acc.* **97**, 251 (1997).
- ⁴⁸M. Zeitz, S. D. Peyerimhoff, and R. J. Buenker, *Chem. Phys. Lett.* **58**, 487 (1978).
- ⁴⁹M. Dupuis and B. Liu, *J. Chem. Phys.* **73**, 337 (1980).
- ⁵⁰J. A. Nichols and J. Simons, *J. Chem. Phys.* **86**, 6972 (1987).
- ⁵¹C. W. Bauschlicher, Jr. and S. R. Langhoff, *J. Chem. Phys.* **87**, 2919 (1987).
- ⁵²T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989); R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *ibid.* **96**, 6769 (1992); D. E. Woon and T. H. Dunning, Jr., *ibid.* **103**, 4752 (1995).
- ⁵³W. D. Allen and A. G. Csaszar, *J. Chem. Phys.* **98**, 2983 (1993).
- ⁵⁴C. W. Bauschlicher, Jr. and H. Partridge, *J. Chem. Phys.* **100**, 4329 (1994).
- ⁵⁵J. M. L. Martin, *Chem. Phys. Lett.* **242**, 343 (1995).
- ⁵⁶M. Douay, *J. Mol. Spectrosc.* **131**, 250 (1988).
- ⁵⁷M. Musial, S. A. Kucharski, and R. J. Bartlett, *Chem. Phys. Lett.* **320**, 542 (2000).