

ARTICLES

Formulation and implementation of the full coupled-cluster method through pentuple excitations

M. Musiał and S. A. Kucharski

*Institute of Chemistry, Silesian University, Szkolna 9, 40-006 Katowice, Poland*R. J. Bartlett^{a)}*Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, Florida 32611*

(Received 2 October 2001; accepted 3 December 2001)

Using the quasilinearized formulation of CC theory in terms of recursively computed intermediates, we present the detailed equations and implementation of coupled-cluster theory with single, double, triple, quadruple, and pentuple excitations, CCSDTQP. We illustrate its results by comparison with several full CI results in double zeta, polarized basis sets (DZP), at different geometries. The maximum error compared to full CI occurs for H₂O at twice R_e which is 0.026 mH. For all other cases, HF, SiH₂, and CH₂ in its singlet state, the largest errors are 0.001 mH. The magnitude of the connected T_5 contribution is as large as 0.35 mH, but usually less than 0.1 mH for these examples.

© 2002 American Institute of Physics. [DOI: 10.1063/1.1445744]

I. INTRODUCTION

In the search for a method to provide benchmark results for the electronic correlation energy of molecules too large for full CI, we present the initial, general purpose implementation of the full coupled-cluster method through pentuple excitations, CCSDTQP. The coupled cluster (CC) method^{1–15} has been recognized as an essential tool in the treatment of electron correlation calculations since its introduction into quantum chemistry by Cizek and Paldus.¹ The first general implementation at the level of CCD (CC Doubles) soon appeared^{3,4} followed by the CCSD approach with the full inclusion of the single and double excitations.^{5–7} The higher rank approaches with approximate⁸ CCSDT-1, and full^{9–11} inclusion of connected triple excitations, CCSDT, were subsequently coded and tested, and its noniterative triples approximations, CCSD[T] and particularly its CCSD(T) variant, have become quite popular.^{12,13} The next step in the development of the CC method was the formulation and implementation of connected quadruple excitations, the T_4 operator. This was first done in an approximate form¹⁴ and then followed by its full CCSDTQ realization.¹⁵ The first applications of the CC method with inclusion of the T_4 operator were compared to FCI results in small basis sets.¹⁵ The recent applications^{16,17} involve larger basis sets, e.g., for the C₂ molecule, calculations with the cc-pVTZ basis set,¹⁷ indicated that the full CCSDTQ method is definitely applicable to larger systems than the FCI approach. In addition, the detailed analysis of possible approaches involving the T_4 operator resulted in the new approximate methods denoted as CCSDT(Q_r) and CCSD(TQ_r)¹⁸ in which the T_4 operator is treated in a factorized way. Both methods have an n^7 scaling

for (Q_r), the same as (T). Demonstrating the efficiency, the former, built upon the full CCSDT, has been applied to basis set limit calculations for N₂ and C₂, where the number of basis functions was ~ 200 .^{16,17}

Following a similar philosophy, we evaluated the correlation contribution due to the operator responsible for connected pentuple excitations, T_5 . We began with its approximate treatment resulting in the CCSDTQ(P) method,¹⁹ in which we solve the full CCSDTQ equations and then we evaluate the noniterative contributions due to the lowest order T_5 operator computed on the basis of the converged T_2 , T_3 , and T_4 amplitudes. The role of pentuples was demonstrated to be essential in achieving $\sim 1 \text{ cm}^{-1}$ accuracy in the vibrational frequency of N₂. Now, in the current work we report the results of the rigorous, full CCSDTQP method.

There are two main barriers to the successful realization of the project. The first corresponds to the very complex form of the amplitude equations that define the CCSDTQP model, which makes the coding of the scheme rather cumbersome. The second pertains to the high rank of the computational procedure, n^{12} , which makes the full T_5 calculation quite costly. We know how to partially circumvent the first difficulty. Although the T_5 equation consists of 99 antisymmetrized diagrams (or, equivalently, 205 diagrams if we choose to work within a Goldstone formalism) the method can still be implemented very efficiently. Following the approach introduced in Refs. 15, 20 we may formulate the CCSDTQP equations in a quasilinear form replacing the complicated 99 diagrammatic terms with 14 linear diagrams (or 205 with 20, for the Goldstone formulation). Obviously, in order to reduce the T_5 equation to its linear form, the proper intermediates must be defined and computed. However, most of them are already constructed at the CCSDTQ level.¹⁵ When adding the T_5 operator, only three new inter-

^{a)} Author to whom correspondence should be addressed.

TABLE I. Schematic form of the quasilinear CCSDTQP equations.

$$\begin{aligned}
 \langle \Phi_i^a | (I_0^a + I_1^a T_1 + I_2^a T_2 + I_2'^a T_1 + I_3'^a T_2 + I_4^a T_3)_c | \Phi_0 \rangle &= 0 \\
 \langle \Phi_{ij}^{ab} | (I_0^{ab} + I_1^{ab} T_2 + I_2^{ab} T_3 + I_1'^{ab} T_1 + I_2'^{ab} T_2 + I_3'^{ab} T_3 + I_4^a T_4)_c | \Phi_0 \rangle &= 0 \\
 \langle \Phi_{ijk}^{abc} | (I_1^a T_3 + I_2^a T_4 + I_1'^{ab} T_2 + I_2'^{ab} T_3 + I_3^a T_4 + I_4^a T_5)_c | \Phi_0 \rangle &= 0 \\
 \langle \Phi_{ijkl}^{abcd} | (I_1^a T_4 + I_2^a T_5 + I_1'^{ab} T_3 + I_2'^{ab} T_4 + I_3'^{ab} T_5 + I_1'^{abc} T_2 + I_2'^{abc} T_3)_c | \Phi_0 \rangle &= 0 \\
 \langle \Phi_{ijklm}^{abcde} | (I_1^a T_5 + I_2^a T_4 + I_2'^{ab} T_5 + I_1'^{abc} T_3 + I_2'^{abc} T_4 + I_1'^{abcd} T_2)_c | \Phi_0 \rangle &= 0
 \end{aligned}$$

mediates need to be defined, while two others require a small modification.

There is no immediate remedy to the second difficulty connected with the T_5 operator, barring further approximation of some type. One observation to be made is that the number of combinations of the hole indices to be considered when solving the T_5 equation is comparable with those occurring for the T_3 and T_4 amplitudes. Within the formalism we present for constructing the T_n amplitude the hole indices should fulfill the condition: $i \geq j \geq k \geq l \geq \dots$ and no more than two indices can have the same value. Due to this we have:

- for $n_o = 3$: $n(T_3) = 7, n(T_4) = 6, n(T_5) = 3$;
- for $n_o = 4$: $n(T_3) = 16, n(T_4) = 19, n(T_5) = 16$;
- for $n_o = 5$: $n(T_3) = 30, n(T_4) = 45, n(T_5) = 51$;
- for $n_o = 6$: $n(T_3) = 50, n(T_4) = 90, n(T_5) = 126$, etc.,

where by $n(X)$ we mean the number of the hole index combinations for the X operator; n_o is the number of occupied orbitals. Thus, the computational steps responsible (for e.g., $T_4 \Rightarrow T_4$ and $T_5 \Rightarrow T_5$) differ—for low numbers of the occu-

ried levels—only by the summations over the virtual levels which scale as n_v^6 and n_v^7 , respectively, for the T_4 and T_5 equations.

The second promising route to more general purpose calculations involving higher cluster operators is indicated in Refs. 21–23, where the higher excitation operators (in the cited works— T_3 and T_4) may be defined within an active subspace of the full one-particle space, the size of which is significantly smaller than that of the full space. This allows for the inclusion of the higher cluster operators in a very efficient way, presumably retaining (due to the properly selected active space) all their important effects. The same scheme can be easily extended to the T_5 operator, and since the active space corresponds to a small number of one-particle states, the T_5 operator can be treated very efficiently. A third route is to explore our factorized ansatz, namely writing the CC wave function as $\exp(T_f)\exp(T_u)$, where the T_f indicates diagrams which may be factorized in the new way we discuss elsewhere,¹⁸ and T_u are those that do not permit such factorization. If the numerical effects of the latter can be proved to be bounded, this ansatz offers a very powerful one that can be applied to quite high-order cluster operators.

In the next section we give a short description of the CCSDTQP method with definitions of the amplitude equations and required intermediates. This will be followed by the results of the test calculations.

II. THEORY

General coupled-cluster equations are obtained by inserting the wave function, expressed through the exponential Ansatz

TABLE II. Coupled cluster equations for the CCSDTQP method in quasilinear form.

	Expression ^a
$D_i^a t_i^a =$	$f_i^a + t_i^a I_f^a - t_n^a I_n^a + t_n^a I_n^a + t_n^a v_{fi}^{na} - \frac{1}{2} t_{no}^a v_{fi}^{no} + \frac{1}{2} t_{ni}^a v_{fg}^{na} + \frac{1}{4} t_{ino}^a v_{fg}^{no}$
$D_{ij}^{ab} t_{ij}^{ab} =$	$v_{ij}^{ab} + P(a/b) I_f^a t_{ij}^{fb} - P(i/j) I_n^a t_{nj}^{ab} + \frac{1}{2} t_{ij}^a I_{fg}^{ab} + \frac{1}{2} t_{no}^a I_{ij}^{no}$ $+ P(a/b) P(i/j) t_{in}^{af} t_{fj}^{nb} - \frac{1}{2} P(a/b) I_{fg}^{ab} t_{ni}^{fb}$ $- \frac{1}{2} P(i/j) I_{fi}^{no} t_{noj}^{fab} + t_{ni}^{ab} I_f^a + P(i/j) t_{ij}^{ab} - P(a/b) t_n^a I_{ij}^{nb} + \frac{1}{4} t_{ino}^{ab} v_{fg}^{no}$
$D_{ijk}^{abc} t_{ijk}^{abc} =$	$P(a/bc) I_f^a t_{ijk}^{fbc} - P(i/jk) I_n^a t_{njkl}^{abc} + \frac{1}{2} P(a/bc) t_{ij}^{af} I_{fg}^{bc} + \frac{1}{2} P(i/jk) t_{ino}^{abc} I_{jk}^{no}$ $+ P(ab/c) P(ij/k) t_{ijn}^{abf} t_{fjk}^{nc} + P(a/bc) P(ij/k) t_{ij}^{af} I_{fjk}^{bc} - P(ab/c) P(ij/k) t_{in}^{ab} t_{ijk}^{nc}$ $+ t_{nij}^{abc} I_f^a + \frac{1}{2} P(a/bc) I_{fg}^{abc} t_{nij}^{bc} - P(i/jk) I_{fi}^{no} t_{noj}^{fab} + \frac{1}{4} t_{ijkno}^{ab} v_{fg}^{no}$
$D_{ijkl}^{abcd} t_{ijkl}^{abcd} =$	$P(a/bcd) I_f^a t_{ijkl}^{abcd} - P(i/jkl) I_n^a t_{njkl}^{abcd} + \frac{1}{2} P(ab/cd) t_{ij}^{abf} t_{fg}^{cd} + \frac{1}{2} P(ij/kl) t_{ino}^{abcd} I_{lm}^{no}$ $+ P(abc/d) P(ijk/l) t_{ijkn}^{abcf} t_{fl}^{nd} + P(ab/cd) P(ijk/l) t_{ijk}^{abf} t_{fl}^{cd} - P(abc/d) P(ij/kl) t_{ijn}^{ab} t_{kl}^{ncd}$ $+ P(a/bcd) P(ij/kl) t_{ij}^{af} t_{fjkl}^{bcd} - P(ab/cd) P(ij/kl) t_{in}^{ab} I_{jkl}^{ncd} + P(ab/cd) P(ij/kl) t_{ij}^{ab} I_{jkl}^{ncd}$ $+ \frac{1}{2} P(abc/d) P(ij/kl) t_{ino}^{abcd} t_{jkl}^{nod} + t_{nijkl}^{abcd} I_f^a + \frac{1}{2} P(a/bcd) I_{fg}^{abcd} t_{nijkl}^{abcd}$ $- \frac{1}{2} P(ij/kl) I_{fi}^{no} t_{nojkl}^{fabcd}$
$D_{ijklm}^{abcde} t_{ijklm}^{abcde} =$	$P(a/bcde) I_f^a t_{ijklm}^{bcde} - P(i/jklm) I_n^a t_{njklm}^{abcde} + \frac{1}{2} P(abc/de) t_{ijklm}^{abcf} t_{de} + \frac{1}{2} P(ijk/lm) t_{ijkno}^{abcde} I_{lm}^{no}$ $+ P(abcde) P(ijkl/m) t_{ijklm}^{abcf} t_{fm}^{ne} + P(abc/de) P(ijk/lm) t_{ijkn}^{abcf} I_{lm}^{nde}$ $+ \frac{1}{2} P(abcde) P(ijklm) t_{ijno}^{abcd} I_{klm}^{noe} + \frac{1}{2} P(ab/cde) P(ijk/lm) t_{ijkl}^{abfg} I_{fgm}^{cde}$ $+ P(abc/de) P(ijk/lm) t_{ijkl}^{abcf} I_{lm}^{cde} - P(abcde) P(ijk/lm) t_{ijkn}^{abcf} t_{lm}^{cde}$ $+ P(ab/cde) P(ijk/lm) t_{ijk}^{abf} t_{flm}^{cde} - P(abc/de) P(ijklm) t_{ijn}^{ab} t_{klm}^{ncde}$ $+ P(a/bcde) P(ijklm) t_{ij}^{af} t_{fjklm}^{bcde} - P(ab/cde) P(ijklm) t_{in}^{ab} I_{jklm}^{ncde}$

^aSummation over repeated indices assumed; ij, \dots , run over occupied one-particle states; ab, \dots , run over virtual one-particle states; $P(i/j)$ or $P(a/b)$ implies sum of two components differing by permutation of i, j and a, b indices, respectively; $P(ab\dots/ef\dots)$ indicates that in addition to the identity permutation the summation should include all possible permutations exchanging labels between subsets $(ab\dots)$ and $(ef\dots)$. The same refers to $P(ij\dots/mn\dots)$. The $I_{ij\dots}^{ab\dots}$ intermediates are defined in Table IV.

TABLE III. Interrelations between various forms of intermediates and contributing standard diagrams.

	I_0^1	I_1^1	$I_1^{1'}$	I_2^1	I_0^2	I_1^2	$I_1^{2'}$	$I_1^{2''}$	I_2^2	$I_2^{2'}$	$I_2^{2''}$	I_3^2	$I_3^{2'}$	I_4^2	$I_1^{3'}$	$I_1^{3''}$	$I_2^{3'}$	$I_2^{3''}$	$I_1^{4'}$
OAI ^a	1	2	2	1	1	2	2	2	3	1	3	2	2	1	2	2	2	3	2
OGI ^a	1	2	2	1	1	2	2	2	4	2	4	2	2	1	2	2	3	4	2
A ^b	1	10	8	2	1	26	8	24	13	1	12	4	2	1	23	30	2	3	17
G ^b	1	13	16	3	1	48	9	45	20	2	19	4	2	1	44	63	5	6	36

^aNumber of oriented-line antisymmetrized (Goldstone) intermediates corresponding to I_k^n arrowless intermediate.

^bNumber of standard antisymmetrized (Goldstone) diagrams contributing to I_k^n arrowless intermediate.

$$\Psi = e^T \Phi_o \quad (1)$$

into the Schrödinger equation which then takes the form

$$H_N e^T |\Phi_o\rangle = E_{CC} e^T |\Phi_o\rangle, \quad (2)$$

$$e^{-T} H_N e^T |\Phi_o\rangle = (H_N e^T)_c |\Phi_o\rangle = E_{CC} |\Phi_o\rangle, \quad (3)$$

where the normal ordered Hamiltonian H_N is defined as

$$\begin{aligned} H_N &= H - \langle \Phi_o | H | \Phi_o \rangle \\ &= \sum_r e_r \{r^\dagger r\} + \sum_{rs} f_{rs} \{r^\dagger s\} + \frac{1}{4} \sum_{rstu} v_{tu}^{rs} \{r^\dagger s^\dagger ut\} \\ &= H_N^o + F_N + W_N. \end{aligned} \quad (4)$$

The e_r are one-particle eigenenergies and F_N and W_N are one- and two-body operators with F_N disappearing for the canonical Hartree–Fock reference state; $v_{tu}^{rs} = \langle rs || tu \rangle$ is an antisymmetrized two-electron integral; E_{CC} is the coupled-cluster correlation energy and the subscript c recognizes that in the commutator expansion of $e^{-T} H_N e^T$ only connected terms (i.e., those which show common indices among T 's and H) survive, as any independently summed term will cancel from the commutator; $|\Phi_o\rangle$ is the reference determinant.

The working equation for the CCSDTQP amplitudes are obtained by projecting Eq. (3) onto the subspaces of excitations. As the cluster operator is expressed as

$$T = T_1 + T_2 + T_3 + T_4 + T_5, \quad (5)$$

where $T_n = (n!)^{-2} \sum_{ij\dots} t_{ij\dots}^{ab\dots} a^\dagger b^\dagger \dots ji$, we obtain five sets of equations for single, double, triple, quadruple, and pentuple excitation amplitudes, respectively,

$$\langle \Phi_i^a | (H_N e^T)_c | \Phi_o \rangle = 0, \quad (6)$$

$$\langle \Phi_{ij}^{ab} | (H_N e^T)_c | \Phi_o \rangle = 0, \quad (7)$$

$$\langle \Phi_{ijk}^{abc} | (H_N e^T)_c | \Phi_o \rangle = 0, \quad (8)$$

$$\langle \Phi_{ijkl}^{abcd} | (H_N e^T)_c | \Phi_o \rangle = 0, \quad (9)$$

$$\langle \Phi_{ijklm}^{abcde} | (H_N e^T)_c | \Phi_o \rangle = 0. \quad (10)$$

The more detailed form of the above equations can be obtained by expanding the exponential Ansatz. For the sake of compactness we explicitly list only those terms which do not occur in the CCSDTQ approach:¹⁵

$$T_1(\text{CCSDTQP}) = T_1(\text{CCSDT}), \quad (11)$$

$$T_2(\text{CCSDTQP}) = T_2(\text{CCSDTQ}), \quad (12)$$

$$T_3(\text{CCSDTQP}) = T_3(\text{CCSDTQ}) + R_3(W_N T_5), \quad (13)$$

$$\begin{aligned} T_4(\text{CCSDTQP}) &= T_4(\text{CCSDTQ}) + R_4(F_N T_5 \\ &\quad + W_N(T_5 + T_1 T_5))_c, \end{aligned} \quad (14)$$


$$\begin{aligned} T_5(\text{CCSDTQP}) &= R_5(F_N(T_5 + T_1 T_5 + T_2 T_4 + T_3^2/2) \\ &\quad + W_N(T_4 + T_5 + T_1 T_4 \\ &\quad + T_1 T_5 + T_2 T_3 + T_2 T_4 + T_2 T_5 + T_3^2/2 \\ &\quad + T_3 T_4 + T_1^2 T_4/2 + T_1^2 T_5/2 + T_1 T_2 T_3 \\ &\quad + T_1 T_2 T_4 + T_1 T_3^2/2 \\ &\quad + T_2^3/6 + T_2^2 T_3/2 + T_1^3 T_4/6 + T_1^2 T_2 T_3/2 \\ &\quad + T_1 T_2^3/6))_c. \end{aligned} \quad (15)$$

Here $T_n(X)$ represents the set of terms occurring in the T_n equation of the X method, $\langle ij\dots |^{ab\dots}$ is an n -tuply excited determinant and the R_n operator is defined as

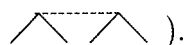
$$R_n(X) = (n!)^{-2} \sum \frac{\langle \Phi_{ij\dots}^{ab\dots} | X | \Phi_o \rangle}{e_i + e_j + \dots - e_b - e_a} a^\dagger b^\dagger \dots ji \quad (16)$$

to ensure the presence of the required denominator and the proper projection subspace for the T_n operator.

The quasilinear form of the CC equations for the current model is defined in the most compact way in Table I. We see that in each term only one cluster operator occurs; all others are “hidden” in the properly defined intermediates, I_k^n , where n and k indicate that this is an n -body intermediate with k annihilation lines. Since the annihilation lines originate exclusively from the two-electron integral, the k index is never greater than 4. For example, the I_1^2 indicates the two-body intermediate with one annihilation line

(diagrammatically denoted as );

the I_4^2 (two-body with four annihilation lines) represents just the two-electron integral

(diagrammatic form: ).

The intermediates introduced in Table I occur either in their complete form (not primed) or in their incomplete one (primed and double-primed). The complete form tells us that the intermediate can be identified with the terms occurring in the \bar{H} expansion:

TABLE IV. Algebraic expression for the intermediates used in the CCS-DTQP equations.

Intermediate	Expression ^a
I_a^i	$f_a^i + t_n^f v_{af}^{in}$
I_b^a	$f_b^a + t_n^f v_{bf}^{an} - \frac{1}{2} t_{no}^{af} v_{bf}^{no} - t_n^{an}$
I_j^i	$I_j^i + t_n^f v_{jf}^{ij}$
I_j^{ii}	$f_j^i + t_n^f v_{jf}^{in} + \frac{1}{2} t_{jn}^{fg} v_{fg}^{in}$
I_{bc}^{ai}	$\chi_{bc}^{ai} - \frac{1}{2} t_n^a v_{bc}^{ni}$
χ_{bc}^{ai}	$v_{bc}^{ai} - \frac{1}{2} t_n^a v_{bc}^{ni}$
I_{ka}^{ij}	$\chi_{ka}^{ij} + \frac{1}{2} t_k^i v_{fa}^{ij}$
χ_{ka}^{ij}	$v_{ka}^{ij} + \frac{1}{2} t_k^i v_{fa}^{ij}$
I_{cd}^{ab}	$I_{cd}^{ab} + \frac{1}{2} t_{no}^{ab} v_{cd}^{no}$
I_{cd}^{ab}	$v_{cd}^{ab} - P(a/b) \chi_{cd}^{an} t_n^b$
I_{kl}^{ij}	$v_{kl}^{ij} + P(k/l) \chi_{kl}^{ij} t_n^f + \frac{1}{2} t_{kl}^{fg} v_{fg}^{ij}$
I_{bj}^{ia}	$\chi_{bj}^{ia} + v_{bj}^{in,af} + \frac{1}{2} t_{bj}^{ia} t_n^f$
χ_{bj}^{ia}	$v_{bj}^{ia} - \frac{1}{2} v_{bj}^{in,a} + \chi_{bj}^{ia} t_n^f$
χ_{bj}^{ia}	$v_{bj}^{ia} - \frac{1}{2} v_{bj}^{in,a} + \frac{1}{2} \chi_{bj}^{ia} t_n^f$
χ_{bj}^{ia}	$v_{bj}^{ia} - v_{bj}^{in,a} + \frac{1}{2} t_{bj}^{ia} t_n^f$
I_{ci}^{ab}	$v_{ci}^{ab} + v_{cf}^{ab} t_n^i - P(a/b) t_n^a \chi_{ci}^{nb} - I_{ci}^{an} t_n^b + P(a/b) I_{cf}^{an} t_n^b + \frac{1}{2} t_{no}^{ab} I_{ci}^{no}$
I_{ci}^{ab}	$v_{ci}^{ab} + \frac{1}{2} v_{cf}^{ab} t_n^i - P(a/b) t_n^a \chi_{ci}^{nb}$
I_{jk}^{ia}	$I_{jk}^{ia} + I_{jk}^{ia} t_n^f$
I_{jk}^{ia}	$v_{jk}^{ia} - v_{jk}^{in,af} + P(j/k) t_n^f \chi_{jk}^{ia} + P(j/k) I_{jj}^{in,af} t_n^f + \frac{1}{2} t_{fg}^{ia} t_n^f$
I_{jk}^{ia}	$v_{jk}^{ia} - \frac{1}{2} v_{fg}^{in,af} + \frac{1}{2} v_{fg}^{in,af} t_n^f$
I_{dij}^{abc}	$\frac{1}{2} P(ab/c) I_{df}^{ab,fc} + P(a/bc) I_{df}^{an,bcf} + \frac{1}{2} P(i/j) I_{di}^{no,abc} - \frac{1}{2} v_{df}^{no,abc}$
I_{dij}^{abc}	$P(ab/c) I_{df}^{ab,fc} + P(bc/a) P(i/j) I_{di}^{an,bc} + \frac{1}{2} P(i/j) I_{di}^{no,abc} + P(a/bc) I_{df}^{an,bcf} - \frac{1}{2} v_{df}^{no,abc}$
I_{jkl}^{iab}	$-\frac{1}{2} P(jk/l) I_{jk}^{ia,ab} - P(a/b) P(kl/j) I_{jj}^{ia,fb} + P(j/k) I_{jj}^{in,abf} + \frac{1}{2} P(a/b) I_{fg}^{ia,fb} + \frac{1}{2} v_{fg}^{in,abf}$
I_{jkl}^{iab}	$-P(jk/l) I_{jk}^{in,ab} - P(a/b) P(kl/j) I_{jj}^{ia,fb} + P(j/k) I_{jj}^{in,abf} + \frac{1}{2} P(a/b) I_{fg}^{ia,fb} + I_{jj}^{in,abf} + \frac{1}{2} v_{fg}^{in,abf}$
I_{klm}^{ija}	$\frac{1}{2} v_{fg}^{ij,fga}$
I_{dei}^{abc}	$\frac{1}{2} v_{no}^{abc,de}$
I_{cjk}^{iab}	$\frac{1}{2} v_{cf}^{in,abf}$
I_{cjk}^{iab}	$v_{cf}^{in,abf}$
χ_{cjk}^{iab}	$I_{cjk}^{iab} + P(a/b) I_{cf}^{ia,fb} - P(j/k) I_{cj}^{in,ab}$
I_{eijk}^{abcd}	$P(bcd/a) I_{fe}^{abcd} + \frac{1}{2} P(i/jk) I_{ei}^{no,abcd} + \frac{1}{2} P(abc/d) P(i/jk) I_{eif}^{abc,fd} - \frac{1}{2} v_{ef}^{no,abcd}$
I_{jklm}^{iabc}	$P(klm/j) I_{jj}^{iabc,klm} + \frac{1}{2} P(a/bc) I_{fg}^{ia,fb} + \frac{1}{2} P(a/bc) P(jkl/m) I_{jkl}^{ian,fb} + P(c/ab) P(lm/jk) \chi_{jkl}^{iab,fc} + \frac{1}{2} v_{fg}^{in,abc}$

^aSummation over repeated indices assumed, summation indices belong to the set $\{f,g,n,o\}$.

$$\bar{H} = e^{-T} H e^T. \tag{17}$$

We see, e.g., that in the T_5 equation the first three intermediates (I_1^1, I_1^2, I_2^2) occur in the complete form, while in the last three terms their incomplete form should be applied ($I_1^{n3}, I_2^{n3}, I_1^{4}$).

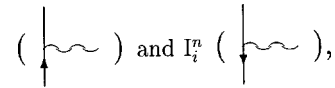
The two incomplete intermediates of the same type usually differ, i.e., a different number of standard diagrams contribute to them, hence we use prime and double-prime to distinguish between them. For example, the incomplete intermediate I_1^2 occurring in the T_2 and T_3 equations takes a

different form in either of the two equations; hence in the T_2 equation we use $I_1^{\prime 2}$, and in the T_3 equation, $I_1^{\prime\prime 2}$.

We should be aware that the intermediates denoted as I_k^n in Table I correspond to the diagrammatic terms expressed in arrowless form (sometimes this kind of diagram is referred to as a skeleton diagram²⁴). On the other hand, the $I_{uv}^{s\dots}$ intermediates, used in the CC equations, see Table II, refer to the regular (i.e., with oriented lines) diagrammatic expressions. For example the I_1^1 intermediate, see Table I can be expressed diagrammatically as



in fact, it corresponds to two terms in Table II: I_f^a



with the diagrammatic forms in parenthesis. In general the intermediates I_1^1 (one annihilation line) and I_3^3 (three annihilation lines, $n > 1$) correspond to two regular diagrammatic terms, and the intermediates I_2^2 (two annihilation lines, $n > 1$) correspond to three regular diagrammatic terms. To clarify this point, we give following the definition of the I_k^n intermediates expressed through their regular forms as used in the T_5 equation of Table II.

$$I_1^1 = \sum I_f^a a^\dagger f + \sum I_i^n n^\dagger i,$$

$$I_1^2 = \sum I_{fm}^{de} d^\dagger e^\dagger m f + \sum I_{lm}^{ne} n^\dagger e^\dagger m l,$$

$$I_2^2 = \sum I_{fg}^{de} d^\dagger e^\dagger g f + \sum I_{lm}^{no} n^\dagger o^\dagger m l + \sum I_{fm}^{ne} n^\dagger e^\dagger m f,$$

$$I_1^{n3} = \sum I_{flm}^{cde} c^\dagger d^\dagger e^\dagger m l f + \sum I_{klm}^{nde} n^\dagger d^\dagger e^\dagger m l k,$$

$$I_2^{n3} = \sum I_{flm}^{nde} n^\dagger d^\dagger e^\dagger m l f + \sum I_{klm}^{noe} n^\dagger o^\dagger e^\dagger m l k$$

$$+ \sum I_{fgm}^{cde} c^\dagger d^\dagger e^\dagger m g f,$$

$$I_1^4 = \sum I_{fklm}^{bcde} b^\dagger c^\dagger d^\dagger e^\dagger m l k f$$

$$+ \sum I_{jklm}^{ncde} n^\dagger c^\dagger d^\dagger e^\dagger m l k j.$$

Thus, although we have only six terms contributing to the T_5 equation in Table I, when writing down the intermediates in their regular (i.e., oriented-line) diagrammatic form, the number of relevant terms increases to 14; see the last equation in Table II.

We should mention also that there exists a very simple formula to evaluate the rank of the computational procedure

TABLE V. Correlation corrections with various CC methods relative to FCI^a values [mH].

		CCSD	CCSDT	CCSDTQ	CCSDTQP
HF(DZP)	R_e	3.006	0.266	0.018	0.000
	$1.5R_e$	5.099	0.646	0.041	0.000
	$2.0R_e$	10.181	1.125	0.062	0.001
H ₂ O(DZP)	R_e	4.122	0.531	0.023	0.002
	$1.5R_e$	10.158	1.784	0.139	0.025
	$2.0R_e$	21.404	-2.472	-0.015	0.026
SiH ₂ (DZP)	R_e	2.843	0.100	0.002	0.001
	$1.5R_e$	6.685	0.058	-0.015	0.001
	$2.0R_e$	14.869	-3.689	-0.346	0.001
CH ₂ (DZP)	R_e	3.544	0.206	0.007	0.000
	$1.5R_e$	6.961	0.310	0.026	0.000
	$2.0R_e^b$	14.648	-1.900	-0.050	0.000

^aThe FCI values, basis sets and geometry taken from Ref. 27 for HF; Ref. 28 for H₂O; Ref. 29 for SiH₂ and Ref. 30 for CH₂.

^bFor CH₂ system the FCI $2R_e$ result is not available; the CCSDTQP value was used as the reference.

for each term in Table I. Namely, for the general term $I_k^n T_m$ the rank of the computational procedure (scaling) is equal to $n^{2(n+m)-k}$.

The relationships between the I_k^n intermediates in Table I and their oriented-line counterparts are presented in Table III. The second (third) row of that Table indicates the number of regular, i.e., with oriented lines, antisymmetrized (Goldstone) forms corresponding to the given intermediate listed in the first row. The fourth (fifth) row gives the number of standard antisymmetrized (Goldstone) diagrams contributing to the I_k^n intermediate. For instance, by summing the numbers (10+26+13+30+3+17=99) corresponding to the intermediates occurring in the T_5 equation ($I_1^1, I_1^2, I_2^2, I_1^3, I_2^3, I_1^4$) we obtain the full number of antisymmetrized diagrams contributing to the T_5 equation written in the standard form. This can be done for each CC equation. Similarly the last row of Table III helps us to evaluate the total number of the Goldstone diagrams contributing to the given CC equation, although in this case one should keep in mind that certain intermediates (having two annihilation lines: particle and hole) can be contracted with T_n in more than one way.

The important thing is that once a given intermediate is constructed it can be used in all equations. For example, the I_1^1 intermediate is composed of 10 antisymmetrized diagrams, and in each equation, the $I_1^1 T_n$ term represents 10

antisymmetrized diagrams contributing to the T_n equation (in case of T_1 the I_1^1 consists of 8 diagrams).

In Table II we give the detailed algebraic expressions of the terms occurring in the CC equation for t_i^a , t_{ij}^{ab} , t_{ijk}^{abc} , t_{ijkl}^{abcd} and t_{ijklm}^{abcde} amplitudes. To each term the proper permutation of the external indices should be applied. The intermediates $I_{uv}^{s\dots}$ are defined in Table IV. Note that we adopted in the last two Tables the tensor notation with implied summation over repeated indices. Throughout the paper the usual notation convention is assumed with letters: a, b, \dots (i, j, \dots) representing the particle (hole) indices, and the r, s, \dots —the general indices. The annihilation operator is represented by the presence of the hole index as the superscript or the particle index as the subscript. Thus e.g., the I_{bj}^{ia} intermediate corresponds to two annihilation operators, $:$ indicated with the indices i (hole and superscript) and b (particle and subscript).

The intermediates collected in Table IV are defined in the recursive way (RGI—recursive generation of intermediates, see Ref. 20), i.e., the lower rank intermediate can be used in the definition of the intermediate of higher rank.

III. RESULTS AND DISCUSSION

In order to verify the correctness of the program as a first step we reproduced the CCSDTQP results of Hirata and

TABLE VI. Net correlation effects due to particular cluster operators [mH].

		$\Delta E(\text{CCSD})$	$\Delta E(T_3)$	$\Delta E(T_4)$	$\Delta E(T_5)$
HF(DZP)	R_e	-200.876	-2.740	-0.248	-0.018
	$1.5R_e$	-222.066	-4.453	-0.605	-0.041
	$2.0R_e$	-253.355	-9.056	-1.063	-0.061
H ₂ O(DZP)	R_e	-211.960	-3.591	-0.508	-0.021
	$1.5R_e$	-260.753	-8.374	-1.645	-0.114
	$2.0R_e$	-348.579	-23.876	2.457	0.041
SiH ₂ (DZP)	R_e	-114.008	-2.743	-0.098	-0.001
	$1.5R_e$	-137.469	-6.627	-0.073	0.016
	$2.0R_e$	-210.241	-18.558	3.343	0.347
CH ₂ (DZP)	R_e	-137.342	-3.338	-0.119	-0.007
	$1.5R_e$	-172.173	-6.651	-0.284	-0.026
	$2.0R_e$	-242.752	-16.548	1.850	0.050

Bartlett,²⁶ obtained with a FCI program for HF and H₂O molecules at the 6-31G basis set level. To investigate the performance of the method for larger basis sets we selected four molecules (HF, H₂O, SiH₂, and CH₂) for which the FCI values are available at the DZP level. The results are collected in Tables V and VI. The values in Table V represent the deviations of the calculated correlation energies from the exact (i.e., FCI) results. Table VI provides the net correlation corrections due to the particular cluster operators.

For the HF molecule the CCSDTQP results are right at the FCI values for all three bond lengths: equilibrium ($r = R_e$), intermediate ($r = 1.5R_e$) and stretched ($r = 2R_e$). It is interesting to observe the improvement in the correlation energy due to each cluster operator. We see that the CCSD method generates errors of ca. 3, 5, and 10 mHartree for the three considered geometries. An inclusion of the T_3 operator (CCSDT method) reduces the error by approximately an order of magnitude to ca. 0.3, 0.6, and 1.1 mHartree. The more sophisticated approach, CCSDTQ, brings the errors down again by somewhat more than one order of magnitude to 0.02, 0.04, and 0.06 mHartree. Finally, the full inclusion of the T_5 operator reduces the error below the accuracy of the reported results, i.e., below 0.001 mHartree (for the $r = 2R_e$ the error rounds up to 0.001 mHartree). We see from the net values of the correlation corrections connected with each cluster operator collected in Table VI, that for the HF molecule the CC method behaves in a quasivariational manner approaching the exact value from above, i.e., each net value is negative.

The next examples refer to the situation where two bonds are simultaneously stretched. A general observation to be made here is that while the T_3 operator substantially reduces the CCSD error for the equilibrium and intermediate distances, it overshoots in the case of the stretched bond (although the absolute deviation for that geometry is also much smaller than that of CCSD). This is particularly apparent in the case of the SiH₂ molecule where the CCSD error of 14.869 mHartree is reduced to -3.689 mHartree. The T_4 correction for this case is large and positive, see Table VI, being equal to 3.343 mHartree. However, this is not enough to compensate the total negative error due to the T_3 operator. The CCSDTQ method still overshoots by 0.346 mHartree. The T_5 correction works in the same direction and, with the value of 0.347 mHartree, reduces the CCSDTQP error to 0.001 mHartree.

That situation is characteristic of all three systems considered with two single bonds. For the equilibrium and intermediate geometries the T_4 and T_5 corrections are both negative, approaching the exact value in a quasivariational manner as in the case of the HF molecule. The CCSDTQP error for these geometries is of the order of 0.001 mHartree or lower, with an exception for the H₂O molecule, where it is equal to 0.002 and 0.025 mHartree, respectively for the R_e and $1.5R_e$ bond lengths. For the stretched geometry both T_4 and T_5 corrections are large and positive, working consistently to cancel the negative error introduced by the T_3 operator. We see that in the case of the H₂O molecule an inclusion of T_5 still leaves some error for the intermediate and stretched geometries. However, the deviations are nearly the

same which means that the CCSDTQP curve stays parallel to the exact one.

IV. CONCLUSIONS

We have shown that with the proper factorization of the diagrammatic terms the implementation of the full CCSDTQP method is feasible. The program requires coding of fourteen terms contributing to the T_5 equation supplemented with small modifications for the T_4 and T_3 equations. Three new intermediates have been defined; two of them are represented by the $I_1'^4$ type, the third one $I_{fgm}'^{cde}$ belongs to the $I_2''^3$ type. A small modification was required also in the case of the I_1^3 intermediate, in the T_5 equation referred to as $I_1''^3$.

Computationally, the method is not yet as efficient as it can be, since although it scales inherently as $n_o^5 n_v^7$, some of the terms contributing to the T_5 equation require quite a large number of permutations. In this, the program can be substantially improved by eliminating permutations among the hole indices of the same value and applying appropriate symmetrization procedures, as is currently being done.

The results obtained with CCSDTQP generate an accuracy of the order of 0.001 mHartree in the majority of the cases studied. This indicates that the approach is a good candidate as a method that can provide high accuracy benchmark values for small molecular systems.

ACKNOWLEDGMENTS

The authors thank Dr. So Hirata from the Quantum Theory Project, University of Florida, for providing high accuracy reference results obtained with his FCI program, which were very helpful in debugging the program. This work has been supported by the United States Air Force Office of Scientific Research under Grant No. AFOSR-F49620-98-0116 and also by the Committee for Scientific Research (KBN, Poland), under Grant No. 3 T09A 087 18.

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).

²For reviews see: R. J. Bartlett, J. Phys. Chem. **93**, 1697 (1989); R. J. Bartlett, in *Modern Electronic Structure Theory*, Part I, edited by D. R. Yarkony (World Scientific Publishing Co., New York, 1995), pp. 1047–1131; R. J. Bartlett and J. F. Stanton, in *Reviews in Computational Chemistry*, Vol. 5, edited by K. B. Lipkowitz and D. B. Boyd (VCH, New York, 1994), pp. 65–169; M. Urban, I. Cernusak, V. Kello, and J. Noga, in *Methods in Computational Chemistry I*, edited by S. Wilson (Plenum, New York, 1987), p. 117; J. Paldus, in *Methods in Computational Molecular Physics*, NATO ASI, 1991.

³R. J. Bartlett and G. D. Purvis III, Int. J. Quantum Chem., Quantum Chem. Symp. **14**, 561 (1978); Phys. Scr. **21**, 255 (1980).

⁴J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem., Quantum Chem. Symp. **14**, 545 (1978).

⁵G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).

⁶G. E. Scuseria, T. J. Lee, and H. F. Schaefer, Chem. Phys. Lett. **130**, 236 (1986).

⁷T. J. Lee and J. E. Rice, Chem. Phys. Lett. **150**, 406 (1988).

⁸Y. S. Lee, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **81**, 5906 (1984).

⁹J. Noga and R. J. Bartlett, J. Chem. Phys. **86**, 7041 (1987); *ibid.* **89**, 3401 (1988).

¹⁰G. Scuseria and H. F. Schaefer III, Chem. Phys. Lett. **152**, 382 (1988).

¹¹J. D. Watts and R. J. Bartlett, J. Chem. Phys. **93**, 6104 (1989).

- ¹²M. Urban, J. Noga, S. J. Cole, and R. J. Bartlett, *J. Chem. Phys.* **83**, 4041 (1985).
- ¹³K. Raghvachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
- ¹⁴S. A. Kucharski and R. J. Bartlett, *Chem. Phys. Lett.* **158**, 550 (1989).
- ¹⁵S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **97**, 4282 (1992).
- ¹⁶S. A. Kucharski, J. D. Watts, and R. J. Bartlett, *Chem. Phys. Lett.* **302**, 295 (1999).
- ¹⁷S. A. Kucharski, M. Kolaski, and R. J. Bartlett, *J. Chem. Phys.* **114**, 692 (2001).
- ¹⁸S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **108**, 9221 (1998).
- ¹⁹M. Musiał, S. A. Kucharski, and R. J. Bartlett, *Chem. Phys. Lett.* **320**, 542 (2000).
- ²⁰S. A. Kucharski and R. J. Bartlett, *Theor. Chim. Acta* **80**, 387 (1991).
- ²¹P. Piecuch, S. A. Kucharski, and R. J. Bartlett, *J. Chem. Phys.* **110**, 6103 (1999).
- ²²P. Piecuch and L. Adamowicz, *J. Chem. Phys.* **100**, 5792 (1994).
- ²³P. Piecuch, N. Oliphant, and L. Adamowicz, *J. Chem. Phys.* **99**, 1875 (1993).
- ²⁴S. A. Kucharski and R. J. Bartlett, *Adv. Quantum Chem.* **18**, 281 (1986).
- ²⁵M. Musiał, S. A. Kucharski, and R. J. Bartlett, *Mol. Phys.* (in press).
- ²⁶S. Hirata, and R. J. Bartlett, *Chem. Phys. Lett.* **321**, 216 (2000).
- ²⁷C. W. Bauschlicher, Jr., S. R. Langhoff, P. R. Taylor, N. C. Handy, and P. J. Knowles, *J. Chem. Phys.* **85**, 1469 (1986).
- ²⁸C. W. Bauschlicher, Jr. and P. R. Taylor, *J. Chem. Phys.* **85**, 2779 (1986).
- ²⁹C. W. Bauschlicher, Jr. and P. R. Taylor, *J. Chem. Phys.* **86**, 1420 (1987).
- ³⁰C. W. Bauschlicher, Jr. and P. R. Taylor, *J. Chem. Phys.* **86**, 2844 (1987).