Quadratic configuration interaction. A general technique for determining electron correlation energies

John A. Pople, Martin Head-Gordon, and Krishnan Raghavachari

Citation: J. Chem. Phys. 87, 5968 (1987); doi: 10.1063/1.453520

View online: http://dx.doi.org/10.1063/1.453520

View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v87/i10

Published by the American Institute of Physics.
Quadratic configuration interaction. A general technique for determining electron correlation energies

John A. Pople and Martin Head-Gordon
Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213
Krishnan Raghavachari
AT&T Bell Laboratories, Murray Hill, New Jersey 07974

(Received 1 June 1987; accepted 14 August 1987)

A general procedure is introduced for calculation of the electron correlation energy, starting from a single Hartree-Fock determinant. The normal equations of (linear) configuration interaction theory are modified by introducing new terms which are quadratic in the configuration coefficients and which ensure size consistency in the resulting total energy. When used in the truncated configuration space of single and double substitutions, the method, termed QCISD, leads to a tractable set of quadratic equations. The relation of this method to coupled-cluster (CCSD) theory is discussed. A simplified method of adding corrections for triple substitutions is outlined, leading to a method termed QCISD(T). Both of these new procedures are tested (and compared with other procedures) by application to some small systems for which full configuration interaction results are available.

I. INTRODUCTION

There is extensive literature on theoretical methods for the electron correlation correction to single-configuration, Hartree-Fock wave functions and energies. Even with orbital expansion in a finite basis, a full solution (full configuration interaction or FCI) is only possible for small systems. Other approximate methods are subject to various advantages and disadvantages which have been debated at some length. Desirable features of a theoretical model chemistry including electron correlation are:

1. It should be well defined, leading to a unique energy for any nuclear configuration and a continuous potential surface.
2. It should be size consistent, so that when applied to an ensemble of isolated molecules, calculated energies should be additive.
3. It should be exact (equivalent to FCI) when applied to a two-electron system.
4. It should be efficient, so that application is possible for large basis sets.
5. It should be accurate enough to give an adequate approximation to the FCI result.
6. It should be variational, so that the computed energy is an upper bound to the correct energy.

No current method satisfies all of these criteria; most compromise by introducing approximations with varying degrees of success.

Of current methods, limited (or truncated) configuration interaction has been used extensively. If all double-substitution configurations from a Hartree-Fock reference are included, the method is called CASSCF. If single substitutions are added, it becomes CASSCF. These methods are variational (condition 6) but not size consistent (condition 2). In applications to a range of molecules, size consistency is generally regarded as being more important than provision of an energy upper bound, so CASSCF results are usually modified by a simple correction, introduced by Langhoff and Davidson, which makes the energies approximately size consistent but no longer variational. However, this Davidson correction fails to give correct answers for a two-electron system (condition 3). The CASSCF method also completely omits effects of triple substitutions, which are known to be important.

A second range of approximate correlation methods is based on many-body perturbation theory, introduced originally by Möller and Plesset (MP). These treat the correlation part of the Hamiltonian as a perturbation on the Hartree-Fock part and truncate the energy expansion at some order. The theory is fully practical to fourth order (MP4), at which level it incorporates the effects of single, double, triple, and quadruple substitutions. The method is size consistent at any order but not variational. Its principal deficiency is that the MP series sometimes converges slowly, so that a truncated expansion may not adequately satisfy condition (5).

The third set of methods utilize coupled cluster theory, introduced into quantum chemistry by Cizek. These methods introduce substituted configurations in a multiplicative (exponential) rather than an additive manner. The configuration coefficients are found by requiring that the projection of the Schrödinger function \( (H - E)\Psi \) onto various determinants is zero. Such procedures are size consistent but not variational. There are indications that they are still effective under circumstances where perturbation theory convergence is slow. When limited to double substitutions, the coupled cluster method is usually denoted by CASSCF. It was originally implemented partially by Taylor, Bacsay, Hush, and Hurley, and fully by Pople, Krishnan, Schlegel, and Binkley and by Bartlett and Purvis. The latter authors later extended their work to include single and double substitutions (CCSD). Like CASSCF, the CASSCF and CCSD methods take inadequate account of triple substitutions, although recently Bartlett and co-workers have implemented the coupled cluster theory including triples (CCSDT) and have also proposed some approximate treatments. In other applications, hybrid methods are used in which triple substitutions are added to CASSCF or CCSD by perturbation methods.
In this paper, we introduce a new series of methods which are, in a sense, intermediate between configuration interaction and coupled cluster theory. The idea is to modify the configuration interaction equations in a simple manner, so as to restore size consistency [condition (2)] at the cost of loss of variational character [condition (6)]. This requires additional terms which are quadratic in the general substitution operators and leads to a treatment that is somewhat simpler than coupled cluster methods, at least for higher levels of substitution. The theory and its implementation for single and double substitutions is outlined in Sec. II. This is followed in Sec. III by an approximate treatment of the triple substitution contributions. Finally, Sec. IV presents some comparisons between the new method and some of the earlier techniques described above.

II. FORMULATION OF QUADRATIC CONFIGURATION INTERACTION

All of the theoretical techniques discussed here begin with the Hartree–Fock single-determinant wave function

$$\Psi_0 = \langle n! \rangle^{-1/2} \text{det}\{\chi_1, \cdots, \chi_n\}, \tag{2.1}$$

where \(\chi_i (i = 1, \cdots, n)\) are occupied spin orbitals. We will consider only the spin-unrestricted (UHF) form of (2.1), for which each \(\chi_i\) is an eigenfunction of the one-electron Fock operator \(F\) with eigenvalue \(\varepsilon_i\). The remaining eigenfunctions and eigenvalues of \(F\) (actual spin orbitals and energies) are \(\chi_a, \varepsilon_a\) \((a = n + 1, n + 2, \ldots, N)\), where \(N\) is the dimension of the basis. Following a widespread convention, we will use symbols \(ij, ik, \ldots\) for occupied and \(a, b, c, \ldots\) for virtual spin orbitals without further identification.

Other determinantal wave functions are derived from \(\Psi_0\) by substitution of occupied spin orbitals by virtual spin orbitals. Following general practice, we define complete single, double, triple, ... substitution operators.

$$T_1 = \sum_{\alpha} a^\alpha \frac{\tilde{\alpha}}{\alpha},$$

$$T_2 = \frac{1}{4} \sum_{\alpha \beta} a^\alpha_{\beta} \frac{\tilde{\beta}}{\gamma} a^\gamma_{\delta}, \tag{2.2}$$

$$T_3 = \frac{1}{36} \sum_{\alpha \beta \gamma \delta} a^\alpha_{\beta} \frac{\tilde{\beta}}{\gamma} a^\gamma_{\delta} \frac{\tilde{\delta}}{\epsilon},$$

where \(\tilde{\alpha} \frac{\tilde{\beta}}{\gamma} \cdots\) are elementary substitution operators and the \(a\) arrays involve coefficients to be determined. The effect of a substitution operator (e.g., \(\tilde{a}^{\alpha}_{\beta}\)) on a determinant where the occupied orbitals \((i\) or \(j)\) are not there or where the virtual orbitals \((a\) or \(b)\) are already there is zero. The operators act only on antisymmetric determinants and hence the \(a\) arrays will be taken to be antisymmetric in both the occupied and virtual suffixes. Thus \(T_2 \Psi_0\), for example, is a general linear combination of all determinants obtained by double substitution in \(\Psi_0\).

Various types of antisymmetric wave functions can be obtained by applying various functions of the \(T\) operators to \(\Psi_0\).

$$\Psi = f(T_0, T_1, \ldots) \Psi_0 \tag{2.3}$$

and then determining the unknown \(a\) coefficients by appropriate projection of the Schrödinger function \((H - E) \Psi\), where \(H\) is the full Hamiltonian, and \(E\) the total energy. Thus

$$\langle \Psi_0 | H - E | \Psi_0 \rangle = 0, \tag{2.4}$$

$$\langle \Psi_0 | H - E | \Psi_0 \rangle = 0, \tag{2.5}$$

$$\langle \Psi_0 | H - E | \Psi_0 \rangle = 0, \tag{2.6}$$

where \(\Psi_0\) is the singly substituted determinant \(T_0 \Psi_0\) and \(T_2 \Psi_0\) and so forth. The number of such projection equations can be chosen to be equal to the number of unknowns. For example, if only single and double substitution operators are used (SD methods), Eqs. (2.4) to (2.6) suffice to determine the unknowns \(a_i, a^a_\beta, \ldots\).

The simplest choice for the configuration function \(f\) is linear, leading to the configuration interaction wave functions

$$\Psi_{\text{CID}} = (1 + T_2) \Psi_0, \tag{2.7}$$

$$\Psi_{\text{CISD}} = (1 + T_1 + T_2) \Psi_0. \tag{2.8}$$

It should be noted that we use the intermediate normalization convention \(\langle \Psi_0 | \Psi \rangle = 1\) throughout. The projection equations (2.4) to (2.6) are then identical with those obtained by minimization of the expectation energy

$$E = \langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle,$$

so that the CID and CISD energies are variational upper bounds to the exact (FCI) result.

If we define

$$H = F + V,$$

$$E_{\text{HF}} = \langle \Psi_0 | H | \Psi_0 \rangle,$$

$$E = E_{\text{HF}} + E_{\text{correlation}},$$

$$\bar{H} = H - E_{\text{HF}},$$

$$V = V - \langle \Psi_0 | V | \Psi_0 \rangle,$$

where \(F\) is the Fock Hamiltonian, then the CISD projection equations can be written

$$\langle \Psi_0 | \bar{H} | T_2 \Psi_0 \rangle = E_{\text{correlation}}, \tag{2.14}$$

$$\langle \Psi_0 | \bar{H} | (T_1 + T_2) \Psi_0 \rangle = a^a_\beta E_{\text{correlation}}, \tag{2.15}$$

$$\langle \Psi_0 | \bar{H} | (1 + T_1 + T_2) \Psi_0 \rangle = a^a_\beta a^a_\delta E_{\text{correlation}}. \tag{2.16}$$

In deriving these equations from the projection conditions (2.4)–(2.6), we have replaced \(\langle \Psi_0^* | H | \Psi_0 \rangle\) by zero for all \(i, a\). This is because \(\Psi_0\) is the optimized Hartree–Fock function (Brillouin’s theorem).

The CISD energy emerging from (2.14) to (2.16) is not size consistent. This is because the right-hand sides of Eqs. (2.15) and (2.16) are quadratic in the \(a\) vectors, while the left-hand sides are linear. Consider, for example, two infinitely separated molecules \(\chi\) and \(\chi\). If \(\tilde{\gamma} \rightarrow a b\) is a double substitution on \(\chi\), the corresponding equations obtained (2.16) becomes

$$\langle \Psi_0 | \bar{H} | (1 + T_1 + T_2) \Psi_0 \rangle = a^a_\beta a^a_\delta E_{\text{correlation}}(\chi) + E_{\text{correlation}}(\chi). \tag{2.17}$$

The term \(a^a_\beta a^a_\delta E_{\text{correlation}}(\chi)\) on the right here implies that the value of an \(a\) coefficient on \(X\) depends on the nature of \(Y\), since no linear term of the left can cancel this. Hence, size-inconsistent results are obtained.

The simplest method of correcting this type of error is to add terms on the left of Eqs. (2.15) and (2.16) which are
quadratic in $T_1$, $T_2$ and which eliminate the term on the right involving $E_{\text{correlation}}$. The equations proposed are

$$\langle \Psi_0 | H | T_2 \Psi_0 \rangle = E_{\text{correlation}},$$

$$\langle \Psi_i^0 | H | (T_1 + T_2 + T_3) \Psi_0 \rangle = a_i^0 E_{\text{correlation}},$$

$$\langle \Psi_{ij}^0 | H | (1 + T_1 + T_2 + T_3 + \frac{1}{2} T_4^2) \Psi_0 \rangle = a_{ij}^0 E_{\text{correlation}}.$$  

(2.20)

These equations lead to size-consistent energies, as will be proved after further development.

For a two-electron system, Eqs. (2.18)-(2.20) are identical with the CISD equations (2.14)-(2.16). This is because the functions $T_1, T_2, \Psi_0$ and $\frac{1}{2} T_4^2 \Psi_0$ are sums of triple and quadruple substitutions, respectively, which vanish if only two electrons are present. Thus the proposed method then gives correct (FCI) answers, satisfying condition (3) mentioned in the Introduction. We shall demonstrate that this and related methods are also efficient and accurate [conditions (4) and (5)] later in this article.

The general concept of quadratic configuration interaction can also be applied to higher levels of substitution. If all substitutions up to $m$-fold are included, the projection equations for all levels up to $(m - 2)$ are retained in the linear form, but quadratic terms $T_{m-1} T_1$ and $T_n T_2$ are added to the equations for levels $(m - 1)$ and $m$. An additional factor 1/2 must be included if the product $T_2 T_3$ appears. Thus the QCISDT equations are

$$\langle \Psi_0 | H | T_2 \Psi_0 \rangle = E_{\text{correlation}},$$

$$\langle \Psi_i^0 | H | (T_1 + T_2 + T_3) \Psi_0 \rangle = a_i^0 E_{\text{correlation}},$$

$$\langle \Psi_{ij}^0 | H | (1 + T_1 + T_2 + T_3 + \frac{1}{2} T_4^2) \Psi_0 \rangle = a_{ij}^0 E_{\text{correlation}},$$

$$\langle \Psi_{ijk}^0 | H | (T_1 + T_2 + T_3 + T_2/T_3) \Psi_0 \rangle = a_{ijk}^0 E_{\text{correlation}}.$$  

(2.23)

The series of methods will terminate with full configuration interaction if $m = n$. This is because the operators $T_{n-1} T_1$ and $T_n T_2$ give zero when applied to $\Psi_0$. At all levels, the quadratic terms introduced serve to cancel the right-hand side for the two highest levels of substitution.

It should be noted that no $T_1 T_2 T_3$ term appears on the left of the doubles equations (2.20) for the QCISD method. This means that the equations do not constitute a full set of projection equations (2.4)-(2.6) with a well-defined wave function $\Psi$. However, this is not necessarily a disadvantage. We note that the defining equations lead to a well-defined energy $E$, so that specification of a full electronic wave function may not be necessary. We also point out that many electron correlation methods such as CEPA,17 CPF,18 CI including Davidson correction, perturbation techniques4 such as MP2, MP3, or MP4, or several approximate CCSDT methods14 also do not have an associated unique wave function.

Before proceeding with further development of the QCISD equations, we will discuss the relation between this method and coupled cluster theory. In the space of single and double substitutions, coupled cluster theory (CCSD) uses an exponential type of wave function.

$$\Psi_{\text{CCSD}} = \exp \left( T_1 + T_2 \right) \Psi_0.$$  

(2.25)

If we define $T_2 = T_2' + \frac{1}{2} T_4'$ and expand the exponential in the projection equations (2.4)-(2.6), we obtain the CCSD equations in the form

$$\langle \Psi_0 | H | T_2 \Psi_0 \rangle = E_{\text{correlation}},$$

$$\langle \Psi_i^0 | H | (T_1 + T_2 + T_3) \Psi_0 \rangle = a_i^0 E_{\text{correlation}},$$

$$\langle \Psi_{ij}^0 | H | (1 + T_1 + T_2 + T_3 - \frac{1}{2} T_4' + \frac{1}{2} T_4'^2) \Psi_0 \rangle = a_{ij}^0 E_{\text{correlation}}.$$  

(2.27)

Equations equivalent to these were first implemented in the CCSD program of Purvis and Bartlett.13 They demonstrated that size consistency followed from the exponential form of the wave function. However, the equations are quartic and considerably more complex than the QCISD equations (2.18)-(2.20).

If single substitutions are omitted, $T_1$ must be removed from the CCSD equations, leading to the familiar CCD equations. The same CCD equations also follow by omitting $T_1$ in Eqs. (2.18)-(2.20). Thus CCD and QCID are identical procedures.

The QCISD equations can be obtained by leaving out certain terms in Eqs. (2.26)-(2.28), so that QCISD could be regarded as a simplified approximate form of CCSD. However, we prefer to treat the quadratic configuration procedures as an independent approach to the problem of electron correlation.

We now return to the QCISD equations (2.18)-(2.20). All matrix elements are easily evaluated in terms of two-electron integrals, using techniques described elsewhere.19,20 The final form of the equations can be written as a simple modification of the CCD equations presented some years ago,11

$$E_{\text{correlation}} = \frac{1}{4} \sum_{pqrs} \langle ij | ab \rangle a^b_a,$$

$$\Delta^a_a + u^a + v^a = 0,$$

$$\langle ab | \langle ij \rangle + \Delta^b_a a^b_a + u^b_a + v^b_a = 0.$$  

(2.29)

where $(ab \parallel ij)$ are the usual antisymmetrized two-electron integrals

$$(pq || rs) = \int \chi_p^* (1) \chi_q^* (2) (1/r_{12}) \left[ \chi_r (1) \chi_s (2) - \chi_r (2) \chi_s (1) \right] dr_1 dr_2$$  

(2.32)

and

$$\Delta^a_a = \epsilon_a - \epsilon_i,$$

$$\Delta^b_a = \epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j,$$

$$u^a = - \frac{1}{2} \sum_{\alpha} \langle ja | ib \rangle a^b_a - \frac{1}{2} \sum_{\beta} \langle ja | bc \rangle a^c_b,$$

$$- \frac{1}{2} \sum_{\kappa} \langle jk | ib \rangle a^b_k,$$  

(2.35)
\[ u_{ij}^{ab} = \sum_{c} \left[ (ab || cj) a_{i}^{c} - (ab || ci) a_{j}^{c} \right] + \sum_{k} \left[ - (kb || ij) a_{k}^{b} + (ka || ij) a_{k}^{a} \right] + \frac{1}{2} \sum_{cd} (ab || cd) a_{d}^{ab} + \frac{1}{2} \sum_{kl} (kl || ij) a_{k}^{a} a_{l}^{b}, \]

\[ - \sum_{kc} \left[ (kb || j) a_{k}^{c} a_{c}^{b} + (ka || j) a_{k}^{a} a_{c}^{b} \right] + (kb || jk) a_{k}^{c} a_{j}^{b} + (ka || jk) a_{k}^{a} a_{j}^{b}, \]

(2.36)

\[ v_{i}^{a} = \frac{1}{2} \sum_{jkbc} \left[ (jk || bc) a_{j}^{b} a_{k}^{a} + a_{j}^{a} a_{k}^{b} + 2 a_{j}^{a} a_{k}^{b} \right], \]

(2.37)

\[ v_{0}^{b} = \frac{1}{4} \sum_{abcd} \left[ (kl || cd) a_{d}^{ab} a_{k}^{b} - 2 a_{d}^{ab} a_{k}^{b} + a_{d}^{ab} a_{k}^{b} \right] + 2 a_{d}^{ab} a_{k}^{b} + 2 a_{d}^{ab} a_{k}^{b} \]

(2.38)

Equations (2.30) and (2.31) may be solved iteratively in the form

\[ a_{i}^{a} = - (\Delta_{i}^{a})^{-1} \left[ u_{i}^{a} + v_{i}^{a} \right], \]

(2.39)

\[ a_{0}^{ab} = - (\Delta_{0}^{ab})^{-1} \left[ (ab || ij) + u_{0}^{ab} + v_{0}^{ab} \right], \]

(2.40)

using \( u \) and \( v \) arrays from the previous iteration, starting with \( u = v = 0 \). The computer program to do this is very close to that previously developed for CISD and CCD. 11,19

The only additional work is evaluation of the singles \( v_{i}^{a} \) array by Eq. (2.37). This can be done using intermediate arrays already computed for \( v_{0}^{b} \). Using a notation explained in Ref. 11, \( v_{i}^{a} \) is obtained from

\[ v_{i}^{a} = \frac{1}{2} \sum_{X} \left[ (a || X)_{i} b_{a}^{b} - \sum_{Y} (j || Y)_{i} a_{j}^{b} \right], \]

(2.41)

The most time-consuming term here is the third, which involves \(-n^{2}N^{2}\) multiplications. This is quite negligible compared with the \(-n^{2}N^{4}\) steps needed in the evaluation of \( u_{0}^{ab} \) by Eq. (2.36). The other terms involving single substitutions, e.g., the evaluation of \( v_{i}^{a} \) by Eq. (2.35), are also smaller, requiring only \( n^{3}N^{3} \) computational steps. Hence, the computational effort per cycle for QCISD is close to that required for QCID or CCD.

We now give a proof of size consistency for the QCISD method. Suppose that we are dealing with two separated molecules \( X \) and \( Y \). Then all substitution operators \( \tau_{i}^{a}(X) \) and \( \tau_{j}^{b}(Y) \) which transfer electrons from one molecule to another (leading to ionic structures such as \( X^{+} + Y^{-} \)) can be ignored, since they lead to wave functions which will have zero matrix elements with \( \Psi_{0} \) for any operator. We are left with:

(1) Singles operators localized on one molecule, \( \tau_{i}^{a}(X) \) or \( \tau_{j}^{b}(Y) \).

(2) Doubles operators localized on one molecule, \( \tau_{ij}^{ab}(XX) \) or \( \tau_{ij}^{ab}(YY) \).

(3) Doubles operators split between the molecules, \( \tau_{ij}^{ab}(XY) \) with \( (ia) \) on \( X \) and \( (jb) \) on \( Y \).

Each of these operators will be associated with an \( a \) coefficient.

The first iteration of Eqs. (2.39) and (2.40), with \( u = v = 0 \), leads to zero values for \( a_{i}^{a}(X) \) and \( a_{0}^{ab}(XY) \), since \((ab || ij)\) is nonzero only if \( i, j, a, b \) are on the same center. For later iterations, the \( u \) and \( v \) arrays must be determined from Eqs. (2.35) to (2.38). This gives nonzero values for \( a_{i}^{a}(X) \) and \( a_{0}^{ab}(XY) \). However, \( a_{0}^{ab}(XY) \) remains zero. This is because if \( a_{0}^{ab}(XY) \) is zero in the previous cycle, \( u_{0}^{ab}(XY) \) and \( v_{0}^{ab}(XY) \) are both zero from Eqs. (2.36) and (2.38). Examination of individual terms shows this result. For example, consider the term \(- \sum_{k} (kb || jk) a_{k}^{c} a_{j}^{b} \) in the expression for \( u_{0}^{ab} \). If \( (ia) \) is on \( X \) and \( (jb) \) is on \( Y \), then the integral is nonvanishing only if the suffixes \( (kc) \) are both on \( Y \). But \( a_{k}^{c}(Y) \) is zero from the previous cycle. Similar considerations apply to all other terms, so the \( a_{0}^{ab}(XY) \) remains zero at every iteration by induction. As a result, the computation of the \( u \) and \( v \) arrays at each iteration splits into separate computations for the isolated molecules \( X \) and \( Y \). This leads to a final \( T_{2} \), which is a sum \( T_{2}(X) + T_{2}(Y) \) leading to energy additivity from Eq. (2.29).

III. SIMPLIFIED METHOD FOR TRIPLE SUBSTITUTIONS

The QCISD theory introduced in the previous section still takes inadequate account of the contributions of triple substitutions. These are known to be important and could be incorporated by proceeding to the next level, QCISD. However, implementation of QCISDT for large systems is impractical at present because of handling the large number of triple–triple Hamiltonian matrix elements. Some simpler procedure is therefore desirable.

A useful approximation for triple substitutions is to treat them as a perturbation on a solution already obtained at the singles–doubles level. This means that the Hamiltonian matrix elements \( V_{tt} \) between \( \Psi_{t} \) (\( s = \) single or double) and \( \Psi_{t} \) (\( t = \) triple) are supposed small. At the same time, the energy denominator that occurs in a perturbation type approach is taken from the eigenvalues of the Fock Hamiltonian matrix or the unperturbed matrix of Møller–Plesset theory. If an initial CISD wave function is, for example,

\[ \Psi = \Psi_{0} + \sum_{T} a_{T} \Psi_{t}, \]

then the leading term in the energy lowering due to triples would be

\[ \Delta E_{T} = \sum_{t} \left( E_{0} - E_{t} \right)^{-1} |\langle \Psi_{t} | V_{tt} | \Psi_{t} \rangle|^{2} \]

\[ = \sum_{t} \sum_{T} \left( E_{0} - E_{t} \right)^{-1} a_{T} V_{tt} a_{T}. \]

(3.2)

Here \( E_{t} - E_{0} \) is the triples excitation energy using the Fock Hamiltonian. Such a procedure is close to that introduced by one of us16 to compute the contribution of both singles and triples, starting from a converged CCD (or QCID) wave function. A similar procedure has also been used by Urban, Noga, Cole, and Bartlett14 to add a triples contribution to a converged CCSD calculation.

Application of Eq. (3.2) to get a triples correction to QCISD cannot be fully justified, since some contribution from triples is already included in the QCISD energy, due to
the presence of the $T_1T_2$ term in Eq. (2.19). In fact, as demonstrated in the Appendix, the part of (3.2) with $s$ and $u$ being single substitutions contains the leading contribution of this $T_1T_2$ quadratic term. We therefore chose to omit the sum $\sum_{\alpha}^\prime$ in (3.2) and propose the remaining triples correction

$$\Delta E_T(QCISD) = \left[ 2 \sum_{\alpha}^\prime + \sum_{\alpha}^\prime \sum_{\beta}^\prime \sum_{\gamma}^\prime (E_{\alpha} - E_i)^{-1} a_{\alpha} V_{\alpha} V_{\alpha} a_{\beta} - \sum_{\alpha}^\prime \sum_{\beta}^\prime \sum_{\gamma}^\prime (E_{\alpha} - E_i)^{-1} a_{\alpha} V_{\alpha} V_{\alpha} a_{\beta} \right],$$

(3.3)

where $u$-array elements are

$$\bar{u}_{\alpha\beta} = a_{\alpha}^*(jk||bc) + a_{\alpha}^*(jk||ca) + a_{\alpha}^*(jk||ab) + a_{\alpha}^*(ki||bc) + a_{\alpha}^*(ki||ca) + a_{\alpha}^*(ki||ab),$$

$$\bar{u}_{\alpha\beta} = \sum_{e} \left[ a_{\alpha}^m(bc||ek) + a_{\alpha}^m(ca||ek) + a_{\alpha}^m(ab||ek) + a_{\alpha}^m(bc||ej) + a_{\alpha}^m(ca||ej) + a_{\alpha}^m(ab||ej) + \sum_{m} \left( a_{m}^{ab}(cm||jk) + a_{m}^{ac}(am||jk) + a_{m}^{ac}(cm||kj) + a_{m}^{ac}(am||kj) + a_{m}^{bc}(bm||ki) + a_{m}^{bc}(cm||ij) + a_{m}^{bc}(am||ij) + a_{m}^{bc}(bm||ij) \right) \right].$$

(3.5)

This procedure is closely related to the treatment of triple substitutions in full fourth-order Möller–Plesset theory (MP4SDTQ). The MP4 result follows if the double coefficients $\bar{u}_{\alpha}^{ab}$ in Eq. (3.5) are replaced by their first-order values, $- (\Delta E_{\alpha})^{-1} (ab||ij)$. The term resulting from Eq. (3.5) is closely related to a term in fifth-order perturbation theory which results from the interaction of single and triple substitutions.

The method in which the triples correction $\Delta E_T$ from Eq. (3.3) is added to the QCISD energy is denoted by QCISD(T), as it represents a compromise in which triples are only treated in a partial manner. A computer program to evaluate $\Delta E_T$ is close to that used in MP4 theory, which already contains code to generate $\bar{u}_{\alpha\beta}^{ab}$ from Eq. (3.6). This step requires $-n^3N^4$ multiplications and is the only part of the computation proportional to the seventh power of the size of the system. The additional terms $\bar{u}_{\alpha\beta}^{ab}$ from Eq. (3.5) require only $-n^2N^3$ steps and form a relatively insignificant part of the total computation. Since the preceding QCISD calculation itself requires a number of iterations, each involving $-n^2N^4$ steps, the addition of the triples correction to give the QCISD(T) energy is usually possible with moderate extra expense.

It remains to confirm that the triples correction (3.4) is still size consistent. This follows by arguments similar to those presented in the previous section. For a compound $XY$ system ($X$, $Y$ infinitely separated), the $\bar{u}_{\alpha\beta}^{ab}$ elements may be divided into $XXX$, $XXY$, $XYX$, and $YYY$ classes. In fact the $XXX$ and $XXY$ terms are zero. To show this, suppose that $(ijab)$ are located on $X$ and $(k)$ on $Y$. Then consider each term in Eq. (3.6) separately. In the first term, $a_{\alpha}^{ab}$ will only be nonzero if $(e)$ belongs to $X$, but then the integral $(bc||ek)$ will be zero since it is spread over $X$ and $Y$. In the sixth term, the integral $(ab||ej)$ will be nonzero if $(e)$ belongs to $X$, but then $a_{\alpha}^{ab}$ will be zero, since it is of the type $a(XY)$ which was shown to vanish from the QCISD equations. A similar argument can be constructed for all 18 terms in Eq. (3.6). It follows that the triples energy correction (3.4) is just the sum of $XXX$ and $YYY$ parts, so that size consistency is retained.

IV. PRELIMINARY APPLICATIONS

In this section, we report the application of the QCISD and QCISD(T) methods to a set of atoms and molecules for which Bauschlicher and co-workers have recently reported full configuration interaction (FCI) results. The objective is to test whether the new methods are as successful or more successful then previous ones in coming close to the correct result (within the space defined by a particular basis set). The results test the applicability of these theoretical methods to chemically interesting problems such as determination of the electron affinity of F, the singlet–triplet separation in CH$_2$, and excited electronic states of NH$_3$.

The energies from 24 such QCISD and QCISD(T) calculations are given in Table I, together with the FCI energies from Ref. 1. In Table II, we give the errors, that is the differences $E$(approximate) $- E$(FCI), together with corresponding results for other approximate methods. These other methods are MP4 (full MP4SDTQ), reported by Cole and Bartlett,9 CISD with the Davidson correction from Bauschlicher et al.,1 and CCSD and CCSDT-1, also due to Cole and Bartlett. At the end of the table, we give the mean absolute errors for each method, using those calculations for which the data is available. Since some of the calculations refer to molecules with stretched bonds, which are sometimes subject to large errors, we also list mean absolute er-

\[ \Delta E_T = -\frac{1}{36} \sum_{ijk} (\Delta E_{ij}^{abc})^{-1} (2\bar{u}_{ijk}^{abc} + \bar{u}_{ijkl}^{abc}). \]
TABLE I. Total energies (hartrees).

<table>
<thead>
<tr>
<th>Molecule/ (state, geom.*)</th>
<th>Procedure* / basis*</th>
<th>QCISD</th>
<th>QCISD(T)</th>
<th>FCI*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>R/4s2p1d</td>
<td>-128.705754</td>
<td>-128.702617</td>
<td>-128.702462</td>
</tr>
<tr>
<td>F</td>
<td>R/6s4p1d</td>
<td>-128.765378</td>
<td>-128.767929</td>
<td>-128.767889</td>
</tr>
<tr>
<td>U/4s3p1d</td>
<td>-99.545157</td>
<td>-99.546631</td>
<td>-99.546620</td>
<td></td>
</tr>
<tr>
<td>F-</td>
<td>R/4s3p1d</td>
<td>-99.651109</td>
<td>-99.653329</td>
<td>-99.653341</td>
</tr>
<tr>
<td>R/4s3p2d</td>
<td>-99.674143</td>
<td>-99.677591</td>
<td>-99.677676</td>
<td></td>
</tr>
<tr>
<td>R/5s4p2d</td>
<td>-99.700755</td>
<td>-99.707057</td>
<td>-99.706690</td>
<td></td>
</tr>
<tr>
<td>HF/R_e</td>
<td>R/DZP</td>
<td>-100.248402</td>
<td>-100.250727</td>
<td>-100.250969</td>
</tr>
<tr>
<td>1.5R_e</td>
<td>R/DZP</td>
<td>-100.156257</td>
<td>-100.159729</td>
<td>-100.160393</td>
</tr>
<tr>
<td>2R_e</td>
<td>R/DZP</td>
<td>-100.072713</td>
<td>-100.079725</td>
<td>-100.081108</td>
</tr>
<tr>
<td>H_2O/R_e</td>
<td>R/DZP</td>
<td>-76.252745</td>
<td>-76.256007</td>
<td>-76.256624</td>
</tr>
<tr>
<td>1.5R_e</td>
<td>R/DZP</td>
<td>-76.062040</td>
<td>-76.069391</td>
<td>-76.071405</td>
</tr>
<tr>
<td>2R_e</td>
<td>R/DZP</td>
<td>-75.930888</td>
<td>-75.953530</td>
<td>-75.952269</td>
</tr>
<tr>
<td>1.5R_e</td>
<td>U/DZP</td>
<td>-76.05706</td>
<td>-76.065701</td>
<td>-76.071405</td>
</tr>
<tr>
<td>2R_e</td>
<td>U/DZP</td>
<td>-75.935541</td>
<td>-75.940310</td>
<td>-75.952269</td>
</tr>
<tr>
<td>NH_2/2B_2Re</td>
<td>U/DZP</td>
<td>-55.739486</td>
<td>-55.742110</td>
<td>-55.742620</td>
</tr>
<tr>
<td>^2B_2Re,1.5R_e</td>
<td>U/DZP</td>
<td>-55.595224</td>
<td>-55.601600</td>
<td>-55.605209</td>
</tr>
<tr>
<td>^2B_2Re,2R_e</td>
<td>U/DZP</td>
<td>-55.494490</td>
<td>-55.498131</td>
<td>-55.505524</td>
</tr>
<tr>
<td>^3A_1,Re</td>
<td>U/DZP</td>
<td>-55.685854</td>
<td>-55.688270</td>
<td>-55.688762</td>
</tr>
<tr>
<td>^2A_1,1.5R_e</td>
<td>U/DZP</td>
<td>-55.510338</td>
<td>-55.515648</td>
<td>-55.517614</td>
</tr>
<tr>
<td>^2A_1,2R_e</td>
<td>U/DZP</td>
<td>-55.393207</td>
<td>-55.405182</td>
<td>-55.415133</td>
</tr>
</tbody>
</table>

* Geometries, basis sets, and FCI energies are all given in Ref. 1.
* Restricted (RHF) reference denoted by R; unrestricted (UHF) reference denoted by U.

TABLE II. Energy differences, \( E(\text{approx}) - E(\text{FCI}) \) (mhartrees).

<table>
<thead>
<tr>
<th>Molecule/ (state, geom.)</th>
<th>Procedure/ basis</th>
<th>MP4</th>
<th>CISD + Dav*</th>
<th>CCSD</th>
<th>QCISD</th>
<th>CCSDT - 1</th>
<th>QCISD(T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>R/4s2p1d</td>
<td>-0.873</td>
<td>+1.300</td>
<td>+2.142</td>
<td>+1.708</td>
<td>-0.227</td>
<td>-0.155</td>
</tr>
<tr>
<td>F</td>
<td>R/6s4p1d</td>
<td>+2.562</td>
<td>+2.030</td>
<td>+2.875</td>
<td>+2.511</td>
<td>-0.095</td>
<td>-0.040</td>
</tr>
<tr>
<td>U/4s3p1d</td>
<td>+1.830</td>
<td>+1.830</td>
<td>+1.463</td>
<td>+0.522</td>
<td>-0.011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U/4s3p2d</td>
<td>+0.265</td>
<td>+2.639</td>
<td>+2.290</td>
<td>-0.074</td>
<td>-0.033</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U/5s4p2d</td>
<td>+0.529</td>
<td>+1.428</td>
<td>+3.335</td>
<td>+3.097</td>
<td>+0.035</td>
<td>+0.047</td>
<td></td>
</tr>
<tr>
<td>F-</td>
<td>R/4s3p1d</td>
<td>-4.748</td>
<td>+2.962</td>
<td>+4.730</td>
<td>+2.232</td>
<td>-1.108</td>
<td>+0.112</td>
</tr>
<tr>
<td>R/4s3p2d</td>
<td>-5.044</td>
<td>+5.806</td>
<td>+3.533</td>
<td>-1.137</td>
<td>+0.085</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R/5s4p2d</td>
<td>-5.398</td>
<td>+5.518</td>
<td>+7.715</td>
<td>+5.935</td>
<td>-1.207</td>
<td>-0.367</td>
<td></td>
</tr>
<tr>
<td>HF/R_e</td>
<td>R/DZP</td>
<td>-1.263</td>
<td>+1.596</td>
<td>+3.007</td>
<td>+2.567</td>
<td>+0.170</td>
<td>+0.242</td>
</tr>
<tr>
<td>1.5R_e</td>
<td>R/DZP</td>
<td>+0.769</td>
<td>+1.411</td>
<td>+5.100</td>
<td>+4.136</td>
<td>+0.490</td>
<td>+0.664</td>
</tr>
<tr>
<td>2R_e</td>
<td>R/DZP</td>
<td>+4.841</td>
<td>-1.291</td>
<td>+10.181</td>
<td>+8.395</td>
<td>+0.222</td>
<td>+1.383</td>
</tr>
<tr>
<td>H_2O/R_e</td>
<td>R/DZP</td>
<td>+0.917</td>
<td>+2.075</td>
<td>+4.122</td>
<td>+3.879</td>
<td>+0.596</td>
<td>+0.617</td>
</tr>
<tr>
<td>1.5R_e</td>
<td>R/DZP</td>
<td>+5.764</td>
<td>+4.402</td>
<td>+10.158</td>
<td>+9.365</td>
<td>+1.991</td>
<td>+1.814</td>
</tr>
<tr>
<td>1.5R_e</td>
<td>U/DZP</td>
<td>+56.150</td>
<td>+13.676</td>
<td>+13.699</td>
<td>+6.252</td>
<td>+5.704</td>
<td></td>
</tr>
<tr>
<td>2R_e</td>
<td>U/DZP</td>
<td>+25.960</td>
<td>+17.385</td>
<td>+16.728</td>
<td>+10.332</td>
<td>+11.959</td>
<td></td>
</tr>
<tr>
<td>NH_2/2B_2Re</td>
<td>U/DZP</td>
<td>+1.900</td>
<td>+0.572</td>
<td>+3.212</td>
<td>+3.134</td>
<td>+0.519</td>
<td>+0.510</td>
</tr>
<tr>
<td>^2B_2Re,1.5R_e</td>
<td>U/DZP</td>
<td>+3.920</td>
<td>+1.584</td>
<td>+9.841</td>
<td>+9.985</td>
<td>+3.735</td>
<td>+3.609</td>
</tr>
<tr>
<td>^3A_1,Re</td>
<td>U/DZP</td>
<td>-1.615</td>
<td>+0.618</td>
<td>+2.992</td>
<td>+2.908</td>
<td>+0.496</td>
<td>+0.492</td>
</tr>
<tr>
<td>^2A_1,1.5R_e</td>
<td>U/DZP</td>
<td>+7.779</td>
<td>+2.403</td>
<td>+7.670</td>
<td>+7.276</td>
<td>+2.116</td>
<td>+1.966</td>
</tr>
<tr>
<td>^2A_1,2R_e</td>
<td>U/DZP</td>
<td>+41.297</td>
<td>+6.866</td>
<td>+20.168</td>
<td>+19.926</td>
<td>+9.519</td>
<td>+9.951</td>
</tr>
<tr>
<td>CH_2^+/B_1</td>
<td>U/DZP</td>
<td>-0.650</td>
<td>+2.054</td>
<td>-0.039</td>
<td>+3.522</td>
<td>+0.867</td>
<td></td>
</tr>
<tr>
<td>^1A_1</td>
<td>R/DZP</td>
<td>-0.903</td>
<td>+3.522</td>
<td>+0.867</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* All CISD calculations in this column are based on an RHF reference function.

rors for the smaller set of atoms and molecules close to equilibrium geometries.

The QCISD total energies reported in Table I all lie above the corresponding full configuration interaction results, even though the method is no longer variational. The errors in this method range from 1.7 mhartree (F atom with the smaller basis) to more than 20 mhartree (RHF-based calculation on H₂O with bonds stretched to twice the equilibrium bond length). The results for molecular states close to equilibrium geometries, however, show only modest errors of 2–4 mhartree, the mean absolute error for this subset being 2.914 mhartree. This is consistent with the observation made in the previous section that QCISD takes only limited account of triple substitutions and their contribution to lowering the total energy. The effect of triples is known to be much larger for stretched bonds, as already demonstrated by perturbation theory. ⁹

Inclusion of the triples correction at the QCISD(T) level leads to substantial improvement in the calculated total energies. For all of the 24 examples, the QCISD(T) energy is closer to the correct (FCI) answer than is the QCISD value. In some cases, the QCISD(T) energy is below the FCI result, but the largest such overshoot is only about 1.2 mhartree. Results on molecules with stretched bonds are improved but still not really satisfactory, the greatest error being 12 mhartree, (UHF-based calculation on H₂O with twice the equilibrium bond lengths). It is interesting to note that QCISD(T) performs quite well even for stretched bond lengths when an RHF reference function is used (e.g., HF and H₂O from Table II). The comparatively larger errors for the UHF-based stretched bond lengths appears to be related to the very large degree of spin contamination in such wave functions, though our preliminary results indicate that the method performs very well in cases with moderate amounts of spin contamination. It should also be noted that for open-shell cases, our calculations involve frozen UHF 1s orbitals, whereas Bauschlicher et al. use frozen RHF 1s orbitals and hence the full CI comparisons are not as definitive for open shells. For atoms and molecules close to equilibrium geometries, the QCISD(T) results are particularly good, the mean absolute error for the subset being only 0.28 mhartree.

The comparison with other methods displayed in Table II shows a number of interesting features. Fourth-order Möller–Plesset energies (MP4), which include contributions from triple substitutions, show particularly poor results for unrestricted (UMP) cases with stretched bonds. The slow convergence of the UMP series under these circumstances has been noted before. ⁶ For geometries close to equilibrium, however, MP4 energies show a mean absolute error of only 2.246 mhartree. This is consistent with the observation reported in the previous section that the full CI comparisons are not as definitive for open-shell cases but only RHF based). For those 11 examples, the mean absolute deviation for the MP4 method is 0.78 mhartree. For the same 11 examples, the mean deviation for the QCISD(T) method is 0.61 mhartree. More examples are needed for a rigorous evaluation but the performance of the QCISD(T) method is clearly excellent.

In conclusion, the quadratic configuration interaction method QCISD(T), introduced in this paper, appears to be simple, efficient, and effective. It should prove valuable in exploring correlation energies for a wide range of molecules.

ACKNOWLEDGMENTS

Part of this work was carried out while one of us (J.A.P) was a Visiting Fellow at the Australian National University, Canberra. Support from the National Science Foundation (Grant CHE-84-09405) is acknowledged. We are indebted to Professor R. J. Bartlett for some valuable comments.

APPENDIX

Here we consider the contribution to the energy that arises from the inclusion of the product term T₁ T₂ in Eq. (2.19). This term gives rise to a triple substitution coefficient which may be explicitly written as follows:

\[
a^{abc}_{ijk} = a^{abc}_{ijk} - a^{abc}_{ijk} + a^{abc}_{ijk} - a^{abc}_{ijk} - a^{abc}_{ijk} + a^{abc}_{ijk}
\]

The nine terms in Eq. (A1) can all be derived from the first term by permuting the indices i,j,k and a,b,c and using the antisymmetric nature of a^{abc}_{ijk}. Thus,

\[
a^{abc}_{ijk} = a^{abc}_{ijk} + \text{permutation terms.}
\]

In terms of a Möller–Plesset perturbation expansion, the
doubles coefficients are of first order and the singles coefficients are of second order and thus the triples coefficients formed as the product \( T_1 T_2 \) are third-order terms. Replacing the doubles coefficients by their first-order values, we get
\[
\alpha^{abc}_{ijk} = a^r_i(jk)bc (-\Delta_{ijkl}^{ab})^{-1} + \text{permutation terms.} \tag{A3}
\]
We now use the identity
\[
(\Delta_{ijkl}^{ab})^{-1} = (\Delta_{ijkl}^{abc})^{-1}[1 + (\Delta_{ijkl}^{a})/\Delta_{ijkl}^{bc}] \tag{A4}
\]
and substitute in Eq. (A3), giving
\[
\alpha^{abc}_{ijk} = [a^r_i(jk)bc (-\Delta_{ijkl}^{abc})^{-1} + \text{permutation terms}] + \text{other terms.} \tag{A5}
\]
We note that the other terms in Eq. (A5) result from the second part of Eq. (A4) and will not be considered further in this discussion. Full details of the contribution of these terms will be given in a later paper. Neglecting such terms, Eq. (A5) can now be recast using the general notation
\[
\alpha^{abc}_{ijk} = a_i = \sum_u^S (E_0 - E_i)^{-1} V_u a_u. \tag{A6}
\]
Evaluating \( \langle \Psi_i^j | H | T_1 T_2 \Psi_0 \rangle \) in Eq. (2.19) now gives
\[
u_i^j = \sum_u^T \sum_u^S (E_0 - E_i)^{-1} V_u V_u a_u. \tag{A7}
\]
\( \nu_i^j \) does not contribute directly to the correlation energy since the single substitutions do not occur in the energy expression (2.29). The contribution to the energy occurs as follows. The \( \nu_i^j \) term [which is closely related to the \( v_i^j \) in Eq. (2.37)] induces a change in the singles coefficients \( a^r_i \) by Eq. (2.30). The \( a^r_i \) in turn yields a contribution to \( \nu^q_i \) [Eq. (2.36)] and hence changes the doubles coefficients \( \alpha^{pq}_{ij} \) by Eq. (2.31). This finally gives an energy correction via Eq. (2.29). In leading order, this energy contribution may be shown to be equivalent to
\[
\delta E = \sum_u^S a_i \nu_i^j, \tag{A8}
\]
where \( a_i \) are the singles coefficients in second order. Substituting Eq. (A7) into Eq. (A8), we finally have
\[
\delta E = \sum_i^T \sum_u^S \sum_u^T (E_0 - E_i)^{-1} a_i V_u V_u a_u. \tag{A9}
\]
This is identical to the term that was neglected in Eq. (3.2). In the above discussion, the term arising from the right-hand side of Eq. (2.19) was not considered. This term can.