A doubles correction to electronic excited states
from configuration interaction in the space of single substitutions

Martin Head-Gordon [a], Rudolph J. Rico [a], Manabu Oumi [a], Timothy J. Lee [b]

[a] *Department of Chemistry, University of California, Berkeley, CA 94720, USA*
[b] *NASA Ames Research Center, Moffett Field, CA 94035, USA*

Abstract

A perturbative correction to the method of configuration interaction with single substitutions (CIS) is presented. This CIS(D) correction approximately introduces the effect of double substitutions which are absent in CIS excited states. CIS(D) is a second-order perturbation expansion of the coupled-cluster excited state method, restricted to single and double substitutions, in a series in which CIS is zeroth order, and the first-order correction vanishes. CIS(D) excitation energies are size consistent and the calculational complexity scales with the fifth power of molecular size, akin to second-order Møller–Plesset theory for the ground state. Calculations on singlet excited states of ethylene, formaldehyde, acetaldehyde, butadiene and benzene show that CIS(D) is a uniform improvement over CIS. CIS(D) appears to be a promising method for examining excited states of large molecules, where more accurate methods are not feasible.

1. Introduction

Configuration interaction in the space of all single substitutions (CIS) is perhaps the simplest molecular orbital theory of excited states with reasonably general validity [1–5]. While CIS (or the Tamm–Dancoff approximation in nuclear physics [6]) has been known for decades, efficient direct algorithms have only recently been developed [5], and the method is now becoming widely used. There is some debate over the extent of applicability of CIS [5,7], since it is not unusual to find nonsystematic errors of 0.5 to 1.0 eV or more in vertical excitation energies for one-electron transitions [5,7–11]. Often this level of accuracy is an "adequate zeroth-order treatment for many of the excited states of molecules" [5], and it is generally attained for excited states that are primarily one-electron transitions from a single reference ground state. For excited states with dominant double excitation character, CIS is, of course, an inappropriate theoretical method.

Much more accurate multireference configuration interaction (MR-CI) methods [12–16] are often a preferred alternative for excited states of small molecules, although another promising class of theories are excited state coupled cluster methods in the space of single and double substitutions [17–21]. CIS represents an excited state treatment which is roughly analogous to the Hartree–Fock method for ground states, since it is inexpensive to apply but yields only qualitative accuracy. As such, the question immediately arises as to whether there are perturbative corrections to CIS which might include the leading order dynamical correlation effects in a way analogous to second-order Møller–Plesset (MP2) theory [22–24] for ground states. Such a method might correct some of the quantitative deficiencies of CIS while still permitting application to reasonably large molecular

systems, although excited states that are double excitations from the ground state could still not be appropriately handled. We note that considerable progress has recently been made on the related problem of obtaining perturbative corrections to MR-CI methods [25–27].

By analogy to MP2 theory, which introduces electron correlation by coupling the single reference to double substitutions [24], we anticipate that double substitutions from the CIS wavefunction will contain the main effects of correlation in the excited state. A proposal along these lines was recently made, termed the CIS-MP2 method [5]. CIS-MP2 is a correction to the total CIS energy of the excited state. By subtracting the ground state MP2 energy, excitation energies in better agreement with experiment than CIS are sometimes obtained. CIS-MP2 has been successful at improving the relative energies of valence versus Rydberg states for formaldehyde and acetaldehyde [10]. However in other cases such as ethylene [9] and butadiene [11], CIS-MP2 does not improve excitation energies. Furthermore it is substantially more complex to apply than the ground state MP2 method (the calculations scale as the sixth rather than fifth power of molecular size). Even more seriously, CIS-MP2 excitation energies are not size consistent, as we show in section 2.

The purpose of this work is to introduce an alternative doubles correction to the CIS method which we believe is preferable. We term this new method CIS(D), consistent with the notation used to denote perturbative triples corrections to ground state singles and doubles methods [28]. CIS(D) theory has two formal advantages over CIS-MP2. First and foremost, it yields excitation energies which are strictly size consistent. Second, the calculational complexity rises as only the fifth power of molecular size, akin to MP2 theory (CIS in the atomic orbital basis scales formally as the fourth power of molecular size if a fixed number of states are sought, before cutoffs). In section 2, CIS-MP2 is briefly reviewed, and the resulting excitation energies are shown to be size inconsistent. In section 3, the new equations defining CIS(D) are presented and shown to be size consistent. Finally in section 4, a series of calculations of singlet excited states are reported at the CIS, CIS-MP2, CIS(D) and CCSD levels of theory to assess

the comparative numerical performance of these methods. Our conclusions are presented in section 5.

2. CIS and the CIS-MP2 correction

By its very nature the Hartree–Fock single determinant Φ_0 is optimized so that the Hamiltonian matrix elements coupling Φ_0 to single substitutions Φ_i^a of any occupied spin orbital i to any unoccupied spin orbital a are zero [24]. This is Brillouin's theorem [29]. Therefore if we seek excited states at a level of accuracy roughly comparable to the HF ground state, we can diagonalize the singles–singles block of the full Hamiltonian. The resulting CIS eigenvalue equations are

$$\langle \Phi_i^a | \bar{H} | U_1 \Phi_0 \rangle = \omega b_i^a , \qquad (1)$$

where $\bar{H} = H - E_{HF}$ and ω is the CIS excitation energy. The operator U_1 generates all single substitutions Φ_i^a from Φ_0 with amplitudes b_i^a, giving the CIS wavefunction, Φ_{CIS},

$$\Phi_{CIS} = U_1 \Phi_0 = \sum_{ia} b_i^a \Phi_i^a .$$

Note that by contrast with U operators and b amplitudes for excited states, we use T and a for the corresponding ground state operators and amplitudes. With this notation the left-hand side of Eq. (1) becomes

$$\langle \Phi_i^a | \bar{H} | U_1 \Phi_0 \rangle = \sum_{jb} b_j^b \langle \Phi_i^a | \bar{H} | \Phi_j^b \rangle$$

$$= \Delta_i^a b_i^a + \sum_{jb} (aj\|ib) b_j^b , \qquad (2)$$

where Δ_i^a is the difference between orbital eigenvalues ϵ_a and ϵ_i, and $(pq\|rs)$ are antisymmetrized two-electron integrals in standard notation [24]. The CIS excitation energies ω are size consistent [5], meaning that excitation energies for a system of two noninteracting fragments include the excitation energies that would be obtained by considering each fragment as a separate system.

To include electron correlation in the ground state in leading order, MP2 theory is frequently employed. The only substituted determinants which directly couple to Φ_0 are doubles, Φ_{ij}^{ab} [24]. They are included by partitioning the full Hamiltonian as

$H = F + \lambda V$, where F, the unperturbed Hamiltonian, is the sum of one-electron Fock operators of which the orbitals are eigenvectors. V is the perturbation potential due to electron correlation, and λ is a formal expansion parameter, which is set to 1 after terms of different order are separated. Performing time-independent perturbation theory to second order then yields [22–24]

$$E^{\mathrm{MP2}} = \langle \Phi_0 | V | T_2 \Phi_0 \rangle$$
$$= \tfrac{1}{4} \sum_{ijab} a_{ij}^{ab} (ij\|ab) , \qquad (3)$$

where the ground state double substitution amplitudes to first order in the correlation perturbation (i.e. λ) are contained in the T_2 operator,

$$T_2 \Phi_0 = \tfrac{1}{4} \sum_{ijab} a_{ij}^{ab} \Phi_{ij}^{ab} , \qquad (4a)$$

$$a_{ij}^{ab} = - \frac{\langle \Phi_{ij}^{ab} | V | \Phi_0 \rangle}{\Delta_{ij}^{ab}} = - \frac{(ab\|ij)}{\Delta_{ij}^{ab}} . \qquad (4b)$$

The ground state MP2 energy defined in this way is size consistent (as is the HF energy of course), meaning that the energies of noninteracting fragments are additive. The evaluation of Eqs. (3) and (4) nominally scales as the fifth power of molecular size.

The CIS-MP2 correction to the CIS excited state energy is based on the presumption that similar physics holds for the excited state, with two modifications relative to MP2 theory. First, double substitutions from Φ_{CIS} ($\Phi_{\mathrm{CIS}} = U_1 \Phi_0$) gives triple substitutions from Φ_0, and second, since Brillouin's theorem does not hold we must also include single substitutions from Φ_{CIS} (i.e. double substitutions from Φ_0). The unperturbed Hamiltonian is defined as the SS block of the full Hamiltonian (with CIS eigenvectors and eigenvalues), together with the (diagonal) elements of F in the DD and TT blocks. The first-order perturbation is then the SD, ST, DD, DT and TT blocks of V. The second-order CIS-MP2 correction to a CIS state is then

$$E^{\mathrm{CIS\text{-}MP2}} = \langle \Phi_{\mathrm{CIS}} | V | U_2 \Phi_0 \rangle$$
$$+ \langle \Phi_{\mathrm{CIS}} | V | U_3 \Phi_0 \rangle , \qquad (5)$$

where the first-order perturbation theory values of the double and triple substitution amplitudes are contained in the U_2 and U_3 operators, and are:

$$b_{ij}^{ab} = - \frac{\langle \Phi_{ij}^{ab} | V | U_1 \Phi_0 \rangle}{\Delta_{ij}^{ab} - \omega} = - \frac{\bar{u}_{ij}^{ab}}{\Delta_{ij}^{ab} - \omega} , \qquad (6a)$$

$$b_{ijk}^{abc} = - \frac{\langle \Phi_{ijk}^{abc} | V | U_1 \Phi_0 \rangle}{\Delta_{ijk}^{abc} - \omega} = - \frac{\bar{u}_{ijk}^{abc}}{\Delta_{ijk}^{abc} - \omega} . \qquad (6b)$$

The state dependence of the energy correction enters through both the CIS amplitudes and the excitation energy. We note that while the above partitioning of the full Hamiltonian is mathematically logical given that CIS excited states are the unperturbed starting point, it cannot be given the same physical $F + \lambda V$ characterization as ground state Møller–Plesset theory. In other words, the unperturbed Hamiltonian is the SS block of the full Hamiltonian, which includes some correlation, and thus V is no longer strictly the correlation potential.

The spin orbital expression for CIS-MP2 may now be written as [5]

$$E^{\mathrm{CIS\text{-}MP2}} = -\tfrac{1}{4} \sum_{ijab} (\bar{u}_{ij}^{ab})^2 / (\Delta_{ij}^{ab} - \omega)$$
$$- \tfrac{1}{36} \sum_{ijkabc} (\bar{u}_{ijk}^{abc})^2 / (\Delta_{ijk}^{abc} - \omega) , \qquad (7)$$

where the $\bar{u}_{ij}^{ab}$ and $\bar{u}_{ijk}^{abc}$ arrays have been defined in the context of ground state correlation theories [28], and are

$$\bar{u}_{ij}^{ab} = \sum_c [(ab\|cj) b_i^c - (ab\|ci) b_j^c]$$
$$+ \sum_k [(ka\|ij) b_k^b - (kb\|ij) b_k^a] , \qquad (8)$$

$$\bar{u}_{ijk}^{abc} = (jk\|bc) b_i^a + (jk\|ca) b_i^b + (jk\|ab) b_i^c$$
$$+ (ki|bc) b_j^a + (ki|ca) b_j^b + (ki\|ab) b_j^c + (ij\|bc) b_k^a$$
$$+ (ij\|ca) b_k^b + (ij|ab) b_k^c . \qquad (9)$$

From the second term of Eq. (7), and Eq. (9), the scaling of CIS-MP2 with molecular size is sixth order. Finally, the CIS-MP2 correction to the excitation energy is given as

$$\omega^{\mathrm{CIS\text{-}MP2}} = E^{\mathrm{CIS\text{-}MP2}} - E^{\mathrm{MP2}} . \qquad (10)$$

Neither the correction to the total energy defined in Eq. (7), nor the correction to the excitation energy defined in Eq. (10) is size consistent. In the language of many-body theory [30], size consistency is violated because Eq. (9) consists of disconnected products which do not occur in the cluster expansion of

the exact wavefunction. The lack of size consistency is also evident from the following counter-example for two infinitely separated systems A and B, of which one, A, is excited with frequency ω in Φ_{CIS}, while the other is in its HF ground state. For size consistency, the CIS-MP2 correlation energy of the combined system must be equal to the CIS-MP2 energy of A alone, plus the ground state MP2 energy of B alone. The latter part is not exactly recovered as the CIS excitation energy of A, ω, affects the amplitudes of all excitations involving B through the denominator of Eqs. (6). Therefore the correlation energy of the fragment B which is not excited depends on the identity (and state) of the excited fragment A. Size consistency is thus not satisfied, and in the limit as $M \to \infty$, a supersystem of A and MB systems will exhibit $\omega^{\text{CIS-MP2}} \to \pm \infty$ for an excitation localized on A. Thus the performance of CIS-MP2 for the somewhat analogous case of an excitation localized on one functional group of a large molecule will be poor.

3. The CIS(D) correction

To introduce the new doubles correction to CIS, we can examine the triples term of CIS-MP2 theory, which is responsible for its failure to be size consistent, and inquire how we might modify it. To generate a triple substitution relative to the ground state requires that first a single replacement be performed to obtain a term of the CIS wavefunction, $\Phi_{\text{CIS}} = U_1 \Phi_0$. A double substitution involving orbitals which did not participate in the first substitution then follows to yield the triple replacement. Cases where the orbitals involved in the single substitution are promoted again will give rise to an overall double replacement relative to the single reference Φ_0, and this is treated in the first term of Eq. (5) or (7).

The fact that the doubles part of the triple substitution term involves pairs of electrons that are essentially "inactive" in the electronic excitation suggests a simple approximation for that term. We suggest that the operator U_3 in Eq. (5) can be replaced by a product of operators $T_2 U_1$, where we use the ground state double substitution amplitudes unmodified (i.e. T_2) since these correspond to pairs of electrons which are essentially unaffected by the excitation. This operator remains first order in correlation if the first order

ground state doubles amplitudes from Eqs. (4) are employed, and thus the modified correction remains second order. More sophisticated theories would allow for some change in the correlation of these "inactive" pairs of electrons due to the excitation, as in coupled cluster theory for excited states [17–21].

The equations defining the new doubles correction are thus a simple modification of just the second term of Eq. (5) for the CIS-MP2 correction:

$$E^{\text{CIS(D)}} = \langle \Phi_{\text{CIS}} | V | U_2 \Phi_0 \rangle$$
$$+ \langle \Phi_{\text{CIS}} | V | T_2 U_1 \Phi_0 \rangle . \tag{11}$$

The unchanged first term is given by the first term of Eq. (7) again, while the second term can be written in spin orbital notation as the sum of a disconnected and a connected term (the latter denoted by a subscript c) [30,31]:

$$\langle \phi_{\text{CIS}} | V | T_2 U_1 \Phi_0 \rangle = \langle \Phi_0 | V | T_2 \Phi_0 \rangle$$
$$+ \langle \Phi_{\text{CIS}} | V | T_2 U_1 \Phi_0 \rangle_c = E^{\text{MP2}} + \sum_{ia} b_i^a v_i^a . \tag{12}$$

The v_i^a array has been defined previously, for example in the context of QCISD theory [28]:

$$v_i^a = \tfrac{1}{2} \sum_{jkbc} (jk \| bc)(b_i^b a_{jk}^{ca} + b_j^a a_{ik}^{cb} + 2b_j^b a_{ik}^{ac}) . \tag{13}$$

It is immediately clear that the CIS(D) correction to excitation energies takes a particularly simple form, because the "disconnected" first term of Eq. (12) yields the MP2 energy, which cancels the second-order correction to the ground state. Therefore

$$\omega^{\text{CIS(D)}} = E^{\text{CIS(D)}} - E^{\text{MP2}}$$
$$= -\tfrac{1}{4} \sum_{ijab} (\bar{u}_{ij}^{ab})^2 / (\Delta_{ij}^{ab} - \omega) + \sum_{ia} b_i^a v_i^a . \tag{14}$$

It is straightforward to show that the evaluation of each of the terms of the above expression for $\omega^{\text{CIS(D)}}$ scales as the fifth power of molecular size, when ordinary canonical orbitals are used. This is one of the two key advantages of the new expression relative to the CIS-MP2 correction, which scales as the sixth power of molecular size.

Let us establish the size consistency of CIS(D). As a preliminary, we note that due to the presence of the disconnected term, which yields E^{MP2} exactly, the counter-example we used to demonstrate the size inconsistency of CIS-MP2 is now satisfied. More gen-

erally, size consistency follows from three facts concerning the expression for the excitation energy, Eq. (14). First, Eq. (14) is connected in the sense that both $\bar{u}_{ij}^{ab}$ and v_i^a are given by tensor products rather than scalar products, by contrast with the disconnected $\bar{u}_{ijk}^{abc}$ intermediate in CIS-MP2. Second, for two noninteracting systems, the size consistent CIS method gives amplitudes which are localized on one system or the other. Third, all integrals and MP2 doubles are zero, unless all indices are on one system. With these facts, it follows that the CIS(D) correction to a CIS excitation on one system will be unchanged by the presence of another infinitely separated system. This is sufficient for size consistency.

To this point, we have presented the CIS(D) method by hypothesis, and have established two desirable properties of the resulting equations. However the same equations arise from a perturbation expansion applied to the CCSD excited state equations. In addition to further justifying the form of the CIS(D) correction, this connection may be quite useful for generating new perturbative approximations in the future. The CCSD excited state equations can be written in fully connected form as follows [32]:

$$\langle \Phi_i^a | \bar{H} | (U_1 + U_2 + T_1 U_1 + T_1 U_2$$
$$+ U_1 T_2 + \tfrac{1}{2} T_1^2 U_1) \Phi_0 \rangle_c = \omega b_i^a , \tag{15a}$$

$$\langle \Phi_{ij}^{ab} | \bar{H} | (U_1 + U_2 + T_1 U_1 + T_1 U_2 + U_1 T_2 + \tfrac{1}{2} T_1^2 U_1$$
$$+ T_2 U_2 + \tfrac{1}{2} T_1^2 U_2 + T_1 T_2 U_1 + \tfrac{1}{6} T_1^3 U_1) \Phi_0 \rangle_c$$
$$= \omega b_{ij}^{ab} . \tag{15b}$$

This is an eigenvalue equation of the form $\mathbf{A}b = \omega b$ for the transition amplitudes (b vectors), which are also contained in the U operators. Note that the T and U operators in the CCSD equations contain contributions from all orders of perturbation theory, and in particular the converged U_1, U_2 and T_2 operators will contain different amplitudes to those used above (in particular the CIS U_1, MP2, T_2, and CIS-MP2 U_2 operators) in our presentation of the CIS(D) method.

We seek a second-order approximation to the CCSD eigenvalue equation, beginning from a CIS eigenvector, $b^{(0)}$, and transition energy $\omega^{(0)}$. The zeroth-order response matrix, $\mathbf{A}^{(0)}$ is the same as in CIS-MP2 theory: the SS block of H, together with the (diagonal) DD block due to F. The first-order matrix, $\mathbf{A}^{(1)}$ contains the SD, DS and DD blocks of V. There-

fore $\omega^{(1)}$ ($b^{(0)\dagger} \mathbf{A}^{(1)} b^{(0)}$) is zero, and the first-order eigenvector correction, $b^{(1)}$, is given by Eq. (6a). Since connected triple substitutions do not appear in the CCSD excited state equations, there is no term like Eq. (6b). Due to the presence of T_1 and T_2 products in the response matrix $\mathbf{A}$ defined by Eqs. (15), there are also second-order and higher terms in the perturbation expansion of $\mathbf{A}$. $\mathbf{A}^{(2)}$ consists of terms that are linear in T_2, and zeroth order in T_1, since T_2 has a first-order expansion (the MP1 wavefunction) and V contributes the other power of λ. T_1 terms do not contribute, as they first enter ground state perturbation theory in the second-order wavefunction due to Brillouin's theorem. The second-order contribution to the excitation energy follows as

$$\omega^{(2)} = b^{(0)\dagger} \mathbf{A}^{(1)} b^{(1)} + b^{(0)\dagger} \mathbf{A}^{(2)} b^{(0)} , \tag{16}$$

which when evaluated with the definitions given above yields the two terms of Eq. (14). Therefore CIS(D) is closely related to the result of the first iteration of the full CCSD equations, and can be justified by this perturbation expansion of CCSD.

4. Comparative calculations

The CIS(D) excited state equations have been implemented as extensions to the Titan programs [*1], as we have recently done for the CCSD and QCISD excited state methods [21,32]. This implementation is currently restricted to closed-shell ground states, and singlet excited states. It is not designed to exploit the potential applicability of CIS(D) to large molecules, but rather is intended to assess its promise for such purposes. We are presently developing direct algorithms analogous to those employed in the ground state HF [33] and MP2 [34,35] methods to address the large molecule problem.

The calculations we report here assess the CIS(D) correction by comparison with first the CIS method that it is designed to improve upon, second the CIS-MP2 method which we anticipate it may be comparable to, and third the CCSD method which is known to yield results within a few tenths of an eV of the exact treatment within a given basis set (full CI) [18–

21]. The molecules chosen for study are several prototypical small organic species, ethylene, formaldehyde, acetaldehyde and butadiene, which have been the subject of extensive CIS and CIS-MP2 calculations [9–11]. These represent cases where CIS is a qualitatively correct starting point, and thus a useful perturbative correction should be expected to improve the results. For consistency with reported calculations, we have employed the 6-311(2+,2+)G** basis [8–11] for ethylene and formaldehyde, and 6-311(2+)G* for acetaldehyde and butadiene. For all molecules we have used MP2/6-31G* optimized geometries, again for consistency.

The results for the excitation energies to low-lying states of these four molecules are collected in Table 1. A total of 43 states are included, and the results may be summarized as follows. The root mean square (rms) deviations between the most accurate method, CCSD, and the three more approximate methods, CIS, CIS(D) and CIS-MP2 are 1.03, 0.43 and 0.41 eV, respectively. For comparison, the rms deviations of the same three methods relative to experiment are 0.93, 0.42 and 0.52 eV, while the rms deviation for CCSD is only 0.20 eV. Overall CIS(D) does achieve its goal of improving markedly on CIS, although its performance is not actually uniformly superior to the more expensive CIS-MP2 correction. Full CCSD theory performs excellently even with a basis set of this relatively modest size, and is clearly the preferred single reference method when feasible.

As a general observation, there are no cases in Table 1 where CIS(D) noticeably worsens a CIS result relative to either CCSD or experiment. Quite substantial corrections to the excited states of formaldehyde and acetaldehyde are qualitatively correct. In fact the success of CIS(D) for the lowest excited state of these molecules is quite striking. CIS(D) appears to overcorrect many of the Rydberg states of formaldehyde and acetaldehyde, while still significantly improving upon CIS. For ethylene and butadiene, the CIS results compare very well with experiment, and as one would hope, the CIS(D) corrections are generally very small. The uniformity with which CIS(D) improves upon CIS suggests that it is a generally valuable addition to a CIS calculation whenever it is feasible. We will probe the validity of CIS(D) with calculations on a much wider range of molecules in the near future, including triplet excited states.

By contrast, CIS-MP2 exhibits more erratic behavior, as has been observed previously [5,9–11]. In some cases CIS-MP2 performs substantially better than CIS(D), for example the lower Rydberg excited states of formaldehyde and acetaldehyde. However in other cases, it is substantially poorer, and indeed can be significantly worse than even CIS. Examples in this class include virtually all of the states of butadiene, and also ethylene to a less dramatic extent, and the lowest excited state of formaldehyde and acetaldehyde. Evidently the failure of CIS-MP2 to be size consistent does indeed lead to inconsistent behavior.

As a final additional test of CIS(D), we chose to perform a preliminary set of calculations on the vertical excitation energies of the benzene molecule. Benzene is widely acknowledged [7,19] to be a challenging test for single reference excited state methods because of the significant contribution of double substitutions to the low-lying $^1B_{1u}$ state [7]. Therefore CIS yields poor results [36] for the relative energies of this state and the lowest singlet excited state, $^1B_{2u}$. Due to the computational limitations of our pilot program, we report results in Table 2 using only the small 6-31+G* basis [38]. This basis lacks the Rydberg functions included in the calculations of Table 1, and thus will not accurately describe states with Rydberg character. However it is small enough that CCSD calculations are feasible, to assess the adequacy of CIS(D).

From Table 2, the splitting of the lowest two excited states is seriously in error at the CIS level, as was already known [36]. The CCSD method substantially improves the splitting, and with further improvement of the basis, this method would come still closer to experiment in all likelihood. Remarkably, the results obtained with the simple CIS(D) correction parallel full CCSD theory very closely, which suggests that the principal correlation effects due to double substitutions are being captured correctly. Apparently this differential correlation effect, while large in magnitude, is well described at only second order. Calculations with larger basis sets are undoubtedly necessary to draw definitive conclusions on the performance of CIS(D) for benzene, but these preliminary results are very encouraging motivation for such studies.

Table 1
Calculations of the excitation energies (in eV) of selected lower singlet excited states of formaldehyde and ethylene using the 6-311(2+, 2+)G** basis [9], and acetaldehyde and butadiene using the 6-311(2+)G* basis [10]. MP2/6-31G* optimized geometries [9–11] are employed

Molecule [a]	State [b]	CIS	CIS(D)	CIS-MP2 [c]	CCSD [d]	Exp. [e]
CH_2O	$^1A_2$ (V)	4.48	3.98	4.58	3.95	4.07
	$^1B_2$ (R)	8.63	6.44	6.85	7.06	7.11
	$^1A_1$ (V)	9.36	7.26	7.66	7.89	7.97
	$^1B_2$ (R)	9.45	8.80	9.19	9.27	
	$^1B_1$ (V)	9.66	8.12	8.47	8.00	8.14
	$^1A_1$ (R)	9.66	9.37	9.97	9.26	
	$^1A_2$ (R)	9.78	7.50	7.83	8.23	8.37
	$^1A_2$ (R)	10.61	8.21	8.46	9.07	8.88
	$^1B_1$ (R)	10.86	8.63	8.94	9.40	
	$^1B_1$ (R)	10.88	8.52	8.75	–	
C_2H_4	$^1B_{3u}$ (V)	7.13	7.21	7.52	7.31	7.11
	$^1B_{1g}$ (V)	7.71	7.84	8.14	7.96	7.80
	$^1B_{1u}$ (V)	7.74	8.04	8.39	8.14	7.60
	$^1B_{2g}$ (V)	7.86	7.86	8.12	7.99	8.01
	^1A_g (R)	8.09	8.18	8.42	8.34	8.29
	$^1B_{3u}$ (V)	8.63	8.69	8.92	8.86	8.62
	^1A_u (V)	8.77	8.80	9.00	9.01	
	$^1B_{3u}$ (V)	8.93	8.96	9.14	9.18	
	$^1B_{1g}$ (R)	9.09	9.12	9.31	–	9.34
	$^1B_{1u}$ (R)	9.09	9.18	9.38	–	9.33
C_2H_4O	$^1A''$ (V)	4.89	4.28	5.27	4.26	4.28
	$^1A'$ (R)	8.51	6.13	6.71	6.78	6.82
	$^1A'$ (R)	9.22	7.04	7.57	7.49	7.46
	$^1A'$ (R)	9.30	7.42	8.00	7.68	7.75
	$^1A''$ (R)	9.37	6.90	7.37	7.64	
	$^1A'$ (R)	9.73	8.50	9.07	–	
	$^1A''$ (V)	9.78	9.34	10.34	–	
	$^1A'$ (R)	10.19	7.70	8.09	8.39	8.43
	$^1A'$ (R)	10.26	7.70	8.08	8.51	8.69
	$^1A''$ (R)	10.31	7.74	8.10	8.57	
C_4H_6	^1B_g (R)	6.11	6.11	6.73	6.20	6.22
	^1B_u (V)	6.21	6.29	7.00	6.42	5.91
	^1A_u (R)	6.45	6.44	7.03	6.53	
	^1A_u (R)	6.61	6.55	7.11	6.67	6.66
	^1B_u (R)	6.99	7.03	7.58	7.17	7.07
	^1A_g (R)	7.19	7.19	7.74	7.10	7.4
	^1B_g (R)	7.22	7.17	7.66	7.31	7.36
	^1B_g (R)	7.25	7.24	7.74	7.39	7.62
	^1B_g (R)	7.39	7.40	7.87	7.55	7.72
	^1A_g (R)	7.45	7.44	7.88	7.61	
	^1A_u (R)	7.78	7.73	6.75	–	8.18
	^1A_u (R)	7.92	7.86	7.66	–	8.21
	^1B_u (R)	8.05	8.01	8.40	–	8.00

[a] Total energies (in au [24]) for the ground state at the HF, MP2 and CCSD levels for the ground state of each molecule are as follows. We list the absolute HF energy, and the correlation energy via MP2 and CCSD. Formaldehyde: $E_{HF} = -113.900047$, $\Delta E_{MP2} = -0.387019$, $\Delta E_{CCSD} = -0.399209$. Ethylene: $E_{HF} = -78.055638$, $\Delta E_{MP2} = -0.332592$, $\Delta E_{CCSD} = -0.361424$. Acetaldehyde: $E_{HF} = -153.953482$, $\Delta E_{MP2} = -0.533791$, $\Delta E_{CCSD} = -0.557338$. Butadiene: $E_{HF} = -154.950804$, $\Delta E_{MP2} = -0.538160$, $\Delta E_{CCSD} = -0.578345$. The correlation energies are obtained with the frozen core approximation for butadiene, and all orbitals correlated for the other species.
[b] The symmetry of the various excited states together with the qualitative description of each state as either valence (V) or Rydberg (R), as given in refs. [9–11].
[c] CIS-MP2 excitation energies are taken from the calculations reported in refs. [9–11].
[d] The missing CCSD values are a consequence of seeking only a fixed number of roots in the calculations. The missing values lie higher in energy than the highest reported CCSD excitation energy for the given molecule.
[e] Experimental excitation energies are taken from the compilations given in refs. [9–11].

Table 2
Vertical excitation energies of benzene, calculated with the CIS, CIS(D) and CCSD methods using the 6-31+G* basis and the frozen core approximation. The geometry employed is the same as for the calculations of ref. [7] (CC=1.395 Å, CH=1.085 Å)

State	CIS [a]	CIS(D) [b]	CCSD [c]	Exp. [d]
$^1B_{2u}$	6.08	5.36	5.23	4.9
$^1B_{1u}$	6.23	6.76	6.68	6.2
$^1E_{1g}$	7.09	6.87	6.84	6.33
$^1A_{2u}$	7.41	7.33	7.32	6.93
$^1E_{2u}$	7.71	7.42	7.44	6.95
$^1E_{1u}$	7.87	7.41	7.47	7.0

[a] The ground state HF energy is −230.706201 au.
[b] The ground state frozen core MP2 correlation energy is −0.769235 au.
[c] The ground state frozen core CCSD correlation energy is −0.804368 au.
[d] Experimental results are taken from ref. [37].

5. Conclusions

In this Letter, we have introduced a new perturbative excited state method, CIS(D), that approximately corrects CIS for the effect of double substitutions. It has two principal formal merits relative to the CIS-MP2 correction proposed earlier [5]:

(1) The CIS(D) correction is strictly size consistent, while CIS-MP2 is not.

(2) The computational complexity of CIS(D) scales as only the fifth power of molecular size, while CIS-MP2 scales as the sixth power.

As such CIS(D) appears to be the simplest size-consistent doubles correction to CIS. In this sense CIS(D) is an excited state analog of MP2 theory for ground states. CIS(D) can be justified based on an order analysis of the excited state coupled-cluster equations carried to second order. This also makes it possible to develop higher-order size-consistent corrections via this formalism. We are currently exploring such ideas [39].

A series of calculations on singlet excited states of small and medium sized molecules using CIS, CIS(D) and CIS-MP2 have been performed. The results were compared to a significantly more accurate theory, CCSD, and to experiment. The CIS(D) results are a substantial overall improvement relative to CIS itself, indicating that the doubles correction is moderately effective in cases where CIS is a reasonable approximation to the state of interest. Preliminary CIS(D) results on benzene where correlation effects are known to be significant also compare well with CCSD. Accordingly CIS(D) appears to be a useful level of theory for examining excited states of large molecules where more accurate techniques such as MR-CI or CCSD methods are not computationally feasible.

Acknowledgement

References

[1] R. Pariser and R.G. Parr, J. Chem. Phys. 21 (1953) 466, 767.
[2] J.A. Pople, Trans. Faraday Soc. 49 (1953) 1375.
[3] J. Del Bene, R. Ditchfield and J.A. Pople, J. Chem. Phys. 55 (1971) 2236.
[4] E.R. Davidson and L.E. McMurchie, in: Excited states, Vol. 5, ed. E.C. Lim (Academic Press, New York, 1982) p. 1.
[5] J.B. Foresman, M. Head-Gordon, J.A. Pople and M.J. Frisch, J. Phys. Chem. 96 (1992) 135.
[6] D.J. Rowe, Nuclear collective motion: models and theory (Methuen, London, 1970).
[7] B.O. Roos, K. Andersson and M.P. Fülscher, Chem. Phys. Letters 192 (1992) 5.
[8] V.A. Walters, C.M. Hadad, Y. Thiel, S.D. Colson, K.B. Wiberg, P.M. Johnson and J.B. Foresman, J. Am. Chem. Soc. 113 (1991) 4782.
[9] K.B. Wiberg, C.M. Hadad, J.B. Foresman and W.A. Chupka, J. Phys. Chem. 96 (1992) 10756.
[10] C.M. Hadad, J.B. Foresman and K.B. Wiberg, J. Phys. Chem. 97 (1993) 4293.
[11] K.B. Wiberg, C.M. Hadad, G.B. Ellison and J.B. Foresman, J. Am. Chem. Soc., in press.
[12] P.J. Bruna and S.D. Peyerimhoff, Advan. Chem. Phys. 67 (1987) 1.
[13] B.O. Roos, Advan. Chem. Phys. 67 (1987) 399.
[14] C.W. Bauschlicher Jr. and S.R. Langhoff, Chem. Rev. 91 (1991) 701.
[15] F.B.C. Machado and E.R. Davidson, J. Chem. Phys. 97 (1992) 1881.
[16] D.R. Yarkony, J. Am. Chem. Soc. 114 (1992) 5406.

[17] H.J. Monkhorst, Intern. J. Quantum Chem. Symp. 11 (1977) 421;
E. Dalgaard and H.J. Monkhorst, Phys. Rev. A 28 (1983) 1217.
[18] J.F. Stanton and R.J. Bartlett, J. Chem. Phys. 98 (1993) 7029;
D.C. Comeau and R.J. Bartlett, Chem. Phys. Letters 207 (1993) 414.
[19] H. Nakatsuji, Acta Chim. Hung. 129 (1992) 719.
[20] H. Koch and P. Jørgensen, J. Chem. Phys. 93 (1990) 3333;
H. Koch, H.J.Aa. Jensen, P. Jørgensen and T. Helgaker, J. Chem. Phys. 93 (1990) 3345.
[21] R.J. Rico and M. Head-Gordon, Chem. Phys. Letters 213 (1993) 224.
[22] C. Møller and M.S. Plesset, Phys. Rev. 46 (1934) 618.
[23] J.A. Pople, J.S. Binkley and R. Seeger, Intern. J. Quantum Chem. Symp. 10 (1976) 1.
[24] A. Szabo and N.S. Ostlund, Modern quantum chemistry: introduction to advanced electronic structure theory (McGraw-Hill, New York, 1989) ch. 6.
[25] K. Wolinski, H.L. Sellers and P. Pulay, Chem. Phys. Letters 140 (1987) 225;
K. Wolinski and P. Pulay, J. Chem. Phys. 90 (1989) 3647.

[26] K. Andersson, P.A. Malmqvist, B.O. Roos, A.J. Sadlej and K. Wolinski, J. Phys. Chem. 94 (1990) 5483.
[27] K. Hirao, Chem. Phys. Letters 201 (1993) 59.
[28] K. Raghavachari, Ann. Rev. Phys. Chem. 42 (1991) 615.
[29] L. Brillouin, Actual. Sci. Ind. 71 (1934) 159.
[30] F.E. Harris, H.J. Monkhorst and D.L. Freeman, Algebraic and diagrammatic methods in many-fermion theory (Oxford Univ. Press, Oxford, 1992).
[31] J. Paldus, J. Čížek and B. Jeziorski, J. Chem. Phys. 90 (1989) 4356; 93 (1990) 1485.
[32] R.J. Rico, T.J. Lee and M. Head-Gordon, Chem. Phys. Letters 218 (1994) 139.
[33] J. Almlöf, K. Fægri Jr. and K. Korsell, J. Comput. Chem. 3 (1982) 385.
[34] M. Head-Gordon, J.A. Pople and M.J. Frisch, Chem. Phys. Letters 153 (1988) 503.
[35] S. Saebø and J. Almlöf, Chem. Phys. Letters 154 (1989) 83.
[36] J.B. Foresman and H.B. Schlegel, in: Recent experimental and computational advances in molecular spectroscopy, ed. R. Fausto (Kluwer, Dordrecht, 1993) pp. 11–26.
[37] M.H. Palmer and I.C. Walker, Chem. Phys. 133 (1989) 113.
[38] M.J. Frisch, J.A. Pople and J.S. Binkley, J. Chem. Phys. 80 (1984) 3265.
[39] R.J. Rico, T.J. Lee and M. Head-Gordon, unpublished.