

An open-shell restricted Hartree–Fock perturbation theory based on symmetric spin orbitals

Timothy J. Lee and Dylan Jayatilaka

NASA Ames Research Center, Moffett Field, CA 94035, USA

Received 25 August 1992; in final form 25 September 1992

A new open-shell perturbation theory is formulated in terms of symmetric spin orbitals. Only one set of spatial orbitals is required, thereby reducing the number of independent coefficients in the perturbed wavefunctions. For second order, the computational cost is shown to be similar to a closed-shell calculation. This formalism is therefore more efficient than the recently developed RMP, ROMP or RMP-MBPT theories. The perturbation theory described herein was designed to have a close correspondence with our recently proposed coupled-cluster theory based on symmetric spin orbitals. The first-order wavefunction contains contributions from only doubly excited determinants. Equilibrium structures and vibrational frequencies determined from second-order perturbation theory are presented for OH, NH, CH, O₂, NH₂ and CH₂.

1. Introduction

Several perturbation theories for including the effects of electron correlation in open-shell systems, which adopt restricted Hartree–Fock (RHF) reference wavefunctions, have been proposed recently [1–7]. The interest in RHF reference functions is due to several factors. Firstly, while unrestricted Hartree–Fock (UHF) energies are lower than RHF energies, it is often the case that electronic energies that approximately include the effects of electron correlation, such as the singles and doubles configuration interaction (CISD) or singles and doubles coupled-cluster (CCSD) levels of theory, are lower for the RHF-based method relative to the UHF treatment. This observation is consistent with similar studies in which the use of a spatial symmetry-broken reference function has been compared with the use of a symmetry constrained reference function. In other words, whether spatial or spin symmetry is concerned, it is better to use a symmetry-constrained reference function [8–11]. Otherwise approximate correlation treatments seem to expend some of their flexibility in correcting this inadequacy in the reference function and are therefore less effective in describing electron correlation.

Secondly, it has been demonstrated [8,9] that the

UHF-based Møller–Plesset perturbation series converges very slowly when the reference function exhibits a large degree of spin contamination. Thus if a reasonably accurate description of electron correlation is desired in low-order perturbation theory it is clearly better to start with an RHF reference wavefunction.

Thirdly, UHF-based perturbation theory is, in general, more expensive than an RHF-based perturbation theory since the different α and β spatial orbitals reduce the symmetry of the molecular orbital (MO) integrals and also the symmetry of the coefficients in the perturbed wavefunctions. It has been demonstrated over the last several years that RHF Møller–Plesset perturbation theory performs well for closed-shell atoms and molecules that are dominated by a single determinant reference function. Thus if UHF-based perturbation theory is employed for open-shell species, the computational cost for open-shell species is out of proportion with respect to the cost of similar calculations on closed-shell species.

Before discussing previous open-shell perturbation theories that are based on RHF reference wavefunctions, it is important to specify the motivations of the present work. As may be inferred from the first three paragraphs, one motivation is to define an open-shell perturbation theory, based on an RHF

reference function, that has a computational cost similar to a closed-shell calculation (based on an RHF reference). However, an equally important consideration for us has been to define a perturbation theory that is consistent with the new open-shell coupled-cluster theory that we have recently proposed [12]. Our specific purpose here is to devise a perturbation theory that can ultimately be used to define a perturbational estimate of connected triple excitations to be added to the singles and doubles coupled-cluster (CCSD) correlation energy – similar to the procedure for closed-shell coupled-cluster theory [13,14]. Nonetheless, it is our belief that the new open-shell perturbation theory presented in this work also has many other advantages over previous theories.

Of the open-shell perturbation theory methods that have been suggested previously, all have their advantages and disadvantages. We limit our discussion here to those methods that have been specifically designed for single-determinant reference wavefunctions. In the method of Hubac and Carsky [1] the zeroth-order Hamiltonian is given by the Roothaan open-shell Fock operator [15]. This method has the advantage that the α and β spatial orbitals are the same and thus there is only one set of MO integrals. However, Amos et al. [3] have shown that the perturbation series from this method converges erratically for relatively well-behaved systems (such as the ground state of NH_2) and diverges for more difficult systems, such as CN. Although based on an RHF reference function, the ROMP and RMP methods proposed by Amos et al. [3] and Knowles et al. [4], respectively, have different α and β spatial orbitals and therefore are more expensive than a closed-shell calculation. It should be noted that this additional cost for the open-shell procedure increases at each order of the perturbation theory. For example, there will be about twice as many unique coefficients for singly excited determinants, but there will be about three times the number of unique coefficients for doubly excited determinants. The ROHF-MBPT method proposed by Lauderdale et al. [5] will also be significantly more expensive for open-shell species than for closed-shell species because of the asymmetry in the α and β spin indices.

The so-called OPT1 and OPT2 methods proposed by Murray and Davidson [6] both adopt only one

set of spatial orbitals and are in fact spin-free formulations. Unfortunately, a consequence of these spin-free formulations is that the annihilation and creation subspaces of the basis are not completely orthogonal. As stated above, one of our goals is to develop an open-shell perturbation theory that can be used to define a perturbational triples correction for our newly proposed open-shell CCSD method, and in this sense the OPT1 and OPT2 methods are not of practical use. This is due to the fact that if the annihilation and creation spaces are not completely orthogonal, then the CCSD equations become considerably more complicated with the addition of thousands of extra terms [16]. It is therefore not obvious that a practical coupled-cluster theory can be formulated along similar lines as OPT1 and/or OPT2. However, since the OPT1 and OPT2 methods are of comparable cost for open- and closed-shell species, it is of interest to compare results from these methods with those from the new perturbation theory proposed here.

Our new open-shell CCSD method [12] is based on a set of spin orbitals that we term “symmetric spin orbitals”. These spin orbitals consist of the standard RHF spatial orbitals together with the following spin functions: the standard α and β spin functions are assigned to doubly occupied and unoccupied orbitals (in the reference function) while the σ^+ and σ^- spin functions are assigned to singly occupied orbitals. The σ^+ and σ^- spin functions are given by

$$\sigma^+ = \frac{1}{\sqrt{2}} (\alpha + \beta), \quad (1)$$

$$\sigma^- = \frac{1}{\sqrt{2}} (\alpha - \beta). \quad (2)$$

Thus the reference function is symmetric with respect to interchange of the α and β spin indices and is a linear combination of all possible M_s states. The reader is referred to ref. [12] for a more complete discussion of the properties of the reference function and of correlated wavefunctions that are based on symmetric spin orbitals. The important points to note here are that (1) perturbation theory (and coupled-cluster theory) formulae based on symmetric spin orbitals possess more symmetry than formulae based on standard spin orbitals; (2) this higher degree of symmetry leads to fewer independent coefficients in

the perturbed wavefunctions which reduces the computational cost, and (3) since the reference function is symmetric with respect to interchange of the α and β spin indices, it is natural to use identical spatial parts for the α and β spin orbitals – again leading to significant computational savings. The main points to be addressed in our definition of a perturbation theory based on symmetric spin orbitals is the choice of the canonicalization of the MOs and the subsequent choice of the zeroth-order Hamiltonian. These two points together with some other relevant aspects are addressed in section 2. Some initial applications of our new perturbation theory (in second order) are presented in section 3 and our conclusions are given in section 4.

2. Theory

In the following equations, I, J, K and L will be used to denote occupied spin orbitals; A, B, C and D will be used to denote unoccupied spin orbitals; P and Q will denote any spin orbital; i, j, k and l denote doubly occupied spatial orbitals; a, b, c and d denote unoccupied spatial orbitals; s, t, u and v will denote singly occupied spatial orbitals; and m, n, p and q will denote any spatial orbital. As discussed previously [12], the symmetric spin orbital basis is complete and orthonormal, and the Fock matrix exhibits much more symmetry than in the standard spin orbital basis. For example, in the spin orbital basis the Fock matrix is given by

$$F_{PQ}^{\mathcal{Q}} = h_{PQ} + \sum_K (PQ\|KK), \quad (3)$$

in which h_{PQ} are the one-electron integrals and the antisymmetrized integral $(PQ\|KK) = (PQ|KK) - (PK|QK)$, where the integral $(PQ|KK)$ is in Mulliken notation. Note that $F_P^{\mathcal{Q}} = F_Q^{\mathcal{P}}$. Examination of the formulae for the various types of Fock matrix elements leads to the following relationships:

$$F_{i\alpha}^{j\alpha} = F_{i\beta}^{j\beta} = \frac{1}{2} (f_{i\alpha}^{j\alpha} + f_{i\beta}^{j\beta}) \\ = h_{ij} + 2v_{ik}^{jk} - v_{ij}^{kk} + v_{is}^s - \frac{1}{2}v_{ij}^{ss}, \quad (4)$$

$$F_{a\alpha}^{b\alpha} = F_{a\beta}^{b\beta} = \frac{1}{2} (f_{a\alpha}^{b\alpha} + f_{a\beta}^{b\beta}) \\ = h_{ab} + 2v_{ak}^{bk} - v_{ab}^{kk} + v_{as}^s - \frac{1}{2}v_{ab}^{ss}, \quad (5)$$

$$F_{i\alpha}^{a\alpha} = F_{i\beta}^{a\beta} = \frac{1}{2} (f_{i\alpha}^{a\alpha} + f_{i\beta}^{a\beta}) \\ = h_{ia} + 2v_{ik}^{ak} - v_{ia}^{kk} + v_{is}^s - \frac{1}{2}v_{ia}^{ss}, \quad (6)$$

$$F_{i\alpha}^{j\beta} = F_{i\beta}^{j\alpha} = \frac{1}{2} (f_{i\alpha}^{j\alpha} - f_{i\beta}^{j\beta}) = -\frac{1}{2}v_{ij}^{ss}, \quad (7)$$

$$F_{a\alpha}^{b\beta} = F_{a\beta}^{b\alpha} = \frac{1}{2} (f_{a\alpha}^{b\alpha} - f_{a\beta}^{b\beta}) = -\frac{1}{2}v_{ab}^{ss}, \quad (8)$$

$$F_{i\alpha}^{a\beta} = F_{i\beta}^{a\alpha} = \frac{1}{2} (f_{i\alpha}^{a\alpha} - f_{i\beta}^{a\beta}) = -\frac{1}{2}v_{ia}^{ss}, \quad (9)$$

$$F_{s\alpha}^{i\alpha} + F_{s\beta}^{i\beta} = \frac{1}{\sqrt{2}} f_{s\alpha}^{i\alpha} \\ = \frac{1}{\sqrt{2}} (h_{is} + 2v_{ik}^{sk} - v_{is}^{kk} + v_{ii}^{si} - v_{is}^{ii}), \quad (10)$$

$$F_{s\alpha}^{a\alpha} + F_{s\beta}^{a\beta} = \frac{1}{\sqrt{2}} f_{s\alpha}^{a\alpha} \\ = \frac{1}{\sqrt{2}} (h_{sa} + 2v_{sk}^{ak} - v_{sa}^{kk} + v_{st}^{at} - v_{sa}^{tt}), \quad (11)$$

$$F_{s\alpha}^{t\alpha} + F_{s\beta}^{t\beta} = h_{st} + 2v_{sk}^{tk} - v_{st}^{kk} + v_{su}^{tu} - v_{st}^{uu}, \quad (12)$$

$$F_{s\alpha}^{i\alpha} - F_{s\beta}^{i\beta} = \frac{1}{\sqrt{2}} f_{s\beta}^{i\beta} \\ = \frac{1}{\sqrt{2}} (h_{is} + 2v_{ik}^{sk} - v_{is}^{kk} + v_{ii}^{si}), \quad (13)$$

$$F_{s\alpha}^{a\alpha} - F_{s\beta}^{a\beta} = \frac{1}{\sqrt{2}} f_{s\beta}^{a\beta} \\ = \frac{1}{\sqrt{2}} (h_{sa} + 2v_{sk}^{ak} - v_{sa}^{kk} + v_{st}^{at}), \quad (14)$$

$$F_{s\alpha}^{t\alpha} - F_{s\beta}^{t\beta} = f_{s\beta}^{t\beta} = h_{st} + 2v_{sk}^{tk} - v_{st}^{kk} + v_{su}^{tu}, \quad (15)$$

$$F_{s\alpha}^{t\alpha} - F_{s\beta}^{t\beta} = F_{s\alpha}^{t\alpha} - F_{s\beta}^{t\beta} = f_{p\alpha}^{q\beta} = f_{p\beta}^{q\alpha} = 0. \quad (16)$$

In eqs. (4)–(16) (and in all following equations) summations are denoted by repeated indices; v_{mn}^{pq} is the $(mp|nq)$ two-electron integral; \mathbf{F} is the Fock matrix in the symmetric spin orbital basis and \mathbf{f} is the Fock matrix in the standard spin orbital basis. Eqs. (4)–(16) contains the symmetry relationships of the symmetric spin orbital Fock matrix elements, the relationships of the symmetric spin orbital Fock matrix elements to the standard spin orbital Fock matrix elements, and also the explicit formula, involving only spatial orbital integrals, for the symmetric spin orbital Fock matrix elements. The high symmetry of the Fock matrix in the symmetric spin orbital basis is quite evident.

For the following discussion, it is convenient to examine the spin orbital Fock matrix in pictorial form:

$$\begin{matrix} & \alpha & \beta & \sigma^+ & \sigma^- & \alpha & \beta \\ \alpha & \left(\begin{array}{cccccc} F_{i\alpha}^{j\alpha} & & & & & \\ F_{i\alpha}^{j\beta} & F_{i\alpha}^{j\alpha} & & & & \\ F_{s\sigma^+}^{i\alpha} & F_{s\sigma^+}^{i\alpha} & F_{s\sigma^+}^{s\sigma^+} & & & \\ F_{s\sigma^-}^{i\alpha} & -F_{s\sigma^-}^{i\alpha} & 0 & F_{s\sigma^-}^{s\sigma^-} & & \\ F_{i\alpha}^{a\alpha} & F_{i\alpha}^{a\beta} & F_{s\sigma^+}^{a\alpha} & F_{s\sigma^-}^{a\alpha} & F_{a\alpha}^{b\alpha} & \\ F_{i\alpha}^{a\beta} & F_{i\alpha}^{a\alpha} & F_{s\sigma^+}^{a\alpha} & -F_{s\sigma^-}^{a\alpha} & F_{a\alpha}^{b\beta} & F_{a\alpha}^{b\alpha} \end{array} \right) \end{matrix} \quad (17)$$

In the above matrix, certain blocks of the Fock matrix have been replaced according to the symmetry relations given in eqs. (4)–(16). As stated in section 1, in order for a perturbation theory to exhibit good convergence properties, it is necessary to make the zeroth-order Hamiltonian (i.e. the Fock matrix) as diagonal as possible. When using an RHF open-shell reference wavefunction, the $F_{s\sigma^+}^{a\alpha}$, $F_{s\sigma^-}^{i\alpha}$ and $F_{i\alpha}^{a\alpha}$ blocks are zero because of the RHF convergence conditions. Furthermore, it is possible to make the $F_{i\alpha}^{j\alpha}$ and $F_{a\alpha}^{b\alpha}$ blocks diagonal by a specific canonicalization of the doubly occupied and unoccupied molecular orbitals that does not mix the different shells (i.e. the RHF energy is left unaffected). Thus it is possible to require that all of the doubly occupied and unoccupied blocks be diagonal while still having the same spatial orbital for the α and β spin orbitals. The operator (in matrix form) that is used to canonicalize the three unique shells (i.e. the doubly occupied, the singly occupied, and the unoccupied shells) is given by

$$\epsilon_{pq} = h_{pq} + (2v_{pk}^{qk} - v_{pq}^{kk}) + (v_{ps}^{qs} - \frac{1}{2}v_{pq}^{ss}) \quad (18)$$

This procedure is equivalent to using $\frac{1}{2}(f_{p\alpha}^{q\alpha} + f_{p\beta}^{q\beta})$ to canonicalize the three shells. Thus, after canonicalization and use of an RHF reference wavefunction, the Fock matrix in pictorial form is given by

$$\begin{matrix} & \alpha & \beta & \sigma^+ & \sigma^- & \alpha & \beta \\ \alpha & \left(\begin{array}{cccccc} \dots & & & & & \\ F_{i\alpha}^{j\beta} & \dots & & & & \\ F_{s\sigma^+}^{i\alpha} & F_{s\sigma^+}^{i\alpha} & F_{s\sigma^+}^{s\sigma^+} & & & \\ 0 & 0 & 0 & F_{s\sigma^-}^{s\sigma^-} & & \\ 0 & F_{i\alpha}^{a\beta} & 0 & F_{s\sigma^-}^{a\alpha} & \dots & \\ F_{i\alpha}^{a\beta} & 0 & 0 & -F_{s\sigma^-}^{a\alpha} & F_{a\alpha}^{b\beta} & \dots \end{array} \right) \end{matrix} \quad (19)$$

Note that since the $F_{i\alpha}^{j\beta}$, $F_{a\alpha}^{b\beta}$, $F_{s\sigma^+}^{i\alpha}$ and $F_{s\sigma^-}^{a\alpha}$ blocks are non-zero, this canonicalization does not strictly obey Koopman's theorem, however, ionization energies and electron affinities should be predicted reasonably well since the diagonal elements of the $F_{s\sigma^+}^{s\sigma^+}$ and $F_{s\sigma^-}^{s\sigma^-}$ blocks are different and have the proper behavior.

It is possible to canonicalize the singly occupied shell such that the $F_{s\sigma^+}^{i\alpha}$ or $F_{s\sigma^-}^{i\alpha}$ blocks will be diagonal. However, it is not possible to make *both* blocks diagonal unless different spatial orbitals are used for the σ^+ and σ^- spin orbitals. As stated earlier, one of our main goals has been to devise an open-shell perturbation theory that is consistent with our recently proposed open-shell coupled-cluster theory – and since the coupled-cluster theory is invariant to a rotation that leads to different spatial orbitals for the σ^+ and σ^- spin orbitals, such a rotation would only serve to complicate unnecessarily the coupled-cluster theory. It must also be emphasized that for *most practical* applications, the $F_{s\sigma^+}^{i\alpha}$ and $F_{s\sigma^-}^{i\alpha}$ blocks will be diagonal anyway. For example, for all doublet electronic states these blocks will automatically be diagonal, and also for all states in which all of the open-shell orbitals belong to different irreducible representations of the molecular point group. In fact, other than the possible exception of transition metal compounds (where it could be argued that perturbation theory is not appropriate anyway), there are not many examples in which the ground electronic state of an atom or molecule possesses more than one open-shell orbital of the same symmetry. For these reasons we have chosen to canonicalize all three unique shells with the matrix operator given in eq. (18). It is worthy of note that this particular canonicalization is identical to that used previously in open-shell configuration gradient theory [17,18]. With this particular canonicalization, neither of the $F_{s\sigma^+}^{i\alpha}$ or $F_{s\sigma^-}^{i\alpha}$ blocks is diagonal, but each is “equally” non-diagonal. That is, the off-diagonal elements $F_{s\sigma^+}^{i\alpha}$ are equal to $-F_{s\sigma^-}^{i\alpha}$, and it can also be shown that

$$F_{s\sigma^+}^{i\alpha} = -\frac{1}{2}v_{su}^{tu} \quad (s \neq t) \quad (20)$$

Having decided to use symmetric spin orbitals, and having chosen a unique canonicalization of the molecular orbitals, we define the zeroth-order Hamil-

tonian following Knowles et al. [4]. That is the zeroth-order Hamiltonian is defined as the diagonal part of the spin orbital Fock matrix, or in matrix operator form it is given by

$$H_0 = \sum_P |P\rangle F_P^P \langle P|. \quad (21)$$

As pointed out by others, the only difference relative to closed-shell theory is that single excitations first appear in the first-order perturbed wavefunction. The RHF energy is given by $E_0 + E_1$, where in terms of spatial orbital integrals E_0 and E_1 are given by

$$E_0 = 2F_{i\alpha}^{i\alpha} + F_{s\sigma^+}^{s\sigma^+}, \quad (22)$$

$$E_1 = -[2v_{ij}^{ij} - v_{ii}^{ii} + 2v_{is}^{is} - v_{ii}^{ss} + \frac{1}{2}(v_{st}^{st} - v_{ss}^{ss})]. \quad (23)$$

The formula for E_2 , in spin orbital notation, is given by

$$E_2 = \frac{|F_I^A|^2}{F_I^I - F_A^A} + \frac{1}{4} \frac{|v_{ij}^{AB} - v_{ji}^{AB}|^2}{4F_I^I + F_J^J - F_A^A - F_B^B}, \quad (24)$$

or in terms of spatial orbital integrals it is given by

$$E_2 = \frac{2(F_{i\alpha}^{i\alpha})^2}{F_{i\alpha}^{i\alpha} - F_{\alpha\alpha}^{\alpha\alpha}} + v_{ij}^{ab} t_{ij}^{ab} + v_{sj}^{ab} t_{sj}^{ab} + v_{ij}^{sb} t_{ij}^{sb} + \frac{1}{4}(v_{st}^{ab} - v_{ts}^{ab}) t_{st}^{ab} + \frac{1}{4}(v_{ij}^{st} - v_{ij}^{ts}) t_{ij}^{st} + v_{sj}^{bt} t_{sj}^{bt}, \quad (25)$$

where the following intermediates have been defined:

$$t_{ij}^{ab} = \frac{2v_{ij}^{ab} - v_{ji}^{ab}}{D_{ij}^{ab}}, \quad (26)$$

$$t_{sj}^{ab} = \frac{2v_{sj}^{ab} - v_{js}^{ab}}{D_{sj}^{ab}}, \quad (27)$$

$$t_{ij}^{sb} = \frac{2v_{ij}^{sb} - v_{ji}^{sb}}{D_{ij}^{sb}}, \quad (28)$$

$$t_{st}^{ab} = \frac{v_{st}^{ab} - v_{ts}^{ab}}{D_{st}^{ab}}, \quad (29)$$

$$t_{ij}^{st} = \frac{v_{ij}^{st} - v_{ji}^{st}}{D_{ij}^{st}}, \quad (30)$$

$$t_{sj}^{bt} = \frac{v_{sj}^{bt}}{D_{sj}^{bt}}, \quad (31)$$

$$D_{mn}^{pq} = F_{m\lambda}^{m\lambda} + F_{n\lambda}^{n\lambda} - F_{p\gamma}^{p\gamma} - F_{q\gamma}^{q\gamma}, \quad (32)$$

where λ is either α or σ^+ depending on m and n , and γ is either α or σ^- depending on p and q . Close examination of eq. (25) reveals that the computational cost of an open-shell second-order perturba-

tion theory calculation is similar to the cost of an analogous closed-shell calculation. Based upon our analysis of the computational cost of open-shell coupled-cluster theory [12], it is likely that higher orders of the perturbation theory proposed here will be somewhat more expensive than an analogous closed-shell calculation, although it will be at most half the cost of other RHF open-shell spin orbital perturbation theories [3–5]. Thus, beyond second order, the method proposed here is the least expensive spin orbital formulation of RHF open-shell perturbation theory. For second order, the method proposed here will be less expensive in the n^4 steps (n is the number of molecular orbitals) involved in the evaluation of E_2 , but as is well known the calculation is dominated by the integral transformation and Knowles et al. [4] have shown that a single n^5 transformation is required for the ROMP and RMP theories provided that the number of unoccupied orbitals is significantly larger than the number of occupied orbitals. It should also be noted that since *all* of the equations for our perturbation theory may be written strictly in terms of spatial orbital integrals, and because there is only one set of spatial orbitals for either the $\{\alpha, \beta\}$ or the $\{\sigma^+, \sigma^-\}$ spin functions, analytical derivative theory applied to this perturbation theory will be significantly simpler than for the ROMP [3], RMP [4] or RMP-MBPT [5] theories. In future studies, explicit formulae for higher orders of the open-shell perturbation theory proposed here, in terms of spatial orbital integrals, will be presented.

The perturbation theory proposed here shall be referred to as Z -averaged perturbation theory or ZAPT since the reference wavefunction is a linear combination of S_z eigenfunctions. Second order ZAPT shall be denoted by ZAPT2. We note that the weight of a particular M_s state in the reference function is given by a formula presented in ref. [12]. As with other open-shell spin orbital RHF perturbation theories, the ZAPT perturbed wavefunctions will not in general be eigenfunctions of S^2 . However, as discussed previously [4], the n th-order energies will have no direct contribution from S^2 contaminants. Should n th-order perturbed wavefunctions that are eigenfunctions of S_z be desired, one procedure is to derive relationships to the ZAPT perturbed wavefunctions by using the S_+ and S_- ladder operators – an alternative procedure is discussed below.

It is important to note some additional features of eq. (25). Since the $F_{\sigma\sigma^+}^{a\alpha}$, $F_{\sigma\sigma^-}^{i\alpha}$ and $F_{i\alpha}^{a\alpha}$, matrix elements are zero due to the RHF convergence conditions, the only singly excited determinants to appear in the first-order wavefunction are of the spin-flip type, i.e. $|i\alpha^{\beta}\rangle$. It has previously been argued [12] that a spin-flip must be counted as an excitation and therefore the $|i\alpha^{\beta}\rangle$ determinant is in fact a double excitation. This argument is based in part on the fact that the $\langle i\alpha^{\beta}|H|0\rangle = F_{i\alpha}^{a\beta}$ matrix element does not possess a one-electron contribution. (The interested reader is referred to ref. [12] for a definition of the interacting space [19] for correlation procedures based on symmetric spin orbitals.) Thus the open-shell perturbation theory defined here has the pleasing property of actually having only double excitations in the first-order perturbed wavefunction.

It is instructive to note that H_0 defined in eq. (21) is invariant to unitary transformation of the $\{\alpha, \beta\}$ spin functions to the $\{\sigma^+, \sigma^-\}$ set (i.e. for the doubly occupied and unoccupied orbitals). In this way, H_0 may be written completely in terms of the $\{\sigma^+, \sigma^-\}$ spin functions or equivalently it may be written completely in terms of the standard $\{\alpha, \beta\}$ spin functions. It should be remembered that in this context the F_P^P quantities in eq. (21) are simply scalar quantities and are unaffected by this transformation (i.e. they remain the Fock matrix elements from the symmetric spin orbital basis). Then by also transforming the spin functions of the doubly occupied orbitals in the reference function Ψ^0 , perturbed wavefunctions Ψ^1, Ψ^2, \dots that are S_z eigenfunctions may be obtained. Thus, because H_0 is invariant to this unitary transformation, the perturbation theory energies E_0, E_1, E_2, \dots order by order will be identical to those obtained by not performing this transformation, but the wavefunctions Ψ^1, Ψ^2, \dots will be different. This observation may be used to simplify the spin integration necessary when determining formulae for E_1, E_2 , etc. However, we note that the term in E_2 involving the coefficients of the singly excited determinants does not easily reduce to the form given in eq. (25) (i.e. it is not immediately obvious that the terms reduce to that given in eq. (25) without knowing in advance that they must). In addition, the concept of the "interacting space" mentioned earlier is lost and therefore the basis for the argument that only doubly excited determinants exist in the first-order wave-

function is also lost. More importantly, however, when formulae for a perturbational triples correction to our recently proposed open-shell CCSD procedure [12] are derived, it will be *necessary* to leave the spin functions of H_0 untouched since it will be necessary to relate the coefficients of the Ψ^1 wavefunction to those in the CCSD wavefunction. In summary, the fact that H_0 is invariant to unitary transformation of the $\{\alpha, \beta\}$ spin functions to the $\{\sigma^+, \sigma^-\}$ set may simplify the spin integration for some terms in the formulae for E_1, E_2, \dots and it will also more easily allow investigation of the convergence properties of the ZAPT perturbation series (via full configuration interaction programs that work in terms of S_z eigenfunctions), but it destroys the concept of the interacting space, and the formulae for E_4 and E_5 (in terms of coefficients multiplied by matrix elements) will not be useful in defining a perturbational correction for connected triple excitations.

3. Test applications

Table 1 contains a comparison of correlation energies determined at second-order ZAPT, ROMP, RMP, OPT1 and OPT2 theories for the $^2\Sigma^+$ state of the CN radical and the 2B_1 state of the NH_2 radical. As expected, the ZAPT2 correlation energies are similar to the RMP, ROMP and OPT2 correlation energies. In fact, the differences between the ZAPT2 energies and the RMP, ROMP and OPT2 energies are between 0.0004–0.0020 E_h . For this small set of examples, the ZAPT2 energies are always smaller in magnitude than any of the other methods, although this set is too small to make any definitive conclusions regarding this point.

One of the great successes of closed-shell perturbation theory has been the accuracy of equilibrium geometries and vibrational frequencies that may be achieved at second order, in conjunction with large one-particle basis sets. Therefore presented in table 2 are the equilibrium geometries, harmonic frequencies and anharmonicities of the first-row hydrides $^2\Pi$ CH, $^3\Sigma^-$ NH, $^2\Pi$ OH and $^1\Sigma^+$ HF for comparison. The DZP basis sets are the usual Dunning–Huzinaga [20,21] [9s5p/4s2p] and [4s/2s] for the heavy atoms and hydrogen, respectively, augmented with a single polarization function ($\alpha_4=1.0, 0.85, 0.80, 0.75$

Table 1
Comparison of second-order perturbation energies (E_h) for ${}^2\Sigma^+$ CN (STO-3G basis) and 2B_1 NH₂ (6-31G basis) ^{a)}

| SCF | ZAPT2 | ROMP | RMP | OPT1 | OPT2 |
|---------------------------------------|----------|----------|----------|----------|----------|
| ${}^2\Sigma^+$ CN | | | | | |
| $R=1.1619, -90.99752$ | -0.15489 | -0.15515 | -0.15685 | - | - |
| 2B_1 NH ₂ | | | | | |
| $R=1.0130, \theta=103.2$ -55.53018 | -0.08673 | -0.08721 | -0.08710 | -0.08963 | -0.08733 |
| $R=1.5195, \theta=103.2$ -55.36773 | -0.12179 | -0.12235 | -0.12224 | -0.12560 | -0.12250 |
| $R=2.0260, \theta=103.2$ -55.18159 | -0.17551 | -0.17612 | -0.17602 | -0.18045 | -0.17677 |

^{a)} RMP results from ref. [4]; ROMP results from ref. [3]; OPT1 and OPT2 results from ref. [6].

Table 2
Total energy (E_h), equilibrium bond distance (Å), harmonic frequency (cm⁻¹) and anharmonicity (cm⁻¹) for ${}^2\Pi$ CH, ${}^3\Sigma^-$ NH, ${}^2\Pi$ OH, and ${}^1\Sigma^+$ HF. Obtained at the ZAPT2 level of theory with various basis sets ^{a)}

| Molecule | Basis | Energy | r_e | ω_e | $\omega_e x_e$ |
|----------|-------|-------------|-------|------------|----------------|
| CH | DZP | -38.354599 | 1.121 | 2976 | 59.4 |
| | TZ2P | -38.375990 | 1.113 | 2946 | 57.7 |
| | TZ2Pf | -38.385467 | 1.114 | 2949 | 56.7 |
| | exp. | - | 1.120 | 2859 | 63.0 |
| NH | DZP | -55.076367 | 1.038 | 3424 | 73.1 |
| | TZ2P | -55.107950 | 1.030 | 3405 | 72.0 |
| | TZ2Pf | -55.121450 | 1.030 | 3403 | 68.9 |
| | exp. | - | 1.036 | 3282 | 78.3 |
| OH | DZP | -75.552255 | 0.976 | 3827 | 83.0 |
| | TZ2P | -75.607646 | 0.967 | 3815 | 81.6 |
| | TZ2Pf | -75.627973 | 0.967 | 3819 | 78.3 |
| | exp. | - | 0.970 | 3738 | 84.9 |
| HF | DZP | -100.235341 | 0.922 | 4209 | 92.9 |
| | TZ2P | -100.318984 | 0.918 | 4156 | 89.4 |
| | TZ2Pf | -100.344406 | 0.918 | 4163 | 86.8 |
| | exp. | - | 0.917 | 4138 | 89.9 |

^{a)} All experimental data from ref. [23].

for F, O, N, and C, and $\alpha_p=1.0$ for H). The hydrogen s functions were scaled by 1.2, as recommended by Dunning. The TZ2P basis sets are also taken from Dunning-Huzinaga [20,22]; [10s6p/5s4p] and [5s/3s] for the heavy atoms and hydrogen, respectively, and augmented with two sets of polarization functions ($\alpha_d=3.107$ and 0.855 for F, 2.314 and 0.645 for O, 1.654 and 0.469 for N, and 1.097 and 0.318

for C; $\alpha_p=1.407$ and 0.388 for H; polarization function exponents were taken from ref. [22]). The TZ2Pf basis set is constructed from the TZ2P basis by adding one set of f functions to the heavy atoms and a set of d functions to hydrogen ($\alpha_f=1.917, 1.428, 1.093,$ and 0.761 for F, O, N, and C, and $\alpha_d=1.057$ for H; also taken from ref. [23]). In these calculations all Cartesian components of the d and f functions were included in the basis sets. In all cases, only valence electrons have been correlated.

As expected, the equilibrium bond distances decrease as the basis set is improved. The harmonic frequencies, on the other hand, initially decrease on going from DZP to TZ2P. For all of the hydrides, the differences between the TZ2P results and the TZ2Pf results are small. It is therefore expected that further improvements in the one-particle basis sets will have only minor effects on the equilibrium bond distances and vibrational frequencies. Comparison of the TZ2Pf results for CH, NH and OH with experiment [24] shows that none of these species are as well described at the ZAPT2 level of theory as is HF (note that ZAPT2 for closed-shell HF is equivalent to second-order Møller-Plesset (MP2) perturbation theory). There is little doubt that the often excellent agreement between experiment and the MP2 level of theory for closed-shell species is due in part to a cancellation of errors. The limited results presented in table 2 suggests that the second-order perturbation level of theory may not exhibit the same fortuitous cancellation of errors for open-shell species.

It was decided to also examine a multiply bonded system, since MP2 generally does not perform as well

Table 3

Total energy (E_h), equilibrium bond distance (\AA), harmonic frequency (cm^{-1}) and anharmonicity (cm^{-1}) for ${}^3\Sigma_g^- \text{O}_2$ and ${}^1\Sigma_g^+ \text{N}_2$. Obtained at the ZAPT2 level of theory with various basis sets ^{a)}

| Molecule | Basis | Energy | r_e | ω_e | $\omega_e x_e$ |
|--------------|-------|-------------|-------|------------|----------------|
| O_2 | DZP | -149.996748 | 1.265 | 1333 | 17.5 |
| | TZ2P | -150.099056 | 1.248 | 1304 | 15.9 |
| | TZ2Pf | -150.139857 | 1.238 | 1369 | 16.4 |
| | exp. | - | 1.208 | 1580 | 12.0 |
| N_2 | DZP | -109.266574 | 1.140 | 2105 | 19.5 |
| | TZ2P | -109.339788 | 1.112 | 2189 | 19.3 |
| | TZ2Pf | -109.368862 | 1.112 | 2201 | 19.1 |
| | exp. | - | 1.098 | 2359 | 14.3 |

^{a)} All experimental data from ref. [23].

for multiply bonded closed-shell species as for singly bonded species. Table 3 contains ZAPT2 equilibrium bond lengths, harmonic frequencies and anharmonicities for ${}^3\Sigma_g^- \text{O}_2$ and also ${}^1\Sigma_g^+ \text{N}_2$ for comparison. The results do not show quite as good convergence with respect to the one-particle basis set, although further improvements are not expected to affect bond distances by more than about 0.02 \AA , and vibrational frequencies by more than about 50 cm^{-1} . Even though the second-order perturbation theory results for N_2 are not as reliable as found for the singly bonded HF, it appears that they are somewhat more reliable than for the open-shell O_2 species – consistent with the above observations for the singly bonded hydrides. The ZAPT2 results for O_2 pre-

sented here are similar to RMP and ROMP results obtained by the Cambridge group [25].

It is also of interest to examine ZAPT2 structures and frequencies for triatomic species, so that bond angles, bending frequencies and more complicated stretching frequencies can be studied. Table 4 contains ZAPT2 equilibrium structures and harmonic frequencies for ${}^3\text{B}_1 \text{CH}_2$ and ${}^2\text{B}_1 \text{NH}_2$ together with the available experimental data [26–28]. Only the TZ2Pf basis set was used for these species, and SCF results are also presented for comparison. The ZAPT2 geometries are clearly an improvement over the SCF quantities. The ZAPT2 harmonic frequencies for NH_2 are also a considerable improvement over the SCF values as compared to the experimental fundamental frequencies. The bending frequencies for both CH_2 and NH_2 are improved at the ZAPT2 level of theory as compared to SCF, but considering that the anharmonic corrections for these modes are probably small (e.g., of the order of 20 cm^{-1}), it appears that the ZAPT2 quantities are still too high.

There are not many results currently available in the literature for the RMP, ROMP, RMP-MBPT, OPT1 or OPT2 levels of theory, but in future studies we plan to examine in more detail the numerical differences between these theories and ZAPT. It is also planned to examine the ZAPT series convergence properties [29]. It is expected that ZAPT will display similar properties to RMP, ROMP, RMP-MBPT and possibly OPT2.

Table 4

Total energy (E_h), equilibrium structure (\AA and deg), and harmonic frequencies (cm^{-1}) for ${}^3\text{B}_1 \text{CH}_2$ and ${}^2\text{B}_1 \text{NH}_2$. Obtained at the SCF and ZAPT2 levels of theory with the TZ2Pf basis set

| Parameter | CH_2 | | | NH_2 | | |
|---------------------|---------------|------------|--------------------|---------------|------------|--------------------|
| | SCF | MP2 | exp. ^{a)} | SCF | MP2 | exp. ^{b)} |
| energy | -38.933498 | -39.056313 | - | -55.583936 | -55.773287 | - |
| r_{AH} | 1.070 | 1.073 | 1.075 | 1.008 | 1.021 | 1.024 |
| $\angle \text{HAH}$ | 129.5 | 133.0 | 133.9 | 104.8 | 103.0 | 103.3 |
| $\omega_1(a_1)$ | 3262 | 3212 | 3021 | 3619 | 3450 | 3219 |
| $\omega_2(a_1)$ | 1299 | 1137 | 963 | 1642 | 1550 | 1497 |
| $\omega_3(b_2)$ | 3471 | 3447 | 3153 | 3708 | 3558 | 3301 |

^{a)} Experimental equilibrium geometry and fundamental frequencies from ref. [25].

^{b)} Experimental vibrationally averaged geometry from ref. [26]; fundamental frequencies from ref. [27].

4. Conclusions

A new perturbation theory for including the effects of electron correlation in open-shell atoms and molecules has been presented. This perturbation theory, ZAPT, is based on the "symmetric spin orbital" basis and on an RHF reference wavefunction, and has been defined so as to be consistent with a recently proposed open-shell coupled-cluster theory [12] that is based on symmetric spin orbitals. In a future study, the open-shell perturbation theory proposed here will be used to define a perturbational estimate of the effects of connected triple excitations for addition to the open-shell CCSD correlation energy. The reference wavefunction is a linear combination of *all* possible M_x states and so it is not an S_z eigenfunction, although two mechanisms, by which wavefunctions that are S_z eigenfunctions may be found, are discussed. The reference function is an eigenfunction of S^2 , although the n th-order perturbed wavefunctions are generally not. However, as has been discussed previously [4], the n th-order perturbational energies do not contain a direct contribution from any S^2 contaminant.

The main advantage of ZAPT is that because the perturbation formulae are symmetric with respect to the α and β spin indices, there is considerably more symmetry present in the spin orbital Fock matrix and consequently also in the perturbed wavefunctions. The symmetry of the spin orbital Fock matrix allows the same spatial orbitals to be used for α and β spin orbitals, even after canonicalization. Thus the computational cost of open-shell second-order ZAPT is about the same as MP2 for closed-shell species. (Note that for closed-shell species ZAPT is equivalent to Møller–Plesset theory.) It is also interesting to note that according to the definition of the interacting space for electron correlation procedures based on symmetric spin orbitals [12], the ZAPT first-order perturbed wavefunction contains only doubly excited determinants and does not contain any purely singly excited determinants.

Sample calculations at second order, i.e. ZAPT2, have been presented. These second-order energies have been compared with other recently proposed RHF open-shell perturbation theories, and ZAPT2 correlation energies are found to be similar to RMP, ROMP, and OPT2 correlation energies (at second

order). Equilibrium structures and vibrational frequencies for CH, NH, OH, HF, O₂, N₂, NH₂ and CH₂ have also been presented. While the ZAPT2 structures and frequencies are an improvement relative to the SCF quantities, the limited results presented here suggest that open-shell second-order perturbation theory may not exhibit the same fortuitous cancellation of errors that has been found for closed-shell species.

Acknowledgement

Professor N.C. Handy, Professor J.A. Pople, Dr. P.J. Knowles and Dr. K. Raghavachari are thanked for illuminating discussions at the 11th Canadian Symposium on Theoretical Chemistry where this work was first presented. Dr. David Schwenke is thanked for suggesting the "ZAPT" acronym. DJ was supported by a National Research Council Post-Doctoral Research Fellowship.

References

- [1] I. Hubac and P. Carsky, *Phys. Rev. A* 22 (1980) 2392.
- [2] K. Wolinski and P. Pulay, *J. Chem. Phys.* 90 (1989) 3647.
- [3] R.D. Amos, J.S. Andrews, N.C. Handy and P.J. Knowles, *Chem. Phys. Letters* 185 (1991) 256.
- [4] P.J. Knowles, J.S. Andrews, R.D. Amos, N.C. Handy and J.A. Pople, *Chem. Phys. Letters* 186 (1991) 130.
- [5] W.J. Lauderdale, J.F. Stanton, J. Gauss, J.D. Watts and R.J. Bartlett, *Chem. Phys. Letters* 187 (1991) 21.
- [6] C. Murray and E.R. Davidson, *Chem. Phys. Letters* 187 (1991) 451.
- [7] K. Andersson, P.-Å. Malmqvist and B.O. Roos, *J. Chem. Phys.* 96 (1992) 1218.
- [8] N.C. Handy, P.J. Knowles and K. Somasundrum, *Theoret. Chim. Acta.* 68 (1985) 87.
- [9] R.H. Nobes, J.A. Pople, L. Radom, N.C. Handy and P.J. Knowles, *Chem. Phys. Letters* 138 (1987) 481.
- [10] R. Murphy, H.F. Schaefer III, R.H. Nobes, L. Radom and R.S. Pitzer, *Intern. Rev. Phys. Chem.* 5 (1986) 229.
- [11] J.D. Watts and R.J. Bartlett, *J. Chem. Phys.* 95 (1991) 6652.
- [12] D. Jayatilaka and T.J. Lee, *Chem. Phys. Letters* 199 (1992) 211.
- [13] K. Raghavachari, G.W. Trucks, J.A. Pople and M. Head-Gordon, *Chem. Phys. Letters* 157 (1989) 479.
- [14] A.P. Rendell, T.J. Lee and A. Komornicki, *Chem. Phys. Letters* 178 (1991) 462.
- [15] C.C.J. Roothaan, *Rev. Mod. Phys.* 32 (1960) 179.
- [16] C.L. Janssen and H.F. Schaefer III, *Theoret. Chim. Acta.* 79 (1991) 1.

- [17] J.E. Rice, R.D. Amos, N.C. Handy, T.J. Lee and H.F. Schaefer III, *J. Chem. Phys.* 85 (1986) 963.
- [18] T.J. Lee, W.D. Allen and H.F. Schaefer III, *J. Chem. Phys.* 87 (1987) 7062.
- [19] A.D. McLean and B. Liu, *J. Chem. Phys.* 58 (1973) 1066.
- [20] S. Huzinaga, *J. Chem. Phys.* 42 (1965) 1293.
- [21] T.H. Dunning, *J. Chem. Phys.* 53 (1970) 2823.
- [22] T.H. Dunning, *J. Chem. Phys.* 55 (1971) 716.
- [23] T.H. Dunning, *J. Chem. Phys.* 90 (1989) 1007.
- [24] K.P. Huber and G. Herzberg, *Constants of diatomic molecules* (Van Nostrand Reinhold, New York, 1979).
- [25] C.W. Murray and N.C. Handy, private communication.
- [26] P. Jensen, P.R. Bunker and A. Hoy, *J. Chem. Phys.* 77 (1982) 5370.
- [27] K. Dressler and D.A. Ramsay, *Phil Trans. Roy. Soc. London A* 251 (1959) 553.
- [28] M.E. Jacox, *J. Phys. Chem. Ref. Data* 17 (1988) 269.
- [29] P.J. Knowles, D. Jayatilaka and T.J. Lee, to be published.