AIP | The Journal of Chemical Physics

On the energy invariance of openshell perturbation theory with respect to unitary transformations of molecular orbitals

T. Daniel Crawford, Henry F. Schaefer III, and Timothy J. Lee

Citation: The Journal of Chemical Physics **105**, 1060 (1996); doi: 10.1063/1.471951
View online: http://dx.doi.org/10.1063/1.471951
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/105/3?ver=pdfcov
Published by the AIP Publishing

On the energy invariance of open-shell perturbation theory with respect to unitary transformations of molecular orbitals

T. Daniel Crawford[a] and Henry F. Schaefer III
Center for Computational Quantum Chemistry, University of Georgia, Athens, Georgia 30602-2556

Timothy J. Lee
NASA Ames Research Center, Moffett Field, California 94035-1000

(Received 29 January 1996; accepted 8 April 1996)

A number of recently proposed single-reference open-shell perturbation theories based on a spin-restricted open-shell Hartree-Fock reference function are examined, with an emphasis on a consistent formalism within which the theories may be compared. In particular, the effect of unitary transformations among the molecular orbitals on the energy is discussed. Of the seven different perturbation theories examined here, the restricted Møller–Plesset theory, open-shell perturbation theory method 1, the method of Hubač and Čársky, Z-averaged perturbation theory, and invariant open-shell perturbation theory methods are found to be invariant to all types of rotations for which the reference wave function is unaffected, though all are invariant to transformations of a more limited nature. Explicit equations for the generalized invariant forms of each perturbation theory are presented, in order to provide working equations for extension of the theories to local correlation schemes or coupled-cluster perturbational corrections, among others. © *1996 American Institute of Physics.* [S0021-9606(96)00327-3]

I. INTRODUCTION

For both practical and aesthetic reasons, a desirable characteristic of the wave function in most methods in electronic structure theory is that of invariance of the energy to certain types of unitary transformations among the reference molecular orbitals (MOs). For example, it has sometimes proved more computationally efficient to use a variety of definitions of MOs in order to improve the convergence of the self-consistent field (SCF) procedure[1–3] (particularly for open-shell systems), though at convergence each set of MOs gives the same energy. Also, localized orbitals (as opposed to SCF canonical orbitals) have been used[4–6] to reduce the magnitude of components contributing to certain types of correlated wave functions, such as configuration interaction (CI), coupled electron-pair approximation (CEPA),[7] and second-order Møller–Plesset perturbation theory (MP2)[8] wave functions, among others. Additionally, it is well known that energy invariance can be used to simplify the construction of analytic energy gradients.[9]

Roothaan's original papers concerning the solution of the Hartree–Fock equations for closed-shell[10] and high-spin open-shell[1] systems note the invariance characteristics of the single-determinant wave function. Specifically, a unitary transformation among the component molecular spin orbitals does not alter the wave function, apart from a trivial phase factor, and, hence, the energy and all properties remain unaffected. In terms of spin-restricted Hartree–Fock wave functions, similar properties hold for transformations among the spatial orbitals alone. For closed-shell wave functions, any rotations among the doubly occupied orbitals or among the virtual orbitals leave the energy unchanged. For high-spin open-shell wave functions, spatial orbital rotations are allowed, provided they are restricted to the doubly occupied space, the virtual space, or the singly occupied space. Rotations in the singly occupied space are allowed due to the fact that all occupied spin orbitals are associated with alpha spin functions, while the unoccupied spin orbitals are associated with beta spin functions, by convention. (This statement also holds, of course, for the so-called symmetric spin orbitals basis,[11] in which different spin functions are used for the open shells. This will be discussed later in this work.) These properties of the SCF wave function have been pointed out by Bobrowicz and Goddard.[12]

Correlated wave functions constructed as a sum of determinants produced through substitution of occupied molecular orbitals from the SCF (reference) wave function by unoccupied orbitals may also, in general, be shown to have similar invariance properties.[9] This includes coupled-cluster (CC) and configuration interaction (CI) wave functions, though it has been recently pointed out[13] that the use of the so-called first-order interacting space[14] in the construction of CI wave functions for high-spin open-shell systems is necessary in order to maintain the general invariance of the energy. On the other hand, coupled-pair functional (CPF) wave functions are not invariant with respect to such orbital rotations.[15] In general, wave functions constructed via many-body perturbation theory (MBPT) also exhibit these invariance properties, although it is sometimes necessary to exercise some care in identifying an appropriate partitioning of the electronic Hamiltonian. Such partitioning requires (implicitly or explicitly) some definition of canonical molecular orbitals.[16]

In the last fifteen years, and especially since 1991, a number of methods for the construction of high-spin open-shell perturbation theory wave functions based on spin-

[a]Electronic mail: crawdad@zopyros.ccqc.uga.edu

restricted open-shell Hartree–Fock (ROHF) reference functions have been introduced in the literature.[17–23] These include: (1) the method of Hubač and Čársky[17] (here referred to as HCPT), (2) restricted open-shell Møller–Plesset theory (ROMP),[18] (3) restricted Møller–Plesset theory (RMP)[19] or (ROHF-MBPT)[20] (independently developed but identical methods), (4) open-shell perturbation theory method 1 (OPT1),[21] (5) open-shell perturbation theory method 2 (OPT2),[21] (6) Z-averaged perturbation theory (ZAPT),[22] and (7) invariant open-shell perturbation theory (IOPT).[23] These methods demonstrate varying convergence properties, and careful analyses have been presented to account for the differences among them.[24,25] However, the general invariance properties of these approaches have not yet been examined in detail.[26]

In this paper we will explicitly examine the effect of unitary transformations among the SCF reference MOs on the energy given by each of the recently presented single-reference open-shell perturbation theories. In section II, we will present a discussion of the invariance properties of closed-shell MP2 energies with respect to such rotations. We will then provide generalized perturbation theory equations, which do not require the zeroth-order Hamiltonian to be diagonal in the N-electron expansion basis. In section III, the partitioning of the Hamiltonian will be given for each method, and iterative equations for the first-order wave function and expressions for the second-order energy will be presented. From these equations, the allowed rotations of the reference MOs may be ascertained. The generalized invariant forms of the zeroth-order Hamiltonian and perturbation presented here may be used as working equations for, for example, the construction of local correlation schemes, such as those developed previously for Møller–Plesset wave functions,[5,6] or for new perturbational corrections to coupled-cluster wave functions, such as the well-known (T) correction.[27,28]

II. GENERAL THEORY

Throughout this discussion, we will use a number of notational conventions. Spin orbitals will be indicated by lowercase letters and spatial orbitals by uppercase letters. Orbital indices p, q, r, and s will refer to general spin orbitals, while i, j, and k (a, b, and c) will refer to spin orbitals occupied (unoccupied) in the reference wave function (with no distinction between the doubly occupied and singly occupied spaces). The indices l, m, and n will refer to spin orbitals in the doubly occupied space, d, e, and f to spin orbitals in the ''doubly unoccupied'' space, and t, u, and v to spin orbitals in the singly occupied space. The uppercase versions of all of the above will apply to spatial orbitals.

In standard closed-shell second-order Møller–Plesset perturbation theory (MP2),[8] the zeroth-order Hamiltonian, $\hat{H}^{(0)}$, is taken to be the spin orbital Fock operator,

$$f_{pq}=h_{pq}+\sum_i \langle pi\|qi\rangle, \tag{1}$$

which is diagonal in the basis of molecular spin orbitals corresponding to the SCF canonical set. Hence, in second quantization

$$\hat{H}^{(0)}=\sum_p \epsilon_p\{p^\dagger p\}=\sum_p f_{pp}\{p^\dagger p\}, \tag{2}$$

where ϵ_p is the pth eigenvalue of the spin orbital Fock matrix, f_{pp}. Throughout this discussion, we will make use of normal-ordered strings of annihilation and creation operators, indicated by {} as in Eq. (2).[29] A diagonal zeroth-order Hamiltonian allows the use of the standard Rayleigh–Schrödinger perturbation theory (RSPT) expressions, which expand the nth-order wave function as a linear combination of the zeroth-order eigenstates, i.e. the set of determinants constructed from the SCF canonical spin orbitals. This leads to the well-known expression for the second-order MP2 energy

$$E_0^{(2)}(\mathrm{MP2})=\frac{1}{4}\sum_{ijab}\frac{|\langle ij||ab\rangle|^2}{\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b}, \tag{3}$$

where $\langle ij||ab\rangle$ is an antisymmetrized two-electron integral in Dirac's notation.

Any non-trivial unitary transformation among the occupied orbitals or the virtual orbitals (including those that mix only the alpha spin orbitals while leaving the beta spin orbitals alone, thereby destroying any spin restriction which may have been imposed on the original orbitals) will result in different values for the ϵ_p in the denominator of Eq. (3). The zeroth-order Hamiltonian (i.e., the Fock matrix) is no longer diagonal in the expansion basis, and thus the standard RSPT expressions do not apply. This does not imply that the MP2 energy is not invariant to such rotations; only that Eq. (3) is valid only for canonical SCF orbitals. In order to construct a set of equations which do not depend on the orbitals, all components of the zeroth-order Hamiltonian must be included in the derivation of the nth-order wave function.

The general zeroth- and nth-order ''Schrödinger equations'' resulting from a many-body perturbation expansion of the wave function and energy may be written

$$\hat{H}^{(0)}|\Psi_0^{(0)}\rangle=E_0^{(0)}|\Psi_0^{(0)}\rangle \tag{4}$$

and

$$\hat{H}^{(0)}|\Psi_0^{(n)}\rangle+\hat{V}|\Psi_0^{(n-1)}\rangle=\sum_{m=0}^{n}E_0^{(m)}|\Psi_0^{(n-m)}\rangle, \tag{5}$$

respectively, where $\hat{V}$ is the perturbation and $|\Psi_0^{(n)}\rangle$ is the nth-order wave function. Expansion of $|\Psi_0^{(n)}\rangle$ in the complete set of substituted determinants, $|\Psi_\nu^{(0)}\rangle$, and left projection of Eq. (5) by $\langle\Psi_0^{(0)}|$ and $\langle\Psi_\mu^{(0)}|$ gives the nth-order energy and wave function equations

$$E_0^{(n)}=\sum_\nu \langle\Psi_0^{(0)}|\hat{V}|\Psi_\nu^{(0)}\rangle\alpha_\nu^{(n-1)} \tag{6}$$

and

$$E_0^{(0)} a_\mu^{(n)} = \sum_\nu \langle \Psi_\mu^{(0)} | \hat{H}^{(0)} | \Psi_\nu^{(0)} \rangle \alpha_\nu^{(n)} + \sum_\nu \langle \Psi_\mu^{(0)} | (\hat{V} - E_0^{(1)}) | \Psi_\nu^{(0)} \rangle a_\nu^{(n-1)} - \sum_{m=2}^{n-1} E_0^{(m)} \alpha_\mu^{(n-m)}, \tag{7}$$

respectively, where the $a_\nu^{(n)}$ are the nth-order expansion coefficients. It should be noted that the nth-order wave function must be determined from Eq. (7) via an iterative procedure if off-diagonal elements of $\hat{H}^{(0)}$ are present.

These generalized perturbation theory equations will be used in the next section as a framework for analyzing the open-shell formalisms mentioned in section I. Specifically, we examine the partitioning of the Hamiltonian and the iterative expressions for the first-order wave functions in order to determine the invariance properties associated with each method.

III. INVARIANCE OF OPEN-SHELL PERTURBATION THEORIES

As discussed in section II, the generalized perturbation theory of Eqs. (6) and (7) may be used in order to correctly determine the invariance properties of the perturbed wave functions given a particular partitioning of the Hamiltonian,

$$\hat{H}_N = \hat{F}_N + \hat{W}_N = \sum_{pq} f_{pq} \{p^\dagger q\} + \tfrac{1}{4} \sum_{pqrs} \langle pq || rs \rangle \{p^\dagger q^\dagger s r\}. \tag{8}$$

The subscript N here refers to the normal-ordered form of this second-quantized operator. For many perturbation theories, the MBPT expansion is carried out by first defining $\hat{H}^{(0)}$ as a sum of certain diagonal blocks of some one-electron operator[30] which provides a convenient definition of molecular orbitals,

$$\hat{H}^{(0)} = \sum_{pq} O_{pq} \{p^\dagger q\} + \sum_{rs} O_{rs} \{r^\dagger s\}, \tag{9}$$

where $\hat{O}$ is some one-electron operator and p and q (r and s) span a subspace of the total orbital space that does not contain r and s (p and q), i.e., the orbital subspaces are disjoint. For example, p and q might be within the doubly occupied space and r and s might be within the singly occupied space. The partitioning is defined, then, by subtracting the diagonal blocks of the operator out of the Hamiltonian,

$$\hat{V} = \sum_{pq} (f_{pq} - O_{pq}) \{p^\dagger q\} + \sum_{rs} (f_{rs} - O_{rs}) \{r^\dagger s\} + \sum_{pr} f_{pr} [\{p^\dagger r\} + \{r^\dagger p\}] + \hat{W}_N. \tag{10}$$

The fundamental concept here is that a perturbation theory based on this partitioning will be invariant to a rotation of the orbitals within a certain subspace if that rotation leaves the partitioning unchanged. Hence, examination of the partitioned Hamiltonian alone will be sufficient to determine the orbital invariance properties of the method, provided the partitioning is written in a sufficiently general form. Therefore, if the rotation

$$|\phi_p\rangle = \sum_{p'} U_{pp'} |\phi_{p'}\rangle \tag{11}$$

is carried out, the zeroth-order Hamiltonian becomes

$$\hat{H}^{(0)} = \sum_{pq} \sum_{p'p''} \sum_{q'q''} U_{qq'}^\dagger U_{pp''}^\dagger O_{p'q'} U_{pp'} U_{qq''} \{p''^\dagger q''\} \tag{12}$$

since $\hat{O}$ is assumed to be Hermitian. Because the transformation is unitary,

$$\hat{H}^{(0)} = \sum_{p'p''} \sum_{q'q''} \delta_{p'p''} \delta_{q'q''} O_{p'q'} \{p''^\dagger q''\} = \sum_{p'q'} O_{p'q'} \{p'^\dagger q'\}. \tag{13}$$

Hence, the spectrum and trace of the zeroth-order Hamiltonian remain unchanged so long as the rotation occurs only within the subspace for which the operator has been defined. Similar arguments may be used for the components of $\hat{V}$. If, on the other hand, this rotation mixes orbitals in separate subspaces, e.g., the rotation

$$|\phi_p\rangle = \sum_{p'} U_{pp'} |\phi_{p'}\rangle + \sum_{r'} U_{pr'} |\phi_{r'}\rangle, \tag{14}$$

where p and r lie in different subspaces, such a rotation will move components from the perturbation into the zeroth-order Hamiltonian and vice versa. Explicitly, the above rotation affects the first term on the right hand side of Eq. (9) as

$$\sum_{pq} O_{pq} \{p^\dagger q\} = \sum_{p'q'} O_{p'q'} \{p'^\dagger q'\} + \sum_{r's'} O_{r's'} \{r'^\dagger s'\} + \sum_{p'r'} O_{p'r'} [\{p'^\dagger r'\} + \{r'^\dagger p'\}]. \tag{15}$$

Hence, components of the perturbation appear in the zeroth-order Hamiltonian, thereby altering the partitioning, and the spectrum and trace of the operator. Therefore, the perturbed energies will not be invariant to rotations of this type.

These general concepts will be useful in the following sections as we examine each of the open-shell perturbation theories in detail. The order in which each method will be presented has been chosen strictly for pedagogical reasons.

A. Restricted Møller–Plesset theory (RMP)

RMP theory[19,20] is perhaps the most straightforward of the open-shell theories investigated here. The zeroth-order Hamiltonian is defined to be the occupied/occupied and virtual/virtual blocks of the spin orbital Fock operator

$$\hat{H}_{\mathrm{RMP}}^{(0)} = \hat{f}_{oo} + \hat{f}_{vv} = \sum_{ij} f_{ij} \{i^\dagger j\} + \sum_{ab} f_{ab} \{a^\dagger b\}. \tag{16}$$

According to this definition, the perturbation consists of the remaining terms in the Hamiltonian [Eq. (8)]

$$\hat{V}_{\mathrm{RMP}} = \hat{f}_{ov} + \hat{W}_N = \sum_{ia} f_{ia}[\{i^\dagger a\} + \{a^\dagger i\}] + \hat{W}_N. \tag{17}$$

Examination of Eqs. (16) and (11)–(13) shows that a unitary transformation of the occupied spin orbitals among themselves or of the unoccupied spin orbitals among themselves will change neither $\hat{f}_{oo}$ nor $\hat{f}_{vv}$. Similar statements may be made for $\hat{V}$. Thus, such rotations leave the energy unchanged. On the other hand, rotations which mix occupied and virtual spaces shift components from $\hat{V}$ into $\hat{H}^{(0)}$ and vice versa, as is clear from Eqs. (14) to (15). Thus, the partitioning changes, and subsequently, the calculated energy changes. Rotations of this type, however, do not leave the reference wave function itself unchanged. Therefore, RMP theory is invariant to all rotations allowed for the ROHF reference wave function.

In the usual application of this method, a diagonal $\hat{H}^{(0)}$ is obtained by diagonalizing the spin orbital Fock matrix in the occupied and virtual subspaces, separately. The resulting orbitals are referred to as ''semi-canonical.'' If such orbitals are not used, Eqs. (6) and (7) must be used to construct iterative expressions for the nth-order energy and wave function. For example, the first-order RMP wave function is constructed iteratively from single and double substitutions via the two identities

$$0 = f_{ia} + \sum_b f_{ab} a_i^{b(1)} - \sum_j f_{ij} a_j^{a(1)} \tag{18}$$

and

$$0 = \langle ij||ab\rangle + \sum_k (f_{kj} a_{ki}^{ab(1)} - f_{ki} a_{kj}^{ab(1)}) + \sum_c (f_{ac} a_{ij}^{cb(1)} - f_{bc} a_{ij}^{ca(1)}), \tag{19}$$

where $a_i^{a(1)}$ and $a_{ij}^{ab(1)}$ are coefficients of the first-order wave function. The second-order energy may then be determined from

$$E_0^{(2)} = \sum_{ia} f_{ia} a_i^{a(1)} + \frac{1}{4} \sum_{ijab} \langle ij||ab\rangle a_{ij}^{ab(1)}. \tag{20}$$

These equations are the same as those presented by Lauderdale *et al.*,[31] who derived them from the coupled-cluster singles and doubles (CCSD) equations. Lauderdale *et al.* also make note of the invariance properties of this method.

Diagonalization of $\hat{H}^{(0)}$ results in a shift of the spatial orbitals, such that different orbitals are associated with different spins (DODS). This orbital set more closely resembles a spin-unrestricted Hartree–Fock wave function (UHF) in structure, though the SCF energy is unchanged. Additionally, $\hat{H}^{(0)}$ is spin dependent, resulting in spin contaminated perturbed wave functions. However, it has been shown[19,25] that the perturbed energies are spin projected, eliminating direct but not indirect spin contamination.[32]

B. Restricted open-shell Møller–Plesset theory (ROMP)

ROMP[18] theory is closely related to RMP[19,20] theory. The zeroth-order Hamiltonian here is defined as the doubly occupied/doubly occupied, singly occupied/singly occupied, and virtual/virtual blocks of the spin orbital Fock operator

$$\hat{H}_{\mathrm{ROMP}}^{(0)} = \hat{f}_{dd} + \hat{f}_{ss} + \hat{f}_{vv} = \sum_{lm} f_{lm}\{l^\dagger m\} + \sum_{tu} f_{tu}\{t^\dagger u\} + \sum_{de} f_{de}\{d^\dagger e\}. \tag{21}$$

The perturbation therefore contains the remaining terms of the Hamiltonian

$$\hat{V}_{\mathrm{ROMP}} = \hat{f}_{dv} + \hat{f}_{ds} + \hat{f}_{sv} + \hat{W}_N = \sum_{ld} f_{ld}[\{l^\dagger d\} + \{d^\dagger l\}] + \sum_{lt} f_{lt}[\{l^\dagger t\} + \{t^\dagger l\}] + \sum_{td} f_{td}[\{t^\dagger d\} + \{d^\dagger t\}] + \hat{W}_N. \tag{22}$$

In this case, more limited rotations leave the perturbed wave functions unaffected. A unitary transformation which mixes the spin orbitals of the doubly occupied, singly occupied, and virtual spaces independently will not change the partitioning. However, unitary transformations which mix spin orbitals in the doubly occupied space together with spin orbitals in the singly occupied space (a rotation of the α spin orbitals, which does not change the reference wave function) will alter the partitioning of the Hamiltonian via the coupling component $\hat{f}_{ds}$. This result follows directly from the analysis given at the beginning of section III. Therefore, ROMP theory is not invariant to general rotations among the occupied spin orbitals. This lack of invariance will not generally affect the calculated perturbed energies, since the ROHF orbitals are constructed with spin-restriction imposed prior to the limited diagonalization of the spin-orbital Fock matrix. However, this can affect the construction of a local correlation scheme for ROMP theory, for example, if the local orbital definition used were based on a spin-dependent (DODS) representation.

Implementation of ROMP theory usually requires first that $\hat{H}^{(0)}$ be diagonalized in the three subspaces separately. This diagonalization results in DODS, just as in RMP theory. Additionally, because the $\hat{H}^{(0)}$ is spin-dependent, the perturbed wave functions are not eigenfunctions of $\hat{S}^2$,[24,25] though the perturbed energies are spin projected.[32]

If $\hat{H}^{(0)}$ is not diagonalized, then iterative expressions based on Eqs. (6) and (7) must be solved. For ROMP, the equations for the single and double substitution contributions to the first-order wave function are

$$0 = f_{ld} + \sum_e f_{de} a_l^{e(1)} - \sum_m f_{lm} a_m^{d(1)} \tag{23}$$

and

$$0=\langle ij||ab\rangle+\sum_{d}\ (a_{ij}^{ad(1)}f_{b\in v,d}-a_{ij}^{bd(1)}f_{a\in v,d})$$
$$+\sum_{t}\ (a_{ij}^{at(1)}f_{b\in s,t}-a_{ij}^{bt(1)}f_{a\in s,t})-\sum_{t}\ (a_{it}^{ab(1)}f_{t,j\in s}$$
$$-a_{jt}^{ab(1)}f_{t,i\in s})-\sum_{l}\ (a_{il}^{ab(1)}f_{l,j\in d}-a_{jl}^{ab(1)}f_{l,i\in d}). \tag{24}$$

The notation, $p\in x$ indicates that the term will not contribute unless the index p corresponds to an orbital in the subspace of orbitals x, where x represents the doubly occupied (d), singly occupied (s), or virtual (v) subspaces. The second-order energy is then calculated from

$$E_0^{(2)}=\sum_{ld}\ f_{ld}a_l^{d(1)}+\tfrac{1}{4}\sum_{ijab}\ \langle ij||ab\rangle a_{ij}^{ab(1)}. \tag{25}$$

Note first that no single substitutions involving singly occupied spin orbitals contribute to the second-order energy for ROMP. As a result, the first-order single substitutions in Eq. (23) look very similar to those of RMP theory, though the summations in ROMP theory involve only doubly occupied and virtual orbitals. Note that the subspace restrictions indicated in Eq. (24) always eliminate two of the summations, depending on the spaces of i, j, a, and b. As a result, the first-order double substitution equations are also similar to those of RMP theory.

C. Open-shell perturbation theory method 1 (OPT1)

In both RMP and ROMP theory, diagonalization of the spin orbital Fock operator fragments in restricted orbital subspaces results in a DODS spin orbital set. OPT1 was introduced as an open-shell perturbation theory that maintains spin restriction on the spatial orbitals.[21] In OPT1, the zeroth-order Hamiltonian is constructed based on an averaged Fock operator, which is defined in terms of spatial orbitals to be

$$\hat{F}^{\mathrm{av}}=\hat{h}+2\hat{J}^c-\hat{K}^c+\hat{J}^o-\tfrac{1}{2}\,\hat{K}^o, \tag{26}$$

where $\hat{h}$ is the usual one-electron Hamiltonian, $\hat{J}^c$ and $\hat{K}^c$ are the Coulomb and exchange operators, respectively, including only the closed-shell orbitals, and $\hat{J}^o$ and $\hat{K}^o$ are these operators including only the singly occupied orbitals. Spin orbital forms of $\hat{F}^{\mathrm{av}}$ allow direct comparison to the original spin orbital Fock operator and subsequent partitioning of the Hamiltonian.

In OPT1, the zeroth-order Hamiltonian is then constructed by writing $\hat{F}^{\mathrm{av}}$ in the spin orbital representation in the doubly occupied, singly occupied, and virtual subspaces,

$$\hat{H}_{\mathrm{OPT1}}^{(0)}=\sum_{LM}\ F_{LM}^{\mathrm{av}}[\{L_\alpha^\dagger\ M_\alpha\}+\{L_\beta^\dagger\ M_\beta\}]$$
$$+\sum_{TU}\ F_{TU}^{av}[\{T_\alpha^\dagger\ U_\alpha\}+\{T_\beta^\dagger\ U_\beta\}]$$
$$+\sum_{DE}\ F_{DE}^{av}[\{D_\alpha^\dagger\ E_\alpha\}+\{D_\beta^\dagger\ E_\beta\}]. \tag{27}$$

The perturbation then consists of the remaining terms from the Hamiltonian

$$\hat{V}_{\mathrm{OPT1}}=\tfrac{1}{2}\sum_{LMT}\ [\langle L_\alpha T_\alpha||M_\alpha T_\alpha\rangle-\langle L_\alpha T_\beta|M_\alpha T_\beta\rangle]\{L_\alpha^\dagger M_\alpha\}+\tfrac{1}{2}\sum_{LMT}\ [\langle L_\beta T_\alpha|M_\beta T_\alpha\rangle-\langle L_\beta T_\beta||M_\beta T_\beta\rangle]\{L_\beta^\dagger M_\beta\}$$
$$+\tfrac{1}{2}\sum_{TUV}\ [\langle T_\alpha V_\alpha||U_\alpha V_\alpha\rangle-\langle T_\alpha V_\beta|U_\alpha V_\beta\rangle]\{T_\alpha^\dagger U_\alpha\}+\tfrac{1}{2}\sum_{TUV}\ [\langle T_\beta V_\alpha|U_\beta V_\alpha\rangle-\langle T_\beta V_\beta||U_\beta V_\beta\rangle]\{T_\beta^\dagger U_\beta\}$$
$$+\tfrac{1}{2}\sum_{DET}\ [\langle D_\alpha T_\alpha||E_\alpha T_\alpha\rangle-\langle D_\alpha T_\beta|E_\alpha T_\beta\rangle]\{D_\alpha^\dagger E_\alpha\}+\tfrac{1}{2}\sum_{DET}\ [\langle D_\beta T_\alpha|E_\beta T_\alpha\rangle-\langle D_\beta T_\beta||E_\beta T_\beta\rangle]\{D_\beta^\dagger E_\beta\}$$
$$+\hat{f}_{ds}+\hat{f}_{dv}+\hat{f}_{sv}+\hat{W}_N. \tag{28}$$

Based on the arguments given in section III earlier, it is clear that this perturbation theory is invariant to rotations which mix the doubly occupied orbitals, the singly occupied orbitals, or the virtual orbitals independently, even if those rotations fail to maintain spin restriction. However, because the zeroth-order Hamiltonian is independent of spin, the nth-order wave function and energy may be written in terms of spatial orbitals alone. Indeed, one of the goals of OPT1 is to construct perturbed wave functions which are eigenfunctions of spin. Hence, spin orbital rotations are not an issue for OPT1 from a practical perspective, since the spin free implementation prevents loss of spin restriction. Therefore, OPT1 is invariant to all spatial orbital rotations which are allowed for the ROHF reference wave function.

Implementation of OPT1 usually requires diagonalization of $\hat{F}^{\mathrm{av}}$ in the three standard subspaces to obtain orbital energies. These energies in the singly occupied space have been described as averages of electron affinities and ionization energies.[21] If this diagonalization is not performed, iterative expressions based on Eq. (7) for the nth-order wave function must be solved to obtain the correct energy. For OPT1, the equations for the single and double substitution contributions to the first-order wave function are

$$0=f_{ld}+\sum_{e} F_{de}^{\mathrm{av}} a_{l}^{e(1)}-\sum_{m} F_{lm}^{\mathrm{av}} a_{m}^{d(1)} \tag{29}$$

and

$$\begin{aligned}0=&\langle ij\|ab\rangle+\sum_{d}(a_{ij}^{ad(1)}F_{b\in v,d}^{\mathrm{av}}-a_{ij}^{bd(1)}F_{a\in v,d}^{\mathrm{av}})\\&+\sum_{t}(a_{ij}^{at(1)}F_{b\in s,t}^{\mathrm{av}}-a_{ij}^{bt(1)}F_{a\in s,t}^{\mathrm{av}})\\&-\sum_{t}(a_{it}^{ab(1)}F_{t,j\in s}^{\mathrm{av}}-a_{jt}^{ab(1)}F_{t,i\in s}^{\mathrm{av}})\\&-\sum_{l}(a_{il}^{ab(1)}F_{l,j\in d}^{\mathrm{av}}-a_{jl}^{ab(1)}F_{l,i\in d}^{\mathrm{av}}),\end{aligned} \tag{30}$$

respectively. The second-order energy is then calculated from

$$E_{0}^{(2)}=\sum_{ld} f_{ld}a_{l}^{d(1)}+\tfrac{1}{4}\sum_{ijab}\langle ij\|ab\rangle a_{ij}^{ab(1)}. \tag{31}$$

Equations (29)–(31) show strong similarities to the analogous equations for ROMP (Eqs. (23)–(25)). The second-order energy expression is in fact identical for the two approaches, though the first-order coefficients are defined differently. This is reasonable since both methods partition the orbital space similarly. The primary difference lies in the use of $\hat{F}^{\mathrm{av}}$ in OPT1.

D. Perturbation theory of Hubač and Čársky (HCPT)

HCPT was the first of the open-shell perturbation theories based on a spin restricted reference wave function,[17] and is closely related to OPT1 in that orbitals are defined in such a way as to maintain this spin restriction. However, HCPT uses different operators from those of OPT1 to define the doubly occupied and virtual orbitals, namely

$$\hat{F}^{D}=\hat{F}^{\mathrm{av}}-\tfrac{1}{2}\hat{K}^{o} \tag{32}$$

and

$$\hat{F}^{V}=\hat{F}^{\mathrm{av}}+\tfrac{1}{2}\hat{K}^{o} \tag{33}$$

where $\hat{F}^{\mathrm{av}}$ and $\hat{K}^{o}$ are the same as defined earlier in section III C. The $\hat{F}^{\mathrm{av}}$ operator is used to define the singly occupied orbitals, as in OPT1. The orbitals constructed from these operators are the same as those originally proposed by Roothaan for high-spin open-shell SCF calculations.[1]

Spin orbital forms of the three operators allow comparison to the original spin orbital Fock matrix. Thus, the zeroth-order Hamiltonian is defined as

$$\begin{aligned}\hat{H}_{\mathrm{HCPT}}^{(0)}=&\sum_{LM}F_{LM}^{D}[\{L_{\alpha}{}^{\dagger}M_{\alpha}\}+\{L_{\beta}{}^{\dagger}M_{\beta}\}]\\&+\sum_{TU}F_{TU}^{\mathrm{av}}[\{T_{\alpha}{}^{\dagger}U_{\alpha}\}\{T_{\beta}{}^{\dagger}U_{\beta}\}]\\&+\sum_{DE}F_{DE}^{V}[\{D_{\alpha}{}^{\dagger}E_{\alpha}\}+\{D_{\beta}{}^{\dagger}E_{\beta}\}].\end{aligned} \tag{34}$$

The perturbation then consists of the remaining terms from the Hamiltonian

$$\begin{aligned}\hat{V}_{\mathrm{HCPT}}=&\tfrac{1}{2}\sum_{LMT}[\langle L_{\alpha}T_{\alpha}|M_{\alpha}T_{\alpha}\rangle-\langle L_{\alpha}T_{\beta}|M_{\alpha}T_{\beta}\rangle]\{L_{\alpha}{}^{\dagger}M_{\alpha}\}+\tfrac{1}{2}\sum_{LMT}[\langle L_{\beta}T_{\alpha}|M_{\beta}T_{\alpha}\rangle-\langle L_{\beta}T_{\beta}|M_{\beta}T_{\beta}\rangle-2\langle L_{\beta}T_{\beta}|T_{\beta}M_{\beta}\rangle]\\&\times\{L_{\beta}{}^{\dagger}M_{\beta}\}+\tfrac{1}{2}\sum_{TUV}[\langle T_{\alpha}V_{\alpha}\|U_{\alpha}V_{\alpha}\rangle-\langle T_{\alpha}V_{\beta}|U_{\alpha}V_{\beta}\rangle]\{T_{\alpha}{}^{\dagger}U_{\alpha}\}+\tfrac{1}{2}\sum_{TUV}[\langle T_{\beta}V_{\alpha}|U_{\beta}V_{\alpha}\rangle-\langle T_{\beta}V_{\beta}\|U_{\beta}V_{\beta}\rangle]\\&\times\{T_{\beta}{}^{\dagger}U_{\beta}\}+\tfrac{1}{2}\sum_{DET}[\langle D_{\alpha}T_{\alpha}|E_{\alpha}T_{\alpha}\rangle-2\langle D_{\alpha}T_{\alpha}|T_{\alpha}E_{\alpha}\rangle-\langle D_{\alpha}T_{\beta}|E_{\alpha}T_{\beta}\rangle]\{D_{\alpha}{}^{\dagger}E_{\alpha}\}-\tfrac{1}{2}\sum_{DET}[\langle D_{\beta}T_{\alpha}|E_{\beta}T_{\alpha}\rangle\\&-\langle D_{\beta}T_{\beta}|E_{\beta}T_{\beta}\rangle]\{D_{\beta}{}^{\dagger}E_{\beta}\}+\hat{f}_{ds}+\hat{f}_{dv}+\hat{f}_{sv}+\hat{W}_{N}.\end{aligned} \tag{35}$$

Similarly to OPT1, HCPT is invariant only to independent rotations of the doubly occupied spin orbitals, singly occupied spin orbitals, or the virtual spin orbitals. However, the fact that $\hat{H}^{(0)}$ is spin independent means that a spin free implementation of the method is possible. Hence, by the same arguments as for OPT1, HCPT is invariant to all spatial orbital rotations which are allowed for the ROHF reference wave function.

Orbital energies are obtained in HCPT by diagonalization of $\hat{F}^{D}$, $\hat{F}^{\mathrm{av}}$, and $\hat{F}^{V}$ in the three standard subspaces. If this diagonalization is not carried out, the iterative expressions based on Eq. (7) must be solved to obtain the perturbed energy. For HCPT, the equations for the single and double substitution components of the first-order wave function are

$$0=f_{ld}+\sum_{e} F_{de}^{V} a_{l}^{e(1)}-\sum_{m} F_{lm}^{D} a_{m}^{d(1)} \tag{36}$$

and

$$0=\langle ij\|ab\rangle+\sum_d \left(a_{ij}^{ad(1)}F_{b\in v,d}^{V}-a_{ij}^{bd(1)}F_{a\in v,d}^{V}\right)$$
$$+\sum_t \left(a_{ij}^{at(1)}F_{b\in s,t}^{\mathrm{av}}-a_{ij}^{bt(1)}F_{a\in s,t}^{\mathrm{av}}\right)$$
$$-\sum_t \left(a_{it}^{ab(1)}F_{t,j\in s}^{\mathrm{av}}-a_{jt}^{ab(1)}F_{t,i\in s}^{\mathrm{av}}\right)$$
$$-\sum_l \left(a_{il}^{ab(1)}F_{l,j\in d}^{D}-a_{jl}^{ab(1)}F_{l,i\in d}^{D}\right), \tag{37}$$

respectively. The second-order energy is then calculated from

$$E_0^{(2)}=\sum_{ld} f_{ld}a_l^{d(1)}+\tfrac{1}{4}\sum_{ijab}\langle ij\|ab\rangle a_{ij}^{ab(1)}. \tag{38}$$

Equations (36)–(38) are nearly identical to the analogous expressions for OPT1 [Eqs. (29)–(31)]. The only differences between the two come from the use of $\hat{F}^D$ and $\hat{F}^V$ for the doubly occupied and virtual orbital subspaces, respectively, in HCPT. As a result, the two methods are identical in terms of computational expense.

E. *Z*-averaged perturbation theory (ZAPT)

Lee and Jayatilaka introduced ZAPT[22] as an attempt to produce nth-order coefficients with higher symmetry than those of RMP and ROMP, and hence, to reduce the computational cost of those approaches. ZAPT requires the redefinition of the spin orbital basis:[11] for each doubly occupied spatial orbital and each unoccupied spatial orbital, the usual α and β spin functions are used, but for the singly occupied orbitals, new spin functions,

$$\sigma^+=\frac{1}{\sqrt{2}}(\alpha+\beta) \tag{39}$$

and

$$\sigma^-=\frac{1}{\sqrt{2}}(\alpha-\beta) \tag{40}$$

are used. σ^+ functions are, by convention, associated with occupied spin orbitals, and σ^- functions with unoccupied spin orbitals. This spin basis is referred to as the symmetric spin basis. In this basis the spin orbital Fock operator is reconstructed. In schematic form, the matrix is

$$\hat{F}^{\mathrm{ZAPT}}=\begin{array}{c} d_\alpha \\ d_{\beta_s} \\ s_{\sigma^+} \\ s_{\sigma^-} \\ v_\alpha \\ v_\beta \end{array}\begin{pmatrix} \hat{F}_{L_\alpha}^{M_\alpha} & \hat{F}_{L_\beta}^{M_\alpha} & \hat{F}_{T_{\sigma^+}}^{L_\alpha} & 0 & 0 & \hat{F}_{L_\alpha}^{D_\beta} \\ \hat{F}_{L_\beta}^{M_\alpha} & \hat{F}_{L_\alpha}^{M_\alpha} & \hat{F}_{T_{\sigma^+}}^{L_\alpha} & 0 & \hat{F}_{L_\alpha}^{D_\beta} & 0 \\ \hat{F}_{T_{\sigma^+}}^{L_\alpha} & \hat{F}_{T_{\sigma^+}}^{L_\alpha} & \hat{F}_{T_{\sigma^+}}^{U_{\sigma+}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{F}_{T_{\sigma^-}}^{U_{\sigma^-}} & \hat{F}_{T_{\sigma^-}}^{U_\alpha} & -\hat{F}_{T_{\sigma^-}}^{D_\alpha} \\ 0 & \hat{F}_{L_\alpha}^{D_\beta} & 0 & \hat{F}_{T_{\sigma^-}}^{D_\alpha} & \hat{F}_{D_\alpha}^{E_\alpha} & \hat{F}_{D_\alpha}^{E_\beta} \\ \hat{F}_{L_\alpha}^{D_\beta} & 0 & 0 & -\hat{F}_{T_{\sigma^-}}^{D_\alpha} & \hat{F}_{D_\alpha}^{E_\beta} & \hat{F}_{D_\alpha}^{E_\alpha} \end{pmatrix}. \tag{41}$$

This form illustrates some of the symmetry of the Fock matrix in the symmetric spin basis. Certain elements of this matrix have been set to zero due to the ROHF convergence conditions or to the orthogonality of the component spin functions.[22] It is interesting to note that elements such as $\hat{F}_{L_\beta}^{M_\alpha}$ are not zero, in general, since the α and β spin functions are not orthogonal to the σ^+ and σ^- spin functions. Additionally, the block diagonal components in the doubly occupied and virtual spaces (e.g. $\hat{F}_{L_\alpha}^{M_\alpha}$) are identical to $\hat{F}^{\mathrm{av}}$ after spin integration of the former. For this reason, Lee and Jayatilaka chose to use the same orbital definition as that of the OPT1 method. $\hat{H}^{(0)}$ is thus defined to include only the block diagonal elements of the spin orbital Fock operator in the doubly occupied and virtual subspaces. In the singly occupied space, however, only the diagonal elements of $\hat{F}^{\mathrm{ZAPT}}$ are included in zeroth-order. Therefore,

$$\hat{H}_{\mathrm{ZAPT}}^{(0)}=\sum_{L_\alpha M_\alpha}\hat{F}_{L_\alpha}^{M_\alpha}\{L_\alpha{}^\dagger M_\alpha\}+\sum_{L_\beta M_\beta}\hat{F}_{L_\beta}^{M_\beta}\{L_\beta{}^\dagger M_\beta\}+\sum_{T_{\sigma^+}}\epsilon_{T_{\sigma^+}}\{T_{\sigma^+}{}^\dagger T_{\sigma^+}\}+\sum_{T_{\sigma^-}}\epsilon_{T_{\sigma^-}}\{T_{\sigma^-}{}^\dagger T_{\sigma^-}\}+\sum_{D_\alpha E_\alpha}\hat{F}_{D_\alpha}^{E_\alpha}\{D_\alpha{}^\dagger E_\alpha\}$$
$$+\sum_{D_\beta E_\beta}\hat{F}_{D_\beta}^{E_\beta}\{D_\beta{}^\dagger E_\beta\} \tag{42}$$

and the perturbation is, therefore,

$$\hat{V}^{\mathrm{ZAPT}}=\sum_{L_\beta M_\alpha}\hat{F}_{L_\beta}^{M_\alpha}\{L_\beta{}^\dagger M_\alpha\}+\sum_{L_\alpha M_\beta}\hat{F}_{L_\alpha}^{M_\beta}\{L_\alpha{}^\dagger M_\beta\}+\sum_{T_{\sigma^+}L_\alpha}\hat{F}_{T_{\sigma^+}}^{L_\alpha}\{T_{\sigma^+}{}^\dagger L_\alpha\}+\sum_{L_\alpha T_{\sigma^+}}\hat{F}_{L_\alpha}^{T_{\sigma^+}}\{L_\alpha{}^\dagger T_{\sigma^+}\}+\sum_{T_{\sigma^+}L_\beta}\hat{F}_{T_{\sigma^+}}^{L_\beta}\{T_{\sigma^+}{}^\dagger L_\beta\}$$

$$+\sum_{L_\beta T_{\sigma^+}} \hat{F}^{T_{\sigma^+}}_{L_\beta}\{L_\beta{}^\dagger T_{\sigma^+}\}+\sum_{L_\beta D_\alpha} \hat{F}^{D_\alpha}_{L_\beta}\{L_\beta{}^\dagger D_\alpha\}+\sum_{D_\alpha L_\beta} \hat{F}^{L_\beta}_{D_\alpha}\{D_\alpha{}^\dagger L_\beta\}+\sum_{T_{\sigma^+}U_{\sigma^+}} (1-\delta_{T_{\sigma^+}U_{\sigma^+}})\hat{F}^{U_{\sigma^+}}_{T_{\sigma^+}}\{T_{\sigma^+}{}^\dagger U_{\sigma^+}\}$$

$$+\sum_{T_{\sigma^-}U_{\sigma^-}} (1-\delta_{T_{\sigma^-}U_{\sigma^-}})\hat{F}^{U_{\sigma^-}}_{T_{\sigma^-}}\{T_{\sigma^-}{}^\dagger U_{\sigma^-}\}+\sum_{L_\alpha D_\beta} \hat{F}^{D_\beta}_{L_\alpha}\{L_\alpha{}^\dagger D_\beta\}+\sum_{D_\beta L_\alpha} \hat{F}^{L_\alpha}_{D_\beta}\{D_\beta{}^\dagger L_\alpha\}+\sum_{D_\alpha T_{\sigma^-}} \hat{F}^{T_{\sigma^-}}_{D_\alpha}\{D_\alpha{}^\dagger T_{\sigma^-}\}$$

$$+\sum_{T_{\sigma^-}D_\alpha} \hat{F}^{D_\alpha}_{T_{\sigma^-}}\{T_{\sigma^-}{}^\dagger D_\alpha\}+\sum_{D_\beta T_{\sigma^-}} \hat{F}^{T_{\sigma^-}}_{D_\beta}\{D_\beta{}^\dagger T_{\sigma^-}\}+\sum_{T_{\sigma^-}D_\beta} \hat{F}^{D_\beta}_{T_{\sigma^-}}\{T_{\sigma^-}{}^\dagger D_\beta\}+\sum_{D_\beta E_\alpha} \hat{F}^{E_\alpha}_{D_\beta}\{D_\beta{}^\dagger E_\alpha\}$$

$$+\sum_{D_\alpha E_\beta} \hat{F}^{E_\beta}_{D_\alpha}\{D_\alpha{}^\dagger E_\beta\}+\hat{W}_N. \tag{43}$$

In Eq. (42), $\epsilon_{T_{\sigma^+}}$ is the diagonal element of $\hat{F}^{U_{\sigma^+}}_{T_{\sigma^+}}$ and $\epsilon_{T_{\sigma^-}}$ that of $\hat{F}^{U_{\sigma^-}}_{T_{\sigma^-}}$, and these same terms have been removed from Eq. (43).

Based on our previous analyses and, in particular, comparison to Eq. (15), ZAPT is not invariant to rotations which mix doubly occupied spin orbitals and singly occupied spin orbitals together, similarly to ROMP theory. However, these rotations do not, in general, have practical implications for ZAPT, due to the orbital canonicalization chosen by Lee and Jayatilaka.[22] Specifically, since the $\hat{F}^{\mathrm{av}}$ operator is used to define orbitals in all three spaces (just as for OPT1), spin restriction of the orbitals is maintained. In a truly spin restricted implementation of the method, then, rotations of doubly occupied and singly occupied spatial orbitals are not allowed to occur since they alter the ROHF reference wave function itself.

Additionally, ZAPT is not generally invariant to rotations which mix singly occupied spin orbitals together, since only the diagonal elements of $\hat{F}^{\mathrm{ZAPT}}$ in this block have been used to define $\hat{H}^{(0)}$. However, as pointed out by Lee and Jayatilaka, in many practical applications, such mixings cannot occur since the singly occupied orbitals are of different spatial symmetries. Hence, the singly occupied block of $\hat{F}^{\mathrm{ZAPT}}$ will be diagonal regardless of the chosen canonicalization conditions. Therefore, it is reasonable to say that from a practical perspective, ZAPT is invariant to all spatial orbital rotations which are allowed for the ROHF reference wave function, though this is not true for all conceivable cases.

The iterative expressions for the single and double substitution contributions to the first-order wave function for ZAPT, which must be solved in the case that the $\hat{F}^{\mathrm{av}}$ operator is not diagonalized in the three standard subspaces, are

$$0=\hat{F}^{L_\alpha}_{D_\beta}+\sum_{E_\beta} \hat{F}^{E_\beta}_{D_\beta}a^{E_\beta(1)}_{L_\alpha}-\sum_{M_\alpha} \hat{F}^{L_\alpha}_{M_\alpha}a^{D_\beta(1)}_{M_\alpha}, \tag{44}$$

$$0=\hat{F}^{L_\beta}_{D_\alpha}+\sum_{E_\alpha} \hat{F}^{E_\alpha}_{D_\alpha}a^{E_\alpha(1)}_{L_\beta}-\sum_{M_\beta} \hat{F}^{L_\beta}_{M_\beta}a^{D_\alpha(1)}_{M_\beta}, \tag{45}$$

and

$$0=\langle ij\|ab\rangle+\sum_{D_\alpha} (a^{aD_\alpha(1)}_{ij}\hat{F}^{D_\alpha}_{b\in v_\alpha}-a^{bD_\alpha(1)}_{ij}\hat{F}^{D_\alpha}_{a\in v_\alpha})+\sum_{D_\beta} (a^{aD_\beta(1)}_{ij}\hat{F}^{D_\beta}_{b\in v_\beta}-a^{bD_\beta(1)}_{ij}\hat{F}^{D_\beta}_{a\in v_\beta})+\sum_{T_{\sigma^-}} (a^{aT_{\sigma^-}(1)}_{ij}\hat{F}^{T_{\sigma^-}}_{b\in s_{\sigma^-}}\delta_{b,T_{\sigma^-}}$$

$$-a^{bT_{\sigma^-}(1)}_{ij}\hat{F}^{T_{\sigma^-}}_{a\in s_{\sigma^-}}\delta_{a,T_{\sigma^-}})-\sum_{T_{\sigma^+}} (a^{ab(1)}_{iT_{\sigma^+}}\hat{F}^{j\in s_{\sigma^+}}_{T_{\sigma^+}}\delta_{j,T_{\sigma^+}}-a^{ab(1)}_{jT_{\sigma^+}}\hat{F}^{i\in s_{\sigma^+}}_{T_{\sigma^+}}\delta_{i,T_{\sigma^+}})-\sum_{L_\alpha} (a^{ab(1)}_{iL_\alpha}\hat{F}^{j\in d_\alpha}_{L_\alpha}-a^{ab(1)}_{jL_\alpha}\hat{F}^{i\in d_\alpha}_{L_\alpha})$$

$$-\sum_{L_\beta} (a^{ab(1)}_{iL_\beta}\hat{F}^{j\in d_\beta}_{L_\beta}-a^{ab(1)}_{jL_\beta}\hat{F}^{i\in d_\beta}_{L_\beta}). \tag{46}$$

The second-order energy is then calculated from

$$E^{(2)}_0=\sum_{L_\alpha D_\beta} \hat{F}^{D_\beta}_{L_\alpha}a^{D_\beta(1)}_{L_\alpha}+\sum_{L_\beta D_\alpha} \hat{F}^{D_\alpha}_{L_\beta}a^{D_\alpha(1)}_{L_\beta}+\tfrac{1}{4}\sum_{ijab} \langle ij\|ab\rangle a^{ab(1)}_{ij}. \tag{47}$$

Equations (44)–(47) appear to be significantly different from both ROMP [Eqs. (23)–(25)] and OPT1 [Eqs. (29)–(31)]. However, the complication here is primarily notational, as we have illustrated the necessary spin combinationsin order to make the relationship to the symmetric spin basis clear. Additionally, Lee and Jayatilaka have pointed out[22] that first-order interacting-space arguments,[14] when applied to the symmetric spin basis, suggest that spin-flip substitutions such as that indicated by $a^{D_\beta(1)}_{L_\alpha}$ should be considered double substitutions since their matrix element with the reference wave function through the Hamiltonian contains no one-electron contributions.[11] Hence, the symmetric spin orbitals obey a form of Brillouin's theorem and the second-order ZAPT energy contains no contributions from singly substituted determinants, similar to the MP2 energy.

F. Open-shell perturbation theory method 2 (OPT2)

OPT2 was introduced by Murray and Davidson[21] as an improvement to the OPT1 method; it is designed to have better convergence. OPT2 chooses the same orbital definition as OPT1, i.e., the eigenfunctions of $\hat{F}^{\mathrm{av}}$. The primary distinction between OPT1 and OPT2 comes in the addition of a two-electron component to the zeroth-order Hamiltonian of OPT1:

$$\hat{H}^{(0)}_{\mathrm{OPT2}} = \hat{H}^{(0)}_{\mathrm{OPT1}} + \sum_T \hat{n}_T(\hat{n}_T - 2)\,\frac{1}{2}\,\hat{K}^o_{TT}, \tag{48}$$

where $\hat{K}^o_{TT}$ is the Tth diagonal spatial orbital component of the usual exchange operator defined over the singly occupied orbitals only, and $\hat{n}_P$ is the operator

$$\hat{n}_P = P_\alpha{}^\dagger P_\alpha + P_\beta{}^\dagger P_\beta . \tag{49}$$

Thus, in normal-ordered operator notation, the OPT2 zeroth-order Hamiltonian becomes,

$$\begin{aligned}\hat{H}^{(0)}_{\mathrm{OPT2}} = {}& \sum_{LM} F^{\mathrm{av}}_{LM}[\{L_\alpha{}^\dagger M_\alpha\} + \{L_\beta{}^\dagger M_\beta\}] + \sum_{DE} F^{\mathrm{av}}_{DE}[\{D_\alpha{}^\dagger E_\alpha\} + \{D_\beta{}^\dagger E_\beta\}] + \sum_{TU} F^{\mathrm{av}}_{TU}[\{T_\alpha{}^\dagger U_\alpha\} + \{T_\beta{}^\dagger U_\beta\}] \\ & - \tfrac{1}{2} \sum_T K^o_{TT}[\{T_\alpha{}^\dagger T_\alpha\} - \{T_\beta{}^\dagger T_\beta\}] + \tfrac{1}{2} \sum_T K^o_{TT}[\{T_\alpha{}^\dagger T_\alpha T_\alpha{}^\dagger T_\alpha\} + \{T_\beta{}^\dagger T_\beta T_\beta{}^\dagger T_\beta\} + \{T_\alpha{}^\dagger T_\alpha T_\beta{}^\dagger T_\beta\} + \{T_\beta{}^\dagger T_\beta T_\alpha{}^\dagger T_\alpha\}].\end{aligned} \tag{50}$$

The purpose of this correction term is to shift the singly occupied orbital energies by a factor of $1/2K^o_{TT}$ depending on whether an excitation is occurring into or out of the orbital. That is, if the excitation occurs into the orbital, the orbital energy will resemble an electron affinity, while if the excitation occurs out of the orbital the orbital energy will resemble an ionization potential.

The inclusion of individual components of $\hat{W}_N$ in the last four terms in the above equation differentiates OPT2 from all of the other perturbation theories examined so far. It is clear from the structure of $\hat{H}^{(0)}_{\mathrm{OPT2}}$ that the energy will be invariant to orbital rotations within the doubly occupied and virtual spaces separately, just as for OPT1. However, an analysis similar to that presented earlier for one-electron operators indicates that rotations of orbitals in the singly occupied space will alter the partitioning of the Hamiltonian, and, therefore, the energy. A practical implication is that different geometric representations (such as rotation of the molecule relative to a space-fixed axis system) can result in different energies; the OPT2 energy expression can be a multivalued function of nuclear geometry even though the mathematical definition of the spatial orbitals is maintained. This has been shown to occur in triplet twisted ethylene,[25] where rotations within the degenerate pair of singly occupied orbitals caused energy variations on the order of 9 kcal/mol.

G. Invariant open-shell perturbation theory (IOPT)

IOPT was introduced by Kozlowski and Davidson[23] as an adjustment to OPT2 to ensure invariance with respect to rotations among the singly occupied orbitals. This involved altering the correction function such that singly occupied orbital energies could be shifted by a rotation-independent term, unlike that in OPT2. The zeroth-order Hamiltonian of IOPT is

$$\begin{aligned}\hat{H}^{(0)}_{\mathrm{IOPT}} = {}& \sum_{LM} F^{\mathrm{av}}_{LM}[\{L_\alpha{}^\dagger M_\alpha\} + \{L_\beta{}^\dagger M_\beta\}] \\ & + \sum_{DE} F^{\mathrm{av}}_{DE}[\{D_\alpha{}^\dagger E_\alpha\} + \{D_\beta{}^\dagger E_\beta\}] \\ & + \sum_{TU} F^{\mathrm{av}}_{TU}[\{T_\alpha{}^\dagger U_\alpha\} + \{T_\beta{}^\dagger U_\beta\}] \\ & + \tfrac{1}{2}k\left|\sum_T [\{T_\alpha{}^\dagger T_\alpha\} + \{T_\beta{}^\dagger T_\beta\}] - N_s\right|,\end{aligned} \tag{51}$$

where N_s is the number of open-shell electrons in the reference wave function and k is a constant given by

$$k = \sum_{TU} K^o_{TU}/N_s . \tag{52}$$

Therefore, as the molecular orbitals are rotated in the singly occupied space, the number operators occurring in the final term of the zeroth-order Hamiltonian remain unaffected, as does the value of k. As a result, IOPT is invariant to all spatial orbital rotations allowed for the reference wave function, just as OPT1 is. The IOPT perturbation is simply the OPT1 perturbation, Eq. (28), less the last term shown in the equation above. The iterative equations for IOPT will therefore be exactly those presented in equations (29) and (30), with orbital-independent correction terms used to shift the diagonal elements of F^{av}_{pq}. Because of their similarity to the iterative expressions for OPT1 already presented, we will omit these equations here. It should be pointed out that the IOPT approach has been shown to contain size-extensivity errors,[33] though the significance of these errors has been questioned.[34]

IV. CONCLUSIONS

Using a consistent notation and formalism, we have explicitly examined the effects of unitary transformations of the reference molecular orbitals on the energy and perturbed wave functions given by a number of spin-restricted single-reference open-shell perturbation theories. Of the seven methods examined here, the RMP, OPT1, HCPT, ZAPT, and IOPT approaches are found to be invariant to all types of rotations which do not alter the reference wave function. The ROMP method, which is implemented in a spin-dependent representation, and the OPT2 method, which includes two-electron components in the zeroth-order Hamiltonian, are found to be invariant to more limited transformations. While this will not necessarily affect the practical application of ROMP (except perhaps in its extension to local correlation schemes), the consequences are more severe for OPT2, whose energy functional may be multivalued for a single nuclear framework.

Additionally, we have presented generalized invariant equations for the partitioned Hamiltonian as well as for the first-order wave function and second-order energy. These equations may be applied with any convenient set of reference molecular orbitals, subject to the constraint that the chosen orbitals are related to the particular method's so-called canonical orbitals by an allowed unitary transformation. It is clear that the standard equations associated with a particular perturbation theory (that is, a particular partitioning of the Hamiltonian) are usually valid only for a specific set of canonical orbitals (such as the semicanonical orbitals associated with RMP theory) and for no other. The use of other orbitals in the canonical orbital expressions will result in a different nth-order energy, since the zeroth-order Hamiltonian is not diagonal in the new orbital basis. The equations presented here, however, are valid for any set of molecular orbitals which are simply a (perhaps limited) unitary transformation away from the canonical orbitals defined for the perturbation theory in question. Such orbitals require, in general, an iterative construction of the perturbed wave functions. Additionally, the generalized equations presented here may serve as working forms for the extension of these methods to local correlation schemes or perturbational corrections to coupled-cluster energies, such as the well-known (T) correction. This latter example is interesting because the two most widely used approaches for open-shell systems[35–37] at present suffer either from a lack of invariance[35] or the use of a less convenient set of orbitals.[36,38] Construction of a new correction, based on ZAPT, for example, would overcome these problems.

ACKNOWLEDGMENTS

This work was supported by the U.S. National Science Foundation, Grant No. CHE-9216754. T.D.C. was supported by a Department of Defense Graduate Fellowship and a Fritz London Graduate Fellowship. T.D.C. would like to thank Professor Rodney J. Bartlett and his research group for their hospitality during a one week visit to the University of Florida's Quantum Theory Project. The authors would also like to thank Professor John F. Stanton, Professor Ernest Davidson, Professor Martin Head-Gordon, Professor Gustavo Scuseria, and Professor Mark Hoffman, and CCQC members Dr. George Vacek, Dr. Justin Fermann, and Brian Kellogg for helpful discussions.

[1] C. C. J. Roothaan, Rev. Mod. Phys. **32**, 179 (1960).
[2] M. F. Guest and V. R. Saunders, Mol. Phys. **28**, 819 (1974).
[3] J. S. Binkley, J. A. Pople, and P. S. Dobosh, Mol. Phys. **28**, 1423 (1974).
[4] S. Sæbø and P. Pulay, Chem. Phys. Lett. **113**, 13 (1985).
[5] P. Pulay and S. Sæbø, Theo. Chim. Acta **69**, 357 (1986).
[6] S. Sæbø and P. Pulay, J. Chem. Phys. **86**, 914 (1987).
[7] W. Meyer, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer (Plenum, New York, 1977), Chap. 11, p. 413.
[8] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).
[9] D. Jayatilaka and N. C. Handy, Int. J. Quantum Chem. **42**, 445 (1992).
[10] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).
[11] D. Jayatilaka and T. J. Lee, Chem. Phys. Lett. **199**, 211 (1992).
[12] F. W. Bobrowicz and W. A. Goddard, in Ref. 7, Chap. 4, p. 79.
[13] C. W. Bauschlicher and H. Partridge, Theo. Chim. Acta **85**, 255 (1993).
[14] A. D. Mclean and B. Liu, J. Chem. Phys. **58**, 1066 (1973).
[15] J. E. Rice, T. J. Lee, and N. C. Handy, J. Chem. Phys. **88**, 7011 (1988).
[16] ''Canonical'' in this context does not necessarily refer to the set of SCF canonical orbitals in the basis of which the spin-orbital Fock matrix is diagonal, but simply refers to the set of orbitals necessary for some specific partitioning of the Hamiltonian.
[17] I. Hubač and P. Čársky, Phys. Rev. A **22**, 2392 (1980).
[18] R. D. Amos, J. S. Andrews, N. C. Handy, and P. J. Knowles, Chem. Phys. Lett. **185**, 256 (1991).
[19] P. J. Knowles, J. S. Andrews, R. D. Amos, N. C. Handy, and J. A. Pople, Chem. Phys. Lett. **186**, 130 (1991).
[20] W. J. Lauderdale, J. F. Stanton, J. Gauss, J. D. Watts, and R. J. Bartlett, Chem. Phys. Lett. **187**, 21 (1991).
[21] C. M. Murray and E. R. Davidson, Chem. Phys. Lett. **187**, 451 (1991).
[22] T. J. Lee and D. Jayatilaka, Chem. Phys. Lett. **201**, 1 (1993).
[23] P. M. Kozlowski and E. R. Davidson, Chem. Phys. Lett. **226**, 440 (1994).
[24] C. M. Murray and N. C. Handy, J. Chem. Phys. **97**, 6509 (1992).
[25] T. J. Lee, A. P. Rendell, K. G. Dyall, and D. Jayatilaka, J. Chem. Phys. **100**, 7400 (1994).
[26] The exception to this statement is the OPT2 method of Murray and Davidson (Ref. 21), which has been shown (Ref. 25) to be non-invariant with respect to rotations among degenerate singly-occupied orbitals. The IOPT method (Ref. 23) was designed to overcome this problem. This is discussed in the text.
[27] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
[28] R. J. Bartlett, J. D. Watts, S. A. Kucharski, and J. Noga, Chem. Phys. Lett. **165**, 513 (1990), erratum: **167**, 609 (1990).
[29] J. Paldus and J. Čížek, Adv. Quantum Chem. **9**, 105 (1975).
[30] This is not true of the OPT2 method, which includes certain two-electron terms in the zeroth-order Hamiltonian. This is discussed in the text.
[31] W. J. Lauderdale, J. F. Stanton, J. Gauss, J. D. Watts, and R. J. Bartlett, J. Chem. Phys. **97**, 6606 (1992).
[32] While the energies are spin projected—the spin free Hamiltonian cannot couple states of different spin — third- and higher-order energies will not be the same as those determined using perturbed wave functions which are spin eigenfunctions.
[33] P.-Å. Malmqvist, Chem. Phys. Lett. **241**, 429 (1995).
[34] E. R. Davidson, Chem. Phys. Lett. **241**, 432 (1995).
[35] G. E. Scuseria, Chem. Phys. Lett. **176**, 27 (1991).
[36] J. Gauss, W. J. Lauderdale, J. F. Stanton, J. D. Watts, and R. J. Bartlett, Chem. Phys. Lett. **182**, 207 (1991).
[37] T. D. Crawford and H. F. Schaefer, J. Chem. Phys. **104**, 6259 (1996).
[38] J. D. Watts, J. Gauss, and R. J. Bartlett, J. Chem. Phys. **98**, 8718 (1993).