

**ANALYTIC ENERGY GRADIENTS FOR SECOND-ORDER  
MULTIREFERENCE PERTURBATION THEORY**

HARUYUKI NAKANO, NORIO OTSUKA AND KIMIHIKO HIRAO

*Department of Applied Chemistry, Graduate School of Engineering, University of Tokyo,  
Tokyo 113-8656, Japan*

An analytic energy gradient method for second-order quasidgenerate perturbation theory with multiconfigurational self-consistent field reference functions (MC-QDPT) is presented. An expression for the energy gradients is derived using the Lagrange multiplier method. The gradients are calculated without solving coupled perturbed equations. Instead, it is necessary to solve eight sets of linear equations for the multipliers. Six of the eight equations reduces to simple partial differential forms which directly give the multipliers, and only the remaining two are large scale linear equations that need iterative procedure. The gradients are given as the product of the first derivative integrals and the effective densities that depends on the obtained multipliers and the parameters. The expression for the conventional quasidgenerate perturbation theory and numerical results for the methylene molecule are also presented.

## 1 Introduction

### 1.1 Multireference perturbation theory

The development of multireference methods represents important progress in electronic structure theory in the last two decades. The Hartree-Fock (HF) method has played a very important role and various single reference based methods starting with the HF wavefunction have been successful. However, for the states in the bond-breaking region or for excited states, the wavefunctions are basically of a multiconfigurational nature. Multireference methods are indispensable for such states.

The complete active space self-consistent field method (CASSCF)<sup>1,2</sup> is an attempt to generalize the HF model to a situation where the multiconfigurational nature is important, while keeping the conceptual simplicity of the HF model as much as possible, and is often used to study potential energy surfaces of chemical reactions. In fact, the CASSCF method has many advantages: (1) it can be applied to excited state as well as the ground state in a single framework; (2) it is size-consistent, and (3) it is well defined on the whole potential energy surface of a chemical reaction if an appropriate active space is chosen. However, CASSCF takes into account only non-dynamic electron correlation and not dynamic correlation. Accuracy in the energy calculation such as excitation energy and dissociation energy does not reach chemical accuracy, that is, within several kcal/mol. A method is required that takes into account both the non-dynamic and dynamic correlations.

There are three approaches in including dynamic correlation based on the CASSCF

method, similar to the single reference methods based on the HF method: multireference configuration interaction (CI), coupled cluster (CC) methods, and perturbation theory (PT).

The most popular is probably the multireference configuration interaction (MRCI) method. Much effort has been devoted to developing the method to make it applicable to larger systems and to increase its efficiency. Calculations can now be performed involving a billion configurations. The accuracy of potential energy surfaces reaches a few kcal/mol for *very small* molecule. This is very useful in constructing a potential surface of small molecules. However, for realistically sized molecules, even today the MRCI is still not practical due to the huge dimension of the CI expansion.

The multireference coupled cluster (MRCC) method has also been studied for many years. Although the cluster expansion method has attractive features such as size-consistency, it has not yet been fully established and there are no program packages readily available.

The multireference perturbation theory (MRPT) is a third technique for including dynamic correlation. At present, MRPT has three forms. The first form is the conventional quasidegenerate perturbation theory (QDPT)<sup>3</sup>, where an effective Hamiltonian is constructed on a pure configuration state function (CSF) basis and then is diagonalized to obtain the energies of the states of interest. The effective Hamiltonian to second order (if unitary normalization is applied) is given by

$$\left(K_{eff}^{(0-2)}\right)_{AB} = \langle A|H|B\rangle + \frac{1}{2} \sum_I \left\{ \frac{\langle A|V|I\rangle\langle I|V|B\rangle}{E_B^{(0)} - E_I^{(0)}} + \frac{\langle B|V|I\rangle\langle I|V|A\rangle}{E_A^{(0)} - E_I^{(0)}} \right\}, \quad (1)$$

where  $H$  and  $V$  denote Hamiltonian and perturbation operators, and  $|A\rangle$  and  $|B\rangle$  indicate the CSFs in the active space,  $|I\rangle$  the CSFs in the outer space, and  $E_A^{(0)}$  are zeroth order energies of the CSFs. This method has also often been called “multiconfigurational perturbation theory”. However, each element of the effective Hamiltonian is calculated with a *CSF basis*; in other words, all reference functions in the perturbation calculation are single-configurational, and they are mixed *after* the diagonalization of the effective Hamiltonian. Thus, the conventional QDPT should be classified as a *single-configuration* basis *multi-state* PT, to distinguish it from under-mentioned PTs. The history of the conventional QDPT is long and dates back to 1950s. In spite of its long history it is not often used by chemists, most probably because QDPT has convergence difficulty because of the intruder states and thus it is not always stable.

The second form is Rayleigh-Schrödinger perturbation theory (RSPT) based on a multiconfigurational reference function<sup>4,5</sup>. The multireference Møller-Plesset perturbation theory (MRMP PT)<sup>5</sup> we proposed previously is one of the MRPTs of this form, where CI coefficients as well as molecular orbitals (MOs) are determined first through CASSCF and then RSPT is applied using the CASSCF wave function as reference. The energy to second order is given by

$$E^{(0-2)} = \langle \alpha|H|\alpha\rangle + \sum_I \frac{\langle \alpha|V|I\rangle\langle I|V|\alpha\rangle}{E_\alpha^{(0)} - E_I^{(0)}}, \quad (2)$$

where  $|\alpha\rangle$  is a CASSCF wave function. In this case, a single wave function that is multi-configurational is used as reference. Thus, this PT is classified as a *multi-configuration basis single-state* PT.

We have also proposed a perturbation theory unifying the above two kinds of PTs, a quasidegenerate perturbation theory with MCSCF reference functions (MC-QDPT)<sup>6</sup>, which is a *multi-configuration basis multi-state* PT. In this PT, state-averaged CASSCF is first performed to set reference functions, then an effective Hamiltonian is constructed, which finally is diagonalized to obtain the energies of interest. The effective Hamiltonian to second order is given by

$$\left(K_{eff}^{(0-2)}\right)_{\alpha\beta} = \langle\alpha|H|\beta\rangle + \frac{1}{2} \sum_I \left\{ \frac{\langle\alpha|V|I\rangle\langle I|V|\beta\rangle}{E_\beta^{(0)} - E_I^{(0)}} + \frac{\langle\beta|V|I\rangle\langle I|V|\alpha\rangle}{E_\alpha^{(0)} - E_I^{(0)}} \right\}. \quad (3)$$

Differing from the conventional QDPT case, the reference functions  $|\alpha\rangle$  and  $|\beta\rangle$  are multiconfigurational. This theory includes MRMP PT (the case that the set of reference functions reduces to a single function) and the conventional QDPT (the case that each multiconfigurational reference function reduces to a single CSF) as special cases.

MC-QDPT was designed to treat a relatively small number of states ( $\sim 20$ ) while in the conventional QDPT the number of states can reach several thousands or more. The codes needed for efficient computation for respective cases are quite different. Although Eq. (3) includes Eq. (1) as a special case, the definition of orbital energies for the conventional QDPT is somewhat different from that of MC-QDPT and MRMP PT. Therefore, the gradients in the conventional QDPT are treated separately in the present article.

Our previous applications<sup>7</sup> have been proved that MRMP PT and MC-QDPT are powerful tools for investigating electronic structures of molecules and potential energy surfaces of chemical reactions. In fact, they are accurate enough to provide chemical accuracy. Although MRMP PT and MC-QDPT to second order do not yield a total energy very close to the exact one, they are well balanced and relative energies like dissociation energies, excitation energies, activation energies, etc. are quite good. Furthermore, they are much more efficient and accessible than MRCI and MRCC methods. The effective Hamiltonian is computed as sums of the product of molecular integrals and coupling constants between a reference state and a CSF divided by energy difference, so the computer resources required do not depend strongly on the size of the active space and basis set. This presents a stark contrast to the case of MRCI and MRCC.

## 1.2 Analytic energy derivatives of multireference perturbation theory

The analytic energy derivative method<sup>8</sup> is an essential technique in modern electronic structure theory. Geometry optimization is a key step in studying chemical reaction mechanisms and molecular structures, and normal mode vibrational analysis is utilized to identify infrared and Raman spectra. The required information is obtained from the first and second derivatives with respect to nuclear coordinates of a molecule.

Analytic gradient methods have been developed for several levels of theory beyond Hartree-Fock. The multiconfigurational self-consistent field (MCSCF) method is a method analytic gradients of which was developed early in the history of the research of the energy derivative method<sup>9</sup>. Now CASSCF, a form of MCSCF, is often used to locate stationary points<sup>8</sup>, since the CASSCF wave function is stable even at distorted molecular structures. However, as was mentioned in the previous sub-section, the CASSCF potential surfaces do not reach chemical accuracy, so that it can predict, for example, a fictitious transition state that disappears upon the addition of dynamic correlation. Moreover, it is difficult to use fully optimized CASSCF wave functions for a pure excited electronic state, that is, a state that is not the lowest of its spin and symmetry, since the CASSCF iterative procedure for such cases frequently diverges. There is a need for an analytic gradient for a multireference wave function that goes beyond CASSCF and that is efficient enough to be broadly applicable.

There are few studies involving numerical examples on the gradient for a multireference wave function. Only the multireference CI energy gradient<sup>10</sup> has been implemented. However, as described in the previous section, the costs are generally so high that MRCI gradients are not yet a general practical alternative of CASSCF one.

In the present article, we present a derivation of the analytic gradients for second-order MC-QDPT. Although we presented a derivation of the analytic gradient in a previous publication<sup>11</sup>, a limitation was put on the derivation, that the set of reference functions be identical to that of the (state-averaged) CASSCF wave functions. This setting of reference functions is often not the case. For example, we use *one* of the state-averaged CASSCF functions as reference, or we use molecular orbitals obtained with a singlet CASSCF calculation for triplet states. In the present article, the gradient method is extended to more general reference settings.

Several methods have been developed for the derivation of efficient formulas for the gradient and higher-order energy derivatives for molecular wave functions. In the mid 1980s, Handy and Schaefer proposed a method now called the Z-vector method<sup>12</sup>, designed to avoid solving time-consuming coupled-perturbed (CP) equations, such as the CP Hartree-Fock (CPHF) equation, in computing gradients of configuration interaction (CI) energies. The Lagrange multiplier method in the response function formalism<sup>13</sup>, developed by Jørgensen, Helgaker, and their co-workers, extends the applicability of the Z-vector method to any order of energy derivatives. This method minimizes the number of coupled perturbed equations in a systematic way, and derives the formulas that automatically satisfy the  $(2n+1)$  rule<sup>14</sup> for the perturbation (nuclear displacements). Moreover, with this method one can also systematically treat many constraining conditions.

In the response function formalism derivation for configuration interaction, coupled cluster, and Møller-Plesset perturbation methods, the second-quantization formalism, as well as the unitary exponential forms for parameter relaxation in the OMO (orthonormal molecular orbital) basis, proved to be a useful way to express the energy and the

constraining equations for the MO and CI coefficients<sup>13,15</sup>. This formalism avoids using redundant parameters, and results in simple formulas for variational (and some nonvariational) wave function methods for which the energy expressions are relatively simple. However, for the nonvariational MC-QDPT method, the energy expression is more complex, so the second-quantization formalism for parameter relaxation complicates, rather than simplifies, the derivation. Therefore it is not used in our derivation.

The contents of the present article is as follows. In Section 2, the Lagrangian multiplier method in the response function formalism is briefly reviewed. In Section 3 the MC-QDPT Lagrangian is defined. In Section 4 the linear equations for determining the Lagrangian multipliers, which are necessary for the gradient calculation used in later sections, are derived. In Section 5 the method used to obtain the gradients is discussed. In Section 6 the derivation of gradients is repeated for the conventional QDPT. Conclusions are given in Section 7.

## 2 Lagrange multiplier method

We review very briefly the Lagrange multiplier method used in the present article. The details of the method have been described elsewhere<sup>13</sup>.

The Lagrangian is defined by

$$L(X, \zeta, C) = W(X, C) + \zeta e(X, C), \quad (4)$$

where  $X$  is a nuclear coordinate,  $C$  represents the molecular orbital, configuration interaction coefficients and other parameters,  $W$  is the energy,  $e$  represents constraints on the parameters  $C$ , and  $\zeta$  is the Lagrangian multiplier. The parameter  $C$  in Eq. (4) is determined by

$$e(X, C) = 0. \quad (5)$$

On the other hand,  $\zeta$  is arbitrary since  $e(X, C)$  is identically zero in Eq. (4). Thus, we may place any constraining condition on  $\zeta$ . In the response function formalism,  $\zeta$  is determined such that the first derivative of the Lagrangian with respect to the  $C$ s is zero,

$$\frac{\partial L}{\partial C} = \frac{\partial W}{\partial C} + \zeta \frac{\partial e}{\partial C} = 0. \quad (6)$$

Differentiating Eqs. (5) and (6) with respect to  $X$ , we obtain the equations for determining the derivatives of  $C$ ,

$$\frac{d^n}{dX^n} e = 0 \quad (n = 0, 1, 2, \dots), \quad (7)$$

and those for determining the derivative of  $\zeta$ ,

$$\frac{d^n}{dX^n} \left( \frac{\partial W}{\partial C} + \zeta \frac{\partial e}{\partial C} \right) = 0 \quad (n = 0, 1, 2, \dots). \quad (8)$$

Eq. (7) leads to

$$e = 0, \quad (9)$$

$$\frac{d}{dX} e = \frac{dC}{dX} \frac{\partial e}{\partial C} + \frac{\partial e}{\partial X} = 0, \quad (10)$$

$$\frac{d^2}{dX^2} e = \frac{d^2 C}{dX^2} \frac{\partial e}{\partial C} + \left( \frac{\partial}{\partial X} + \frac{dC}{dX} \frac{\partial}{\partial C} \right)^2 e = 0, \quad (11)$$

..... ,

and Eq. (8) to

$$\frac{\partial W}{\partial C} + \zeta \frac{\partial e}{\partial C} = 0, \quad (12)$$

$$\frac{d}{dX} \left( \frac{\partial W}{\partial C} + \zeta \frac{\partial e}{\partial C} \right) = \left( \frac{\partial}{\partial X} + \frac{dC}{dX} \frac{\partial}{\partial C} \right) \left( \frac{\partial W}{\partial C} + \zeta \frac{\partial e}{\partial C} \right) + \frac{d\zeta}{dX} \frac{\partial e}{\partial C} = 0, \quad (13)$$

..... .

Using Eqs. (9)-(13), the derivatives of the Lagrangian with respect to  $X$ ,

$$\frac{dL}{dX} = \frac{\partial}{\partial X} (W + \zeta e) + \frac{dC}{dX} \left( \frac{\partial W}{\partial C} + \zeta \frac{\partial e}{\partial C} \right) + \frac{d\zeta}{dX} e, \quad (14)$$

$$\begin{aligned} \frac{d^2 L}{dX^2} &= \left( \frac{\partial}{\partial X} + \frac{dC}{dX} \frac{\partial}{\partial C} \right)^2 (W + \zeta e) + \frac{d^2 C}{dX^2} \left( \frac{\partial W}{\partial C} + \zeta \frac{\partial e}{\partial C} \right) \\ &\quad + \frac{d^2 \zeta}{dX^2} e + 2 \frac{d\zeta}{dX} \frac{de}{dX}, \end{aligned} \quad (15)$$

$$\begin{aligned} \frac{d^3 L}{dX^3} &= \left( \frac{\partial}{\partial X} + \frac{dC}{dX} \frac{\partial}{\partial C} \right)^3 (W + \zeta e) + \frac{d^3 C}{dX^3} \left( \frac{\partial W}{\partial C} + \zeta \frac{\partial e}{\partial C} \right) \\ &\quad + 3 \frac{d^2 C}{dX^2} \frac{d}{dX} \left( \frac{\partial W}{\partial C} + \zeta \frac{\partial e}{\partial C} \right) + \frac{d^3 \zeta}{dX^3} e + 3 \frac{d^2 \zeta}{dX^2} \frac{de}{dX} \\ &\quad + 3 \frac{d\zeta}{dX} \left( \frac{\partial}{\partial X} + \frac{dC}{dX} \frac{\partial}{\partial C} \right)^2 e, \end{aligned} \quad (16)$$

$$\frac{d^4 L}{dX^4} = \dots, \quad (17)$$

..... ,

reduce to more compact forms,

$$\frac{dL}{dX} = \frac{\partial}{\partial X} (W + \zeta e), \quad (18)$$

$$\frac{d^2 L}{dX^2} = \left( \frac{\partial}{\partial X} + \frac{dC}{dX} \frac{\partial}{\partial C} \right)^2 (W + \zeta e), \quad (19)$$

$$\frac{d^3 L}{dX^3} = \left( \frac{\partial}{\partial X} + \frac{dC}{dX} \frac{\partial}{\partial C} \right)^3 (W + \zeta e) + 3 \frac{d\zeta}{dX} \left( \frac{\partial}{\partial X} + \frac{dC}{dX} \frac{\partial}{\partial C} \right)^2 e, \quad (20)$$

$$\begin{aligned} \frac{d^4 L}{dX^4} &= \left( \frac{\partial}{\partial X} + \frac{dC}{dX} \frac{\partial}{\partial C} \right)^4 (W + \zeta e) + 3 \left( \frac{d^2 C}{dX^2} \right)^2 \left( \frac{\partial^2 W}{\partial C^2} + \zeta \frac{\partial^2 e}{\partial C^2} \right) \\ &\quad + 6 \frac{d^2 C}{dX^2} \left[ \left( \frac{\partial}{\partial X} + \frac{dC}{dX} \frac{\partial}{\partial C} \right)^2 \left( \frac{\partial W}{\partial C} + \zeta \frac{\partial e}{\partial C} \right) + 2 \frac{d\zeta}{dX} \left( \frac{\partial}{\partial X} + \frac{dC}{dX} \frac{\partial}{\partial C} \right) \frac{\partial e}{\partial C} \right] \\ &\quad + 4 \frac{d\zeta}{dX} \left( \frac{\partial}{\partial X} + \frac{dC}{dX} \frac{\partial}{\partial C} \right)^3 e, \end{aligned}$$

(21)

For the first derivatives, neither the first derivatives of the parameters  $dC/dX$  nor those of Lagrange multipliers  $d\zeta/dX$  are necessary. The gradients of energy  $W$  may be written in terms of only the zeroth derivative parameters  $C$ , the zeroth derivative multipliers  $\zeta$  determined by Eq. (18), and the first derivative molecular integrals.

### 3 MC-QDPT Lagrangian

In this section we derive the Lagrangian of second-order MC-QDPT. Our previous derivation<sup>11</sup> contained a limitation that reference functions be identical to the state averaged CASSCF functions. This limitation is, however, too strong. In fact, it is not the case in some applications. For example, in an excited state calculation, molecular orbitals are often determined with state-averaged CASSCF to avoid the convergence difficulty with excited states and some of the solutions are used as reference functions in perturbation calculation, while in triplet state calculations, the molecular orbitals optimized for singlet states are employed. The Lagrangian in this section is more general than previous one (reference 11): the reference functions are solutions of some CI Hamiltonian that are independent of the state-averaged CASSCF functions. This makes computation a little complicated, but does not give rise to a serious difficulty.

#### 3.1 The MC-QDPT energy expression to the second order

The effective Hamiltonian to second order in MC-QDPT is given<sup>6</sup> by

$$(K_{eff})_{\alpha\beta} = \langle \alpha | H | \beta \rangle + \frac{1}{2} \left\{ \sum_I \frac{\langle \alpha | V | I \rangle \langle I | V | \beta \rangle}{E_\alpha^{(0)} - E_I^{(0)}} + (\alpha \leftrightarrow \beta) \right\}, \quad (22)$$

where  $\{|I\rangle\}$  is the set of all singly and doubly excited configurations from the reference configurations in the complete active space (CAS). (Eq. (22) is equivalent to Eq. (3).) The reference functions  $|\alpha\rangle$  and  $|\beta\rangle$  are CAS-CI eigenfunctions, and the notation  $\alpha \leftrightarrow \beta$  means interchange  $\alpha$  with  $\beta$  from the first term in curly brackets. The first term on the right-hand side (rhs) of Eq. (22) is a diagonal matrix, diagonal elements of which are the CAS-CI energies. If we use the eigenvectors  $\{D_\alpha\}$  of the effective Hamiltonian and substitute the second-quantized operator,

$$\begin{aligned} V &= \sum_{pq} (h_{pq} - \varepsilon_p \delta_{pq}) E_{pq} + \frac{1}{2} \sum_{pqrs} (pq | rs) E_{pq,rs} \\ &= \sum_{pq} v_{pq} E_{pq} + \frac{1}{2} \sum_{pqrs} (pq | rs) E_{pq,rs}, \end{aligned} \quad (23)$$

the energy expression  $W$  is given by

$$W = \sum_{\alpha\beta} D_\alpha D_\beta (K_{eff})_{\alpha\beta} / \sum_\alpha D_\alpha^2$$

$$\begin{aligned}
&= \frac{1}{2} \sum_{\alpha\beta} D_\alpha D_\beta \left\{ \langle \alpha | H | \beta \rangle - \sum_{pq,B} \langle \alpha | E_{pq} | B \rangle R_B(\beta) \sum_e \frac{(p|v|e)(e|v|q)}{\varepsilon_e - \varepsilon_q + \Delta E_{B\alpha}} \right. \\
&\quad - \sum_{pqrs,B} \langle \alpha | E_{pq,rs} | B \rangle R_B(\beta) \left[ \sum_e \frac{(p|v|e)(eq|rs)}{\varepsilon_e - \varepsilon_q + \varepsilon_r - \varepsilon_s + \Delta E_{B\alpha}} \right. \\
&\quad \left. \left. + \sum_e \frac{(pe|rs)(e|v|q)}{\varepsilon_e - \varepsilon_q + \Delta E_{B\alpha}} + \frac{1}{2} \sum_{(a,b)} \frac{(pa|rb)(aq|bs)}{\varepsilon_a - \varepsilon_q + \varepsilon_b - \varepsilon_s + \Delta E_{B\alpha}} \right] \right. \\
&\quad \left. - \sum_{pqrstu,B} \langle \alpha | E_{pq,rs,tu} | B \rangle R_B(\beta) \sum_e \frac{(pe|rs)(eq|tu)}{\varepsilon_e - \varepsilon_q + \varepsilon_t - \varepsilon_u + \Delta E_{B\alpha}} + (\alpha \leftrightarrow \beta) \right\} \\
&\quad \left/ \sum_\alpha D_\alpha^2 \right. .
\end{aligned} \tag{24}$$

In Eq. (24),  $B$  refers to a configuration state function (CSF) in the reference wave function,  $R_B(\beta)$  are CAS-CI coefficients for CSF  $B$  in reference state  $\beta$ ,  $E_{pq}$  is a unitary group generator,

$$E_{pq} = \sum_{\sigma=\alpha,\beta} a_{p\sigma}^+ a_{q\sigma}, \tag{25}$$

and

$$E_{pq,rs} = E_{pq} E_{rs} - \delta_{qr} E_{ps} = \sum_{\sigma\sigma'} a_{p\sigma}^+ a_{r\sigma'}^+ a_{s\sigma} a_{q\sigma}, \tag{26}$$

$$E_{pq,rs,tu} = E_{pq,rs} E_{tu} - \delta_{qt} E_{pu,rs} - \delta_{st} E_{pq,ru} = \sum_{\sigma\sigma'\sigma''} a_{p\sigma}^+ a_{r\sigma'}^+ a_{t\sigma''}^+ a_{u\sigma''} a_{s\sigma'} a_{q\sigma}. \tag{27}$$

$\varepsilon_p$  are orbital energies. Active orbitals are indicated by the indices  $p, q, r, s, t, u$ , while the external (virtual) orbitals are indicated by indices  $e, f$ . The symbol  $(a, b)$  in the summation in the third term in the curly bracket means that  $a$  and  $b$  run over both active and external orbitals, but that  $a$  and  $b$  cannot both be external orbitals simultaneously.  $\Delta E_{B\alpha}$  is the difference between the energies of the zeroth-order state  $\alpha$  and CSF  $B$ ,

$$\Delta E_{B\alpha} = E_B^{(0)} - E_\alpha^{(0)}. \tag{28}$$

In the present article we employ an energy formula omitting for simplicity the doubly occupied orbitals. While it is straightforward to include these in practice, they unnecessarily complicate the derivation.

### 3.2 The constraining conditions for determining the parameters

The calculation of MC-QDPT energies has four steps:

- (i) the determination of molecular orbitals,
- (ii) the determination of the reference CAS-CI wave functions,
- (iii) the construction of the effective Hamiltonian, and
- (iv) the diagonalization of the effective Hamiltonian.

In Steps (i), (ii), and (iv) parameters (molecular orbital coefficients, CAS-CI coefficients and CAS-CI energies, and eigenvector of the effective Hamiltonian and the total energy) are



defined. The equation determining the parameters are used as constraining conditions in the Lagrange multiplier method. In this subsection we list these conditions.

### 3.2.1 Determination of molecular orbitals

The molecular orbitals are determined *uniquely* by (1) state-averaged CASSCF equations, and (2) the orbital canonicalization condition with (3) the orthonormal condition assumed implicitly.

(1) The state-averaged CASSCF equations

The variational condition for the CI coefficients is the secular equation,

$$\sum_D (H_{CD} - \delta_{CD} E^{\text{CAS}}(\gamma)) C_D(\gamma) = 0 \quad \text{and} \quad \sum_C C_C^2(\gamma) = 1. \quad (29)$$

and that for the MO coefficients is a symmetry condition on the matrix  $x_{pq}$ ,

$$x_{pq} = x_{qp}, \quad (30)$$

where  $x_{pq}$  is defined by

$$\begin{aligned} x_{pq} &= \sum_{\gamma} w^{\text{CAS}}(\gamma) x_{pq}(\gamma) \\ &= \sum_{\gamma} w^{\text{CAS}}(\gamma) \left( \sum_i h_{pi} \langle \gamma | E_{qi} | \gamma \rangle + \sum_{ijk} (pi | jk) \langle \gamma | E_{qi,jk} | \gamma \rangle \right), \quad (31) \\ (w^{\text{CAS}}(\gamma): \text{the average weight of the } \gamma\text{th state; } |\gamma\rangle &= \sum_C C_C(\gamma) |C\rangle) \end{aligned}$$

Note that the CI coefficients obtained through these equation as well as the following conditions (2) and (3) are independent of the *reference* CI coefficients.

(2) The orbital canonicalization condition

Following the CASSCF optimization, canonicalization removes the rotational freedom of the CASSCF orbitals within the doubly occupied, active, and external orbital subspaces.

For the doubly occupied, active, and external diagonal block ( $D$ ),

$$F_{pq} = (p | h | q) + \sum_{rs} D_{rs}^{\text{CAS}} [(pq | rs) - (pr | sq)/2] = 0 \quad (p > q \in D), \quad (32)$$

where  $D_{rs}^{\text{CAS}}$  is a averaged one-particle density matrix for the CASSCF states,

$$D_{rs}^{\text{CAS}} = \sum_{\gamma} w^{\text{CAS}}(\gamma) \langle \gamma | E_{rs} | \gamma \rangle. \quad (33)$$

(3) The orthonormalization condition for the molecular orbitals

The orthonormality condition,

$$(p | q) = \delta_{pq} \quad (p \geq q), \quad (34)$$

is assumed implicitly for the molecular orbitals. It must be included in the Lagrangian explicitly.

### 3.2.2 Determination of the reference functions

After the determination of the MOs, they are then used to set reference functions. The CI coefficients of the reference functions are determined by the CAS-CI equation and the

normalization condition,

$$\sum_B (H_{AB} - \delta_{AB} E^{REF}(\alpha)) R_B(\alpha) = 0 \text{ and } \sum_A R_A^2(\alpha) = 1. \quad (35)$$

We do not put the limitation that the set of reference functions is not identical to that of CASSCF functions as mentioned above. Thus Eqs. (29) and (35) are treated as independent conditions.

### 3.2.3 Definition of the orbital energies

Using the averaged one-particle density for the reference functions,

$$D_{rs}^{REF} = \sum_{\alpha} w^{REF} \langle \alpha | E_{rs} | \alpha \rangle, \quad (36)$$

$$(w^{REF}(\alpha): \text{the average weight of the } \alpha\text{th state; } |\alpha\rangle = \sum_A R_A(\alpha) |A\rangle)$$

(note that not the CASSCF density  $D_{rs}^{CAS}$ ), the orbital energies are computed by

$$\varepsilon_p = (p | h | p) + \sum_{rs} D_{rs}^{REF} [(pp | rs) - (pr | sp)/2]. \quad (37)$$

This definition of orbital energies is also used as the constraining condition for *parameters*  $\varepsilon_p$ .

### 3.2.4 The diagonalization condition of the effective Hamiltonian

In the last step of the MC-QDPT calculation, after the construction of the effective Hamiltonian, we diagonalize it to obtain the energy of the target state. This condition is expressed by

$$\sum_{\beta} \left\{ (K_{eff})_{\alpha\beta} - \delta_{\alpha\beta} E \right\} D_{\beta} = 0 \text{ and } \sum_{\alpha} D_{\alpha}^2 = 1, \quad (38)$$

The total energy  $E$  is a parameter constrained by the above equations. Note that  $E$  is a parameter that arises from the diagonalization condition, while  $W$  (see Eq. (24)) is a function of the parameters and nuclear coordinates.

### 3.2.5 The definition of some auxiliary parameters

Now all the parameters that appear in the energy expression, Eq. (24), is defined. We further introduce the following auxiliary parameters:

$$u_{pq} = (p | h | q) = \sum_{\mu\nu} c_{\mu p} c_{\nu q} (\mu | h | \nu) \quad (39)$$

$$= (p | v | q) \quad (p \neq q),$$

$$g_{pqrs} = (pq | rs) = \sum_{\mu\nu\rho\sigma} c_{\mu p} c_{\nu q} c_{\rho r} c_{\sigma s} (\mu\nu | \rho\sigma). \quad (40)$$

and

$$\Delta E_{B\alpha} = \sum_p \left\{ \langle B | E_{pp} | B \rangle - \langle \alpha | E_{pp} | \alpha \rangle \right\} \varepsilon_p. \quad (41)$$

These equations may then be used as constraining conditions that determine the *parame-*

ters  $u_{pq}$ ,  $g_{pqrs}$ , and  $\Delta E_{B\alpha}$ .

### 3.3 The MC-QDPT Lagrangian

Using the energy expression Eq. (24) and the constraining conditions listed in the previous subsection, we may write the Lagrangian,

$$\begin{aligned}
L = & \sum_{\alpha\beta} D_\alpha D_\beta (K_{eff})_{\alpha\beta} / \sum_{\alpha} D_\alpha^2 \\
& + \sum_{C\gamma} \zeta_{CAS,CI}^{C\gamma} \left\{ \sum_D (H_{CD} - \delta_{CD} E^{CAS}(\gamma)) C_D(\gamma) \right\} + \sum_{\gamma} \zeta_{CAS,E}^{\gamma} \left\{ 1 - \sum_C C_C^2(\gamma) \right\} \\
& + \sum_{p>q \in O} \zeta_{CAS,MO}^{pq} (x_{pq} - x_{qp}) \\
& + \sum_{p>q \in D} \zeta_{CAS,MO}^{pq} \left\{ (p|h|q) + \sum_{rs} D_{rs}^{CAS} [(pq|rs) - (pr|sq)/2] \right\} \\
& + \sum_{p \geq q} \zeta_S^{pq} \left\{ (p|q) - \delta_{pq} \right\} \\
& + \sum_{A\alpha} \zeta_{REF,CI}^{A\alpha} \left\{ \sum_B (H_{AB} - \delta_{AB} E^{REF}(\alpha)) R_B(\alpha) \right\} + \sum_{\alpha} \zeta_{REF,E}^{\alpha} \left\{ 1 - \sum_A R_A^2(\alpha) \right\} \\
& + \sum_p \zeta_{OE}^p \left\{ (p|h|p) + \sum_{rs} D_{rs}^{REF} [(pp|rs) - (pr|sp)/2] - \epsilon_p \right\} \\
& + \sum_{\alpha} \zeta_K^{\alpha} \left\{ \sum_{\beta} [(K_{eff})_{\alpha\beta} - \delta_{\alpha\beta} E] D_{\beta} \right\} + \zeta_K \left( 1 - \sum_{\alpha} D_{\alpha}^2 \right) \\
& + \sum_{B\alpha} \zeta_{\Delta E}^{B\alpha} \left\{ \sum_p [\langle B|E_{pp}|B\rangle - \langle \alpha|E_{pp}|\alpha\rangle] \epsilon_p - \Delta E_{B\alpha} \right\} \\
& + \sum_{pq} \zeta_u^{pq} \left\{ (p|h|q) - u_{pq} \right\} + \sum_{pqrs} \zeta_g^{pqrs} \left\{ (pq|rs) - g_{pqrs} \right\}.
\end{aligned} \tag{42}$$

The energy expression  $W$  is now expressed only in terms of the molecular orbital coefficients, (CAS-)CI coefficients,  $u_{pq}$ ,  $g_{pqrs}$  etc., but not explicitly in terms of the molecular integrals. In other words, it does not depend on the nuclear coordinates explicitly.

## 4 The linear equations for the Lagrange multipliers

The first step of the gradient calculation is to solve the linear equation for determining the Lagrangian multipliers  $\zeta$ s,

$$\frac{\partial W}{\partial C} = 0 \quad \leftrightarrow \quad \frac{\partial W}{\partial C} + \zeta \frac{\partial e}{\partial C} = 0. \tag{43}$$

This may be decoupled into eight sets of equations corresponding to the step-wise wave function determination described in Subsection. 3.2.

#### 4.1 The multipliers for diagonalization of the effective Hamiltonian

Since the effective Hamiltonian is diagonalized exactly, the energy is stationary with respect to changes in all  $D_\alpha$ ,

$$\partial W / \partial D_\alpha = 0 \text{ and } \partial W / \partial E = 0. \quad (44)$$

The solutions for these linear equations are clearly

$$\zeta_u^\alpha = 0 \text{ and } \zeta_K = 0. \quad (45)$$

#### 4.2 The multipliers for the one- and two-electron integrals

The multipliers for the one- and two-electron integrals are obtained with simple partial differentiation as

$$\zeta_u^{ab} = \partial W / \partial u_{ab} = \sum_\alpha D_\alpha^2 \langle \alpha | E_{ab} | \alpha \rangle + 2d_{ab}^{W(2)}, \quad (46)$$

$$\zeta_g^{abcd} = \partial W / \partial g_{abcd} = \frac{1}{2} \sum_\alpha D_\alpha^2 \langle \alpha | E_{ab,cd} | \alpha \rangle + 2D_{abcd}^{W(2)}, \quad (47)$$

where the first terms on the rhs of Eqs. (46) and (47) are the one- and two-electron density for the zeroth- plus first-order energy, respectively, and  $d_{ab}^{W(2)}$  and  $D_{abcd}^{W(2)}$  indicate the one- and two-electron densities in the MO basis for the second-order energy,

$$\begin{aligned} d_{ab}^{W(2)} &= \partial W^{(2)} / \partial u_{ab} \\ &= \frac{1}{4} \sum_{\alpha\beta} D_\alpha D_\beta \left\{ - \sum_{pq,B} \langle \alpha | E_{pq} | B \rangle R_{B\beta} \sum_e \frac{\delta_{pa} \delta_{eb} u_{eq} + u_{pe} \delta_{ea} \delta_{qb}}{\epsilon_e - \epsilon_q + \Delta E_{B\alpha}} \right. \\ &\quad - \sum_{pqrs,B} \langle \alpha | E_{pq,rs} | B \rangle R_{B\beta} \left[ \sum_e \frac{\delta_{pa} \delta_{eb} g_{eqrs}}{\epsilon_e - \epsilon_q + \epsilon_r - \epsilon_s + \Delta E_{B\alpha}} \right. \\ &\quad \left. \left. + \sum_e \frac{g_{pers} \delta_{ea} \delta_{qb}}{\epsilon_e - \epsilon_q + \Delta E_{B\alpha}} \right] + (\alpha \leftrightarrow \beta) \right\} \end{aligned} \quad (48)$$

and

$$\begin{aligned} D_{abcd}^{W(2)} &= \partial W^{(2)} / \partial g_{abcd} \\ &= \frac{1}{4} \sum_{\alpha\beta} D_\alpha D_\beta \left\{ - \sum_{pqrs,B} \langle \alpha | E_{pq,rs} | B \rangle R_{B\beta} \left[ \sum_e \frac{u_{pe} \delta_{ea} \delta_{qb} \delta_{rc} \delta_{sd}}{\epsilon_e - \epsilon_q + \epsilon_r - \epsilon_s + \Delta E_{B\alpha}} \right. \right. \\ &\quad + \sum_e \frac{\delta_{pa} \delta_{eb} \delta_{rc} \delta_{sd} u_{eq}}{\epsilon_e - \epsilon_q + \Delta E_{B\alpha}} \\ &\quad \left. \left. + \frac{1}{2} \sum_{(a',b')} \frac{\delta_{pa} \delta_{a'b} \delta_{rc} \delta_{b'd} g_{a'qb's} + g_{pa'rb'} \delta_{a'a} \delta_{qb} \delta_{b'c} \delta_{sd}}{\epsilon_{a'} - \epsilon_q + \epsilon_{b'} - \epsilon_s + \Delta E_{B\alpha}} \right] \right. \\ &\quad \left. + \sum_{pqrsu,B} \langle \alpha | E_{pq,rs,tu} | B \rangle R_{B\beta} \sum_e \frac{\delta_{pa} \delta_{eb} \delta_{rc} \delta_{sd} g_{eqtu} + g_{pers} \delta_{ea} \delta_{qb} \delta_{tc} \delta_{ud}}{\epsilon_e - \epsilon_q + \epsilon_t - \epsilon_u + \Delta E_{B\alpha}} + (\alpha \leftrightarrow \beta) \right\}, \end{aligned} \quad (49)$$

respectively.

To simplify subsequent equations, it is convenient to employ the symmetrized effec-

tive densities:

$$d_{ab}^{E^{(2)}} \leftarrow \frac{1}{2} \left( d_{ab}^{E^{(2)}} + d_{ba}^{E^{(2)}} \right), \quad (50)$$

$$D_{abcd}^{E^{(2)}} \leftarrow \frac{1}{8} \left( D_{abcd}^{E^{(2)}} + D_{abdc}^{E^{(2)}} + D_{bacd}^{E^{(2)}} + D_{badc}^{E^{(2)}} + D_{cdab}^{E^{(2)}} + D_{cdba}^{E^{(2)}} + D_{dcab}^{E^{(2)}} + D_{dcba}^{E^{(2)}} \right). \quad (51)$$

### 4.3 The multipliers for the definition of energy shift

The equation for obtaining the multipliers for the energy shifts  $\Delta E_{B\alpha}$ ,

$$\partial L / \partial \Delta E_{B\alpha} = 0 \quad (52)$$

is easily reduced to a formula that gives the multipliers directly,

$$\begin{aligned} \zeta_{\Delta E}^{B\alpha} &= \partial W / \partial \Delta E_{B\alpha} \\ &= \frac{1}{2} \sum_{\alpha\beta} D_\alpha D_\beta \left\{ \sum_{pq,B} \langle \alpha | E_{pq} | B \rangle R_B(\beta) \sum_e \frac{u_{pe} u_{eq}}{(\epsilon_e - \epsilon_q + \Delta E_{B\alpha})^2} + (\alpha \leftrightarrow \beta) \right\} \\ &\quad + (2\text{-} + 3\text{- body terms}). \end{aligned} \quad (53)$$

The two- and three-body terms in Eq. (53) are readily derived by analogy to the one-body term, since  $\partial / \partial \Delta E_{B\alpha}$  operates only on the energy denominators in the Lagrangian. The structure of the creation-annihilation operators is the same as that for the energy expression Eq. (24), so that the computation is performed using the same approach as for the effective Hamiltonian.

### 4.4 The multipliers for the orbital energies

The multipliers for  $\Delta E_{B\alpha}$  are again given by partial differentiation of the energy expression,

$$\partial L / \partial \epsilon_p = 0 \quad \leftrightarrow \quad \zeta_{OE}^p = - \sum_{B\alpha} \zeta_{\Delta E}^{B\alpha} \left[ \langle B | E_{pp} | B \rangle - \langle \alpha | E_{pp} | \alpha \rangle \right] + \partial W / \partial \epsilon_p. \quad (54)$$

As noted above for the multipliers for  $\Delta E_{B\alpha}$ , the second term in Eq. (54) has the same creation-annihilation operator structure as the effective Hamiltonian, so that the calculation is again straightforward,

$$\begin{aligned} \partial W / \partial \epsilon_m &= \frac{1}{2} \sum_{\alpha\beta} D_\alpha D_\beta \\ &\quad \times \left\{ - \sum_{pq,B} \langle \alpha | E_{pq} | B \rangle R_B(\beta) \sum_e (\delta_{mq} - \delta_{me}) \frac{u_{pe} u_{eq}}{(\epsilon_e - \epsilon_q + \Delta E_{B\alpha})^2} \right. \\ &\quad \left. + (\alpha \leftrightarrow \beta) \right\} + (2\text{-} + 3\text{- body terms}). \end{aligned} \quad (55)$$

### 4.5 The linear equations for the multipliers for the CI coefficients of the reference

functions

Unlike the above multipliers, the multipliers for the CI coefficients of the reference functions  $\zeta_{REF,CI}^{A\alpha}$  and the reference energies  $\zeta_{REF,E}^\alpha$  are determined by  $N$  independent sets of linear equations ( $N$  is the number of the reference states),

$$\begin{cases} \partial L / \partial R_A(\alpha) = 0 \\ \partial L / \partial E^{REF}(\alpha) = 0 \end{cases} \quad (\alpha = 1, 2, \dots, N) \quad (56)$$

Equation (56) may be written in matrix form as,

$$\begin{pmatrix} \mathbf{A}_{REF,REF} & \mathbf{A}_{REF,E} \\ \mathbf{A}_{E,REF} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{x}_{REF,CI} \\ \mathbf{x}_{REF,E} \end{pmatrix} = \begin{pmatrix} \mathbf{b}_{REF,CI} \\ \mathbf{0} \end{pmatrix}. \quad (57)$$

for each  $\alpha$ . The nonzero parts of the coefficient matrix  $\mathbf{A}$  and the vector  $\mathbf{b}$  are given by

$$A_{REF,REF}^{A\alpha,B\alpha} = H_{BA} - \delta_{BA} E^{REF}(\alpha), \quad (58)$$

$$A_{REF,E}^{A\alpha,\alpha} = -2R_A(\alpha), \quad (59)$$

$$A_{E,REF}^{\alpha,B\alpha} = -R_B(\alpha), \quad (60)$$

and

$$\begin{aligned} b_{REF,CI}^{A\alpha} = & -W^{REF}(\alpha) \sum_p \zeta_{OE}^p \sum_{rs} \langle \alpha | E_{rs} + E_{sr} | A \rangle P_{pprs} \\ & + 2R_A(\alpha) E_A^{(0)} \sum_B \zeta_{\Delta E}^{B\alpha} - \partial W / \partial R_A(\alpha) \end{aligned}, \quad (61)$$

where

$$\begin{aligned} \partial W / \partial R_A(\alpha) = & 2D_\alpha^2 E^{REF}(\alpha) R_A(\alpha) \\ & + D_\alpha \sum_\beta D_\beta \left\{ - \sum_{pq} \langle \beta | E_{pq} | A \rangle \sum_e \left[ \frac{u_{pe} u_{eq}}{\epsilon_e - \epsilon_q + \Delta E_{A\beta}} + (\beta \rightarrow \alpha) \right] \right\} \\ & + (2- + 3\text{-body terms}). \end{aligned} \quad (62)$$

The dimension of the linear equations for each  $a$  is the dimension of the reference CAS-CI space plus one.

#### 4.6 The multipliers for the MO rotations in the invariant doubly occupied and external orbital subspaces

The multipliers for the orbital rotations in the doubly occupied and external subspaces are obtained from the following equation,

$$(\partial / \partial U_{ab} - \partial / \partial U_{ba}) L = 0 \quad (a > b \in doc, ext), \quad (63)$$

where

$$c_{\mu i}(X) = \sum_m c_{\mu m}(X_0) U_{mi}. \quad (64)$$

We use  $U_{mi}$  rather than molecular coefficients themselves as in the conventional energy derivative methods. Eq. (63) is written as

$$\sum_{p>q \in D} \zeta_{CAS,MO}^{pq} A^{ab,pq} = b^{ab} \quad (a > b \in doc, ext) \quad (65)$$

where

$$\begin{aligned} A^{ab,pq} &= \varepsilon_a (\delta_{pa} \delta_{qb} + \delta_{qa} \delta_{pb}) + 2 \sum_i D_{ib}^{CAS} P_{pqia} - (a \leftrightarrow b) \\ &= (\varepsilon_a - \varepsilon_b) \delta_{pa} \delta_{qb} \end{aligned} \quad (66)$$

as the matrix elements, and

$$\begin{aligned} b^{ab} &= - \left\{ \sum_{A\alpha} \zeta_{REF,CI}^{A\alpha} \sum_B X_{ab}^{AB} R_B(\alpha) + 2 \sum_i \zeta_{OE}^i \sum_j D_{jb}^{REF} P_{ijja} \right. \\ &\quad \left. + 2 \sum_i \zeta_u^{ib} u_{ia} + 4 \sum_{ijk} \zeta_g^{ijk} (ij | ka) - (a \leftrightarrow b) \right\} \end{aligned} \quad (67)$$

as the vector elements.  $P_{pqrs}$  are the Roothaan-Bagus supermatrix integrals<sup>16</sup>

$$P_{pqrs} = (pq | rs) - (pr | qs)/4 - (ps | rq)/4. \quad (68)$$

Since the coefficient matrix is diagonal for doubly occupied and external subspaces, the  $\zeta_{CAS,MO}^{ab}$  for the doubly occupied and external subspaces can be determined without having to solve linear equations,

$$\zeta_{CAS,MO}^{ab} = b_{ab} / (\varepsilon_a - \varepsilon_b). \quad (a > b \in doc, ext) \quad (69)$$

#### 4.7 The linear equations for the multipliers for the MO rotations and the CI coefficients of the CASSCF functions

The remaining multipliers for the orbital rotation mixing and the multipliers for the CI coefficients are obtained by solving coupled linear equations corresponding to the state-averaged CASSCF equation. The dimension of these linear equations is large, so their solution represents the most time-consuming part of the problem:

$$\begin{cases} (\partial/\partial U_{ab} - \partial/\partial U_{ba})L = 0 & (a > b \in O, act) \\ \partial L/\partial C_c(\gamma) = 0 \\ \partial L/\partial E^{CAS}(\gamma) = 0 \end{cases}, \quad (70)$$

where  $a > b \in O$  means that orbitals  $a$  and  $b$  are in different orbital subspaces (doubly occupied, active, or external). Equation (70) may be written in matrix form as,

$$\begin{pmatrix} \mathbf{A}_{MO,MO} & \mathbf{A}_{MO,CI} & \mathbf{0} \\ \mathbf{A}_{CI,MO} & \mathbf{A}_{CI,CI} & \mathbf{A}_{CI,E} \\ \mathbf{0} & \mathbf{A}_{E,CI} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{x}_{CAS,MO}^{O,act} \\ \mathbf{x}_{CAS,CI} \\ \mathbf{x}_{CAS,E} \end{pmatrix} = \begin{pmatrix} \mathbf{b}_{CAS,MO}^{O,act} \\ \mathbf{b}_{CAS,CI} \\ \mathbf{0} \end{pmatrix}. \quad (71)$$

The coefficient matrix  $\mathbf{A}$  consists of six nonzero parts,

$$A_{MO,MO}^{ab,pq} = \begin{cases} (\delta_{bp} x_{aq} + Y_{pqab} - (p \leftrightarrow q)) - (a \leftrightarrow b) & (p > q \in O) \\ (\varepsilon_a - \varepsilon_b) \delta_{pa} \delta_{qb} + 2 \sum_i (D_{ib}^{CAS} P_{pqia} - D_{ia}^{CAS} P_{pqib}) & (p > q \in act) \end{cases}, \quad (72)$$

$$A_{MO,CI}^{ab,C\gamma} = \sum_D X_{ab}^{CD} C_D(\gamma) - (a \leftrightarrow b), \quad (73)$$

( $a > b \in O, act$ )

$$A_{CI,MO}^{C\gamma,pq} = \begin{cases} w^{CAS}(\gamma) \sum_D C_D(\gamma) (X_{pq}^{CD} - X_{qp}^{CD}) & (p > q \in O) \\ w^{CAS}(\gamma) \sum_{rs} \langle \gamma | E_{rs} + E_{sr} | C \rangle P_{pqrs} & (p > q \in act) \end{cases}, \quad (74)$$

$$A_{CI,CI}^{C\gamma,D\delta} = [H_{DC} - \delta_{DC} E^{CAS}(\gamma)] \delta_{\gamma\delta}, \quad (75)$$

$$A_{CI,E}^{C\gamma,\delta} = -2C_C(\gamma) \delta_{\gamma\delta}, \quad (76)$$

$$A_{E,CI}^{\gamma,D\delta} = -C_D(\gamma) \delta_{\gamma\delta}, \quad (77)$$

where  $X_{ab}^{AB}$  and  $Y_{pqab}$  are defined by

$$X_{pq}^{AB} = \sum_i (p|h|i) [\langle A|E_{qi}|B \rangle + \langle B|E_{qi}|A \rangle] \\ + \sum_{ijk} (pi|jk) [\langle A|E_{qi,jk}|B \rangle + \langle B|E_{qi,jk}|A \rangle], \quad (78)$$

$$Y_{pqab} = \sum_{\gamma} w^{CAS}(\gamma) Y_{pqab}(\gamma) \\ = \sum_{\gamma} w^{CAS}(\gamma) \left\{ (p|h|a) \langle \gamma | E_{qb} | \gamma \rangle + \sum_{mn} [(pa|mn) \langle \gamma | E_{qb,mn} | \gamma \rangle \right. \\ \left. + (pm|an) \langle \gamma | E_{qm,bn} | \gamma \rangle + (pm|an) \langle \gamma | E_{qm,nb} | \gamma \rangle] \right\}. \quad (79)$$

The vector  $\mathbf{b}$  on the rhs of Eq. (71) has two nonzero parts,

$$b_{CAS,MO}^{ab} = - \left\{ \sum_p \zeta_{OE}^p \left( 2\delta_{bp} F_{ap} + 2 \sum_i D_{ib}^{REF} P_{ppia} \right) \right. \\ + \sum_{p>q \in doc, ext} \zeta_{CAS,MO}^{pq} \left( \delta_{bp} F_{aq} + \delta_{bq} F_{ap} + 2 \sum_i D_{ib}^{CAS} P_{pqia} \right) \\ + \sum_{A\alpha} \zeta_{REF,CI}^{A\alpha} \sum_B X_{ab}^{AB} R_B(\alpha) + 2 \sum_i \zeta_u^{ib} u_{ia} + 4 \sum_{ijk} \zeta_g^{ijkb} (ij|ka) \\ \left. - (a \leftrightarrow b) \right\} \quad (80)$$

and

$$b_{CAS,CI}^{C\gamma} = -w^{CAS}(\gamma) \sum_{p>q \in doc, ext} \zeta_{CAS,MO}^{pq} \sum_{rs} \langle \gamma | E_{rs} + E_{sr} | C \rangle P_{pqrs}. \quad (81)$$

#### 4.8 The multipliers for the MO orthonormalization conditions

The final step in solving the linear equations for the multipliers is the computation of the multipliers for the orthonormality conditions of the orbitals,

$$(\partial/\partial U_{ab} + \partial/\partial U_{ba})L = 0, \quad (a \geq b) \quad (82)$$

which is written as

$$\zeta_s^{ab} = -2^{-1} (1 + \delta_{ab})^{-1} \xi_{ab}. \quad (83)$$



The  $\xi_{ab}$  are obtained by changing the sign with respect to interchange of  $a$  with  $b$  from minus to plus in Subsections 4.6-7 and collecting them,

$$\begin{aligned}
\xi_{ab} = & \sum_{\gamma} \sum_C \zeta_{CAS,CI}^{C\gamma} \sum_D X_{ab}^{CD} C_D(\gamma) + \sum_{p>q \in O} \zeta_{CAS,MO}^{pq} (\delta_{bp} x_{aq} + Y_{pqab} - (p \leftrightarrow q)) \\
& + \sum_{p>q \in D} \zeta_{CAS,MO}^{pq} \left( \delta_{bp} F_{aq} + \delta_{bq} F_{ap} + 2 \sum_i D_{ib}^{CAS} P_{pqia} \right) \\
& + \sum_p \zeta_{OE}^p \left( 2\delta_{bp} F_{ap} + 2 \sum_i D_{ib}^{CAS} P_{ppia} \right) \\
& + \sum_{\alpha} \sum_A \zeta_{REF,CI}^{A\alpha} \sum_B X_{ab}^{AB} R_B(\alpha) + 2 \left( \sum_i \zeta_u^{ib} u_{ia} + 2 \sum_{ijk} \zeta_g^{ijk} (ij | ka) \right) \\
& + (a \leftrightarrow b).
\end{aligned} \tag{84}$$

## 5 Molecular energy gradient

The Lagrangian multipliers obtained in the previous section may now be combined with the molecular integrals to compute the molecular energy gradients:

$$\begin{aligned}
\frac{dL}{dX} = & \frac{\partial W}{\partial X} + \zeta \frac{\partial e}{\partial X} = \zeta \frac{\partial e}{\partial X} \\
= & \sum_{C\gamma} \zeta_{CAS,CI}^{C\gamma} \sum_D \frac{\partial H_{CD}}{\partial X} C_D(\gamma) + \sum_{A\alpha} \zeta_{CAS,CI}^{A\alpha} \sum_B \frac{\partial H_{AB}}{\partial X} R_B(\alpha) \\
& + \sum_{p>q \in O} \zeta_{CAS,MO}^{pq} \frac{\partial}{\partial X} (x_{pq} - x_{qp}) + \sum_{p>q \in D} \zeta_{CAS,MO}^{pq} \frac{\partial F_{pq}}{\partial X} \\
& + \sum_p \zeta_{OE}^p \frac{\partial}{\partial X} \left\{ (p|h|p) + \sum_{rs} D_{rs}^{REF} [(pp|rs) - (pr|sp)/2] \right\} \\
& + \sum_{p \geq q} \zeta_s^{pq} (p|q)^X + \sum_{pq} \zeta_u^{pq} (p|v|q)^X + \sum_{pqrs} \zeta_g^{pqrs} (pq|rs)^X \\
= & \sum_{pq} d_h^{pq} (p|h|q)^X + \sum_{pw} d_s^{pq} (p|q)^X + \sum_{pqrs} D_g^{pqrs} (pq|rs)^X,
\end{aligned} \tag{85}$$

where  $(p|h|q)^X$ ,  $(p|q)^X$ , and  $(pq|rs)^X$  are transformed derivative integrals in the MO basis

$$(p|q)^X = \sum_{\mu\nu} c_{\mu p} c_{\nu q} \frac{d}{dX} (\mu|\nu), \tag{86}$$

$$(p|h|q)^X = \sum_{\mu\nu} c_{\mu p} c_{\nu q} \frac{d}{dX} (\mu|h|\nu), \tag{87}$$

$$(pq|rs)^X = \sum_{\mu\nu\rho\sigma} c_{\mu p} c_{\nu q} c_{\rho r} c_{\sigma s} \frac{d}{dX} (\mu\nu|\rho\sigma), \tag{88}$$

and  $d_h^{pq}$ ,  $d_s^{pq}$ , and  $D_g^{pqrs}$  are effective densities for the Lagrangian in the MO basis<sup>17</sup>.

These densities are given by

$$\begin{aligned}
d_h^{pq} = & \sum_{\gamma} \sum_C \zeta_{CAS,CI}^{C\gamma} \langle C | E_{pq} | \gamma \rangle + \sum_{\alpha} \sum_A \zeta_{REF,CI}^{A\alpha} \langle A | E_{pq} | \alpha \rangle \\
& + \sum_{i>j \in O} \zeta_{CAS,MO}^{ij} \sum_{\gamma} w^{CAS}(\gamma) (\delta_{pi} \langle \gamma | E_{jq} | \gamma \rangle - (i \leftrightarrow j)) + \zeta_u^{pq} \\
& + \left\{ \begin{array}{ll} \zeta_{OE}^p & (p = q) \\ \zeta_{CAS,MO}^{pq} & (p > q \in D) \\ 0 & (\text{otherwise}) \end{array} \right\},
\end{aligned} \tag{89}$$

$$d_s^{pq} = \zeta_s^{pq}, \tag{90}$$

and

$$\begin{aligned}
D_g^{pqrs} = & \sum_{\gamma} \sum_C \zeta_{CAS,CI}^{C\gamma} \langle C | E_{pq,rs} | \gamma \rangle / 2 + \sum_{\alpha} \sum_A \zeta_{REF,CI}^{A\alpha} \langle A | E_{pq,rs} | \alpha \rangle / 2 \\
& + \sum_{i>j \in O} \zeta_{CAS,MO}^{ij} \sum_{\gamma} w^{CAS}(\gamma) (\delta_{pi} \langle \gamma | E_{jq,rs} | \gamma \rangle - (i \leftrightarrow j)) \\
& + \sum_{i>j \in D} \zeta_{CAS,MO}^{ij} (\delta_{pi} \delta_{qj} D_{rs}^{CAS} - \delta_{pi} \delta_{sj} D_{qr}^{CAS} / 2) \\
& + \sum_i \zeta_{OE}^i (\delta_{pi} \delta_{qi} D_{rs}^{REF} - \delta_{pi} \delta_{si} D_{pr}^{REF} / 2) \\
& + \zeta_g^{pqrs}.
\end{aligned} \tag{91}$$

These are transformed back into the atomic orbital (AO) basis, and the energy gradients may be obtained by multiplying the densities in the AO basis by the first-derivative integrals,

$$\frac{dL}{dX} = \sum_{\mu\nu} d_h^{\mu\nu} \frac{d}{dX} (\mu | h | \nu) + \sum_{\mu\nu} d_s^{\mu\nu} \frac{d}{dX} (\mu | \nu) + \sum_{\mu\nu\rho\sigma} D_g^{\mu\nu\rho\sigma} \frac{d}{dX} (\mu\nu | \rho\sigma). \tag{92}$$

The following is a summary of the computational steps:

*Step 1* Determination of the energy of the target state

- (a) Compute the (state-averaged) CASSCF wave functions.
- (b) Canonicalize the CASSCF MOs and transform the integrals to the MO basis.
- (c) Compute the reference functions for the canonical Fock MOs by solving CAS-CI (The active space is not necessarily the same as (a)).
- (d) Compute the MC-QDPT effective Hamiltonian and obtain the final energy by diagonalizing it.

*Step 2* Determination of the Lagrange multipliers

- (a) Compute the multipliers  $\zeta_K$ ,  $\zeta_u$ ,  $\zeta_g$ ,  $\zeta_{\Delta E}$ , and  $\zeta_{OE}$  in order. These are simply obtained with partial differentiations.
- (b) Solve linear equations for the multipliers  $\zeta_{REF,CI}$  and  $\zeta_{REF,E}$  for each reference state.
- (c) Compute the multipliers  $\zeta_{CAS,MO}$  for the doubly occupied and external orbital subspace.
- (d) Solve the linear equation for the  $\zeta_{CAS,CI}$  and the remaining part of the  $\zeta_{CAS,MO}$ .

(e) Compute the  $\zeta_s$ .

*Step 3* Evaluation of energy gradients

- (a) Compute the effective densities of the Lagrangian for the overlap, one-, and two-electron integrals on the MO basis.
- (b) Transform the densities back to the AO basis.
- (c) Multiply the densities and the derivative integrals to obtain the gradient, which is looped over all the nuclear coordinates.

## 6 Conventional quasidegenerate perturbation theory (QDPT) case

We could extend the discussions in Sections 3 to 5 to the conventional quasidegenerate perturbation theory (QDPT)<sup>3</sup> case that constructs an effective Hamiltonian in the *CSF* basis, with a slight modification of the MC-QDPT Lagrangian. However, we do not adopt such a line, since the general formulas derived in such a way are so complicated that the codes become inefficient. In the MC-QDPT case, the dimension of the effective Hamiltonian is usually small ( $\sim 20$ ) even for a large CAS, while in the conventional QDPT case, the dimension is equal to that of CAS. Accordingly, the algorithms of the MC- and conventional QDPTs are quite different from each other. Thus we present the formula for gradients of the conventional QDPT separately in this section.

### 6.1 The QDPT molecular gradients

The energy expression Eq. (24) in Section 3.1 reduces to the conventional QDPT one if we set reference functions ( $|\alpha\rangle$ ) as being all the CSFs in the CAS. Some parameters take definite values in that case,

$$R_A(\alpha) = 1, \quad (93)$$

$$\Delta E_{B\alpha} = 0. \quad (94)$$

In this section we treat the case that the molecular orbitals are determined by restricted (open- or closed-shell) Hartree-Fock equation and orbital energies are defined by the OPT1 scheme<sup>18</sup>. In this case the CASSCF optimization of molecular orbitals and the canonicalization scheme are applicable:

$$x_{pq} = x_{qp} \quad (p > q \in O), \quad (95)$$

$$(x_{pq} = \sum_i h_{pi} D_{qi}^{HF} + \sum_{ijk} (pi | ij) D_{qi,ij}^{HF}) \quad (96)$$

$$F_{pq} = (p | h | q) + \sum_{rs} D_{rs}^{HF} [(pq | rs) - (pr | sq)/2] = 0 \quad (p > q \in D) \quad (97)$$

where  $D_{ij}^{HF}$  and  $D_{ij,kl}^{HF}$  indicate the one- and two-particle Hartree-Fock density matrices defined by

$$D_{ij}^{HF} = \langle HF | E_{ij} | HF \rangle = \delta_{ij} \langle HF | E_{ii} | HF \rangle = \begin{cases} 2\delta_{ij} & (i: \text{doubly occupied}) \\ \delta_{ij} & (i: \text{singly occupied}) \\ 0 & (i: \text{virtual}) \end{cases} \quad (98)$$

and

$$D_{ij,kl}^{HF} = D_{ij}^{HF} D_{kl}^{HF} - \delta_{jk} D_{il}^{HF} = \delta_{ij} \delta_{kl} D_{ii}^{HF} D_{kk}^{HF} - \delta_{jk} \delta_{il} D_{ii}^{HF}. \quad (99)$$

The symbol  $D$  stands for the doubly occupied, singly occupied, and virtual diagonal blocks and the symbol  $O$  for the off-diagonal block. The CI equation Eqs. (29) vanishes.

Using these equations we obtain the QDPT Lagrangian,

$$\begin{aligned} L = & \sum_{\alpha\beta} D_\alpha D_\beta (K_{eff})_{\alpha\beta} / \sum_\alpha D_\alpha^2 \\ & + \sum_{p>q \in O} \zeta_{HF,MO}^{pq} (x_{pq} - x_{qp}) \\ & + \sum_{p>q \in D} \zeta_{HF,MO}^{pq} \left\{ (p|h|q) + \sum_{rs} D_{rs}^{HF} [(pq|rs) - (pr|sq)/2] \right\} \\ & + \sum_{p \geq q} \zeta_S^{pq} \left\{ (p|q) - \delta_{pq} \right\} \\ & + \sum_p \zeta_{OE}^p \left\{ (p|h|p) + \sum_{rs} D_{rs}^{HF} [(pp|rs) - (pr|sp)/2] - \varepsilon_p \right\} \\ & + \sum_\alpha \zeta_K^\alpha \left\{ \sum_\beta [(K_{eff})_{\alpha\beta} - \delta_{\alpha\beta} E] D_\beta \right\} + \zeta_K \left( 1 - \sum_\alpha D_\alpha^2 \right) \\ & + \sum_{pq} \zeta_u^{pq} \left\{ (p|h|q) - u_{pq} \right\} + \sum_{pqrs} \zeta_g^{pqrs} \left\{ (pq|rs) - g_{pqrs} \right\}. \end{aligned} \quad (100)$$

The multipliers  $\zeta_K$ ,  $\zeta_u$ ,  $\zeta_g$ , and  $\zeta_{OE}$  are determined in the same manner as the MC-QDPT case:

$$\zeta_K^\alpha = 0 \text{ and } \zeta_K = 0, \quad (101)$$

$$\zeta_u^{ab} = \partial W / \partial u_{ab} = \sum_\alpha D_\alpha^2 \langle \alpha | E_{ab} | \alpha \rangle + 2d_{ab}^{W(2)}, \quad (102)$$

$$\zeta_g^{abcd} = \partial W / \partial g_{abcd} = \frac{1}{2} \sum_\alpha D_\alpha^2 \langle \alpha | E_{ab,cd} | \alpha \rangle + 2D_{abcd}^{W(2)}, \quad (103)$$

and

$$\begin{aligned} \zeta_{OE}^p &= \partial W / \partial \varepsilon_p \\ &= \frac{1}{2} \sum_{\alpha\beta} D_\alpha D_\beta \left\{ - \sum_{pq} \langle \alpha | E_{pq} | \beta \rangle \sum_e (\delta_{mq} - \delta_{me}) \frac{u_{pe} u_{eq}}{(\varepsilon_e - \varepsilon_q)^2} + (\alpha \leftrightarrow \beta) \right\} \\ &+ (2- + 3- \text{ body terms}). \end{aligned} \quad (104)$$

The multipliers for the orbital rotations in the doubly occupied, singly occupied, and virtual subspaces are obtained from the following equation,

$$(\partial / \partial U_{ab} - \partial / \partial U_{ba}) L = 0 \quad (a > b \in D), \quad (105)$$

This is written as

$$\zeta_{HF,MO}^{ab} = -\frac{2}{\varepsilon_a - \varepsilon_b} \left\{ \sum_i \zeta_{OE}^i \sum_j D_{jb}^{HF} P_{ija} + \sum_i \zeta_u^{ib} u_{ia} + 2 \sum_{ijk} \zeta_g^{ikb} (ij | ka) - (a \leftrightarrow b) \right\}. \quad (106)$$

$$(a > b \in D)$$

One difference from the MC-QDPT case is that the multipliers are given for all three diagonal blocks (doubly occupied, singly occupied, and virtual orbital blocks, which correspond to doubly occupied, active, and external orbital blocks, respectively, in the CASSCF case).

The multipliers for the orbital rotation mixing among different subspaces are obtained by solving a linear equation,

$$(\partial/\partial U_{ab} - \partial/\partial U_{ba})L = 0 \quad (a > b \in O). \quad (107)$$

This may be written in matrix form as

$$\mathbf{A}_{MO,MO} \mathbf{x}_{HF,MO}^O = \mathbf{b}_{HF,MO}^O. \quad (108)$$

The coefficient matrix  $\mathbf{A}$  is given by

$$A_{MO,MO}^{ab,pq} = (\delta_{bp} x_{aq} + Y_{pqab} - (p \leftrightarrow q)) - (a \leftrightarrow b), \quad (109)$$

where

$$Y_{pqab} = (p | h | a) D_{qb}^{HF} + \sum_{mn} [(pa | mn) D_{qb,mn}^{HF} + (pm | an) D_{qm,bn}^{HF} + (pm | an) D_{qm,nb}^{HF}], \quad (110)$$

and the vector  $\mathbf{b}$

$$b_{HF,MO}^{ab} = - \left\{ 2 \sum_p \zeta_{OE}^p (\delta_{bp} F_{ap} + D_{bb}^{HF} P_{ppba}) + \sum_{p>q \in D} \zeta_{HF,MO}^{pq} \left( \delta_{bp} F_{aq} + \delta_{bq} F_{ap} + 2 \sum_i D_{ib}^{HF} P_{ppia} \right) + 2 \sum_i \zeta_u^{ib} u_{ia} + 4 \sum_{ijk} \zeta_g^{ikb} (ij | ka) - (a \leftrightarrow b) \right\}. \quad (111)$$

Finally the multipliers for the orthonormality condition of the orbitals are obtained as

$$\zeta_S^{ab} = -2^{-1} (1 + \delta_{ab})^{-1} \xi_{ab}. \quad (112)$$

where

$$\begin{aligned} \xi_{ab} = & \sum_{p>q \in O} \zeta_{HF,MO}^{pq} (\delta_{bp} x_{aq} + Y_{pqab} - (p \leftrightarrow q)) \\ & + \sum_{p>q \in D} \zeta_{HF,MO}^{pq} \left( \delta_{bp} F_{aq} + \delta_{bq} F_{ap} + 2 \sum_i D_{ib}^{HF} P_{ppia} \right) \\ & + \sum_p \zeta_{OE}^p \left( 2 \delta_{bp} F_{ap} + 2 \sum_i D_{ib}^{HF} P_{ppia} \right) + 2 \left( \sum_i \zeta_u^{ib} u_{ia} + 2 \sum_{ijk} \zeta_g^{ikb} (ij | ka) \right) \\ & + (a \leftrightarrow b). \end{aligned}$$

(113)

The obtained Lagrangian multipliers may now be combined with the molecular integrals to compute the molecular energy gradients:

$$\begin{aligned}
\frac{dL}{dX} &= \frac{\partial W}{\partial X} + \zeta \frac{\partial e}{\partial X} = \zeta \frac{\partial e}{\partial X} \\
&= \sum_{p>q \in O} \zeta_{HF,MO}^{pq} \frac{\partial}{\partial X} (x_{pq} - x_{qp}) + \sum_{p>q \in D} \zeta_{HF,MO}^{pq} \frac{\partial F_{pq}}{\partial X} \\
&\quad + \sum_p \zeta_{OE}^p \frac{\partial}{\partial X} \left\{ (p|h|p) + \sum_{rs} D_{rs}^{HF} [(pp|rs) - (pr|sp)/2] \right\} \\
&\quad + \sum_{p \geq q} \zeta_S^{pq} (p|q)^X + \sum_{pq} \zeta_u^{pq} (p|v|q)^X + \sum_{pqrs} \zeta_g^{pqrs} (pq|rs)^X \\
&= \sum_{pq} d_h^{pq} (p|h|q)^X + \sum_{pw} d_s^{pq} (p|q)^X + \sum_{pqrs} D_g^{pqrs} (pq|rs)^X,
\end{aligned} \tag{114}$$

$d_h^{pq}$ ,  $d_s^{pq}$ , and  $D_g^{pqrs}$  are effective densities for the Lagrangian in the MO basis. They are given by

$$d_h^{pq} = \sum_{i>j \in O} \zeta_{HF,MO}^{ij} (\delta_{pi} D_{jq}^{HF} - (i \leftrightarrow j)) + \zeta_u^{pq} + \begin{cases} \zeta_{OE}^p & (p = q) \\ \zeta_{HF,MO}^{pq} & (p > q \in D) \\ 0 & (\text{otherwise}) \end{cases}, \tag{115}$$

$$d_s^{pq} = \zeta_S^{pq}, \tag{116}$$

and

$$\begin{aligned}
D_g^{pqrs} &= \sum_{i>j \in O} \zeta_{HF,MO}^{ij} \sum_{\gamma} (\delta_{pi} D_{jq,rs}^{HF} - (i \leftrightarrow j)) \\
&\quad + \sum_{i>j \in D} \zeta_{HF,MO}^{ij} (\delta_{pi} \delta_{qj} D_{rs}^{HF} - \delta_{pi} \delta_{sj} D_{qr}^{HF} / 2) \\
&\quad + \sum_i \zeta_{OE}^i (\delta_{pi} \delta_{qi} D_{rs}^{HF} - \delta_{pi} \delta_{si} D_{qr}^{HF} / 2) \\
&\quad + \zeta_g^{pqrs}.
\end{aligned} \tag{117}$$

These densities are transformed back into the atomic orbital (AO) basis, and the energy gradients may be obtained by multiplying the densities in the AO basis by the first-derivative integrals as in the MC-QDPT case,

$$\frac{dL}{dX} = \sum_{\mu\nu} d_h^{\mu\nu} \frac{d}{dX} (\mu|h|\nu) + \sum_{\mu\nu} d_s^{\mu\nu} \frac{d}{dX} (\mu|\nu) + \sum_{\mu\nu\rho\sigma} D_g^{\mu\nu\rho\sigma} \frac{d}{dX} (\mu\nu|\rho\sigma). \tag{118}$$

The formulas in this section are further reducible using Eqs. (98) and (99). However, they are remained as they are to maintain the similarity to the formulas in the previous sections.

## 6.2 An example calculation: methylene $CH_2$

As an example application of the energy gradient method of the conventional QDPT, we

determined the geometries of the ground ( $1^3B_2$ ) and two excited ( $1^1A_1$  and  $1^1B_2$ ) states of methylene  $CH_2$ . These states have the following main configurations:

$$1^1A_1 \text{ — } (1a_1)^2(2a_1)^2(1b_1)^2(3a_1)^2(1b_2)^0 \cdots \text{ HF}$$

$$1^3B_2 \text{ — } (1a_1)^2(2a_1)^2(1b_1)^2(3a_1)^1(1b_2)^1 \cdots 3a_1 \rightarrow 1b_2$$

$$1^1B_2 \text{ — } (1a_1)^2(2a_1)^2(1b_1)^2(3a_1)^1(1b_2)^1 \cdots 3a_1 \rightarrow 1b_2.$$

The basis set used are Dunning's correlation consistent valence double zeta (cc-pVDZ) set<sup>19</sup>. The closed shell Hartree-Fock orbitals, corresponding to  $1^1A_1$  state, were used for all the states. The active space was one constructed by distributing eight electrons to five orbitals, CAS(8e,5o).

Results are shown in Table 1. Internally contracted multireference configuration interaction (MRCI)<sup>20</sup> results are also given in the table for comparison. (The orbitals are optimized with CASSCF for each state.) The geometries using MRCI were determined with a non-gradient numerical optimization method. The ground and first excited state structures of QDPT are in very good agreement with those of MRCI. The deviations in bond length and bond angle are within 0.01 Å and 1 degree, respectively, in both states. The adiabatic excitation energy using QDPT is 15.31 kcal/mol, which is somewhat larger than that using MRCI, 11.74 kcal/mol. For  $1^1B_2$  state, the structure using QDPT is also close to that obtained using MRCI. The deviations from the MRCI values are 0.006 Å and 3.4 degrees. The excitation energy using QDPT is almost equal to the MRCI value. In Table 1 the excitation energies to the first order (reference space CI) computed at the QDPT optimized structures are also shown. For the  $1^1A_1$  state the excitation energy computed is too small, 0.47 kcal/mol, and for  $1^1B_2$  state it is also small, 35.89 kcal/mol, which indicates that a second-order perturbation treatment is essential for these states.

**Table 1:** Optimized geometries and adiabatic excitation energies in  $CH_2$ .

	r(C-H) (Å)	∠HCH (degree)	ΔE (kcal/mol)
$1^3B_2$			
QDPT	1.090	131.4	–
MRCI	1.092	131.9	–
$1^1A_1$			
QDPT	1.119	101.7	15.31
MRCI	1.124	101.0	11.74
CI(8e,5o)	(1.119)	(101.7)	0.47
HF	1.107	102.7	–
$1^1B_2$			
QDPT	1.083	144.0	38.94
MRCI	1.089	140.6	38.54
CI(8e,5o)	(1.083)	(144.0)	35.89

Note that the program used to compute the numbers in this section is a preliminary one, coded in early 1990s.

## 7 Concluding remarks

We have derived an analytic expression for the analytic energy gradients for MC-QDPT using the Lagrange multiplier method. With this method, we can obtain gradients without solving coupled perturbed equations. Instead, it is necessary to solve eight sets of linear equations for the multipliers. Six of the eight equations reduces to simple partial differential forms that directly give the multipliers, and only the remaining two are large scale linear equations that need iterative procedure. The gradients are given as the product of the first derivative integrals and the effective density that depends the obtained multipliers and the parameters.

The terms derived from the second-order energy may be expressed through diagrams. Since we did not perform any operations that change the structure of the creation-annihilation operators in the derivation, the diagrams do not change from those of the second-order effective Hamiltonian. Thus, the set of (25 Goldstone) diagrams used in the calculation of the effective Hamiltonian is applicable to the relevant terms if we adopted appropriate rules for transferring a diagram to an algebraic formula. This is why we did not adopted the second quantized form of parameter relaxation, which introduces more than hundred (156) Goldstone diagrams for orbital relaxation.

The gradient method of the conventional QDPT and some numerical results are also presented: these were not included in our previous paper. At present various definition of MOs and their energies are used. The present Lagrangian is not applicable to such general cases, since the orbitals are limited to the Hartree-Fock MOs. However, the extension is readily feasible by changing some terms in the Lagrangian.

Finally, we note that the code for the conventional QDPT gradients used in the Section 6, which is an old preliminary one, is independent of the more powerful code for MC-QDPT gradients we are currently developing. The new code will be released in some *ab initio* program packages, GAMESS<sup>21</sup>, HONDO<sup>22</sup>, and MR2D<sup>23</sup>.

## Acknowledgment

The authors were supported by the Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan. The conventional QDPT gradient and MRCI calculations were performed with KOTO92<sup>24</sup> and MOLPRO96<sup>25</sup>, respectively.

## References

1. P. E. Siegbahn, A. Heiberg, B. O. Roos, and B. Levy, *Physica Scripta*, **21**, 323 (1980); B. O. Roos, P. R. Taylor, and P. E. Siegbahn, *Chem. Phys.* **48**, 157 (1980); B. O. Roos, *Intern. J. Quantum Chem.* **S14**, 175 (1980).



2. K. Ruedenberg and K. R. Sundberg, in *Quantum Science*, eds. J. L. Calais, O. Goscinski, J. Linderberg, and Y. Öhrn (Plenum Press, New York, 1976) p. 505; L. M. Cheung, K. R. Sundberg, and K. Ruedenberg, *Intern.J. Quantum Chem.* **16**, 1103 (1979).
3. B. H. Brandow, *Rev. Mod. Phys.* **39**, 771 (1967); I. Lindgren, *Intern. J. Quantum Chem.* **S12**, 33 (1978); B. H. Brandow, *Intern. J. Quantum Chem.* **15**, 207 (1979); G. Hose and U. Kaldor, *J. Phys. B* **12**, 3827 (1981); K. F. Freed and M. G. Sheppard, *J. Chem. Phys.* **75**, 4507, 4525 (1981); H. Primas, *Helv. Phys. Acta* **34**, 331 (1961); B. Kirtman, *J. Chem. Phys.* **49**, 3890 (1968); P. R. Certain and J. O. Hirschfelder, *J. Chem. Phys.* **52**, 5977 (1970); J. O. Hirschfelder, *Chem. Phys. Lett.* **54**, 1 (1978).
4. K. Wolinski, H. L. Sellers, and P. Pulay, *Chem. Phys. Lett.* **140**, 225 (1987); J. J. W. McDouall, K. Peasley, and M. A. Robb, *Chem. Phys. Lett.* **148**, 183 (1988); K. Andersson, P. Malmqvist, B. O. Roos, A. J. Sadlej, K. Wolinski, *J. Phys. Chem.* **94**, 5483 (1990); K. Andersson, P. Malmqvist, B. O. Roos, *J. Chem. Phys.* **96**, 1218 (1992); R. B. Murphy and R. P. Messmer, *Chem. Phys. Lett.* **183**, 443 (1991); *J. Chem. Phys.* **97**, 4170 (1992).
5. K. Hirao, *Chem. Phys. Lett.* **190**, 374 (1992); **196**, 397 (1992); **201**, 59 (1993); *Intern. J. Quantum Chem.* **S26**, 517 (1992); H. Nakano, M. Yamanishi, and K. Hirao, *Trends in Chem. Phys.*, **6**, 167 (1997).
6. H. Nakano, *J. Chem. Phys.* **99**, 7983 (1993); H. Nakano, *Chem. Phys. Lett.* **207**, 372 (1993).
7. T. Tsuneda, H. Nakano, and K. Hirao, *J. Chem. Phys.* **103**, 6520 (1995); K. Hirao, H. Nakano, and T. Hashimoto, *Chem. Phys. Lett.* **235**, 430 (1995); H. Nakano, T. Tsuneda, T. Hashimoto, and K. Hirao, *J. Chem. Phys.* **104**, 2312 (1996); T. Hashimoto, H. Nakano, and K. Hirao, *J. Chem. Phys.* **104**, 6244 (1996); Y. Kawashima, K. Nakayama, H. Nakano, and K. Hirao, *Chem. Phys. Lett.* **267**, 82 (1997); H. Nakano, K. Nakayama, K. Hirao, and M. Dupuis, *J. Chem. Phys.* **106**, 4912 (1997); K. Nakayama, H. Nakano, and K. Hirao, *Intern. J. Quantum Chem.* **66**, 157 (1998).
8. Y. Yamaguchi, Y. Osamura, J. D. Goddard, and H. F. Schaefer III, *A New Dimension to Quantum Chemistry, Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory* (Oxford University Press, Oxford, 1994); P. Pulay, *Adv. Chem. Phys.* **69**, 241 (1987); P. Pulay, in *Modern Electronic Structure Theory Part II*, ed. D. R. Yarkony (World Scientific, Singapore, 1995), p. 1191.
9. S. Kato and K. Morokuma, *Chem. Phys. Lett.* **65**, 19 (1979).
10. M. Page, P. Saxe, G. F. Adams, and B. H. Lengsfeld III, *J. Chem. Phys.* **81**, 434 (1984); R. Shepard, H. Lischka, P. G. Szalay, T. Kovar, and M. Ernzerhof, *J. Chem. Phys.* **96**, 2085 (1992); R. Shepard, in *Modern Electronic Structure Theory Part II*, ed. D. R. Yarkony (World Scientific, Singapore, 1995), p. 345.

11. H. Nakano, K. Hirao, and M. S. Gordon, *J. Chem. Phys.* **108**, 5660 (1998).
12. N. C. Handy and H. F. Schaefer III, *J. Chem. Phys.* **81**, 5031 (1984).
13. T. U. Helgaker and P. Jørgensen, *Adv. Quantum Chem.* **19**, 183 (1988); P. Jørgensen and T. U. Helgaker, *J. Chem. Phys.* **89**, 1560 (1988).
14. P. Pulay, *J. Chem. Phys.* **78**, 4043 (1983).
15. P. Jørgensen and J. Simons, *J. Chem. Phys.* **79**, 334 (1983); T. U. Helgaker and J. Almlöf, *Intern. J. Quantum Chem.* **26**, 275 (1984).
16. C. C. J. Roothaan and P. S. Bagus, *Methods Comput. Phys.* **2**, 47 (1963).
17. J. E. Rice and R. D. Amos, *Chem. Phys. Lett.* **122**, 585 (1985).
18. C. W. Murray and E. R. Davidson, *Chem. Phys. Lett.* **187**, 451 (1991); P. M. Kozłowski and E. R. Davidson, *J. Chem. Phys.* **100**, 3672 (1994).
19. T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
20. H. -J. Werner and P. J. Knowles, *J. Chem. Phys.* **89**, 5803 (1988); P. J. Knowles and H. -J. Werner, *Chem. Phys. Lett.* **145**, 514 (1988).
21. M. W. Schmidt, K. K. Baldrige, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, Jr., *J. Comp. Chem.*, **14**, 1347 (1993).
22. M. Dupuis, S. Chin, and A. Marquez, in *Relativistic and Electron Correlation Effects in Molecules and Clusters*, NATO ASI Series, ed. G. L. Malli (Plenum, New York, 1992).
23. MR2D Ver. 2, H. Nakano, University of Tokyo, 1995.
24. KOTO92, S. Obara, M. Honda, H. Nakano, F. Sakano, S. Takada, Y. Miyake, K. Sato, and T. Nakajima, Kyoto University, 1992.
25. MOLPRO is a package of *ab initio* programs written by H. -J. Werner and P. J. Knowles, with contributions from J. Almlöf, R. D. Amos, M. J. O. Deegan, S. T. Elbert, C. Hampel, W. Meyer, K. Peterson, R. Pitzer, A. J. Stone, P. R. Taylor and R. Lindh.