

# Analytic first-order derivatives of partially contracted $n$ -electron valence state second-order perturbation theory (PC-NEVPT2)

Yoshio Nishimoto<sup>1, a)</sup>

*Fukui Institute for Fundamental Chemistry, Kyoto University,  
34-4 Takano Nishihiraki-cho, Sakyo-ku, Kyoto 606-8103,  
Japan*

(Dated: 1 August 2019)

A balanced treatment of dynamic and static electron correlation is important in computational chemistry, and multireference perturbation theory (MRPT) is able to do this at a reasonable computational cost. In this paper, analytic first-order derivatives, specifically gradients and dipole moments, are developed for a particular MRPT method, state-specific partially contracted  $n$ -electron valence state second-order perturbation theory (PC-NEVPT2). Only one linear equation needs to be solved for the derivative calculation if the Z-vector method is employed, which facilitates the practical application of this approach. Comparison of the calculated results with experimental geometrical parameters of  $\text{O}_3$  indicates excellent agreement, although the calculated results for  $\text{O}_3^-$  are slightly outside the experimental error bars. The 0–0 transition energies of various methylpyrimidines and *trans*-polyacetylene are calculated by performing geometry optimizations and seminumerical second-order geometrical derivative calculations. In particular, the deviations of 0–0 transition energies of *trans*-polyacetylene from experimental values are consistently less than 0.1 eV with PC-NEVPT2, indicating the reliability of the method. These results demonstrate the importance of adding dynamic electron correlation on top of methods dominated by static electron correlation and of developing analytic derivatives for highly accurate methods.

---

<sup>a)</sup>E-mail:nishimoto.yoshio@fukui.kyoto-u.ac.jp

## I. INTRODUCTION

A central topic in computational chemistry is the development of methods that capture electron correlation as much as possible at a reasonable computational cost. It is well known that the full configuration interaction (FCI) gives an exact solution to the nonrelativistic and time-independent Schrödinger equation. The FCI method treats all possible excited determinants or configuration state functions, and electron correlation is fully recovered. However, the associated computational cost is already prohibitively high even for small systems. In contrast, the Hartree–Fock (HF) method uses only one determinant, which usually obeys the Aufbau principle, and is known to completely lack electron correlation, as defined by Löwdin.<sup>1</sup> Therefore, to achieve qualitatively—and hopefully quantitatively—correct descriptions, it is nowadays popular to take account of dynamic electron correlation through the use of post-HF methods.

To achieve practical accuracy, a number of post-HF methods have been developed.<sup>2</sup> The most successful single-reference (SR) method is the coupled-cluster method with single and double excitations including perturbative triples [CCSD(T)], which is sometimes considered to be the “gold standard” method. CCSD(T) gives a very accurate description of what we call dynamic electron correlation, which arises from instantaneous repulsion of electrons and which is not considered in the HF method. However, SR methods do not adequately consider the effect of what we call static or nondynamic electron correlation, which arises from near-degenerate determinants. Density functional theory (DFT) with standard exchange–correlation functionals belongs to the class of SR methods, but its philosophy is quite different from that of the post-HF methods. Nevertheless, properties predicted with DFT agree well with experimental results in many cases, with much reduced computational cost, if appropriate exchange–correlation functionals are employed.

Another class, namely, multiconfiguration (MC) methods,<sup>3</sup> employ more than one Slater determinant, allowing near-degeneracy of determinants to be considered. A widely used method is the complete active space self-consistent field (CASSCF) method, in which FCI is performed within a restricted (active) space and molecular orbitals (MOs) are also updated iteratively. CASSCF captures static electron correlation to a large extent, but electron promotions and excitations outside the active space are completely ignored, so a large portion of the dynamic electron correlation is missing. Therefore, in analogy to the SR

post-HF methods, powerful multireference (MR) post-CASSCF methods have been developed. Among these, MR coupled-cluster<sup>4-6</sup> and configuration interaction (CI)<sup>7</sup> methods give extremely accurate results. However, the drawback of these methods is their high computational cost, and they cannot be applied to other than small systems. Finding stationary points, i.e., geometry optimization, is even more challenging. The cheapest post-CASSCF methods exploit perturbation theory in the form of an MR perturbation theory (MRPT). The most well-known MRPT is probably complete active space second-order perturbation theory (CASPT2).<sup>8-10</sup> Other MRPTs, such as (extended)<sup>11</sup> multiconfiguration quasidegenerate second-order perturbation theory [(X)MCQDPT2],<sup>12</sup>  $n$ -electron valence state second-order perturbation theory (NEVPT2),<sup>13-15</sup> and generalized van Vleck second-order perturbation theory (GVVPT2),<sup>16</sup> are also employed for a number of tasks.

MRPT methods often use internally contracted functions, and this significantly reduces the computational complexity, particularly when the active space is large. However, it also makes the formulation complicated. The expressions for the first-order wavefunction and the second-order energy correction are already complicated, and the further development of analytic derivatives is even more challenging. Indeed, for the above MRPT methods, only a limited number of software packages are available that allow analytic derivatives to be performed (for CASPT2<sup>17-21</sup> and for GVVPT2<sup>22-25</sup>), although a wider range of software packages can realize energy calculations. For MCQDPT2, there is a formulation for analytic derivatives,<sup>26</sup> but it seems that the actual code is not available (at least not publicly). There have been no previous reports of such software for NEVPT2 (although, during the submission process of this paper, Park<sup>27</sup> reported the implementation of the partially and strongly contracted variants of the NEVPT2 gradient).

Looking at MRPT methods in terms of analytic derivative techniques,<sup>28</sup> the first challenge is posed by the fact that most MRPT energies are not variationally optimized with respect to wavefunction parameters. This necessitates the evaluation of derivatives of nonvariational parameters, which are orbital and CI (when internally contracted) coefficients. A straightforward approach (although much more complicated than in the SR case) is to evaluate these derivatives directly by solving the coupled-perturbed (CP) MCSCF equations as in Ref. 22. In spite of the simplicity of this strategy, the number of CP-MCSCF equations to be solved increases linearly with the number of atoms in the system, so it is not suitable for treating even medium-sized systems. Another, more practical, approach relies on the Z-vector<sup>29</sup> or

Lagrangian<sup>30</sup> methods. Both of these arrive at the  $2n+1$  rule for wavefunction parameters at the first-order derivative, but the latter can straightforwardly exploit a stronger  $2n+2$  rule for Lagrange multipliers<sup>31</sup> for higher-order derivatives. The advantage of these methods is the reduced computational requirement: solving only one perturbation-independent equation is sufficient to evaluate response contributions for the first-order derivative. This strategy was realized for CASPT2 in 2003<sup>17</sup> and for GVVPT2 in 2011.<sup>24</sup> It is useful to point out the difference between the two analytic derivatives for the CASPT2 method, namely, those in Refs. 17 and 19. These studies employed the partially and fully internal contraction schemes, respectively. The former scheme contracts only the doubly external configurations (two electron excitations into virtual orbitals). As a result, the development of analytic derivatives is somewhat simplified, since we evaluate only the derivatives of one- and two-particle reduced density matrices (RDMs). In contrast, the fully internal contraction scheme contracts all the excitation classes and additionally requires three- and four-particle RDMs (depending on the implementation). The latter scheme is more efficient, but, owing to the complexity of its development, MacLeod *et al.* employed an automatic code generation technique.<sup>19</sup> In passing, it should be noted that the partial contraction used in NEVPT2 is equivalent to the fully internal contraction, but not to the partially internal contraction.<sup>32</sup>

In this paper, the development of analytic first-order derivatives (gradients and dipole moments) of the state-specific partially contracted variant of NEVPT2 (partially contracted NEVPT2: PC-NEVPT2) is reported. The response contribution is evaluated through the Z-vector method, so only one response equation needs to be solved.

## II. DERIVATION

With regard to notation, the following indices are used:

- doubly occupied MOs:  $i$  and  $j$  (referred to as “docc”);
- active MOs:  $a$ ,  $a'$ ,  $b$ ,  $c$ ,  $d$ , and  $e$  (“act”);
- external (virtual) MOs:  $r$  and  $s$  (“vir”);
- general MOs:  $p$ ,  $q$ ,  $m$ ,  $n$ , and  $o$ ;
- internally contracted configurations:  $\phi$  and  $\chi$ ;

- atomic orbitals (AOs):  $\mu, \nu, \rho$ , and  $\sigma$ ;
- Slater determinants:  $I$  and  $J$ ;
- internal states:  $S, T$ , and  $U$  ( $\in \Omega$ );
- external states:  $P$  and  $Q$  ( $\in \Lambda$ );
- general states:  $A$ .

The internal and external states<sup>33</sup> are those averaged in state-averaged CASSCF (SA-CASSCF) and those not included in SA-CASSCF (orthogonal complement to the internal states), respectively. The SA-CASSCF energy is given by

$$E^{\text{SA-CAS}} = \sum_S^{\Omega} \omega_S E_S^{\text{CAS}} , \quad (1)$$

where  $\omega_S$  ( $0 < \omega_S \leq 1$ ) is the weighting factor of the state averaging and  $E_S$  is the state-specific (SS) energy of state  $S$ .  $\Omega + \Lambda (= N)$  spans the full configuration space.

### A. CASSCF energy and gradient

Before going into the details of NEVPT2, it is convenient to derive the CASSCF energy and first-order derivatives. The SS-CASSCF energy for the internal state  $S$  is given by

$$E_S^{\text{CAS}} = \sum_{pq} h_{pq} \gamma_{pq}^S + \frac{1}{2} \sum_{pqmn} (pq|mn) \Gamma_{pqmn}^S , \quad (2)$$

where  $h_{pq}$  and  $(pq|mn)$  are the usual one- and two-electron (chemists' notation<sup>34</sup>) integrals, respectively, and  $\gamma_{pq}^S$  and  $\Gamma_{pqmn}^S$  are the one- and two-particle reduced density matrix (1-RDM and 2-RDM) elements, respectively, defined by

$$\gamma_{pq}^S = R_{pq}^{S,(1)} = \sum_{IJ}^N c_I^S c_J^S \gamma_{pq}^{IJ} , \quad (3)$$

$$\Gamma_{pqmn}^S = R_{pm,qn}^{S,(2)} = \sum_{IJ}^N c_I^S c_J^S \Gamma_{pqmn}^{IJ} , \quad (4)$$

where  $c_I^S$  is the configuration coefficient,  $\gamma_{pq}^{IJ}$  and  $\Gamma_{pqmn}^{IJ}$  are the one- and two-particle coupling coefficients, respectively, and  $R_{pq}^{S,(1)}$  and  $R_{pm,qn}^{S,(2)}$  are used in the derivation below. The coupling

coefficients are independent of the derivative parameter, so the derivatives of the RDMs are dependent only on the derivative of the configuration coefficient. The derivative of the SS-CASSCF energy with respect to a derivative parameter  $\alpha$  can be calculated by

$$\frac{dE_S^{\text{CAS}}}{d\alpha} = \sum_{\mu\nu} d_{\mu\nu}^{S,\text{CAS}} \frac{dh_{\mu\nu}}{d\alpha} - \sum_{\mu\nu} X_{\mu\nu}^{S,\text{CAS}} \frac{dS_{\mu\nu}}{d\alpha} + \sum_{\mu\nu\rho\sigma} D_{\mu\nu\rho\sigma}^{S,\text{CAS}} \frac{d(\mu\nu|\rho\sigma)}{d\alpha} + \sum_{pq}^{\text{indep}} \mathcal{L}_{pq}^{S,\text{CAS},o} \tilde{U}_{pq}^\alpha, \quad (5)$$

where  $S_{\mu\nu}$  is the overlap matrix, and  $d_{\mu\nu}^{S,\text{CAS}}$ ,  $X_{\mu\nu}^{S,\text{CAS}}$ , and  $D_{\mu\nu\rho\sigma}^{S,\text{CAS}}$  are defined by

$$d_{\mu\nu}^{S,\text{CAS}} = \sum_{pq} C_{\mu p} C_{\nu q} \gamma_{pq}^S, \quad (6)$$

$$X_{\mu\nu}^{S,\text{CAS}} = \sum_p^{\text{all}} \sum_q^{\text{docc+act}} C_{\mu p} C_{\nu q} \varepsilon_{pq}^S, \quad (7)$$

and

$$D_{\mu\nu\rho\sigma}^{S,\text{CAS}} = \frac{1}{2} \sum_{pqmn} C_{\mu p} C_{\nu q} C_{\rho m} C_{\sigma n} \Gamma_{pqmn}^S, \quad (8)$$

respectively, where  $C_{\mu p}$  is the MO coefficient matrix and the Lagrangian matrix  $\varepsilon_{pq}^S$  is defined in the next paragraph. The last term in Eq. (5) is due to the orbital response contribution, which comes from the derivatives of the MO coefficients. These derivatives are expressed by a unitary transformation:

$$\frac{dC_{\mu q}}{d\alpha} = \sum_p^{\text{all}} C_{\mu p} U_{pq}^\alpha. \quad (9)$$

In the actual implementation in GAMESS-US, a slightly modified expression is employed:<sup>35</sup>

$$\tilde{U}_{pq}^\alpha = U_{pq}^\alpha + \frac{1}{2} S_{pq}^\alpha, \quad (10)$$

where

$$S_{pq}^\alpha = \sum_{\mu\nu} C_{\mu p} C_{\nu q} \frac{dS_{\mu\nu}}{d\alpha}. \quad (11)$$

The scaling factor in the second term of  $\tilde{U}_{pq}^\alpha$  seems to be arbitrary, but the choice of this factor affects the definition in Eq. (7). In the summation in the last term of Eq. (5), the shorthand notation “indep” is used to denote the independent rotations that do not change the total energy. This term can be formally written as the sum of four orbital subspaces:

$$\sum_{pq}^{\text{indep}} = \sum_p^{\text{act}} \sum_q^{\text{docc}} \oplus \sum_p^{\text{vir}} \sum_q^{\text{docc}} \oplus \sum_p^{\text{vir}} \sum_q^{\text{act}} \oplus \sum_{p>q}^{\text{act}}. \quad (12)$$

The last term here vanishes if the energy is invariant under rotations among active orbitals, so it vanishes for CASSCF wavefunctions, although not for general MCSCF. Therefore, since only CASSCF wavefunctions are employed in this study, the independent rotation pairs consist of  $(p, q) \in (\text{act}, \text{docc}), (\text{vir}, \text{docc}),$  and  $(\text{vir}, \text{act})$  orbital blocks.

The SS-CASSCF Lagrangian matrix in Eq. (7) is defined as

$$\varepsilon_{pq}^S = \sum_m h_{pm} \gamma_{qm}^S + \sum_{mno} (pm|no) \Gamma_{qmmo}^S, \quad (13)$$

and the partial derivative matrix with respect to the MO coefficients  $\mathcal{L}_{pq}^{R,\text{CAS},o}$  may then be written as

$$\mathcal{L}_{pq}^{S,\text{CAS},o} = (1 - \hat{\tau}_{pq}) \sum_{\mu} C_{\mu p} \frac{\partial E_S^{\text{CAS}}}{\partial C_{\mu q}} = 2 (\varepsilon_{pq}^S - \varepsilon_{qp}^S), \quad (14)$$

where  $\hat{\tau}_{pq}$  permutes the indices  $p$  and  $q$ . The orbital derivative is zero for single-state CASSCF but generally nonzero for SA-CASSCF, because the generalized Brillouin condition is satisfied only for the state-averaged derivative matrix, and so

$$\sum_S^{\Omega} \omega_S \mathcal{L}_{pq}^{S,\text{CAS},o} = 0. \quad (15)$$

For this reason, to obtain the derivatives of SS-CASSCF energies from SA-CASSCF wavefunctions, the last term in Eq. (5) has to be evaluated by solving either the CP or Z-vector equations. The response due to the first-order derivatives of CI coefficients is zero for SS- and SA-CASSCF, because CI coefficients are optimized for each internal state.

## B. Outline of the PC-NEVPT2 energy

A full description has already been given in the original NEVPT2 papers,<sup>13–15</sup> but some details are given here with the later derivation in mind. In some ways, PC-NEVPT2 is similar to (fully) internally contracted CASPT2. However, they differ in the definition of the zeroth-order Hamiltonian, which is defined by projecting a model Hamiltonian: NEVPT2 and CASPT2 use the Dyall Hamiltonian<sup>36</sup> (bielectronic in the active space) and the generalized Fock operator (monoelectronic) as their respective model Hamiltonians. Another difference, related to the above, is the spectral decomposition in the zeroth-order Hamiltonian. NEVPT2 does not include coupling between different excitation classes, so the wavefunction is usually parameterized through a diagonalization (for PC-NEVPT2) or summation (for SC-NEVPT2,

see below). In contrast, the amplitude equation is iteratively solved for CASPT2. A more detailed comparison between CASPT2 and NEVPT2 may be found in Ref. 37. It is known that, owing to its sophisticated definition of the zeroth-order Hamiltonian, NEVPT2 does not exhibit the intruder state problem and is a size-consistent method.<sup>13</sup> One familiar drawback of the NEVPT2 method is its sensitivity to the quality of the active space. As discussed by Schapiro *et al.*,<sup>38</sup> this sensitivity is due to the different treatment of inactive and active orbitals, by one- and two-particle operators, respectively. In contrast, CASPT2 employs a one-particle operator in all the orbital regions, so it is less affected by the choice of the active space. In this study, analytic first-order derivatives in the PC-NEVPT2 method are studied. There is another variant of contracted NEVPT2, namely, strongly contracted NEVPT2 (SC-NEVPT2), but this is not considered in the present study.

In NEVPT2, the second-order perturbative correction to the energy  $E_S^{\text{PT2}}$  for internal state  $S$  is given by the sum of the eight excitation contributions:

$$E_S^{\text{PT2}} = E_S^{(0)} + E_S^{(-1)} + E_S^{(+1)} + E_S^{(-2)} + E_S^{(+2)} + E_S^{(0)'} + E_S^{(-1)'} + E_S^{(+1)'} , \quad (16)$$

where the terms have been defined in previous publications.<sup>13–15</sup> The numbers in the superscript parentheses represent the number of electrons promoted to or removed from the active space, and the prime distinguishes unique excitation classes. The total NEVPT2 energy is given by the sum of the reference (CASSCF) energy and the perturbative correction:  $E_S^{\text{NEVPT2}} = E_S^{\text{CAS}} + E_S^{\text{PT2}}$ . The index of the internal state  $S$  is omitted hereinafter for simplicity unless needed. In the following derivation, we focus on the  $E_S^{(-1)}$  term only.  $E_S^{(0)}$  is quite similar to the SR MP2 energy and gradients. For the other terms, one can derive working equations by following a similar procedure to that presented below.

Following Ref. 15 with a slight modification, the second-order perturbative correction of the  $E^{(-1)}$  term for PC-NEVPT2 is given by

$$E^{(-1)} = -\frac{1}{2} \sum_{rs} \sum_i \sum_{\phi} \frac{((rs, \phi i))^2 + ((rs, \phi i)')^2}{\varepsilon_r + \varepsilon_s - \varepsilon_i + e_{\phi}} , \quad (17)$$

where

$$(rs, \phi i) = \frac{1}{\sqrt{2}} \sum_a ((ri|sa) + (ra|si)) \mathcal{S}_{\phi a} \quad (18)$$

and

$$(rs, \phi i)' = \sqrt{\frac{3}{2}} \sum_a ((ri|sa) - (ra|si)) \mathcal{S}_{\phi a} , \quad (19)$$



with

$$\mathcal{S}_{\phi a} = \sum_{a'} \mathcal{C}_{a'\phi} \mathcal{R}_{a'a} . \quad (20)$$

$\varepsilon_p$  is the orbital energy of the  $p$ th MO, obtained by block-diagonalizing the averaged Fock matrix.<sup>39</sup> The internally contracted coefficient  $\mathcal{C}_{a'\phi}$  and eigenvalue  $e_\phi$  are obtained by solving the generalized eigenvalue equation

$$\mathcal{K}\mathcal{C} = \mathcal{R}\mathcal{C}\mathbf{e} , \quad (21)$$

where

$$\mathcal{K}_{a'a} = K_{a'a} = - \sum_c h_{ac}^{\text{eff}} R_{a'c}^{(1)} - \sum_{cde} (ad|ce) R_{a'c,de}^{(2)} \quad (22)$$

and

$$\mathcal{R}_{a'a} = R_{a'a}^{(1)} \quad (23)$$

for this excitation class. The effective Hamiltonian is

$$h_{pq}^{\text{eff}} = h_{pq} + \sum_i (2(pq|ii) - (pi|qi)) . \quad (24)$$

There are no couplings between other excitation classes in Eq. (21), so the eigenvalue equation is solved by direct diagonalization, but not iteratively as in CASPT2. In the contracted NEVPT2, up to four-particle RDMs (1-, 2-, 3-, and 4-RDMs) are needed.

### C. Derivative of the $E^{(-1)}$ term

Since the PC-NEVPT2 (and SC-NEVPT2) energy is not stationary with respect to changes in the MO coefficients  $C_{\mu p}$  and configuration coefficients  $c_I$ , derivatives of the coefficients remain in the first-order derivative. As a consequence, one has to evaluate the partial derivatives of the energy with respect to configuration,  $\mathcal{L}_{SI}^{\text{PT2,c}}$ , and state,  $\mathcal{L}_{ST}^{\text{PT2,s}}$ , as well as orbital,  $\mathcal{L}_{pq}^{\text{PT2,o}}$ . This is slightly different from SA-CASSCF, in which the energy is stationary with respect to changes in configuration coefficients. Therefore, the first task is to find an expression such that

$$\begin{aligned} \frac{dE^{(-1)}}{d\alpha} = & \sum_{\mu\nu} d_{\mu\nu}^{\text{PT2}} \frac{dh_{\mu\nu}}{d\alpha} - \sum_{\mu\nu} X_{\mu\nu}^{\text{PT2}} \frac{dS_{\mu\nu}}{d\alpha} + \sum_{\mu\nu\rho\sigma} D_{\mu\nu\rho\sigma}^{\text{PT2}} \frac{d(\mu\nu|\rho\sigma)}{d\alpha} \\ & + \sum_{pq}^{\text{indep}} U_{pq}^\alpha \mathcal{L}_{pq}^{\text{PT2,o}} + \sum_S^\Omega \sum_I^N V_{SI}^\alpha \mathcal{L}_{SI}^{\text{PT2,c}} + \sum_{S<T}^\Omega W_{ST}^\alpha \mathcal{L}_{ST}^{\text{PT2,s}} . \end{aligned} \quad (25)$$

$d_{\mu\nu}^{\text{PT2}}$ ,  $X_{\mu\nu}^{\text{PT2}}$ , and  $D_{\mu\nu\rho\sigma}^{\text{PT2}}$  are the unrelaxed (without response) one-electron, energy-weighted, and two-electron density matrices of the  $E^{(-1)}$  term, respectively, and  $U_{pq}^\alpha$ ,  $V_{RI}^\alpha$ , and  $W_{RS}^\alpha$  are the responses of orbitals, configuration coefficients, and states, respectively [which are given in Eqs. (9), (B6) (see Appendix B), and (74) (see below), respectively]. The derivation below is apparently different from those for CASPT2<sup>17</sup> and GVVPT2<sup>24</sup> (although it is equivalent), because it is based mostly on a variant of the Z-vector<sup>29</sup> method, while the derivations for CASPT2 and GVVPT2 are based on the Lagrangian<sup>30</sup> method. In the latter case, one first defines the Lagrangian as in Eq. (30) of Ref. 17 or Eq. (37) of Ref. 24, and then obtains the Z-vector equation to be solved by taking the partial derivative of the Lagrangian with respect to the wavefunction parameters. Since the CP-MCSCF equation<sup>33</sup> is obtained from the stationary condition on the Lagrangian, the Z-vector equations and final analytic derivatives derived in the two approaches are equivalent.

First, it is convenient to rewrite the energy as

$$E^{(-1)} = \sum_i \sum_a \sum_{rs} \sum_\phi \tilde{T}_{rs,i}^\phi \mathcal{S}_{\phi a}(ri|sa) , \quad (26)$$

where

$$\tilde{T}_{rs,i}^\phi = \sum_a \frac{2(ri|sa) - (ra|si)}{\varepsilon_i - \varepsilon_r - \varepsilon_s - e_\phi} \mathcal{S}_{\phi a} . \quad (27)$$

By taking the derivative of the energy with respect to a derivative parameter  $\alpha$ , we easily arrive at the following expression:

$$\begin{aligned} \frac{dE^{(-1)}}{d\alpha} &= 2 \sum_i \sum_a \sum_{rs} \sum_\phi \tilde{T}_{rs,i}^\phi \mathcal{S}_{\phi a} \frac{d(ri|sa)}{d\alpha} \\ &\quad - \sum_i \sum_{rs} \sum_\phi \tilde{T}_{rs,i}^\phi T_{rs,i}^\phi \left( \frac{d\varepsilon_i}{d\alpha} - \frac{d\varepsilon_r}{d\alpha} - \frac{d\varepsilon_s}{d\alpha} \right) \\ &\quad + 2 \sum_i \sum_a \sum_{rs} \sum_\phi \tilde{T}_{rs,i}^\phi \frac{d\mathcal{S}_{\phi a}}{d\alpha}(ri|sa) + \sum_i \sum_{rs} \sum_\phi \tilde{T}_{rs,i}^\phi T_{rs,i}^\phi \frac{de_\phi}{d\alpha} , \end{aligned} \quad (28)$$

where

$$T_{rs,i}^\phi = \sum_a \frac{(ri|sa)}{\varepsilon_i - \varepsilon_r - \varepsilon_s - e_\phi} \mathcal{S}_{\phi a} . \quad (29)$$

If the summation over  $\phi$  in the first and second terms is taken, the resulting expression is quite similar to those found during the derivation of the MP2 gradient.<sup>40</sup> The contracted

MP2-like density matrix can be constructed in PC-NEVPT2 as well:

$$d_{pq}^{\text{PT2,inact}} = \begin{cases} -\sum_{rs} \sum_{\phi} \tilde{T}_{rs,p}^{\phi} T_{rs,q}^{\phi} & (p, q) \in (\text{docc}, \text{docc}) , \\ \sum_i \sum_r \sum_{\phi} \left( \tilde{T}_{pr,i}^{\phi} T_{qr,i}^{\phi} + \tilde{T}_{rp,i}^{\phi} T_{rq,i}^{\phi} \right) & (p, q) \in (\text{vir}, \text{vir}) , \\ 0 & \text{otherwise} . \end{cases} \quad (30)$$

Following the standard procedure, we can easily transform the first two terms (a brief derivation is given in Appendix A):

$$\begin{aligned} \frac{dE^{(-1)}}{d\alpha} &= \sum_{\mu\nu} d_{\mu\nu}^{\text{PT2,inact}} \frac{dh_{\mu\nu}}{d\alpha} \\ &+ \sum_{\mu\nu\rho\sigma} \left( D_{\mu\nu\rho\sigma}^{\text{PT2,inact}} + d_{\mu\nu}^{\text{PT2,inact}} d_{\rho\sigma}^{\text{SA}} - \frac{1}{4} d_{\mu\rho}^{\text{PT2,inact}} d_{\nu\sigma}^{\text{SA}} - \frac{1}{4} d_{\mu\sigma}^{\text{PT2,inact}} d_{\nu\rho}^{\text{SA}} \right) \frac{d(\mu\nu|\rho\sigma)}{d\alpha} \\ &+ \sum_{pq}^{\text{all}} \mathcal{L}_{pq}^{\text{PT2,inact}} U_{pq}^{\alpha} + \sum_{a'a} \tilde{R}_{a'a}^{(1),\text{inact}} \frac{d\gamma_{a'a}^{\text{SA}}}{d\alpha} \\ &+ 2 \sum_i \sum_a \sum_{rs} \sum_{\phi} \tilde{T}_{rs,i}^{\phi} \frac{d\mathcal{S}_{\phi a}}{d\alpha} (ri|sa) + \sum_i \sum_{rs} \sum_{\phi} \tilde{T}_{rs,i}^{\phi} T_{rs,i}^{\phi} \frac{de_{\phi}}{d\alpha} , \end{aligned} \quad (31)$$

where

$$D_{\mu\nu\rho\sigma}^{\text{PT2,inact}} = 2 \sum_i \sum_a \sum_{rs} C_{\mu r} C_{\nu i} C_{\rho s} C_{\sigma a} \tilde{T}_{rs,i}^a , \quad (32)$$

$d_{\mu\nu}^{\text{SA}}$  is the state-averaged density matrix in the AO basis [Eq. (A12)],

$$\mathcal{L}_{pq}^{\text{PT2,inact}} = \begin{cases} 2 \sum_a \sum_{rs} (rp|sa) \tilde{T}_{rs,q}^a + 2\tilde{d}_{pq} + 4G_{pq} [\mathbf{d}^{\text{PT2,inact}}] & q \in \text{docc} , \\ 2 \sum_i \sum_{rs} (ri|sp) \tilde{T}_{rs,i}^q + 2 \sum_a^{\text{act}} G_{pa} [\mathbf{d}^{\text{PT2,inact}}] \gamma_{aq}^{\text{SA}} & q \in \text{act} , \\ 2 \sum_i \sum_a \sum_r \left( (pi|ra) \tilde{T}_{qr,i}^a + (ri|pa) \tilde{T}_{rq,i}^a \right) + 2\tilde{d}_{pq} & q \in \text{vir} , \end{cases} \quad (33)$$

$$\tilde{T}_{rs,i}^a = \sum_{\phi} \tilde{T}_{rs,i}^{\phi} \mathcal{S}_{\phi a} , \quad (34)$$

and

$$\tilde{R}_{a'a}^{(1),\text{inact}} = G_{a'a} [\mathbf{d}^{\text{PT2,inact}}] = \sum_{\mu\nu} C_{\mu a'} C_{\nu a} G_{\mu\nu} [\mathbf{d}^{\text{PT2,inact}}] . \quad (35)$$

The Fock-like two-electron transformation is defined by

$$G_{\mu\nu} [\mathbf{d}] = \sum_{\rho\sigma} \left\{ (\mu\nu|\rho\sigma) - \frac{1}{4}(\mu\rho|\nu\sigma) - \frac{1}{4}(\mu\sigma|\nu\rho) \right\} d_{\rho\sigma} . \quad (36)$$

The Fock-weighted density matrix  $\tilde{d}_{pq}$  and the state-averaged density matrix  $\gamma_{pq}^{\text{SA}}$  are defined in Appendix A. The first and second terms in Eq. (31) are the integral derivatives of the one- and two-electron integral terms, respectively, and a similar term appears in MP2. The third term contains “explicit” orbital response contributions. With this transformation, it is possible to avoid division by the difference of orbital energies within doubly occupied and virtual orbitals. This property (which is related to rotational invariance in diagonal blocks) is important for molecules with high symmetry operations.

Up to here, the derivation is quite closely related to that in MP2. Actually, a similar procedure can be applied to the other seven excitation classes. The remaining difficulty lies in the fifth and sixth terms in Eq. (31). Obviously, the fifth term can be separated into two contributions using Eq. (20):

$$\begin{aligned} & 2 \sum_i \sum_a \sum_{rs} \sum_{\phi} \tilde{T}_{rs,i}^{\phi} \frac{d\mathcal{S}_{\phi a}}{d\alpha}(ri|sa) \\ &= \sum_{a'a} \check{R}_{a'a}^{(1),\text{exp}} \frac{dR_{a'a}^{(1)}}{d\alpha} + 2 \sum_i \sum_{a'a} \sum_{rs} \sum_{\phi} \tilde{T}_{rs,i}^{\phi} \frac{d\mathcal{C}_{a'\phi}}{d\alpha} \mathcal{R}_{a'a}(ri|sa) , \end{aligned} \quad (37)$$

by defining an “explicit” 1-RDM derivative matrix

$$\check{R}_{a'a}^{(1),\text{exp}} = 2 \sum_i \sum_{rs} \sum_{\phi} \tilde{T}_{rs,i}^{\phi} \mathcal{C}_{a'\phi}(ri|sa) . \quad (38)$$

The remaining contributions are the derivatives of  $\mathcal{C}_{a\phi}$  [the second term in Eq. (37)] and  $e_{\phi}$  [the sixth term in Eq. (31)].

#### D. Derivatives of $\mathcal{C}$ and $e_{\phi}$

The remaining terms to be transformed are those involving the derivatives of  $\mathcal{C}$  and  $e_{\phi}$ :

$$2 \sum_i \sum_{a'a} \sum_{rs} \sum_{\phi} \tilde{T}_{rs,i}^{\phi} \frac{d\mathcal{C}_{a'\phi}}{d\alpha} \mathcal{R}_{a'a}(ri|sa) + \sum_i \sum_{rs} \sum_{\phi} \tilde{T}_{rs,i}^{\phi} T_{rs,i}^{\phi} \frac{de_{\phi}}{d\alpha} , \quad (39)$$

which are the sixth term in Eq. (31) and the second term in Eq. (37). Let us first analyze the derivative of  $\mathcal{C}$ .

As in the case of the MO coefficient derivative [Eq. (9)], the derivative of the internally contracted configuration coefficient is also written as a unitary transformation:

$$\frac{d\mathcal{C}_{a\phi}}{d\alpha} = \sum_{\chi}^{\text{all}} \mathcal{C}_{a\chi} \mathcal{U}_{\chi\phi}^{\alpha} . \quad (40)$$

From the orthonormality condition, we obtain

$$\mathcal{U}_{\phi\chi}^\alpha + \mathcal{U}_{\chi\phi}^\alpha + \mathcal{R}_{\phi\chi}^\alpha = 0, \quad (41)$$

where

$$\mathcal{R}_{\phi\chi}^\alpha = \sum_{a'a} \mathcal{C}_{a'\phi} \mathcal{C}_{a\chi} \frac{d\mathcal{R}_{a'a}}{d\alpha}. \quad (42)$$

Clearly, because of Eq. (41), the upper-triangular part of  $\mathcal{U}_{\phi\chi}^\alpha$  is redundant. From the diagonality of the Fock-like matrix,

$$\sum_{a'a} \mathcal{C}_{a'\phi} \mathcal{C}_{a\chi} \mathcal{K}_{a'a} = e_{\phi\chi} = \delta_{\phi\chi} e_\phi, \quad (43)$$

where  $\delta_{\phi\chi}$  is the Kronecker delta, we can derive

$$\mathcal{U}_{\phi\chi}^\alpha = \frac{1}{e_\chi - e_\phi} (e_{\phi\chi}^\alpha - e_\chi \mathcal{R}_{\phi\chi}^\alpha), \quad (44)$$

where

$$e_{\phi\chi}^\alpha = \sum_{a'a} \mathcal{C}_{a'\phi} \mathcal{C}_{a\chi} \frac{d\mathcal{K}_{a'a}}{d\alpha}. \quad (45)$$

It should be noted that the expressions that appear in this subsection are very similar to those of  $U_{pq}^\alpha$  for HF.<sup>39</sup>

Using the above expressions, the first term in Eq. (39) can be transformed:

$$\begin{aligned} & 2 \sum_i \sum_{a'a} \sum_{rs} \sum_\phi \tilde{T}_{rs,i}^\phi \frac{d\mathcal{C}_{a'\phi}}{d\alpha} \mathcal{R}_{a'a}(ri|sa) \\ &= 2 \sum_i \sum_{rs} \sum_{\phi < \chi} (e_\phi - e_\chi) \tilde{T}_{rs,i}^\phi T_{rs,i}^\chi \mathcal{U}_{\chi\phi}^\alpha - \sum_i \sum_a \sum_{rs} \sum_{\phi\chi} \tilde{T}_{rs,i}^\phi \mathcal{S}_{\chi a}(ri|sa) \mathcal{R}_{\phi\chi}^\alpha \\ &= \sum_i \sum_{rs} \sum_{\phi \neq \chi} \tilde{T}_{rs,i}^\phi T_{rs,i}^\chi \left( e_{\phi\chi}^\alpha - \frac{e_\phi + e_\chi}{2} \mathcal{R}_{\phi\chi}^\alpha \right) - \sum_i \sum_a \sum_{rs} \sum_{\phi\chi} \tilde{T}_{rs,i}^\phi \mathcal{S}_{\chi a}(ri|sa) \mathcal{R}_{\phi\chi}^\alpha, \quad (46) \end{aligned}$$

where

$$\sum_{a'a} \left( \tilde{T}_{rs,i}^\phi \mathcal{C}_{a'\chi} - \tilde{T}_{rs,i}^\chi \mathcal{C}_{a'\phi} \right) \mathcal{R}_{a'a}(ri|sa) = (e_\phi - e_\chi) \tilde{T}_{rs,i}^\phi T_{rs,i}^\chi \quad (47)$$

has been used. On adding the second term in Eq. (39) to Eq. (46), Eq. (39) becomes

$$\begin{aligned} & 2 \sum_i \sum_{a'a} \sum_{rs} \sum_\phi \tilde{T}_{rs,i}^\phi \frac{d\mathcal{C}_{a'\phi}}{d\alpha} \mathcal{R}_{a'a}(ri|sa) + \sum_i \sum_a \sum_{rs} \sum_\phi \tilde{T}_{rs,i}^\phi T_{rs,i}^\phi \frac{de_\phi}{d\alpha} \\ &= \sum_{\phi\chi} \mathcal{D}_{\phi\chi} \left( e_{\phi\chi}^\alpha - \frac{e_\phi + e_\chi}{2} \mathcal{R}_{\phi\chi}^\alpha \right) - \sum_i \sum_a \sum_{rs} \sum_{\phi\chi} \tilde{T}_{rs,i}^\phi \mathcal{S}_{\chi a}(ri|sa) \mathcal{R}_{\phi\chi}^\alpha, \quad (48) \end{aligned}$$

where

$$\mathcal{D}_{\phi\chi} = \sum_i \sum_{rs} \tilde{T}_{rs,i}^{\phi} T_{rs,i}^{\chi} . \quad (49)$$

By defining the “implicit” 1-RDM derivative contribution

$$\check{R}_{a'a}^{(1),\text{imp}} = - \sum_{\phi\chi} \mathcal{C}_{a'\phi} \mathcal{C}_{a\chi} \left( \frac{e_{\phi} + e_{\chi}}{2} \mathcal{D}_{\phi\chi} + \sum_i \sum_a \sum_{rs} \tilde{T}_{rs,i}^{\phi} \mathcal{S}_{\chi a}(ri|sa) \right) \quad (50)$$

and a contracted active orbital density matrix

$$\bar{d}_{a'a} = \sum_{\phi\chi} \mathcal{C}_{a'\phi} \mathcal{C}_{a\chi} \mathcal{D}_{\phi\chi} , \quad (51)$$

we obtain a simple expression for Eq. (39):

$$\begin{aligned} & 2 \sum_i \sum_{a'a} \sum_{rs} \sum_{\phi} \tilde{T}_{rs,i}^{\phi} \frac{d\mathcal{C}_{a'\phi}}{d\alpha} \mathcal{R}_{a'a}(ri|sa) + \sum_i \sum_a \sum_{rs} \sum_{\phi} \tilde{T}_{rs,i}^{\phi} T_{rs,i}^{\phi} \frac{de_{\phi}}{d\alpha} \\ &= \sum_{a'a} \bar{d}_{a'a} \frac{dK_{a'a}}{d\alpha} + \sum_{a'a} \check{R}_{a'a}^{(1),\text{imp}} \frac{dR_{a'a}^{(1)}}{d\alpha} . \end{aligned} \quad (52)$$

It should be noted that the derivation presented here can be performed directly by applying the noncanonical approach.<sup>41,42</sup> Briefly speaking, the rotation parameter  $\mathcal{U}_{\phi\chi}^{\alpha}$  for the blocks in which rotations do not affect the energy (rotational invariance) can be substituted by the integral derivative of the overlap matrix ( $\mathcal{R}_{\phi\chi}^{\alpha}$ ) usually scaled by  $-1/2$ . The derivative of the eigenvalue matrix is then no longer a diagonal matrix (noncanonical). This approach has been extensively employed in developing correlated SR methods (such as MP2<sup>41,42</sup>). The same applies in the internally contracted space, and the final expression is useful for avoiding division by a difference between orbital energies, which may cause numerical problems when two eigenvalues are nearly degenerate. In addition to the simplified expression, we also find that the perturbation energy is invariant under rotations of internally contracted configurations.

## E. Derivative of K

In this subsection, the first term in Eq. (52) is further transformed. First, it is expanded as

$$\begin{aligned} \sum_{a'a} \bar{d}_{a'a} \frac{dK_{a'a}}{d\alpha} &= \sum_{ab} d_{ab}^{\text{PT2,act}} \frac{dh_{ab}^{\text{eff}}}{d\alpha} + \sum_{abcd} D_{abcd}^{\text{PT2,act}} \frac{d(ab|cd)}{d\alpha} \\ &\quad + \sum_{ab} \check{R}_{ab}^{(1),\text{act}} \frac{dR_{ab}^{(1)}}{d\alpha} + \sum_{abcd} \check{R}_{ab,cd}^{(2),\text{act}} \frac{dR_{ab,cd}^{(2)}}{d\alpha} . \end{aligned} \quad (53)$$

From Eq. (22), we easily find that

$$d_{ab}^{\text{PT2,act}} = - \sum_c \bar{d}_{ca} R_{cb}^{(1)} , \quad (54)$$

$$D_{abcd}^{\text{PT2,act}} = - \sum_e \bar{d}_{ea} R_{ec,bd}^{(2)} , \quad (55)$$

$$\check{R}_{ab}^{(1),\text{act}} = - \sum_c \bar{d}_{ac} h_{cb}^{\text{eff}} , \quad (56)$$

$$\check{R}_{ab,cd}^{(2),\text{act}} = - \sum_e \bar{d}_{ae} (ec|bd) . \quad (57)$$

Further transformations of RDM derivatives will be discussed in the next subsection.

Then, the derivatives of  $h_{ab}^{\text{eff}}$  and the two-electron integrals have to be further transformed:

$$\begin{aligned} \sum_{ab} d_{ab}^{\text{PT2,act}} \frac{dh_{ab}^{\text{eff}}}{d\alpha} &= \sum_{ab} d_{ab}^{\text{PT2,act}} \left( h_{ab}^\alpha + \sum_i (2(ab|ii)^\alpha - (ai|bi)^\alpha) \right) \\ &\quad + \sum_p \sum_{ab} (h_{pb}^{\text{eff}} U_{pa}^\alpha + h_{ap}^{\text{eff}} U_{pb}^\alpha) d_{ab}^{\text{PT2,act}} \\ &\quad + \sum_p \sum_i U_{pi}^\alpha \sum_{ab} (4(ab|pi) - (ap|bi) - (ai|bp)) d_{ab}^{\text{PT2,act}} , \end{aligned} \quad (58)$$

where  $(pq|mn)^\alpha$  is the skeleton derivative of the two-electron integral defined by

$$(pq|mn)^\alpha = \sum_{\mu\nu\rho\sigma} C_{\mu p} C_{\nu q} C_{\rho m} C_{\sigma n} \frac{d(\mu\nu|\rho\sigma)}{d\alpha} . \quad (59)$$

In practice, we may first transform  $d_{ab}^{\text{PT2,act}}$  to the AO basis:

$$d_{\mu\nu}^{\text{PT2,act}} = \sum_{ab} C_{\mu a} C_{\nu b} d_{ab}^{\text{PT2,act}} . \quad (60)$$

The third term in Eq. (58) can be computed by a Fock-like transformation [Eq. (36)] followed by an AO-to-MO transformation:

$$\begin{aligned} \sum_{ab} d_{ab}^{\text{PT2,act}} \frac{dh_{ab}^{\text{eff}}}{d\alpha} &= \sum_{\mu\nu} d_{\mu\nu}^{\text{PT2,act}} \frac{dh_{\mu\nu}}{d\alpha} \\ &+ \sum_{\mu\nu\rho\sigma} \left( d_{\mu\nu}^{\text{PT2,act}} d_{\rho\sigma}^{\text{core}} - \frac{1}{4} d_{\mu\rho}^{\text{PT2,act}} d_{\nu\sigma}^{\text{core}} - \frac{1}{4} d_{\mu\sigma}^{\text{PT2,act}} d_{\nu\nu}^{\text{core}} \right) \frac{d(\mu\nu|\rho\sigma)}{d\alpha} \\ &+ 4 \sum_p^{\text{all}} \sum_q^{\text{docc}} G_{pq} [\mathbf{d}^{\text{PT2,act}}] U_{pq}^\alpha + 2 \sum_p^{\text{all}} \sum_q^{\text{act}} \sum_a^{\text{act}} h_{pa}^{\text{eff}} d_{aq}^{\text{PT2,act}} U_{pq}^\alpha, \end{aligned} \quad (61)$$

where  $d_{\mu\nu}^{\text{core}}$  is the core density matrix:

$$d_{\mu\nu}^{\text{core}} = 2 \sum_i C_{\mu i} C_{\nu i}. \quad (62)$$

The derivatives of the two-electron integrals can be transformed as

$$\begin{aligned} \sum_{abcd} D_{abcd}^{\text{PT2,act}} \frac{d(ab|cd)}{d\alpha} &= \sum_{abcd} D_{abcd}^{\text{PT2,act}} (ab|cd)^\alpha \\ &+ \sum_p^{\text{all}} \sum_q^{\text{act}} \sum_{abc} \hat{v}_{pq}^{abc}(pa|bc) D_{qabc}^{\text{PT2,act}} U_{pq}^\alpha, \end{aligned} \quad (63)$$

where  $\hat{v}_{pq}^{abc}$  generates the following cyclic permutation:

$$\hat{v}_{pq}^{abc}(pa|bc) D_{qabc}^{\text{PT2,act}} = (pa|bc) D_{qabc}^{\text{PT2,act}} + (ap|bc) D_{aqbc}^{\text{PT2,act}} + (ab|pc) D_{abqc}^{\text{PT2,act}} + (ab|cq) D_{abcq}^{\text{PT2,act}}. \quad (64)$$

The dimension of  $(pb|cd)$  is  $N_{\text{MO}} \times N_{\text{act}}^3$ , so it can be easily stored in memory for medium-sized systems.

Finally, the derivative derived in this subsection can be written as

$$\begin{aligned} \sum_{a'a} \bar{d}_{a'a} \frac{dK_{a'a}}{d\alpha} &= \sum_{\mu\nu} d_{\mu\nu}^{\text{PT2,act}} \frac{dh_{\mu\nu}}{d\alpha} \\ &+ \sum_{\mu\nu\rho\sigma} \left( D_{\mu\nu\rho\sigma}^{\text{PT2,act}} + d_{\mu\nu}^{\text{PT2,act}} d_{\rho\sigma}^{\text{core}} - \frac{1}{4} d_{\mu\rho}^{\text{PT2,act}} d_{\nu\sigma}^{\text{core}} - \frac{1}{4} d_{\mu\sigma}^{\text{PT2,act}} d_{\nu\nu}^{\text{core}} \right) \frac{d(\mu\nu|\rho\sigma)}{d\alpha} \\ &+ \sum_{pq} \mathcal{L}_{pq}^{\text{PT2,act}} U_{pq}^\alpha + \sum_{ab} \check{R}_{ab}^{(1),\text{act}} \frac{dR_{ab}^{(1)}}{d\alpha} + \sum_{abcd} \check{R}_{ab,cd}^{(2),\text{act}} \frac{dR_{ab,cd}^{(2)}}{d\alpha}, \end{aligned} \quad (65)$$

where

$$D_{\mu\nu\rho\sigma}^{\text{PT2,act}} = \sum_{abcd}^{\text{act}} C_{\mu a} C_{\nu b} C_{\rho c} C_{\sigma d} D_{abcd}^{\text{PT2,act}} \quad (66)$$



and

$$\mathcal{L}_{pq}^{\text{PT2,act}} = \begin{cases} 4G_{pq} [\mathbf{d}^{\text{PT2,act}}] & q \in \text{docc} , \\ 2 \sum_a h_{pa}^{\text{eff}} d_{aq}^{\text{PT2,act}} + \sum_{abc} \hat{v}_{pq}^{abc} (pa|bc) D_{qabc}^{\text{PT2,act}} & q \in \text{act} . \end{cases} \quad (67)$$

## F. Derivatives of RDMs

In this subsection, we have to distinguish states, so they are indicated explicitly. From Eqs. (31), (37), (52), and (65), the derivative of  $E_S^{(-1)}$  is transformed as

$$\begin{aligned} \frac{dE_S^{(-1)}}{d\alpha} &= \sum_{\mu\nu} d_{\mu\nu}^{S,\text{PT2}} \frac{dh_{\mu\nu}}{d\alpha} + \sum_{\mu\nu\rho\sigma} D_{\mu\nu\rho\sigma}^{S,\text{PT2}} \frac{d(\mu\nu|\rho\sigma)}{d\alpha} \\ &+ \sum_{pq} \mathcal{L}_{pq}^{S,\text{PT2}} U_{pq}^\alpha + \sum_{ab} \sum_T \check{R}_{ab}^{S,T,(1)} \frac{dR_{ab}^{T,(1)}}{d\alpha} + \sum_{abcd} \check{R}_{ab,cd}^{S,(2)} \frac{dR_{ab,cd}^{S,(2)}}{d\alpha} , \end{aligned} \quad (68)$$

where

$$d_{\mu\nu}^{S,\text{PT2}} = d_{\mu\nu}^{S,\text{PT2,inact}} + d_{\mu\nu}^{S,\text{PT2,act}} , \quad (69)$$

$$\begin{aligned} D_{\mu\nu\rho\sigma}^{S,\text{PT2}} &= D_{\mu\nu\rho\sigma}^{S,\text{PT2,inact}} + D_{\mu\nu\rho\sigma}^{S,\text{PT2,act}} \\ &+ d_{\mu\nu}^{S,\text{PT2,inact}} d_{\rho\sigma}^{\text{SA}} - \frac{1}{4} d_{\mu\rho}^{S,\text{PT2,inact}} d_{\nu\sigma}^{\text{SA}} - \frac{1}{4} d_{\mu\sigma}^{S,\text{PT2,inact}} d_{\nu\rho}^{\text{SA}} \\ &+ d_{\mu\nu}^{S,\text{PT2,act}} d_{\rho\sigma}^{\text{core}} - \frac{1}{4} d_{\mu\rho}^{S,\text{PT2,act}} d_{\nu\sigma}^{\text{core}} - \frac{1}{4} d_{\mu\sigma}^{S,\text{PT2,act}} d_{\nu\rho}^{\text{core}} , \end{aligned} \quad (70)$$

$$\mathcal{L}_{pq}^{S,\text{PT2,o}} = \mathcal{L}_{pq}^{S,\text{PT2,inact}} + \mathcal{L}_{pq}^{S,\text{PT2,act}} , \quad (71)$$

$$\check{R}_{ab}^{S,T,(1)} = \omega_T \check{R}_{ab}^{T,(1),\text{inact}} + \delta_{ST} \left( \check{R}_{ab}^{S,(1),\text{exp}} + \check{R}_{ab}^{S,(1),\text{imp}} + \check{R}_{ab}^{S,(1),\text{act}} \right) , \quad (72)$$

$$\check{R}_{ab,cd}^{S,(2)} = \check{R}_{ab,cd}^{S,(2),\text{act}} . \quad (73)$$

In this subsection, the last two terms in Eq. (68) are transformed. In evaluating the  $E_S^{(-1)}$  term, the 1-RDM and 2-RDM are sufficient, but RDMs up to 4-RDM have to be evaluated for other terms. Although this subsection presents the derivative of 1-RDM only, the extension to higher-order RDMs is conceptually straightforward.

Following Ref. 33 with a slight modification, the derivative of  $c_I^S$  is written in terms of the state rotation parameter  $W_{AB}^\alpha$ :

$$\frac{dc_I^A}{d\alpha} = \sum_B^N c_I^B W_{AB}^\alpha . \quad (74)$$

From the orthonormality of states,

$$\sum_I^N c_I^A c_I^B = \delta_{AB} , \quad (75)$$

one can easily derive the relation to be satisfied:

$$W_{AB}^\alpha + W_{BA}^\alpha = 0 . \quad (76)$$

Rotations for  $A \geq B$  are redundant, so they are eliminated when solving the CP or Z-vector equations.

Following the derivation given in Appendix B, the derivative of 1-RDM can be written as

$$\frac{dR_{ab}^{S,(1)}}{d\alpha} = \sum_{IJ} (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) c_J^S V_{SI}^\alpha + \sum_T^\Omega s_{ST} \left( R_{ab}^{ST,(1)} + R_{ab}^{TS,(1)} \right) W_{ST}^\alpha , \quad (77)$$

where  $s_{ST} = -\hat{\tau}_{ST}$ , 1, and 0 for  $S > T$ ,  $S < T$ , and  $S = T$ , respectively, and  $R_{ab}^{ST,(1)}$  is the one-particle transition RDM. For higher-order RDMs,  $\gamma_{ab}^{IJ}$  and  $R_{ab}^{ST,(1)}$  are replaced with the coupling coefficients and RDMs for the corresponding particle number.

With these quantities defined, we can expand the derivatives of  $R_{ab}^{S,(1)}$  and  $R_{abcd}^{S,(2)}$ :

$$\sum_{ab} \check{R}_{ab}^{S,(1)} \frac{dR_{ab}^{S,(1)}}{d\alpha} + \sum_{abcd} \check{R}_{ab,cd}^{S,(2)} \frac{dR_{ab,cd}^{S,(2)}}{d\alpha} = \sum_S^\Omega \sum_I^N \mathcal{L}_{SI}^{\text{PT2,c}} V_{SI}^\alpha + \sum_{S<T}^\Omega \mathcal{L}_{ST}^{\text{PT2,s}} W_{ST}^\alpha , \quad (78)$$

where

$$\mathcal{L}_{SI}^{\text{PT2,c}} = \sum_{ab} \check{R}_{ab}^{S,(1)} (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) c_J^S + \sum_{abcd} \check{R}_{abcd}^{S,(2)} (\Gamma_{acbd}^{IJ} + \Gamma_{acbd}^{JI}) c_J^S \quad (79)$$

and

$$\mathcal{L}_{ST}^{\text{PT2,s}} = s_{ST} \sum_{ab} \check{R}_{ab}^{S,(1)} \left( R_{ab}^{ST,(1)} + R_{ab}^{TS,(1)} \right) + s_{ST} \sum_{abcd} \check{R}_{ab,cd}^{S,(2)} \left( R_{ab,cd}^{ST,(2)} + R_{ab,cd}^{TS,(2)} \right) . \quad (80)$$

Clearly, the state rotation term [the second term in Eq. (78)] is evaluated only when the reference wavefunction is state-averaged.

## G. Transformation to independent orbital rotation

Up to this point, we have not distinguished between independent and non-independent rotation pairs of  $U_{pq}^\alpha$ , and  $\mathcal{L}_{pq}^{\text{PT2,o}}$  still contain redundant ( $p \leq q$ ) variables. Non-independent

pairs are rotations that do not change the energy. It is possible to solve the CP or Z-vector equations involving non-independent and redundant pairs in some cases, but the solution procedure can exhibit convergence problems and is computationally inefficient, so non-independent pairs should be eliminated. In the standard derivations, it appears as if they are eliminated by using the orthonormality condition,

$$U_{pq}^\alpha + U_{qp}^\alpha + S_{pq}^\alpha = 0, \quad (81)$$

as in the derivation of the MP2 gradient.<sup>40</sup> In the present derivation, however, there are so many terms that it is rather complicated to apply this protocol, and therefore the following term is manipulated in the computer code after all the contributions have been computed:

$$\sum_{pq}^{\text{all}} \mathcal{L}_{pq}^{\text{PT2,o}} U_{pq}^\alpha = \sum_{pq}^{\text{indep}} \tilde{\mathcal{L}}_{pq}^{\text{PT2,o}} \tilde{U}_{pq}^\alpha - \frac{1}{2} \sum_{pq}^{\text{all}} \mathcal{L}_{pq}^{\text{PT2,o}} S_{pq}^\alpha, \quad (82)$$

where

$$\tilde{\mathcal{L}}_{pq}^{\text{PT2,o}} = (1 - \hat{\tau}_{pq}) \mathcal{L}_{pq}^{\text{PT2,o}}. \quad (83)$$

The  $(1 - \hat{\tau}_{pq})$  operation removes the redundant variables.

The first-order derivative of the PC-NEVPT2 energy can be written as

$$\begin{aligned} \frac{dE^{\text{PC-NEVPT2}}}{d\alpha} &= \sum_{\mu\nu} (d_{\mu\nu}^{\text{CAS}} + d_{\mu\nu}^{\text{PT2}}) \frac{dh_{\mu\nu}}{d\alpha} - \sum_{\mu\nu} (X_{\mu\nu}^{\text{CAS}} + X_{\mu\nu}^{\text{PT2}}) \frac{dS_{\mu\nu}}{d\alpha} \\ &\quad + \sum_{\mu\nu\rho\sigma} (D_{\mu\nu\rho\sigma}^{\text{CAS}} + D_{\mu\nu\rho\sigma}^{\text{PT2}}) \frac{d(\mu\nu|\rho\sigma)}{d\alpha} \\ &\quad + \sum_{pq}^{\text{indep}} \mathcal{L}_{pq}^{\text{o}} \tilde{U}_{pq}^\alpha + \sum_S^\Omega \sum_I^N \mathcal{L}_{SI}^{\text{c}} V_{SI}^\alpha + \sum_{S<T}^\Omega \mathcal{L}_{ST}^{\text{s}} W_{ST}^\alpha, \end{aligned} \quad (84)$$

where the unrelaxed energy-weighted PT2 density matrix is

$$X_{\mu\nu}^{\text{PT2}} = \frac{1}{2} \sum_{pq}^{\text{all}} C_{\mu p} C_{\nu q} \mathcal{L}_{pq}^{\text{PT2,o}}. \quad (85)$$

The partial derivatives of the energy with respect to orbital, configuration, and state are also defined by  $\mathcal{L}_{pq}^{\text{o}} = \mathcal{L}_{pq}^{\text{CAS,o}} + \tilde{\mathcal{L}}_{pq}^{\text{PT2,o}}$ ,  $\mathcal{L}_{SI}^{\text{c}} = \mathcal{L}_{SI}^{\text{PT2,c}}$ , and  $\mathcal{L}_{ST}^{\text{s}} = \mathcal{L}_{ST}^{\text{PT2,s}}$ , respectively.

## H. CP-MCSCF

At this point, the first-order derivative of the energy is given by the simplified derivative expression in Eq. (25) or (84) using unrelaxed density matrices, although the actual implementation and computation have to include the terms of the other seven excitation classes in

the density matrices and partial derivative contributions. The remaining task is to evaluate the response contributions, i.e., the last three terms in Eq. (25) or (84).

A CP-MCSCF is much more complicated than an SR CP-HF. This is because the variational parameters (MO and configuration coefficients) are coupled, so both parameters have to be simultaneously optimized. In addition, if different weighting factors are adopted, state rotations have to be taken into account to satisfy the orthonormality of states (see Sec. II F), so the additional state rotation parameters  $\mathbf{W}^\alpha$  must be considered. As a consequence, one can derive a CP-MCSCF in the form of a super-matrix expression:<sup>33</sup>

$$\begin{pmatrix} \mathbf{A}^{\text{oo}} & \mathbf{A}^{\text{oc}} & \mathbf{A}^{\text{os}} \\ \mathbf{A}^{\text{co}} & \mathbf{A}^{\text{cc}} & 0 \\ \mathbf{A}^{\text{so}} & 0 & \mathbf{A}^{\text{ss}} \end{pmatrix} \begin{pmatrix} \mathbf{U}^\alpha \\ \mathbf{V}^\alpha \\ \mathbf{W}^\alpha \end{pmatrix} = \begin{pmatrix} \mathbf{B}^{\alpha,\text{o}} \\ \mathbf{B}^{\alpha,\text{c}} \\ \mathbf{B}^{\alpha,\text{s}} \end{pmatrix}, \quad (86)$$

where the  $\mathbf{A}$  and  $\mathbf{B}^\alpha$  matrices are all defined in Ref. 33 or, for SS-MCSCF, in Ref. 35. Usually, the number of CP-MCSCF equations to be solved scales as the number of atoms, since the responses of the wavefunction parameters depend on the derivative parameter  $\alpha$ . However, this can be avoided by solving the Z-vector equation instead,

$$\begin{pmatrix} \mathbf{A}^{\text{oo}} & \mathbf{A}^{\text{oc}} & \mathbf{A}^{\text{os}} \\ \mathbf{A}^{\text{co}} & \mathbf{A}^{\text{cc}} & 0 \\ \mathbf{A}^{\text{so}} & 0 & \mathbf{A}^{\text{ss}} \end{pmatrix} \begin{pmatrix} \mathbf{Z}^{\text{o}} \\ \mathbf{Z}^{\text{c}} \\ \mathbf{Z}^{\text{s}} \end{pmatrix} = \begin{pmatrix} \mathcal{L}^{\text{o}} \\ \mathcal{L}^{\text{c}} \\ \mathcal{L}^{\text{s}} \end{pmatrix}, \quad (87)$$

because the super-matrix  $\mathbf{A}$  is symmetric. Once we have obtained  $\mathbf{Z}^{\text{o}}$ ,  $\mathbf{Z}^{\text{c}}$ , and  $\mathbf{Z}^{\text{s}}$  as solutions, the response contribution to the gradient can be finally computed by

$$(\mathcal{L}^{\text{o}})^{\text{T}} \cdot \mathbf{U}^\alpha + (\mathcal{L}^{\text{c}})^{\text{T}} \cdot \mathbf{V}^\alpha + (\mathcal{L}^{\text{s}})^{\text{T}} \cdot \mathbf{W}^\alpha = (\mathbf{Z}^{\text{o}})^{\text{T}} \cdot \mathbf{B}^{\alpha,\text{o}} + (\mathbf{Z}^{\text{c}})^{\text{T}} \cdot \mathbf{B}^{\alpha,\text{c}} + (\mathbf{Z}^{\text{s}})^{\text{T}} \cdot \mathbf{B}^{\alpha,\text{s}}, \quad (88)$$

where the superscript “T” indicates the transpose of a matrix or vector. Since the  $\mathbf{B}^\alpha$  matrices consist of derivatives of integrals, no additional response equations are solved.

If the reference wavefunction is an SS-CASSCF or SA-CASSCF with different weighting factors, no further complications arise. On the other hand, if the reference wavefunction is an SA-CASSCF with identical weighting factors, then the elements of  $\mathbf{A}^{\text{so}}$ ,  $\mathbf{A}^{\text{os}}$ ,  $\mathbf{A}^{\text{ss}}$ , and  $\mathbf{B}^{\alpha,\text{s}}$  for equally weighted states will be zero, and a convergence problem appears. This is because the residue evaluated during the Z-vector iteration does not decrease, since there is no choice of  $\mathbf{Z}^{\text{s}}$  for equal weighting factors that results in nonzero values. In such a case, as

discussed in Ref. 33, the rotation between internal states with equal weighting factor should be written as

$$W_{ST}^\alpha = \frac{-1}{E_T - E_S} \left\{ \sum_{IJ} c_I^S c_J^T H_{IJ}^\alpha - \frac{1}{2} \sum_{pq} S_{pq}^\alpha \varepsilon_{pq}^{ST} + \sum_{pq}^{\text{indep}} (\varepsilon_{pq}^{ST} - \varepsilon_{qp}^{ST}) \tilde{U}_{pq}^\alpha \right\}, \quad (89)$$

where  $\varepsilon_{pq}^{ST}$  is the Lagrangian matrix for states  $S$  and  $T$ :

$$\varepsilon_{pq}^{ST} = \sum_m h_{pm} \gamma_{qm}^{ST} + \sum_{mno} (pm|no) \Gamma_{qmno}^{ST}. \quad (90)$$

Equation (89) necessitates division by the difference in CASSCF energies, so it may blow up when two energies are almost the same. At the beginning of the Z-vector iteration, the orbital derivative  $\mathcal{L}_{pq}^o$  is modified using the last term in Eq. (89):

$$\mathcal{L}_{pq}^o \leftarrow \mathcal{L}_{pq}^o + \sum_{\substack{S < T \\ \omega_S = \omega_T}} \frac{-\mathcal{L}_{pq}^s}{E_T - E_S} (\varepsilon_{pq}^{ST} - \varepsilon_{qp}^{ST}). \quad (91)$$

Once the Z-vector equation has been solved, the integral derivative contributions, according to the first two terms, are evaluated:

$$Z_{ST}^s B_{ST}^{\alpha,s} = \frac{-\mathcal{L}_{ST}^s}{E_T - E_S} \left\{ \sum_{IJ} c_I^S c_J^T H_{IJ}^\alpha - \frac{1}{2} \sum_{pq} S_{pq}^\alpha \varepsilon_{pq}^{ST} \right\} \quad \text{for } \omega_S = \omega_T. \quad (92)$$

Finally, the first-order derivative of the PC-NEVPT2 energy is transformed as

$$\begin{aligned} \frac{dE^{\text{PC-NEVPT2}}}{d\alpha} &= \sum_{\mu\nu} (d_{\mu\nu}^{\text{CAS}} + d_{\mu\nu}^{\text{PT2}}) \frac{dh_{\mu\nu}}{d\alpha} - \sum_{\mu\nu} (X_{\mu\nu}^{\text{CAS}} + X_{\mu\nu}^{\text{PT2}}) \frac{dS_{\mu\nu}}{d\alpha} \\ &\quad + \sum_{\mu\nu\rho\sigma} (D_{\mu\nu\rho\sigma}^{\text{CAS}} + D_{\mu\nu\rho\sigma}^{\text{PT2}}) \frac{d(\mu\nu|\rho\sigma)}{d\alpha} \\ &\quad + (\mathbf{Z}^o)^T \cdot \mathbf{B}^{\alpha,o} + (\mathbf{Z}^c)^T \cdot \mathbf{B}^{\alpha,c} + (\mathbf{Z}^s)^T \cdot \mathbf{B}^{\alpha,s}. \end{aligned} \quad (93)$$

In principle, it is possible to evaluate the response contributions [the last three terms in Eq. (93)] by transforming the density matrices to the AO basis, but here they are computed using the dot product of the  $\mathbf{Z}$  and  $\mathbf{B}^\alpha$  matrices, following the existing implementation.

## I. Implementation

The derived equations have been (manually) implemented in a local version of GAMESS-US.<sup>43</sup> Previously, analytic derivatives of the fully internally contracted CASPT2 were im-

plemented with automatic derivation and code generation,<sup>19</sup> but the expressions for PC-NEVPT2 seem to be somewhat simpler and amenable to manual derivation and implementation. This is probably because couplings with other excitation classes are zero and each excitation energy can be computed independently in PC-NEVPT2. The gradient computed analytically usually agrees with the numerical gradient up to  $10^{-6}$  hartree/bohr, if numerical gradients are carefully evaluated. The reduced density matrices are constructed using Slater determinants, instead of configuration state functions, using the resolution of the identity technique.<sup>44</sup> The frozen core approximation excludes core orbitals from correlation during NEVPT2 calculations. This option was also implemented and required some minor modifications as given in Appendix D.

In PC-NEVPT2, the linear dependence of the overlap matrix in the internally contracted basis ( $\mathcal{R}$ ; expressions for each excitation class are given in Ref. 15) have to be removed before solving the generalized eigenvalue problem [Eq. (21)]. At present, the linear dependence threshold is set to  $10^{-6}$ . Although the choice of this threshold affects the energy, the accuracy of the gradient is not affected, because all post-energy processes for analytic derivatives are executed using only linearly independent vectors. It should be noted that it is possible to employ the following relation to simplify some expressions in Sec. IID:  $\mathcal{R}^{-1} = \mathcal{C}\mathcal{C}^\dagger$ . This relation is, however, applicable only if the linearly independent basis is complete, so it is employed only for  $E^{(-1)}$  and  $E^{(+1)}$  (no linear dependence is assumed<sup>15</sup>) in the present implementation. Whether the energy is continuous during a geometry optimization is a different matter. This issue has not been investigated carefully, but no such difficulties have been encountered in this study.

In principle, it is possible to develop analytic derivatives for the SC-NEVPT2 method in a similar manner, but a serious obstacle arises because SC-NEVPT2 is not invariant under rotations within inactive orbitals. This has been pointed out previously<sup>45</sup> and has been verified numerically.<sup>32</sup> It may necessitate division by differences in orbital energies within doubly occupied and virtual orbitals, and this can cause a numerical problem when the system has exactly or nearly degenerate orbitals. Although SC-NEVPT2 (energy only) has also been implemented, no SC-NEVPT2 calculations have been performed in this study, so the prefix “PC-” is omitted hereinafter unless there is a need to distinguish this case.

With an SA-CASSCF reference wavefunction, molecular orbitals are canonicalized with respect to the state-averaged density matrix in the present implementation. Another canon-

icalization, which is believed to be more accurate, uses a SS density matrix, but this has not been implemented. Extension to computation of the NEVPT2 energy should be straightforward, although this would take significant computational time with a large basis set. Fortunately, the overall difference between the two canonicalization approaches is not significant.<sup>38</sup>

Once gradients have been computed, it is straightforward to compute dipole moments analytically. It would be convenient to compute the relaxed (one-electron) density matrix  $d_{\mu\nu}^{\text{relax}}$  as a sum of the “state-specific” CASSCF density matrix  $d_{\mu\nu}^{\text{CAS}}$ , the perturbation density matrix  $d_{\mu\nu}^{\text{PT2}}$ , and the response density matrix  $d_{\mu\nu}^{\text{resp}}$  (given in Appendix C). Then, the dipole moment of the density-dependent part can be easily computed as the trace of  $-\mathbf{d}^{\text{relax}} \cdot \mathbf{x}$ , where  $\mathbf{x}$  is a dipole integral.

The computational cost of the present implementation is very much dependent on the number of AOs/MOs and active orbitals. If the number of AOs/MOs is large, then evaluation of the orbital derivative  $\mathcal{L}_{pq}^o$  is dominant. However, in practice, CASSCF, specifically the four-index transformation of two-electron integrals, is more time-consuming than the PC-NEVPT2 (+ SC-NEVPT2) calculation itself. Another time-consuming step is the transformation of two-electron integral derivatives in AOs to active MOs before solving the Z-vector. The transformed MO integrals are used in constructing  $\mathbf{B}^\alpha$  matrices, but this transformation can be avoided in principle. The present implementation relies on an existing one,<sup>35</sup> and this part may be improved or rewritten in the future. In the opposite case, when the number of active orbitals ( $N_{\text{act}}$ ) is greater than 10 for instance, it is the evaluations of the terms involving 4-RDM that are the most time-consuming, since the computational cost of these terms formally scales as  $\mathcal{O}(N_{\text{act}}^9)$ . The partial derivative calculations for NEVPT2 ( $\mathcal{L}_{pq}^o$ ,  $\mathcal{L}_{SI}^c$ , and  $\mathcal{L}_{ST}^s$ ) are approximately three times more expensive than the energy calculation for NEVPT2. Nevertheless, computing two-electron integral derivatives and solving the Z-vector usually takes more time than the NEVPT2 calculation itself does. For instance, a single-point gradient calculation of  $\text{C}_6\text{H}_{14}$  (*trans*-polyacetylene; 250 AOs) at the two-state averaged NEVPT2(12e,12o)/cc-pVDZ level of theory took 2110, 2689, 234, and 10 301 s for CASSCF, NEVPT2, two-electron integral derivative, and Z-vector calculations, respectively, using six CPU cores of an Intel Xeon E5-1650 v3 CPU. In any case, if SA-CASSCF geometry optimizations can be performed, so can PC-NEVPT2, in principle.

### III. COMPUTATIONAL DETAILS

All CASSCF, NEVPT2, and DFT calculations were performed with a locally modified version of GAMESS-US.<sup>43</sup> All geometry optimizations using the fully internally contracted CASPT2 method were performed with BAGEL.<sup>46,47</sup> With SA-CASSCF as reference, the extended multistate (XMS)<sup>20,48</sup> CASPT2 method within the SS-SR contraction scheme was employed. Density fitting (DF)<sup>18</sup> was applied in all (XMS-)CASPT2 calculations, but the prefix “DF-” is omitted here. No symmetry constraints were applied. The frozen core approximation was applied to all NEVPT2 and CASPT2 calculations unless otherwise stated. Second-order derivatives were numerically computed by a two-point stencil approach using the first-order analytic derivative.

First, single-state CASSCF, NEVPT2, and CASPT2 calculations were performed for ozone ( $\text{O}_3$ ) and ozonide ( $\text{O}_3^-$ ) using the (aug-)cc-pVXZ [ $X = \text{T}$  and  $\text{Q}$ ; abbreviated as (A)VXZ] basis sets.<sup>49,50</sup> Some DFT calculations using B3LYP,<sup>51,52</sup> CAM-B3LYP,<sup>53</sup> M06,<sup>54</sup> and  $\omega\text{B97X-D}$ <sup>55</sup> functionals were also performed using the VTZ basis set. All of valence active space was employed in the MC and MR methods: 18 electrons in 12 orbitals for  $\text{O}_3$  (18e, 12o), and 19 in 12 for  $\text{O}_3^-$  (19e, 12o), respectively.

Second, 0–0 transition energies of 2-, 4-, and 5-methylpyrimidine were computed. The reference CASSCF consisted of 10 electrons in 8 orbitals (10e, 8o), and NEVPT2 and XMS-CASPT2 calculations were successively performed. Since the method developed in this study is based on the state-specific NEVPT2, although it would be natural to compare NEVPT2 results with state-specific CASPT2 ones, the XMS-CASPT2 method has been employed here for technical reasons. A formal comparison would require the development of analytic derivatives of the quasidegenerate NEVPT2 method.<sup>56</sup> In all cases, the ground and target ( $\text{S}_1$ ;  $n \rightarrow \pi^*$  excitation) excited-state densities were averaged in solving the SA-CASSCF equation. In this calculation, the imaginary level-shift technique<sup>57</sup> was employed with a shift parameter of 0.2i in the XMS-CASPT2 calculations with SA-CASSCF as reference using the recent implementation<sup>58</sup> so that the CASPT2 iteration converged smoothly. DFT calculations with CAM-B3LYP and LC-BLYP<sup>59</sup> ( $\mu = 0.33$ ) were also performed. Geometry optimizations and energy refinement calculations were performed with the cc-pVTZ (VTZ) and aug-cc-pVTZ (AVTZ) basis sets, respectively.

Finally, CASSCF, NEVPT2, and XMS-CASPT2 calculations were performed for *trans*-



polyacetylene ( $\text{C}_{2n}\text{H}_{2n+2}$ ;  $n = 2, 3, 4, 5, 6$ ). All the geometry optimizations were performed for the plane molecule, and, as a result, optimized geometries of  $n = 2$  and 3 at the excited state were identified to be higher- and first-order saddle points, respectively. The two lowest states, corresponding to the ground and double-electron excited states, were averaged, and the active space consisted of all the  $\pi$  orbitals:  $2n$  electrons in  $2n$  orbitals, ( $2ne$ ,  $2no$ ). In the XMS-CASPT2 calculation, an imaginary shift parameter of 0.2i was employed. The calculations were performed with the cc-pVDZ (VDZ) basis set.

Single-point XMS-CASPT2 (fully internally contracted) calculations using the ionization potential–electron affinity (IPEA) shift technique<sup>60</sup> for optimized XMS-CASPT2 geometries were also performed with OpenMolcas version 18.0<sup>61</sup> using the default parameter of 0.25 and an imaginary shift parameter of 0.2i for the methylpyrimidine derivatives and *trans*-polyacetylenes. Note that analytic derivatives of CASPT2 had not been implemented in (Open)Molcas, while the IPEA shift was not available in BAGEL. The Cholesky decomposition (CD) technique<sup>62</sup> was applied to save disk space and computational cost using the HIGH CHOLESKY keyword, but the prefix “CD-” is omitted here.

Optimized coordinates of the methylpyrimidines and *trans*-polyacetylene are provided in the Supplementary Material.

## IV. RESULTS AND DISCUSSION

### A. Application to $\text{O}_3$

First, geometrical parameters, dipole moments, and harmonic vibrational frequencies of  $\text{O}_3$  were computed using CASSCF, NEVPT2, CASPT2, and DFT with several well-known exchange–correlation functionals, and the results are compared with experimental data in Table I.

CASSCF, NEVPT2, and CASPT2 with any basis sets agreed very well with the experimental O–O bond length, O–O–O angle, and dipole moment. In particular, NEVPT2/VTZ gave the best agreement with the experiments by Barbe *et al.*:<sup>63</sup> the largest deviations of the harmonic frequencies were only +2.1, −1.6, and +9.9  $\text{cm}^{-1}$  for  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$ , respectively. Larger basis sets did not always result in better agreement; a medium-sized basis set sometimes benefits from fortuitous error cancellation. CASPT2 was of similar quality for

the first-order properties (O–O distance, O–O–O angle, and dipole moment), but the agreement with the experimental harmonic vibrational frequencies, in particular  $\omega_3$ , was rather unsatisfactory, with deviations as large as  $100\text{ cm}^{-1}$ .

DFT calculations were not satisfactorily accurate. The deviations of the bond length and angle might be acceptable, but the dipole moment was overestimated by approximately 20% and the vibrational frequency by up to  $230\text{ cm}^{-1}$  (20%). Among the four DFT calculations, B3LYP/VTZ gave the best vibrational frequency. In DFT, it is common to scale computed frequencies with an empirical parameter. According to the Computational Chemistry Comparison and Benchmark DataBase (CCCBDB),<sup>64</sup> the scaling factors for B3LYP/VTZ and  $\omega$ B97X-D/VTZ are 0.965 and 0.956, respectively. Even if the computed frequencies are scaled with these numbers, the deviations are still fairly large. The scaled  $\omega_2$  of B3LYP agrees rather well, but the scaled  $\omega_1$  and  $\omega_3$  are still more than  $50\text{ cm}^{-1}$  away from the experimental results.

## B. Application to $\text{O}_3^-$

In a similar way, geometrical parameters and harmonic vibrational frequencies of  $\text{O}_3^-$  were computed, and the results are compared with experimental data in Table II. For DFT calculations, a restricted open-shell electron configuration was employed to obtain a spin-contamination-free wavefunction.

All the CASSCF, NEVPT2, and CASPT2 calculations predicted a bond length within the error bar of the experiment. However, the experimental O–O–O angle of  $111.7^\circ$  could not be predicted with any of the calculation methods employed in this study, all of which predicted a wider angle of  $115^\circ$ . A previous study<sup>67</sup> using CASPT2 with a smaller active space also reported a similar angle. Calculations without the frozen core approximation were also performed with the aug-cc-pwCVQZ basis set, but the predicted bond angle was always greater than  $115^\circ$ . It is likely that even higher-level calculations will need to be performed to reproduce the experimental bond angle.

CASSCF predicted a rather small  $\omega_3$  frequency, but the correction of the perturbative treatment was significant, and the resulting  $\omega_3$  predicted with NEVPT2 and CASPT2 agreed well with experiment. However, for the other frequencies  $\omega_1$  and  $\omega_2$ , CASSCF showed slightly better agreement with experiment than NEVPT2 did. Nevertheless, the  $\omega_2$  values predicted

TABLE I. Deviations of calculated values from experimental<sup>65,66</sup> bond length, angle, dipole moment  $\mu$ , and harmonic vibrational frequencies for  $\text{O}_3$ .

Method	Length ( $\text{\AA}$ )	Angle ( $^\circ$ )	$\mu$ (D)	$\omega_1$ ( $\text{cm}^{-1}$ )	$\omega_2$ ( $\text{cm}^{-1}$ )	$\omega_3$ ( $\text{cm}^{-1}$ )
CASSCF/VTZ	+0.013	+0.20	+0.019	-65.7	-3.2	-57.1
CASSCF/AVTZ	+0.012	+0.28	+0.021	-36.0	-11.8	-55.1
CASSCF/VQZ	+0.009	+0.31	+0.022	-32.0	-11.8	-59.9
CASSCF/AVQZ	+0.009	+0.37	+0.026	-43.8	-45.3	-8.9
NEVPT2/VTZ	+0.003	+0.20	-0.005	+2.1	-1.6	+9.9
NEVPT2/AVTZ	+0.001	+0.39	+0.019	+5.6	-1.9	+8.2
NEVPT2/VQZ	-0.001	+0.35	+0.003	+11.0	+3.9	+9.9
NEVPT2/AVQZ	-0.002	+0.45	+0.022	+14.6	+7.5	+19.4
CASPT2/VTZ	+0.016	+0.04	-0.035	-45.3	-25.1	-84.4
CASPT2/AVTZ	+0.017	+0.13	-0.011	-51.1	-27.3	-96.4
CASPT2/VQZ	+0.010	+0.21	-0.023	-31.1	-17.4	-65.6
CASPT2/AVQZ	+0.010	+0.29	-0.004	-32.2	-17.5	-68.6
B3LYP/VTZ	-0.015	+1.68	+0.112	+115.6	+29.3	+103.2
CAM-B3LYP/VTZ	-0.030	+1.42	+0.162	+181.8	+59.0	+203.0
M06/VTZ	-0.032	+1.37	+0.152	+193.4	+66.4	+197.8
$\omega$ B97X-D/VTZ	-0.031	+1.49	+0.161	+199.7	+65.4	+230.0
Experiment	1.2717 <sup>a</sup>	116.47 <sup>a</sup>	0.53 $\pm$ 0.02 <sup>b</sup>	1134.9 <sup>a</sup>	716.0 <sup>a</sup>	1089.2 <sup>a</sup>

<sup>a</sup>Ref. 66. <sup>b</sup>Ref. 65.

by both NEVPT2 and CASPT2 were still within the experimental error bar, as was  $\omega_1$  from CASPT2. NEVPT2 predicted a rather high  $\omega_1$ , but the deviation from the upper bound of the experimental error was not significant.

Again, the prediction of DFT was not as good as those of the methods based on MC wavefunctions, but the deviation was smaller than in the case of  $\text{O}_3$ . This is partly because the MR character of  $\text{O}_3^-$  is smaller than that of  $\text{O}_3$ :  $\text{O}_3^-$  and  $\text{O}_3$  exhibit MR characters of 90% and 80%, respectively, according to the CI coefficient of the most dominant Slater determinant.

TABLE II. Deviations of calculated values from experimental<sup>68</sup> bond length, angle, and harmonic vibrational frequencies for  $\text{O}_3^-$ .

Method	Length ( $\text{\AA}$ )	Angle ( $^\circ$ )	$\omega_1$ ( $\text{cm}^{-1}$ )	$\omega_2$ ( $\text{cm}^{-1}$ )	$\omega_3$ ( $\text{cm}^{-1}$ )
CASSCF/VTZ	+0.012	+3.86	+13.7	+19.4	-129.4
CASSCF/AVTZ	+0.009	+3.77	+16.4	+22.3	-189.5
CASSCF/VQZ	+0.008	+3.83	+19.8	+22.3	-160.0
CASSCF/AVQZ	+0.006	+3.80	+20.5	+26.4	-182.6
NEVPT2/VTZ	-0.003	+3.66	+53.0	+34.1	+6.7
NEVPT2/AVTZ	-0.006	+3.55	+56.1	+32.1	-12.8
NEVPT2/VQZ	-0.007	+3.60	+45.6	+37.4	+7.5
NEVPT2/AVQZ	-0.010	+3.55	+53.6	+39.0	-1.3
CASPT2/VTZ	+0.007	+3.90	+31.0	+21.1	-14.1
CASPT2/AVTZ	+0.006	+3.60	+15.4	+17.0	-30.1
CASPT2/VQZ	+0.001	+3.60	+34.7	+26.1	-8.6
CASPT2/AVQZ	+0.000	+3.61	+26.4	+24.6	-16.9
B3LYP/VTZ	-0.004	+3.80	+83.9	+44.4	+21.9
CAM-B3LYP/VTZ	-0.023	+3.42	+142.7	+73.5	+90.3
M06/VTZ	-0.027	+3.29	+152.2	+81.7	+102.0
$\omega$ B97X-D/VTZ	-0.026	+3.51	+160.8	+83.1	+123.3
Experiment <sup>a</sup>	$1.36 \pm 0.02$	$111.7 \pm 2.0$	$975.0 \pm 50$	$550.0 \pm 50$	$880.0 \pm 50$

<sup>a</sup>Ref. 68.

### C. Application to 0–0 transition energies of methylpyrimidine

0–0 transition energies can be experimentally measured as crossing points of absorption and fluorescence spectra. To simulate 0–0 transition energies computationally, one has to first optimize the system at both the ground and excited states, whose energies are denoted here by  $E^{\text{GS}}$  and  $E^{\text{ES}}$ , respectively. The difference between these energies at the optimized geometry is referred to as the adiabatic excitation energy (AEE):  $E^{\text{AEE}} = E^{\text{ES}} - E^{\text{GS}}$ . Then, at each minimum, vibrational frequency calculations are performed to obtain the zero-point

vibrational energy (ZPVE; within harmonic approximations):  $E_{\text{ZPVE}}^{\text{GS}}$  and  $E_{\text{ZPVE}}^{\text{ES}}$  for the ground and excited states, respectively. The 0–0 transition energy  $E^{0-0}$  is defined as the difference between the ZPVE-corrected energies:  $E^{0-0} = (E^{\text{ES}} + E_{\text{ZPVE}}^{\text{ES}}) - (E^{\text{GS}} + E_{\text{ZPVE}}^{\text{GS}})$ .

In this study, 2-, 4-, and 5-methylpyrimidine (MP) molecules, for which experimental data<sup>69,70</sup> are available, were selected, and CASSCF, NEVPT2, XMS-CASPT2, and some linear-response time-dependent DFT (TD-DFT)<sup>71</sup> with B3LYP, CAM-B3LYP, and LC-BLYP calculations were performed. First, the structures were optimized using the VTZ basis set, and single-point energy calculations were performed using the AVTZ basis set. For CASSCF, NEVPT2, and XMS-CASPT2, the active space consisted of 10 electrons in 8 orbitals (10e, 8o). For technical reasons, vibrational frequency calculations could not be performed with XMS-CASPT2 for this system, so ZPVE corrections of NEVPT2 were added to the XMS-CASPT2 results.

Overall, the best agreement with experiment was achieved by the XMS-CASPT2 using the IPEA shift technique among the wavefunction methods presented in Table III. On the other hand, if the IPEA shift was not employed, then XMS-CASPT2 underestimated the 0–0 transition energy by 0.25–0.38 eV, which was rather systematic. The impact of the IPEA shift is as large as 0.30–0.45 eV, and the resulting agreement is better than NEVPT2 and comparable to CAM-B3LYP. NEVPT2 overshoot by 0.07–0.13 eV, which is still within the range of expected accuracy and close to the value from SCS-CC2.<sup>72</sup> The agreement of the semiempirical methods OM2/MRCI and OM3/MRCI was rather good, but they wrongly predicted that 2-MP gives higher 0–0 transition energies than 5-MP. The best agreement was achieved with LC-BLYP, with the largest deviation being only 0.04 eV.

Here, the effect of the real shift parameter for XMS-CASPT2 is briefly investigated. First, it was not possible to optimize the geometry, because of the intruder-state problem that occurred in the absence of a shift parameter. Even with shift parameters of 0.1 or 0.2, the CASPT2 iteration did not fully converge to the default threshold, although the convergence criteria for geometry optimization were satisfied. A larger shift value results in a larger 0–0 transition energy. Nevertheless, the XMS-CASPT2 result with a shift parameter of 0.5 still underestimated the experimental value by more than 0.135 eV, which is slightly larger than the deviation with NEVPT2. We can also see that on increasing the shift value by 0.1, the 0–0 transition energy rose by 0.02–0.03 eV (except for shift = 0.1). For the XMS-CASPT2 calculation in the next subsection, a shift parameter of 0.2 was employed, since this value

TABLE III. Deviations of calculated values from experimental<sup>69,70</sup> 0–0 transition energies (eV) of 2-, 4-, and 5-methylpyrimidine (MP).

	2-MP	4-MP	5-MP
CASSCF	+0.238	+0.233	+0.265
NEVPT2	+0.129	+0.071	+0.131
XMS-CASPT2 <sup>a</sup>	−0.245	−0.256	−0.382
XMS-CASPT2 <sup>a,b</sup>	+0.057	+0.079	−0.067
B3LYP	−0.182	−0.209	−0.198
CAM-B3LYP	+0.084	+0.071	+0.082
LC-BLYP	+0.012	−0.038	+0.004
ADC(2) <sup>72</sup>	−0.205		−0.168
CC2 <sup>72</sup>	−0.131		−0.122
SCS-MP2 <sup>72</sup>	+0.097		+0.105
SOS-MP2 <sup>72</sup>	+0.258		+0.258
OM2/MRCI <sup>74</sup>	+0.234		+0.141
OM3/MRCI <sup>74</sup>	+0.084		−0.029
Experiment	3.776 <sup>c</sup>	3.920 <sup>d</sup>	3.819 <sup>c</sup>

<sup>a</sup>Using ZPVE corrections obtained with NEVPT2.

<sup>b</sup>Single-point calculation using the IPEA shift (0.25).

<sup>c</sup>Ref. 69. <sup>d</sup>Ref. 70.

was employed in one of the earliest CASPT2 calculations using the shift technique.<sup>73</sup>

#### D. Application to double excitation in *trans*-polyacetylene

One of the drawbacks of SR methods (e.g., TD-DFT) is the lack of a description of double excitations. The double excitation in *trans*-polyacetylene corresponds to two-electron excitation from HOMO to LUMO (the 1 <sup>1</sup>A<sub>g</sub> to 2 <sup>1</sup>A<sub>g</sub> transition, although symmetry constraints were not applied in the actual computation). One interesting point in a series of *trans*-polyacetylene is that experimental studies demonstrated the doubly excited 2 <sup>1</sup>A<sub>g</sub> state is energetically lower than the 1 <sup>1</sup>B<sub>u</sub> state, which is a HOMO → LUMO single-electron

excitation.<sup>75</sup> One reason of the relatively low transition energies of the  $2\ ^1A_g$  state is attributed to the mixing of the HOMO-1  $\rightarrow$  LUMO and HOMO  $\rightarrow$  LUMO+1 single-electron excitations. Actually, although the leading configuration of the  $2\ ^1A_g$  state is the HOMO  $\rightarrow$  LUMO double-electron excitation, this state additionally comprises these single-electron excitations in a similar weight. A detailed discussion in this aspect can be found in Refs. 76 and 77. Some previous studies<sup>78,79</sup> also computed adiabatic excitation energies (without ZPVE corrections), but the experimental results<sup>80-84</sup> presented in Table IV are 0-0 transition energies, so one must compute second-order geometrical derivatives and include the ZPVE corrections for a fair comparison. For the CASSCF, NEVPT2, and XMS-CASPT2 calculations, the lowest two states ( $1\ ^1A_g$  and  $2\ ^1A_g$ ) were averaged in all calculations, and the active space consisted of  $2n$  electrons in  $2n$  orbitals.

Again, ZPVE corrections computed with NEVPT2 were applied to  $E^{0-0}$  with XMS-CASPT2, in addition to DMRG<sup>78</sup> and EOM-pCCSD-LCCSD,<sup>79</sup> for which only AEEs were presented. In principle, adding the ZPVE correction reduces the transition energy ( $E^{AEE} > E^{0-0}$ ), although the reduction decreases as  $n$  increases. The second-order derivative calculation for  $n = 6$  was too expensive, so extrapolated values of  $-0.135$  and  $-0.107$  eV were employed for CASSCF and NEVPT2, respectively. These values were obtained by fitting the ZPVEs using the quadratic function  $y = an^2 + bn + c$ .

Overall, the difference between NEVPT2 and experiment was almost constant and less than 0.083 eV, while CASSCF underestimated for small systems ( $-0.299$  eV for  $n = 2$ ) and overestimated for large systems ( $+0.118$  eV for  $n = 6$ ). XMS-CASPT2 underestimated the 0-0 transition energies somewhat, by  $-0.157$  to  $-0.229$  eV; the calculation in the previous section with MP derivatives implied such an underestimation. The IPEA shift for the XMS-CASPT2 method (only single-point calculation) raised the transition energy by more than 0.3 eV. As a result, agreement for  $n = 2$  and 3 became better, but that for  $n = 4$  and 5 became slightly worse. As discussed in Ref. 85, the IPEA shift sometimes increases the error. The difference between CASSCF and DMRG can be attributed to state averaging: CASSCF in this study averaged two states, while the DMRG calculation<sup>78</sup> was state-specific. The latter calculation is physically reasonable, but it sometimes exhibits an SCF convergence problem.

To access the differences in optimized geometries with different methods, single-point energy calculations for  $n = 5$  were performed using the CASSCF, NEVPT2, and XMS-

TABLE IV. Calculated adiabatic excitation ( $E^{\text{AEE}}$ ) and 0-0 transition ( $E^{0-0}$ ) energies and experimental<sup>80-84</sup>  $E^{0-0}$  of  $\text{C}_{2n}\text{H}_{2n+2}$  ( $n = 2, 3, 4, 5$ , and 6).

	$E^{\text{AEE}}$ (eV)					$E^{0-0}$ (eV)				
	$n = 2$	3	4	5	6	$n = 2$	3	4	5	6 <sup>h</sup>
CASSCF/VDZ	5.395 <sup>d</sup>	4.368 <sup>f</sup>	3.699	3.231	2.888	5.101 <sup>d</sup>	4.144 <sup>f</sup>	3.549	3.110	2.807
NEVPT2/VDZ	5.726 <sup>e</sup>	4.533	3.759	3.218	2.821	5.471 <sup>e</sup>	4.345	3.624	3.109	2.746
XMS-CASPT2/VDZ	5.426	4.280	3.527	3.002		5.171 <sup>g</sup>	4.092 <sup>g</sup>	3.392 <sup>g</sup>	2.893 <sup>g</sup>	
XMS-CASPT2/VDZ <sup>a</sup>	5.726	4.594	3.864	3.354		5.471 <sup>g</sup>	4.406 <sup>g</sup>	3.728 <sup>g</sup>	3.246 <sup>g</sup>	
DMRG <sup>b</sup>				3.36	2.99				3.25 <sup>g</sup>	2.92 <sup>g</sup>
EOM-pCCD-LCCSD <sup>c</sup>				2.97	2.75				2.86 <sup>g</sup>	2.68 <sup>g</sup>
Experiment						5.4 <sup>i</sup>	4.264 <sup>j</sup>	3.541 <sup>k</sup>	3.05 <sup>l</sup>	2.689 <sup>m</sup>

<sup>a</sup>Single-point calculation using the IPEA shift (0.25).

<sup>b</sup>Ref. 78. <sup>c</sup>Ref. 79.

<sup>d</sup>Excited-state geometry is identified to be a fourth-order saddle point.

<sup>e</sup>Excited-state geometry is identified to be a second-order saddle point.

<sup>f</sup>Excited-state geometry is identified to be a first-order saddle point (transition state).

<sup>g</sup>Using ZPVE corrections obtained with NEVPT2.

<sup>h</sup>Using extrapolated ZPVE corrections.

<sup>i</sup>Ref. 80. <sup>j</sup>Ref. 81. <sup>k</sup>Ref. 82. <sup>l</sup>Ref. 83. <sup>m</sup>Ref. 84.

CASPT2 methods. The calculated vertical ( $E^{\text{vert}}$ ) and adiabatic ( $E^{\text{AEE}}$ ) excitation energies are presented in Table V.  $E^{\text{vert}}$  is quite sensitive to the geometry: for instance, the values of  $E^{\text{vert}}$  computed with CASSCF using the CASSCF, NEVPT2, and XMS-CASPT2 geometries are 4.381, 4.206, and 4.133 eV, respectively, so the dependence on geometry can be as large as 0.25 eV. The dependence is slightly decreased with the MRPT methods: approximately 0.21 eV with NEVPT2 and XMS-CASPT2. On the other hand,  $E^{\text{AEE}}$  is not much affected by differences in geometry: the dependence on geometry is less than 0.01 eV. It seems that the potential energy surface of the  $2^1\text{A}_g$  state around the Franck-Condon region (ground-state geometry) is rather steep, while that around the excited-state minimum is rather flat.

To compare geometrical parameters, C-C bond lengths for  $n = 5$  are compiled in Tables 1



TABLE V. Calculated vertical ( $E^{\text{vert}}$ ) and adiabatic excitation ( $E^{\text{AEE}}$ ) energies of  $\text{C}_{10}\text{H}_{12}$  at the  $X/\text{VDZ}/Y/\text{VDZ}$  level of theory, where  $X$  and  $Y$  are methods employed in the single-point energy calculation and geometry optimization, respectively.

	$E^{\text{vert}}$ (eV)			$E^{\text{AEE}}$ (eV)		
	$Y = \text{CASSCF}$	$\text{NEVPT2}$	$\text{XMS-CASPT2}$	$Y = \text{CASSCF}$	$\text{NEVPT2}$	$\text{XMS-CASPT2}$
$X = \text{CASSCF}$	4.381	4.206	4.133	3.231	3.227	3.224
$X = \text{NEVPT2}$	4.260	4.108	4.044	3.213	3.218	3.217
$X = \text{XMS-CASPT2}$	3.998	3.853	3.792	2.995	3.001	3.002
$X = \text{XMS-CASPT2}^a$	4.359	4.208	4.144	3.353	3.357	3.354

<sup>a</sup>Using the IPEA shift (0.25).

and 3 in the Supplementary Material. The ground-state geometries optimized with NEVPT2 and XMS-CASPT2 are almost the same: the maximum difference between the bond lengths is only 0.005 Å. The small difference reflects the relatively small impact on  $E^{\text{vert}}$  ( $\approx 0.06$  eV) in Table V. On the other hand, the bond lengths of the five shorter (single) bonds ( $\text{C}_1\text{--}\text{C}_2$ ,  $\text{C}_3\text{--}\text{C}_4$ , ...,  $\text{C}_9\text{--}\text{C}_{10}$ ) of the CASSCF geometry at the ground state are shorter by 0.008 and 0.011 Å than those of the NEVPT2 and XMS-CASPT2 geometries, respectively. At the excited state, the difference between NEVPT2 and XMS-CASPT2 is small, but  $\text{C}_2\text{--}\text{C}_3$ ,  $\text{C}_5\text{--}\text{C}_6$ , and  $\text{C}_8\text{--}\text{C}_9$  of the CASSCF structure are shorter by approximately 0.007 and 0.010 Å. Another aspect is the degree of the bond-length alternation. If we calculate the difference between the neighboring C–C bond lengths (Tables 2 and 4 in the Supplementary Material), we notice that the difference at the ground state is larger for CASSCF than for NEVPT2 and XMS-CASPT2. On the other hand, the difference at the excited state is somewhat more complicated: larger at the edge but smaller at the middle position.

From these results, the perturbative correction, the difference between CASSCF and NEVPT2, was as large as 0.37 eV (for  $n = 2$ ), and this indicates the importance of post-CASSCF treatment for the 0–0 transition energies. In addition, the ZPVE correction for  $n = 2$  reduced the transition energy by  $-0.255$  eV, and that for  $n = 5$  was still as large as  $-0.109$  eV, so one cannot neglect the ZPVE correction when comparing with experimental results. The computation of second-order geometrical derivatives is a rather expensive

procedure (up to  $36N_{\text{atom}}^2$  energy or  $6N_{\text{atom}}$  gradient calculations are required), so analytic derivatives are essential. This computation is much more expensive than the geometry optimization itself, so it is not straightforward to compare with the size employed in this and earlier studies without a vibrational frequency analysis. With the present implementation, if only geometry optimizations are required, then 500–600 AOs and/or (12e,12o) active space may be employed at a reasonable computational cost using a single computer node.

## V. CONCLUSION

In this paper, the development of first-order analytic derivatives of the state-specific PC-NEVPT2 method has been outlined. These derivatives are computed by exploiting the Z-vector method, so the additional computational cost is not explicitly dependent on the number of atoms. The derived expressions do not involve any vanishing denominators resulting from orbital degeneracies.

The methodology has been implemented in a local version of GAMESS-US and applied to typical systems: the geometrical parameters of  $\text{O}_3$  and  $\text{O}_3^-$  and the 0–0 transition energies of methylpyrimidines and a series of *trans*-polyacetylenes. The results show the importance of perturbative corrections and the necessity for development of analytic derivatives. Comparison with experiment suggests that the accuracies of NEVPT2 and CASPT2 are roughly comparable: CASPT2 with the IPEA shift outperforms other wavefunction methods for methylpyrimidine derivatives, while NEVPT2 does so for *trans*-polyacetylene.

CASPT2 calculations for the state-averaged CASSCF reference required a shift technique<sup>73</sup> to ensure convergence of the amplitude equation. The 0–0 transition energies of methylpyrimidines computed with various values of the shift parameter for XMS-CASPT2 indicated that the transition energy increased by 0.02–0.03 eV as the shift increased by 0.1. The 0–0 transition energies computed using the XMS-CASPT2 method with and without the IPEA shift indicated that the impact of the IPEA shift is as large as 0.3 eV. The IPEA shift improved the agreement with experiment for methylpyrimidine derivatives, but slightly degraded it for the larger *trans*-polyacetylene. These two shift techniques may impart a slight arbitrariness to multistate CASPT2 calculations, but NEVPT2 is an intruder-state-free method, and no such empirical parameters are required as long as an appropriate active space is used. In addition, since NEVPT2 is size-consistent, it can be applied to chemical reactions.

Extension to quasidegenerate NEVPT2<sup>56</sup> and the computation of derivative coupling<sup>21</sup> are essential tasks for the future, in particular in the study of photochemical processes, for example, in the identification of conical intersections. Algorithmic improvements are also highly desirable; for instance, CASSCF in GAMESS-US does not utilize any matrix decomposition techniques, which severely restricts the number of orbitals that can be employed in CASSCF and NEVPT2 calculations.

## Appendix A: Derivative of doubly occupied orbital in $E^{(-1)}$

Here, we give a brief description of the transformation of the derivatives of  $C_{\mu i}$  and  $\varepsilon_i$  in Eq. (28):

$$2 \sum_i \sum_a \sum_{rs} \sum_{\phi} \sum_{\mu} \tilde{T}_{rs,i}^{\phi} \mathcal{S}_{\phi a} \frac{dC_{\mu i}}{d\alpha}(r\mu|sa) - \sum_i \sum_{rs} \sum_{\mu} \tilde{T}_{rs,i}^{\phi} T_{rs,i}^{\phi} \frac{d\varepsilon_i}{d\alpha}. \quad (\text{A1})$$

This is relevant to the transformation of the first two terms in Eq. (28) into the first five terms in Eq. (31). A similar procedure can be applied to virtual orbital derivatives.

The first term is

$$\begin{aligned} & 2 \sum_i \sum_a \sum_{rs} \sum_{\phi} \sum_{\mu} \tilde{T}_{rs,i}^{\phi} \mathcal{S}_{\phi a} \frac{dC_{\mu i}}{d\alpha}(r\mu|sa) \\ &= 2 \sum_p^{\text{act+vir}} \sum_q^{\text{docc}} \sum_a \sum_{rs} (rp|sa) \tilde{T}_{rs,q}^a U_{pq}^{\alpha} \\ &+ 2 \sum_{p>q}^{\text{docc}} \sum_{\phi} \sum_{rs} (\varepsilon_p - \varepsilon_q) \tilde{T}_{rs,p}^{\phi} T_{rs,q}^{\phi} U_{pq}^{\alpha} \\ &- \sum_{pq}^{\text{docc}} \sum_a \sum_{rs} (rp|sa) \tilde{T}_{rs,q}^a S_{pq}^{\alpha}, \end{aligned} \quad (\text{A2})$$

where

$$\sum_a \left( (rp|sa) \tilde{T}_{rs,q}^a - (rq|sa) \tilde{T}_{rs,p}^a \right) = (\varepsilon_p - \varepsilon_q) \sum_{\phi} \tilde{T}_{rs,p}^{\phi} T_{rs,q}^{\phi} \quad (\text{A3})$$

has been used. A similar transformation is detailed for MP2 in Ref. 40. Note that the contracted PT2 density [Eq. (30)] is symmetric. When  $p$  and  $q$  are in the diagonal doubly occupied or virtual block,  $U_{pq}^{\alpha}$  can be written explicitly as<sup>39</sup>

$$U_{pq}^{\alpha} = \frac{1}{\varepsilon_q - \varepsilon_p} \left\{ F_{pq}^{\alpha} - \varepsilon_q S_{pq}^{\alpha} + (1 + \hat{\tau}_{pq}) \sum_m^{\text{d.s.}} F_{pm} U_{mq}^{\alpha} \right\}, \quad (\text{A4})$$

where  $\hat{\tau}_{pq}$  permutes the indices  $p$  and  $q$ , and “d.s.” (= different shell) is act + vir for  $(p, q) \in (\text{docc}, \text{docc})$  and docc + act for  $(p, q) \in (\text{vir}, \text{vir})$ , and

$$F_{pq}^\alpha = \sum_{\mu\nu} C_{\mu p} C_{\nu q} \frac{dF_{\mu\nu}}{d\alpha}, \quad (\text{A5})$$

where  $F_{\mu\nu}$  is defined in Eq. (A11). Using this relation and the definition of the contracted density for the doubly occupied orbital [Eq. (30)], the first term in Eq. (A1) is

$$\begin{aligned} & 2 \sum_i \sum_a \sum_{rs} \sum_\phi \sum_\mu \tilde{T}_{rs,i}^\phi \mathcal{S}_{\phi a} \frac{dC_{\mu i}}{d\alpha} (r\mu|sa) \\ &= 2 \sum_p^{\text{act+vir}} \sum_q^{\text{docc}} \sum_a \sum_{rs} (rp|sa) \tilde{T}_{rs,q}^a U_{pq}^\alpha \\ &+ 2 \sum_{p>q}^{\text{docc}} d_{pq}^{\text{PT2,inact}} \left\{ F_{pq}^\alpha - \varepsilon_q S_{pq}^\alpha + (1 + \hat{\tau}_{pq}) \sum_m^{\text{act+vir}} F_{pm} U_{mq}^\alpha \right\} \\ &- \sum_{pq}^{\text{docc}} \sum_a \sum_{rs} (rp|sa) \tilde{T}_{rs,q}^a S_{pq}^\alpha \end{aligned} \quad (\text{A6})$$

$$\begin{aligned} &= 2 \sum_p^{\text{act+vir}} \sum_q^{\text{docc}} \sum_a \sum_{rs} (rp|sa) \tilde{T}_{rs,q}^a U_{pq}^\alpha \\ &+ \sum_{p \neq q}^{\text{docc}} d_{pq}^{\text{PT2,inact}} \left\{ F_{pq}^\alpha - \frac{\varepsilon_p + \varepsilon_q}{2} S_{pq}^\alpha + (1 + \hat{\tau}_{pq}) \sum_m^{\text{act+vir}} F_{pm} U_{mq}^\alpha \right\} \\ &- \sum_{pq}^{\text{docc}} \sum_a \sum_{rs} (rp|sa) \tilde{T}_{rs,q}^a S_{pq}^\alpha. \end{aligned} \quad (\text{A7})$$

Adding the eigenvalue derivative term [the second term in Eq. (A1)], Eq. (A1) becomes

$$\begin{aligned} & 2 \sum_i \sum_a \sum_{rs} \sum_\phi \sum_\mu \tilde{T}_{rs,i}^\phi \mathcal{S}_{\phi a} \frac{dC_{\mu i}}{d\alpha} (r\mu|sa) - \sum_i \sum_{rs} \sum_\mu \tilde{T}_{rs,i}^\mu T_{rs,i}^\mu \frac{d\varepsilon_i}{d\alpha} \\ &= 2 \sum_p^{\text{act+vir}} \sum_q^{\text{docc}} \left( \sum_a \sum_{rs} (rp|sa) \tilde{T}_{rs,q}^a + \sum_i F_{pi} d_{iq}^{\text{PT2,inact}} \right) U_{pq}^\alpha \\ &- \sum_{pq}^{\text{docc}} \left( \sum_a \sum_{rs} (rp|sa) \tilde{T}_{rs,q}^a + \frac{\varepsilon_p + \varepsilon_q}{2} d_{pq}^{\text{PT2,inact}} \right) S_{pq}^\alpha \\ &+ \sum_{pq}^{\text{docc}} d_{pq}^{\text{PT2,inact}} F_{pq}^\alpha. \end{aligned} \quad (\text{A8})$$

If we define a Fock-weighted density matrix:

$$\tilde{d}_{pq} = \begin{cases} \frac{\varepsilon_p + \varepsilon_q}{2} d_{pq}^{\text{PT2,inact}} & (p, q) \in (\text{docc}, \text{docc}) \text{ or } (\text{vir}, \text{vir}) , \\ \sum_{i \in \text{docc}} F_{pi} d_{iq}^{\text{PT2,inact}} & p \notin \text{docc} \text{ and } q \in \text{docc} , \\ \sum_{r \in \text{vir}} F_{pr} d_{rq}^{\text{PT2,inact}} & p \notin \text{vir} \text{ and } q \in \text{vir} \end{cases} \quad (\text{A9})$$

and assume that  $U_{pq}^\alpha$  for  $(p, q) \in (\text{docc}, \text{docc})$  can be written as  $U_{pq}^\alpha = -S_{pq}^\alpha/2$ , then the above expression can be written as

$$\begin{aligned} & 2 \sum_i \sum_a \sum_{rs} \sum_\phi \sum_\mu \tilde{T}_{rs,i}^\phi \mathcal{S}_{\phi a} \frac{dC_{\mu i}}{d\alpha} (r\mu|sa) - \sum_i \sum_{rs} \sum_\mu \tilde{T}_{rs,i}^\mu T_{rs,i}^\mu \frac{d\varepsilon_i}{d\alpha} \\ & = 2 \sum_p^{\text{act+vir}} \sum_q^{\text{docc}} \left( \sum_a \sum_{rs} (rp|sa) \tilde{T}_{rs,q}^a + \tilde{d}_{pq} \right) U_{pq}^\alpha + \sum_{pq}^{\text{docc}} d_{pq}^{\text{PT2,inact}} F_{pq}^\alpha . \end{aligned} \quad (\text{A10})$$

A similar transformation should be applied to virtual orbitals too.

The second term in Eq. (A10) has to be further transformed. The Fock matrix of CASSCF in this study is defined by

$$F_{\mu\nu} = h_{\mu\nu} + G_{\mu\nu} [\mathbf{d}^{\text{SA}}] , \quad (\text{A11})$$

where the state-averaged density matrix is

$$d_{\mu\nu}^{\text{SA}} = \sum_{pq}^{\text{docc+act}} C_{\mu p} C_{\nu q} \gamma_{pq}^{\text{SA}} . \quad (\text{A12})$$

In particular,  $\gamma_{pq}^{\text{SA}} = 2\delta_{pq}$  for  $(p, q) \in (\text{docc}, \text{docc})$ ,  $\gamma_{pq}^{\text{SA}} = 0$  for  $(p, q) \in (\text{docc}, \text{act})$  or  $(\text{act}, \text{docc})$ , and

$$\gamma_{a'a}^{\text{SA}} = \sum_S^\Omega \omega_S \gamma_{a'a}^S \quad (\text{A13})$$

for  $(p, q) \in (\text{act}, \text{act})$ . The two-electron integral contribution to the Fock matrix has been defined in Eq. (36). With these expressions, the second term of Eq. (A10) can be written as

$$\begin{aligned} \sum_{pq}^{\text{docc}} d_{pq}^{\text{PT2,inact}} F_{pq}^\alpha & = \sum_{\mu\nu} d_{\mu\nu}^{\text{PT2,inact}} \frac{dh_{\mu\nu}}{d\alpha} \\ & + \sum_{\mu\nu\rho\sigma} \left( d_{\mu\nu}^{\text{PT2,inact}} d_{\rho\sigma}^{\text{SA}} - \frac{1}{4} d_{\mu\rho}^{\text{PT2,inact}} d_{\nu\sigma}^{\text{SA}} - \frac{1}{4} d_{\mu\sigma}^{\text{PT2,inact}} d_{\nu\rho}^{\text{SA}} \right) \frac{d(\mu\nu|\rho\sigma)}{d\alpha} \\ & + 2 \sum_p^{\text{all}} \sum_q^{\text{docc+act}} \sum_m^{\text{docc+act}} G_{pm} [\mathbf{d}^{\text{PT2,inact}}] \gamma_{mq}^{\text{SA}} U_{pq}^\alpha + \sum_{a'a}^{\text{act}} G_{a'a} [\mathbf{d}^{\text{PT2,inact}}] \frac{d\gamma_{a'a}^{\text{SA}}}{d\alpha} . \end{aligned} \quad (\text{A14})$$

where  $d_{\mu\nu}^{\text{PT2,inact}}$  is obtained by transforming  $d_{pq}^{\text{PT2,inact}}$  into the AO basis:

$$d_{\mu\nu}^{\text{PT2,inact}} = \sum_{pq} C_{\mu p} C_{\nu q} d_{pq}^{\text{PT2,inact}} , \quad (\text{A15})$$

while  $G_{pm}[\mathbf{d}^{\text{PT2,inact}}]$  is obtained by transforming  $G_{\mu\nu}[\mathbf{d}^{\text{PT2,inact}}]$  into the MO basis. The last two terms in Eq. (A14) arise from the derivative of the state-averaged density matrix. Using Eqs. (A10) and (A14) and collecting the contributions for virtual orbitals, we obtain the full expression of Eq. (31).

## Appendix B: Derivation of RDM derivatives

Here, Eq. (77) is derived from Eqs. (3), (74), and (76). Taking the derivative of Eq. (3), we straightforwardly obtain

$$\begin{aligned} \frac{dR_{ab}^{S,(1)}}{d\alpha} &= \sum_{IJ} \left( \frac{dc_I^S}{d\alpha} c_J^S + c_I^S \frac{dc_J^S}{d\alpha} \right) \gamma_{ab}^{IJ} \\ &= \sum_{IJ} \sum_A^{\Omega+\Lambda} (c_I^A c_J^S + c_I^S c_J^A) \gamma_{ab}^{IJ} W_{SA}^\alpha \\ &= \sum_{IJ} \sum_T^\Omega c_I^S c_J^T (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) W_{ST}^\alpha + \sum_{IJ} \sum_P^\Lambda c_I^S c_J^P (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) W_{SP}^\alpha , \end{aligned} \quad (\text{B1})$$

where  $\Omega$  and  $\Lambda$  represent internal and external states (see Sec. II), respectively.

The first term in Eq. (B1) can be split into three cases:  $S < T$ ,  $S > T$ , and  $S = T$ . For  $S < T$ ,  $W_{ST}^\alpha$  is a nonredundant rotation, so it can be simply transformed as

$$\sum_{IJ} \sum_{S < T}^\Omega (c_I^T c_J^S + c_I^S c_J^T) \gamma_{ab}^{IJ} W_{ST}^\alpha = \sum_{S < T}^\Omega \left( R_{ab}^{ST,(1)} + R_{ab}^{TS,(1)} \right) W_{ST}^\alpha . \quad (\text{B2})$$

For  $S > T$ ,  $W_{ST}^\alpha$  is a redundant rotation, so  $W_{ST}^\alpha$  is replaced with  $-W_{TS}^\alpha$ :

$$\sum_{IJ} \sum_{S > T}^\Omega (c_I^T c_J^S + c_I^S c_J^T) \gamma_{ab}^{IJ} W_{ST}^\alpha = - \sum_{S > T}^\Omega \left( R_{ab}^{ST,(1)} + R_{ab}^{TS,(1)} \right) W_{TS}^\alpha . \quad (\text{B3})$$

For  $S = T$ , it is clear that Eq. (76) is satisfied only when  $W_{ST}^\alpha = 0$ . These three cases may be written compactly as

$$\sum_{IJ} \sum_T^\Omega c_I^S c_J^T (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) W_{ST}^\alpha = \sum_T^\Omega s_{ST} \left( R_{ab}^{ST,(1)} + R_{ab}^{TS,(1)} \right) W_{ST}^\alpha , \quad (\text{B4})$$

where  $s_{ST}$  is defined in the main text.

The second term in Eq. (B1) is not a practically useful expression, because one has to provide CI coefficients of external states to compute it and to solve the Z-vector equations. Therefore, using the orthonormality of states [Eq. (75)], one can insert  $\delta_{AB}$  and transform as follows:

$$\begin{aligned}
& \sum_{IJ} \sum_P^\Lambda c_I^S c_J^P (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) W_{SP}^\alpha \\
&= \sum_{IJ} \sum_{PQ}^\Lambda c_I^S c_J^P \delta_{PQ} (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) W_{SQ}^\alpha \\
&= \sum_{IJK} \sum_{PQ}^\Lambda c_I^S c_J^P c_K^P c_K^Q (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) W_{SQ}^\alpha \\
&= \sum_{IJK} \sum_P^\Lambda c_I^S c_J^P c_K^P (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) V_{SK}^\alpha , \tag{B5}
\end{aligned}$$

where

$$V_{SK}^\alpha = \sum_Q^\Lambda W_{SQ}^\alpha c_K^Q . \tag{B6}$$

The CI coefficients of external states can be eliminated using the resolution of the identity

$$\sum_S^\Omega c_I^S c_J^S + \sum_P^\Lambda c_I^P c_J^P = \delta_{IJ} , \tag{B7}$$

and one may further transform the above expression:

$$\begin{aligned}
& \sum_{IJ} \sum_P^\Lambda c_I^S c_J^P (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) W_{SP}^\alpha \\
&= \sum_{IJK} c_I^S \left( \delta_{JK} - \sum_T^\Omega c_J^T c_K^T \right) (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) V_{SK}^\alpha \\
&= \sum_{IJ} (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) c_J^S V_{SI}^\alpha - \sum_I \sum_T^\Omega \left( R_{ab}^{ST,(1)} + R_{ab}^{TS,(1)} \right) c_I^T V_{SI}^\alpha . \tag{B8}
\end{aligned}$$

However, the second term does not contribute to the final gradient, because of the orthonormalization conditions on  $Z_{SI}^C$  and  $c_I^S$ .

Thus, from Eqs. (B4) and (B8), the derivative of 1-RDM can be written as

$$\frac{dR_{ab}^{S,(1)}}{d\alpha} = \sum_{IJ} (\gamma_{ab}^{IJ} + \gamma_{ab}^{JI}) c_J^S V_{SI}^\alpha + \sum_T^\Omega s_{ST} \left( R_{ab}^{ST,(1)} + R_{ab}^{TS,(1)} \right) W_{ST}^\alpha , \tag{B9}$$

which is given in Eq. (77) in the main text.

## Appendix C: Calculation of the relaxed density matrix

The one-electron relaxed density matrix of the internal state  $S$  can be written as

$$d_{\mu\nu}^{S,\text{relax}} = d_{\mu\nu}^{S,\text{CAS}} + d_{\mu\nu}^{S,\text{PT2}} + d_{\mu\nu}^{S,\text{resp}} . \quad (\text{C1})$$

The “state-specific” unrelaxed density matrix from the reference wavefunction is

$$d_{\mu\nu}^{S,\text{CAS}} = 2 \sum_i C_{\mu i} C_{\nu i} + \sum_{a'a} C_{\mu a'} C_{\nu a} \gamma_{a'a}^{SS} , \quad (\text{C2})$$

where  $\gamma_{a'a}^{ST}$  is the 1-RDM ( $S = T$ ) or the transition 1-RDM ( $S \neq T$ ). The perturbation density matrix  $d_{\mu\nu}^{S,\text{PT2}}$  has been defined in Eq. (69). One also needs the similar contributions for the other seven excitation classes.

The response density matrix is computed using the solutions of the Z-vector equation,  $Z_{pq}^o$ ,  $Z_{SI}^c$ , and  $Z_{ST}^s$ , which come from orbital, configuration, and state rotations, respectively:

$$\begin{aligned} d_{\mu\nu}^{S,\text{relax}} = & 4 \sum_p \sum_i C_{\mu p} C_{\nu i} Z_{pi}^{S,o} + 2 \sum_p \sum_{a'} C_{\mu p} C_{\nu a'} \sum_a \gamma_{a'a}^{\text{SA}} Z_{pa}^{S,o} \\ & + \sum_{a'a} C_{\mu a'} C_{\nu a} \sum_{IJ} \sum_T^\Omega \omega_T \tilde{\gamma}_{a'a}^{S,T} \\ & + \sum_{a'a} C_{\mu a'} C_{\nu a} \sum_{TU}^\Omega (\omega_T - \omega_U) Z_{TU}^{S,s} (\gamma_{a'a}^{TU} + \gamma_{a'a}^{UT}) , \end{aligned} \quad (\text{C3})$$

where

$$\tilde{\gamma}_{a'a}^{S,T} = \sum_{IJ} \left( \tilde{z}_{IT}^{S,c} c_J^T + c_I^T \tilde{z}_{JT}^{S,c} \right) \gamma_{a'a}^{IJ} \quad (\text{C4})$$

and

$$\tilde{z}_{IT}^{S,c} = Z_{IT}^{S,c} - \sum_U c_I^U \sum_J Z_{JT}^{S,c} c_J^U . \quad (\text{C5})$$

If the weighting factors are the same ( $\omega_S = \omega_T$ ), then the last term in Eq. (C3) for the pair  $S$  and  $T$  requires the replacement

$$(\omega_S - \omega_T) Z_{ST}^{S,s} \rightarrow \frac{\mathcal{L}_{ST}^{S,s}}{2(E_S^{\text{CAS}} - E_T^{\text{CAS}})} . \quad (\text{C6})$$

## Appendix D: Frozen core approximation

The use of the frozen core approximation requires minor modifications to the procedure described in the main text. In computing energies, one simply has to exclude core orbitals



from the summation over  $i$ . In computing gradients, it should be noted that the relation in Eq. (A3) no longer holds if either  $p$  or  $q$  belongs to a core orbital. This is because  $i$  runs over nonfrozen doubly occupied orbitals, while  $p$  runs over all orbitals, including core orbitals.

In this case,  $d_{pq}^{\text{PT2,inact}}$  given in Eq. (30) has to be defined differently:

$$d_{pq}^{\text{PT2,inact}} = \begin{cases} 0 & (p, q) \in (\text{cor}, \text{cor}) , \\ \frac{1}{\varepsilon_p - \varepsilon_q} \sum_a \sum_{rs} (rp|sa) \tilde{T}_{rs,q}^a & (p, q) \in (\text{cor}, \text{val}) \text{ or } (\text{val}, \text{cor}) , \\ - \sum_{rs} \sum_{\phi} \tilde{T}_{rs,p}^{\phi} T_{rs,q}^{\phi} & (p, q) \in (\text{val}, \text{val}) , \\ \sum_i \sum_r \sum_{\phi} \left( \tilde{T}_{pr,i}^{\phi} T_{qr,i}^{\phi} + \tilde{T}_{rp,i}^{\phi} T_{rq,i}^{\phi} \right) & (p, q) \in (\text{vir}, \text{vir}) , \\ 0 & \text{otherwise} , \end{cases} \quad (\text{D1})$$

where doubly occupied orbital space is split into “cor” (core) and “val” (doubly occupied valence) orbitals. One has to divide by the difference of orbital energies, but the gap between the core and valence orbital subspaces is usually large, so there should be no singularity (vanishing denominator).

The above expression indicates that the use of the frozen core approximation collapses the rotational invariance of PC-NEVPT2 energy within doubly occupied orbitals. Nevertheless, the non-invariance appears only in rotations between core and valence orbitals, so there is still invariance under rotations within core and valence orbital subspaces. This is different from the case of SC-NEVPT2 (regardless of the frozen core approximation), where there is also non-invariance under rotations within core and valence orbital subspaces.

## SUPPLEMENTARY MATERIAL

See supplementary material for optimized coordinates of the methylpyrimidines and *trans*-polyacetylene.

## ACKNOWLEDGMENTS

This work was partially supported by JSPS KAKENHI Grant Number 17K14436. The author wishes to acknowledge Professors Yuki Kurashige (Kyoto University) and Yuji Mochizuki (Rikkyo University) for their helpful advice.

## REFERENCES

- <sup>1</sup>P.-O. Löwdin, “Quantum theory of many-particle systems. iii. extension of the hartree-fock scheme to include degenerate systems and correlation effects,” *Phys. Rev.* **97**, 1509–1520 (1955).
- <sup>2</sup>T. Helgaker, P. Jørgensen, and J. Olsen, eds., Molecular Electronic-Structure Theory (John Wiley & Sons, Ltd., England, 2000).
- <sup>3</sup>B. O. Roos, R. Lindh, P. Å. Malmqvist, V. Veryazov, and P.-O. Widmark, eds., Multiconfigurational Quantum Chemistry (John Wiley & Sons, Ltd., Hoboken, New Jersey, 2016).
- <sup>4</sup>M. Musiał, A. Perera, and R. J. Bartlett, “Multireference coupled-cluster theory: The easy way,” *J. Chem. Phys.* **134**, 114108 (2011).
- <sup>5</sup>D. I. Lyakh, M. Musiał, V. F. Lotrich, and R. J. Bartlett, “Multireference nature of chemistry: The coupled-cluster view,” *Cosmic Res.* **112**, 182–243 (2012).
- <sup>6</sup>A. Köhn, M. Hanauer, L. A. Mück, T.-C. Jagau, and J. Gauss, “State-specific multireference coupled-cluster theory,” *WIREs Comput. Mol. Sci.* **3**, 176–197 (2013).
- <sup>7</sup>P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, and R. Shepard, “Multiconfiguration self-consistent field and multireference configuration interaction methods and applications,” *Cosmic Res.* **112**, 108–181 (2012).
- <sup>8</sup>B. O. Roos, P. Linse, P. E. Siegbahn, and M. R. Blomberg, “A simple method for the evaluation of the second-order-perturbation energy from external double-excitations with a casscf reference wavefunction,” *Contemp. Phys.* **66**, 197 – 207 (1982).
- <sup>9</sup>K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, “Second-order perturbation theory with a casscf reference function,” *J. Phys. Chem.* **94**, 5483–5488 (1990).
- <sup>10</sup>K. Andersson, P. Malmqvist, and B. O. Roos, “Second-order perturbation theory with a complete active space self-consistent field reference function,” *J. Chem. Phys.* **96**, 1218–1226 (1992).
- <sup>11</sup>A. A. Granovsky, “Extended multi-configuration quasi-degenerate perturbation theory: The new approach to multi-state multi-reference perturbation theory,” *J. Chem. Phys.* **134**, 214113 (2011).

- <sup>12</sup>H. Nakano, “Quasidegenerate perturbation theory with multiconfigurational self-consistent-field reference functions,” *J. Chem. Phys.* **99**, 7983–7992 (1993).
- <sup>13</sup>C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu, “Introduction of  $n$ -electron valence states for multireference perturbation theory,” *J. Chem. Phys.* **114**, 10252–10264 (2001).
- <sup>14</sup>C. Angeli, R. Cimiraglia, and J.-P. Malrieu, “ $n$ -electron valence state perturbation theory: a fast implementation of the strongly contracted variant,” *Chem. Phys. Lett.* **350**, 297 – 305 (2001).
- <sup>15</sup>C. Angeli, R. Cimiraglia, and J.-P. Malrieu, “ $n$ -electron valence state perturbation theory: A spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants,” *J. Chem. Phys.* **117**, 9138–9153 (2002).
- <sup>16</sup>Y. G. Khait, J. Song, and M. R. Hoffmann, “Explication and revision of generalized van vleck perturbation theory for molecular electronic structure,” *J. Chem. Phys.* **117**, 4133–4145 (2002).
- <sup>17</sup>P. Celani and H.-J. Werner, “Analytical energy gradients for internally contracted second-order multireference perturbation theory,” *J. Chem. Phys.* **119**, 5044–5057 (2003).
- <sup>18</sup>W. Győrffy, T. Shiozaki, G. Knizia, and H.-J. Werner, “Analytical energy gradients for second-order multireference perturbation theory using density fitting,” *J. Chem. Phys.* **138**, 104104 (2013).
- <sup>19</sup>M. K. MacLeod and T. Shiozaki, “Communication: Automatic code generation enables nuclear gradient computations for fully internally contracted multireference theory,” *J. Chem. Phys.* **142**, 051103 (2015).
- <sup>20</sup>B. Vlaisavljevich and T. Shiozaki, “Nuclear energy gradients for internally contracted complete active space second-order perturbation theory: Multistate extensions,” *J. Chem. Theory Comput.* **12**, 3781–3787 (2016).
- <sup>21</sup>J. W. Park and T. Shiozaki, “Analytical derivative coupling for multistate caspt2 theory,” *J. Chem. Theory Comput.* **13**, 2561–2570 (2017).
- <sup>22</sup>T. J. Dudley, Y. G. Khait, and M. R. Hoffmann, “Molecular gradients for the second-order generalized van vleck variant of multireference perturbation theory,” *J. Chem. Phys.* **119**, 651–660 (2003).
- <sup>23</sup>D. Theis, Y. G. Khait, S. Pal, and M. R. Hoffmann, “Molecular electric dipole moments using the gvvpt2 variant of multireference perturbation theory,” *Chem. Phys. Lett.* **487**,

- 116 – 121 (2010).
- <sup>24</sup>D. Theis, Y. G. Khait, and M. R. Hoffmann, “Gvvpt2 energy gradient using a lagrangian formulation,” *J. Chem. Phys.* **135**, 044117 (2011).
- <sup>25</sup>Y. G. Khait, D. Theis, and M. R. Hoffmann, “Nonadiabatic coupling terms for the gvvpt2 variant of multireference perturbation theory,” *Contemp. Phys.* **401**, 88 – 94 (2012).
- <sup>26</sup>H. Nakano, K. Hirao, and M. S. Gordon, “Analytic energy gradients for multiconfigurational self-consistent field second-order quasidegenerate perturbation theory (mc-qdpt),” *J. Chem. Phys.* **108**, 5660–5669 (1998).
- <sup>27</sup>J. W. Park, “Analytical gradient theory for strongly contracted (sc-) and partially contracted (pc-) n-electron valence state perturbation theory (nevpt2),” arXiv:1907.10180.
- <sup>28</sup>P. Pulay, “Analytical derivatives, forces, force constants, molecular geometries, and related response properties in electronic structure theory,” *WIREs Comput. Mol. Sci.* **4**, 169–181 (2014).
- <sup>29</sup>N. C. Handy and H. F. Schaefer III, “On the evaluation of analytic energy derivatives for correlated wave functions,” *J. Chem. Phys.* **81**, 5031–5033 (1984).
- <sup>30</sup>T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen, and K. Ruud, “Recent advances in wave function-based methods of molecular-property calculations,” *Cosmic Res.* **112**, 543–631 (2012).
- <sup>31</sup>K. Kristensen, P. Jørgensen, A. J. Thorvaldsen, and T. Helgaker, “Efficient elimination of response parameters in molecular property calculations for variational and nonvariational energies,” *J. Chem. Phys.* **129**, 214103 (2008).
- <sup>32</sup>Y. Guo, K. Sivalingam, E. F. Valeev, and F. Neese, “Sparsemaps—a systematic infrastructure for reduced-scaling electronic structure methods. iii. linear-scaling multireference domain-based pair natural orbital n-electron valence perturbation theory,” *J. Chem. Phys.* **144**, 094111 (2016).
- <sup>33</sup>K. L. Bak, J. Boatz, and J. Simons, “First-order geometrical response equations for state-averaged multiconfigurational self-consistent field (sa-mcscf) wave functions,” *Int. J. Quantum Chem.* **40**, 361–378 (1991).
- <sup>34</sup>A. Szabo and N. S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure (Macmillan Publishing Co., Inc., New York, 1982).
- <sup>35</sup>T. J. Dudley, R. M. Olson, M. W. Schmidt, and M. S. Gordon, “Parallel coupled perturbed casscf equations and analytic casscf second derivatives,” *J. Comput. Chem.* **27**, 352–362

- (2006).
- <sup>36</sup>K. G. Dyall, “The choice of a zeroth-order hamiltonian for second-order perturbation theory with a complete active space self-consistent-field reference function,” *J. Chem. Phys.* **102**, 4909–4918 (1995).
- <sup>37</sup>B. Helmich-Paris and S. Knecht, “Laplace-transformed multi-reference second-order perturbation theories in the atomic and active molecular orbital basis,” *J. Chem. Phys.* **146**, 224101 (2017).
- <sup>38</sup>I. Schapiro, K. Sivalingam, and F. Neese, “Assessment of  $n$ -electron valence state perturbation theory for vertical excitation energies,” *J. Chem. Theory Comput.* **9**, 3567–3580 (2013).
- <sup>39</sup>Y. Yamaguchi, H. F. Schaefer III, Y. Osamura, and J. Goddard, A New Dimension to Quantum Chemistry: Analytical Derivative Methods in Ab Initio Molecular Elect (Oxford University Press, New York, 1994).
- <sup>40</sup>C. M. Aikens, S. P. Webb, R. L. Bell, G. D. Fletcher, M. W. Schmidt, and M. S. Gordon, “A derivation of the frozen-orbital unrestricted open-shell and restricted closed-shell second-order perturbation theory analytic gradient expressions,” *Theor. Chem. Acc.* **110**, 233–253 (2003).
- <sup>41</sup>R. Moccia, “Variable bases in scf mo calculations,” *Chem. Phys. Lett.* **5**, 260 – 264 (1970).
- <sup>42</sup>N. Handy, R. Amos, J. Gaw, J. Rice, and E. Simandiras, “The elimination of singularities in derivative calculations,” *Chem. Phys. Lett.* **120**, 151 – 158 (1985).
- <sup>43</sup>M. W. Schmidt, K. K. Baldridge, 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, “General atomic and molecular electronic structure system,” *J. Comput. Chem.* **14**, 1347–1363 (1993).
- <sup>44</sup>P. J. Knowles and H.-J. Werner, “An efficient method for the evaluation of coupling coefficients in configuration interaction calculations,” *Chem. Phys. Lett.* **145**, 514 – 522 (1988).
- <sup>45</sup>C. Angeli, M. Pastore, and R. Cimiraglia, “New perspectives in multireference perturbation theory: The  $n$ -electron valence state approach,” *Theor. Chem. Acc.* **117**, 743–754 (2007).
- <sup>46</sup>“Bagel, brilliantly advanced general electronic-structure library. <http://www.nubakery.org> under the gnu general public license., (accessed on june 20, 2019),”.

- <sup>47</sup>T. Shiozaki, “Bagel: Brilliantly advanced general electronic-structure library,” *WIREs Comput. Mol. Sci.* **8**, e1331 (2018).
- <sup>48</sup>T. Shiozaki, W. Győrffy, P. Celani, and H.-J. Werner, “Communication: Extended multi-state complete active space second-order perturbation theory: Energy and nuclear gradients,” *J. Chem. Phys.* **135**, 081106 (2011).
- <sup>49</sup>T. H. Dunning, “Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen,” *J. Chem. Phys.* **90**, 1007–1023 (1989).
- <sup>50</sup>R. A. Kendall, T. H. Dunning, and R. J. Harrison, “Electron affinities of the first-row atoms revisited. systematic basis sets and wave functions,” *J. Chem. Phys.* **96**, 6796–6806 (1992).
- <sup>51</sup>A. D. Becke, “A new mixing of hartree-fock and local density-functional theories,” *J. Chem. Phys.* **98**, 1372–1377 (1993).
- <sup>52</sup>A. D. Becke, “Density-functional thermochemistry. iii. the role of exact exchange,” *J. Chem. Phys.* **98**, 5648–5652 (1993).
- <sup>53</sup>T. Yanai, D. P. Tew, and N. C. Handy, “A new hybrid exchange–correlation functional using the coulomb-attenuating method (cam-b3lyp),” *Chem. Phys. Lett.* **393**, 51 – 57 (2004).
- <sup>54</sup>Y. Zhao and D. G. Truhlar, “The m06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four m06-class functionals and 12 other functionals,” *Theor. Chem. Acc.* **120**, 215–241 (2008).
- <sup>55</sup>J.-D. Chai and M. Head-Gordon, “Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections,” *Phys. Chem. Chem. Phys.* **10**, 6615–6620 (2008).
- <sup>56</sup>C. Angeli, S. Borini, M. Cestari, and R. Cimiraglia, “A quasidegenerate formulation of the second order  $n$ -electron valence state perturbation theory approach,” *J. Chem. Phys.* **121**, 4043–4049 (2004).
- <sup>57</sup>N. Forsberg and P. Åke Malmqvist, “Multiconfiguration perturbation theory with imaginary level shift,” *Chem. Phys. Lett.* **274**, 196 – 204 (1997).
- <sup>58</sup>J. W. Park, R. Al-Saadon, N. E. Strand, and T. Shiozaki, “Imaginary shift in caspt2 nuclear gradient and derivative coupling theory,” *J. Chem. Theory Comput.* **15**, 4088–4098 (2019).

- <sup>59</sup>Y. Tawada, T. Tsuneda, S. Yanagisawa, T. Yanai, and K. Hirao, “A long-range-corrected time-dependent density functional theory,” *J. Chem. Phys.* **120**, 8425–8433 (2004).
- <sup>60</sup>G. Ghigo, B. O. Roos, and P.-Å. Malmqvist, “A modified definition of the zeroth-order hamiltonian in multiconfigurational perturbation theory (caspt2),” *Chem. Phys. Lett.* **396**, 142 – 149 (2004).
- <sup>61</sup>I. Fdez. Galván, M. Vacher, A. Alavi, C. Angeli, J. Autschbach, J. J. Bao, S. I. Bokarev, N. A. Bogdanov, R. K. Carlson, L. F. Chibotaru, and et al., “Openmolcas: From source code to insight,” (2019).
- <sup>62</sup>F. Aquilante, P.-Å. Malmqvist, T. B. Pedersen, A. Ghosh, and B. O. Roos, “Cholesky decomposition-based multiconfiguration second-order perturbation theory (cd-caspt2): Application to the spin-state energetics of coiii(diiminato)(nph),” *J. Chem. Theory Comput.* **4**, 694–702 (2008).
- <sup>63</sup>A. Barbe, C. Secroun, and P. Jouve, “Infrared spectra of  $^{16}\text{O}_3$  and  $^{18}\text{O}_3$ : Darling and dennison resonance and anharmonic potential function of ozone,” *J. Mol. Spectrosc.* **49**, 171 – 182 (1974).
- <sup>64</sup>“Nist computational chemistry comparison and benchmark database, nist standard reference database number 101 release 19, april 2018, editor: Russell d. johnson iii <http://cccbdb.nist.gov/> (accessed on june 20, 2019),”.
- <sup>65</sup>R. Trambarulo, S. N. Ghosh, C. A. Burrus, and W. Gordy, “The molecular structure, dipole moment, and g factor of ozone from its microwave spectrum,” *J. Chem. Phys.* **21**, 851–855 (1953).
- <sup>66</sup>A. Barbe, C. Secroun, and P. Jouve, “Infrared spectra of  $^{16}\text{O}_3$  and  $^{18}\text{O}_3$ : Darling and dennison resonance and anharmonic potential function of ozone,” *J. Mol. Spectr.* **49**, 171–182 (1974).
- <sup>67</sup>P. Borowski, B. O. Roos, S. C. Racine, T. J. Lee, and S. Carter, “The ozonide anion: A theoretical study,” *J. Chem. Phys.* **103**, 266–273 (1995).
- <sup>68</sup>D. W. Arnold, C. Xu, E. H. Kim, and D. M. Neumark, “Study of low-lying electronic states of ozone by anion photoelectron spectroscopy of  $\text{o}_3^-$ ,” *J. Chem. Phys.* **101**, 912–922 (1994).
- <sup>69</sup>L. Alvarez-Valtierra, X.-Q. Tan, and D. W. Pratt, “On the role of methyl torsional modes in the intersystem crossing dynamics of isolated molecules,” *J. Phys. Chem. A* **111**, 12802–12809 (2007).

- <sup>70</sup>R. E. Bandy, J. Nash, and T. S. Zwier, "The spectroscopic and photophysical effects of the position of methyl substitution. i. 4- and 5-methylpyrimidine," *J. Chem. Phys.* **95**, 2317–2335 (1991).
- <sup>71</sup>M. E. Casida, "Time-dependent density functional response theory for molecules," in *Recent Advances in Density Functional Methods*, edited by D. P. Chong (World Scientific, Singapore, 1995) Chap. 5, pp. 155–192.
- <sup>72</sup>N. O. C. Winter, N. K. Graf, S. Leutwyler, and C. Hättig, "Benchmarks for 0–0 transitions of aromatic organic molecules: Dft/b3lyp, adc(2), cc2, sos-cc2 and scs-cc2 compared to high-resolution gas-phase data," *Phys. Chem. Chem. Phys.* **15**, 6623–6630 (2013).
- <sup>73</sup>B. O. Roos and K. Andersson, "Multiconfigurational perturbation theory with level shift — the cr<sub>2</sub> potential revisited," *Chem. Phys. Lett.* **245**, 215 – 223 (1995).
- <sup>74</sup>D. Tuna, Y. Lu, A. Kosłowski, and W. Thiel, "Semiempirical quantum-chemical orthogonalization-corrected methods: Benchmarks of electronically excited states," *J. Chem. Theory Comput.* **12**, 4400–4422 (2016).
- <sup>75</sup>G. Orlandi, F. Zerbetto, and M. Z. Zgierski, "Theoretical analysis of spectra of short polyenes," *Cosmic Res.* **91**, 867–891 (1991).
- <sup>76</sup>C. Angeli and M. Pastore, "The lowest singlet states of octatetraene revisited," *J. Chem. Phys.* **134**, 184302 (2011).
- <sup>77</sup>J. H. Starcke, M. Wormit, J. Schirmer, and A. Dreuw, "How much double excitation character do the lowest excited states of linear polyenes have?" *Contemp. Phys.* **329**, 39 – 49 (2006).
- <sup>78</sup>W. Hu and G. K.-L. Chan, "Excited-state geometry optimization with the density matrix renormalization group, as applied to polyenes," *J. Chem. Theory Comput.* **11**, 3000–3009 (2015).
- <sup>79</sup>K. Boguslawski, "Targeting doubly excited states with equation of motion coupled cluster theory restricted to double excitations," *J. Chem. Theory Comput.* **15**, 18–24 (2019).
- <sup>80</sup>R. R. Chadwick, D. P. Gerrity, and B. S. Hudson, "Resonance raman spectroscopy of butadiene: Demonstration of a 2 <sup>1</sup>a<sub>g</sub> state below the 1 <sup>1</sup>b<sub>u</sub> v state," *Chem. Phys. Lett.* **115**, 24 – 28 (1985).
- <sup>81</sup>W. J. Buma, B. E. Kohler, and K. Song, "Location of the 2 <sup>1</sup>a<sub>g</sub> state in hexatriene," *J. Chem. Phys.* **92**, 4622–4623 (1990).



- <sup>82</sup>M. F. Granville, G. R. Holtom, and B. E. Kohler, “High-resolution one and two photon excitation spectra of trans, trans-1,3,5,7-octatetraene,” *J. Chem. Phys.* **72**, 4671–4675 (1980).
- <sup>83</sup>K. L. D’Amico, C. Manos, and R. L. Christensen, “Electronic energy levels in a homologous series of unsubstituted linear polyenes,” *J. Am. Chem. Soc.* **102**, 1777–1782 (1980).
- <sup>84</sup>R. Snyder, E. Arvidson, C. Foote, L. Harrigan, and R. L. Christensen, “Electronic energy levels in long polyenes:  $S_2 \rightarrow s_0$  emission in all-trans-1,3,5,7,9,11,13-tetradecaheptaene,” *J. Am. Chem. Soc.* **107**, 4117–4122 (1985).
- <sup>85</sup>J. P. Zobel, J. J. Nogueira, and L. González, “The ipea dilemma in caspt2,” *Chem. Sci.* **8**, 1482–1499 (2017).