Analytic First-Order Derivatives of CASPT2 with IPEA Shift

Yoshio Nishimoto[1, a)]

Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

(Dated: 13 April 2023)

Complete active space second-order perturbation theory (CASPT2) is useful for accurately predicting properties of complex electronic structures, but it is well known that it systematically underestimates excitation energies. The underestimation can be corrected using the ionization potential–electron affinity (IPEA) shift. In this study, analytic first-order derivatives of CASPT2 with IPEA shift are developed. CASPT2-IPEA is not invariant with respect to rotations among active molecular orbitals, and two additional constraint conditions are necessary in the CASPT2 Lagrangian to formulate analytic derivatives. The method developed here is applied to methylpyrimidine derivatives and cytosine, and minimum energy structures and conical intersections are located. By comparing energies relative to the closed-shell ground state, we find that the agreement with experiments and high-level calculations is indeed improved by inclusion of the IPEA shift. The agreement of geometrical parameters with high-level calculations may also be improved in some cases.

[a)]Electronic mail: nishimoto@kuchem.kyoto-u.ac.jp.

I. INTRODUCTION

Capturing the electron correlation as fully as possible is essential to accurately describe the complex electronic structures of atoms and molecules. The electron correlation is formally split into "static" and "dynamic" components, which result from near-degenerate configuration state functions (CSFs) and instantaneous repulsion of electrons, respectively. The static electron correlation is described using the multiconfiguration self-consistent field (MCSCF) method, in which the wavefunction is expressed as a linear combination of the selected CSFs. A particularly useful variant of the MCSCF method is the complete active space SCF (CASSCF) method. We define an active space and express the wavefunction by performing the full configuration interaction within the active space. However, the static electron correlation is not sufficient to describe complex electronic structures; we also require the dynamic electron correlation, which is treated with post-MCSCF methods. These are typically referred to as multireference (MR) methods and include MR coupled-cluster[1–3] and MR configuration interaction (MRCI)[4] approaches. Although both of these MR approaches are accurate, their computational cost is high. Another type of MR approach is MR perturbation theory (MRPT). Owing to recent developments of analytic derivative theories,[5] MRPT is a practical choice for accurately determining molecular structures or even performing molecular dynamics (MD) simulations.

The complete active space second-order perturbation theory (CASPT2)[6–8] is one of the most well-known MRPTs. Other MRPTs, including (extended)[9] multiconfiguration quasi-degenerate PT2 [(X)MCQDPT2],[10] n-electron valence state PT2 (NEVPT2),[11–13] generalized Van Vleck PT2 (GVVPT2),[14] retaining the excitation degree PT (REPT),[15] unitary group adapted state-specific MRPT (UGA-SSMRPT),[16] Jeziorski–Monkhorst MRPT2,[17] and driven similarity renormalization group state-averaged MRPT (SA-DSRG-MRPT2),[18] are employed in many applications. As they can capture static and dynamic electron correlations in a balanced manner, we typically expect the prediction of excitation energies using MRPTs to be within an error margin of 0.2 eV compared with experimental results.

The initial implementation of (full) CASPT2 goes back to 1992,[8] but the complexity of this MRPT has hindered the development of its analytic derivative. The first analytic gradient implementation for CASPT2 was achieved in 2003,[19] and that for GVVPT2 was developed in the same year[20] by evaluating the derivative of wavefunction parameters explic-

itly (solving coupled-perturbed equations). The former method employs a partial internal contraction scheme; this avoids the evaluation of higher-order reduced density matrices (RDMs), and therefore the computational demand may be high for large active spaces. After more than 10 years, analytic gradients for the fully internally contracted CASPT2 were implemented using an automatic code-generation technique.[21] Despite the complexity of the formalism, development of analytic derivatives is ongoing, and they have been proposed for many MRPTs: CASPT2,[19,21–24] GVVPT2,[25] NEVPT2,[26,27] (X)MCQDPT2,[28] and (SA-)DSRG-MRPT2.[29,30] Some of the above implementations can be used for locating minimum energy conical intersections (MECIs)[22,28,30–36] and performing MD simulations.[37]

However, it is well recognized that CASPT2 using the standard Fock operator tends to underestimate excitation energies.[38] The origin of this underestimation has been attributed to an overestimation of open-shell (typically excited) states, for which the ionization potential–electron affinity (IPEA) shift has been suggested as a remedy.[39] Some studies in the literature have statistically proved that vertical excitation energies are indeed improved with IPEA shift.[40,41] It has also been reported that geometrical parameters[42,43] and molecular properties[44,45] may be improved. In spite of its potential utility, no analytic derivatives using IPEA shift have yet been developed; this may be partially due to the requirement for introducing an empirical parameter into "*ab initio*" methodology rather than technical difficulties.

This paper describes the development of analytic first-order derivatives (gradient and derivative coupling vectors) for CASPT2 with IPEA shift (CASPT2-IPEA) as well as restricted active space (RAS)[46] perturbation theory (RASPT2[-IPEA])[47,48]. The modification required to compute the energy with IPEA shift is relatively simple,[39] but CASPT2-IPEA loses an invariant character of the original CASPT2 method, and two constraint conditions are also needed for formulating analytic derivatives. The origin of the problem is identified and solutions are discussed below in detail.

II. METHODOLOGY

In this section, the following indices are used:

- general molecular orbitals (MOs): p, q, r, s;

- inactive (doubly occupied) MOs: i, j;
- active MOs: t, u, v, w;
- secondary (virtual) MOs: a, b, c, d;
- atomic orbitals (AOs): μ, ν;
- CSFs: I, J;
- internal states: α, β, γ, δ (also used as Kronecker delta), θ ($\in \mathcal{P}$);
- external states: Γ, Δ, Θ ($\notin \mathcal{P}$).

Here, $\mathcal{P}$ is the reference space, which usually spans the states averaged in the reference SCF calculation. We assume that all states are equally averaged in SCF and included in the PT2 calculation, although this assumption is not in principle necessary.

For internally contracted basis (ICB) and internally contracted configurations (ICCs), which are introduced later, the following indices are also used:

- non-orthogonal ICB: η, ι, π, ρ, σ;
- orthogonal ICB : τ ($\underset{\sim}{\tau}$), υ ($\tilde{\upsilon}$);
- orthogonal ICCs : ϕ, χ.

A tilde under or over the symbol indicates linearly dependent or independent vectors, respectively. The orthogonal ICCs are always computed in the independent vector space, so the tilde over ϕ or χ is omitted.

A. Brief overview of CASPT2

The zeroth-order (reference) wavefunction for the internal state $\gamma \in \mathcal{P}$ is defined by $\left|\Psi_\gamma^{(0)}\right\rangle$, which is typically obtained by solving state-averaged CASSCF (SA-CASSCF) equations. In quasi-degenerate perturbation theory, the zeroth-order Hamiltonian is typically defined by

$$\hat{H}^{(0)} = \hat{P}\hat{F}\hat{P} + \hat{Q}\hat{F}\hat{Q}, \tag{1}$$

where $\hat{P} = \sum_{\gamma \in \mathcal{P}} \left|\Psi_\gamma^{(0)}\right\rangle\left\langle\Psi_\gamma^{(0)}\right|$ is the projector onto the reference space $\mathcal{P}$ and $\hat{Q} = 1 - \hat{P}$ is that onto the complementary space. In CASPT2, the Fock operator $\hat{F}$ is defined by

$$\hat{F} = \sum_{pq} f_{pq} \hat{E}_{pq}, \tag{2}$$

where $\hat{E}_{pq}$ is the one-electron spin-averaged excitation operator and f_{pq} is the Fock matrix element, which is defined later. The definition of the zeroth-order Hamiltonian strongly affects the smoothness of the potential energy surface (PES). In particular, the invariance with respect to rotations of the reference wavefunction[9] is a desirable property for MRPTs and is essential for smooth PESs. By operating the projector $\hat{P}$, the first term in Eq. (1) can be explicitly written as

$$\hat{P}\hat{F}\hat{P} = \sum_{\gamma,\delta \in \mathcal{P}} \left|\Psi_\gamma^{(0)}\right\rangle\left\langle\Psi_\gamma^{(0)}\right|\hat{F}\left|\Psi_\delta^{(0)}\right\rangle\left\langle\Psi_\delta^{(0)}\right|. \tag{3}$$

The original multistate CASPT2 (MS-CASPT2) and MCQDPT2 neglect the off-diagonal elements of $\left\langle\Psi_\gamma^{(0)}\right|\hat{F}\left|\Psi_\delta^{(0)}\right\rangle$; this approximation introduces severe discontinuities in the PES near state-crossing regions.[9,31,33,36,49,50] Granovsky[9] and Shiozaki[51] suggested that the artifact caused by the non-invariant character of MRPTs could be fixed by rotating the reference state without sacrificing efficiency. The resultant XMCQDPT2[9] and extended MS-CASPT2 (XMS-CASPT2)[51] can indeed describe uniform PESs, and recently the methodology was further extended to a modified zeroth-order Hamiltonian, in the form of extended dynamically weighted CASPT2 (XDW-CASPT2)[52] and rotated MS-CASPT2 (RMS-CASPT2).[50] In XMS-, XDW-, and RMS-CASPT2, the reference state is rotated to satisfy

$$\left\langle \widetilde{\Psi}_\gamma^{(0)} \left| \hat{F}^{\mathrm{SA}} \right| \widetilde{\Psi}_\delta^{(0)} \right\rangle = 0 \tag{4}$$

for $\gamma \neq \delta \in \mathcal{P}$. The rotated reference states $\left|\widetilde{\Psi}_\gamma^{(0)}\right\rangle$ are obtained by a unitary transformation of the reference state:

$$|\widetilde{\Psi}_\gamma^{(0)}\rangle = \sum_{\delta \in \mathcal{P}} |\Psi_\delta^{(0)}\rangle U_{\delta\gamma}. \tag{5}$$

In the code, the rotated reference states are obtained by rotating the reference configuration interaction coefficients $c_{I,\delta}$ similarly:

$$\widetilde{c}_{I,\gamma} = \sum_{\delta \in \mathcal{P}} c_{I,\delta} U_{\delta\gamma}. \tag{6}$$

For single-state CASPT2 (SS-CASPT2) and MS-CASPT2, the unitary matrix $U_{\delta\gamma}$ is just a unit matrix.

The Fock matrix element f_{pq}, appearing in Eq. (2), is

$$\begin{aligned} f_{pq} &= h_{pq} + \sum_{rs} \left((pq|rs) - \frac{1}{2}(pr|qs) \right) D_{rs} \\ &= h_{pq} + g_{pq}(\mathbf{D}), \end{aligned} \tag{7}$$

where h_{pq} and $(pq|rs)$ are the one-electron and electron repulsion integrals, respectively, and D_{rs} is the one-electron density matrix. In addition to the rotation of the reference state, one can choose different definitions for the density matrix and thus different CASPT2 flavors. SS-CASPT2 and MS-CASPT2[53] employ the state-specific density matrix:

$$D^{\gamma}_{pq} = \left\langle \Psi^{(0)}_{\gamma} \middle| \hat{E}_{pq} \middle| \Psi^{(0)}_{\gamma} \right\rangle. \tag{8}$$

RMS-CASPT2[50] also does so but with a rotated reference state:

$$\widetilde{D}^{\gamma}_{pq} = \left\langle \widetilde{\Psi}^{(0)}_{\gamma} \middle| \hat{E}_{pq} \middle| \widetilde{\Psi}^{(0)}_{\gamma} \right\rangle = \sum_{IJ} \widetilde{c}_{I,\gamma} \widetilde{c}_{J,\gamma} \left\langle I \middle| \hat{E}_{pq} \middle| J \right\rangle. \tag{9}$$

On the other hand, XMS-CASPT2 uses the state-averaged density matrix

$$D^{\mathrm{SA}}_{pq} = \frac{1}{N_{\mathrm{state}}} \sum_{\gamma \in \mathcal{P}} \widetilde{D}^{\gamma}_{pq}, \tag{10}$$

where N_{state} is the number of the reference state. In XDW-CASPT2, the density matrix used in the construction of the Fock operator for the perturbed state γ is averaged using dynamical weights ω^{δ}_{γ} as follows:

$$\overline{\mathbf{D}}^{\gamma} = \sum_{\delta \in \mathcal{P}} \omega^{\delta}_{\gamma} \widetilde{\mathbf{D}}^{\delta}, \tag{11}$$

where the dynamical weight is defined using the following Boltzmann-like function:

$$\omega^{\delta}_{\gamma} = \frac{\mathrm{e}^{-\zeta(\Delta_{\gamma\delta})^2}}{\sum_{\theta \in \mathcal{P}} \mathrm{e}^{-\zeta(\Delta_{\gamma\theta})^2}}. \tag{12}$$

The parameter ζ regulates the sharpness of the transition between mixed-density and state-specific regimes. Three expressions have been suggested for $\Delta_{\gamma\delta}$,[50,52] and the initial expression is defined as a difference in the rotated reference state energy. The Fock matrix element for XDW-CASPT2 can be obtained by

$$f_{pq} \rightarrow \overline{f}^{\gamma}_{pq} = h_{pq} + g_{pq}\left(\overline{\mathbf{D}}^{\gamma}\right), \tag{13}$$

and the MO coefficient matrix is obtained by block-diagonalizing $\overline{f}^{\gamma}_{pq}$ (quasi-canonical MOs). Thus, for SS-CASPT2, MS-CASPT2, RMS-CASPT2, and XDW-CASPT2, N_{state} sets of the MO coefficient matrix are defined. XMS-CASPT2 and RMS-CASPT2 are obtained by taking the limits $\zeta \to 0$ and $\to \infty$, respectively; therefore, in the following derivation, we use the zeroth-order Hamiltonian of XDW-CASPT2, although no XDW-CASPT2 calculations are performed in this study.

In CASPT2, the (variational) second-order perturbation energy $E^{(\text{PT2})}$ is obtained as a minimum of the Hylleraas functional:

$$E^{(\text{PT2})}_{\gamma} \overset{\text{min}}{=} E^{(2)}_{\gamma} = 2\left\langle \Psi^{(1)}_{\gamma} \middle| \hat{H} \middle| \widetilde{\Psi}^{(0)}_{\gamma} \right\rangle + \left\langle \Psi^{(1)}_{\gamma} \middle| \hat{H}^{(0)}_{\gamma} - E^{(0)}_{\gamma} + E_{\text{shift}} \middle| \Psi^{(1)}_{\gamma} \right\rangle, \tag{14}$$

where the first-order correction to the wavefunction $|\Psi^{(1)}_{\gamma}\rangle$ is generally defined by applying two-electron excitations to the (rotated) reference wavefunction:

$$|\Psi^{(1)}_{\gamma}\rangle = \sum_{pqrs} T^{\gamma}_{pqrs} \hat{E}_{pq}\hat{E}_{rs}|\widetilde{\Psi}^{(0)}_{\gamma}\rangle = \sum_{pqrs} T^{\gamma}_{pqrs}|\Phi^{\gamma}_{pqrs}\rangle. \tag{15}$$

Here, T^{γ}_{pqrs}, $\hat{E}_{pq}\hat{E}_{rs}$, and $|\Phi^{\gamma}_{pqrs}\rangle$ are the excitation amplitude, the two-electron excitation operator, and the doubly excited configuration, respectively. The excitation amplitude T^{γ}_{pqrs} is determined by solving the amplitude equation:

$$\frac{1}{2}\frac{\partial E^{(2)}_{\gamma}}{\partial T^{\gamma}_{pqrs}} = \left\langle \Phi^{\gamma}_{pqrs} \middle| \hat{H} \middle| \widetilde{\Psi}^{(0)}_{\gamma} \right\rangle + \left\langle \Phi^{\gamma}_{pqrs} \middle| \hat{H}^{(0)}_{\gamma} - E^{(0)}_{\gamma} + E_{\text{shift}} \middle| \Psi^{(1)}_{\gamma} \right\rangle = 0. \tag{16}$$

The dimension of this equation is generally large, and it is solved by an iterative procedure. The second term in Eq. (16) can be small, and the equation can be ill-conditioned. This issue is commonly referred to as the intruder state problem, and the level shift term E_{shift} has been introduced to avoid ill-conditioned equations. The term can be expressed as

$$E_{\text{shift}} = E^{\text{real}}_{\text{shift}} + \frac{\left(E^{\text{imaginary}}_{\text{shift}}\right)^2}{\hat{H}^{(0)}_{\text{D}} - E^{(0)}}, \tag{17}$$

where either $E^{\text{real}}_{\text{shift}}$ and $E^{\text{imaginary}}_{\text{shift}}$ is provided as a parameter, and $\hat{H}^{(0)}_{\text{D}}$ is the diagonal part of the zeroth-order Hamiltonian. As detailed below, this shift term has nothing to do with the IPEA shift. The unshifted second-order perturbation energy E^{PT2}_{γ}, with which the SS-CASPT2 energy is computed, is obtained by removing the shift correction:

$$E^{\text{PT2}}_{\gamma} = 2\left\langle \Psi^{(1)}_{\gamma} \middle| \hat{H} \middle| \widetilde{\Psi}^{(0)}_{\gamma} \right\rangle + \left\langle \Psi^{(1)}_{\gamma} \middle| \hat{H}^{(0)}_{\gamma} - E^{(0)}_{\gamma} \middle| \Psi^{(1)}_{\gamma} \right\rangle. \tag{18}$$

This energy is not variational with respect to changes in the amplitude $\mathbf{T}$, if the level shift technique has been employed.

Interactions between different perturbed states are considered in MS-CASPT2 variants ([X]MS-CASPT2, XDW-CASPT2, and RMS-CASPT2). We construct the (symmetrized) effective Hamiltonian

$$\overline{H}^{\mathrm{eff}}_{\gamma\delta} = \left\langle \widetilde{\Psi}^{(0)}_{\gamma} \middle| \hat{H} \middle| \widetilde{\Psi}^{(0)}_{\delta} \right\rangle + \frac{1}{2}\left(H^{(2)}_{\gamma\delta} + H^{(2)}_{\delta\gamma} \right), \tag{19}$$

where

$$H^{(2)}_{\gamma\delta} = \left\langle \Psi^{(1)}_{\gamma} \middle| \hat{H} \middle| \widetilde{\Psi}^{(0)}_{\delta} \right\rangle. \tag{20}$$

The effective Hamiltonian is then diagonalized to obtain the MS-CASPT2 energy in the form of eigenvalues:

$$E^{\mathrm{CASPT2}}_{\alpha} = \sum_{\gamma\delta\in\mathcal{P}} R_{\gamma\alpha} \overline{H}^{\mathrm{eff}}_{\gamma\delta} R_{\delta\alpha}. \tag{21}$$

B. IPEA shift

1. *Level shift and IPEA shift*

First, it will be convenient to distinguish between two types of shift technique employed in CASPT2: *level* and *IPEA* shift techniques. The first type includes the real[54] and imaginary[55] level shift technique, which has been expressed as Eq. (17). Very recently, a new level shift technique was proposed.[56] The main purpose of these level shift techniques is to avoid the intruder state problem involving an ill-conditioned amplitude equation (Eq. (16)), with minimal effect on the computed result. Perturbation energies computed with real level shift are generally sensitive to the shift parameter, hence the development of the imaginary level shift[55] and the regularized CASPT2[56]: these level shift techniques have lower dependence on the chosen shift parameter. Interested readers should refer to Ref. 56 for details of these three level shift techniques including the numerical assessment.

The other shift technique is the IPEA shift technique,[39] the main focus of this paper. In contrast to level shift techniques, its main purpose is to affect the computed result; specifically, to reduce the systematic underestimation of excitation energies by CASPT2 with the standard Fock operator. As this underestimation can be attributed to overestimation of the correlation energy for open-shell systems,[39,57] it can be alleviated by correcting the

denominator of the correlation energy (the second term in Eq. (18)) through modifying the zeroth-order Hamiltonian. The IPEA shift technique often also alleviates the intruder state problem, although this is not its primary objective.

2. *Brief overview of IPEA shift*

The idea of the IPEA shift is based on Koopmans' theorem for the closed-shell case, for which the orbital energies of the inactive and secondary orbitals can be written as the negative energies of the ionization potential (IP) and electron affinity (EA), respectively. In analogy, we want the orbital energies of the active orbitals to be the negatives of IP and EA for excitation out of and into partially occupied orbitals, respectively. Ghigo *et al.*[39] expressed the Fock matrix as a weighted average of the IP and EA:

$$\overline{f}_{pp} = -\frac{1}{2}\left(\widetilde{D}_{pp}\,(\mathrm{IP})_p + (2-\widetilde{D}_{pp})\,(\mathrm{EA})_p\right). \tag{22}$$

This expression satisfies Koopmans' theorem for the closed-shell case ($\widetilde{D}_{pp} = 0$ or 2). However, this identity does not hold for active MOs. In an extreme case, $\widetilde{D}_{tt} = 1$, the Fock matrix element can be written as an average of IP and EA, but this characteristic of the Fock matrix leads to overestimation of the perturbation energy for open-shell states.[39] To satisfy the Koopmans' theorem-like identity, Ghigo *et al.* introduced a shift. When an electron is excited into a partially occupied orbital, we want the Fock matrix element to be the negative of EA:

$$-(\mathrm{EA})_t = \overline{f}_{tt} + \frac{1}{2}\widetilde{D}_{tt}\left((\mathrm{IP})_t - (\mathrm{EA})_t\right) = \overline{f}_{tt} + \frac{1}{2}\widetilde{D}_{tt}B_{\mathrm{shift}}, \tag{23}$$

where the difference between IP and EA is replaced with the parameter B_{shift}. On the other hand, when an electron is excited out, the element is to be the negative of IP:

$$-(\mathrm{IP})_t = \overline{f}_{tt} - \frac{1}{2}(2-\widetilde{D}_{tt})\left((\mathrm{IP})_t - (\mathrm{EA})_t\right) = \overline{f}_{tt} - \frac{1}{2}(2-\widetilde{D}_{tt})B_{\mathrm{shift}}. \tag{24}$$

The IP and EA values, which have dependencies on the index of the active MO, cannot be straightforwardly defined; we therefore replace them with the shift parameter B_{shift}. The parameter was determined to minimize the error of the dissociation energy for 49 molecules[39]: $B_{\mathrm{shift}} = 0.25$ a.u., which is the default value in OpenMolcas. Hence, the IPEA shift adds $\frac{1}{2}\widetilde{D}_{tt}B_{\mathrm{shift}}$ or $-\frac{1}{2}(2-\widetilde{D}_{tt})B_{\mathrm{shift}}$ for Fock matrix elements in the active orbital when electrons

are excited in or out, respectively. Thus, the zeroth-order Hamiltonian will have a positive shift (with $B_{\text{shift}} > 0$) for open-shell character configurations, and the denominator part of the amplitude equation (the second term in Eq. (18)) shifts to a larger value.[57] Thus, the overestimation of the second-order perturbative contribution of these configurations is reduced. As a side effect, the positive shift often alleviates the intruder state problem.

The *posterior* correction breaks an inherent property of CASPT2: CASPT2-IPEA has an additional source of non-invariance. Specifically, the perturbation energy of CASPT2-IPEA is not invariant with respect to rotations among active MOs, whereas that of the original CASPT2 (including all MS-CASPT2 variants) is. The same non-invariance can be found in other MRPTs.[10,14,18] However, CASPT2 employs an internal contraction scheme, whereas the other MRPTs do not, so the non-invariance in CASPT2 poses a unique problem. As long as we are concerned with the perturbation energy, this non-invariance can only be detected by careful examination. This is because CASPT2 usually employs quasi-canonical orbitals in the perturbation part, and the Fock matrix in the active orbital block is thus diagonal; the perturbation energy is therefore uniquely determined for the given molecular and electronic structure unless there are no degeneracies in the active orbital. However, the non-invariance has to be carefully addressed when developing analytic derivatives, because part of the derivation can be simplified with the invariance property (i.e., the non-canonical approach).[58,59] Therefore, we cannot obtain analytic derivatives of CASPT2-IPEA just by evaluating the derivative contribution of the additional term (the third term of Eq. (25), for instance). Note that the non-invariance discussed here is totally different from that with respect to rotations among reference states, where the original MS-CASPT2 is not suitable for finding MECIs. The choice of internal contraction scheme (single-state single-reference (SS-SR) or multistate MR) is also irrelevant. In OpenMolcas, only the SS-SR internal contraction scheme is implemented, so the implemented XMS-CASPT2 is not fully invariant, although the original theory is.

3. Non-invariance of CASPT2-IPEA

To illustrate the origin of the non-invariance problem, let us consider how the IPEA shift is introduced in the formulation of CASPT2. The Hamiltonian matrix elements for the second simplest excitation class (G; $\hat{E}_{ai}\hat{E}_{bt}|\Psi_\gamma^{(0)}\rangle$) among eight[7] can be explicitly written

with quasi-canonical MOs:

$$\mathcal{B}^{\mathrm{G},\gamma}_{tu} = \mathcal{F}^{\mathrm{G},\gamma}_{tu} - E^{\gamma}_{\mathrm{sum}}\mathcal{S}^{\mathrm{G},\gamma}_{tu} + \frac{1}{2}\left(2 - \widetilde{D}^{\gamma}_{tu}\right) B_{\mathrm{shift}}\mathcal{S}^{\mathrm{G},\gamma}_{tu}\delta_{tu}, \tag{25}$$

where

$$\delta_{ac}\delta_{ij}\delta_{cd}\mathcal{F}^{\mathrm{G},\gamma}_{tu} = \left\langle \Phi^{\gamma}_{aibt} \middle| \sum_{vw} \overline{f}^{\gamma}_{vw}\hat{E}_{vw} \middle| \Phi^{\gamma}_{cjdu} \right\rangle = \sum_{v} \overline{f}^{\gamma}_{vv} \left\langle \widetilde{\Psi}^{(0)}_{\gamma} \middle| a^{\dagger}_{t}a^{\dagger}_{v}a_{v}a_{u} \middle| \widetilde{\Psi}^{(0)}_{\gamma} \right\rangle \tag{26}$$

$$E^{\gamma}_{\mathrm{sum}} = \sum_{t} \overline{f}^{\gamma}_{tt}\widetilde{D}^{\gamma}_{tt} \tag{27}$$

$$\delta_{ac}\delta_{ij}\delta_{bd}\mathcal{S}^{\mathrm{G},\gamma}_{tu} = \left\langle \Phi^{\gamma}_{aibt} \middle| \Phi^{\gamma}_{cjdu} \right\rangle = \left\langle \widetilde{\Psi}^{(0)}_{\gamma} \middle| a^{\dagger}_{t}a_{u} \middle| \widetilde{\Psi}^{(0)}_{\gamma} \right\rangle. \tag{28}$$

These matrices are necessary when the second term in Eqs. (14), (16), and (18) is evaluated. In this excitation class, one electron is excited out from the active block. This excitation class does not actually give linear dependence as long as the active space is sensible, as the overlap matrix $\boldsymbol{\mathcal{S}}^{\mathrm{G},\gamma}_{tu}$ is just the one-particle RDM, and its eigenvalue is the natural occupation number. Nevertheless, the origin of the non-invariance is already clear; the third term in Eq. (25) has a quadratic dependence on both indices t and u, so the term, specific to CASPT2-IPEA, cannot be written in a quadratic form and the basis cannot be uniquely transformed by simply multiplying a transformation matrix. CASPT2-IPEA is therefore non-invariant with respect to rotations in the basis: non-orthogonal ICB and active MOs, specifically. Consequently, first, the perturbation energy depends on the choice of the basis, and second, even if the derivative contribution of the IPEA shift term (the third term in Eq. (25)) is evaluated, derivatives computed using algorithms that rely on the invariance with respect to these rotations can have a relatively large error of up to 10^{-4} hartrees/bohr even for a simple molecule; see "Analytic0" in Table II. Necessary modifications to formulate analytic derivatives are discussed in Sections II C 3 and II C 4.

We next consider how the basis of matrix $\boldsymbol{\mathcal{B}}$ (indices for the excitation class and perturbed state γ are omitted) is transformed in the computation procedure to further analyze the non-invariance of CASPT2-IPEA. $\boldsymbol{\mathcal{B}}$ is initially computed in the active MO basis (see Eq. (25), for instance). However, the basis is generally non-orthonormal and can be linearly dependent, so we have to diagonalize the overlap matrix in the non-orthogonal basis first:

$$\left(\boldsymbol{\mathcal{X}}^{0}\right)^{\dagger}\boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{X}}^{0} = \mathbf{s}, \tag{29}$$

where $\mathbf{s}$ is a diagonal matrix whose elements are the eigenvalues of $\boldsymbol{\mathcal{S}}$. The vectors whose eigenvalues are less than a threshold (10^{-8} by default in OpenMolcas) are linearly depen-

dent (τ) and discarded. Using Löwdin's canonical orthogonalization,[60,61] we can define the transformation matrix $\boldsymbol{\mathcal{X}}$ only for linearly independent vectors ($\tilde{v}$):

$$\mathcal{X}_{\eta\tilde{v}} = \mathcal{X}^0_{\eta\tilde{v}} \cdot \frac{1}{\sqrt{s_{\tilde{v}}}}, \tag{30}$$

which satisfies $\boldsymbol{\mathcal{X}}^\dagger \boldsymbol{\mathcal{S}} \boldsymbol{\mathcal{X}} = \mathbf{I}$, where $\mathbf{I}$ is a unit matrix. Depending on the excitation class, the Hamiltonian (Eq. (25)) and overlap (Eq. (28)) have zero, two, four, or six indices of the active MO, but the explicit expression of the matrices is not essential in the derivation, so the indices are collectively written as η, ι, ...: $\mathcal{B}_{\eta\iota}$ and $\mathcal{S}_{\eta\iota}$. To distinguish between the two cases, the basis of the collective active MO (η, ι, ...) is referred to as a *non-orthogonal ICB* in this study, whereas the orthonormalized basis is an *orthogonal ICB* and is indexed with $\tilde{v}$ and $\tilde{\tau}$ for linearly independent and dependent vectors, respectively. To generate orthogonal configurations, we have to solve the following generalized eigenvalue equation:

$$\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{C}}\varepsilon, \tag{31}$$

where $\boldsymbol{\mathcal{C}}$ is the internally contracted coefficient. To solve this equation, we first transform the non-orthogonal basis to an orthogonal basis using the transformation matrix obtained above:[60]

$$\boldsymbol{\mathcal{B}}'\boldsymbol{\mathcal{C}}' = \boldsymbol{\mathcal{C}}'\varepsilon, \tag{32}$$

with unitary matrix $\boldsymbol{\mathcal{C}}'$ and

$$\boldsymbol{\mathcal{B}}' = \boldsymbol{\mathcal{X}}^\dagger \boldsymbol{\mathcal{B}} \boldsymbol{\mathcal{X}}. \tag{33}$$

The basis of $\boldsymbol{\mathcal{B}}'$ should span the linearly independent space. Finally, the standard eigenvalue equation is solved to obtain *orthogonal ICCs* (indexed with ϕ and χ):

$$\varepsilon = \boldsymbol{\mathcal{C}}^\dagger \boldsymbol{\mathcal{B}} \boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{C}}'^\dagger \boldsymbol{\mathcal{B}}' \boldsymbol{\mathcal{C}}', \tag{34}$$

where $\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{C}}'$. The left-hand side of Eq. (34) is in practice obtained as eigenvalues of Eq. (32). Two transformations, Eqs. (33) and (34), are associated with the non-invariance of CASPT2-IPEA. As the above algorithm is a standard procedure for solving generalized eigenvalue equations, the equation can be numerically solved. However, the basis of non-quadratic forms cannot be linearly transformed by the above matrix multiplications, so the solution ($\boldsymbol{\varepsilon}$ and $\boldsymbol{\mathcal{C}}$) is not obtained with the basis defined rigorously.

An interesting consequence of the non-invariance is the dependence on the orthonormalization procedure: perturbation energies are different with different orthonormalization

procedures when linear dependency exists, even if the number of independent vectors are the same and the diagonal approximation for the Fock operator (CASPT2-D)[7] is employed. The difference is noticeable even when only exact linear dependency exists; in other words, the first-order interacting space spans the same space. By contrast, the dependence is negligibly small without IPEA shift. A numerical example is summarized in Tables S1 and S2 (Supplementary Material). In this study, Löwdin's canonical orthonormalization is employed.[60,61] Other orthonormalizations may be used as long as certain conditions hold true (Eqs. (56)–(58)). Another consequence is the further violation of the size-extensivity; this is briefly discussed in Section IV A with some numerical results. Note that the CASPT2-IPEA energy is invariant with respect to rotations among orthogonal ICCs, as the IPEA shift does not modify the procedure for solving CASPT2 equations.

C. Analytic derivatives of non-invariant CASPT2

As discussed above, the IPEA shift breaks the non-invariant character of CASPT2, and thus the transformation from non-orthogonal to orthogonal ICB or from orthogonal ICB to ICC is not unique. Although both transformation steps are relevant, we can separately discuss two modifications for each transformation step. Before discussing the solution, we need to define the Lagrangian used in this study. In addition, the approach analytic derivatives are computed using the Lagrangian method[62], which is also briefly reviewed.

1. CASPT2-IPEA Lagrangian

Gradients are the derivative of the energy with respect to nuclear displacements. Based on the chain rule of the derivative, we may think that the derivatives of the wavefunction parameters (such as MO coefficients) have to be evaluated. This can be done by solving coupled-perturbed SCF equations; however, the evaluation is usually as complex as solving the corresponding SCF equation, so doing this for all nuclear displacements would be prohibitively expensive even for medium-sized systems. By employing the Lagrangian[62] or Z-vector[63] methods, we can avoid explicitly evaluating the dependence of the wavefunction parameters on the displacement. Similar to previous analytic derivative theories for MRPT methods, we use the Lagrangian technique.[62] The total Lagrangian $\mathcal{L}_{\alpha\beta}^{\mathrm{CASPT2}}$ is defined as the sum of

the SA-CASSCF and PT2 Lagrangians, $\mathcal{L}_{\alpha\beta}^{\mathrm{CASSCF}}$ and $\mathcal{L}_{\alpha\beta}^{\mathrm{PT2}}$: $\mathcal{L}_{\alpha\beta}^{\mathrm{CASPT2}} = \mathcal{L}_{\alpha\beta}^{\mathrm{CASSCF}} + \mathcal{L}_{\alpha\beta}^{\mathrm{PT2}}$. The SA-CASSCF part of the Lagrangian is similar to that described in a previous study[24] and consists of constraint conditions imposed when solving the SA-CASSCF equation:[64]

$$\mathcal{L}_{\alpha\beta}^{\mathrm{CASSCF}} = \sum_{pq} \overline{\kappa}_{pq}^{\alpha\beta} \frac{\partial E^{\mathrm{SA}}}{\partial \kappa_{pq}} + \sum_{\gamma\Gamma} \overline{P}_{\gamma\Gamma}^{\alpha\beta\perp} \frac{\partial E^{\mathrm{SA}}}{\partial P_{\gamma\Gamma}^{\perp}}, \tag{35}$$

where E^{SA} is the SA-CASSCF energy:

$$E^{\mathrm{SA}} = \sum_{\gamma\in\mathcal{P}} \omega_\gamma \left\langle \Psi_\gamma^{(0)} \middle| \hat{H} \middle| \Psi_\gamma^{(0)} \right\rangle = \sum_{\gamma\in\mathcal{P}} \omega_\gamma E_\gamma^{\mathrm{SCF}}. \tag{36}$$

The energy of the SCF (zeroth-order wavefunction) part is included in $\mathcal{L}_{\alpha\beta}^{\mathrm{PT2}}$, because it is dependent on the rotation matrix $R_{\alpha\beta}$ (see Eq. (21)). κ_{pq} and $P_{\gamma\Gamma}^{\perp}$ are the orbital rotation and state rotation (out of the $\mathcal{P}$ space) parameters, respectively. $\overline{\kappa}_{pq}^{\alpha\beta}$ and $\overline{P}_{\gamma\Gamma}^{\alpha\beta\perp}$ are the Lagrange multipliers, determined by solving Z-vector equations, as described later. Following the formulation in Ref. 64, the orthonormalization conditions of MOs are not explicitly written in the Lagrangian, but the so-called connection term,[65] which is a contraction between the overlap derivative and effective Fock matrices, is evaluated when taking the partial derivative with respect to nuclear displacements.

In this work, the PT2 part of the Lagrangian is defined as follows:

$$\mathcal{L}_{\alpha\beta}^{\mathrm{PT2}} = \sum_{\gamma\in\mathcal{P}} \mathcal{L}_{\gamma,\alpha\beta}^{\mathrm{PT2}} + \sum_{\gamma\neq\delta\in\mathcal{P}} w_{\gamma\delta}^{\alpha\beta} \langle \widetilde{\Psi}_\gamma^{(0)} | \hat{F}^{\mathrm{SA}} | \widetilde{\Psi}_\delta^{(0)} \rangle. \tag{37}$$

The first term is the PT2 Lagrangian, which is explicitly dependent on MOs for state γ, whereas the second term, which is added only for XMS-, XDW-, and RMS-CASPT2 and -RASPT2, is the constraint condition for the rotation of reference states (Eq. (4)). The first term in Eq. (37) is explicitly written as

$$\begin{aligned}
\mathcal{L}_{\gamma,\alpha\beta}^{\mathrm{PT2}} = {} & \frac{1}{2} \sum_{\delta\in\mathcal{P}} \overline{H}_{\gamma\delta}^{\mathrm{eff}} \left(R_{\gamma\alpha} R_{\delta\beta} + R_{\gamma\beta} R_{\delta\alpha} \right) \\
& + \sum_{pqrs} \lambda_{pqrs}^{\gamma,\alpha\beta} \left(\left\langle \Phi_{pqrs}^{\gamma} \middle| \hat{H} \middle| \widetilde{\Psi}_\gamma^{(0)} \right\rangle + \left\langle \Phi_{pqrs}^{\gamma} \middle| \hat{H}_\gamma^{(0)} - E_\gamma^{(0)} + E_{\mathrm{shift}} \middle| \Psi_\gamma^{(1)} \right\rangle \right) \\
& + \sum_{i}^{\mathrm{core}} \sum_{j}^{\mathrm{inactive}} \overline{\kappa}_{ij}^{\mathrm{core},\gamma,\alpha\beta} \overline{f}_{ij}^{\gamma} + \sum_{e} \sum_{\tilde{\tau}\tilde{v}\in e} \mathcal{W}_{\tilde{\tau}\tilde{v}}^{e,\gamma,\alpha\beta} \mathcal{S}_{\tilde{\tau}\tilde{v}}^{\mathrm{X},e,\gamma} + \sum_{t>u} \overline{\kappa}_{tu}^{\mathrm{act},\gamma,\alpha\beta} \overline{f}_{tu}^{\gamma}.
\end{aligned} \tag{38}$$

In Eq. (38), the first term is the CASPT2 energy ($\alpha = \beta$; Eq. (21)) or the coupling between states ($\alpha \neq \beta$). The second term corresponds to the variational condition for the amplitude

T^{γ}_{pqrs} (Eq. (16)), and the third term is the canonical condition for frozen MOs, which is required when the frozen core approximation is employed. The fourth term ensures the orthogonality between the independent and dependent orthogonal vectors in ICB (Eq. (29)), and the overlap matrix is defined by

$$\mathcal{S}^{\mathrm{X}}_{\underset{\sim}{\tau}\tilde{\upsilon}} = \sum_{\eta\iota} \mathcal{X}^{0*}_{\eta\underset{\sim}{\tau}} \mathcal{S}_{\eta\iota} \mathcal{X}^{0}_{\iota\tilde{\upsilon}} = 0. \tag{39}$$

The indices e and γ are omitted for clarity. In the summation, e (= A, B, C, D, E, F, G, H) corresponds to the eight excitation classes (Eqs. (1a), (1b), ..., and (1h) in Ref. 7). The fifth term is the canonical condition for active MOs. Compared with the previous formulation for MS-CASPT2 variants without IPEA shift,[36] the last two terms are newly added in this work. $\mathcal{W}^{e,\gamma,\alpha\beta}_{\underset{\sim}{\tau}\tilde{\upsilon}}$ and $\overline{\kappa}^{\mathrm{act},\gamma,\alpha\beta}_{tu}$ are Lagrange multipliers and are detailed in Sections II C 3 and II C 4, respectively.

In addition, it is useful for later derivations to express the PT2 energy (Eq. (18)) and thus the PT2 correction to the diagonal part of the effective Hamiltonian, in a similar manner to non-orthogonal second-order Møller–Plesset perturbation theory:[66,67]

$$\left\langle \Psi^{(1)}_{\gamma} \middle| \hat{H} \middle| \widetilde{\Psi}^{(0)}_{\gamma} \right\rangle = E^{\mathrm{PT2}}_{\gamma} = 2 \left\langle \Psi^{(1)}_{\gamma} \middle| \hat{H} \middle| \widetilde{\Psi}^{(0)}_{\gamma} \right\rangle + \sum_{pq} D^{\gamma,(2)}_{pq} \overline{f}^{\gamma}_{pq} + \sum_{e} \sum_{\phi\chi\in e} \mathcal{D}^{e,\gamma}_{\phi\chi} \mathcal{B}^{e,\gamma}_{\phi\chi}, \tag{40}$$

where the (unrelaxed) second-order correlated density matrix is

$$D^{\gamma,(2)}_{pq} = \left\langle \Psi^{(1)}_{\gamma} \middle| \hat{E}_{pq} \middle| \Psi^{(1)}_{\gamma} \right\rangle, \tag{41}$$

and the density in orthogonal ICCs is

$$\mathcal{D}^{c,\gamma}_{\phi\chi} = \left\langle \Psi^{(1)}_{\gamma} \middle| \hat{E}_{\phi\chi} \middle| \Psi^{(1)}_{\gamma} \right\rangle, \tag{42}$$

with the pseudo-excitation operator for orthogonal ICCs $\hat{E}_{\phi\chi}$. For instance, $\mathcal{D}^{e,\gamma}_{\phi\chi}$ (e = G; $\hat{E}_{ai}\hat{E}_{bt}$) can be evaluated by

$$\mathcal{D}^{\mathrm{G},\gamma}_{\phi\chi} = \sum_{i} \sum_{ab} T^{\mathrm{G},\gamma}_{aib\phi} T^{\mathrm{G},\gamma}_{aib\chi}. \tag{43}$$

All matrix elements of $D^{\gamma,(2)}_{pq}$ can be computed with the excitation amplitude if no IPEA shift is employed. For CASPT2-IPEA, however, the off-diagonal elements in the active (RAS1, RAS2, and RAS3 for RASPT2-IPEA) block cannot be computed similarly, as the first-order correction to the wavefunction is dependent on the choice of MOs, and the non-canonical approach cannot be applied.

2. *Lagrangian method*

To formulate analytic derivatives using the Lagrangian method, the Lagrangian must be made stationary to all independent wavefunction parameters (Ξ). In this study, $\Xi = \left\{\boldsymbol{\kappa},\ \mathbf{P}^{\perp},\ \mathbf{U},\ \boldsymbol{\mathcal{X}}^{0},\ \mathbf{T}\right\}$; if we can determine all Lagrange multipliers ($\overline{\kappa}^{\alpha\beta}_{pq}$, $\overline{P}^{\alpha\beta\perp}_{\gamma\Gamma}$, $w^{\alpha\beta}_{\gamma\delta}$, $\lambda^{\gamma,\alpha\beta}_{pqrs}$, $\overline{\kappa}^{\mathrm{core},\gamma,\alpha\beta}_{ij}$, $\mathcal{W}^{e,\gamma,\alpha\beta}_{\tilde{\tau}\tilde{v}}$, $\overline{\kappa}^{\mathrm{act},\gamma,\alpha\beta}_{tu}$) that satisfy

$$\frac{\partial \mathcal{L}^{\mathrm{CASPT2}}_{\alpha\beta}}{\partial \Xi} = 0, \tag{44}$$

the derivative of the CASPT2 energy can be computed with integral derivatives only:

$$\frac{\mathrm{d}E^{\mathrm{CASPT2}}_{\alpha\beta}}{\mathrm{d}x} = \frac{\mathrm{d}\mathcal{L}^{\mathrm{CASPT2}}_{\alpha\beta}}{\mathrm{d}x} = \frac{\partial \mathcal{L}^{\mathrm{CASPT2}}_{\alpha\beta}}{\partial x} + \cancel{\frac{\partial \mathcal{L}^{\mathrm{CASPT2}}_{\alpha\beta}}{\partial \Xi}} \cdot \frac{\mathrm{d}\Xi}{\mathrm{d}x} = \frac{\partial \mathcal{L}^{\mathrm{CASPT2}}_{\alpha\beta}}{\partial x}. \tag{45}$$

Thus, we can avoid explicitly evaluating the dependence of the wavefunction parameters on x (such as nuclear displacements). We just solve a few linear equations, which are usually referred to as Z-vector or λ-equations, so the computational cost for analytic derivatives is typically a few times more expensive than the energy calculation.

In solving the SA-CASSCF equation, the orbital and configuration interaction parameters are variationally optimized:

$$\frac{\partial E^{\mathrm{SA}}}{\partial \kappa_{pq}} = 0 \tag{46}$$

and

$$\frac{\partial E^{\mathrm{SCF}}_{\gamma}}{\partial P^{\perp}_{\gamma\Gamma}} = 0 \qquad \forall \qquad \gamma \in \mathcal{P}. \tag{47}$$

The (independent) orbital rotation parameter κ_{pq} does not span all MOs. The SA-CASSCF energy is invariant to rotations among inactive (doubly occupied), active, and secondary (virtual) orbitals; thus, rotations in the diagonal subspaces do not change the SA-CASSCF energy. Furthermore, the orthonormalization condition of the MO guarantees that the rotation of pq and qp is redundant. The pq for κ_{pq} therefore spans only the lower (or upper) part of the off-diagonal blocks: $(p, q) \in (\mathrm{active}, \mathrm{inactive}),\ (\mathrm{secondary}, \mathrm{inactive}),\ (\mathrm{secondary}, \mathrm{active})$. These rotations, which can change the SA-CASSCF energy, are referred to as independent orbital rotations. As the standard Z-vector equation is derived using the above two conditions, it cannot be used to determine the orbital rotation parameter for the diagonal and upper (or lower) off-diagonal blocks, because they have linear dependency.

A similar simplification applies to the PT2 part. The standard CASPT2 (i.e., without IPEA shift) is invariant with respect to rotations among inactive, active, and secondary

MOs, so we naturally expect that

$$\frac{\partial E_\alpha^{\mathrm{PT2}}}{\partial \kappa_{ij}} = \frac{\partial E_\alpha^{\mathrm{PT2}}}{\partial \kappa_{tu}} = \frac{\partial E_\alpha^{\mathrm{PT2}}}{\partial \kappa_{ab}} = 0 \qquad (B_{\mathrm{shift}} = 0). \tag{48}$$

The same holds true for E_α^{SCF}; therefore,

$$\frac{\partial E_\alpha^{\mathrm{CASPT2}}}{\partial \kappa_{ij}} = \frac{\partial E_\alpha^{\mathrm{CASPT2}}}{\partial \kappa_{tu}} = \frac{\partial E_\alpha^{\mathrm{CASPT2}}}{\partial \kappa_{ab}} = 0. \tag{49}$$

This expression indicates that there is no need to evaluate the orbital rotation ($\overline{\kappa}_{pq}$) for the diagonal blocks. By contrast, if an IPEA shift is employed, because of the non-invariance with respect to rotations among active MOs, the above equality no longer holds true:

$$\frac{\partial E_\alpha^{\mathrm{PT2}}}{\partial \kappa_{ij}} = \frac{\partial E_\alpha^{\mathrm{PT2}}}{\partial \kappa_{ab}} = 0 \qquad \frac{\partial E_\alpha^{\mathrm{PT2}}}{\partial \kappa_{tu}} \neq 0 \qquad (B_{\mathrm{shift}} \neq 0). \tag{50}$$

However, it is still possible to satisfy $\partial \mathcal{L}_{\alpha\beta}^{\mathrm{PT2}} / \partial \kappa_{tu} = 0$. This is the reason we need an additional constraint condition, and a Lagrange multiplier (the fifth term in Eq. (38)) is introduced. This is further detailed in Section II C 4. A similar invariance property can be used for the rotation among orthogonal ICCs:

$$\frac{\partial E_\alpha^{\mathrm{CASPT2}}}{\partial \kappa_{\phi\chi}^{e,\gamma}} = \frac{\partial E_\alpha^{\mathrm{PT2}}}{\partial \kappa_{\phi\chi}^{e,\gamma}} = 0 = (1 - \hat{\tau}_{\phi\chi}) \sum_\eta \mathcal{C}_{\eta\phi}^{e,\gamma} \frac{\partial E_\alpha^{\mathrm{PT2}}}{\partial \mathcal{C}_{\eta\chi}^{e,\gamma}} \qquad \forall \qquad e,\ \phi\chi \in e,\ \gamma \in \mathcal{P}, \tag{51}$$

where $\hat{\tau}_{\phi\chi}$ permutes the indices ϕ and χ. Therefore, the derivative of $\boldsymbol{\mathcal{C}}$ is not explicitly evaluated.

Among the Lagrange multipliers, we usually determine $\lambda_{pqrs}^{\gamma,\alpha\beta}$ first. If we employ the real or imaginary (not IPEA) shift or MS-CASPT2 variants, the CASPT2 energy (Eq. (21)) is no longer variational with respect to amplitude changes. Consequently, we need to solve the λ-equation for each γ:

$$\frac{\partial \mathcal{L}_{\alpha\beta}^{\mathrm{CASPT2}}}{\partial T_{pqrs}^{\gamma}} = \frac{\partial \mathcal{L}_{\gamma,\alpha\beta}^{\mathrm{PT2}}}{\partial T_{pqrs}^{\gamma}} = 0 \tag{52}$$

This equation is similar in the structure to the amplitude equation used in computing the CASPT2 energy (Eq. (14)), so it can be solved using the existing implementation.

3. Non-invariance with respect to orthogonal transformation

In this subsection, we discuss how the derivative of the internally contracted coefficient $\boldsymbol{\mathcal{C}}$ is evaluated with IPEA shift. First, as mentioned above, there is no need to explicitly

evaluate the derivative of the internally contracted coefficient $\boldsymbol{\mathcal{C}}$. We instead evaluate partial derivatives of the Lagrangian with respect to $\boldsymbol{\mathcal{C}}$ and obtain contributions to the derivative of the configuration interaction coefficient:

$$-\frac{1}{2}\sum_{\eta}\mathcal{C}^{e,\gamma}_{\eta\phi}\frac{\partial\mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial\mathcal{C}^{e,\gamma}_{\eta\chi}}\frac{\partial\mathcal{S}^{e,\gamma}_{\phi\chi}}{\partial P^{\gamma\perp}_{\delta\Delta}}. \tag{53}$$

This term has been computed in earlier analytic derivative studies of internally contracted MRPTs (CASPT2 and NEVPT2), and the origin of the term is similar to the AO derivative contribution: the so-called connection term[64,65] or the Pulay force.[68] In the context of ICCs, it may be considered that the term arises in order to keep the orthogonal ICCs orthogonalized after perturbative displacements of atoms. More explicitly, we can take the partial derivative of, for example, the third term in Eq. (40) and obtain

$$\sum_{\eta}\mathcal{C}^{e,\gamma}_{\eta\phi}\frac{\partial}{\partial\mathcal{C}^{e,\gamma}_{\eta\chi}}\sum_{e}\sum_{\phi'\chi'\in e}\mathcal{D}^{e,\gamma}_{\phi'\chi'}\mathcal{B}^{e,\gamma}_{\phi'\chi'}=(\mathcal{B}^{e,\gamma}_{\phi\phi}+\mathcal{B}^{e,\gamma}_{\chi\chi})\mathcal{D}^{e,\gamma}_{\phi\chi}. \tag{54}$$

Naturally, the orthogonality between the independent and dependence vectors should be maintained. However, extra care is needed in the case with IPEA shift. The above transformation is numerically straightforward, but this term is sufficient for analytic derivatives only when the basis of $\boldsymbol{\mathcal{B}}^{e,\gamma}$ can be uniquely transformed linearly. This is not the case with IPEA shift. Considering the fact that perturbation energies are dependent on the orthonormalization procedure only when linear dependent vectors exist as mentioned in Section II B 3, it may be hypothesized that dependent vector space is mistakenly exploited because of the non-invariance of CASPT2-IPEA. Therefore, we need an additional constraint condition that makes the independent and dependent vector spaces orthogonal. This is the reason that the additional term is added to the PT2 Lagrangian (the fourth term in Eq. (38)). If there are no linearly dependent vectors, although the dependence on active MOs (discussed in Section II C 4) remains, the orthogonal ICB spans exactly the same space as the non-orthogonal one. Therefore, the derivative contribution in the orthogonal ICB can be exactly expressed in the non-orthogonal ICB, and the following modification is not necessary; after derivation, we can confirm that no contributions arise.

To determine $\mathcal{W}_{\tilde{\tau}\tilde{\upsilon}}$, we solve the equation

$$\sum_{\eta}\mathcal{X}^{0}_{\eta\tilde{\upsilon}}\frac{\partial\mathcal{L}^{\mathrm{PT2}}}{\partial\mathcal{X}^{0}_{\eta\tilde{\tau}}}=0. \tag{55}$$

The indices for excitation classes (e), perturbed states (γ), and target states ($\alpha\beta$) are omitted for clarity. As the perturbation energy is not explicitly dependent on the linear dependent vector $\mathcal{X}^0_{\eta\underline{\tau}}$, a straightforward partial derivative is trivial. However, using the following identities:[69]

$$\sum_{\underline{\tau}} \mathcal{X}^0_{\eta\underline{\tau}} \mathcal{X}^0_{\iota\underline{\tau}} + \sum_{\tilde{\upsilon}} \mathcal{X}^0_{\eta\tilde{\upsilon}} \mathcal{X}^0_{\iota\tilde{\upsilon}} = \delta_{\eta\iota} \tag{56}$$

$$\sum_{\eta} \mathcal{X}^0_{\eta\underline{\tau}} \mathcal{C}_{\eta\phi} = 0 \tag{57}$$

$$\sum_{\tilde{\upsilon}} \sum_{\iota} \mathcal{X}^0_{\eta\tilde{\upsilon}} \mathcal{X}^0_{\iota\tilde{\upsilon}} \mathcal{C}_{\iota\chi} = \mathcal{C}_{\eta\chi}, \tag{58}$$

the transformation matrix $\boldsymbol{\mathcal{C}}$ can be written in a form dependent on the linear dependent vectors:

$$\mathcal{C}_{\eta\phi} = \left(\delta_{\eta\iota} - \sum_{\underline{\tau}} \mathcal{X}^0_{\eta\underline{\tau}} \mathcal{X}^0_{\iota\underline{\tau}} \right) \mathcal{C}_{\iota\phi}. \tag{59}$$

Using this expression and Eqs. (40) and (57), Eq. (55) can be solved directly. For the diagonal ($\delta = \gamma$; no contributions from $\delta \neq \gamma$) term of the first term in Eq. (38), using a non-orthogonal formulation (Eq. (40)) and taking its derivative:

$$\mathcal{W}_{\underline{\tau}\tilde{\upsilon}} = 2 \sum_{\eta\rho\sigma} \mathcal{B}_{\underline{\tau}\phi} \mathcal{D}_{\phi\chi} \mathcal{C}_{\eta\chi} \cdot \frac{1}{s_{\tilde{\upsilon}}} \mathcal{X}^0_{\eta\tilde{\upsilon}}, \tag{60}$$

where

$$\mathcal{B}_{\underline{\tau}\phi} = \sum_{\eta\theta} \mathcal{X}^0_{\eta\underline{\tau}} \mathcal{B}_{\eta\theta} \mathcal{C}_{\theta\phi}. \tag{61}$$

Note that although the transformation matrix $\boldsymbol{\mathcal{C}}$ is involved in the second term in Eq. (38) and the first term in Eq. (40), these terms do not contribute to the final derivatives (confirmed by actual computations), because these terms do not exhibit non-invariance, and contributions can be uniquely transformed into the non-orthogonal independent vector space. In practice, the contribution to the derivative of the overlap matrix is computed by back-transforming $\mathcal{W}_{\underline{\tau}\tilde{\upsilon}}$ into the non-orthogonal ICB:

$$\mathcal{W}_{\eta\iota} = 2 \sum_{\pi\rho\sigma} \left(\sum_{\underline{\tau}} \mathcal{X}^0_{\eta\underline{\tau}} \mathcal{X}^0_{\pi\underline{\tau}} \right) \mathcal{B}_{\pi\rho} \mathcal{D}_{\rho\sigma} \left(\sum_{\tilde{\upsilon}} \mathcal{X}_{\sigma\tilde{\upsilon}} \mathcal{X}_{\iota\tilde{\upsilon}} \right). \tag{62}$$

The contribution to partial derivatives is thus the second term of the following expression:

$$\frac{\partial \mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial P^{\gamma\perp}_{\delta\Delta}} \rightarrow \frac{\partial \mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial P^{\gamma\perp}_{\delta\Delta}} + \sum_{e} \sum_{\eta\iota\in e} \mathcal{W}^{e,\gamma,\alpha\beta}_{\eta\iota} \frac{\partial \mathcal{S}^{e,\gamma}_{\eta\iota}}{\partial P^{\gamma\perp}_{\delta\Delta}}. \tag{63}$$

Some additional contributions must be evaluated if λ-equations are to be solved or $E_{\text{shift}}^{\text{imaginary}} \neq 0$, but this can be done similarly using a slightly modified density matrix that has already been used in previous developments. If there are no linearly dependent vectors, $\sum_{\underset{\sim}{\tau}} \mathcal{X}^0_{\eta\underset{\sim}{\tau}} \mathcal{X}^0_{\pi\underset{\sim}{\tau}}$ in Eq. (62) spans the null space, so this term does not contribute to derivatives at all, as mentioned above.

The correction term does not solve the problem arising when the number of linearly independent vectors varies in different geometries. In that case, we may need other approaches, such as perturbative correction.[69] We may need to consider correction for MD simulation, as there is more chance of variation in the number of linearly dependent vectors than for geometry optimizations.

4. Non-invariance with respect to active MOs

Next, we need to determine $\overline{\kappa}_{tu}^{\text{act},\gamma,\alpha\beta}$. As mentioned, we require that the CASPT2 Lagrangian

$$\frac{\partial \mathcal{L}_{\alpha\beta}^{\text{PT2}}}{\partial \kappa_{tu}} = 0. \tag{64}$$

However, the CASPT2-IPEA energy is not invariant with respect to rotations among active MOs (Eq. (50)), so we need to obtain a term that is capable of making the Lagrangian stationary with respect to rotations among active MOs and also determine $\overline{\kappa}_{tu}^{\text{act},\gamma,\alpha\beta}$ to satisfy the above equation. These orbital rotations are redundant at the SCF level, and therefore the rotation cannot be uniquely determined by solving the standard Z-vector equation introduced later (Eq. (75)). Moreover, in contrast to the inactive and secondary orbitals, the rotation parameter (Lagrange multiplier) cannot be determined with

$$\overline{\kappa}_{tu}^{\text{act},\gamma,\alpha\beta} \neq -\frac{1-\hat{\tau}_{tu}}{\overline{f}_{tt}^{\gamma} - \overline{f}_{uu}^{\gamma}} \frac{\partial \mathcal{L}_{\gamma,\alpha\beta}^{\text{PT2}}}{\partial \kappa_{tu}^{\gamma}} \tag{65}$$

because of the linear dependence between orbital and state rotation parameters. On the other hand, $\overline{\kappa}_{ij}^{\text{core},\gamma,\alpha\beta}$ can be determined with

$$\overline{\kappa}_{ij}^{\text{core},\gamma,\alpha\beta} = -\frac{1-\hat{\tau}_{ij}}{\overline{f}_{ii}^{\gamma} - \overline{f}_{jj}^{\gamma}} \frac{\partial \mathcal{L}_{\gamma,\alpha\beta}^{\text{PT2}}}{\partial \kappa_{ij}^{\gamma}}. \tag{66}$$

The same non-invariance problem of MRPTs has been discussed in Refs. 20,28–30,35, and the Lagrange multiplier $\overline{\kappa}^{\text{act}}$ is determined by solving a modified Z-vector equation.

However, this approach is applicable only when $\overline{f}^{\gamma}_{pq}$ is constructed with the state-averaged density matrix (i.e., SS-CASPT2 with single-state CASSCF reference or XMS-CASPT2 using all internal states). In this study, our aim was to develop analytic derivatives using state-dependent zeroth-order Hamiltonian (the original MS-CASPT2, XDW-CASPT2, and RMS-CASPT2), so we instead determine $\overline{\kappa}^{\mathrm{act},\gamma,\alpha\beta}_{tu}$ by the iterative method introduced by Song *et al.* for CASPT2.[23]

First, we take the partial derivative of $\mathcal{L}^{\mathrm{CASPT2}}_{\gamma,\alpha\beta}$ with respect to κ^{γ}_{pq}, $P^{\gamma\perp}_{\delta\Delta}$, and $\kappa^{\mathrm{act},\gamma}_{tu}$ and obtain the matrix linear equation:

$$\begin{pmatrix} \frac{\partial^2 E^{\mathrm{SA}}}{\partial\kappa^{\gamma}_{pq}\partial\kappa^{\gamma}_{rs}} & \frac{\partial^2 E^{\mathrm{SA}}}{\partial\kappa^{\gamma}_{pq}\partial P^{\gamma\perp}_{\theta\Theta}} & \frac{\partial^2 E^{\mathrm{SA}}}{\partial\kappa^{\gamma}_{pq}\partial\kappa^{\mathrm{act},\gamma}_{vw}} \\ \frac{\partial^2 E^{\mathrm{SA}}}{\partial P^{\gamma\perp}_{\delta\Delta}\partial\kappa^{\gamma}_{rs}} & \frac{\partial^2 E^{\mathrm{SA}}}{\partial P^{\gamma\perp}_{\delta\Delta}\partial P^{\gamma\perp}_{\theta\Theta}} & \frac{\partial^2 E^{\mathrm{SA}}}{\partial P^{\gamma\perp}_{\delta\Delta}\partial\kappa^{\mathrm{act},\gamma}_{vw}} \\ \frac{\partial^2 E^{\mathrm{SA}}}{\partial\kappa^{\mathrm{act},\gamma}_{tu}\partial\kappa^{\gamma}_{rs}} & \frac{\partial^2 E^{\mathrm{SA}}}{\partial\kappa^{\mathrm{act},\gamma}_{tu}\partial P^{\gamma\perp}_{\theta\Theta}} & \frac{\partial^2 E^{\mathrm{SA}}}{\partial\kappa^{\mathrm{act},\gamma}_{tu}\partial\kappa^{\mathrm{act},\gamma}_{vw}} \end{pmatrix} \begin{pmatrix} \overline{\kappa}^{\gamma,\alpha\beta}_{rs} \\ \overline{P}^{\gamma,\alpha\beta\perp}_{\theta\Theta} \\ \overline{\kappa}^{\mathrm{act},\gamma,\alpha\beta}_{vw} \end{pmatrix} = - \begin{pmatrix} \frac{\partial\mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial\kappa^{\gamma}_{pq}} \\ \frac{\partial\mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial P^{\gamma\perp}_{\delta\Delta}} \\ \frac{\partial\mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial\kappa^{\mathrm{act},\gamma}_{tu}} \end{pmatrix}. \tag{67}$$

Here, pq and rs consist of the independent orbital rotations at the SCF level. As the rotations of active MOs and configuration interaction coefficients are linearly dependent, we assume the existence of a transformation matrix $\mathcal{T}^{\gamma}_{tu,\delta\Delta}$ that satisfies the following relation:

$$\frac{\partial E^{\mathrm{SA}}}{\partial\kappa^{\mathrm{act},\gamma}_{tu}} = \sum_{\delta\Delta}\mathcal{T}^{\gamma}_{tu,\delta\Delta}\frac{\partial E^{\mathrm{SA}}}{\partial P^{\gamma\perp}_{\delta\Delta}}. \tag{68}$$

Using this relation and the second row of Eq. (67), the third row of Eq. (67) can be written as

$$\sum_{vw}\left(\frac{\partial^2 E^{\mathrm{SA}}}{\partial\kappa^{\mathrm{act},\gamma}_{tu}\partial\kappa^{\mathrm{act},\gamma}_{vw}} - \sum_{\delta\Delta}\mathcal{T}^{\gamma}_{tu,\delta\Delta}\frac{\partial^2 E^{\mathrm{SA}}}{\partial P^{\gamma\perp}_{\delta\Delta}\partial\kappa^{\mathrm{act},\gamma}_{vw}}\right)\overline{\kappa}^{\mathrm{act},\gamma,\alpha\beta}_{vw} = -\frac{\partial\mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial\kappa^{\mathrm{act},\gamma}_{tu}} + \sum_{\delta\Delta}\mathcal{T}^{\gamma}_{tu,\delta\Delta}\frac{\partial\mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial P^{\gamma\perp}_{\delta\Delta}}, \tag{69}$$

which can be simplified to yield[23]

$$\overline{\kappa}^{\mathrm{act},\gamma,\alpha\beta}_{tu} = -\frac{1}{\overline{f}^{\gamma}_{tt} - \overline{f}^{\gamma}_{uu}}\left((1-\hat{\tau}_{tu})\frac{\partial\mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial\kappa^{\mathrm{act},\gamma}_{tu}} - \sum_{\delta\Delta}\mathcal{T}^{\gamma}_{tu,\delta\Delta}\frac{\partial\mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial P^{\gamma\perp}_{\delta\Delta}}\right). \tag{70}$$

The partial derivatives on the right-hand side were computed in earlier analytic derivative work, so we need to evaluate the transformation matrix $\mathcal{T}^{\gamma}_{tu,\delta\Delta}$. Unfortunately, there seem to be no analytic expressions, so we first solve the following linear equation:

$$\sum_{\theta\Theta}\frac{\partial^2 E^{\mathrm{SA}}}{\partial P^{\gamma\perp}_{\delta\Delta}\partial P^{\gamma\perp}_{\theta\Theta}}\overline{Q}^{\gamma,\alpha\beta\perp}_{\theta\Theta} = \frac{\partial\mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial P^{\gamma\perp}_{\delta\Delta}}, \tag{71}$$

where $\overline{Q}^{\gamma,\alpha\beta\perp}_{\theta\Theta}$ is the solution to this equation, and the second derivative of E^{SA} with respect to the state rotation parameter appears in an earlier analytic derivative study for

SA-CASSCF,[64] so this equation can be solved iteratively in a conventional way. Once we obtain the solution, we can easily transform the latter term in the parentheses of Eq. (70):

$$\sum_{\delta\Delta} \mathcal{T}^{\gamma}_{tu,\delta\Delta} \frac{\partial \mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial P^{\gamma\perp}_{\delta\Delta}} = \sum_{\delta\Delta,\theta\Theta} \mathcal{T}^{\gamma}_{tu,\delta\Delta} \frac{\partial^2 E^{\mathrm{SA}}}{\partial P^{\gamma\perp}_{\delta\Delta} \partial P^{\gamma\perp}_{\theta\Theta}} \overline{Q}^{\gamma,\alpha\beta\perp}_{\theta\Theta} = \sum_{\theta\Theta} \frac{\partial^2 E^{\mathrm{SA}}}{\partial \kappa^{\mathrm{act},\gamma}_{tu} \partial P^{\gamma\perp}_{\theta\Theta}} \overline{Q}^{\gamma,\alpha\beta\perp}_{\theta\Theta} \tag{72}$$

using Eqs. (68) and (71). The second derivative on the right-hand side of Eq. (72) also appears in analytic derivatives for SA-CASSCF, so it can also be contracted in a conventional way. Thus, $\overline{\kappa}^{\mathrm{act},\gamma,\alpha\beta}_{tu}$ can be determined with Eq. (70).

Solving Eq. (71) (for each $\gamma \in \mathcal{P}$) represents the major additional computational cost. After the application of some projections,[70] $\overline{Q}^{\gamma,\alpha\beta\perp}_{\theta\Theta}$ has a dimension of $N_{\mathrm{state}} \times N_{\mathrm{CSF}}$, where N_{CSF} is the number of CSFs, so the computational cost is strongly dependent on the length of the CSF expansion for the wavefunction (i.e., the active space). Without IPEA shift, $\overline{\kappa}^{\mathrm{act},\gamma,\alpha\beta}_{tu}$ computed with the iterative approach reproduces the off-diagonal elements of the unrelaxed second-order correlated density matrix (Eq. (41)) in the active block, but for the active space, it may be computed by

$$D^{\gamma,\alpha\beta,(2)}_{tu} = \frac{\partial \mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial \overline{f}^{\gamma}_{tu}}. \tag{73}$$

This expression can be used for computing the diagonal elements in the active space in either case.

The above formulation implies that N_{state} sets of the Z-vector equation have to be solved to obtain $\overline{\boldsymbol{\kappa}}$ and $\overline{\mathbf{P}}^{\perp}$; however, $\dfrac{\partial \mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial \kappa^{\gamma}_{pq}}$ and $\dfrac{\partial \mathcal{L}^{\mathrm{PT2}}_{\gamma,\alpha\beta}}{\partial P^{\gamma\perp}_{\delta\Delta}}$ can be readily transformed to a selected basis, so the Z-vector equation defined later (Eq. (75)) is solved only once per $\alpha\beta$. In the actual implementation, these partial derivatives are transformed from a state-dependent quasi-canonical to a natural orbital basis.

5. Final CASPT2 derivatives

Up to this point, we have determined the multipliers $\lambda^{\gamma,\alpha\beta}_{pqrs}$, $\overline{\kappa}^{\mathrm{core},\gamma,\alpha\beta}_{ij}$, $\mathcal{W}^{c,\gamma,\alpha\beta}_{\tau\tilde{v}}$, and $\overline{\kappa}^{\mathrm{act},\gamma,\alpha\beta}_{tu}$. The multiplier for the XMS condition $w^{\alpha\beta}_{\gamma\delta}$ in Eq. (37) can be determined by[71]

$$w^{\alpha\beta}_{\gamma\delta} = \frac{1}{2} \frac{1}{E^{(0)}_{\delta} - E^{(0)}_{\gamma}} \sum_{I} \left(\tilde{c}_{I,\gamma} \frac{\partial \mathcal{L}^{\mathrm{PT2}}_{\alpha\beta}}{\partial \tilde{c}_{I,\delta}} - \frac{\partial \mathcal{L}^{\mathrm{PT2}}_{\alpha\beta}}{\partial \tilde{c}_{I,\gamma}} \tilde{c}_{I,\delta} \right), \tag{74}$$

where $E^{(0)}_{\gamma}$ is the zeroth-order energy.

The remaining multipliers ($\overline{\kappa}^{\alpha\beta}_{pq}$ and $\overline{P}^{\alpha\beta\perp}_{\gamma\Gamma}$) can be determined by solving the Z-vector equation:

$$\frac{\partial \mathcal{L}^{\mathrm{CASPT2}}_{\alpha\beta}}{\partial \kappa_{pq}} = \frac{\partial \mathcal{L}^{\mathrm{CASPT2}}_{\alpha\beta}}{\partial P^{\perp}_{\gamma\Gamma}} = 0, \tag{75}$$

which is sometimes written as a super-matrix form:

$$\begin{pmatrix} \frac{\partial^2 E^{\mathrm{SA}}}{\partial \kappa_{pq} \partial \kappa_{rs}} & \frac{\partial^2 E^{\mathrm{SA}}}{\partial \kappa_{pq} \partial P^{\perp}_{\theta\Theta}} \\ \frac{\partial^2 E^{\mathrm{SA}}}{\partial P^{\perp}_{\delta\Delta} \partial \kappa_{rs}} & \frac{\partial^2 E^{\mathrm{SA}}}{\partial P^{\perp}_{\delta\Delta} \partial P^{\perp}_{\theta\Theta}} \end{pmatrix} \begin{pmatrix} \overline{\kappa}^{\alpha\beta}_{rs} \\ \overline{P}^{\alpha\beta\perp}_{\theta\Theta} \end{pmatrix} = - \begin{pmatrix} \frac{\partial \mathcal{L}^{\mathrm{CASPT2}}_{\gamma,\alpha\beta}}{\partial \kappa_{pq}} \\ \frac{\partial \mathcal{L}^{\mathrm{CASPT2}}_{\gamma,\alpha\beta}}{\partial P^{\perp}_{\delta\Delta}} \end{pmatrix}. \tag{76}$$

As mentioned above, the right-hand side is obtained by transforming a state-dependent quasi-canonical basis to a natural orbital basis. By solving the equation for $\overline{\kappa}_{rs}$ and $\overline{P}^{\perp}_{\theta\Theta}$, $\mathcal{L}^{\mathrm{CASPT2}}_{\alpha\beta}$ is stationary with respect to all wavefunction parameters Ξ, so the derivative of the CASPT2 energy with respect to nuclear coordinates x (i.e., gradient vector) is evaluated as a partial derivative of the Lagrangian:

$$\mathbf{g}^x_\alpha := \frac{\mathrm{d}E^{\mathrm{CASPT2}}_\alpha}{\mathrm{d}x} = \frac{\partial \mathcal{L}^{\mathrm{CASPT2}}_{\alpha\alpha}}{\partial x}. \tag{77}$$

The derivative coupling for (X)MS-CASPT2[31]

$$\mathbf{d}^x_{\alpha\beta} = \frac{1}{2}\left[\left\langle \widetilde{\Psi}^{(0)}_\alpha \middle| \frac{\mathrm{d}\left(\widetilde{\Psi}^{(0)}_\beta + \Psi^{(1)}_\beta\right)}{\mathrm{d}\mathbf{x}} \right\rangle + \left\langle \widetilde{\Psi}^{(0)}_\alpha + \Psi^{(1)}_\alpha \middle| \frac{\mathrm{d}\widetilde{\Psi}^{(0)}_\beta}{\mathrm{d}\mathbf{x}} \right\rangle \right] \tag{78}$$

can also be computed in a similar manner. However, for locating MECIs, we may omit the so-called CSF or determinant term $\mathbf{d}^{\mathrm{CSF}}_{\alpha\beta}$[31] (no modifications are required for CASPT2-IPEA) and use the non-adiabatic coupling (NAC) matrix elements:

$$\mathbf{h}^x_{\alpha\beta} := \frac{\partial \mathcal{L}^{\mathrm{CASPT2}}_{\alpha\beta}}{\partial x} \qquad (\alpha \neq \beta). \tag{79}$$

The gradient $\mathbf{g}^x_\alpha$ and NAC $\mathbf{h}^x_{\alpha\beta}$ vectors are the derivatives of the diagonal and off-diagonal elements of the diagonalized effective Hamiltonian, so the algorithm for computing $\mathbf{g}^x_\alpha$ can be used for $\mathbf{h}^x_{\alpha\beta}$ as well.

For RASSCF and RASPT2, the independent orbital rotation is differently defined. The one for CAS is (p, q) = (active, inactive) + (secondary, inactive) + (secondary, active), whereas that for RAS is additionally (p, q) += (RAS2, RAS1) + (RAS3, RAS1) + (RAS3, RAS2), because rotations between different RASs also change the energy.[72] Further, MOs are individually canonicalized in each RAS space, so the constraint condition for active MOs (the

last term in Eq. (38)) is differently defined:

$$\sum_{t>u}^{\text{active}} \overline{\kappa}_{tu}^{\text{act},\gamma,\alpha\beta} \overline{f}_{tu}^{\gamma} \to \sum_{t>u}^{\text{RAS1}} \overline{\kappa}_{tu}^{\text{act},\gamma,\alpha\beta} \overline{f}_{tu}^{\gamma} + \sum_{t>u}^{\text{RAS2}} \overline{\kappa}_{tu}^{\text{act},\gamma,\alpha\beta} \overline{f}_{tu}^{\gamma} + \sum_{t>u}^{\text{RAS3}} \overline{\kappa}_{tu}^{\text{act},\gamma,\alpha\beta} \overline{f}_{tu}^{\gamma}. \tag{80}$$

The algorithm to determine $\overline{\kappa}_{tu}^{\text{act},\gamma,\alpha\beta}$ is the same as that for CAS (Section II C 4), and the linear equation (Eq. (71)) is solved once per perturbed state γ. The independent orbital rotations are determined by solving the Z-vector equation (Eq. (75)) as in the previous RASPT2 study.

In addition to the two linear equations (SCF and the amplitude equation [Eq. (16)]) solved for CASPT2 energies, three additional linear equations (Eqs. (52), (71), and (75)) are solved for analytic derivatives. Of course, the choice of shift parameters (level-shift and IPEA shift parameters) can affect the convergence rate. In the Supplementary Material, the dependence of the convergence rate on the choice of the imaginary level-shift and IPEA parameters is briefly discussed. In short, as long as the parameters are chosen reasonably (e.g., less than 1.0), the convergence of these linear equations is smooth and the behavior is marginally affected.

III. COMPUTATIONAL DETAILS

The theory discussed in Section II C has been implemented in a development version of OpenMolcas.[73,74] All CASSCF and CASPT2 results were obtained with it using the atomic compact Cholesky decomposition[75,76] to generate an on-the-fly auxiliary basis set for the resolution-of-the-identity treatment of the electron repulsion integrals. All CASPT2 calculations were performed with the frozen core approximation and an imaginary level shift of 0.2i, and the IPEA shift parameter was set to either 0.00 (no IPEA) or 0.25 (IPEA) unless otherwise noted. The symmetry constraints were not used. Löwdin's canonical orthonormalization was employed for orthonormalizing the non-orthogonal ICB. Although any MS-CASPT2 variants could be used, XMS-CASPT2 and RMS-CASPT2 were employed in this study.

First, we briefly discuss the size-extensivity of single-state CASPT2 with and without IPEA shift using up to five ethylene molecules, each of which has been optimized at the Hartree–Fock/cc-pVDZ level of theory. For these calculations only, the imaginary level shift parameter was set to zero. Each ethylene molecule was separated by 100.0 Å, and the active

space for each monomer consisted of two electrons in two (π and π^*) orbitals (CAS(2e,2o) to CAS(10e,10o)). Then, the accuracy of the implemented analytic gradients is discussed using a comparison between numerical and "analytic" gradients. We first show that the two additional constraint conditions, discussed in Sections II C 3 and II C 4, are necessary for fully analytic derivatives using hydrogen fluoride with a separation of 1.60 Å at the XMS-CASPT2(2e,2o)/cc-pVDZ level of theory. The active space contained one σ and σ^* orbitals, and the first two states (SA2) were averaged in the reference CASSCF calculation. A similar comparison was made using planar butadiene at the XMS- and RMS-CASPT2 or RASPT2 levels as in Ref. 36.

Next, the smoothness of the PES is discussed using the MECI of pyramidalized ethylene, transoid butadiene (structures **b** and **k**, respectively, in Ref. 77), and a twisted-type MECI of cytosine. The reference wavefunctions were obtained at the three-state averaged (SA3)-CASSCF(6e,4o), SA2-CASSCF(4e,4o), and SA2-CASSCF(12e,9o) levels of theory, respectively.

Following an earlier work,[26] 0–0 transition energies of methylpyrimidine (MP) derivatives were computed using an SA2-CAS(10e,8o) active space consisting of three π, three π^*, and two n_N orbitals. Geometry optimizations and vibrational frequency calculations were performed with the cc-pVTZ basis set, whereas single-point energies were obtained with the aug-cc-pVTZ basis set.

Last, selected minimum energy and MECI structures of cytosine were determined with CASSCF and CASPT2, and relative energies were compared with MRCI calculations from Ref. 78. For comparison, the same basis set (cc-pVDZ) and active space (CAS(12e,9o) with four π, three π^*, one n_N, and one n_O orbitals) were selected, and the first four states (SA4) were equally averaged in the reference CASSCF calculation.

Some molecular structures are visualized in Supplementary Material (Figure S1).

IV. RESULTS AND DISCUSSION

A. Size-extensivity of single-state CASPT2-IPEA

Here, we briefly discuss the size-extensivity of CASPT2-IPEA. CASPT2 is not a size-extensive method even without IPEA shift. The theoretical basis for this has been already

TABLE I. Size-extensivity error (hartrees) of single-state CASSCF and CASPT2[a] with active spaces of CAS($2n$e,$2n$o) and the cc-pVDZ basis set.

Number of ethylenes (n)	CASSCF	CASPT2(0.00)	CASPT2(0.25)	CASPT2(0.50)	CASPT2(1.00)
2	0.000 000	−0.000 001	0.000 033	0.000 049	0.000 062
3	0.000 000	−0.000 002	0.000 065	0.000 097	0.000 123
4	0.000 000	−0.000 002	0.000 098	0.000 146	0.000 186
5	0.000 000	−0.000 005	0.000 129	0.000 194	0.000 247

[a] The values in parentheses for CASPT2 indicate the IPEA shift employed.

explained[72,79]; it is because the zeroth-order Hamiltonian is not separable to each cluster component. An IPEA shift increases the size-extensive error in many practical calculations but for a different reason; this is attributed to the non-invariance discussed in this study. Table I shows the size-extensivity error of CASSCF and CASPT2 with the cc-pVDZ basis set using up to five ethylene molecules. The weight of the closed-shell configuration for the monomer is 95.8%, being dominated by this CSF. Owing to the quasi-single configuration character, the size-extensive error of the CASPT2 calculation without IPEA shift is very small, as expected.[72] However, the error is noticeably increased with IPEA shift and is even larger with higher IPEA shift values. Therefore, applying CASPT2 to association or dissociation reactions may require extra care, in particular, when IPEA shift is employed. Note that RASSCF (unless with a quasi-CAS) and thus RASPT2 are also not size-extensive, but the size-extensive error for these RAS-based methods has a different origin; they are truncated configuration interaction-based methods.

B. Accuracy of implemented gradients

Next, we discuss the fact that two constraint conditions, the fourth and fifth terms in Eq. (38), are both necessary for accurate gradients. Table II shows the gradients of the H atom of HF (along the principal axis) and the deviation from the numerical gradients. "Analytic0" refers to the gradients using the original XMS-CASPT2 implementation[36] with the appropriate derivative contributions of the IPEA shift term (i.e., derivatives of the third

term in Eq. (25)) without the two constraint conditions mentioned above. "Analytic1" and "Analytic2" refer to the gradients with the first (Section II C 3) and second (Section II C 4) constraint conditions, respectively. "Analytic" includes all contributions, so we expect the computed gradients are fully analytic.

When no IPEA shift is employed, the errors from the numerical gradient with both "Analytic0" and "Analytic" are of the order of 10^{-7} hartree/bohr, which should be close to the limit of machine precision. This indicates that the original implementation[36] gives accurate gradients, and the additional constraint conditions do not contribute to the computed gradients, because XMS-CASPT2 (no IPEA)[51] is an invariant theory. Note that the implementation of XMS-CASPT2 in OpenMolcas has small non-invariance owing to the non-uniform first-order interacting space. On the other hand, when IPEA shift is employed, only the full analytic gradient ("Analytic" entry) reproduces the numerical gradient. Other approximate gradients ("Analytic0", "Analytic1", and "Analytic2" entries) can have errors of more than 3.0×10^{-5} hartrees/bohr for relatively simple molecular and electronic structures, so they do not give analytic gradients. Moreover, considering that the error for S_1 is large, the error may be even larger for more complex electronic structures. The magnitude of the error of the "Analytic" gradients is similar to that in the case without IPEA shift.

Geometry optimizations were also performed for verification. In Table III, optimized bond distances with the IPEA shift are summarized. As the error for S_0 gradients (Table II) was rather small, the optimized bond distances (and energies) for all gradients agreed. On the other hand, geometry optimizations using approximate gradients for S_1 did not fully converge. The number of independent vectors was constant during geometry optimizations, indicating that the failure of the geometry optimization was caused by the inaccuracy of the gradient. The distances obtained with fully analytic gradients reproduced those obtained with numerical gradients.

Note that the approach developed in this study cannot be applied to molecules that have degenerate active orbitals, for instance, benzene. In the course of determining $\overline{\kappa}_{tu}^{\mathrm{act},\gamma,\alpha\beta}$, we divide non-vanishing contributions by the difference of orbital energies (with quasi-canonical orbitals), which can be singular (Eq. (70)). A similar difficulty has been found with analytic derivatives of the strongly contracted variant of NEVPT2[27,34] which is not invariant with respect to rotations among inactive and secondary orbitals. In both cases, the origin of the problem is the arbitrariness of the perturbation energy. Hence, if the target system has

TABLE II. Numerical and analytic gradients (hartrees/bohr) of the hydrogen atom of hydrogen fluoride at SA2-XMS-CASPT2(2e,2o)/cc-pVDZ

	S_0	S_0 deviation	S_1	S_1 deviation
		no IPEA		
Numerical	−0.084 860 2	—	0.030 999 2	—
Analytic0	−0.084 860 2	−0.000 000 0	0.030 999 4	0.000 000 2
Analytic	−0.084 860 2	−0.000 000 0	0.030 999 4	0.000 000 2
		IPEA		
Numerical	−0.086 322 3	—	0.035 336 7	—
Analytic0	−0.086 357 9	−0.000 035 7	0.035 429 9	0.000 093 1
Analytic1	−0.086 282 5	0.000 039 8	0.035 644 5	0.000 307 8
Analytic2	−0.086 353 4	−0.000 031 1	0.035 236 9	−0.000 099 8
Analytic	−0.086 322 3	−0.000 000 0	0.035 336 9	0.000 000 2

TABLE III. (Un)optimized bond distances (Å) of hydrogen fluoride at SA2-XMS-CASPT2(2e,2o)/cc-pVDZ ($B_{\text{shift}} = 0.25$ a.u.) with numerical and analytic gradients

	S_0	S_0 deviation	S_1	S_1 deviation
Numerical	0.920 34	—	1.933 83	—
Analytic0	0.920 34	0.000 00	1.936 40[a]	0.002 57[a]
Analytic1	0.920 34	0.000 00	1.914 79[a]	−0.019 04[a]
Analytic2	0.920 34	0.000 00	1.935 60[a]	0.001 77[a]
Analytic	0.920 34	0.000 00	1.933 83	0.000 00

[a]Did not fully converge.

a high spatial symmetry or can have accidental degeneracies in the active orbital region, optimizing it with IPEA shift is not recommended.

For further verification, including applicability to RASPT2, the accuracy of the implemented gradient was tested by a comparison between analytic and numerical gradients with the XMS- and RMS-CASPT2 and RASPT2 methods for planar butadiene; see Table IV.

TABLE IV. Root mean square (R-M-S)[a] differences between analytic and numerical gradients (hartrees/bohr) at the SA2-XMS- and RMS-CASPT2 or RASPT2/cc-pVDZ levels of theory with an IPEA shift value of 0.25

	XMS-CASPT2/RASPT2		RMS-CASPT2/RASPT2	
Active space	S_0	S_1	S_0	S_1
CAS(4e,4o)	3.13×10^{-7}	2.71×10^{-7}	6.20×10^{-7}	5.77×10^{-7}
RAS(4e,4o)/(0e,0o)/2	7.21×10^{-7}	6.42×10^{-7}	9.49×10^{-7}	4.76×10^{-7}
RAS(4e,4o)/(2e,2o)/1	2.03×10^{-6}	1.80×10^{-6}	1.71×10^{-6}	2.32×10^{-6}

[a] Forces perpendicular to the planar axis are excluded from the R-M-S difference.

The definition of the RAS is provided in Ref. 36. The deviation from the numerical gradients was overall very similar to that of the case without IPEA shift,[36] indicating that all terms concerned with IPEA shift have been correctly derived and implemented.

C. Smoothness of the PES

MS-CASPT2 suffers from the non-invariance problem and has been considered to be unsuitable for finding MECIs. As shown in past studies,[31,49,50] the PES described with MS-CASPT2 can be rough, in particular, near CIs at CASSCF. Sources of the non-invariance have been solved using the modification by Shiozaki[51] based on the idea of XMCQDPT2 by Granovsky.[9] As the IPEA shift introduces another source of non-invariance, it is useful to check the quality of PESs described with CASPT2-IPEA and to confirm that the PESs do not have (significant) irregularities.

Figure 1 shows the PES of S_0 and S_1 around the MECI of pyramidalized ethylene located using the XMS-CASPT2 and RMS-CASPT2 methods (left panel) and the energy difference around the one located using SA3-CASSCF (right panel). The $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ vectors (in atomic units) were obtained as a linear combination of the gradient difference and NAC vectors following the definition given in Ref. 77, forming the branching plane. The vectors computed at each CASPT2 level were used in the left panel, whereas those at SA3-CASSCF were used in the right panel. In the left panel, the PES of XMS-CASPT2 (Figure 1 (A)) has no irregularities in the region of the plot, but that of RMS-CASPT2 (Figure 1 (B)) has one small discontinuity ($\sim$5 kcal/mol at $(\hat{\mathbf{x}}, \hat{\mathbf{y}}) = (0.0115, -0.0096)$) near the crossing point.

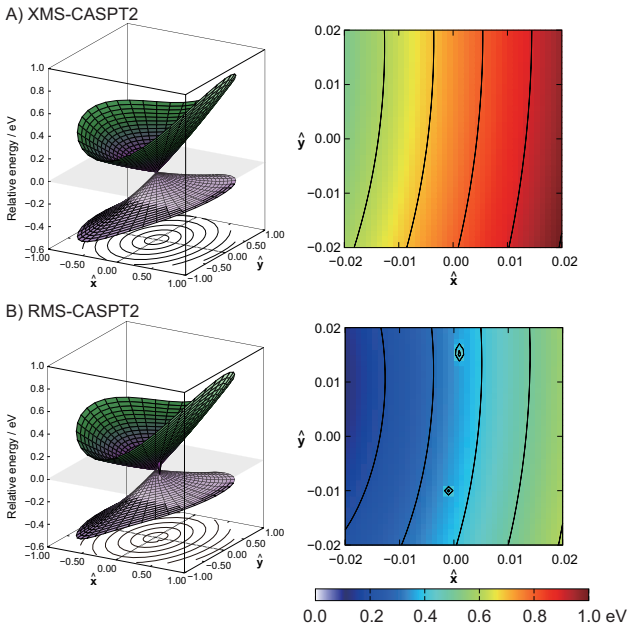

FIG. 1. PESs of pyramidalized ethylene around the CASPT2 MECIs (left) and energy difference between S_0 and S_1 around the SA3-CASSCF MECIs (right) with (A) XMS- and (B) RMS-CASPT2(6e,4o)-IPEA/cc-pVDZ methods.

Close investigation indicates that the discontinuity can be attributed to the difference in the zeroth-order Hamiltonian rather than the use of IPEA shift. XMS- and RMS-CASPT2 use the state-averaged and state-specific density matrix, respectively, for defining the Fock matrix. The use of the state-specific density matrix can be an additional source of non-invariance, so it is natural that RMS-CASPT2 can give less smooth PESs. Nevertheless, the PES is overall smooth, and the discontinuity is noticeable only at one point, so both XMS- and RMS-CASPT2 can be used for locating MECIs. A similar argument applies to the right panel; there are no irregularities in the energy difference described with XMS-CASPT2, whereas there are two (∼4 kcal/mol) with RMS-CASPT2. In spite of these small irregularities, there were no (significant) problems in determining minimum energy and MECI structures in the later sections.

As the non-invariance introduced by the IPEA shift is relevant to rotations among active MOs, we can assume that the non-invariance effect may be stronger with a larger active space. To further check this hypothesis, PESs of cytosine for $CI01_{twist}$ (see Section IV E) at the XMS-CASPT2(12e,9o)/cc-pVDZ level of theory with and without IPEA shift are

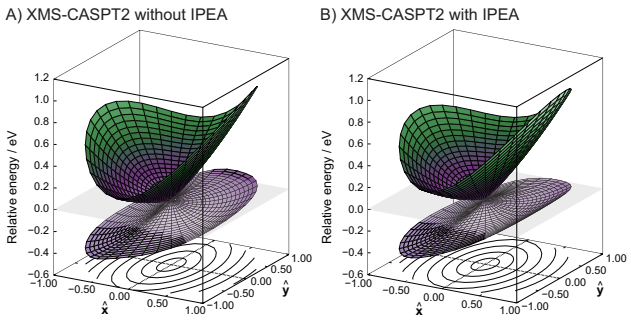

FIG. 2. PESs of $CI01_{twist}$ around the SA2-XMS-CASPT2(12e,9o)/cc-pVDZ MECIs (A) without and (B) with IPEA shift.

plotted in Figure 2. Although the same active space as described in Section IV E was used, only two states were averaged. The PESs were quite smooth, and there were no apparent discontinuities. The size of the active space may thus be a less important factor for the smoothness of PESs.

The smoothness of XMS-RASPT2 was also investigated using three active spaces for a transoid MECI of butadiene (structure **k** in Ref. 77): CAS(4e,4o), RAS(4e,4o)/(0e,0o)/2, and RAS(4e,4o)/(2e,2o)/1. We found small distortions even with CASPT2-IPEA, in particular, for a relatively large $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$. This was not seen in the previous XMS-CASPT2 and -RASPT2 result for the same system without IPEA shift.[36] Irregularities in the left panel were rather small, and geometry optimizations for these MECIs using CASPT2-IPEA and RASPT2-IPEA indeed did not exhibit any difficulties. In RAS(4e,4o)/(0e,0o)/2 (right panel in Figure 3 (B)), there were three noticeable discontinuities (∼5 kcal/mol). As expected, PESs described with RASPT2 could have more irregularities and discontinuities than those described with CASPT2 and more than those with IPEA shift, although it is possible to find MECIs using RASPT2-IPEA. In this study, no further RASPT2 calculations were performed; the applicability of RASPT2-IPEA for MECIs should be further investigated in future work.

D. Application to methylpyrimidine (MP) derivatives

The 0–0 transition energies of 2-, 4-, and 5-MP were computed at the CASPT2 level with and without IPEA shift, similar to a previous study.[26] The character of the S_1 transition

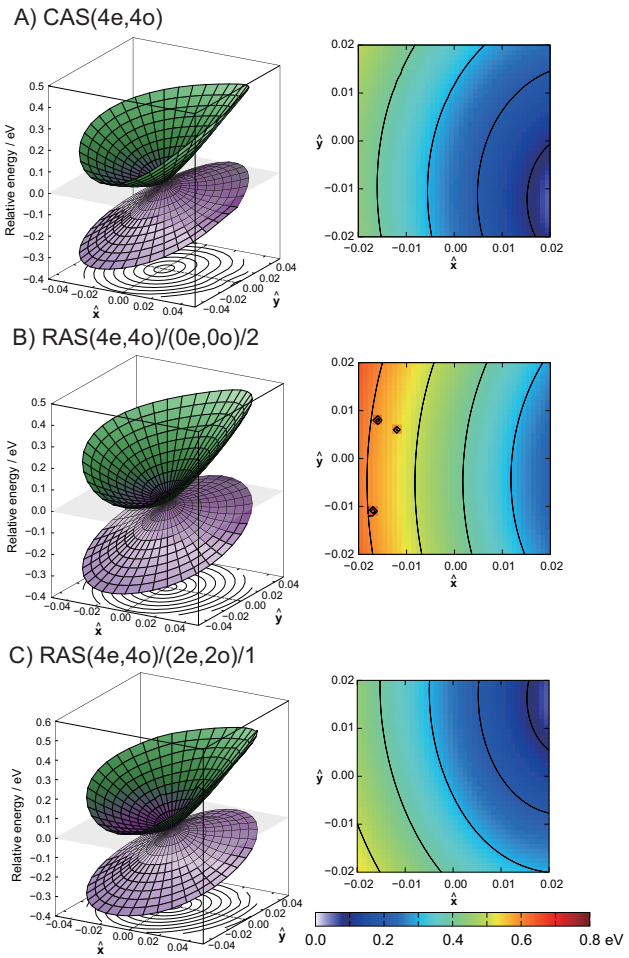

FIG. 3. PESs of butadiene around the CASPT2/RASPT2 MECIs (left) and energy difference between S_0 and S_1 around the SA2-CASSCF/RASSCF MECIs (right) with SA2-XMS-CASPT2-IPEA/cc-pVDZ or SA2-XMS-RASPT2-IPEA/cc-pVDZ using different active spaces: (A) CAS(4e,4o), (B) RAS(4e,4o)/(0e,0o)/2, and (C) RAS(4e,4o)/(2e,2o)/1.

of these MP derivatives is n $\rightarrow \pi^*$ transition. To obtain 0–0 transition energies, we first optimized the system at both the ground and excited states; the respective energies are denoted here by E^{GS} and E^{ES}. Then, at each minimum, vibrational frequency calculations were performed to obtain the zero-point vibrational energy (ZPVE; within harmonic approximations): $E^{\mathrm{GS}}_{\mathrm{ZPVE}}$ and $E^{\mathrm{ES}}_{\mathrm{ZPVE}}$. The 0–0 transition energy $E^{0\text{–}0}$ is defined as the difference between ZPVE-corrected energies: $E^{0\text{–}0} = (E^{\mathrm{ES}} + E^{\mathrm{ES}}_{\mathrm{ZPVE}}) - (E^{\mathrm{GS}} + E^{\mathrm{GS}}_{\mathrm{ZPVE}})$. The Hessian matrix, needed for vibrational frequency analysis, was obtained by a semi-numerical approach, in which analytic first-order geometrical derivatives were computed for displaced geome-

tries. In the previous study,[26] ZPVEs obtained with NEVPT2 were used for CASPT2, but they were computed using individual CASPT2 methods in this study. The deviations from the experimental 0–0 transition energies with CASSCF, (partially contracted) NEVPT2, XMS-CASPT2, and RMS-CASPT2 are summarized in Table V.

The best agreement with experiment was achieved by XMS-CASPT2 (IPEA) among the four CASPT2 methods. The maximum deviation was 0.092 eV. The improvement with IPEA shift was particularly significant for the n $\rightarrow \pi^*$ transition.[41] On the other hand, when IPEA shift was not employed, XMS-CASPT2 underestimated the 0–0 transition energy by 0.19–0.25 eV, which was expected given the well-known systematic underestimation. RMS-CASPT2 without IPEA shift underestimated the 0–0 transition energies even more, by roughly 0.15 eV. Use of the IPEA shift did not completely resolve the underestimation, but the deviation with IPEA shift was up to 0.11 eV, and the agreement with experiment was satisfactory, considering the expected accuracy of MRPTs. For this transition energy, the impact of the IPEA shift was 0.24–0.32 eV, and the deviation from experiment was reduced by approximately 0.2 and 0.3 eV for XMS- and RMS-CASPT2, respectively. The deviation of NEVPT2 from experiment was 0.07–0.13 eV. Considering that there are no empirical parameters in the formulation of NEVPT2 (and that it is size-extensive and free from the intruder state problem), NEVPT2 is also a promising and practical MRPT for determining molecular structures and excitation energies.

E. Application to cytosine

To further investigate the performance and applicability of CASPT2-IPEA, two minimum energy and five MECI structures of cytosine ($C_4H_5N_3O$) were obtained and compared with MRCI results.[78] In Table VI, energies relative to the S_0 state are summarized. For the two minimum energy structures, relative energies of the S_0 and S_1 are shown. For MECIs, the nomenclature used in Ref. 78 is employed, and only relative energies of the crossing states are shown; for instance, the crossing states of "CI12" are S_1 and S_2. In Ref. 78, three levels of MRCI calculations were performed. "MRCI1" included single-excitation CSF generated from the CAS, with the core 1s, σ orbitals, and one oxygen lone pair always remaining frozen. The next level "MRCI$\sigma\pi$1" further included the excitations from 14 σ orbitals and the second oxygen lone pair. The highest level "MRCI$\sigma\pi$2" included single excitations from

TABLE V. Deviations of computed[a] 0–0 transition energies (eV) of 2-, 4-, and 5-methylpyrimidine (MP) from experiment[80,81]

	2-MP	4-MP	5-MP
CASSCF[b]	+0.238	+0.233	+0.265
NEVPT2[b]	+0.129	+0.071	+0.131
XMS-CASPT2 (no IPEA)	−0.187	−0.249	−0.228
XMS-CASPT2 (IPEA)	+0.092	+0.060	+0.015
RMS-CASPT2 (no IPEA)	−0.349	−0.409	−0.395
RMS-CASPT2 (IPEA)	−0.047	−0.107	−0.075
Experiment	3.776[c]	3.920[d]	3.819[c]

[a]Using SA2-CASSCF(10e,8o) reference wavefunctions and cc-pVTZ and aug-cc-pVTZ basis sets for geometry optimization and single-point energy calculations, respectively.

[b]Ref. 26.

[c]Ref. 80.

[d]Ref. 81.

the σ electrons and one oxygen lone pair plus single and double excitations from the CAS orbitals into the virtual orbitals. In Table VI, the energies obtained with the highest level in the geometry optimization are shown, even if single-point calculations at a higher level were available. The character of the S_1 state for the optimized geometry at the ground state was $\pi \rightarrow \pi^*$ at the CASPT2 and MRCI levels, whereas S_1 and S_2 corresponded to the $n_N \rightarrow \pi^*$ (5.435 eV) and $\pi \rightarrow \pi^*$ (5.516 eV) transitions, respectively, at the CASSCF level. A similar reversal between CASSCF and CASPT2 has been found in previous studies,[82] and the order of the states can be sensitive to the choice of the theoretical level.[78] For the same reason, geometry optimizations for the S_1 ($\pi \rightarrow \pi^*$) state were performed with S_1 at CASPT2 (consistent with MRCI) and with S_2 at CASSCF.

Overall, CASPT2 without IPEA shift underestimated the relative energies as expected. CASPT2-IPEA predicted lower absorption energies than MRCI (5.136 eV at MRCI$\sigma\pi$2[78]). However, considering that the experimental absorption energy is approximately 4.7[83] or 4.5–4.6 eV (from the discussion in Ref. 82), RMS-CASPT2-IPEA predicted a reasonable vertical excitation energy. On the other hand, RMS-CASPT2 without IPEA shift under-

TABLE VI. Relative energies (eV) of cytosine using SA4-CASSCF(12e,9o)/cc-pVDZ reference wavefunctions

	CASSCF	XMS (no IPEA)	XMS (IPEA)	RMS (no IPEA)	RMS (IPEA)	MRCI[e]
S_1 energy [a]	5.516[c]	4.454	4.820	4.340	4.661	5.101[f]
S_0 energy [b]	2.223[d]	0.841	1.101	1.010	1.134	1.038[g]
S_1 energy [b]	4.160[c,d]	3.923	4.194	3.710	4.085	4.035[g]
$CI01_{sofa}$	3.758	4.103	4.283	3.915	4.206	4.406[g]
$CI01_{twist}$	4.370	3.541	3.704	3.558	3.799	4.260[g]
CI12	3.975	4.105	4.330	3.846	4.243	4.207[f]
CI12′	4.615	4.340	4.621	4.217	4.590	4.638[f]
CI23	4.722	4.573	4.815	4.535	4.850	4.847[f]

[a] Optimized at S_0.

[b] Optimized at S_1.

[c] S_2 energy.

[d] Optimized at S_2.

[e] Ref. 78.

[f] MRCI1.

[g] MRCI$\sigma\pi$1.

estimated significantly. The vertical absorption energies computed with the cc-pVTZ basis set (including geometry optimizations) were 4.454, 4.855, 4.297, and 4.703 eV at XMS (no IPEA), XMS (IPEA), RMS (no IPEA), and RMS (IPEA), respectively, so the impact of the basis set reached approximately 0.05 eV. Considering the perfect agreement of the 0–0 transition energy with MRCI1,[78] the reference adiabatic excitation energy (S_1 energy at the S_1 geometry) was expected to be approximately 4.0–4.1 eV. Again, RMS-CASPT2-IPEA predicted a reasonable transition energy, whereas RMS-CASPT2 without IPEA shift clearly underestimated. The experimental absorption and transition energies lie between the XMS-CASPT2 energies with and without IPEA shift, and it was somewhat difficult to judge the effectiveness of the IPEA shift.

Regarding the relative energies of MECIs, CASSCF failed to reproduce the energetic

order between $CI01_{sofa}$ and $CI01_{twist}$, which can be important MECIs for ultrafast radiationless decay of cytosine.[78] These MECIs cross between the closed-shell singlet and $\pi \rightarrow \pi^*$ states. By contrast, for higher-state MECIs, in particular, CI12′ (crossing between $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$) and CI23 ($n_O \rightarrow \pi^*$ and $n_N \rightarrow \pi^*$), CASSCF could predict the relative energies rather well, and the deviation from the reference MRCI result was approximately 0.1 eV. On the other hand, with CASPT2, the deviation from the reference MRCI result for $CI01_{twist}$ is larger than the expected accuracy of MRPT. However, the average energies of S_0 and S_1 in the single-point energy calculation at a higher level, MRCI$\sigma\pi$2,[78] were 4.269 and 3.980 eV for $CI01_{sofa}$ and $CI01_{twist}$, respectively, so CASPT2-IPEA was close to the expected accuracy. The energy gaps between S_0 and S_1 in the single-point calculation at MRCI$\sigma\pi$2 were as large as 0.455 and 0.166 eV,[78] respectively; accurate estimation of the relative energy for these MECIs seems to be challenging, and it would be desirable for geometries and energies to be obtained with even higher-level calculations. For higher-state MECIs (CI12 [crossing between $\pi \rightarrow \pi^*$ and $n_O \rightarrow \pi^*$], CI12′, and CI23), CASPT2 (in particular RMS-CASPT2) without IPEA shift predicted much lower energies than the reference MRCI, and the deviations could be much larger than those of CASSCF. On the other hand, CASPT2-IPEA (in particular, RMS-CASPT2-IPEA) predicted quite well with a deviation less than 0.05 eV for the higher-state MECIs.

It was still difficult to identify the better MS-CASPT2 variants, but RMS-CASPT2-IPEA seemed to predict better relative energies in this investigation overall. One merit of RMS-CASPT2 over XMS-CASPT2 is a weaker dependence on the number of averaged states in the reference CASSCF calculation[36] because of the state-specific nature of the zeroth-order Hamiltonian. However, comparisons with other theoretical results may require extra care. In the present comparison, MRCI used the state-averaged density matrix for constructing the Fock matrix, so the uniformity of the Fock matrix should be better reproduced with XMS-CASPT2. Moreover, RMS-CASPT2 has an additional source of non-invariance, so geometry optimizations with RMS-CASPT2 can be less stable than those performed with XMS-CASPT2. Although no rigorous comparison has been conducted, some extra geometry optimization steps were needed for RMS-CASPT2 compared with XMS-CASPT2. Regarding the difference between behaviors of geometry optimization without and with IPEA shift, geometry optimizations with IPEA shift were slightly less stable than those without it and required some extra steps, even though the instability was not critical for determining min-

imum energy and MECI structures.

Geometrical parameters were also compared with MRCI results. The R-M-S deviations of distance and angle for heavier atoms (C, N, and O elements) are separately summarized in Table VII. Although relative energies were largely affected by IPEA shift, the impact on geometrical parameters was not significant. In most cases, the use of the IPEA shift changed the bond distance by just a few mÅ, but it improved some entries significantly (XMS-CASPT2 for S_1 geometry and CI12′ and RMS-CASPT2 for CI12′). However, for S_0 minimum, CI12, and CI12′, the agreement of CASSCF was already extremely good, and the PT2 correction was significantly detrimental to the agreement. Nevertheless, the theoretical level employed for geometry optimization in Ref. 78 was not the highest, and further refinement of geometrical parameters will be desirable.

V. CONCLUSIONS

Analytic gradient and derivative coupling vectors for CASPT2 and its RASPT2 counterparts with IPEA shift were derived and implemented in the open-source package OpenMolcas. IPEA shift introduces an additional source of non-invariance: non-invariance with respect to rotations among non-orthogonal ICB and active MOs. Two additional constraints were introduced to formulate analytic derivatives: orthogonality of the independent and dependent vectors of the orthogonal transformation matrix and of the active MOs (the fourth and fifth terms in Eq. (38), respectively). Both constraint conditions are necessary for fully analytic derivatives. In spite of the additional non-invariance, PESs described with XMS-CASPT2-IPEA are generally smooth, but there is a greater chance of irregularities occurring compared with XMS-CASPT2 without IPEA shift. PESs described with RMS-CASPT2 can have a few discontinuities, but these are introduced by the use of the RMS approach and are not severe when locating minimum energy and MECI structures.

The developed method was applied to MP derivatives and cytosine. The use of an IPEA shift usually raised excited-state energies relative to the closed-shell ground state by 0.1–0.4 eV and improved the agreement with experiment or high-level calculations in most cases. The results of the limited comparison of geometrical parameters for cytosine suggest that IPEA shift may improve these too, but further investigations are necessary, in particular, with systems containing transition metal atoms.

TABLE VII. R-M-S deviations of geometrical parameters from MRCI results using SA4-CASSCF(12e,9o)/cc-pVDZ reference wavefunctions

	CASSCF	XMS (no IPEA)	XMS (IPEA)	RMS (no IPEA)	RMS (IPEA)
			Distance		
S_0 geometry[a]	0.010	0.019	0.018	0.015	0.016
S_1 geometry[b]	0.023[c]	0.024	0.012	0.017	0.016
$CI01^{b}_{sofa}$	0.021	0.016	0.016	0.012	0.014
$CI01^{b}_{twist}$	0.020	0.021	0.021	0.021	0.019
CI12[a]	0.007	0.017	0.016	0.021	0.015
CI12′[a]	0.005	0.024	0.016	0.043	0.025
CI23[a]	0.052	0.045	0.046	0.044	0.046
			Angle		
S_0 geometry[a]	0.6	1.1	1.1	0.9	0.9
S_1 geometry[b]	1.4[c]	1.6	0.9	1.3	1.2
$CI01^{b}_{sofa}$	2.1	1.8	1.7	1.8	1.6
$CI01^{b}_{twist}$	2.1	3.0	3.0	2.8	2.6
CI12[a]	0.6	1.0	1.2	1.6	1.8
CI12′[a]	0.6	1.7	1.0	3.9	2.2
CI23[a]	2.1	2.2	2.3	2.1	2.2

[a] Compared with MRCI1 results[78].

[b] Compared with MRCI$\sigma\pi$1 results[78].

[c] Optimized at S_2.

SUPPLEMENTARY MATERIAL

See the supplementary material for the dependence of the perturbation energy on the orthonormalization procedure, the dependence of the convergence behavior on shift parameters, and optimized coordinates of pyramidalized ethylene, 1,3-butadiene, methylpyrimidine derivatives, and cytosine.

ACKNOWLEDGMENTS

Y.N. thanks Professor Hayashi and Professor Kurashige (Kyoto University) for kind support. This work was supported by JSPS KAKENHI (grant no. 20K15230).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]M. Musiał, A. Perera, and R. J. Bartlett, "Multireference coupled-cluster theory: The easy way," J. Chem. Phys. **134**, 114108 (2011).

[2]D. I. Lyakh, M. Musiał, V. F. Lotrich, and R. J. Bartlett, "Multireference nature of chemistry: The coupled-cluster view," Chem. Rev. **112**, 182–243 (2012).

[3]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).

[4]P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, and R. Shepard, "Multiconfiguration self-consistent field and multireference configuration interaction methods and applications," Chem. Rev. **112**, 108–181 (2012).

[5]J. W. Park, R. Al-Saadon, M. K. MacLeod, T. Shiozaki, and B. Vlaisavljevich, "Multireference electron correlation methods: Journeys along potential energy surfaces," Chem. Rev. **120**, 5878–5909 (2020).

[6]B. O. Roos, P. Linse, P. E. M. Siegbahn, and M. R. A. Blomberg, "A simple method for the evaluation of the second-order-perturbation energy from external double-excitations with a CASSCF reference wavefunction," Chem. Phys. **66**, 197 – 207 (1982).

[7]K. Andersson, P.-Å. 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).

[8]K. Andersson, P.-Å. 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).

[9]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).

[10]H. Nakano, "Quasidegenerate perturbation theory with multiconfigurational self-consistent-field reference functions," J. Chem. Phys. **99**, 7983–7992 (1993).

[11]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).

[12]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).

[13]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).

[14]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).

[15]R. F. Fink, "Two new unitary-invariant and size-consistent perturbation theoretical approaches to the electron correlation energy," Chem. Phys. Lett. **428**, 461–466 (2006).

[16]A. Sen, S. Sen, P. K. Samanta, and D. Mukherjee, "Unitary group adapted state specific multireference perturbation theory: Formulation and pilot applications," J. Comput. Chem. **36**, 670–688 (2015).

[17]E. Giner, C. Angeli, Y. Garniron, A. Scemama, and J.-P. Malrieu, "A Jeziorski–Monkhorst fully uncontracted multi-reference perturbative treatment. I. Principles, second-order versions, and tests on ground state potential energy curves," J. Chem. Phys. **146**, 224108 (2017).

[18]C. Li and F. A. Evangelista, "Driven similarity renormalization group for excited states: A state-averaged perturbation theory," J. Chem. Phys. **148**, 124106 (2018).

[19]P. Celani and H.-J. Werner, "Analytical energy gradients for internally contracted second-order multireference perturbation theory," J. Chem. Phys. **119**, 5044–5057 (2003).

[20]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).

[21]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).

[22]T. Mori and S. Kato, "Dynamic electron correlation effect on conical intersections in photochemical ring-opening reaction of cyclohexadiene: MS-CASPT2 study," Chem. Phys. Lett. **476**, 97–100 (2009).

[23]C. Song, J. B. Neaton, and T. J. Martínez, "Reduced scaling formulation of CASPT2 analytical gradients using the supporting subspace method," J. Chem. Phys. **154**, 014103 (2021).

[24]Y. Nishimoto, "Analytic gradients for restricted active space second-order perturbation theory (RASPT2)," J. Chem. Phys. **154**, 194103 (2021).

[25]D. Theis, Y. G. Khait, and M. R. Hoffmann, "GVVPT2 energy gradient using a Lagrangian formulation," J. Chem. Phys. **135**, 044117 (2011).

[26]Y. Nishimoto, "Analytic first-order derivatives of partially contracted n-electron valence state second-order perturbation theory (PC-NEVPT2)," J. Chem. Phys. **151**, 114103 (2019).

[27]J. W. Park, "Analytical gradient theory for strongly contracted (SC) and partially contracted (PC) n-electron valence state perturbation theory (NEVPT2)," J. Chem. Theory Comput. **15**, 5417–5425 (2019).

[28]J. W. Park, "Analytical first-order derivatives of second-order extended multiconfiguration quasi-degenerate perturbation theory (XMCQDPT2): Implementation and application," J. Chem. Theory Comput. **16**, 5562–5571 (2020).

[29]S. Wang, C. Li, and F. A. Evangelista, "Analytic energy gradients for the driven similarity renormalization group multireference second-order perturbation theory," J. Chem. Theory Comput. **17**, 7666–7681 (2021).

[30]J. W. Park, "Analytical gradient theory for spin-free state-averaged second-order driven similarity renormalization group perturbation theory (SA-DSRG-MRPT2) and its applications for conical intersection optimizations," J. Chem. Theory Comput. **18**, 2233–2245 (2022).

[31]J. W. Park and T. Shiozaki, "Analytical derivative coupling for multistate CASPT2 theory," J. Chem. Theory Comput. **13**, 2561–2570 (2017).

[32]J. W. Park, "Single-state single-reference and multistate multireference zeroth-order hamiltonians in MS-CASPT2 and conical intersections," J. Chem. Theory Comput. **15**, 3960–3973 (2019).

[33]J. W. Park, "Analytical gradient theory for quasidegenerate n-electron valence state perturbation theory (QD-NEVPT2)," J. Chem. Theory Comput. **16**, 326–339 (2020).

[34]Y. Nishimoto, "Locating conical intersections using the quasidegenerate partially and strongly contracted NEVPT2 methods," Chem. Phys. Lett. **744**, 137219 (2020).

[35]J. W. Park, "Analytical gradient theory for resolvent-fitted second-order extended multiconfiguration perturbation theory (XMCQDPT2)," J. Chem. Theory Comput. **17**, 6122–6133 (2021).

[36]Y. Nishimoto, S. Battaglia, and R. Lindh, "Analytic first-order derivatives of (X)MS, XDW, and RMS variants of the CASPT2 and RASPT2 methods," J. Chem. Theory Comput. **18**, 4269–4281 (2022).

[37]J. W. Park and T. Shiozaki, "On-the-fly CASPT2 surface-hopping dynamics," J. Chem. Theory Comput. **13**, 3676–3683 (2017).

[38]K. Andersson and B. O. Roos, "Multiconfigurational second-order perturbation theory: A test of geometries and binding energies," Int. J. Quantum Chem. **45**, 591–607 (1993).

[39]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).

[40]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).

[41]R. Sarkar, P.-F. Loos, M. Boggio-Pasqua, and D. Jacquemin, "Assessing the performances of CASPT2 and NEVPT2 for vertical excitation energies," J. Chem. Theory Comput. **18**, 2418–2436 (2022).

[42]F. Ruipérez, F. Aquilante, J. M. Ugalde, and I. Infante, "Complete vs restricted active space perturbation theory calculation of the Cr_2 potential energy surface," J. Chem. Theory Comput. **7**, 1640–1646 (2011).

[43]T. Shiozaki and B. Vlaisavljevich, "Computational spectroscopy of the Cr–Cr bond in coordination complexes," Inorg. Chem. **60**, 19219–19225 (2021).

[44]M. Kepenekian, V. Robert, and B. Le Guennic, "What zeroth-order hamiltonian for CASPT2 adiabatic energetics of Fe(II)N_6 architectures?" J. Chem. Phys. **131**, 114702 (2009).

[45]S. Vela, M. Fumanal, J. Ribas-Ariño, and V. Robert, "On the zeroth-order hamiltonian for CASPT2 calculations of spin crossover compounds," J. Comput. Chem. **37**, 947–953 (2016).

[46]P.-Å. Malmqvist, A. Rendell, and B. O. Roos, "The restricted active space self-consistent-field method, implemented with a split graph unitary group approach," J. Phys. Chem. **94**, 5477–5482 (1990).

[47]P. Celani and H.-J. Werner, "Multireference perturbation theory for large restricted and selected active space reference wave functions," J. Chem. Phys. **112**, 5546–5557 (2000).

[48]P.-Å. Malmqvist, K. Pierloot, A. R. M. Shahi, C. J. Cramer, and L. Gagliardi, "The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO_2 and Cu_2O_2 systems," J. Chem. Phys. **128**, 204109 (2008).

[49]S. Gozem, F. Melaccio, A. Valentini, M. Filatov, M. Huix-Rotllant, N. Ferré, L. M. Frutos, C. Angeli, A. I. Krylov, A. A. Granovsky, R. Lindh, and M. Olivucci, "Shape of multireference, equation-of-motion coupled-cluster, and density functional theory potential energy surfaces at a conical intersection," J. Chem. Theory Comput. **10**, 3074–3084 (2014).

[50]S. Battaglia and R. Lindh, "On the role of symmetry in XDW-CASPT2," J. Chem. Phys. **154**, 034102 (2021).

[51]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).

[52]S. Battaglia and R. Lindh, "Extended dynamically weighted CASPT2: The best of two worlds," J. Chem. Theory Comput. **16**, 1555–1567 (2020).

[53]J. Finley, P.-Å. Malmqvist, B. O. Roos, and L. Serrano-Andrés, "The multi-state CASPT2 method," Chem. Phys. Lett. **288**, 299–306 (1998).

[54]B. O. Roos and K. Andersson, "Multiconfigurational perturbation theory with level shift — the Cr_2 potential revisited," Chem. Phys. Lett. **245**, 215–223 (1995).

[55]N. Forsberg and P.-Å. Malmqvist, "Multiconfiguration perturbation theory with imaginary level shift," Chem. Phys. Lett. **274**, 196–204 (1997).

[56]S. Battaglia, L. Fransén, I. Fdez. Galván, and R. Lindh, "Regularized CASPT2: An intruder-state-free approach," J. Chem. Theory Comput. **18**, 4814–4825 (2022).

[57]J. P. Zobel, J. J. Nogueira, and L. González, "The IPEA dilemma in CASPT2," Chem. Sci. **8**, 1482–1499 (2017).

[58]R. Moccia, "Variable bases in SCF MO calculations," Chem. Phys. Lett. **5**, 260–264 (1970).

[59]N. C. Handy, R. D. Amos, J. F. Gaw, J. E. Rice, and E. D. Simandiras, "The elimination of singularities in derivative calculations," Chem. Phys. Lett. **120**, 151–158 (1985).

[60]A. Szabo and N. S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structu (Macmillan Publishing Co., Inc., New York, 1982).

[61]I. Mayer, Simple Theorems, Proofs, and Derivations in Quantum Chemistry (Springer New York, New York, 2003).

[62]T. Helgaker and P. Jørgensen, "Configuration-interaction energy derivatives in a fully variational formulation," Theor. Chem. Acc. **75**, 111–127 (1989).

[63]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).

[64]J. Stålring, A. Bernhardsson, and R. Lindh, "Analytical gradients of a state average MCSCF state and a state average diagnostic," Mol. Phys. **99**, 103–114 (2001).

[65]T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen, and K. Ruud, "Recent advances in wave function-based methods of molecular-property calculations," Chem. Rev. **112**, 543–631 (2012).

[66]A. E. Azhary, G. Rauhut, P. Pulay, and H.-J. Werner, "Analytical energy gradients for local second-order Møller–Plesset perturbation theory," J. Chem. Phys. **108**, 5185–5193 (1998).

[67]M. Schütz, H.-J. Werner, R. Lindh, and F. R. Manby, "Analytical energy gradients for local second-order Møller–Plesset perturbation theory using density fitting approximations," J. Chem. Phys. **121**, 737–750 (2004).

[68]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).

[69]L. Joubert-Doriol, "Variational approach for linearly dependent moving bases in quantum dynamics: Application to Gaussian functions," J. Chem. Theory Comput. **18**, 5799–5809 (2022).

[70]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).

[71]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).

[72]T. Helgaker, P. Jørgensen, and J. Olsen, eds., Molecular Electronic-Structure Theory (John Wiley & Sons, Ltd., England, 2000).

[73]I. Fdez. Galván, M. Vacher, A. Alavi, C. Angeli, F. Aquilante, J. Autschbach, J. J. Bao, S. I. Bokarev, N. A. Bogdanov, R. K. Carlson, L. F. Chibotaru, J. Creutzberg, N. Dattani, M. G. Delcey, S. S. Dong, A. Dreuw, L. Freitag, L. M. Frutos, L. Gagliardi, F. Gendron, A. Giussani, L. González, G. Grell, M. Guo, C. E. Hoyer, M. Johansson, S. Keller, S. Knecht, G. Kovačević, E. Källman, G. Li Manni, M. Lundberg, Y. Ma, S. Mai, J. P. Malhado, P.-Å. Malmqvist, P. Marquetand, S. A. Mewes, J. Norell, M. Olivucci, M. Oppel, Q. M. Phung, K. Pierloot, F. Plasser, M. Reiher, A. M. Sand, I. Schapiro, P. Sharma, C. J. Stein, L. K. Sørensen, D. G. Truhlar, M. Ugandi, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov, O. Weser, T. A. Wesołowski, P.-O. Widmark, S. Wouters, A. Zech, J. P. Zobel, and R. Lindh, "OpenMolcas: From source code to insight," J. Chem. Theory Comput. **15**, 5925–5964 (2019).

[74]F. Aquilante, J. Autschbach, A. Baiardi, S. Battaglia, V. A. Borin, L. F. Chibotaru, I. Conti, L. De Vico, M. Delcey, I. Fdez. Galván, N. Ferré, L. Freitag, M. Garavelli, X. Gong, S. Knecht, E. D. Larsson, R. Lindh, M. Lundberg, P. Å. Malmqvist, A. Nenov, J. Norell, M. Odelius, M. Olivucci, T. B. Pedersen, L. Pedraza-González, Q. M. Phung, K. Pierloot, M. Reiher, I. Schapiro, J. Segarra-Martí, F. Segatta, L. Seijo, S. Sen, D.-C. Sergentu, C. J. Stein, L. Ungur, M. Vacher, A. Valentini, and V. Veryazov, "Modern quantum chemistry with [Open]Molcas," J. Chem. Phys. **152**, 214117 (2020).

[75]F. Aquilante, R. Lindh, and T. Bondo Pedersen, "Unbiased auxiliary basis sets for accurate two-electron integral approximations," J. Chem. Phys. **127**, 114107 (2007).

[76]F. Aquilante, L. Gagliardi, T. B. Pedersen, and R. Lindh, "Atomic cholesky decompositions: A route to unbiased auxiliary basis sets for density fitting approximation with tunable accuracy and efficiency," J. Chem. Phys. **130**, 154107 (2009).

[77]I. Fdez. Galván, M. G. Delcey, T. B. Pedersen, F. Aquilante, and R. Lindh, "Analytical state-average complete-active-space self-consistent field nonadiabatic coupling vectors: Implementation with density-fitted two-electron integrals and application to conical intersections," J. Chem. Theory Comput. **12**, 3636–3653 (2016).

[78]K. A. Kistler and S. Matsika, "Radiationless decay mechanism of cytosine: An ab initio study with comparisons to the fluorescent analogue 5-methyl-2-pyrimidinone," J. Phys. Chem. A **111**, 2650–2661 (2007).

[79]Y. A. Aoto, A. Bargholz, D. Kats, H.-J. Werner, and A. Köhn, "Perturbation expansion of internally contracted coupled-cluster theory up to third order," J. Chem. Theory Comput. **15**, 2291–2305 (2019).

[80]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).

[81]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).

[82]M. P. Fuelscher and B. O. Roos, "Theoretical study of the electronic spectrum of cytosine," J. Am. Chem. Soc. **117**, 2089–2095 (1995).

[83]F. Žaloudek, J. S. Novros, and L. B. Clark, "The electronic spectrum of cytosine," J. Am. Chem. Soc. **107**, 7344–7351 (1985).