

Analytic First-Order Derivatives of (X)MS, XDW, and RMS Variants of the CASPT2 and RASPT2 Methods

Yoshio Nishimoto,^{*,†} Stefano Battaglia,[‡] and Roland Lindh[‡]

[†]*Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan*

[‡]*Department of Chemistry - BMC, Uppsala University, P.O. Box 576, SE-75123 Uppsala,
Sweden*

E-mail: nishimoto@kuchem.kyoto-u.ac.jp

Abstract

Crossings between states involve complex electronic structures, making the accurate characterization of the crossing point difficult. In this study, the analytic derivatives of three complete active space second-order perturbation theory (CASPT2) variants as well as an extension of the restricted active space (RASPT2) are developed. These variants are applied to locating minimum energy conical intersections. Our results demonstrate that the three CASPT2 variants predict qualitatively similar results, but a recently developed variant, the rotated multistate CASPT2 (RMS-CASPT2), is least sensitive to the number of states considered in the calculation. We demonstrate that CASPT2 and the reference self-consistent field calculations predict qualitatively different energetics and bond lengths.

1 Introduction

Crossings between potential energy surfaces (PESs) of different states play an important role in photochemistry.¹⁻³ Typically, those between the same-spin states and spatial symmetries are referred to as conical intersections (CIs) and are responsible for non-radiative transitions between states. In particular, minimum energy CIs (MECIs) act as an efficient “funnel” around the CI region, with higher states efficiently and non-radiatively decaying to lower states. Because the electronic structure at crossing points is considerably more complex than the one at a minimum energy region of a PES, it is necessary to select a method that appropriately describes the electron correlation. One useful and well-known approach is the multiconfiguration self-consistent field (MCSCF) method, in particular the complete active space (CASSCF) version. Electron correlation described using MCSCF and CASSCF is typically referred to as “static” or “non-dynamic” electron correlation, which is attributed to near-degenerate configuration state functions (CSFs) or determinants, whereas “dynamic” electron correlation, which results from instantaneous repulsion of electrons, is missing. A large portion of the dynamic electron correlation can be recovered using post-MCSCF treatments, which are typically referred to as multireference (MR) methods, such as MR coupled-cluster⁴⁻⁶ and configuration interaction⁷ approaches. However, in terms of computational cost vs accuracy, MR perturbation theory (MRPT) may be the most balanced approach. In particular, when one wants to determine minimum energy structures or perform molecular dynamics (MD) simulations, MRPT is a practical choice for medium-sized systems among the abovementioned MR methods, because these calculations are more time-consuming than a single-point energy evaluation.

The most well-known MRPT is probably the CAS second-order PT (CASPT2) method,⁸⁻¹⁰ but many other versions exist, such as (extended)¹¹ multiconfiguration quasi-degenerate PT2 [(X)MCQDPT2],¹² n -electron valence state PT2 (NEVPT2),¹³⁻¹⁵ generalized Van Vleck PT2 (GVVPT2),¹⁶ retaining the excitation degree PT,¹⁷ unitary group adapted state-specific MRPT,¹⁸ and driven similarity renormalization group state-averaged MRPT (SA-DSRG-

MRPT2).¹⁹ There is no doubt that analytic derivatives are most convenient to locate minimum energy structures, such that analytic gradient theories and programs have been developed for a variety of MRPTs, namely CASPT2,^{20–23} GVVPT2,²⁴ NEVPT2,^{25,26} (X)MCQDPT2,²⁷ and (SA-)DSRG-MRPT2.^{28,29} Another type of multiconfigurational approach, such as multi-configuration pair-density functional theory (MC-PDFT),³⁰ has analytic derivative theories³¹ as well.

It is important to consider interactions between states during MECI searches. State interactions are important to describe the avoided crossing of, for example, LiF with MRPTs;^{12,32,33} the lack of them leads to a notorious double-crossing PES. Regarding CASPT2, the multi-state CASPT2 (MS-CASPT2)³² can treat state interactions and has been used in geometry optimization (primarily via numerical differentiation).^{34,35} However, MS-CASPT2 has a severe non-invariance problem and is unsuitable for MECI searches. Certain studies reported that PESs around MECIs at the MS-CASPT2 level have unphysical behaviors.^{36–38} One of the sources of non-invariance can be solved using the modification technique by Shiozaki³⁹ based on the idea of the XMCQDPT2 by Granovsky,¹¹ and the resulting extended MS-CASPT2 (XMS-CASPT2) has been used in locating MECIs and non-adiabatic MD simulations in several occasions.^{26,37,40,41}

Although XMS-CASPT2 smoothens the rough PES of MS-CASPT2 at CIs, the use of the state-averaged Fock operator *may* in general degrade the accuracy of MS-CASPT2 with a state-specific operator elsewhere. Recently, an alternative approach named extended dynamically weighted CASPT2 (XDW-CASPT2)⁴² has been suggested as a bridge between MS-CASPT2 and XMS-CASPT2. XDW-CASPT2 employs a dynamically weighted density matrix in the Fock operator, in association with reference states that are rotated to remove the interaction between them, following the same approach used in XMS-CASPT2. As a result, one can select a completely state-specific operator that interpolates between state-specific and state-average regimes. The idea of XDW-CASPT2 is similar to the preceding dynamically weighted DSRG-MRPT2⁴³ and SCF⁴⁴ methods. A special case of this variant,

which shares a similar idea to that in Ref. 45, is referred to as rotated multistate CASPT2 (RMS-CASPT2)³⁸ and only performs an initial rotation of the states, thereby avoiding the dynamical weighting scheme. The applicability of XDW-CASPT2 and RMS-CASPT2 has been discussed in two recent studies,^{38,42} demonstrating that PESs near MECIs are smooth. Similar to the standard geometry optimization, MECIs can be located with analytic derivatives. Recent studies⁴⁶ on CASPT2 have enabled the exploration of CI regions using analytic derivative coupling vectors³⁷ at the XMS-CASPT2 level of theory. Non-adiabatic MD simulations can be performed as well.^{40,47} GVVPT2¹⁶ may also be used to explore CIs, and a formulation for derivative coupling vectors exists;⁴⁸ however, an actual implementation has not been achieved so far. NEVPT2,^{49,50} MCQDPT2,^{27,51} and SA-DSRG-MRPT2²⁹ can also be used for MECI search, while MC-PDFT has been used for studying intersystem crossings.⁵²

Although OpenMolcas^{53,54} is the program package with the original implementation of CASPT2, its analytic gradient has not been developed for 30 years. Recently, one of the authors has succeeded in this task for the single-state version of CASPT2.²³ However, the code for the multistate extension is not yet available. Furthermore, it would be interesting to see how the new XDW-CASPT2 and RMS-CASPT2 approaches perform in MECI searches. In this study, we outline the analytic derivative theory for (X)MS-, XDW-, and RMS-CASPT2, as well as their RASPT2 extension.^{55,56} The developed method is used to locate minimum energy structures and MECIs of typical systems, and the smoothness of PESs around MECIs, as well as their dependence on the different flavors of CASPT2/RASPT2, are discussed to evaluate the applicability of these methods.

2 Methodology

In this section, the following indices are used, unless otherwise stated:

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

- Atomic orbitals (AOs): μ, ν
- CSFs: I
- States: $\alpha, \beta, \gamma, \delta$
- Internally contracted configurations: ϕ, χ

For the remainder of this manuscript, shorthand acronyms will be used as defined below, unless otherwise stated:

- MS-CASPT2: MS
- XMS-CASPT2: XMS
- XDW-CASPT2: XDW
- RMS-CASPT2: RMS

2.1 MS-CASPT2 and XMS-CASPT2

In perturbation theory, the electronic Hamiltonian \hat{H} is partitioned into the zeroth-order $\hat{H}^{(0)}$ and perturbation potential \hat{V} :

$$\hat{H} = \hat{H}^{(0)} + \hat{V} . \quad (1)$$

Depending on the definition of the zeroth-order part, one can define a series of (MR)PT methods. In quasi-degenerate perturbation theory, the zeroth-order Hamiltonian is typically defined by

$$\hat{H}^{(0)} = \hat{P}\hat{F}\hat{P} + \hat{Q}\hat{F}\hat{Q} , \quad (2)$$

where $\hat{P} = \sum_{\gamma \in \mathcal{P}} |\Psi_{\gamma}^{(0)}\rangle\langle\Psi_{\gamma}^{(0)}|$ is the projector onto the reference space, and $\hat{Q} = 1 - \hat{P}$ is the projector onto the complementary space. The space of the reference states, which are typically obtained by solving state-averaged CASSCF (SA-CASSCF) equations, is represented

by the symbol \mathcal{P} . The Fock operator \hat{F} can be either state-specific or state-averaged, and one can define several CASPT2 variants depending on the operator choice, as outlined below.

In MS,³² the Fock operator is state-specific, and the zeroth-order Hamiltonian for state α is defined by

$$\begin{aligned} \hat{H}_\alpha^{(0)} &= \sum_{\gamma \in \mathcal{P}} |\Psi_\gamma^{(0)}\rangle \langle \Psi_\gamma^{(0)} | \hat{F}^\alpha | \Psi_\gamma^{(0)}\rangle \langle \Psi_\gamma^{(0)} | \\ &+ \sum_{k \in \mathcal{P}^\perp} |\Psi_k^{(0)}\rangle \langle \Psi_k^{(0)} | \hat{F}^\alpha | \Psi_k^{(0)}\rangle \langle \Psi_k^{(0)} | \\ &+ \hat{Q}_{\text{SD}}^\alpha \hat{F}^\alpha \hat{Q}_{\text{SD}}^\alpha + \hat{Q}_{\text{TQ}\dots}^\alpha \hat{F}^\alpha \hat{Q}_{\text{SD}\dots}^\alpha, \end{aligned} \quad (3)$$

where \mathcal{P}^\perp includes all CAS states orthogonal to \mathcal{P} , $|\Psi_\gamma^{(0)}\rangle$ is the zeroth-order (reference) wavefunction, and $\hat{Q}_{\text{SD}}^\alpha$ and $\hat{Q}_{\text{TQ}\dots}^\alpha$ project onto the first-order and higher-order interacting spaces, respectively. The Fock operator $\hat{F}^\alpha = \sum_{pq} f_{pq}^\alpha \hat{E}_{pq}$, along with its element f_{pq}^α , is state-specific and defined by

$$\begin{aligned} f_{pq}^\alpha &= h_{pq} + \sum_{rs} \left((pq|rs) - \frac{1}{2}(pr|qs) \right) D_{rs}^\alpha \\ &= h_{pq} + g_{pq}(\mathbf{D}^\alpha). \end{aligned} \quad (4)$$

h_{pq} and $(pq|rs)$ are the one-electron and electron repulsion integrals, respectively, and $g_{pq}(\mathbf{D})$ is the multiplication of Coulomb and exchange integrals with \mathbf{D} . The elements of the one-electron (reduced) density matrix D_{rs}^α are calculated by $D_{rs}^\alpha = \langle \Psi_\alpha^{(0)} | \hat{E}_{rs} | \Psi_\alpha^{(0)} \rangle$. \hat{E}_{pq} is the one-electron spin-averaged excitation operator.

As aforementioned, MS has a severe problem when interactions between reference states are not negligible. This is due to the non-invariance of the theory with respect to unitary rotations among the reference states, whose major cause is attributed to neglecting the off-diagonal elements of the Fock operator $\langle \Psi_\alpha^{(0)} | \hat{F}^\alpha | \Psi_\beta^{(0)} \rangle$ for $\alpha \neq \beta \in \mathcal{P}$ in Eq. (3).¹¹ These elements are no longer trivial when interactions between different states are significant. This

is known as the diagonal approximation (which is different from the “diagonal approximation” in CASPT2-D⁹), and is obvious from the structure of the first term in Eq. (2) and \hat{P} . In XMS,³⁹ the diagonal approximation in \hat{H}_0 is removed:

$$\begin{aligned}\hat{H}^{(0)} &= \sum_{\gamma, \delta \in \mathcal{P}} |\Psi_\gamma^{(0)}\rangle \langle \Psi_\gamma^{(0)}| \hat{F}^{\text{SA}} |\Psi_\delta^{(0)}\rangle \langle \Psi_\delta^{(0)}| \\ &+ \sum_{k \in \mathcal{P}^\perp} |\Psi_k^{(0)}\rangle \langle \Psi_k^{(0)}| \hat{F}^{\text{SA}} |\Psi_k^{(0)}\rangle \langle \Psi_k^{(0)}| \\ &+ \hat{Q}_{\text{SD}} \hat{F}^{\text{SA}} \hat{Q}_{\text{SD}} + \hat{Q}_{\text{TQ}} \dots \hat{F}^{\text{SA}} \hat{Q}_{\text{SD}} \dots\end{aligned}\quad (5)$$

The operator is no longer state-specific (but the \hat{Q} operators are state-specific in the actual implementation of OpenMolcas). This form of the zeroth-order Hamiltonian is inconvenient for its direct use in MRPTs, and so the reference states are rotated to satisfy

$$\langle \tilde{\Psi}_\alpha^{(0)} | \hat{F}^{\text{SA}} | \tilde{\Psi}_\beta^{(0)} \rangle = 0 \quad (6)$$

for $\alpha \neq \beta \in \mathcal{P}$. The rotated reference states $|\tilde{\Psi}_\alpha^{(0)}\rangle$ (tilde denotes rotated quantities as per the XMS procedure) is simply obtained by diagonalizing $\langle \Psi_\alpha^{(0)} | \hat{F}^{\text{SA}} | \Psi_\beta^{(0)} \rangle$ and by transforming the reference state with a unitary rotation:

$$|\tilde{\Psi}_\alpha^{(0)}\rangle = \sum_{\beta \in \mathcal{P}} |\Psi_\beta^{(0)}\rangle U_{\beta\alpha}. \quad (7)$$

This procedure corresponds to applying the orthonormal transformation to the reference configuration interaction coefficients, $c_{I,\beta}$:

$$\tilde{c}_{I,\alpha} = \sum_{\beta \in \mathcal{P}} c_{I,\beta} U_{\beta\alpha}. \quad (8)$$

In MS, the unitary matrix $U_{\beta\alpha}$ is the identity matrix. Using the rotated reference states $|\tilde{\Psi}_\alpha^{(0)}\rangle$, the zeroth-order Hamiltonian for XMS (Eq. (5)) is written as an expression similar

to that for MS (Eq. (3)):

$$\begin{aligned}
\hat{H}^{(0)} &= \sum_{\gamma \in \mathcal{P}} |\tilde{\Psi}_\gamma^{(0)}\rangle \langle \tilde{\Psi}_\gamma^{(0)}| \hat{F}^{\text{SA}} |\tilde{\Psi}_\gamma^{(0)}\rangle \langle \tilde{\Psi}_\gamma^{(0)}| \\
&+ \sum_{k \in \mathcal{P}^\perp} |\Psi_k^{(0)}\rangle \langle \Psi_k^{(0)}| \hat{F}^{\text{SA}} |\Psi_k^{(0)}\rangle \langle \Psi_k^{(0)}| \\
&+ \hat{Q}_{\text{SD}} \hat{F}^{\text{SA}} \hat{Q}_{\text{SD}} + \hat{Q}_{\text{TQ}\dots} \hat{F}^{\text{SA}} \hat{Q}_{\text{SD}\dots}
\end{aligned} \tag{9}$$

Unlike MS, the elements of the Fock operator f_{pq}^{SA} (and thus the zeroth-order Hamiltonian) are no longer state-specific and are instead defined using the (equally weighted) state-averaged density matrix:

$$f_{pq}^{\text{SA}} = h_{pq} + g_{pq} (\mathbf{D}^{\text{SA}}) , \tag{10}$$

with

$$\mathbf{D}^{\text{SA}} = \frac{1}{N_{\text{state}}} \sum_{\alpha \in \mathcal{P}} \tilde{\mathbf{D}}^\alpha , \tag{11}$$

where N_{state} is the number of states averaged in the reference self-consistent field (SCF) calculation, assuming that the interaction between all states is considered in the multistate perturbation calculation, and $\tilde{\mathbf{D}}^\alpha$ is the density matrix analog calculated using the rotated state: $\tilde{D}_{rs}^\alpha = \langle \tilde{\Psi}_\alpha^{(0)} | \hat{E}_{rs} | \tilde{\Psi}_\alpha^{(0)} \rangle$.

The first-order wavefunction is written as follows:

$$|\Psi_\alpha^{(1)}\rangle = \sum_{pqrs} \hat{E}_{pq} \hat{E}_{rs} T_{pqrs}^{\alpha\alpha} |\tilde{\Psi}_\alpha^{(0)}\rangle \tag{12}$$

where $\hat{E}_{pq} \hat{E}_{rs}$ is the two-electron excitation operator, and $T_{pqrs}^{\alpha\alpha}$ is the corresponding excitation amplitude. Due to the different orbital subspaces, the definition of $\hat{E}_{pq} \hat{E}_{rs}$ comprises eight distinct excitation classes, see Refs. 9 and 10. The excitation amplitudes are determined by iteratively solving the following linear equation:

$$\langle \Phi_{pqrs} | \hat{H} | \tilde{\Psi}_\alpha^{(0)} \rangle + \langle \Phi_{pqrs} | \hat{H}^{(0)} - E_\alpha^{(0)} + E_{\text{shift}} | \Psi_\alpha^{(1)} \rangle = 0 , \tag{13}$$

where $|\Phi_{pqrs}\rangle$ is a doubly excited configuration, and E_{shift} is the real⁵⁷ or imaginary⁵⁸ level shift parameter to avoid the intruder state problem:

$$E_{\text{shift}} = E_{\text{shift}}^{\text{real}} + \frac{\left(E_{\text{shift}}^{\text{imaginary}}\right)^2}{\hat{H}_{\text{D}}^{(0)} - E^{(0)}} , \quad (14)$$

where $E_{\text{shift}}^{\text{real}}$ and $E_{\text{shift}}^{\text{imaginary}}$ are provided as parameters and cannot be non-zero simultaneously, and $\hat{H}_{\text{D}}^{(0)}$ is the diagonal part of $\hat{H}^{(0)}$. The real and imaginary level shift techniques modify the denominator in the second-order energy expression $1/\Delta_{pqrs}$ by $1/(\Delta_{pqrs} + \varepsilon)$ and $1/(\Delta_{pqrs} + i\varepsilon)$, respectively, where $\Delta_{pqrs} = \langle \Phi_{pqrs} | \hat{H}^{(0)} - E^{(0)} | \Phi_{pqrs} \rangle$ and ε is a small parameter. Although the complex number appears in the denominator of the latter expression, we extract the real part of the amplitude so that the energy remains real:

$$\text{Re} \left(\frac{1}{\Delta_{pqrs} + i\varepsilon} \right) = \frac{1}{\Delta_{pqrs} + \frac{\varepsilon^2}{\Delta_{pqrs}}} \quad (15)$$

We thus do not need complex arithmetic. Note that the imaginary level shift is equivalent to the intruder state avoidance technique in (X)MCQDPT2.⁵⁹

After solving the linear equation, we construct the (symmetrized) effective Hamiltonian

$$\bar{H}_{\alpha\beta}^{\text{eff}} = \langle \tilde{\Psi}_{\alpha}^{(0)} | \hat{H} | \tilde{\Psi}_{\beta}^{(0)} \rangle + \frac{1}{2} \left(H_{\alpha\beta}^{(2)} + H_{\beta\alpha}^{(2)} \right) , \quad (16)$$

where

$$H_{\alpha\beta}^{(2)} = \langle \tilde{\Psi}_{\alpha}^{(0)} | \hat{H} | \Psi_{\beta}^{(1)} \rangle . \quad (17)$$

The (symmetrized) effective Hamiltonian is then diagonalized to obtain the (X)MS energy:

$$E_{\alpha}^{\text{CASPT2}} = \sum_{\gamma\delta \in \mathcal{P}} R_{\gamma\alpha} \bar{H}_{\gamma\delta}^{\text{eff}} R_{\delta\alpha} \quad (18)$$

Note that the implementation of XMS in OpenMolcas is not completely invariant even

though the original theory is. This is because the internal contraction scheme used is the so-called single-state single-reference (SS-SR) one, in which the zeroth-order Hamiltonian (specifically, \hat{Q}_{SD} and \hat{Q}_{TQ} in Eq. (5)) and the first-order wavefunction are state-specific and parameterized according to Eq. (12). In contrast, the multistate MR (MS-MR) scheme, used in the original XMS method,³⁹ defines the first-order wavefunction as

$$|\Psi_{\alpha}^{(1)}\rangle = \sum_{pqrs} \sum_{\beta \in \mathcal{P}} \hat{E}_{pq} \hat{E}_{rs} T_{pqrs}^{\alpha\beta} |\tilde{\Psi}_{\beta}^{(0)}\rangle . \quad (19)$$

The effect of non-invariance due to the different internal contraction schemes has been reported in Ref. 49 in the context of QD-NEVPT2. Although our developed algorithm is not completely invariant, XMS using the SS-SR internal contraction is very robust,^{26,37,49} and characterizing MECIs with it is generally stable.

2.2 XDW-CASPT2 and RMS-CASPT2

Although XMS can produce smooth PESs, using the state-averaged density matrix in $\hat{H}^{(0)}$ may in general decrease the accuracy compared to MS. The idea of interpolating between them has motivated the development of XDW⁴² and RMS.³⁸

First, similarly to XMS, the Fock operator is diagonalized to produce a set of rotated reference states (Eq. (6)). In this transformation, the (equally) state-averaged Fock operator is used; therefore, it is totally equivalent to XMS. In XDW, the density matrix used in the construction of the Fock operator for the perturbed state α is averaged using dynamical weights ω_{α}^{β} as follows:

$$\bar{\mathbf{D}}^{\alpha} = \sum_{\beta \in \mathcal{P}} \omega_{\alpha}^{\beta} \tilde{\mathbf{D}}^{\beta} , \quad (20)$$

where each weight ω_{α}^{β} is defined using the following Boltzmann-like function

$$\omega_{\alpha}^{\beta} = \frac{e^{-\zeta(\Delta_{\alpha\beta})^2}}{\sum_{\gamma \in \mathcal{P}} e^{-\zeta(\Delta_{\alpha\gamma})^2}} . \quad (21)$$

The parameter ζ regulates the sharpness of the transition between mixed-density and state-specific regimes.⁴² Many expressions for $\Delta_{\alpha\beta}$ have been proposed.^{38,42} However, in this study, the original definition has been used:

$$\Delta_{\alpha\beta} = \left\langle \tilde{\Psi}_{\alpha}^{(0)} \left| \hat{H} \right| \tilde{\Psi}_{\alpha}^{(0)} \right\rangle - \left\langle \tilde{\Psi}_{\beta}^{(0)} \left| \hat{H} \right| \tilde{\Psi}_{\beta}^{(0)} \right\rangle \quad (22)$$

If the energies of the two reference states α and β are similar, the $\Delta_{\alpha\beta}$ value is small and the exponent term becomes large, and therefore the density matrix of the state β significantly contributes to that of the perturbed state α . The element of the Fock operator, \bar{f}_{pq}^{α} , can be obtained by substituting \mathbf{D}^{α} with $\bar{\mathbf{D}}^{\alpha}$ in Eq. (4):

$$\bar{f}_{pq}^{\alpha} = h_{pq} + g_{pq}(\bar{\mathbf{D}}^{\alpha}) . \quad (23)$$

In the limit $\zeta \rightarrow 0$, the exponent term does not contribute to the energy weighting, and each state contributes equally; therefore the limit corresponds to the XMS method. On the other hand, in the limit of $\zeta \rightarrow \infty$, the exponent term contributes only when $\alpha = \beta$, and the density matrix is state-specific; this corresponds to the RMS method.

Although explicitly writing the zeroth-order Hamiltonian for XDW and RMS is not straightforward, if the following approximation is valid:

$$\left\langle \tilde{\Psi}_{\alpha}^{(0)} \left| \hat{F}^{\gamma} \right| \tilde{\Psi}_{\beta}^{(0)} \right\rangle \approx 0 , \quad (24)$$

the zeroth-order Hamiltonian for XDW and RMS can be written as a state-dependent form:

$$\begin{aligned} \hat{H}_{\alpha}^{(0)} &\approx \sum_{\gamma \in \mathcal{P}} |\tilde{\Psi}_{\gamma}^{(0)}\rangle \langle \tilde{\Psi}_{\gamma}^{(0)} | \hat{F}^{\alpha} | \tilde{\Psi}_{\gamma}^{(0)}\rangle \langle \tilde{\Psi}_{\gamma}^{(0)} | \\ &+ \sum_{k \in \mathcal{P}^{\perp}} |\Psi_k^{(0)}\rangle \langle \Psi_k^{(0)} | \hat{F}^{\alpha} | \Psi_k^{(0)}\rangle \langle \Psi_k^{(0)} | \\ &+ \hat{Q}_{\text{SD}}^{\alpha} \hat{F}^{\alpha} \hat{Q}_{\text{SD}}^{\alpha} + \hat{Q}_{\text{TQ}\dots}^{\alpha} \hat{F}^{\alpha} \hat{Q}_{\text{SD}\dots}^{\alpha} . \end{aligned} \quad (25)$$

For RMS, the structure of the zeroth-order Hamiltonian is similar to that of MS (Eq. 3). However, because the reference states are rotated, it is expected that the state-specific Fock operator couplings between the corresponding zeroth-order wavefunctions are approximately zero:

$$\langle \tilde{\Psi}_\alpha^{(0)} | \hat{F}^\gamma | \tilde{\Psi}_\beta^{(0)} \rangle \approx 0 \quad (26)$$

for $\alpha \neq \beta \in \mathcal{P}$. It is the use of these rotated states that distinguishes RMS from MS.

Note that RMS is similar to Park’s SS-SR-MS-CASPT2.⁴⁵ The difference between the two methods is in how the reference states are rotated. Rotated states are generated to satisfy Eq. (6) in RMS, whereas the following condition is used in Park’s SS-SR-MS-CASPT2:

$$\left\langle \tilde{\Psi}_\alpha^{(0)} \left| \frac{1}{2} (\hat{F}^\alpha + \hat{F}^\beta) \right| \tilde{\Psi}_\beta^{(0)} \right\rangle = 0 \quad (27)$$

for $\alpha \neq \beta \in \mathcal{P}$. The Fock matrix is thus diagonalized using the state-specific Fock operator inside the bracket. Both RMS and Park’s SS-SR-MS-CASPT2 are approximately invariant. Ref. 45 demonstrated that SS-SR-MS-CASPT2 does not suffer from the shortcomings of the original MS approach and can locate MECIs. The performance of RMS has been investigated in Refs. 38 and 42, and PESs of RMS around MECIs of CASSCF are smooth.

2.3 Analytic Derivatives of (X)MS-, XDW-, and RMS-CASPT2

We assume the zeroth-order Hamiltonian of XDW because other variants are derived considering the abovementioned limits.

2.3.1 Nuclear energy gradient

Similar to previous analytic derivative theories for MRPT methods, we use the Lagrangian technique.⁶⁰ The total Lagrangian $\mathcal{L}^{\text{CASPT2}}$ is defined as the sum of the SA-CASSCF and PT2 Lagrangians, $\mathcal{L}^{\text{CASSCF}}$ and \mathcal{L}^{PT2} : $\mathcal{L}^{\text{CASPT2}} = \mathcal{L}^{\text{CASSCF}} + \mathcal{L}^{\text{PT2}}$. The SA-CASSCF part of the Lagrangian is similar to that of the previous study²³ and comprises constraint conditions

imposed when solving the SA-CASSCF equation:

$$\begin{aligned} \mathcal{L}_{\alpha\beta}^{\text{CASSCF}} = & \frac{1}{2} \text{Tr} [\mathbf{Z}^{\alpha\beta} (\mathbf{A} - \mathbf{A}^\dagger)] - \frac{1}{2} \text{Tr} [\mathbf{X}^{\alpha\beta} (\mathbf{S} - \mathbf{I})] \\ & + \frac{1}{N_{\text{state}}} \sum_{\gamma \in \mathcal{P}} \left[\sum_I z_{I,\gamma}^{\alpha\beta} \langle I | \hat{H} - E_\gamma^{\text{ref}} | \Psi_\gamma^{(0)} \rangle - \frac{1}{2} x_\gamma^{\alpha\beta} (\langle \Psi_\gamma^{(0)} | \Psi_\gamma^{(0)} \rangle - 1) \right], \end{aligned} \quad (28)$$

where the first term is the generalized Brillouin condition with the orbital gradient $\mathbf{A} - \mathbf{A}^\dagger$,^{20,39} the second term is the requirement for MO orthonormalization with the overlap matrix \mathbf{S} in the MO basis, and the third term is the CI condition. In the third term, N_{state} is the number of (equally) averaged states and E_γ^{ref} is the CASSCF energy for state $\gamma \in \mathcal{P}$. $\mathbf{Z}^{\alpha\beta}$, $\mathbf{X}^{\alpha\beta}$, $\mathbf{z}^{\alpha\beta}$, and $\mathbf{x}^{\alpha\beta}$ are Lagrangian multipliers which are determined by solving the Z-vector equation described later (Eq. (35)).

The PT2 part of the Lagrangian can be written as follows:

$$\begin{aligned} \mathcal{L}_{\alpha\beta}^{\text{PT2}} = & \sum_{\gamma\delta \in \mathcal{P}} R_{\gamma\alpha} \bar{H}_{\gamma\delta}^{\text{eff}} R_{\delta\beta} + \sum_{\gamma \in \mathcal{P}} \sum_c \left(\sum_{\phi \neq \chi} \zeta_{\phi\chi}^{c,\gamma,\alpha\beta} \Lambda_{\phi\chi}^{\text{B},c,\gamma} - \sum_{\phi\chi} \xi_{\phi\chi}^{c,\gamma,\alpha\beta} (\Lambda_{\phi\chi}^{\text{S},c,\gamma} - \mathbf{I}) \right) \\ & + \sum_{\gamma \in \mathcal{P}} \sum_{pqrs} \lambda_{pqrs,\gamma}^{\alpha\beta} \left(\langle \Phi_{pqrs} | \hat{H} | \tilde{\Psi}_\gamma^{(0)} \rangle + \langle \Phi_{pqrs} | \hat{H}_\gamma^{(0)} - E_\gamma^{(0)} + E_{\text{shift}} | \Psi_\gamma^{(1)} \rangle \right) \\ & + \sum_{\gamma \in \mathcal{P}} \sum_p \sum_q^{\text{core inactive}} Z_{pq}^{\text{core},\gamma,\alpha\beta} \bar{f}_{pq}^\gamma + \sum_{\gamma \neq \delta \in \mathcal{P}} w_{\gamma\delta}^{\alpha\beta} \langle \tilde{\Psi}_\gamma^{(0)} | \hat{F}^{\text{SA}} | \tilde{\Psi}_\delta^{(0)} \rangle \end{aligned} \quad (29)$$

The first term is the CASPT2 energy ($\alpha = \beta$; Eq. (18)) or the coupling between states ($\alpha \neq \beta$). The second term results from the diagonalization of the zeroth-order Hamiltonian and orthonormalization of the overlap matrix in the active space. In the summation, c corresponds to the eight excitation classes (Eqs. (1a), (1b), ..., and (1h) in Ref. 9). $\Lambda_{\phi\chi}^{\text{B},c,\gamma}$ and $\Lambda_{\phi\chi}^{\text{S},c,\gamma}$ correspond to Eqs. (19) and (26) in Ref. 9. The additional Lagrangian multipliers $\zeta_{\phi\chi}^{c,\gamma,\alpha\beta}$ (different from the parameter for XDW) and $\xi_{\phi\chi}^{c,\gamma,\alpha\beta}$ are introduced. They can be straightforwardly determined by considering the partial derivative of the above Lagrangian with respect to the internal contraction coefficients after solving the λ -equation introduced below as Eq. (30). The third term in Eq. (29) corresponds to the variational condition

of the CASPT2 energy (Eq. (13)), the fourth term for the frozen core approximation is the semi-canonical condition of MOs, and the fifth term is the rotation of reference states (Eq. (6)).

The excitation amplitude $T_{pqrs}^{\alpha\alpha}$ is optimized for each diagonal element of $\overline{\mathbf{H}}^{\text{eff}}$, whereas off-diagonal elements of $\overline{\mathbf{H}}^{\text{eff}}$ contribute to the PT2 Lagrangian, and therefore the PT2 Lagrangian is not variational with respect to amplitude changes. Consequently, we need to solve the λ -equation for each γ :

$$\frac{\partial \mathcal{L}_{\alpha\beta}^{\text{PT2}}}{\partial T_{pqrs}^{\gamma\gamma}} = 0 \quad (30)$$

The λ -equation is similar in structure to the amplitude equation (Eq. (13)), and therefore it can be solved in a similar fashion. Note that the λ -equation is solved with the level shift term E_{shift} , being consistent with Eq. (13). However, unlike the single-state case, it has to be solved even without the level shift E_{shift} for energy calculations, because the MS energy (Eq. (18)) is no longer variationally minimized with respect to the excitation amplitude. Only when the coupling between different excitation classes are neglected, known as CASPT2-D, it can be done in a non-iterative manner. Once $\lambda_{pqrs,\gamma}^{\alpha\beta}$ is determined, first- and second-order correlated density matrices²⁰ can be constructed and used in computing $Y_{pq}^{\alpha\beta}$ and $\tilde{y}_{I,\gamma}^{\alpha\beta}$ defined later.

In XDW, the weight of the density matrix ω_{α}^{β} (Eq. (21)) has to be differentiated as well. Using the second-order one-particle correlated density matrix $\mathbf{D}^{(2)}$,²⁰ the contribution of the derivative through $\langle \tilde{\Psi}_{\alpha}^{(0)} | \hat{H} | \tilde{\Psi}_{\alpha}^{(0)} \rangle$ can be written as follows:

$$\check{D}_{\alpha} = -2\zeta \sum_{\beta \in \mathcal{P}} \left(\omega_{\alpha}^{\beta} \left(\check{\omega}_{\alpha}^{\beta} - \sum_{\gamma \in \mathcal{P}} \omega_{\alpha}^{\gamma} \check{\omega}_{\alpha}^{\gamma} \right) + \omega_{\beta}^{\alpha} \left(\check{\omega}_{\beta}^{\alpha} - \sum_{\gamma \in \mathcal{P}} \omega_{\gamma}^{\alpha} \check{\omega}_{\gamma}^{\alpha} \right) \right) \Delta_{\alpha\beta} \quad (31)$$

where $\check{\omega}_{\alpha}^{\beta} = \sum_{pq \in \text{act}} g_{pq}(\mathbf{D}^{(2),\alpha}) \tilde{D}_{pq}^{\beta}$. \check{D}_{α} is always zero for (X)MS and RMS because ω_{α}^{β} for these theories is either $1/N_{\text{state}}$ or $\delta_{\alpha\beta}$ (Kronecker delta). ζ is a parameter, and therefore its derivative is always zero.

The so-called right-hand side vectors are obtained by computing the partial derivatives

of the total Lagrangian with respect to the wavefunction parameters:

$$Y_{pq}^{\alpha\beta} := \frac{\partial \mathcal{L}_{\alpha\beta}^{\text{PT2}}}{\partial \kappa_{pq}} \quad (32)$$

$$\tilde{y}_{I,\gamma}^{\alpha\beta} := \frac{\partial \mathcal{L}_{\alpha\beta}^{\text{PT2}}}{\partial \tilde{c}_{I,\gamma}}, \quad (33)$$

where κ_{pq} is the orbital rotation parameter. These partial derivatives can be computed similarly to the previously developed scheme in Ref. 23. The multiplier $w_{\gamma\delta}^{\alpha\beta}$ is specific to CASPT2 variants with rotated reference states and is always zero for MS. It can be simply determined by:⁶¹

$$w_{\gamma\delta}^{\alpha\beta} = \frac{1}{2} \frac{1}{E_{\delta}^{(0)} - E_{\gamma}^{(0)}} \sum_I \left(\tilde{c}_{I,\gamma} \tilde{y}_{I,\delta}^{\alpha\beta} - \tilde{y}_{I,\gamma}^{\alpha\beta} \tilde{c}_{I,\delta} \right). \quad (34)$$

The numerical stability of $w_{\delta\gamma}^{\alpha\beta}$ is discussed in Ref. 61.

The Lagrange multipliers specific to CASSCF are determined by solving the Z-vector equation:

$$\frac{\partial \mathcal{L}_{\alpha\beta}^{\text{CASPT2}}}{\partial \kappa_{pq}} = \frac{\partial \mathcal{L}_{\alpha\beta}^{\text{CASPT2}}}{\partial c_{I,\gamma}} = 0 \quad (35)$$

for all independent rotations and $\gamma \in \mathcal{P}$, yielding $\mathbf{Z}^{\alpha\beta}$, $\mathbf{X}^{\alpha\beta}$, $\mathbf{z}^{\alpha\beta}$, and $\mathbf{x}^{\alpha\beta}$. Once the Z-vector equation is solved, the derivative of the CASPT2 energy with respect to the nuclear displacement x is evaluated as a partial derivative of the Lagrangian:

$$(\mathbf{g}_{\alpha})^x := \frac{dE_{\alpha}^{\text{CASPT2}}}{dx} = \frac{\partial \mathcal{L}_{\alpha\alpha}^{\text{CASPT2}}}{\partial x}. \quad (36)$$

It would be informative to compare the computational cost for the four MS-CASPT2 variants in the nuclear gradient calculation. There are three sources of difference. First, we need the derivative of the weight (Eq. (31)) for XDW, but this step is negligible, so the computational cost for XDW and RMS is almost identical. Second, XMS, XDW, and RMS require additional evaluations that arise from the rotation of the reference states. This evaluation is not negligible, but it is needed only once, so the cost is relatively small. Third,

the Fock matrix at the MS, XDW, and RMS levels is defined with a state-dependent density matrix, while the one at the SCF level with the state-averaged density matrix. Due to this difference, the derivative of the Fock matrix at the PT2 level cannot be computed in a similar fashion to the one at the SCF level. To be specific, the derivative of the $g_{pq}(\mathbf{D})$ for the state-dependent Fock matrix is split following the equality:

$$\sum_{pq} \bar{\mathbf{D}}_{pq}^{\alpha} \frac{\partial g_{pq}(\mathbf{D}^{(2),\alpha})}{\partial x} = \sum_{pq} (\bar{\mathbf{D}}_{pq}^{\alpha} - \mathbf{D}_{pq}^{\text{SA}}) \frac{\partial g_{pq}(\mathbf{D}^{(2),\alpha})}{\partial x} + \sum_{pq} \mathbf{D}_{pq}^{\text{SA}} \frac{\partial g_{pq}(\mathbf{D}^{(2),\alpha})}{\partial x}, \quad (37)$$

where the first term in the right hand side is additionally evaluated for MS, XDW, and RMS for each state α . Fortunately, we just need to back transform the density matrices, and the additional computational cost is less than the cost for the major part of Eq. (32). We do not need this evaluation for XMS, because the Fock matrix is defined with the state-averaged density. We estimate the following order of the computational cost: XMS < MS < RMS \approx XDW.

2.3.2 Derivative Coupling

The derivative coupling $\mathbf{d}_{\alpha\beta}$ for states α and β can be written as a sum of the so-called CI^{CI} $\mathbf{d}_{\alpha\beta}$ and CSF contributions^{CSF} $\mathbf{d}_{\alpha\beta}$:

$$\mathbf{d}_{\alpha\beta} = \text{CI}^{\text{CI}}\mathbf{d}_{\alpha\beta} + \text{CSF}^{\text{CSF}}\mathbf{d}_{\alpha\beta} = \frac{\mathbf{h}_{\alpha\beta}}{\Delta E_{\alpha\beta}} + \text{CSF}^{\text{CSF}}\mathbf{d}_{\alpha\beta}. \quad (38)$$

where $\mathbf{h}_{\alpha\beta}$ is the interstate coupling vector, and $\Delta E_{\alpha\beta} = E_{\beta}^{\text{CASPT2}} - E_{\alpha}^{\text{CASPT2}}$. The interstate coupling vector $\mathbf{h}_{\alpha\beta}$ is computed as the first-order derivative of the off-diagonal elements of the diagonalized effective Hamiltonian:^{62,63}

$$h_{\alpha\beta}^x = \sum_{\gamma\delta} R_{\gamma\alpha} \frac{d\bar{H}_{\gamma\delta}^{\text{eff}}}{dx} R_{\delta\beta} = \frac{\partial \mathcal{L}_{\alpha\beta}^{\text{CASPT2}}}{\partial x}. \quad (39)$$

Because the difference between $\mathbf{h}_{\alpha\beta}$ and the gradient ($\mathbf{g}_\alpha = \mathbf{h}_{\alpha\alpha}$) vectors is the element of the differentiated symmetrized effective Hamiltonian, the algorithm for computing $\mathbf{h}_{\alpha\beta}$ is almost equal to that for the gradient vector, and the algorithm in the previous section is used without additional complications.

The CSF contribution ${}^{\text{CSF}}\mathbf{d}_{\alpha\beta}$ can be computed as follows:

$${}^{\text{CSF}}d_{\alpha\beta}^x = \frac{1}{2} \sum_{\gamma\delta \in \mathcal{P}} (R_{\gamma\alpha}R_{\delta\beta} - R_{\delta\alpha}R_{\gamma\beta}) \left\langle \tilde{\Psi}_\gamma^{(0)} + \Psi_\gamma^{(1)} \left| \frac{d\tilde{\Psi}_\delta^{(0)}}{dx} \right. \right\rangle . \quad (40)$$

Here, we mostly followed the procedure outlined in Ref. 37, apart that the algorithm for computing the CSF term at the SA-CASSCF level is somewhat different. In OpenMolcas, the pseudo-density is contracted via the antisymmetric overlap derivative, i.e.,⁶⁴

$$S_{\mu\nu}^{x,-} = \frac{1}{2} \left(\left\langle \mu \left| \frac{\partial\nu}{\partial x} \right. \right\rangle - \left\langle \frac{\partial\mu}{\partial x} \left| \nu \right. \right\rangle \right) , \quad (41)$$

whereas in Ref. 37 it is contracted with the single overlap derivative:

$$\sigma_{\mu\nu}^x = \left\langle \mu \left| \frac{\partial\nu}{\partial x} \right. \right\rangle \quad (42)$$

Consequently, we had to modify one of the source terms in the Z-vector equation:

$$Y_{pq}^{\alpha\beta} = \frac{\partial \mathcal{L}_{\alpha\beta}^{\text{PT2}}}{\partial \kappa_{pq}} - \Delta E_{\alpha\beta} D_{pq}^{\alpha\beta} , \quad (43)$$

where

$$D_{pq}^{\alpha\beta} = \frac{1}{2} \sum_{\gamma\delta} (R_{\gamma\alpha}R_{\delta\beta} - R_{\delta\alpha}R_{\gamma\beta}) \left(\tilde{D}_{pq}^{\gamma\delta} + \langle \Psi_\gamma^{(1)} | \hat{E}_{pq} | \tilde{\Psi}_\delta^{(0)} \rangle \right) , \quad (44)$$

$\tilde{D}_{pq}^{\gamma\delta}$ is the one-particle transition reduced density matrix. The final contraction is

$$d_{\alpha\beta}^x = \frac{1}{\Delta E_{\alpha\beta}} \frac{\partial \mathcal{L}_{\alpha\beta}^{\text{CASPT2}}}{\partial x} + \frac{1}{2} \sum_{pq} (D_{pq}^{\alpha\beta} - D_{qp}^{\alpha\beta}) S_{pq}^{x,-} . \quad (45)$$

Because this difference is only caused by the way the contraction with the pseudo-density matrix is performed, the numerical value of $\mathbf{d}_{\alpha\beta}$ obtained with a development version of OpenMolcas seems to be equivalent to that of BAGEL.^{65,66}

Although the complete derivative coupling has been implemented, the interstate coupling vector $\mathbf{h}_{\alpha\beta}$ is used for MECI searches because $\mathbf{d}_{\alpha\beta}$ is singular and undefined at CI points ($\Delta E_{\alpha\beta} = 0$). Moreover, the CSF contribution $^{\text{CSF}}\mathbf{d}_{\alpha\beta}$ is translationally and rotationally non-invariant, so using the derivative coupling is unsuitable, or at least not helpful, to locate MECIs.

3 Computational Details

The theory discussed in the previous section, including the extension to RASPT2, has been implemented in a development version of OpenMolcas.^{53,54} In the following, we present and discuss the numerical results obtained with it. All calculations employed the frozen core approximation in association with the atomic compact Cholesky decomposition^{67,68} to generate an on-the-fly auxiliary basis set for the resolution-of-identity treatment of the electron repulsion integrals. The ionization potential–electron affinity (IPEA) shift⁶⁹ and symmetry constraints were not used, but an imaginary level shift parameter of 0.2i was selected for all calculations. For XDW, the parameter ζ is set to 50 for all calculations. The selection of the parameter affects the computed results, but the impact of the parameter has been examined in the past literature,^{38,42} and therefore our primary interest is not to examine this influence. In the reference SA-CASSCF calculation, the weight for each state is equal regardless of the number of averaged states.

With RASs, the active space is partitioned into three subspaces: RAS1, RAS2, and RAS3. The RAS2 space is treated as the CAS space, and full configuration interaction is performed. The numbers of holes and electrons, respectively, in RAS1 and RAS3 are limited by setting maximum values. The nomenclature for RAS is the same as in the previous

study:²³ $(ie,jo)/(ke,lo)/m$, where i and j are the numbers of electrons and orbitals in the entire active space (RAS1+RAS2+RAS3), respectively; k and l are those in the RAS2 space, respectively; and m is the maximum number of electrons excited from RAS1 or into RAS3. It is possible to independently select the maximum numbers for RAS1 and RAS3, but we have selected the same number. The design of the active spaces are described in detail in association with each case discussed below. The cc-pVDZ basis set⁷⁰ was used in all studies, except for the benzene and the dioxetanone systems, respectively, in which the cc-pVTZ⁷⁰ and the ANO-RCC-VTZP^{71,72} basis sets were used.

4 Results and Discussion

In this section, the quality of the implementation will be assessed in terms of the accuracy of the gradients for butadiene. This will be followed by three comparative studies of the different flavors of the CASPT2 and the standard RASPT2 methods. First, an investigation of the smoothness of the PESs at a CI in ethylene and butadiene is presented. Second, the quantitative and qualitative differences in the energetics of benzene are studied. Here the comparison is extended to include the partially contracted NEVPT2 (PC-NEVPT2) and XMCQDPT2 approaches. Third, the effect of the state averaging in ethylene is put forward. Finally, in the remaining two subsections, the methods will be applied to the dioxetanone molecule, where the last of the two is especially devoted to an analysis of the qualitative and quantitative differences between the CASPT2 and RASPT2 methods.

4.1 Accuracy of Implemented Gradient

The accuracy of the implemented gradient is confirmed by a comparison between the analytic and numerical gradients at the XMS-, XDW-, and RMS-CASPT2 and RASPT2 methods for planar butadiene, see Table 1. The numerical gradient was evaluated using the two-point stencil method.

Table 1: Root-Mean-Square (R-M-S)^a Differences between Analytic and Numerical Gradients (unit in a.u./bohr) for the Ground State at the XMS-, XDW-, and RMS-CASPT2 or RASPT2 Levels of Theory.

Active space	XMS	XDW	RMS
CAS(4e,4o)	3.07×10^{-7}	5.15×10^{-7}	6.13×10^{-7}
RAS(4e,4o)/(0e,0o)/2	8.95×10^{-7}	9.16×10^{-7}	9.68×10^{-7}
RAS(4e,4o)/(2e,2o)/1	1.95×10^{-6}	2.01×10^{-6}	2.04×10^{-6}

^a Forces perpendicular to the planar axis are excluded in the R-M-S difference

The difference with RAS is somewhat larger than that with CAS, reflecting the more complex description of the electronic structure provided by the former. This is consistent with previous single-state CASPT2 and RASPT2 results.²³ The deviation is similar to the single-state case and significantly smaller than the convergence criteria for geometry optimizations, and therefore the implementation can be confidently applied to geometry optimizations and MECI searches using any CASPT2 and RASPT2 variant.

4.2 Smoothness of the Potential Energy Surface

The smoothness of the PES described by MRPTs is important, in particular, for locating MECIs, but also in general for molecular structure optimizations. Here, MS-, XMS-, XDW-, and RMS-CASPT2 calculations were performed using a three-state averaged CASSCF(6e,4o) reference for the MECI of pyramidalized ethylene (structure **(b)** in Ref. 64). First, Figure 1 shows the PES of S_0 and S_1 around the MECI located using the individual CASPT2 methods. The $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ vectors (in atomic units) are obtained as a linear combination of the gradient difference and interstate coupling vectors following the definition given in Ref. 64, forming the branching plane. Although the PES of MS (Figure 1 A) is considerably affected by artifacts and the surfaces are significantly distorted, those of the other CASPT2 variants (Figure 1 B, C, and D) are similar and exhibit no apparent irregularities. The local topology of MECIs can be identified using two parameters \mathcal{P} and \mathcal{B} ,⁶⁴ which represent the asymmetry of the CI — the relative tilt and its direction at the MECI structure. As per these parameters, the MECI at XMS, XDW, and RMS is a peaked single-path CI. This topological character

agrees well with the plots shown in Figure 1. The MECI at MS is predicted to be peaked bifurcating; however, the PES is strongly distorted at this point, and so it is difficult to characterize it from Figure 1.

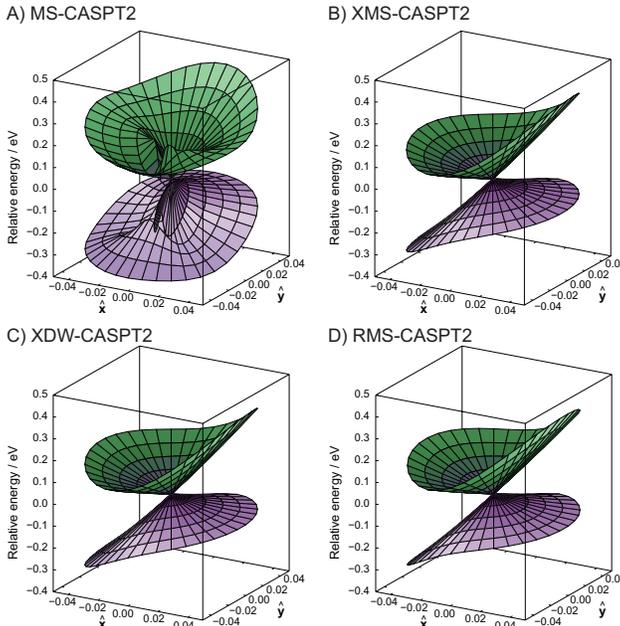


Figure 1: PESs at the MECI of pyramidalized ethylene with A) MS-, B) XMS-, C) XDW-, and D) RMS-CASPT2 methods.

Because the non-invariance character of MS, XDW, and RMS can be significant around the MECI structure obtained at the SA-CASSCF level of theory, it is important to verify the PES around that point. In Figure 2, the energy difference between the first two states (S_0 and S_1) is plotted as a function of displacement in the branching space around the MECI. Again, the PES of MS is unreliable. In particular, in the vicinity of this CI from SA-CASSCF ($\hat{x} = \hat{y} = 0.00$), it exhibits a strong discontinuity. On the contrary, the other CASPT2 variants do not show any severe problems, even though a few discontinuities appear for RMS.

XMS, XDW, and RMS are regulated using the ζ parameter in Eq. (21). At the limit $\zeta \rightarrow 0$, XDW corresponds to XMS, whereas at the limit $\zeta \rightarrow \infty$ to RMS. As $\zeta \rightarrow 0$, the density matrix used in the Fock operator (Eq. (10)) is averaged with equal weights for all

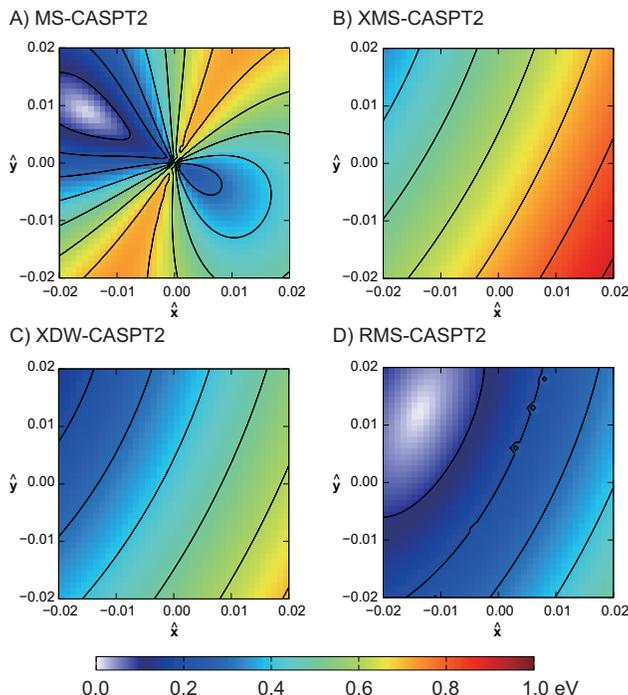


Figure 2: Energy difference between S_0 and S_1 for ethylene around the SA-CASSCF MECI with A) MS-, B) XMS-, C) XDW-, and D) RMS-CASPT2 methods.

states in the perturbation calculation. Consequently, the zeroth-order Hamiltonian (Eq. (5)) has zero off-diagonal elements in the reference space, which results in an extremely smooth PES as shown in Figure 2 (despite the slight non-invariance introduced by the SS-SR internal contraction scheme). As ζ increases, the density matrix and the Fock operator progressively gain a state-dependent character. Consequently, the $H^{(0)}$ off-diagonal coupling between the rotated states may become different from zero (see Eq. (24)), potentially making XDW and RMS less smooth than XMS. In this particular case, the PES described using XDW (Figure 2 C) is extremely smooth. On the other hand, at the $\zeta \rightarrow \infty$ limit, RMS has a few light irregularities in Figure 2 D. The distortion in this case is up to 4 kcal/mol, but RMS remains considerably smoother than MS despite the similar state-specific Fock operator.

Because of its artificial discontinuity, we will not use MS in the following calculations. In certain cases, it is possible to locate stationary points using MS, but its stability is generally significantly lower than that of other variants, in particular, for locating MECIs.

Nevertheless, the similar location of the minima in Figures 2 A and D suggests that RMS is a good alternative to MS, albeit providing considerably smoother PESs.

The smoothness of RASPT2 is investigated using three active spaces for a transoid MECI of butadiene (structure (**k**) in Ref. 64): CAS(4e,4o), RAS(4e,4o)/(0e,0o)/2, and RAS(4e,4o)/(2e,2o)/1. Plots of the branching space around XMS-CASPT2/RASPT2 MECIs and contour plots of the energy difference around the SA-CASSCF/RASSCF CIs are plotted in Figure 3. In all entries, the MECIs were located for each individual active space. Although there are small irregularities in a few instances (right panel in CAS(4e,4o) and left panel in RAS(4e,4o)/(2e,2o)/1), the PESs obtained using XMS-CASPT2 and XMS-RASPT2 are generally well behaved. These CIs are identified as a peaked single-path as per the \mathcal{P} and \mathcal{B} parameters. These results show that it is possible to use RASPT2 to locate minimum energy structures and intersections. However, in practice, it is often difficult to reach a satisfactory SCF convergence.

4.3 Difference between Three CASPT2 Variants

The energetics of the three CASPT2 variants are investigated using benzene. Following Ref. 50, we identify seven MECIs and perform reference calculations using a two-state averaged CASSCF(6e,6o) wavefunction. The entry 5 in Ref. 50 is excluded because it was predicted to be a saddle point in the intersection space. Table 2 summarizes relative energies of S_0 at the D_{6h} (global minimum) structure, S_1 for the same geometry, and the seven MECIs obtained with SA-CASSCF (SCF in Table 2), three CASPT2 variants (XMS, XDW, and RMS), and PC-NEVPT2⁵⁰ and XMCQDPT2⁴⁹ previously reported in the literature. In general, our XMS results are comparable to previous⁵⁰ ones obtained with a different program package (and a different approximation for the electron repulsion integrals). Geometry optimizations converged at the first step for most entries because the initial structures are obtained from Ref. 50. This fact is a sensible evidence that the implemented derivatives, including the interstate coupling vector, are accurate.

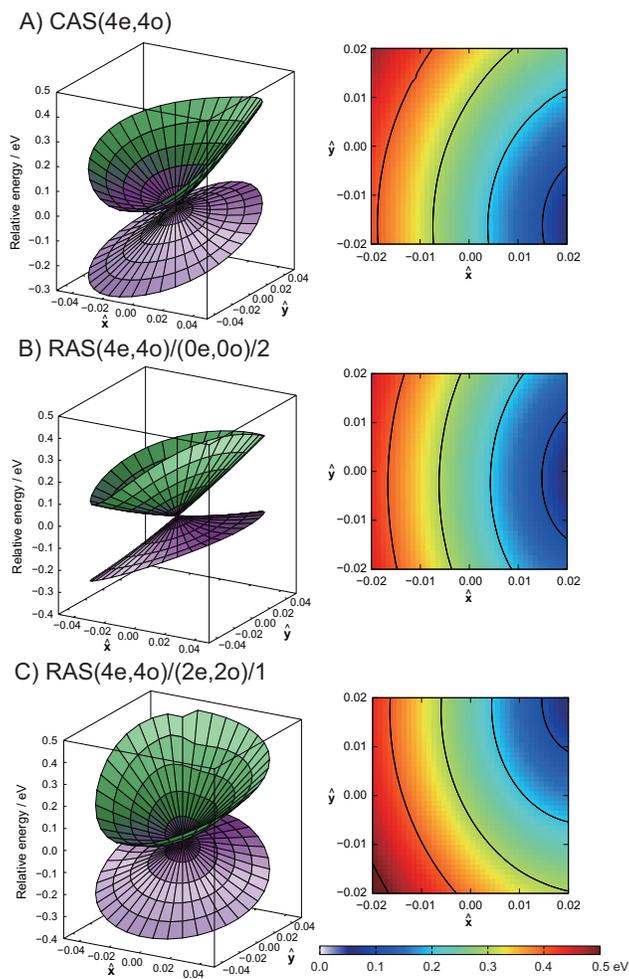


Figure 3: PESs of butadiene around the CASPT2/RASPT2 MECIs (left) and energy difference between S_0 and S_1 around the SA-CASSCF/RASSCF MECIs (right) with XMS-CASPT2 or XMS-RASPT2 using different active spaces: A) CAS(4e,4o), B) RAS(4e,4o)/(0e,0o)/2, and C) RAS(4e,4o)/(2e,2o)/1.

Table 2: Relative Energies of Benzene with Respect to the S_0 Energy at the D_{6h} Geometry and the Seven MECIs (unit in eV).

	SCF	XMS	XDW	RMS	PC-NEVPT2 ^a	XMCQDPT2 ^b
S_0 at D_{6h}	0.00	0.00	0.00	0.00	0.00	0.00
S_1 at D_{6h}	4.98	4.87	4.73	4.68	5.35	4.83
1	5.55	4.96	4.87	4.93	5.30	4.99
2	6.34	5.02	4.93	5.17	5.15	5.10
3	5.90	5.12	5.03	5.10	5.46	5.12
4	5.83	5.46	5.36	5.42	5.84	5.47
6	6.07	5.15	5.06	5.22	5.39	5.30
7	6.32	6.34	6.25	6.19	6.69	6.44
8	6.44	6.49	6.39	6.35	6.81	6.56

^a Ref. 50 ^b Ref. 49

The experimental vertical excitation energy at the D_{6h} geometry is 4.9 eV,⁷³ and so XMS is the closest. RMS deviates by 0.2 eV, but a systematic investigation is required to determine which CASPT2 variant yields the best general agreement with experiment.

Among the seven MECIs, PC-NEVPT2 predicts that **2** is more stable than **1** by 0.15 eV, while all three CASPT2 variants predict the opposite. Among all MRPT approaches, RMS is the only one stabilizing **3** over **2**. In any case, the difference between the three CASPT2 variants is at most 0.25 eV, which is in the range of accuracy expected for MRPT methodologies, such that any variant is potentially a valid option in practice.

The difference between XMS and XDW is consistently 0.1 eV for all MECI entries. In the vicinity of the CI at the PT2 level, the reference CASSCF energy is quasi-degenerate. This means that $\Delta_{\alpha\beta}$ (Eq. (22)) becomes nearly zero, and the density matrix of the α and β states contribute almost equally in the energy weighting scheme. In particular, when two states are averaged, the energy-weighted density matrix for XDW is roughly equivalent to that of XMS, and therefore it is rather natural that the relative energy for both methods is similar. Consequently, the difference between MECI entries of XMS and XDW in Table 2 is primarily due to the difference in the S_0 energy at the D_{6h} structure. As discussed below, that will not be the case if more states are averaged.

4.4 Dependence on State-Averaging

As mentioned in the previous section, the computed results depend on the number of states (N_{state}) included in the model space, and this effect is herein discussed for XMS-, XDW-, and RMS-CASPT2(6e,4o) using ethylene. For this system and active space, the number of CSFs is ten, and calculations up to the complete model space can be performed. In Table 3, we show the excitation energies from the ground state at the D_{2h} structure to the Franck–Condon (FC) point, $S_1(\text{FC})$, and to the pyramidalized MECI as a function of N_{state} .

Table 3: Relative Energies of the S_1 State at the $S_1(\text{FC})$ Structure and Pyramidalized MECI for Ethylene with Respect to S_0 at the D_{2h} Ground State Geometry (unit in eV)

N_{state}	XMS		XDW		RMS	
	$S_1(\text{FC})$	MECI	$S_1(\text{FC})$	MECI	$S_1(\text{FC})$	MECI
2	8.140	4.557	8.089	4.433	8.089	4.646
3	8.010	4.555	7.883	4.437	8.003	4.665
4	7.902	4.586	7.799	4.519	7.944	4.716
5	8.003	4.445	7.762	4.432	7.932	4.691
6	8.128	4.410	7.753	4.424	7.928	4.651
7	8.132	4.344	7.732	4.404	7.901	4.644
8	8.107	4.314	7.733	4.411	7.899	4.660
9	8.076	4.270	7.714	4.391	7.881	4.675
10	8.047	4.243	7.659	4.399	7.860	4.698
Average	8.060	4.413	7.792	4.428	7.937	4.672
Max Difference	0.239	0.344	0.430	0.128	0.229	0.072
R-M-S Deviation	0.074	0.122	0.120	0.036	0.066	0.024

Using XMS, the excitation energy to $S_1(\text{FC})$ almost randomly varies as a function of N_{state} , while XDW and RMS exhibit a decreasing trend. A comparison between the latter two reveals that XDW is considerably more sensitive to N_{state} , and that the magnitude of the deviation is significantly larger than that of XMS and RMS. The different ways in which the excitation energy depends on the number of states for the three methodologies can be rationalized as follows. Because the reference energy (i.e., the one used for $\Delta_{\alpha\beta}$ in Eq. (22)) of S_0 is very different from that of the excited states, the zeroth-order Hamiltonian in XDW has a very pronounced state-specific character, which makes it similar to that of RMS. In

fact, the *absolute* S_0 energy for both of these methods differ by at most 0.045 eV, and their ground state properties are therefore expected to be comparable. In contrast, S_0 from XMS is considerably different, because the Fock operator is averaged over all states with equal weights regardless of the gap between the ground and the excited states. The situation is different for S_1 and higher excited states. For instance, for $N_{\text{state}} > 4$, S_1 , S_2 and S_3 are energetically very similar, and thus the Fock operators for S_1 of XDW and XMS are roughly equivalent, providing comparable *absolute* S_1 energies (different by at most 0.073 eV). Their relative energy with respect to the ground state differ instead significantly due to the different descriptions of the S_0 state. On the other hand, the excited states obtained with RMS are not equivalent to the other two variants, due to the purely state-specific nature of its Fock operators. In short, XDW describes the ground state like RMS, but the excited states like XMS, at least with the choice of exponent made in Eq. (22) (for other choices, see Ref. 38). Hence, even though XDW is formally interpolating between XMS and RMS in terms of the ζ parameter, the predicted relative energies will generally not interpolate between these two extremes.

The situation with the MECI energy is instead slightly different. XMS has a stronger dependence on N_{state} than XDW and RMS, and its decreasing trend is more pronounced than in the case of the excitation energies. The R-M-S deviation of RMS is instead only 0.024 eV, while that for XDW is also relatively small; this is because the zeroth-order description of S_0 and S_1 is a mixture of the two states for both, regardless of the value of N_{state} .

From the above analysis, we can summarize that the dependence of the vertical excitation energy on N_{state} is smaller with XMS and RMS than with XDW, while the dependence of the MECI energy is small when using XDW and RMS since the zeroth-order Hamiltonian is mostly described using the crossing states. Thus, these results suggest that the energetics predicted by RMS are the least sensitive to N_{state} and that this variant is a sensible choice when averaging over many states. At last, we shall note that the dependence of the XDW energy with respect to N_{state} is likely to change for different values of ζ .

Representative geometrical parameters are summarized in Tables S1–S3 (Supporting Information). These do not vary significantly for the global minimum structure between the CASPT2 variants, and they are quite insensitive with respect to N_{state} . The C–C distance varied only 0.005 Å with XMS, whereas it was 0.002 Å with RMS, and the C–H distance remained almost unaffected. The ground state is energetically separated from the other states, and so the global minimum structure determined using XDW and RMS are almost the same. The C–C distance for the MECI structure is slightly more sensitive to N_{state} . The maximum difference observed for this bond with XMS and RMS is 0.037 Å and 0.006 Å, respectively. The C–C distance predicted using XMS is slightly longer than that predicted using XDW and RMS. For a large value of N_{state} , the state-averaged Fock operator (XMS) will have more contributions from highly excited states, which typically involve additional electron occupation numbers in anti-bonding orbitals, and so the C–C bond distance becomes longer with a large N_{state} .

4.5 Application to Dioxetanone

In this subsection, we present the results obtained by applying the developed methodologies to dioxetanone. This compound has been examined^{35,74–76} using multiconfigurational methods as a model system of firefly bioluminescence. In this study, two minimum energy and two MECI structures were located using SA-CASSCF and the three CASPT2 variants. We selected an active space of CAS(16e,13o) with $N_{\text{state}} = 2$. We label the two minimum energy structures on the S_0 surface as (σ, σ^*) -Min and (n, σ^*) -Min, following the nomenclature of Ref. 35. The MECIs are labeled as MECI-1 and MECI-2, and correspond to CI_{OO} and CI_{CC} of Ref. 76, respectively, with the subscript specifying the reaction coordinate for which the PES crossings occur. The relative energies for these stationary points are summarized in Table 4.

The considerable difference between SA-CASSCF and CASPT2 in the prediction of the MECI energies has been known for more than ten years.^{35,74} In previous studies, single-point

Table 4: Relative Energies of Dioxetanone (unit in kcal/mol)

	(σ, σ^*) -Min		MECI-1	(n, σ^*) -Min		MECI-2
	S ₀	S ₁	S ₀ /S ₁	S ₀	S ₁	S ₀ /S ₁
SA-CASSCF	0.00	98.1	14.5	12.6	16.6	19.5
XMS	0.00	104.0	26.9	25.3	30.0	28.2
XDW	0.00	92.2	23.5	21.9	26.7	24.8
RMS	0.00	98.3	24.4	22.8	27.4	25.8

CASPT2 calculations using SA-CASSCF structures predicted significantly higher energies (more than 10 kcal/mol) than SA-CASSCF. As per the results obtained in this study, geometry optimizations at the PT2 level support the single-point CASPT2 energies. However, our CASPT2 results predicted that the energy for MECI-2 is lower than for S₁ at (n, σ^*) -Min, while the opposite is true for SA-CASSCF. The difference is about five kcal/mol, and this may have a non-negligible effect on the ratio of the non-radiative decay. Previous MS-CASPT2//SA-CASSCF calculations³⁵ predicted the same trend as in this SA-CASSCF calculation, such that dynamic electron correlation may play an important role in the geometry optimization for this particular system.

The reaction coordinate of MECI-2 is reported to be along a C–C bond elongation. The predicted C–C distance is 1.64 and 1.60 Å (Table S7 in the Supporting Information) using SA-CASSCF and all CASPT2 variants, respectively, such that there is only a small difference between the methods. However, MECI-1 is reported along the reaction coordinate of O–O elongation, and the predicted bond lengths are 2.47 and 2.60 Å (Table S5 in the Supporting Information). Hence, in this case, the perturbative correction is more significant, resulting in bond distances differing by more than 0.1 Å.

All four methods predicted that the local topology of MECI-1 is peaked bifurcating as per the \mathcal{P} and \mathcal{B} parameters (Table S5). However, this was not the case for MECI-2 (Table S7), where the SA-CASSCF approach characterizes it as a peaked bifurcating CI ($\mathcal{P} = 0.05$ and $\mathcal{B} = 0.39$), while the CASPT2 variants as sloped single-path ($5.00 < \mathcal{P} < 5.50$ and

1.83 < \mathcal{B} < 1.90). Our CASPT2 result agrees with the findings in Ref. 74. As observed in many situations, CASSCF and CASPT2 predict qualitatively different electronic structures, and although the analysis using \mathcal{P} and \mathcal{B} is a local approximation,⁶⁴ the predicted topology agrees with the computed PESs in the branching plane, e.g., in Figures 1 and 3.

4.6 Dioxetanone with XMS-RASPT2

Large active spaces lead to very long expansions in terms of CSFs or determinants, which are computationally very expensive. Hence, it is often convenient to work in reduced active spaces, or to restrict the number of allowed excitations within them as done in RAS-based methods. The RASPT2 implementation in OpenMolcas⁵⁶ does not fully account all excitations within the active space (such as excitations from RAS1 to RAS2) in the perturbation calculation, and so it is generally better to prefer large RAS2 spaces or CAS whenever possible. Moreover, the quasi-canonicalization of MOs in RASPT2 is different from that used in CASPT2. In CASPT2, all MOs in the active space are simultaneously orthonormalized, whereas, in RASPT2, the MOs are separately orthonormalized in RAS1, RAS2, and RAS3. Therefore, RASPT2 using the same number of CSFs (N_{CSF}) as in CAS does not reproduce the corresponding CASPT2 result. It is, therefore, important to examine the reliability of RASPT2. The energetics of various active spaces is summarized in Table 5. All calculations were performed using XMS-CASPT2 or XMS-RASPT2 ($N_{\text{state}} = 2$), and each energy is obtained after relaxing the geometry for each individual active space. In a few cases, it was impossible to locate the MECIs, possibly due to either the non-invariance of XMS-CASPT2 when used with the SS-SR contraction scheme, or because of rotations between different orbital blocks.

The most compact wavefunction expansions, CAS(4e,3o) and RAS(16e,13o)/(0e,0o)/2, are expressed with less than 1000 CSFs, and these calculations predicted different energetics from the reference CAS(16e,13o) energy. The best agreement is achieved using RAS(16e,13o)/(12e,9o)/1 whose deviation is less than 1.5 kcal/mol, and the improvement

Table 5: Relative Energies of Dioxetanone with XMS-CASPT2 Using Different Active Spaces (unit in kcal/mol)

Partitioning	N_{CSF}	(σ, σ^*) -Min		MECI-1	(n, σ^*) -Min		MECI-2
		S_0	S_1	S_0/S_1	S_0	S_1	S_0/S_1
CAS(16e,13o)	429,429	0.0	104.0	26.9	25.3	30.0	28.2
CAS(14e,11o)	32,670	0.0	104.5	25.4	23.7	28.4	31.7
CAS(12e, 9o)	2,520	0.0	104.4	21.4	19.6	24.5	— ^a
CAS(4e, 3o)	6	0.0	116.2	34.0	31.1	37.5	33.5
RAS(16e,13o)/(12e,9o)/1	42,084	0.0	103.7	28.0	24.9	30.0	26.9
RAS(16e,13o)/(4e,3o)/2	4,505	0.0	107.2	— ^a	29.7	35.3	29.2
RAS(16e,13o)/(0e,0o)/4	42,301	0.0	104.1	20.9	19.1	24.4	20.8
RAS(16e,13o)/(0e,0o)/2	861	0.0	135.7	55.5	53.4	59.2	56.0

^a Could not locate the stationary point.

over CAS(12e,9o) is more than five kcal/mol for other than the S_1 energy at the (σ, σ^*) -Min conformation, but with more than ten times N_{CSF} . Although N_{CSF} is similar to RAS(16e,13o)/(12e,9o)/1, RAS(16e,13o)/(0e,0o)/4 is disappointingly different from the reference: it deviates by more than five kcal/mol for other than the S_1 energy at the (σ, σ^*) -Min conformation, and very similar to the CAS(12e,9o) result. However, RAS(16e,13o)/(12e,9o)/1 predicted the opposite energetic order for MECI-1 and MECI-2. The second best agreement was achieved using CAS(14e,11o): the largest difference from CAS(16e,13o) is about 2.5 kcal/mol despite the overestimation for MECI-2. Although it is difficult to finely control RASPT2, it can be used as a cost-effective alternative, especially considering the general expected accuracy of MRPTs. However, if N_{CSF} for a CAS and RAS is similar, the CAS-based approach should be preferred.

5 Conclusions

The analytic gradient and derivative coupling vectors for (X)MS-, XDW-, and RMS-CASPT2, as well as their RASPT2 counterparts, were derived and implemented in the open-source package OpenMolcas. As for other quasi-degenerate MRPT approaches, the derivation is

based on the Lagrangian formalism and the implementation was validated against numerical gradients.

First, we carried out a series of calculations to assess the qualitative and quantitative differences between various flavors of multistate CASPT2 in terms of the smoothness of the PES, the accuracy of excitation energies and the robustness against the number of states included in the model space. Second, we applied the developed methodologies to dioxetanone and appraised the paradigm of correcting lower-level results with single point energy calculations, e.g. CASPT2 energies on CASSCF structures. Moreover, the same molecular system was used to assess the ability of RASPT2 to be a substitute for prohibitively large complete active spaces.

Our findings can be summarized as follows. For ethylene, we have found that the PES around a minimum energy conical intersection obtained with XMS and XDW is smooth, while uncompromising artifacts were observed for RMS. As already documented in previous works, MS failed in this respect. Smoothness is observed in the case of XMS-RASPT2, while the qualitative characterization of the conical intersection strongly depends on the parametrization of the RAS. Vertical transition energies for the benzene molecule computed with several flavors of MRPT vary within the expected accuracy for this family of approximations, such that no definitive conclusion can be made in favor of any variant. Returning to ethylene, now looking at $S_0 \rightarrow S_1$ excitation energies at the Franck–Condon structure, we observed that XMS and RMS are less sensitive to the number of states included as compared to XDW. On the other hand, for the relative energies to the MECI, XMS and XDW are more sensitive to N_{state} than RMS. In the case of dioxetanone it was found that CASPT2 molecular structures may be qualitatively different from those obtained with CASSCF, and could induce different relative energies between stationary points. Additionally, the (16e,13o) active space applied to this molecular system was utilized as a reference in assessing the appropriateness of smaller complete and restricted active spaces. For the case considered here, it is relatively straightforward to construct compacter CASs which provide consistent results

with the reference, whereas this is a much more delicate task in the case of RASs.

To conclude, this work opens the door for non-adiabatic molecular dynamics simulations with a variety of multistate CASPT2 flavors, while the availability of relaxed densities now offers simple development roads towards the calculation of accurate molecular properties, as for example, oscillator and rotatory strengths.

Acknowledgement

Y.N. thanks Professor Hayashi and Professor Kurashige (Kyoto University) for their kind support. This work was supported by JSPS KAKENHI (Grant No. 19H02682 and 20K15230). S.B. acknowledges the Swiss National Science Foundation (SNSF) for the funding received through the Postdoc Mobility fellowship (grant number 199192). R.L. acknowledges the Swedish Research Council (VR, Grant No. 2020-03182).

Supporting Information Available

The Supporting Information is available free of charge at

optimized coordinates for structures determined in this study and representative structural parameters for ethylene and dioxetanone.

References

- (1) Yarkony, D. R. Diabolical Conical Intersections. *Rev. Mod. Phys.* **1996**, *68*, 985–1013.
- (2) Yarkony, D. R. Conical Intersections: The New Conventional Wisdom. *J. Phys. Chem. A* **2001**, *105*, 6277–6293.
- (3) Domcke, W.; Yarkony, D. R. Role of Conical Intersections in Molecular Spectroscopy and Photoinduced Chemical Dynamics. *Annu. Rev. Phys. Chem.* **2012**, *63*, 325–352.

- (4) Musiał, M.; Perera, A.; Bartlett, R. J. Multireference Coupled-Cluster Theory: The Easy Way. *J. Chem. Phys.* **2011**, *134*, 114108.
- (5) Lyakh, D. I.; Musiał, M.; Lotrich, V. F.; Bartlett, R. J. Multireference Nature of Chemistry: The Coupled-Cluster View. *Chem. Rev.* **2012**, *112*, 182–243.
- (6) Köhn, A.; Hanauer, M.; Mück, L. A.; Jagau, T.-C.; Gauss, J. State-Specific Multireference Coupled-Cluster Theory. *WIREs Comput. Mol. Sci.* **2013**, *3*, 176–197.
- (7) Szalay, P. G.; Müller, T.; Gidofalvi, G.; Lischka, H.; Shepard, R. Multiconfiguration Self-Consistent Field and Multireference Configuration Interaction Methods and Applications. *Chem. Rev.* **2012**, *112*, 108–181.
- (8) Roos, B. O.; Linse, P.; Siegbahn, P. E. M.; Blomberg, M. R. A. A Simple Method for the Evaluation of the Second-Order-Perturbation Energy from External Double-Excitations with a CASSCF Reference Wavefunction. *Chem. Phys.* **1982**, *66*, 197 – 207.
- (9) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-Order Perturbation Theory with a CASSCF Reference Function. *J. Phys. Chem.* **1990**, *94*, 5483–5488.
- (10) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-Order Perturbation Theory with a Complete Active Space Self-Consistent Field Reference Function. *J. Chem. Phys.* **1992**, *96*, 1218–1226.
- (11) Granovsky, A. A. Extended Multi-Configuration Quasi-Degenerate Perturbation Theory: The New Approach to Multi-State Multi-Reference Perturbation Theory. *J. Chem. Phys.* **2011**, *134*, 214113.
- (12) Nakano, H. Quasidegenerate Perturbation Theory with Multiconfigurational Self-Consistent-Field Reference Functions. *J. Chem. Phys.* **1993**, *99*, 7983–7992.

- (13) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of n -Electron Valence States for Multireference Perturbation Theory. *J. Chem. Phys.* **2001**, *114*, 10252–10264.
- (14) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. N -Electron Valence State Perturbation Theory: a Fast Implementation of the Strongly Contracted Variant. *Chem. Phys. Lett.* **2001**, *350*, 297–305.
- (15) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. 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.* **2002**, *117*, 9138–9153.
- (16) Khait, Y. G.; Song, J.; Hoffmann, M. R. Explication and Revision of Generalized Van Vleck Perturbation Theory for Molecular Electronic Structure. *J. Chem. Phys.* **2002**, *117*, 4133–4145.
- (17) Fink, R. F. Two New Unitary-Invariant and Size-Consistent Perturbation Theoretical Approaches to the Electron Correlation Energy. *Chem. Phys. Lett.* **2006**, *428*, 461–466.
- (18) Sen, A.; Sen, S.; Samanta, P. K.; Mukherjee, D. Unitary Group Adapted State Specific Multireference Perturbation Theory: Formulation and Pilot Applications. *J. Comput. Chem.* **2015**, *36*, 670–688.
- (19) Li, C.; Evangelista, F. A. Driven Similarity Renormalization Group for Excited States: A State-Averaged Perturbation Theory. *J. Chem. Phys.* **2018**, *148*, 124106.
- (20) Celani, P.; Werner, H.-J. Analytical Energy Gradients for Internally Contracted Second-Order Multireference Perturbation Theory. *J. Chem. Phys.* **2003**, *119*, 5044–5057.
- (21) MacLeod, M. K.; Shiozaki, T. Communication: Automatic Code Generation Enables Nuclear Gradient Computations for Fully Internally Contracted Multireference Theory. *J. Chem. Phys.* **2015**, *142*, 051103.

- (22) Song, C.; Neaton, J. B.; Martínez, T. J. Reduced Scaling Formulation of CASPT2 Analytical Gradients Using the Supporting Subspace Method. *J. Chem. Phys.* **2021**, *154*, 014103.
- (23) Nishimoto, Y. Analytic Gradients for Restricted Active Space Second-Order Perturbation Theory (RASPT2). *J. Chem. Phys.* **2021**, *154*, 194103.
- (24) Theis, D.; Khait, Y. G.; Hoffmann, M. R. GVVPT2 Energy Gradient Using a Lagrangian Formulation. *J. Chem. Phys.* **2011**, *135*, 044117.
- (25) Nishimoto, Y. Analytic First-Order Derivatives of Partially Contracted n -Electron Valence State Second-Order Perturbation Theory (PC-NEVPT2). *J. Chem. Phys.* **2019**, *151*, 114103.
- (26) Park, J. W.; Al-Saadon, R.; Strand, N. E.; Shiozaki, T. Imaginary Shift in CASPT2 Nuclear Gradient and Derivative Coupling Theory. *J. Chem. Theory Comput.* **2019**, *15*, 4088–4098.
- (27) Park, J. W. Analytical First-Order Derivatives of Second-Order Extended Multiconfiguration Quasi-Degenerate Perturbation Theory (XMCQDPT2): Implementation and Application. *J. Chem. Theory Comput.* **2020**, *16*, 5562–5571.
- (28) Wang, S.; Li, C.; Evangelista, F. A. Analytic Energy Gradients for the Driven Similarity Renormalization Group Multireference Second-Order Perturbation Theory. *J. Chem. Theory Comput.* **2021**, *17*, 7666–7681.
- (29) Park, J. W. 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.* **2022**, *18*, 2233–2245.

- (30) Li Manni, G.; Carlson, R. K.; Luo, S.; Ma, D.; Olsen, J.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory. *J. Chem. Theory Comput.* **2014**, *10*, 3669–3680.
- (31) Sand, A. M.; Hoyer, C. E.; Sharkas, K.; Kidder, K. M.; Lindh, R.; Truhlar, D. G.; Gagliardi, L. Analytic Gradients for Complete Active Space Pair-Density Functional Theory. *J. Chem. Theory Comput.* **2018**, *14*, 126–138.
- (32) Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. The multi-state CASPT2 method. *Chem. Phys. Lett.* **1998**, *288*, 299–306.
- (33) Angeli, C.; Borini, S.; Cestari, M.; Cimiraglia, R. A Quasidegenerate Formulation of the Second Order n -Electron Valence State Perturbation Theory Approach. *J. Chem. Phys.* **2004**, *121*, 4043–4049.
- (34) Levine, B. G.; Coe, J. D.; Martínez, T. J. Optimizing Conical Intersections without Derivative Coupling Vectors: Application to Multistate Multireference Second-Order Perturbation Theory (MS-CASPT2). *J. Phys. Chem. B* **2008**, *112*, 405–413.
- (35) Liu, F.; Liu, Y.; De Vico, L.; Lindh, R. Theoretical Study of the Chemiluminescent Decomposition of Dioxetanone. *J. Am. Chem. Soc.* **2009**, *131*, 6181–6188.
- (36) Gozem, S.; Melaccio, F.; Valentini, A.; Filatov, M.; Huix-Rotllant, M.; Ferré, N.; Frutos, L. M.; Angeli, C.; Krylov, A. I.; Granovsky, A. A.; Lindh, R.; Olivucci, M. Shape of Multireference, Equation-of-Motion Coupled-Cluster, and Density Functional Theory Potential Energy Surfaces at a Conical Intersection. *J. Chem. Theory Comput.* **2014**, *10*, 3074–3084.
- (37) Park, J. W.; Shiozaki, T. Analytical Derivative Coupling for Multistate CASPT2 Theory. *J. Chem. Theory Comput.* **2017**, *13*, 2561–2570.

- (38) Battaglia, S.; Lindh, R. On the Role of Symmetry in XDW-CASPT2. *J. Chem. Phys.* **2021**, *154*, 034102.
- (39) Shiozaki, T.; Gyórfy, W.; Celani, P.; Werner, H.-J. Communication: Extended Multi-State Complete Active Space Second-Order Perturbation Theory: Energy and Nuclear Gradients. *J. Chem. Phys.* **2011**, *135*, 081106.
- (40) Polyak, I.; Hutton, L.; Crespo-Otero, R.; Barbatti, M.; Knowles, P. J. Ultrafast Photoinduced Dynamics of 1,3-Cyclohexadiene Using XMS-CASPT2 Surface Hopping. *J. Chem. Theory Comput.* **2019**, *15*, 3929–3940.
- (41) Heindl, M.; González, L. A XMS-CASPT2 Non-Adiabatic Dynamics Study on Pyrrole. *Comput. Theor. Chem.* **2019**, *1155*, 38–46.
- (42) Battaglia, S.; Lindh, R. Extended Dynamically Weighted CASPT2: The Best of Two Worlds. *J. Chem. Theory Comput.* **2020**, *16*, 1555–1567.
- (43) Li, C.; Lindh, R.; Evangelista, F. A. Dynamically Weighted Multireference Perturbation Theory: Combining the Advantages of Multi-State and State-Averaged Methods. *J. Chem. Phys.* **2019**, *150*, 144107.
- (44) Deskevich, M. P.; Nesbitt, D. J.; Werner, H.-J. Dynamically Weighted Multiconfiguration Self-Consistent Field: Multistate Calculations for $F + H_2O \rightarrow HF + OH$ Reaction Paths. *J. Chem. Phys.* **2004**, *120*, 7281–7289.
- (45) Park, J. W. Single-State Single-Reference and Multistate Multireference Zeroth-Order Hamiltonians in MS-CASPT2 and Conical Intersections. *J. Chem. Theory Comput.* **2019**, *15*, 3960–3973.
- (46) Park, J. W.; Al-Saadon, R.; MacLeod, M. K.; Shiozaki, T.; Vlaisavljevich, B. Multireference Electron Correlation Methods: Journeys along Potential Energy Surfaces. *Chem. Rev.* **2020**, *120*, 5878–5909.

- (47) Park, J. W.; Shiozaki, T. On-the-Fly CASPT2 Surface-Hopping Dynamics. *J. Chem. Theory Comput.* **2017**, *13*, 3676–3683.
- (48) Khait, Y. G.; Theis, D.; Hoffmann, M. R. Nonadiabatic Coupling Terms for the GVVPT2 Variant of Multireference Perturbation Theory. *Chem. Phys.* **2012**, *401*, 88–94.
- (49) Park, J. W. Analytical Gradient Theory for Quasidegenerate N-Electron Valence State Perturbation Theory (QD-NEVPT2). *J. Chem. Theory Comput.* **2020**, *16*, 326–339.
- (50) Nishimoto, Y. Locating Conical Intersections Using the Quasidegenerate Partially and Strongly Contracted NEVPT2 Methods. *Chem. Phys. Lett.* **2020**, *744*, 137219.
- (51) Park, J. W. Analytical Gradient Theory for Resolvent-Fitted Second-Order Extended Multiconfiguration Perturbation Theory (XMCQDPT2). *J. Chem. Theory Comput.* **2021**, *17*, 6122–6133.
- (52) Calio, P. B.; Truhlar, D. G.; Gagliardi, L. Nonadiabatic Molecular Dynamics by Multiconfiguration Pair-Density Functional Theory. *J. Chem. Theory Comput.* **2022**, *18*, 614–622.
- (53) Fdez. Galván, I.; Vacher, M.; Alavi, A.; Angeli, C.; Aquilante, F.; Autschbach, J.; Bao, J. J.; Bokarev, S. I.; Bogdanov, N. A.; Carlson, R. K.; Chibotaru, L. F.; Creutzberg, J.; Dattani, N.; Delcey, M. G.; Dong, S. S.; Dreuw, A.; Freitag, L.; Frutos, L. M.; Gagliardi, L.; Gendron, F.; Giussani, A.; González, L.; Grell, G.; Guo, M.; Hoyer, C. E.; Johansson, M.; Keller, S.; Knecht, S.; Kovačević, G.; Källman, E.; Li Manni, G.; Lundberg, M.; Ma, Y.; Mai, S.; Malhado, J. P.; Malmqvist, P.-Å.; Marquetand, P.; Mewes, S. A.; Norell, J.; Olivucci, M.; Oppel, M.; Phung, Q. M.; Pierloot, K.; Plasser, F.; Reiher, M.; Sand, A. M.; Schapiro, I.; Sharma, P.; Stein, C. J.; Sørensen, L. K.; Truhlar, D. G.; Ugandi, M.; Ungur, L.; Valentini, A.; Vancoillie, S.; Veryazov, V.; Weser, O.; Wesołowski, T. A.; Widmark, P.-O.; Wouters, S.; Zech, A.;

- Zobel, J. P.; Lindh, R. OpenMolcas: From Source Code to Insight. *J. Chem. Theory Comput.* **2019**, *15*, 5925–5964.
- (54) Aquilante, F.; Autschbach, J.; Baiardi, A.; Battaglia, S.; Borin, V. A.; Chibotaru, L. F.; Conti, I.; De Vico, L.; Delcey, M.; Fdez. Galván, I.; Ferré, N.; Freitag, L.; Garavelli, M.; Gong, X.; Knecht, S.; Larsson, E. D.; Lindh, R.; Lundberg, M.; Malmqvist, P. Å.; Nenov, A.; Norell, J.; Odelius, M.; Olivucci, M.; Pedersen, T. B.; Pedraza-González, L.; Phung, Q. M.; Pierloot, K.; Reiher, M.; Schapiro, I.; Segarra-Martí, J.; Segatta, F.; Seijo, L.; Sen, S.; Sergentu, D.-C.; Stein, C. J.; Ungur, L.; Vacher, M.; Valentini, A.; Veryazov, V. Modern quantum chemistry with [Open]Molcas. *J. Chem. Phys.* **2020**, *152*, 214117.
- (55) Celani, P.; Werner, H.-J. Multireference Perturbation Theory for Large Restricted and Selected Active Space Reference Wave Functions. *J. Chem. Phys.* **2000**, *112*, 5546–5557.
- (56) Malmqvist, P.-Å.; Pierloot, K.; Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. The Restricted Active Space Followed by Second-Order Perturbation Theory Method: Theory and Application to the Study of CuO₂ and Cu₂O₂ Systems. *J. Chem. Phys.* **2008**, *128*, 204109.
- (57) Roos, B. O.; Andersson, K. Multiconfigurational Perturbation Theory with Level Shift — the Cr₂ Potential Revisited. *Chem. Phys. Lett.* **1995**, *245*, 215–223.
- (58) Forsberg, N.; Malmqvist, P.-Å. Multiconfiguration Perturbation Theory with Imaginary Level Shift. *Chem. Phys. Lett.* **1997**, *274*, 196–204.
- (59) Witek, H. A.; Choe, Y.-K.; Finley, J. P.; Hirao, K. Intruder State Avoidance Multireference Møller–Plesset Perturbation Theory. *J. Comput. Chem.* **2002**, *23*, 957–965.
- (60) Helgaker, T.; Jørgensen, P. Configuration-Interaction Energy Derivatives in a Fully Variational Formulation. *Theor. Chem. Acc.* **1989**, *75*, 111–127.

- (61) Vlaisavljevich, B.; Shiozaki, T. Nuclear Energy Gradients for Internally Contracted Complete Active Space Second-Order Perturbation Theory: Multistate Extensions. *J. Chem. Theory Comput.* **2016**, *12*, 3781–3787.
- (62) Mori, T.; Kato, S. Dynamic Electron Correlation Effect on Conical Intersections in Photochemical Ring-Opening Reaction of Cyclohexadiene: MS-CASPT2 Study. *Chem. Phys. Lett.* **2009**, *476*, 97–100.
- (63) Mori, T.; Nakano, K.; Kato, S. Conical Intersections of Free Energy Surfaces in Solution: Effect of Electron Correlation on a Protonated Schiff Base in Methanol Solution. *J. Chem. Phys.* **2010**, *133*, 064107.
- (64) Fdez. Galván, I.; Delcey, M. G.; Pedersen, T. B.; Aquilante, F.; Lindh, R. 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.* **2016**, *12*, 3636–3653.
- (65) BAGEL, Brilliantly Advanced General Electronic-structure Library. <http://www.nubakery.org> under the GNU General Public License., (Accessed on March 18, 2021).
- (66) Shiozaki, T. BAGEL: Brilliantly Advanced General Electronic-structure Library. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1331.
- (67) Aquilante, F.; Lindh, R.; Bondo Pedersen, T. Unbiased Auxiliary Basis Sets for Accurate Two-Electron Integral Approximations. *J. Chem. Phys.* **2007**, *127*, 114107.
- (68) Aquilante, F.; Gagliardi, L.; Pedersen, T. B.; Lindh, R. Atomic Cholesky Decompositions: A Route to Unbiased Auxiliary Basis Sets for Density Fitting Approximation with Tunable Accuracy and Efficiency. *J. Chem. Phys.* **2009**, *130*, 154107.

- (69) Ghigo, G.; Roos, B. O.; Malmqvist, P.-Å. A Modified Definition of the Zeroth-Order Hamiltonian in Multiconfigurational Perturbation Theory (CASPT2). *Chem. Phys. Lett.* **2004**, *396*, 142 – 149.
- (70) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (71) Widmark, P.-O.; Malmqvist, P.-Å.; Roos, B. O. Density Matrix Averaged Atomic Natural Orbital (ANO) Basis Sets for Correlated Molecular Wave Functions. *Theoret. Chim. Acta* **1990**, *77*, 291–306.
- (72) Roos, B. O.; Lindh, R.; Malmqvist, P.-Å.; Veryazov, V.; Widmark, P.-O. Main Group Atoms and Dimers Studied with a New Relativistic ANO Basis Set. *J. Phys. Chem. A* **2004**, *108*, 2851–2858.
- (73) Lassette, E. N.; Skerbele, A.; Dillon, M. A.; Ross, K. J. High-Resolution Study of Electron-Impact Spectra at Kinetic Energies between 33 and 100 eV and Scattering Angles to 16° . *J. Chem. Phys.* **1968**, *48*, 5066–5096.
- (74) Chung, L. W.; Hayashi, S.; Lundberg, M.; Nakatsu, T.; Kato, H.; Morokuma, K. Mechanism of Efficient Firefly Bioluminescence via Adiabatic Transition State and Seam of Sloped Conical Intersection. *J. Am. Chem. Soc.* **2008**, *130*, 12880–12881.
- (75) Greenman, L.; Mazziotti, D. A. Strong Electron Correlation in the Decomposition Reaction of Dioxetanone with Implications for Firefly Bioluminescence. *J. Chem. Phys.* **2010**, *133*, 164110.
- (76) Ma, Y. Elucidating the Multi-Configurational Character of the Firefly Dioxetanone Anion and Its Prototypes in the Biradical Region Using Full Valence Active Spaces. *Phys. Chem. Chem. Phys.* **2020**, *22*, 4957–4966.

Graphical TOC Entry

