Derivative couplings of Mixed-Reference Spin-Flip TDDFT Excited States

Zheng Pei* and Yihan Shao†
Department of Chemistry, Brandeis University, Waltham, MA 02453

Arnab Chakraborty
Department of Chemistry, University of Southern California, Los Angeles, CA 90089
(Dated: January 31, 2026)

The non-adiabatic decay of same-spin excited states to the ground state (e.g., from S_1 to S_0) at conical intersections requires state-specific methods due to the near-degeneracy that challenges traditional self-consistent field-based response solutions. Mixed-reference spin-flip (MRSF) time-dependent density functional theory (TDDFT), as proposed by Choi et al. (and other spin-flip variants), offers significant advances by treating both ground and excited states on an equal footing while maintaining the spin purity of the states. This makes it a promising tool for simulating such decay processes. In this work, we derive the working equations of the MRSF-TDDFT method using a second-quantization operator formalism and provide analytic expressions for derivative couplings (DCs), along with a simplified state overlap formula for finite-difference DC evaluations. As an example, the DCs for XXX are computed.

I. INTRODUCTION

The non-adiabatic decay of excited states, such as the transition from the first excited state S_1 to the ground state S_0, is a critical process in photochemistry and photophysics, particularly near conical intersections. These decay processes are characterized by near-degeneracies in the electronic energies of the involved states, which present significant challenges for traditional approaches based on self-consistent field (SCF) theory for ground-state and (linear) response theory for excited-states. In this regime, the conventional linear-response approaches often fail to capture the true electronic character of the system, necessitating the development of more advanced, state-specific methodologies that can describe the complex interplay between excited and ground states.

Among the promising approaches, Mixed-Reference Spin-Flip (MRSF) Time-Dependent Density Functional Theory (TDDFT) has emerged as a powerful tool. Initially introduced by Choi et al., MRSF-TDDFT belongs to a broader family of spin-flip methods that provide a practical yet effective framework for treating excited-state dynamics. This method distinguishes itself by allowing for the equal treatment of ground and excited states while maintaining the spin purity of the states involved. By introducing a mixed-reference state, MRSF-TDDFT avoids the spin contamination typically encountered in traditional spin-flip TDDFT methods, making it a compelling choice for simulating non-adiabatic decay processes where accurate description of both ground and excited states is paramount.

Despite its advantages, the current formalism of MRSF-TDDFT is presented by introducing spin-paring couplings rather than a formal derivation. In this work, we derive the working equations of the MRSF-TDDFT method within the framework of second-quantization operators. This formulation not only provides a clear and concise representation of the method but also facilitates the derivation of analytic expressions for derivative couplings (DCs), which are essential for simulating the transition dynamics between electronic states. Furthermore, we introduce a simplified state overlap formula to efficiently evaluate the DCs using finite-difference methods. To illustrate the applicability of the method, we present an example calculation of the DCs for XXX, showcasing the practical use of the MRSF-TDDFT framework in non-adiabatic molecular dynamics simulations.

By leveraging these analytic tools, we aim to bridge the gap between the computational efficiency of spin-flip approaches and the accuracy required for simulating photochemical processes, thus enabling more realistic and scalable simulations of radiationless decay mechanisms, including those associated with conical intersections.

II. METHODS

A. Mixed-Reference Open-Shell State

The mixed-reference spin-flip (MRSF)-DFT was proposed to solve the spin contamination issue of original SF-DFT, which utilizes a higher spin state as a reference and treats the real ground-state and excited-states at equal footing (linear response framework). The mixed-reference ground state wavefunction is the equal combination of two Slater determinants

$$\Psi^{\text{ref}} = \frac{1}{\sqrt{2}} \left(\Psi_{M_s=+1} + \Psi_{M_s=-1} \right). \tag{1}$$

Here we only consider the real ground-state as closed-shell singlet state and restrict the α and β orbitals to be same. The $\Psi_{M_s=+1}$ is the optimized open-shell triplet state with spin projection number as $M_s = +1$, ie, two

* zheng.pei.c@gmail.com
† yihanshao@brandeis.edu

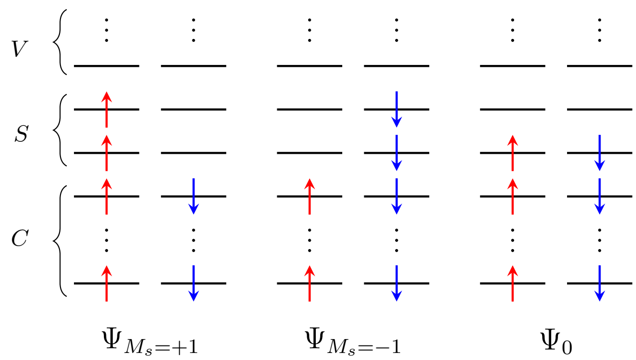

FIG. 1. Schematic electron occupation diagram for $\Psi_{M_s=+1}$ (left), $\Psi_{M_s=-1}$ (middle), and Ψ_0 (right) states. Orbitals are optimized for $\Psi_{M_s=+1}$ state. C, S, and V labels doubly occupied, singly occupied, and virtual orbitals, respectively. In particular, S space includes two orbitals in our current study, indexed by s and t from bottom-up. Red and blue arrows indicate spin-up (α) and spin-down (β) electrons, respectively.

more α-electrons; while $\Psi_{M_s=-1}$ is its spin image by flipping the spins of all the electrons such that it has two more β-electrons. We would use $\Psi_{\pm 1}$ to label the $\Psi_{M_s=\pm 1}$ triplet state as a short-hand notation without confusion.

The underlaying orbitals are then divided into doubly occupied core (C), singly occupied (S), and virtual (V) spaces, which are labeled by $i, j \in C \oplus S$, $s, t \in S$, and $a, b \in S \oplus V$ for individual orbitals in the following. Whether s or t is occupied or virtual orbital depends on the spin labels. In $\Psi_{M_s=+1}$, for instance, s_α is occupied while s_β is virtual orbital. In addition, the electron spins (α, β) are labeled by Greek letters $\sigma, \tau, \varsigma, \kappa$. The triplet states can then be written as

$$\begin{aligned} \left|\Psi_{M_s=+1}\right\rangle &= \hat{a}^\dagger_{t_\alpha} \hat{a}_{s_\beta} \left|\Psi_0\right\rangle = \hat{b}^{t_\alpha}_{s_\beta} \left|\Psi_0\right\rangle = \left|0^{t_\alpha}_{s_\beta}\right\rangle, \\ \left|\Psi_{M_s=-1}\right\rangle &= \hat{a}^\dagger_{t_\beta} \hat{a}_{s_\alpha} \left|\Psi_0\right\rangle = \hat{b}^{t_\beta}_{s_\alpha} \left|\Psi_0\right\rangle = \left|0^{t_\beta}_{s_\alpha}\right\rangle, \end{aligned} \tag{2}$$

with fermionic creator $\hat{a}^\dagger_p$ and annihilator $\hat{a}_p$ of p-th orbital, and the bosonic excitation operator $\hat{b}^p_q = \{\hat{a}^\dagger_p \hat{a}_q\}$ (in normal-ordering and $p \neq q$) and its conjugate $\left(\hat{b}^p_q\right)^\dagger = \hat{b}^q_p$. While $|\Psi_0\rangle$, short-noted as $|0\rangle$ (ie, Fermi vacuum), is a closed-shell Slater determinate with $n_\alpha = n_\beta = n$ electrons, constructed from optimized orbitals of ROKS calculation of $\Psi_{M_s=+1}$. With the above spin-flip excitation operator (as generalized spin lowering or raising operator), the reference state given in Eq. (1) can be written as

$$\left|\Psi^{\text{ref}}\right\rangle = \frac{1}{\sqrt{2}} \sum_\sigma^{\alpha,\beta} \hat{b}^{t_{\bar\sigma}}_{s_\sigma} \left|\Psi_0\right\rangle = \frac{1}{\sqrt{2}} \sum_\sigma^{\alpha,\beta} \left|0^{t_{\bar\sigma}}_{s_\sigma}\right\rangle, \tag{3}$$

which has same S_z value as of the real ground-state Ψ_0, where σ indicates α or β spin. We fix $s = n$ as the lower energy singly occupied orbital while $t = n + 1 \neq s$ is the second singly occupied orbital. The states are illustrated by the orbital diagrams in Fig. 1.

The MRSF-DFT excited-states can be obtained by linear combinations of the spin-flip excitations as

$$\begin{aligned} \hat{T}^I_\pm = \hat{T}^I_{\alpha\to\beta} \pm \hat{T}^I_{\beta\to\alpha} &= \sum_{ai} X^I_{ai} \left(\hat{a}^\dagger_{a_\beta} \hat{a}_{i_\alpha} \pm \hat{a}^\dagger_{a_\alpha} \hat{a}_{i_\beta}\right) \\ &= \sum_{ai} X^I_{ai} \left(\sum_\sigma^{\alpha,\beta} c_\sigma \hat{b}^{a_{\bar\sigma}}_{i_\sigma}\right), \end{aligned} \tag{4}$$

for I-th singlet ($+$) or triplet ($-$) excited-state with transition amplitudes X^I_{ai} (assumed to be real), including the α occupied orbital to β virtual orbital transitions ($\alpha \to \beta$) and their spin images ($\beta \to \alpha$). Here we defined the coefficients $c_\alpha = 1$ and $c_\beta = \pm 1$.

The normal-ordered Hamiltonian $\hat{H}$ with respect only to the $\Psi_{M_s=+1}$ state (energy) is given by

$$\begin{aligned} \hat{H} &= \sum_{pq} f_{pq} \{\hat{a}^\dagger_p \hat{a}_q\} + \frac{1}{4} \sum_{pq} \sum_{r\ell} \langle pr||q\ell\rangle \{\hat{a}^\dagger_p \hat{a}^\dagger_r \hat{a}_\ell \hat{a}_q\} \\ &= \sum_{pq} f_{pq} \hat{b}^p_q + \frac{1}{4} \sum_{pq} \sum_{r\ell} \langle pr||q\ell\rangle \, \hat{b}^{pr}_{q\ell}, \end{aligned} \tag{5}$$

including Fock and two-electron integrals, with p, q, r, ℓ labeling all the molecular spin orbitals. The I-th excitation energy ω_I of can be found by solving the equation of motions (EOM) formalism[3]

$$\left\langle \Psi^{\text{ref}} \middle| \left[\hat{T}^{I,\dagger}_\pm, \hat{H}, \hat{T}^J_\pm\right] \middle| \Psi^{\text{ref}} \right\rangle = \omega_I \left\langle \Psi^{\text{ref}} \middle| \left[\hat{T}^{I,\dagger}_\pm, \hat{T}^J_\pm\right] \middle| \Psi^{\text{ref}} \right\rangle, \tag{6}$$

from the Hamiltonian and excitation operators defined above. Here $[\hat{\mathcal{A}}, \hat{\mathcal{B}}] = \hat{\mathcal{A}}\hat{\mathcal{B}} - \hat{\mathcal{B}}\hat{\mathcal{A}}$ indicates a commutator, and $[\hat{\mathcal{A}}, \hat{\mathcal{B}}, \hat{\mathcal{C}}] = \frac{1}{2}\left([[\hat{\mathcal{A}}, \hat{\mathcal{B}}], \hat{\mathcal{C}}] + [\hat{\mathcal{A}}, [\hat{\mathcal{B}}, \hat{\mathcal{C}}]]\right) = \hat{\mathcal{A}}\hat{\mathcal{B}}\hat{\mathcal{C}} + \hat{\mathcal{C}}\hat{\mathcal{B}}\hat{\mathcal{A}} - \frac{1}{2}[\hat{\mathcal{B}}, [\hat{\mathcal{A}}, \hat{\mathcal{C}}]_+]_+$ the double commutator.[3] The above equation are able to be casted into a generalized eigenvalue equation

$$\Lambda^\pm X = \Delta^\pm X \Omega^\pm, \tag{7}$$

where the electronic response super-matrix $\Lambda^\pm$ and metric super-matrix $\Delta^\pm$ have elements

$$\begin{aligned} \Lambda^\pm_{ai,bj} &= \left\langle \Psi^{\text{ref}} \middle| \left[\hat{T}^\dagger_{\pm,ai}, \hat{H}, \hat{T}_{\pm,bj}\right] \middle| \Psi^{\text{ref}} \right\rangle, \\ \Delta^\pm_{ai,bj} &= \left\langle \Psi^{\text{ref}} \middle| \left[\hat{T}^\dagger_{\pm,ai}, \hat{T}_{\pm,bj}\right] \middle| \Psi^{\text{ref}} \right\rangle, \end{aligned} \tag{8}$$

respectively, where $\hat{T}_{\pm,ai} = \sum_\sigma c_\sigma \hat{b}^{a_{\bar\sigma}}_{i_\sigma}$. The amplitudes are orthonormalized as $X^{I,\dagger} \Delta^\pm X^J = \delta_{IJ}$. As a side note, the above approach should be equivalent to the following set of definitions

$$\begin{aligned} \left|\Psi^\pm\right\rangle &= \frac{1}{\sqrt{2}} \sum_\sigma^{\alpha,\beta} c_{\bar\sigma} \hat{b}^{t_{\bar\sigma}}_{s_\sigma} \left|\Psi_0\right\rangle, \\ \hat{T}^I &= \sum_{ai} X^I_{ai} \sum_\sigma \hat{b}^{a_{\bar\sigma}}_{i_\sigma}, \end{aligned} \tag{9}$$

where the excitation operator conserves the spin multiplicity while whether the excited-state is singlet $(-)$ or triplet $(+)$ is determined by the reference state spin. We would rather use the first definition through out the rest derivations, ie, carry the $\pm$ sign on the excitation operators of the excited-states.

Wick's theorem of particle-hole formalism would be used, from which in the normal ordering pattern, the quasiparticle creation operators ($\hat{p}^\dagger$ and $\hat{h}$) are standing at the left of the annihilators ($\hat{p}$ and $\hat{h}^\dagger$, where $p = a, b, c, \ldots$ labels virtual orbitals while $h = i, j, k, \ldots$ labels occupied orbitals).[4, 5] The fermionic and bosonic operators (with p and q for general orbitals) follow the relations

$$\begin{aligned}
\hat{a}_p^\dagger \hat{a}_q &= \delta_q^p n_q + \{\hat{a}_p^\dagger \hat{a}_q\} = \delta_q^p n_q + \hat{b}_q^p, \\
\hat{a}_q \hat{a}_p^\dagger &= \delta_q^p (1 - n_p) - \{\hat{a}_p^\dagger \hat{a}_q\} = \delta_q^p (1 - n_p) - \hat{b}_q^p,
\end{aligned}$$

in where $p \neq q$ for the bosonic operator $\hat{b}_q^p$.

B. Right-Hand Side of the EOM

The right-hand side of the EOM involves the metric matrix of the excitation operators. First, we can evaluate the commutators of the orbital excitation operators as

$$\begin{aligned}
\left[\hat{b}^{i_\sigma}_{a_{\bar\sigma}}, \hat{b}^{b_{\bar\lambda}}_{j_\lambda}\right] = \delta_{\sigma\lambda}\Big(&\delta^{b_{\bar\sigma}}_{a_{\bar\sigma}}(1 - n_{a_{\bar\sigma}})\hat{b}^{i_\sigma}_{j_\sigma} - \delta^{i_\sigma}_{j_\sigma} n_{i_\sigma} \hat{b}^{b_{\bar\sigma}}_{a_{\bar\sigma}} \\
&+ \delta^{b_{\bar\sigma}}_{a_{\bar\sigma}} \delta^{i_\sigma}_{j_\sigma} n_{i_\sigma}(1 - n_{a_{\bar\sigma}})\Big),
\end{aligned} \tag{10}$$

where $n_{i_\sigma} = 0$ when $i_\sigma > n_\sigma \in S$; and similarly $1 - n_{a_{\bar\sigma}} = 0$ when $a_{\bar\sigma} < n_{\bar\sigma} + 1 \in S$ of the reference state Ψ_0. Hence the metric is independent to the spin multiplicity of the targeted excited-states, such that $\Delta = \Delta^\pm$, whose matrix element in the mixed-reference ground state reads

$$\begin{aligned}
\Delta_{ai,bj} &= \left\langle \Psi^{\mathrm{ref}} \middle| \left[\hat{T}^\dagger_{\pm,ai}, \hat{T}_{\pm,bj}\right] \middle| \Psi^{\mathrm{ref}} \right\rangle \\
&= \frac{1}{2} \sum_{\sigma\lambda\eta} \Big(\Big\langle \hat{b}^{s_\lambda}_{t_{\bar\lambda}} \Big(\delta^{b_{\bar\sigma}}_{a_{\bar\sigma}}(1 - n_{a_{\bar\sigma}}) \hat{b}^{i_\sigma}_{j_\sigma} - \delta^{i_\sigma}_{j_\sigma} n_{i_\sigma} \hat{b}^{b_{\bar\sigma}}_{a_{\bar\sigma}} \Big) \hat{b}^{t_{\bar\eta}}_{s_\eta} \Big\rangle \\
&\qquad + \delta^{b_{\bar\sigma}}_{a_{\bar\sigma}} \delta^{i_\sigma}_{j_\sigma} n_{i_\sigma}(1 - n_{a_{\bar\sigma}}) \left\langle \hat{b}^{s_\lambda}_{t_{\bar\lambda}} \hat{b}^{t_{\bar\eta}}_{s_\eta} \right\rangle \Big) \\
&= \sum_\sigma \delta^{b_{\bar\sigma}}_{a_{\bar\sigma}} \delta^{i_\sigma}_{j_\sigma} \Big((1 - n_{a_{\bar\sigma}})(n_{i_\sigma} + \delta_{i_\sigma t_\sigma} - \delta_{i_\sigma s_\sigma}) \\
&\qquad + n_{i_\sigma}((1 - n_{a_{\bar\sigma}}) - \delta_{a_{\bar\sigma} t_{\bar\sigma}} + \delta_{a_{\bar\sigma} s_{\bar\sigma}}) \Big),
\end{aligned} \tag{11}$$

The above results are derived from the following contraction patterns

$$\{\hat{a}^\dagger_{s_\lambda} \hat{a}_{t_{\bar\lambda}}\}\{\hat{a}^\dagger_p \hat{a}_q\}\{\hat{a}^\dagger_{t_{\bar\eta}} \hat{a}_{s_\eta}\}; \qquad \{\hat{a}^\dagger_{s_\lambda} \hat{a}_{t_{\bar\lambda}}\}\{\hat{a}^\dagger_p \hat{a}_q\}\{\hat{a}^\dagger_{t_{\bar\eta}} \hat{a}_{s_\eta}\};$$

which gives

$$\left\langle \{\hat{a}^\dagger_{s_\lambda} \hat{a}_{t_{\bar\lambda}}\}\{\hat{a}^\dagger_p \hat{a}_q\}\{\hat{a}^\dagger_{t_{\bar\eta}} \hat{a}_{s_\eta}\} \right\rangle = \delta_{\lambda\eta} \delta^p_q \left(\delta^p_{t_{\bar\lambda}} - \delta^p_{s_\lambda} \right); \tag{12}$$

and the trivial contraction $\left\langle \{\hat{a}^\dagger_{s_\lambda} \hat{a}_{t_{\bar\lambda}}\}\{\hat{a}^\dagger_{t_{\bar\eta}} \hat{a}_{s_\eta}\} \right\rangle = \delta_{\lambda\eta}$.

C. Left-Hand Side of the EOM

Recall that the double commutator here reads[3]

$$[\hat{T}_1^\dagger, \hat{H}, \hat{T}_2] = \frac{1}{2} \left([[\hat{T}_1^\dagger, \hat{H}], \hat{T}_2] + [\hat{T}_1^\dagger, [\hat{H}, \hat{T}_2]] \right). \tag{13}$$

For the one-electron Fock operator, we have the elements

$$\begin{aligned}
\left[\hat{b}^{i_\sigma}_{a_{\bar\sigma}}, \hat{b}^p_q, \hat{b}^{b_{\bar\varsigma}}_{j_\varsigma}\right] &= \delta^p_{j_\varsigma} \delta^{i_\sigma}_q n_{i_\sigma} n_{j_\varsigma} \left(\hat{b}^{b_{\bar\varsigma}}_{a_{\bar\sigma}} - \delta_{\sigma\varsigma} \delta^{b_{\bar\sigma}}_{a_{\bar\sigma}} (1 - n_{a_{\bar\sigma}}) \right) \\
&+ \delta^p_{a_{\bar\sigma}} \delta^{b_{\bar\varsigma}}_q (1 - n_{a_{\bar\sigma}})(1 - n_{b_{\bar\varsigma}}) \left(\hat{b}^{i_\sigma}_{j_\varsigma} + \delta_{\sigma\varsigma} \delta^{i_\sigma}_{j_\sigma} n_{i_\sigma} \right) \\
&- \frac{1}{2} \delta_{\sigma\varsigma} \Big(\delta^{i_\sigma}_{j_\sigma} n_{i_\sigma} \left(\delta^{b_{\bar\sigma}}_q (1 - n_{b_{\bar\sigma}}) \hat{b}^p_{a_{\bar\sigma}} + \delta^p_{a_{\bar\sigma}} (1 - n_{a_{\bar\sigma}}) \hat{b}^{b_{\bar\sigma}}_q \right) \\
&\qquad + \delta^{b_{\bar\sigma}}_{a_{\bar\sigma}} (1 - n_{a_{\bar\sigma}}) \left(\delta^{i_\sigma}_q n_{i_\sigma} \hat{b}^p_{j_\sigma} + \delta^p_{j_\sigma} n_{j_\sigma} \hat{b}^{i_\sigma}_q \right) \Big)
\end{aligned} \tag{14}$$

Then its expectation value in the mixed-reference ground-state becomes

$$\sum_{\lambda\eta}\left\langle \Psi_{s_\lambda}^{t_{\bar\lambda}}\middle|\left[\hat T^\dagger_{\pm,ai},\hat F_{\rm N},\hat T_{\pm,bj}\right]\middle|\Psi_{s_\eta}^{t_{\bar\eta}}\right\rangle$$
$$= \frac{1}{2}\sum_\sigma \Bigg(f_{a_{\bar\sigma}b_{\bar\sigma}}\delta_{j_\sigma}^{i_\sigma}\Bigg((1-n_{a_{\bar\sigma}})(1-n_{b_{\bar\sigma}})\left(2n_{i_\sigma}+\delta_{t_\sigma}^{i_\sigma}-\delta_{s_\sigma}^{i_\sigma}\right)-\frac{1}{2}n_{i_\sigma}\left((1-n_{a_{\bar\sigma}})\left(\delta_{t_{\bar\sigma}}^{b_{\bar\sigma}}-\delta_{s_{\bar\sigma}}^{b_{\bar\sigma}}\right)+(1-n_{b_{\bar\sigma}})\left(\delta_{a_{\bar\sigma}}^{t_{\bar\sigma}}-\delta_{a_{\bar\sigma}}^{s_{\bar\sigma}}\right)\right)\Bigg)$$
$$-f_{j_\sigma i_\sigma}\delta_{a_{\bar\sigma}}^{b_{\bar\sigma}}\left(n_{i_\sigma}n_{j_\sigma}\left(2(1-n_{a_{\bar\sigma}})-\delta_{a_{\bar\sigma}}^{t_{\bar\sigma}}+\delta_{a_{\bar\sigma}}^{s_{\bar\sigma}}\right)-\frac{1}{2}(1-n_{a_{\bar\sigma}})\left(n_{i_\sigma}\left(\delta_{j_\sigma}^{t_\sigma}-\delta_{j_\sigma}^{s_\sigma}\right)+n_{j_\sigma}\left(\delta_{t_\sigma}^{i_\sigma}-\delta_{s_\sigma}^{i_\sigma}\right)\right)\right)\Bigg). \tag{15}$$

Similarly for the two-electron operators, we have

$$\left[\hat b_{a_{\bar\sigma}}^{i_\sigma},\hat b_{qs}^{pr},\hat b_{j_\varsigma}^{b_{\bar\varsigma}}\right] = -2\delta_{\sigma\varsigma}\delta_{j_\sigma}^{i_\sigma}n_{i_\sigma}\left(\delta_{a_{\bar\sigma}}^{p}(1-n_{a_{\bar\sigma}})\hat b_{q\ell}^{b_{\bar\sigma}r}+\delta_\ell^{b_{\bar\sigma}}(1-n_{b_{\bar\sigma}})\hat b_{qa_{\bar\sigma}}^{pr}\right)-2\delta_{\sigma\varsigma}\delta_{a_{\bar\sigma}}^{b_{\bar\sigma}}(1-n_{a_{\bar\sigma}})\left(\delta_q^{i_\sigma}n_{i_\sigma}\hat b_{j_\sigma\ell}^{pr}+\delta_{j_\sigma}^{r}n_{j_\sigma}\hat b_{q\ell}^{pi_\sigma}\right)$$
$$+4\delta_{a_{\bar\sigma}}^{p}(1-n_{a_{\bar\sigma}})\delta_{j_\varsigma}^{r}n_{j_\varsigma}\hat b_{q\ell}^{i_\sigma b_{\bar\varsigma}}+4\delta_q^{i_\sigma}n_{i_\sigma}\delta_\ell^{b_{\bar\varsigma}}(1-n_{b_{\bar\varsigma}})\hat b_{a_{\bar\sigma}j_\varsigma}^{pr}-8\delta_{a_{\bar\sigma}}^{p}(1-n_{a_{\bar\sigma}})\delta_\ell^{b_{\bar\varsigma}}(1-n_{b_{\bar\varsigma}})\hat b_{qj_\varsigma}^{i_\sigma r}-8\delta_q^{i_\sigma}n_{i_\sigma}\delta_{j_\varsigma}^{r}n_{j_\varsigma}\hat b_{a_{\bar\sigma}\ell}^{pb_{\bar\varsigma}}$$
$$+4\delta_{a_{\bar\sigma}}^{p}(1-n_{a_{\bar\sigma}})\delta_q^{i_\sigma}n_{i_\sigma}\left(\delta_\ell^{b_{\bar\varsigma}}(1-n_{b_{\bar\varsigma}})\hat b_{j_\varsigma}^{r}-\delta_{j_\varsigma}^{r}n_{j_\varsigma}\hat b_\ell^{b_{\bar\varsigma}}+\delta_{j_\varsigma}^{r}n_{j_\varsigma}\delta_\ell^{b_{\bar\varsigma}}(1-n_{b_{\bar\varsigma}})\right)$$
$$+4\delta_{j_\varsigma}^{r}n_{j_\varsigma}\delta_\ell^{b_{\bar\varsigma}}(1-n_{b_{\bar\varsigma}})\left(\delta_{a_{\bar\sigma}}^{p}(1-n_{a_{\bar\sigma}})\hat b_q^{i_\sigma}-\delta_q^{i_\sigma}n_{i_\sigma}\hat b_{a_{\bar\sigma}}^{p}+\delta_{a_{\bar\sigma}}^{p}(1-n_{a_{\bar\sigma}})\delta_q^{i_\sigma}n_{i_\sigma}\right) \tag{16}$$

which has utilized the symmetry of two-electron integrals. We can then derive for two-electron terms

$$\sum_{\lambda\eta}\left\langle \Psi_{s_\lambda}^{t_{\bar\lambda}}\middle|\left[\hat T^\dagger_{\pm,ai},\hat V_{\rm N},\hat T_{\pm,bj}\right]\middle|\Psi_{s_\eta}^{t_{\bar\eta}}\right\rangle$$
$$=\delta_{j_\sigma}^{i_\sigma}n_{i_\sigma}\left((1-n_{a_{\bar\sigma}})\left(\delta_{t_{\bar\sigma}}^{b_{\bar\sigma}}\langle a_{\bar\sigma}s_\sigma||t_{\bar\sigma}s_\sigma\rangle-\delta_{s_{\bar\sigma}}^{b_{\bar\sigma}}\langle a_{\bar\sigma}t_\sigma||t_\sigma s_{\bar\sigma}\rangle\right)+(1-n_{b_{\bar\sigma}})\left(\delta_{a_{\bar\sigma}}^{t_{\bar\sigma}}\langle s_\sigma t_{\bar\sigma}||s_\sigma b_{\bar\sigma}\rangle-\delta_{a_{\bar\sigma}}^{s_{\bar\sigma}}\langle s_{\bar\sigma}t_\sigma||t_\sigma b_{\bar\sigma}\rangle\right)\right)$$
$$+\delta_{a_{\bar\sigma}}^{b_{\bar\sigma}}(1-n_{a_{\bar\sigma}})\left(n_{j_\sigma}\left(\delta_{s_\sigma}^{i_\sigma}\langle t_{\bar\sigma}j_\sigma||t_{\bar\sigma}s_\sigma\rangle-\delta_{t_\sigma}^{i_\sigma}\langle s_{\bar\sigma}j_\sigma||t_\sigma s_{\bar\sigma}\rangle\right)+n_{i_\sigma}\left(\delta_{j_\sigma}^{s_\sigma}\langle s_\sigma t_{\bar\sigma}||i_\sigma t_{\bar\sigma}\rangle-\delta_{j_\sigma}^{t_\sigma}\langle s_{\bar\sigma}t_\sigma||i_\sigma s_{\bar\sigma}\rangle\right)\right)$$
$$+c_\sigma c_\varsigma(1-n_{a_{\bar\sigma}})n_{j_\varsigma}\left(\delta_{s_\sigma}^{i_\sigma}\delta_{t_{\bar\varsigma}}^{b_{\bar\varsigma}}\langle a_{\bar\sigma}j_\varsigma||t_{\bar\sigma}s_\varsigma\rangle-\delta_{t_\sigma}^{i_\sigma}\delta_{s_{\bar\varsigma}}^{b_{\bar\varsigma}}\langle a_{\bar\sigma}j_\varsigma||t_\varsigma s_{\bar\sigma}\rangle\right)$$
$$+c_\sigma c_\varsigma n_{i_\sigma}(1-n_{b_{\bar\varsigma}})\left(\delta_{a_{\bar\sigma}}^{t_{\bar\sigma}}\delta_{j_\varsigma}^{s_\varsigma}\langle s_\sigma t_{\bar\varsigma}||i_\sigma b_{\bar\varsigma}\rangle-\delta_{j_\varsigma}^{t_\varsigma}\delta_{a_{\bar\sigma}}^{s_{\bar\sigma}}\langle s_{\bar\varsigma}t_\sigma||i_\sigma b_{\bar\varsigma}\rangle\right)$$
$$-2c_\sigma c_\varsigma(1-n_{a_{\bar\sigma}})(1-n_{b_{\bar\varsigma}})\left(\delta_{s_\sigma}^{i_\sigma}\delta_{j_\varsigma}^{s_\varsigma}\langle a_{\bar\sigma}t_{\bar\varsigma}||t_{\bar\sigma}b_{\bar\varsigma}\rangle+\delta_{t_\sigma}^{i_\sigma}\delta_{j_\varsigma}^{t_\varsigma}\langle a_{\bar\sigma}s_{\bar\varsigma}||s_{\bar\sigma}b_{\bar\varsigma}\rangle\right)$$
$$+2(1-n_{a_{\bar\sigma}})(1-n_{b_\sigma})\left(\delta_{s_\sigma}^{i_\sigma}\delta_{j_{\bar\sigma}}^{t_{\bar\sigma}}\langle a_{\bar\sigma}t_\sigma||s_{\bar\sigma}b_\sigma\rangle+\delta_{t_\sigma}^{i_\sigma}\delta_{j_{\bar\sigma}}^{s_{\bar\sigma}}\langle a_{\bar\sigma}s_\sigma||t_{\bar\sigma}b_\sigma\rangle\right)$$
$$-2c_\sigma c_\varsigma n_{i_\sigma}n_{j_\varsigma}\left(\delta_{s_{\bar\varsigma}}^{b_{\bar\varsigma}}\delta_{a_{\bar\sigma}}^{s_{\bar\sigma}}\langle t_\sigma j_\varsigma||i_\sigma t_\varsigma\rangle+\delta_{t_{\bar\varsigma}}^{b_{\bar\varsigma}}\delta_{a_{\bar\sigma}}^{t_{\bar\sigma}}\langle s_\sigma j_\varsigma||i_\sigma s_\varsigma\rangle\right)$$
$$+2n_{i_\sigma}n_{j_{\bar\sigma}}\left(\delta_{s_{\bar\varsigma}}^{b_{\bar\varsigma}}\delta_{a_{\bar\sigma}}^{t_{\bar\sigma}}\langle t_\sigma j_{\bar\sigma}||i_\sigma s_{\bar\sigma}\rangle+\delta_{t_\sigma}^{b_\sigma}\delta_{a_{\bar\sigma}}^{s_{\bar\sigma}}\langle s_\sigma j_{\bar\sigma}||i_\sigma t_{\bar\sigma}\rangle\right)$$
$$+(1-n_{a_{\bar\sigma}})n_{i_\sigma}\left((1-n_{b_{\bar\varsigma}})\left(\delta_{t_\varsigma}^{j_\varsigma}-\delta_{s_\varsigma}^{j_\varsigma}\right)-n_{j_\varsigma}\left(\delta_{t_{\bar\varsigma}}^{b_{\bar\varsigma}}-\delta_{s_{\bar\varsigma}}^{b_{\bar\varsigma}}\right)+n_{j_\varsigma}(1-n_{b_{\bar\varsigma}})\right)\langle a_{\bar\sigma}j_\varsigma||i_\sigma b_{\bar\varsigma}\rangle$$
$$+n_{j_\varsigma}(1-n_{b_{\bar\varsigma}})\left((1-n_{a_{\bar\sigma}})\left(\delta_{t_\sigma}^{i_\sigma}-\delta_{s_\sigma}^{i_\sigma}\right)-n_{i_\sigma}\left(\delta_{t_{\bar\sigma}}^{a_{\bar\sigma}}-\delta_{s_{\bar\sigma}}^{a_{\bar\sigma}}\right)+n_{i_\sigma}(1-n_{a_{\bar\sigma}})\right)\langle a_{\bar\sigma}j_\varsigma||i_\sigma b_{\bar\varsigma}\rangle \tag{17}$$

Here the contractions of two-electron excitations are given from the following patterns

$$\{\hat a_{s_\lambda}^\dagger\hat a_{t_{\bar\lambda}}\}\{\hat a_p^\dagger\hat a_r^\dagger\hat a_\ell\hat a_q\}\{\hat a_{t_{\bar\eta}}^\dagger\hat a_{s_\eta}\};$$
$$\{\hat a_{s_\lambda}^\dagger\hat a_{t_{\bar\lambda}}\}\{\hat a_p^\dagger\hat a_r^\dagger\hat a_\ell\hat a_q\}\{\hat a_{t_{\bar\eta}}^\dagger\hat a_{s_\eta}\};$$
$$\{\hat a_{s_\lambda}^\dagger\hat a_{t_{\bar\lambda}}\}\{\hat a_p^\dagger\hat a_r^\dagger\hat a_\ell\hat a_q\}\{\hat a_{t_{\bar\eta}}^\dagger\hat a_{s_\eta}\};$$
$$\{\hat a_{s_\lambda}^\dagger\hat a_{t_{\bar\lambda}}\}\{\hat a_p^\dagger\hat a_r^\dagger\hat a_\ell\hat a_q\}\{\hat a_{t_{\bar\eta}}^\dagger\hat a_{s_\eta}\};$$

or equivalently

$$
\begin{aligned}
&\left\langle \{\hat{a}^\dagger_{s_\lambda}\hat{a}_{t_{\bar{\lambda}}}\}\{\hat{a}^\dagger_p\hat{a}^\dagger_r\hat{a}_\ell\hat{a}_q\}\{\hat{a}^\dagger_{t_{\bar{\eta}}}\hat{a}_{s_\eta}\}\right\rangle \\
&\quad = \left(\delta^p_{s_\eta}\delta^r_{t_{\bar{\lambda}}} - \delta^p_{t_{\bar{\lambda}}}\delta^r_{s_\eta}\right)\left(\delta^{t_{\bar{\eta}}}_q\delta^{s_\lambda}_\ell - \delta^{s_\lambda}_q\delta^{t_{\bar{\eta}}}_\ell\right). \qquad (18)
\end{aligned}
$$

All in all, the Hamiltonian for excited-states is

$$
\mathbf{I}\Omega^\pm = X^\dagger\Lambda^\pm X. \qquad (19)
$$

D. Electronic Configurations

By expanding the double commutator, we have

$$
\begin{aligned}
\Lambda^\pm = &\left\langle \hat{T}_\pm\Psi^{\mathrm{ref}}\middle|\hat{H}\middle|\hat{T}_\pm\Psi^{\mathrm{ref}}\right\rangle + \left\langle \hat{T}^\dagger_\pm\Psi^{\mathrm{ref}}\middle|\hat{H}\middle|\hat{T}^\dagger_\pm\Psi^{\mathrm{ref}}\right\rangle \\
&- \left\langle \Psi^{\mathrm{ref}}\middle|\hat{H}\middle|\hat{T}^\dagger_\pm\hat{T}_\pm\Psi^{\mathrm{ref}}\right\rangle - \left\langle \hat{T}_\pm\hat{T}^\dagger_\pm\Psi^{\mathrm{ref}}\middle|\hat{H}\middle|\Psi^{\mathrm{ref}}\right\rangle. \qquad (20)
\end{aligned}
$$

First we show the wavefunction configurations to be coupled by the Hamiltonian. For the first two terms, we have the configurations

$$
\begin{aligned}
\left|\hat{T}_{\pm,ai}\Psi^{\mathrm{ref}}\right\rangle &= \frac{1}{\sqrt{2}}\sum_{\sigma\lambda} c_\sigma\hat{a}^\dagger_{a_{\bar{\sigma}}}\hat{a}_{i_\sigma}\hat{a}^\dagger_{t_{\bar{\lambda}}}\hat{a}_{s_\lambda}\left|\Psi_0\right\rangle \\
&= \frac{1}{\sqrt{2}}\sum_{\sigma\lambda} c_\sigma\left(\delta_{i_\sigma t_\sigma}\hat{a}^\dagger_{a_{\bar{\sigma}}}\hat{a}_{s_{\bar{\sigma}}} + \hat{a}^\dagger_{t_{\bar{\lambda}}}\hat{a}^\dagger_{a_{\bar{\sigma}}}\hat{a}_{i_\sigma}\hat{a}_{s_\lambda}\right)\left|\Psi_0\right\rangle \\
&= \frac{1}{\sqrt{2}}\sum_{\sigma\lambda} c_\sigma\left(\delta_{i_\sigma t_\sigma}\left|0^{a_{\bar{\sigma}}}_{s_{\bar{\sigma}}}\right\rangle + \left|0^{a_{\bar{\sigma}}t_{\bar{\lambda}}}_{s_\lambda i_\sigma}\right\rangle\right), \qquad (21) \\
\left|\hat{T}^\dagger_{\pm,ai}\Psi^{\mathrm{ref}}\right\rangle &= \frac{1}{\sqrt{2}}\sum_{\sigma\lambda} c_\sigma\hat{a}^\dagger_{i_\sigma}\hat{a}_{a_{\bar{\sigma}}}\hat{a}^\dagger_{t_{\bar{\lambda}}}\hat{a}_{s_\lambda}\left|\Psi_0\right\rangle \\
&= \frac{1}{\sqrt{2}}\sum_{\sigma\lambda} c_\sigma\left(\delta_{a_{\bar{\sigma}} t_{\bar{\sigma}}}\hat{a}^\dagger_{i_\sigma}\hat{a}_{s_\sigma} + \hat{a}^\dagger_{t_{\bar{\lambda}}}\hat{a}^\dagger_{i_\sigma}\hat{a}_{a_{\bar{\sigma}}}\hat{a}_{s_\lambda}\right)\left|\Psi_0\right\rangle \\
&= \frac{1}{\sqrt{2}}\sum_{\sigma\lambda} c_\sigma\left(\delta_{a_{\bar{\sigma}} t_{\bar{\sigma}}}\left|0^{i_\sigma}_{s_\sigma}\right\rangle + \left|0^{i_\sigma t_{\bar{\lambda}}}_{s_\lambda a_{\bar{\sigma}}}\right\rangle\right), \qquad (22)
\end{aligned}
$$

half of which are illustrated in Fig. 2.

Moreover, we can derive the configurations for the last two terms in Eq. (20) as

$$
\begin{aligned}
&\left|\hat{T}^\dagger_{\pm,ai}\hat{T}_{\pm,bj}\Psi^{\mathrm{ref}}\right\rangle \\
&= \frac{1}{\sqrt{2}}\sum_{\sigma\lambda\varsigma} c_\sigma c_\varsigma\hat{a}^\dagger_{i_\sigma}\hat{a}_{a_{\bar{\sigma}}}\hat{a}^\dagger_{b_{\bar{\varsigma}}}\hat{a}_{j_\varsigma}\hat{a}^\dagger_{t_{\bar{\lambda}}}\hat{a}_{s_\lambda}\left|\Psi_0\right\rangle \\
&= \frac{1}{\sqrt{2}}\sum_{\sigma\lambda}\Big(\delta_{a_{\bar{\sigma}}b_{\bar{\sigma}}}\left(\delta_{j_\sigma t_\sigma}\hat{a}^\dagger_{i_\sigma}\hat{a}_{s_{\bar{\sigma}}} + \hat{a}^\dagger_{t_{\bar{\lambda}}}\hat{a}^\dagger_{i_\sigma}\hat{a}_{j_\sigma}\hat{a}_{s_\lambda}\right) \\
&\qquad + c_\sigma c_{\bar{\lambda}}\delta_{j_{\bar{\lambda}}t_{\bar{\lambda}}}\hat{a}^\dagger_{b_\lambda}\hat{a}^\dagger_{i_\sigma}\hat{a}_{a_{\bar{\sigma}}}\hat{a}_{s_\lambda} + c_\sigma c_\lambda\delta_{a_{\bar{\sigma}}t_{\bar{\sigma}}}\hat{a}^\dagger_{i_\sigma}\hat{a}^\dagger_{b_{\bar{\lambda}}}\hat{a}_{j_\lambda}\hat{a}_{s_\sigma} \\
&\qquad + c_\sigma c_\varsigma\hat{a}^\dagger_{i_\sigma}\hat{a}^\dagger_{t_{\bar{\lambda}}}\hat{a}^\dagger_{b_{\bar{\varsigma}}}\hat{a}_{a_{\bar{\sigma}}}\hat{a}_{j_\varsigma}\hat{a}_{s_\lambda}\Big)\left|\Psi_0\right\rangle \\
&= \frac{1}{\sqrt{2}}\sum_{\sigma\lambda}\Big(\delta_{a_{\bar{\sigma}}b_{\bar{\sigma}}}\left(\delta_{j_\sigma t_\sigma}\left|0^{i_\sigma}_{s_{\bar{\sigma}}}\right\rangle + \left|0^{i_\sigma t_{\bar{\lambda}}}_{s_\lambda j_\sigma}\right\rangle\right) \\
&\qquad + c_\sigma c_{\bar{\lambda}}\delta_{j_{\bar{\lambda}}t_{\bar{\lambda}}}\hat{a}^\dagger\left|0^{i_\sigma b_\lambda}_{s_\lambda a_{\bar{\sigma}}}\right\rangle + c_\sigma c_\lambda\delta_{a_{\bar{\sigma}}t_{\bar{\sigma}}}\left|0^{b_{\bar{\lambda}}i_\sigma}_{s_\sigma j_\lambda}\right\rangle \\
&\qquad + c_\sigma c_\varsigma\left|0^{b_{\bar{\varsigma}}t_{\bar{\lambda}}i_\sigma}_{s_\lambda j_\varsigma a_{\bar{\sigma}}}\right\rangle\Big), \qquad (23) \\
&\left|\hat{T}_{\pm,ai}\hat{T}^\dagger_{\pm,bj}\Psi^{\mathrm{ref}}\right\rangle \\
&= \frac{1}{\sqrt{2}}\sum_{\sigma\lambda\varsigma} c_\sigma c_\varsigma\hat{a}^\dagger_{a_{\bar{\sigma}}}\hat{a}_{i_\sigma}\hat{a}^\dagger_{j_\varsigma}\hat{a}_{b_{\bar{\varsigma}}}\hat{a}^\dagger_{t_{\bar{\lambda}}}\hat{a}_{s_\lambda}\left|\Psi_0\right\rangle \\
&= \frac{1}{\sqrt{2}}\sum_{\sigma\lambda\varsigma}\Big(\delta_{i_\sigma j_\sigma}\left(\delta_{b_{\bar{\sigma}}t_{\bar{\sigma}}}\hat{a}^\dagger_{a_{\bar{\sigma}}}\hat{a}_{s_\sigma} + \hat{a}^\dagger_{t_{\bar{\lambda}}}\hat{a}^\dagger_{a_{\bar{\sigma}}}\hat{a}_{b_{\bar{\sigma}}}\hat{a}_{s_\lambda}\right) \\
&\qquad + c_\sigma c_\lambda\delta_{b_{\bar{\lambda}}t_{\bar{\lambda}}}\hat{a}^\dagger_{j_\lambda}\hat{a}^\dagger_{a_{\bar{\sigma}}}\hat{a}_{i_\sigma}\hat{a}_{s_\lambda} + c_\sigma c_\lambda\delta_{i_\sigma t_\sigma}\hat{a}^\dagger_{a_{\bar{\sigma}}}\hat{a}^\dagger_{j_\lambda}\hat{a}_{b_{\bar{\lambda}}}\hat{a}_{s_{\bar{\sigma}}} \\
&\qquad + c_\sigma c_\varsigma c_{\bar{\lambda}}\hat{a}^\dagger_{a_{\bar{\sigma}}}\hat{a}^\dagger_{t_{\bar{\lambda}}}\hat{a}^\dagger_{j_\varsigma}\hat{a}_{i_\sigma}\hat{a}_{b_{\bar{\varsigma}}}\hat{a}_{s_\lambda}\Big)\left|\Psi_0\right\rangle \\
&= \frac{1}{\sqrt{2}}\sum_{\sigma\lambda\varsigma}\Big(\delta_{i_\sigma j_\sigma}\left(\delta_{b_{\bar{\sigma}}t_{\bar{\sigma}}}\left|0^{a_{\bar{\sigma}}}_{s_\sigma}\right\rangle + \left|0^{a_{\bar{\sigma}}t_{\bar{\lambda}}}_{s_\lambda b_{\bar{\sigma}}}\right\rangle\right) \\
&\qquad + c_\sigma c_\lambda\delta_{b_{\bar{\lambda}}t_{\bar{\lambda}}}\left|0^{a_{\bar{\sigma}}j_\lambda}_{s_\lambda i_\sigma}\right\rangle + c_\sigma c_\lambda\delta_{i_\sigma t_\sigma}\left|0^{j_\lambda a_{\bar{\sigma}}}_{s_{\bar{\sigma}}b_{\bar{\lambda}}}\right\rangle \\
&\qquad + c_\sigma c_\varsigma c_{\bar{\lambda}}\left|0^{j_\varsigma t_{\bar{\lambda}}a_{\bar{\sigma}}}_{s_\lambda b_{\bar{\varsigma}}i_\sigma}\right\rangle\Big), \qquad (24)
\end{aligned}
$$

E. Analytical Gradient

The gradient of the I-th excitation energy from Eq. (19) is

$$
\omega^x_{I,\pm} = X^{I,\dagger}\Lambda^{\pm,x}X^I, \qquad (25)
$$

in where superscript x represents nuclear derivative. Note the derivatives of amplitudes do not contribute due to their self-consistency.

After the amplitudes are solved, on the other hand, the I-th excitation energy can be calculated in AO basis by

$$
\omega_I = \mathrm{tr}\big(\mathbf{P}^I\mathbf{F}\big) + \mathrm{tr}\big(\mathbf{R}^{I,\dagger}\mathbf{G}\mathbf{R}^I\big), \qquad (26)
$$

with the difference density $\mathbf{P}^I$ and transition density $\mathbf{R}^I$.

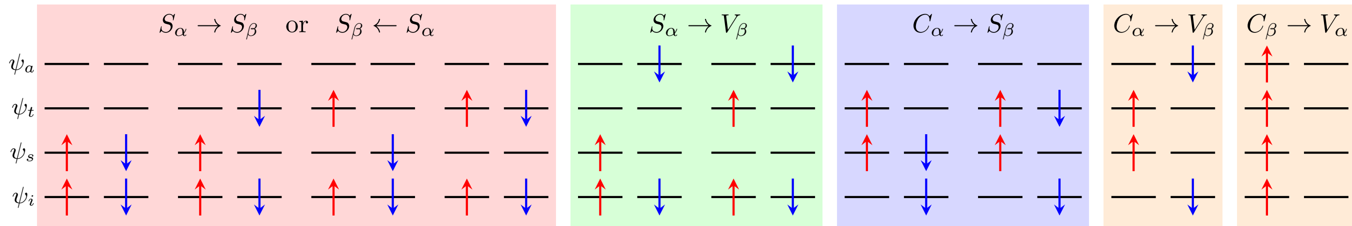

FIG. 2. Schematic electron occupation diagram for one-electron spin-flip excitation and de-excitation configurations (applying $\hat{T}_{\alpha\to\beta}$, $\hat{T}_{\beta\to\alpha}$, $\hat{T}^\dagger_{\alpha\to\beta}$, and $\hat{T}^\dagger_{\beta\to\alpha}$ operators) from $|\Psi_{M_s=+1}\rangle$ reference state of three α and one β electrons. Their spin images (swapping α and β spins) give excitation configurations from $\Psi_{M_s=-1}$ state. The C, S, and V labels doubly occupied, singly occupied, and virtual orbitals, respectively. In particular, S space includes two orbitals in our current study, indexed by s and t from bottom-up. Red and blue arrows indicate spin-up (α) and spin-down (β) electrons, respectively. Right (left) arrows indicate (de-)excitations.

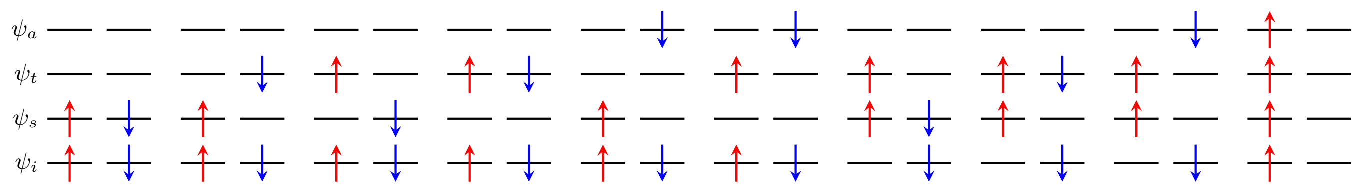

FIG. 3. Schematic electron occupation diagram for one-electron spin-flip double excitation and de-excitation configurations (applying $\hat{T}^\dagger_{\alpha\to\beta}\hat{T}_{\alpha\to\beta}$, $\hat{T}^\dagger_{\alpha\to\beta}\hat{T}_{\beta\to\alpha}$, $\hat{T}^\dagger_{\beta\to\alpha}\hat{T}_{\alpha\to\beta}$, and $\hat{T}^\dagger_{\beta\to\alpha}\hat{T}_{\beta\to\alpha}$ operators) from $|\Psi_{M_s=+1}\rangle$ reference state of three α and one β electrons. Their spin images (swapping α and β spins) give excitation configurations from $\Psi_{M_s=-1}$ state. The C, S, and V labels doubly occupied, singly occupied, and virtual orbitals, respectively. In particular, S space includes two orbitals in our current study, indexed by s and t from bottom-up. Red and blue arrows indicate spin-up (α) and spin-down (β) electrons, respectively. Right (left) arrows indicate (de-)excitations.

F. Derivative Couplings

The derivative coupling between two excited states can be given by the following commutator

$$\left[\hat{T}_I^\dagger, \frac{\partial}{\partial x}, \hat{T}_J\right] = \hat{T}_I^\dagger\left(\frac{\partial}{\partial x}\hat{T}_J\right) - \frac{1}{2}\left(\frac{\partial}{\partial R}\hat{T}_I^\dagger\hat{T}_J\right), \qquad (27)$$

where $\hat{T}$ does not have de-excitation amplitudes in the current formalism. Here $\frac{\partial}{\partial x}$ is the nuclear derivative operator with respective to the nuclear (direction) x, which is equivalent to a one-electron operator. In MO representation, we have

$$\begin{aligned}
\hat{T}_I &= \sum_\sigma \sum_{a_{\bar\sigma} i_\sigma} X^I_{a_{\bar\sigma} i_\sigma} \hat{P}_{a_{\bar\sigma} i_\sigma}, \\
\frac{\partial}{\partial x}\hat{T}_I &= \sum_\sigma \sum_{a_{\bar\sigma} i_\sigma} \left(X^{I,x}_{a_{\bar\sigma} i_\sigma}\hat{P}_{a_{\bar\sigma} i_\sigma} + X^I_{a_{\bar\sigma} i_\sigma}\hat{P}^x_{a_{\bar\sigma} i_\sigma}\right), \\
\frac{\partial}{\partial x}\hat{T}_I^\dagger\hat{T}_J &= \sum_{\sigma\varsigma} \sum_{a_{\bar\sigma} i_\sigma} \sum_{b_{\bar\varsigma} j_\varsigma} \left(\left(X^{I,\dagger}_{a_{\bar\sigma} i_\sigma} X^J_{b_{\bar\varsigma} i_\varsigma}\right)^x \hat{P}_{i_\sigma a_{\bar\sigma}}\hat{P}_{b_{\bar\varsigma} j_\varsigma}\right. \\
&\qquad \left. + X^{I,\dagger}_{a_{\bar\sigma} i_\sigma} X^J_{b_{\bar\varsigma} i_\varsigma}\left(\hat{P}_{i_\sigma a_{\bar\sigma}}\hat{P}_{b_{\bar\varsigma} j_\varsigma}\right)^x\right),
\end{aligned} \qquad (28)$$

with the MO density matrix projector $\hat{P}_{pq} = |\psi_p\rangle\langle\psi_q|$ and its derivative $\hat{P}_{pq} = |\psi_p^x\rangle\langle\psi_q| + |\psi_p\rangle\langle\psi_q^x|$. In particular, the expectations of projectors read $\left\langle\hat{P}_{pq}\right\rangle = \sum_r \langle\psi_r|\psi_p\rangle\langle\psi_q|\psi_r\rangle = \langle\psi_q|\psi_p\rangle = S_{qp}$, $\left\langle\hat{P}^x_{pq}\right\rangle = S^x_{qp} = 0$, $\left\langle\hat{P}_{pq}\hat{P}_{rs}\right\rangle = S_{qr}S_{ps}$, $\left\langle\hat{P}_{pq}\hat{P}^x_{rs}\right\rangle = S_{sp}\langle\psi_q|\psi_r^x\rangle - S_{qr}\langle\psi_s|\psi_p^x\rangle$, and $\left\langle(\hat{P}_{pq}\hat{P}_{rs})^x\right\rangle = 0$. Then the derivative coupling $d_{IJ} = \left\langle\left[\hat{T}_I^\dagger, \frac{\partial}{\partial x}, \hat{T}_J\right]\right\rangle$ with the commutator in Eq. (27) can be derived as

$$\begin{aligned}
d_{IJ} = &\frac{1}{2}\sum_\sigma \sum_{i_\sigma a_{\bar\sigma}}\left(X^I_{a_{\bar\sigma} i_\sigma}X^{J,x}_{a_{\bar\sigma} i_\sigma} - X^{I,x}_{a_{\bar\sigma} i_\sigma}X^J_{a_{\bar\sigma} i_\sigma}\right) \\
&+ \sum_\sigma \sum_{a_{\bar\sigma} b_{\bar\sigma}} \left(\mathbf{X}^J\mathbf{X}^{I,\dagger}\right)_{b_{\bar\sigma} a_{\bar\sigma}} \langle\psi_{a_{\bar\sigma}}|\psi^x_{b_{\bar\sigma}}\rangle \\
&- \sum_\sigma \sum_{i_\sigma j_\sigma} \left(\mathbf{X}^{I,\dagger}\mathbf{X}^J\right)_{i_\sigma j_\sigma} \langle\psi_{j_\sigma}|\psi^x_{i_\sigma}\rangle .
\end{aligned} \qquad (29)$$

In parallel, finite-difference of state overlaps are often used because it is easy to implement in different software. The naive calculation of linear response excited-state overlap, however, requires a four-loop procedure in the occupied and virtual molecular orbital spaces. For TDDFT states, for instance, the overlap reads

$$S_{IJ} = \sum_{ai,bj}\sum_\sigma \left(X^{I,\dagger}_{ai}X^J_{bj}\det\left(\tilde{\mathbf{S}}^{ab}_{ij}\right) - Y^{I,\dagger}_{ai}Y^J_{bj}\det\left(\tilde{\mathbf{S}}^{ba}_{ji}\right)\right). \qquad (30)$$

for same-spin excitation configurations, where one has to calculate the determinant of MO overlap matrix for each

a, i and b, j pairs. Such computational bottleneck can be alleviated by the matrix determinant lemma, which gives the simplified equation involve ground-state MO overlap matrix determinant with a prefactor from one matrix inverse and matrix multiplications. As derived in Appendix B, the above equation becomes

$$S_{IJ} = \det(\mathbf{S}_{\text{oo}}) \left(\operatorname{tr}\left(\mathbf{X}^{I,\dagger}\mathbf{S}_{\text{vo}}\mathbf{S}_{\text{oo}}^{-1}\right) \operatorname{tr}\left(\mathbf{X}^{J}\mathbf{S}_{\text{oo}}^{-1}\mathbf{S}_{\text{ov}}\right) + \operatorname{tr}\left(\mathbf{S}_{\text{oo}}^{-1}\mathbf{X}^{I,\dagger}\left(\mathbf{S}_{\text{vv}} - \mathbf{S}_{\text{vo}}\mathbf{S}_{\text{oo}}^{-1}\mathbf{S}_{\text{ov}}\right)\mathbf{X}^{J}\right)\right), \tag{31}$$

for X part.

G. Spin Square

The spin square operator is $\hat{\vec{S}}^2 = \hat{S}_z^2 + \hat{S}_z + \hat{S}_-\hat{S}_+$ where $\hat{\vec{S}} = \sum_p \hat{\vec{s}}_p$ sums over each spin operator, which in second-quantization reads

$$\begin{aligned} \vec{s}_p &= \left(\hat{a}^\dagger_{p_\alpha},\ \hat{a}^\dagger_{p_\beta}\right)\vec{\sigma}\begin{pmatrix}\hat{a}_{p_\alpha}\\ \hat{a}_{p_\beta}\end{pmatrix} \\ &= \left(\hat{a}^\dagger_{p_\alpha}\hat{a}_{p_\beta},\ \hat{a}^\dagger_{p_\beta}\hat{a}_{p_\alpha},\ \textstyle\sum_\sigma s_\sigma \hat{a}^\dagger_{p_\sigma}\hat{a}_{p_\sigma}\right) \end{aligned} \tag{32}$$

where $\vec{\sigma} = (\sigma_+, \sigma_-, \sigma_z)$ is the Pauli matrices, and $s_\alpha = \frac{1}{2}$ and $s_\beta = -\frac{1}{2}$ as the spins of α and β electron, respectively. Since the spin lowering and raising operators involve different orbitals, they should be written as in general

$$\hat{S}_+ = \sum_{pq} S^{\alpha\beta}_{pq}\hat{a}^\dagger_{p_\alpha}\hat{a}_{q_\beta}, \quad \hat{S}_- = \sum_{pq} S^{\beta\alpha}_{qp}\hat{a}^\dagger_{q_\beta}\hat{a}_{p_\alpha}, \tag{33}$$

respectively, where $S^{\alpha\beta}_{pq} = \left\langle \psi^\alpha_p \middle| \psi^\beta_q \right\rangle$ as the orbital overlap, which would be orthonormal in restricted case. Then we can write the operator

$$\begin{aligned} \hat{\vec{S}}^2 = &\sum_p \sum_\lambda \left(s_\lambda + \frac{1}{4}\right)\hat{a}^\dagger_{p_\lambda}\hat{a}_{p_\lambda} + \sum_{q\ell} D^{\beta\beta}_{q\ell}\hat{a}^\dagger_{q_\beta}\hat{a}_{\ell_\beta} \\ &+ \sum_{pq}\sum_{\lambda\eta} s_\lambda s_\eta \hat{a}^\dagger_{p_\lambda}\hat{a}^\dagger_{q_\eta}\hat{a}_{q_\eta}\hat{a}_{p_\lambda} \\ &- \sum_{pq}\sum_{r\ell} S^{\beta\alpha}_{qp}S^{\alpha\beta}_{r\ell}\hat{a}^\dagger_{q_\beta}\hat{a}^\dagger_{r_\alpha}\hat{a}_{p_\alpha}\hat{a}_{\ell_\beta}, \end{aligned} \tag{34}$$

with $D^{\beta\beta}_{q\ell} = \sum_p S^{\beta\alpha}_{qp}S^{\alpha\beta}_{p\ell}(1 - n_{p_\alpha})$. The I-th MR-SF excited-state reads

$$\left|\Psi^I\right\rangle = \sum_{ai}\sum_\sigma X^I_{ai}\hat{b}^{a_{\bar{\sigma}}}_{i_\sigma}\left|\Psi_0\right\rangle, \tag{35}$$

Then the fully contracted patters for the S^2 value are

$$\left\langle \hat{b}^{i_\sigma}_{a_{\bar{\sigma}}}\hat{\vec{S}}^2\hat{b}^{b_{\bar{\varsigma}}}_{j_\varsigma}\right\rangle = \tag{36}$$

$$
\begin{aligned}
\left[\hat{b}^{i_\sigma}_{a_{\bar\sigma}}, \hat{\vec{S}}^2, \hat{b}^{b_{\bar\varsigma}}_{j_\varsigma}\right] = {}& \delta_{\sigma\varsigma} \left(\left(\frac{1}{4} - s_\sigma - \frac{1}{2} n_{i_\sigma}\right)(1 - n_{a_{\bar\sigma}}) - \left(\frac{1}{4} + s_\sigma\right) n_{i_\sigma} \right) \left(\delta^{i_\sigma}_{j_\sigma} n_{i_\sigma} \delta^{b_{\bar\sigma}}_{a_{\bar\sigma}} (1 - n_{a_{\bar\sigma}}) + \delta^{b_{\bar\sigma}}_{a_{\bar\sigma}} (1 - n_{a_{\bar\sigma}}) \hat{b}^{i_\sigma}_{j_\sigma} - \delta^{i_\sigma}_{j_\sigma} n_{i_\sigma} \hat{b}^{b_{\bar\sigma}}_{a_{\bar\sigma}} \right) \\
& + \sum_\ell \delta_{\sigma\alpha} \delta_{\varsigma\alpha} (1 - n_{a_\beta}) D^{\beta\beta}_{a\ell} \left(\delta^{i_\alpha}_{j_\alpha} n_{j_\alpha} \delta^{b_\beta}_{\ell_\beta} (1 - n_{b_\beta}) + \delta^{b_\beta}_{\ell_\beta} (1 - n_{b_\beta}) \hat{b}^{i_\alpha}_{j_\alpha} - \delta^{i_\alpha}_{j_\alpha} n_{j_\alpha} \hat{b}^{b_\beta}_{\ell_\beta} \right) \\
& - \sum_q \delta_{\sigma\beta} \delta_{\varsigma\beta} n_{i_\beta} D^{\beta\beta}_{qi} \left(\delta^{q_\beta}_{j_\beta} n_{j_\beta} \delta^{b_\alpha}_{a_\alpha} (1 - n_{b_\alpha}) + \delta^{b_\alpha}_{a_\alpha} (1 - n_{b_\alpha}) \hat{b}^{q_\beta}_{j_\beta} - \delta^{q_\beta}_{j_\beta} n_{j_\beta} \hat{b}^{b_\alpha}_{a_\alpha} \right) \\
& + 2 s_{\bar\sigma} \delta_{\sigma\varsigma} (1 - n_{a_{\bar\sigma}} + n_{i_\sigma}) \left(s_{\bar\sigma} n_{i_\sigma} (1 - n_{a_{\bar\sigma}}) \left(\delta^{i_\sigma}_{j_\sigma} \hat{b}^{b_{\bar\sigma}}_{a_{\bar\sigma}} - \delta^{b_{\bar\sigma}}_{a_{\bar\sigma}} \hat{b}^{i_\sigma}_{j_\sigma} \right) \right. \\
& \qquad \left. + \sum_q \sum_\eta s_\eta \left(\delta_{\sigma\varsigma} \delta^{i_\sigma}_{j_\sigma} n_{i_\sigma} \delta^{b_{\bar\sigma}}_{a_{\bar\sigma}} (1 - n_{a_{\bar\sigma}}) \hat{b}^{q_\eta}_{q_\eta} + \delta^{b_{\bar\sigma}}_{a_{\bar\sigma}} (1 - n_{a_{\bar\sigma}}) \hat{b}^{i_\sigma q_\eta}_{j_\sigma q_\eta} - \delta^{i_\sigma}_{j_\sigma} n_{i_\sigma} \hat{b}^{b_{\bar\sigma} q_\eta}_{a_{\bar\sigma} q_\eta} \right) \right) \\
& + 2 s_{\bar\sigma} s_\varsigma (1 - n_{a_{\bar\sigma}} + n_{i_\sigma})(1 - n_{b_{\bar\varsigma}} + n_{j_\varsigma}) \hat{b}^{i_\sigma b_{\bar\varsigma}}_{j_\varsigma a_{\bar\sigma}} \\
& - \delta_{\sigma\varsigma} S^{\bar\sigma\sigma}_{a_{\bar\sigma} i_\sigma} n_{i_\sigma} (1 - n_{a_{\bar\sigma}}) \left(S^{\sigma\bar\sigma}_{j_\sigma b_{\bar\sigma}} + \sum_r S^{\sigma\bar\sigma}_{r b_{\bar\sigma}} (1 - n_{b_{\bar\sigma}}) \hat{b}^r_{j_\sigma} - \sum_\ell S^{\sigma\bar\sigma}_{j_\sigma \ell} n_{j_\sigma} \hat{b}^{b_{\bar\sigma}}_\ell \right) \\
& + 2 s_\sigma \delta_{\sigma\bar\varsigma} \sum_{r\ell} S^{\sigma\bar\sigma}_{r\ell} \left(S^{\bar\sigma\sigma}_{j_{\bar\sigma} i_\sigma} n_{i_\sigma} n_{j_{\bar\sigma}} \hat{b}^{b_\sigma r_\sigma}_{\ell_{\bar\sigma} a_{\bar\sigma}} + S^{\bar\sigma\sigma}_{a_{\bar\sigma} b_\sigma} (1 - n_{a_{\bar\sigma}})(1 - n_{b_\sigma}) \hat{b}^{i_\sigma r_\sigma}_{\ell_{\bar\sigma} j_{\bar\sigma}} \right) \\
& - 2 s_\sigma \delta_{\sigma\varsigma} n_{i_\sigma} \sum_q S^{\bar\sigma\sigma}_{q i_\sigma} \left(n_{j_\sigma} (1 - n_{b_{\bar\sigma}}) \left(S^{\sigma\bar\sigma}_{j_\sigma b_{\bar\sigma}} \hat{b}^{q_{\bar\sigma}}_{a_{\bar\sigma}} + \sum_\ell \delta^{b_{\bar\sigma}}_{a_{\bar\sigma}} S^{\sigma\bar\sigma}_{j_\sigma \ell} \hat{b}^{q_{\bar\sigma}}_{\ell_{\bar\sigma}} \right) \right. \\
& \qquad \left. + \sum_\ell S^{\sigma\bar\sigma}_{j_\sigma \ell} n_{j_\sigma} \hat{b}^{q_{\bar\sigma} b_{\bar\sigma}}_{\ell_{\bar\sigma} a_{\bar\sigma}} + (1 - n_{b_{\bar\sigma}}) \left(\sum_r S^{\sigma\bar\sigma}_{r b_{\bar\sigma}} \hat{b}^{q_{\bar\sigma} r_\sigma}_{j_\sigma a_{\bar\sigma}} - \sum_{r\ell} \delta^{b_{\bar\sigma}}_{a_{\bar\sigma}} S^{\sigma\bar\sigma}_{r\ell} \hat{b}^{q_{\bar\sigma} r_\sigma}_{\ell_{\bar\sigma} j_\sigma} \right) \right) \\
& - 2 s_\sigma \delta_{\sigma\varsigma} (1 - n_{a_{\bar\sigma}}) \sum_p S^{\bar\sigma\sigma}_{a_{\bar\sigma} p} \left(n_{j_\sigma} (1 - n_{b_{\bar\sigma}}) \left(S^{\sigma\bar\sigma}_{j_\sigma b_{\bar\sigma}} \hat{b}^{i_\sigma}_{p_\sigma} + \sum_r \delta^{i_\sigma}_{j_\sigma} S^{\sigma\bar\sigma}_{r b_{\bar\sigma}} \hat{b}^{r_\sigma}_{p_\sigma} \right) \right. \\
& \qquad \left. + \sum_r S^{\sigma\bar\sigma}_{r b_{\bar\sigma}} (1 - n_{b_{\bar\sigma}}) \hat{b}^{i_\sigma r_\sigma}_{j_\sigma p_\sigma} + n_{j_\sigma} \left(\sum_\ell S^{\sigma\bar\sigma}_{j_\sigma \ell} \hat{b}^{i_\sigma b_{\bar\sigma}}_{\ell_{\bar\sigma} p_\sigma} - \sum_{r\ell} \delta^{i_\sigma}_{j_\sigma} S^{\sigma\bar\sigma}_{r\ell} \hat{b}^{b_{\bar\sigma} r_\sigma}_{\ell_{\bar\sigma} p_\sigma} \right) \right)
\end{aligned}
\tag{37}
$$

Its expectation value on the MRSF-DFT excited-states becomes

$$
\sum_{ai,bj} \sum_{\sigma\varsigma} X^I_{ai} X^I_{bj} \left\langle \Psi^{\mathrm{ref}} \middle| \left[\hat{b}^{i_\sigma}_{a_{\bar\sigma}}, \hat{\vec{S}}^2, \hat{b}^{b_{\bar\varsigma}}_{j_\varsigma}\right] \middle| \Psi^{\mathrm{ref}} \right\rangle = \tag{38}
$$

III. RESULTS AND DISCUSSION

IV. CONCLUSION

Conclusion goes here.

Appendix A: Commutations of Operator Strings

The commutators among spin-conserving and spin-flipping single-excitation operators are

$$
\begin{aligned}
[\hat{b}^{p_\sigma}_{q_\sigma}, \hat{b}^{r_\varsigma}_{\ell_\varsigma}] &= \delta_{\sigma\varsigma} \left(\delta^{p_\sigma}_{\ell_\sigma} n_{p_\sigma} \delta^{r_\sigma}_{q_\sigma} (1 - n_{r_\sigma}) \right. \\
& \qquad \left. + \delta^{r_\sigma}_{q_\sigma} (1 - n_{r_\sigma}) \hat{a}^\dagger_{p_\sigma} \hat{a}_{\ell_\sigma} - \delta^{p_\sigma}_{\ell_\sigma} n_{p_\sigma} \hat{a}^\dagger_{r_\sigma} \hat{a}_{q_\sigma} \right), \\
[\hat{b}^{p_\sigma}_{q_{\bar\sigma}}, \hat{b}^{r_\varsigma}_{\ell_{\bar\varsigma}}] &= \delta_{\sigma\bar\varsigma} \left(\delta^{p_\sigma}_{\ell_\sigma} n_{p_\sigma} \delta^{r_{\bar\sigma}}_{q_{\bar\sigma}} (1 - n_{r_{\bar\sigma}}) \right. \\
& \qquad \left. + \delta^{r_{\bar\sigma}}_{q_{\bar\sigma}} (1 - n_{r_{\bar\sigma}}) \hat{a}^\dagger_{p_\sigma} \hat{a}_{\ell_\sigma} - \delta^{p_\sigma}_{\ell_\sigma} n_{p_\sigma} \hat{a}^\dagger_{r_{\bar\sigma}} \hat{a}_{q_{\bar\sigma}} \right), \\
[\hat{b}^{p_\sigma}_{q_\sigma}, \hat{b}^{r_\varsigma}_{\ell_{\bar\varsigma}}] &= \delta_{\sigma\varsigma} \delta^{r_\sigma}_{q_\sigma} (1 - n_{r_\sigma}) \hat{a}^\dagger_{p_\sigma} \hat{a}_{\ell_{\bar\sigma}} - \delta_{\sigma\bar\varsigma} \delta^{p_\sigma}_{\ell_\sigma} n_{p_\sigma} \hat{a}^\dagger_{r_{\bar\sigma}} \hat{a}_{q_\sigma}.
\end{aligned}
\tag{A1}
$$

Appendix B: Excited-State Overlap Calculations

The overlap between two normal TDA excited-states I and J at two different geometries reads

$$S_{IJ} = \sum_{ai,bj} \sum_{\sigma} \Big(X_{ai}^{I,\sigma,\dagger} X_{bj}^{J,\sigma} \left\langle \Psi_0 | \Psi_0' \right\rangle^{\bar{\sigma}} \left\langle \Psi_i^a \middle| \Psi_j'^b \right\rangle^{\sigma} + X_{ai}^{I,\sigma,\dagger} X_{bj}^{J,\bar{\sigma}} \left\langle \Psi_i^a | \Psi_0' \right\rangle^{\sigma} \left\langle \Psi_0 \middle| \Psi_j'^b \right\rangle^{\bar{\sigma}} \Big). \quad \text{(B1)}$$

Here the overlaps of electron configurations are given by the determinants of MO overlaps: $\langle \Psi_0 | \Psi_0' \rangle = \det(\mathbf{S}_{\text{oo}})$, $\langle \Psi_i^a | \Psi_0' \rangle = \det(\mathbf{S}_{\text{o}'\text{o}})$, $\left\langle \Psi_0 \middle| \Psi_j'^b \right\rangle = \det(\mathbf{S}_{\text{oo}''})$, and $\left\langle \Psi_i^a \middle| \Psi_j'^b \right\rangle = \det(\mathbf{S}_{\text{o}'\text{o}''})$, with $\mathbf{S} = \mathbf{C}^\dagger \mathbf{S}_{\text{AO}}^{\text{inter}} \mathbf{C}'$.

The overlap calculations shown above involve explicit i, a and j, b loops, which can be very time-consuming. One can apply tricks from determinant properties. The *Cramer's rule* claims the relation $\det\left(\tilde{\mathbf{S}}_j^b\right) = \det(\mathbf{S}) \mathbf{v}_j^b$ where $\mathbf{S} \equiv \mathbf{S}_{\text{oo}}$ is the original overlap matrix between occupied orbitals, while in $\tilde{\mathbf{S}}_j^b$, the j-th column is replaced by the overlap elements from occupied and b-th orbitals. Here $\mathbf{v}_j^b$ is element of the solution of linear equation $\mathbf{S}_{\text{oo}} \mathbf{v} = \mathbf{S}_{\text{vo}}$. This is a particular situation of the *matrix determinant lemma*, which states

$$\det\left(\mathbf{A} + \mathbf{u}\mathbf{v}^\dagger\right) = \left(1 + \mathbf{v}^\dagger \mathbf{A}^{-1} \mathbf{u}\right) \det(\mathbf{A}), \quad \text{(B2)}$$

for the perturbation $\mathbf{u}\mathbf{v}^\dagger$ on matrix $\mathbf{A}$ with $\mathbf{u}$ and $\mathbf{v}$ are both column vectors.

Rank-1 perturbation. The row replacement $(\tilde{\mathbf{S}}_{\text{o}'\text{o}})_i^a$ gives the i-th row differences of $\mathbf{S}_{\text{oo}}$ as $\mathbf{v}_a^\dagger = (\mathbf{S}_{\text{vo}})_{a:} - (\mathbf{S}_{\text{oo}})_{i:}$, where : represents the column dimension. Then $\det\left(\tilde{\mathbf{S}}_i^a\right) = \det\left(\mathbf{S}_{\text{oo}} + \mathbf{e}_i \mathbf{v}_a^\dagger\right) = \left(1 + \mathbf{v}_a^\dagger \mathbf{S}_{\text{oo}}^{-1} \mathbf{e}_i\right) \det(\mathbf{S}_{\text{oo}}) = \left((\mathbf{S}_{\text{vo}})_{a:} \mathbf{S}_{\text{oo}}^{-1} \mathbf{e}_i\right) \det(\mathbf{S}_{\text{oo}})$, with matrix base-vector $\mathbf{e}_i$ where only the i-th element is 1. In a compact matrix form $\mathcal{S}_{\text{vo}} = \det(\mathbf{S}_{\text{oo}}) \mathbf{S}_{\text{vo}} S_{\text{oo}}^{-1}$ containing $\det\left(\tilde{\mathbf{S}}_i^a\right)$ as its element, so that the X amplitude contribution to S_{m0} in Eq. (**??**) is $\operatorname{tr}\left(\mathbf{X}^{m,\dagger} \mathbf{S}_{\text{vo}} \mathbf{S}_{\text{oo}}^{-1}\right) \operatorname{tr}(\mathbf{S}_{\text{oo}})$ and similarly for Y terms.

Rank-2 perturbation. For $\tilde{\mathbf{S}}_{ij}^{ab}$ matrix is obtained by replacing one row and one column together. Define the row difference as $\mathbf{v}_a^\dagger = (\mathbf{S}_{\text{vo}})_{a:} - (\mathbf{S}_{\text{oo}})_{i:}$, and the column difference as $\mathbf{u}_b = (\mathbf{S}_{\text{ov}})_{:b} - (\mathbf{S}_{\text{oo}})_{:j}$. Notice that the element $\left(\tilde{\mathbf{S}}_{ij}^{ab}\right)_{ij} = (\mathbf{S}_{\text{vv}})_{ab}$ as second-order replacement, we have the total difference as $\tilde{\mathbf{S}}_{ij}^{ab} - \mathbf{S}_{\text{oo}} = \mathbf{e}_i \mathbf{v}_a^\dagger + \mathbf{u}_b \mathbf{e}_j^\dagger + r_{ij} \mathbf{e}_i \mathbf{e}_j^\dagger$ with the number $r_{ij} = (\mathbf{S}_{\text{vv}})_{ab} - (\mathbf{S}_{\text{vo}})_{aj} - (\mathbf{S}_{\text{ov}})_{ib} + (\mathbf{S}_{\text{oo}})_{ij}$. Furthermore, define $\tilde{\mathbf{v}}_a = \mathbf{v}_a + r_{ij} \mathbf{e}_j$, and $\mathbf{U} = [\mathbf{e}_i, \mathbf{u}_b]$ and $\mathbf{V} = [\tilde{\mathbf{v}}_a, \mathbf{e}_j]$, we have $\det\left(\tilde{\mathbf{S}}_{ij}^{ab}\right) = \det\left(\mathbf{S}_{\text{oo}} + \mathbf{U}\mathbf{V}^\dagger\right) = \det\left(\mathbf{I}_2 + \mathbf{V}^\dagger \mathbf{S}_{\text{oo}}^{-1} \mathbf{U}\right) \det(\mathbf{S}_{\text{oo}})$. Here the first determinant reads

$$\begin{aligned}
\det\left(\mathbf{I}_2 + \mathbf{V}^\dagger \mathbf{S}_{\text{oo}}^{-1} \mathbf{U}\right) &= \det \begin{pmatrix} 1 + \tilde{\mathbf{v}}_a^\dagger \mathbf{S}_{\text{oo}}^{-1} \mathbf{e}_i & \tilde{\mathbf{v}}_a^\dagger \mathbf{S}_{\text{oo}}^{-1} \mathbf{u}_b \\ \mathbf{e}_j^\dagger \mathbf{S}_{\text{oo}}^{-1} \mathbf{e}_i & 1 + \mathbf{e}_j^\dagger \mathbf{S}_{\text{oo}}^{-1} \mathbf{u}_b \end{pmatrix} \\
&= \det \begin{pmatrix} (\mathbf{S}_{\text{vo}})_{a:} \mathbf{S}_{\text{oo}}^{-1} \mathbf{e}_i + r_{ij}^\dagger \left(\mathbf{S}_{\text{oo}}^{-1}\right)_{ji} & \mathbf{v}_a^\dagger \mathbf{S}_{\text{oo}}^{-1} \mathbf{u}_b + r_{ij}^\dagger \mathbf{e}_j^\dagger \mathbf{S}_{\text{oo}}^{-1} \mathbf{u}_b \\ \left(\mathbf{S}_{\text{oo}}^{-1}\right)_{ji} & \mathbf{e}_j^\dagger \mathbf{S}_{\text{oo}}^{-1} (\mathbf{S}_{\text{ov}})_{:b} \end{pmatrix} \\
&= \left((\mathbf{S}_{\text{vo}})_{a:} \mathbf{S}_{\text{oo}}^{-1} \mathbf{e}_i\right) \left(\mathbf{e}_j^\dagger \mathbf{S}_{\text{oo}}^{-1} (\mathbf{S}_{\text{ov}})_{:b}\right) \\
&\quad + \left(\mathbf{S}_{\text{oo}}^{-1}\right)_{ji} \left((\mathbf{S}_{\text{vv}})_{ab} - (\mathbf{S}_{\text{vo}})_{a:} \mathbf{S}_{\text{oo}}^{-1} (\mathbf{S}_{\text{ov}})_{:b}\right).
\end{aligned} \quad \text{(B3)}$$

Then the overlap S_{mn}^s between same-spin excited-state configurations in Eq. (**??**) from X amplitude is calculated by $\det(\mathbf{S}_{\text{oo}}) \left(\operatorname{tr}\left(\mathbf{X}^{m,\dagger} \mathbf{S}_{\text{vo}} \mathbf{S}_{\text{oo}}^{-1}\right) \operatorname{tr}\left(\mathbf{X}^n \mathbf{S}_{\text{oo}}^{-1} \mathbf{S}_{\text{ov}}\right) + \operatorname{tr}\left(\mathbf{S}_{\text{oo}}^{-1} \mathbf{X}^{m,\dagger} \left(\mathbf{S}_{\text{vv}} - \mathbf{S}_{\text{vo}}\mathbf{S}\right.\right.\right.$ Here the first term is same as the different-spin overlaps.

[1] T. Kowalczyk, T. Tsuchimochi, P.-T. Chen, L. Top, and T. Van Voorhis, Excitation energies and stokes shifts from a restricted open-shell kohn-sham approach, The Journal of Chemical Physics **138**, 10.1063/1.4801790 (2013).

[2] L. Zhao and E. Neuscamman, Density functional extension to excited-state mean-field theory, Journal of Chemical Theory and Computation **16**, 164–178 (2019).

[3] D. J. ROWE, Equations-of-motion method and the extended shell model, Reviews of Modern Physics **40**, 153–166 (1968).

[4] W. Kutzelnigg and D. Mukherjee, Normal order and extended wick theorem for a multiconfiguration reference wave function, The Journal of Chemical Physics **107**, 432–449 (1997).

[5] A. M. Tokmachev, Separation of metric in wick's theorem, The Journal of Chemical Physics **159**, 10.1063/5.0174998 (2023).