Analytic derivative couplings in time-dependent density functional theory: Quadratic response theory versus pseudo-wavefunction approach

Xing Zhang and John M. Herbert

Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, USA

(Received 10 November 2014; accepted 22 January 2015; published online 11 February 2015)

We revisit the formalism for analytic derivative couplings between excited states in time-dependent density functional theory (TDDFT). We derive and implement these couplings using quadratic response theory, then numerically compare this response-theory formulation to couplings implemented previously based on a pseudo-wavefunction formalism and direct differentiation of the Kohn-Sham determinant. Numerical results, including comparison to full configuration interaction calculations, suggest that the two approaches perform equally well for many molecular systems, provided that the underlying DFT method affords accurate potential energy surfaces. The response contributions are found to be important for certain systems with high symmetry, but can be calculated with only a moderate increase in computational cost beyond what is required for the pseudo-wavefunction approach. In the case of spin-flip TDDFT, we provide a formal proof that the derivative couplings obtained using response theory are identical to those obtained from the pseudo-wavefunction formulation, which validates our previous implementation based on the latter formalism. © 2015 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4907376]

I. INTRODUCTION

The Born-Oppenheimer approximation breaks down when the energy gap between electronic states becomes small, where electronic and nuclear degrees of freedom are strongly coupled and nuclear motions can induce electronic transitions. Nonadiabatic dynamics methods can be applied to go beyond the Born-Oppenheimer approximation and describe these non-radiative transitions.1 and in these methods, transition probabilities between electronic states are governed by the first-order nonadiabatic coupling matrix elements (derivative couplings). Derivative couplings are also important for methods designed to locate minimum-energy crossing points along conical seams.2,3 Such methods are useful to investigate photochemical processes in cases where ab initio nonadiabatic dynamics simulations are not affordable.

Analytic formulations of the derivative couplings can be obtained in straightforward fashion for waveform-based methods, via direct differentiation of the electronic wavefunctions with respect to the nuclear coordinates. Derivative couplings for multireference configuration interaction4–7 (MRCI) and equation-of-motion coupled-cluster theory8,9 have been derived in this way. However, these are computationally expensive electronic structure models that can only be applied to small systems. Recently, analytic derivative couplings have also been implemented for the configuration-interaction singles (CIS) method,3,10 which is computationally inexpensive but fails to provide even a qualitatively correct description in many cases, owing to lack of dynamical correlation. Time-dependent density functional theory (TDDFT) is another inexpensive ab initio method for excited states, which often provides reasonable excited-state properties at a cost comparable to CIS.

The development of analytic derivative couplings for TDDFT is therefore important insofar as nonadiabatic ab initio molecular dynamics methods based on TDDFT may be efficient and accurate enough for large molecules. The most popular implementation of TDDFT is the version based on linear response (LR) of the electron density or density matrix for the noninteracting Kohn-Sham (KS) reference system.12–14 As such, the electronic wavefunction is not defined in LR-TDDFT, which prevents the use of direct differentiation to calculate the derivative couplings. Development of the formalism for TDDFT derivative couplings has been quite active recently,3,15–25 but among these developments, only Send and Furche19 and Li et al.20,21 have put forward consistent formulations of TDDFT derivative couplings by using response theory exclusively. These formulations capture all of the “Pulay terms” arising from atom-centered basis functions. Send and Furche19 have implemented their formalism to obtain the TDDFT derivative couplings between the ground and excited states, whereas Li and Liu20,21 presented a conceptual derivation (but no implementation) of the TDDFT derivative couplings between excited states. Very recently (while the present work was under review), the first implementation was reported.21

It is known that TDDFT fails to provide the correct dimensionality of the branching space for conical intersections that involve the reference state (which is usually the ground state),26 which is caused by the imbalanced treatment of ground- versus excited-state electron correlation. No such topological issue exists for conical intersections between excited states.26 As such, there is merit in implementing formally exact analytic derivative couplings between TDDFT excited states that are derived solely from quadratic response theory.
In this work, we derive and implement the derivative couplings between TDDFT excited states based on quadratic response theory. Numerical examples will compare these properties to those derived based on a pseudo-wavefunction approach\textsuperscript{3,25,27} (PWA), in which one treats the KS determinant as a many-electron wavefunction and computes analytic derivative couplings by direct differentiation.\textsuperscript{3,25,27} Finally, we show that for spin-flip TDDFT,\textsuperscript{28} the PWA formalism for the derivative couplings is formally equivalent to quadratic response theory. This validates our recent implementation of the spin-flip TDDFT derivative couplings.\textsuperscript{3,25,27}

II. THEORY

The following notation is used throughout this work. Occupied and virtual KS orbitals are labeled \( \phi_{i,\alpha}, \phi_{i,\beta}, \phi_{g,\alpha}, \phi_{g,\beta}, \phi_{v,\alpha}, \phi_{v,\beta}, \ldots \) and \( \phi_{a,\alpha}, \phi_{a,\beta}, \phi_{a',\alpha}, \phi_{a',\beta}, \ldots \), respectively, whereas \( \phi_{p,\alpha}, \phi_{p,\beta}, \phi_{r,\alpha}, \phi_{r,\beta}, \ldots \) index arbitrary (occupied or virtual) KS orbitals. Greek letters \( \mu, \nu, \lambda, \sigma, \ldots \) index atomic orbitals. All two-electron integrals will be written in physicists’ notation.

A. Analytic derivative couplings between TDDFT excited states

In this section, we present a compact derivation of the analytic formulation of derivative couplings between two TDDFT excited states, based on the density matrix response theory. Similar derivations have been given previously by Send and Furche\textsuperscript{19} and by Li and Liu.\textsuperscript{20}

1. Quadratic response functions for exact states

The derivative coupling between two exact electronically excited states \( |I\rangle \) and \( |J\rangle \) is

\[
\mathbf{d}_{IJ} = \langle I| \hat{\nabla}_R |J\rangle = \frac{\langle I| \hat{\nabla}_R \hat{H} |J\rangle}{E_J - E_I},
\]

where \( \hat{H} \) is the electronic Hamiltonian, \( |J\rangle \) and \( |I\rangle \) are the occupied KS orbitals, and \( \hat{\nabla}_R \) is the nuclear derivative operator, respectively.

For any time-independent operator \( \hat{A} \), the transition properties \( \langle I| \hat{A} |J\rangle \) can be extracted from the residues of the quadratic response functions of \( \hat{A} \).\textsuperscript{20,29} The response functions of \( \hat{A} \) are those characterizing the time evolution of the average values

\[
\langle A(t) \rangle = \langle 0(t) | \hat{A} | 0(t) \rangle,
\]

where \( |0(t)\rangle \) is the time-dependent electronic ground state when a general time-dependent field \( \mathbf{W}(t) \) is applied to the electronic system. The interaction between the field and the electronic system can be resolved into Fourier components \( \mathbf{V}^{\omega} \) according to\textsuperscript{29}

\[
\mathbf{V}(t) = \int_{-\infty}^{\infty} \mathbf{V}^{\omega} e^{-i\omega t} d\omega
\]

and the quadratic response functions of \( \hat{A} \) at frequencies \( \omega_\alpha \) and \( \omega_\beta \) are then \textsuperscript{29}

\[
\langle \langle A; \mathbf{V}^{\omega_\alpha} , \mathbf{V}^{\omega_\beta} \rangle \rangle = \hat{\mathcal{P}}(\alpha, \beta) \sum_{I,J} \left[ \frac{\langle 0| \hat{A} |I\rangle \langle I| (\mathbf{V}^{\omega_\alpha} - \langle 0| \mathbf{V}^{\omega_\alpha} |0\rangle) |J\rangle \langle J| \mathbf{V}^{\omega_\beta} |0\rangle}{(\omega_\alpha + \omega_\beta - \omega_J)(\omega_\beta - \omega_J)} \right. \\
+ \frac{\langle I| \hat{A} |0\rangle \langle 0| (\mathbf{V}^{\omega_\alpha} - \langle 0| \mathbf{V}^{\omega_\alpha} |0\rangle) |J\rangle \langle J| \mathbf{V}^{\omega_\beta} |0\rangle}{(\omega_\alpha + \omega_\beta + \omega_J)(\omega_\beta + \omega_J)} \\
- \frac{\langle 0| \mathbf{V}^{\omega_\alpha} |I\rangle \langle I| (\hat{A} - \langle 0| \hat{A} |0\rangle) |J\rangle \langle J| \mathbf{V}^{\omega_\beta} |0\rangle}{(\omega_\alpha + \omega_\beta - \omega_J)} \] 

(4)

In this equation, \( |0\rangle \) is the static electronic ground state, without the perturbation from the external field \( \mathbf{W}(t) \); \( \mathbf{V}^{\omega_\alpha} \) and \( \mathbf{V}^{\omega_\beta} \) are the Fourier transform of \( \mathbf{V}(t) \) in Eq. (3) at frequencies \( \omega_\alpha \) and \( \omega_\beta \), respectively; \( \hat{\mathcal{P}}(\alpha, \beta) \) is the permutation operator that generates all the permutations of \( \alpha \) and \( \beta \) and \( \omega_\alpha \) and \( \omega_\beta \) are the excitation energies for the excited states \( |I\rangle \) and \( |J\rangle \).

For \( I \neq J \), the quantity \( \langle I| \hat{A} |J\rangle \) can be obtained from the residue of the quadratic response function

\[
\langle I| \hat{A} |J\rangle = \lim_{\omega_\alpha \to -\omega_\beta} \frac{(\langle A; \mathbf{V}^{\omega_\alpha} , \mathbf{V}^{\omega_\beta} \rangle)}{(0|\mathbf{V}^{\omega_\alpha} |I\rangle \langle J| \mathbf{V}^{\omega_\beta} |0\rangle)},
\]

(5)

If we choose \( \hat{A} = \hat{\nabla}_R \) as the nuclear derivative operator, then the derivative coupling \( \mathbf{d}_{IJ} \) in Eq. (1) can be calculated using Eq. (5).

2. Quadratic response functions in time-dependent Kohn-Sham theory

In order to calculate the derivative coupling \( \mathbf{d}_{IJ} \) between TDDFT excited states, we need to derive the quadratic response functions of \( \hat{\nabla}_R \) for time-dependent Kohn-Sham (TDKS) systems, whereas Eq. (4) gives the response function for exact states. In the TDKS system, the time-dependent ground state is approximated as a single determinant \( \Phi(t) \) that provides the correct electron density at time \( t \), and the expectation value of \( \hat{\nabla}_R \) can be expressed as

\[
\mathbf{D}^{\text{KS}}_R(t) \equiv \langle \Phi(t)| \hat{\nabla}_R |\Phi(t)\rangle = \sum_i \langle \psi_i(t)| \hat{\nabla}_R |\psi_i(t)\rangle,
\]

(6)

where the \( |\psi_i(t)\rangle \) are the occupied TDKS orbitals.

Given the perturbation from the external scalar potentials,

\[
\mathbf{V}(t) = \lambda_\alpha^{\omega_\alpha} e^{-i\omega_\alpha t} + \lambda_\beta^{\omega_\beta} e^{-i\omega_\beta t}.
\]

(7)

the TDKS orbitals may be expanded up to the second order in \( \lambda \),

\[
|\psi_i(t)\rangle = e^{-i\epsilon_i t} \left( |\phi_i\rangle + \lambda_\alpha^{\omega_\alpha} |\phi_i^{(\alpha)}(t)\rangle + \lambda_\beta^{\omega_\beta} |\phi_i^{(\beta)}(t)\rangle + \lambda_\alpha^{\omega_\alpha} \lambda_\beta^{\omega_\beta} |\phi_i^{(\alpha\beta)}(t)\rangle \right) + \lambda_\alpha^{\omega_\alpha} \lambda_\beta^{\omega_\beta} |\phi_i^{(\alpha\beta)}(t)\rangle,
\]

(8)

where the \( |\phi_i\rangle \) are the static KS orbitals, with orbital energies \( \epsilon_i \) in the absence of the perturbation. Hereafter, we will set \( \lambda_\alpha = \lambda_\beta = 1 \) for simplicity.

The orbital \( |\psi_i^{(\alpha)}(t)\rangle \) in Eq. (8) may be expanded in the basis of virtual static KS orbitals \( |\phi_d\rangle \),

\[
|\psi_i^{(\alpha)}(t)\rangle = \sum_d \left( X_{ad}^{\omega_\alpha} e^{i\omega_\alpha t} + Y_{ad}^{\omega_\alpha} e^{-i\omega_\alpha t} \right) |\phi_d\rangle,
\]

(9)
where $X^{(a)}$ and $Y^{(a)}$ are the virtual-occupied (VO) and occupied-virtual (OV) blocks of the linear density matrix response. These response functions satisfy the TDKS linear response equations:

$$\( (\Lambda - \omega_{a}\Lambda) |X^{(a)}, Y^{(a)}\rangle = -|P^{(a)}, Q^{(a)}\rangle \).$$

Here,

$$\Lambda = \begin{pmatrix} A & B \\ B & A \end{pmatrix}$$

is the orbital Hessian whose matrix elements are

$$A_{ai,bj} = (\epsilon_{ai} - \epsilon_{bj}) \delta_{i,j} \delta_{a,b} + \langle \phi_{a} | \phi_{b} | \phi_{j} \rangle$$

and

$$B_{ai,bj} = \langle \phi_{a} | \phi_{b} | \phi_{i} \rangle - C_{HF} \langle \phi_{a} | \phi_{b} | \phi_{i} \rangle + \langle \phi_{a} | f^{xc} \phi_{b} | \phi_{j} \rangle$$

for a hybrid functional within the adiabatic approximation. In Eqs. (12) and (13), $C_{HF}$ is the fraction of the Hartree-Fock (HF) exchange, and $f^{xc}$ is the exchange-correlation functional kernel. The matrix $\Lambda$ in Eq. (10) is defined as

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and $|P^{(a)}, Q^{(a)}\rangle$ represents the perturbation potential whose matrix elements are

$$P^{(a)}_{ai} = \langle \phi_{ai} | V^{\omega_{a}} | \phi_{i} \rangle,$$

$$Q^{(a)}_{ai} = \langle \phi_{i} | V^{\omega_{a}} | \phi_{ai} \rangle.$$

Similarly, the quantity $|\psi^{(a,b)}(t)\rangle$ in Eq. (8) may be expanded as

$$|\psi^{(a,b)}(t)\rangle = \sum_{a} \left( X^{(a)}_{ai}(t) e^{i(\omega_{a} + \omega_{b}) t} + Y^{(a)}_{ai}(t) e^{-i(\omega_{a} + \omega_{b}) t} \right) |\phi_{i}\rangle$$

$$- \sum_{aj} \left( X^{(a)}_{aj}(t) e^{i(\omega_{a} + \omega_{b}) t} + Y^{(a)}_{aj}(t) e^{-i(\omega_{a} + \omega_{b}) t} \right) |\phi_{j}\rangle,$$

where $X^{(a)}$ and $Y^{(a)}$ satisfy the TDKS quadratic response equations:

$$[\Lambda - (\omega_{a} + \omega_{b}) \Lambda] |X^{(a,b)}, Y^{(a,b)}\rangle = -|R^{(a,b)}, S^{(a,b)}\rangle.$$  

More details about Eq. (18) are presented in the Appendix.

The TDKS density operator can be calculated from the TDKS orbitals as

$$\hat{\rho}(t) = \sum_{i} |\psi_{i}(t)\rangle \langle \psi_{i}(t)|.$$

Using Eqs. (8), (9), and (17), it is easy to calculate the linear and the quadratic response functions of the density operator. The linear response can be obtained by collecting the terms which are multiplied by $e^{i(\omega_{a} + \omega_{b}) t}$,

$$\hat{\gamma} = \sum_{ai} \left( X^{(a)}_{ai}(t) |\phi_{ai}\rangle \langle \phi_{i}| + Y^{(a)}_{ai}(t) |\phi_{i}\rangle \langle \phi_{ai}| \right).$$

Likewise, collecting the terms that are multiplied by $e^{i(\omega_{a} + \omega_{b}) t}$ gives us the quadratic response function

$$\hat{\gamma}^{(a)} = \sum_{ai} \left( X^{(a)}_{ai}(t) |\phi_{ai}\rangle \langle \phi_{ai}| + Y^{(a)}_{ai}(t) |\phi_{i}\rangle \langle \phi_{ai}| \right).$$

Equations (20) and (21) have the correct idempotent forms of the linear and the quadratic response functions of the density operator, which validates the expansions in Eqs. (8), (9), and (17).

The quantity $D^{(a,b),KS}_{R}(t)$ in Eq. (6) can now be calculated by direct differentiation, using the formulas presented in Eqs. (8), (9), and (17). In order to get the quadratic response function of $D^{(a,b),KS}_{R}(t)$, we collect all the terms multiplied by $e^{i(\omega_{a} + \omega_{b}) t}$ as before, which gives us

$$D^{(a,b),KS}_{R} = \sum_{ai} \left( X^{(a)}_{ai}(t) - Y^{(a)}_{ai}(t) \right) |\phi_{ai}\rangle \langle \phi_{ai}|$$

$$- \sum_{aj} \left( X^{(a)}_{aj}(t) + X^{(b)}_{ai}(t) \right) |\phi_{ai}\rangle \langle \phi_{aj}|$$

$$+ \sum_{ai} \left( Y^{(a)}_{ai}(t) + Y^{(b)}_{ai}(t) \right) |\phi_{ai}\rangle \langle \phi_{ai}|$$

$$+ \sum_{ai} \left( Y^{(a)}_{ai}(t) - Y^{(b)}_{ai}(t) \right) |\phi_{ai}\rangle \langle \phi_{ai}|.$$

### 3. Derivative couplings between TDDFT excited states

Having derived the quadratic response function of $D^{(a,b),KS}_{R}(t)$, we just need to extract the derivative couplings from the residues of $D^{(a,b),KS}_{R}$ in Eq. (22) following the same procedure shown in Eq. (5).

It is well known that by using Eq. (10) and the spectral expansion, the quantity $|X^{(a)}, Y^{(a)}\rangle$ can be expressed as

$$|X^{(a)}, Y^{(a)}\rangle = \sum_{l} \left( |X_{l}, Y_{l}\rangle \langle X_{l}, Y_{l}| \right) |P^{(a)}, Q^{(a)}\rangle,$$

where $|X_{l}, Y_{l}\rangle$ and $|Y_{l}, X_{l}\rangle$ satisfy the following pseudo-eigenvalue equations:

$$(A - \omega_{a} I) |X_{l}, Y_{l}\rangle = 0,$$

$$(A + \omega_{b} I) |Y_{l}, X_{l}\rangle = 0,$$

subject to the orthonormality conditions

$$|X_{l}, Y_{l}| A |X_{l}, Y_{l}\rangle = \delta_{ll},$$

$$A |Y_{l}, X_{l}| A |Y_{l}, X_{l}\rangle = -\delta_{ll}.$$

From Eqs. (15), (16), and (23), the residues of $|X^{(a)}, Y^{(a)}\rangle$ and $|X^{(b)}, Y^{(b)}\rangle$ at frequencies $-\omega_{a}$ and $\omega_{a}$ can be written as

$$\lim_{\omega_{a} \rightarrow -\omega_{a}} \langle X^{(a)}, Y^{(a)}\rangle = -|Y_{l}, X_{l}\rangle \langle Y_{l}, X_{l}| P^{(-\omega_{a})} Q^{(-\omega_{a})} = -|Y_{l}, X_{l}| \langle Y_{l}, X_{l}| 0 |V^{(-\omega_{a})}| I \rangle.$$
and
\[
\lim_{\omega_R \to \omega_J} (\omega_R - \omega_J) |X^{(\beta)}_J, Y^{(\beta)}_J\rangle = |X_J, Y_J\rangle (X_J, Y_J | P_{a,J}, Q^{a,J}\rangle \\
= |X_J, Y_J\rangle (J|V_{a,J}\rangle 0).
\] (27)

Finally, we can extract the derivative coupling between two TDDFT excited states from the residues of \(D_{R}^{\beta J}\) in Eq. (22) by using Eqs. (26), (27), and \(|X_J, Y_J\rangle\) as derived in the Appendix. The result is
\[
d_{IJ}^{KS} = \sum_{ai}(X_{ai}^{IJ} - Y_{ai}^{IJ})(\phi_i| \nabla_R | \phi_a) \\
- \sum_{ija}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_j) \\
+ \sum_{abi}(X_{bi}^{J}X_{bi}^{J} + Y_{bi}^{J}Y_{bi}^{J})(\phi_a| \nabla_R | \phi_b) \\
+ \sum_{ai} \omega_{a,J} \langle \phi_i| \nabla_R | \phi_a\rangle \\
- \sum_{ai} \omega_{a,J} \langle \phi_i| \nabla_R | \phi_a\rangle \\
+ \sum_{ai} \omega_{a,J} \langle \phi_i| \nabla_R | \phi_a\rangle \\
- \sum_{ai} \omega_{a,J} \langle \phi_i| \nabla_R | \phi_a\rangle \\
+ \sum_{ai} \omega_{a,J} \langle \phi_i| \nabla_R | \phi_a\rangle \\
- \sum_{ai} \omega_{a,J} \langle \phi_i| \nabla_R | \phi_a\rangle.
\] (28)

The last term in Eq. (28) can be calculated following the procedure given by Li and Liu;\textsuperscript{20} details are shown in the Appendix. By substituting Eq. (A14) into Eq. (28), the final expression for the derivative coupling between two TDDFT excited states reads
\[
d_{IJ}^{KS} = \sum_{ai}(X_{ai}^{IJ} - Y_{ai}^{IJ})(\phi_i| \nabla_R | \phi_a) \\
- \sum_{ija}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_j) \\
+ \sum_{abi}(X_{bi}^{J}X_{bi}^{J} + Y_{bi}^{J}Y_{bi}^{J})(\phi_a| \nabla_R | \phi_b) \\
+ \sum_{ijab}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_a) \\
- \sum_{ijab}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_a) \\
+ \sum_{ijab}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_a) \\
- \sum_{ijab}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_a).
\] (29)

In Eq. (29), the nuclear derivatives of the orbital rotation Hessians \(\nabla_R A\) and \(\nabla_R B\) can be obtained from the conventional TDDFT gradient formalism.\textsuperscript{3,10,27} The nuclear derivatives of the KS orbitals, \(\nabla_R \phi_a\), can be calculated as in previous work.\textsuperscript{3,10,27}

**B. Nonadiabatic coupling vectors between TDDFT excited states**

The nonadiabatic coupling vector (NACV) along with the energy difference gradient vector can be used to determine the branching plane at conical intersections.\textsuperscript{32} For states \(|I\rangle\) and \(|J\rangle\) that are exact eigenstates of \(\hat{H}\), the NACV may be defined as\textsuperscript{32}
\[
h_{IJ} = \langle I| \nabla_R H | J\rangle = d_{IJ}(E_J - E_I) \\
= \left\{ I\left[ \sum_{i=1}^{N} \nabla_{ei}(i) \right] J \right\},
\] (30)

where \(\nabla_{ei}(i)\) is the electron-nucleus Coulomb potential for the ith electron.

Since \(d_{IJ}^{KS}\) in Eq. (29) was derived from response theory and is therefore formally exact, we can simplify the NACV between TDDFT excited states as
\[
h_{IJ}^{KS} = (\omega_J - \omega_I) d_{IJ}^{KS}.
\] (31)

In addition, we can also derive \(h_{IJ}^{KS}\) from response theory, where we replace the operator \(\nabla_R\) in Eq. (6) by \(\nabla_R \nabla_{ei}\). Following the same procedure used to derive \(d_{IJ}^{KS}\), we easily obtain
\[
h_{IJ}^{KS} = \sum_{ai}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_a) \\
- \sum_{ija}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_j) \\
+ \sum_{abi}(X_{bi}^{J}X_{bi}^{J} + Y_{bi}^{J}Y_{bi}^{J})(\phi_a| \nabla_R | \phi_b) \\
+ \sum_{ijab}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_a) \\
- \sum_{ijab}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_a) \\
+ \sum_{ijab}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_a).
\] (32)

Equations (31) and (32) should be equivalent in the limit that an exact density functional and frequency-dependent functional kernel are used. Of course, the functionals used in practice are approximate, and the frequency-independent adiabatic approximation\textsuperscript{30} is almost always invoked. As such, the NACVs obtained from Eqs. (31) and (32) will generally differ. Note that Eq. (32) does not involve nuclear derivatives of \(A\) or \(B\), which is very different from the NACV defined within the CIS theory.\textsuperscript{3}

**C. Discussion**

Equation (29) for the derivative coupling \(d_{IJ}^{KS}\) between two TDDFT excited states is equivalent to the expression derived by Li and Liu\textsuperscript{20} [Eq. (125) of Ref. 20]. Those authors started from the equation-of-motion formalism, and obtained the derivative couplings for arbitrary excitation subspaces. Here, for TDDFT, the excitation subspace is limited to single excitations, considering terms up to second-order response.

The first three lines on the right side of Eq. (29) are the VO, OV, and VO blocks of the transition density matrix between two excited states (where O means “occupied” and V means “virtual”), multiplied by the “half derivative” of the corresponding KS orbital overlap matrices, \(\langle \phi_p | \nabla_R | \phi_q\rangle\). These terms are similar to the “configuration state function” (CSF) contribution in the formulation of MRCI derivative couplings.\textsuperscript{4,7} The terms involving the nuclear derivatives of the orbital rotation Hessian in Eq. (29) resemble the “CI contribution,” in the language of MRCI derivative couplings.

In CIS theory, the VO and OV blocks of the transition density matrix between two CIS states is zero because only single excitations are considered. By directly following the CIS procedure\textsuperscript{32} to calculate TDDFT derivative couplings (pseudo-wavefunction approach), the terms including \(|X_J, Y_J\rangle\) are neglected. This affords derivative couplings
\[
d_{IJ}^{PWA} = \sum_{ija}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_j) \\
+ \sum_{ab}(X_{bi}^{J}X_{bi}^{J} + Y_{bi}^{J}Y_{bi}^{J})(\phi_a| \nabla_R | \phi_b) \\
+ \sum_{ijab}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_a) \\
+ \sum_{ijab}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_a) \\
+ \sum_{ijab}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_a) \\
+ \sum_{ijab}(X_{ai}^{J}X_{ai}^{J} + Y_{ai}^{J}Y_{ai}^{J})(\phi_i| \nabla_R | \phi_a).
\] (33)
Interestingly, for spin-flip TDDFT (SF-TDDFT) with either collinear\textsuperscript{28} or non-collinear\textsuperscript{33} kernels, the right side of Eq. (A9) vanishes. For spin-flipping excitations, it is straightforward to show that \( \partial^2F_{pq}/\partial r_s\partial r_{s'} = 0 \) and \( \partial^2F_{pq}/\partial P_{rs}\partial P_{r's'} = 0 \) in Eqs. (A6) and (A7).\textsuperscript{28,33} Thus, the VO and OV blocks of the transition density matrix between two SF-TDDFT states is zero. In other words, for SF-TDDFT, the derivative couplings calculated by the PWA\textsuperscript{3} is exactly the same as the one calculated by the quadratic response approach (QRA) shown in Eq. (29).

The cost of computing \( d_{ij}^{\text{ks}} \) in Eq. (29) is greater than the cost of computing \( d_{ij}^{\text{pwa}} \), since an additional set of linear equations [namely, Eq. (A9)] needs to be solved. The cost of solving Eq. (A9) is about the same as a single TDDFT excited-state gradient calculation. As such, it is important to quantify any differences between \( d_{ij}^{\text{ks}} \) and \( d_{ij}^{\text{pwa}} \), in order to determine situations in which the PWA is capable of providing accurate derivative couplings. This is the topic of the next section.

### III. NUMERICAL EXAMPLES

TDDFT derivative couplings \( d_{ij}^{\text{ks}} \) from quadratic response theory [Eq. (29)] have been implemented in a locally modified version of the Q-Chem program \textsuperscript{34}, whereas derivative couplings with the PWA have been implemented previously.\textsuperscript{3,25,27} Here, we numerically compare the QRA and PWA results in order to determine the importance of the contribution from the VO and OV blocks of the transition density matrix between TDDFT excited states.

We will also compare TDDFT derivative couplings to full configuration interaction (FCI) results, and to CIS results. The effects of the Tamm-Dancoff approximation\textsuperscript{35} (TDA) on the TDDFT results are also considered. For derivative couplings \( d_{ij}^{\text{ks}} \), we take the TDA to mean that the vectors \( Y_I^I, Y_J^J \), and \( Y_{ij}^I \) in Eqs. (29) and (A9) are set to zero.

The CASSCF module\textsuperscript{36,37} of the Molpro program\textsuperscript{38} was used to perform the FCI calculations. The derivative couplings between the FCI excited states were calculated using finite central differences with a step size of 0.01 a.u. All the other calculations were performed using Q-Chem.\textsuperscript{34}

#### A. Difference between \( d_{ij}^{\text{ks}} \) and \( d_{ij}^{\text{pwa}} \)

Table I. Line integrals of the TDDFT derivative couplings along closed circular loops that enclose a conical intersection. These loops are centered either at the minimum-energy crossing point (\( \mathbf{R}_{\text{cross}} \)) or at a point displaced 0.1 Å in the \( \mathbf{g}_{ij} \) direction. Geometric phases \( \phi \) are given in units of \( \pi \), and values in parenthesis are computed using the Tamm-Dancoff approximation.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>( \phi^{\text{QRA}}/\pi )</th>
<th>( \phi^{\text{PWA}}/\pi )</th>
<th>( \phi^{\text{QRA}}/\pi )</th>
<th>( \phi^{\text{PWA}}/\pi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{H}_2\text{O} )</td>
<td>0.9997 (1.0001)</td>
<td>1.0003 (1.0000)</td>
<td>0.0043 (0.0119)</td>
<td>0.0110 (0.0120)</td>
</tr>
<tr>
<td>Uracil</td>
<td>0.9978 (0.9999)</td>
<td>0.9978 (0.9999)</td>
<td>0.0013 (0.0014)</td>
<td>0.0014 (0.0014)</td>
</tr>
</tbody>
</table>

FIG. 1. Differences between \( d_{ij}^{\text{ks}} \) and \( d_{ij}^{\text{pwa}} \) at various energy gaps. The derivative couplings were calculated by TDDFT/TDA at the PBE0/6-31G** level. Full (non-TDA) TDDFT results are similar and are omitted here. The magnitude difference is defined in Eq. (34) and \( \cos \theta \) in Eq. (35).

From Fig. 1, we see that there is almost no difference between \( d_{ij}^{\text{ks}} \) and \( d_{ij}^{\text{pwa}} \) for systems with energy gaps \(<1 \text{ eV} \). Only for larger gaps does the magnitude difference [Eq. (34)]
FIG. 2. (a) Energies of the 2 $^1\Sigma^+_g$ and 3 $^1\Sigma^+_g$ states of H$_2$ computed at the FCI, CIS, and TDDFT/TDA levels, using the aug-cc-pVDZ basis set and the PBE0 functional for TDDFT. (b) Absolute value of the $z$-component of the derivative coupling between these two states, computed using the same methods.

FIG. 3. (a) Energies of the 2 $^1\Sigma^+_g$ and 3 $^1\Sigma^+_g$ states of H$_2$ computed at the FCI, TDHF, and TDDFT levels, using the aug-cc-pVDZ basis set and the PBE0 functional for TDDFT. (b) Absolute value of the $z$-component of the derivative coupling between these two states, computed using the same methods.

exceed 2%, and even in these cases the vectors $d_{KS}^{IJ}$ and $d_{PW A}^{IJ}$ are nearly parallel. This suggests that for optimizations of minimum-energy crossing points along conical seams, where only the direction of the derivative coupling is important, $d_{PW A}^{IJ}$ can be safely used with lower computational cost. As such, we conclude that for molecules with low symmetry, $d_{KS}^{IJ}$ and $d_{PW A}^{IJ}$ can usually be used interchangeably between the states with energy gaps as large as 6 eV. (For highly symmetric molecules, such as Li$_2$ as considered below, $d_{KS}^{IJ}$ and $d_{PW A}^{IJ}$ may exhibit larger differences even for small-gap systems, and the choice between the two must be considered more carefully.)

B. Geometric phase effect around conical intersections

According to the geometric phase effect, the nuclear wavefunction accumulates an additional (geometric) phase as it is transported around a path that encloses a conical intersection. This compensates for the phase change of the adiabatic electronic wavefunction. The accumulated phase $\phi$ is related to the derivative couplings according to

$$\phi = \int_C d_{IJ} \cdot d\mathbf{R} = \pi,$$

where the loop $C$ encloses a conical intersection. (Here, we define $\phi$ as the indicated integral around $C$. For exact derivative couplings, $\phi = \pi$, though a different value of the integral might be obtained using approximate derivative couplings.)

It has been shown previously that $d_{PW A}^{IJ}$ satisfies Eq. (36) for both full TDDFT calculations and TDDFT/TDA calculations. This is hardly surprising, since the PWA is by nature a wavefunction-based method. It is not clear whether $d_{KS}^{IJ}$ derived from response theory should satisfy Eq. (36), although given that the difference between $d_{KS}^{IJ}$ and $d_{PW A}^{IJ}$ is small near crossing points, we might anticipate this relationship is satisfied for $d_{KS}^{IJ}$ as well.

In this work, we calculated the geometric phase $\phi$ for H$_2$O and for uracil. For H$_2$O, we computed $\phi$ at the B3LYP/6-31G** level for the $S_1/S_4$ conical intersection, and for uracil, we calculated $\phi$ at the PBE0/6-31G** level for the $S_1/S_2$ intersection. The loop $C$ was chosen as a circle in the branching plane with a radius of 0.001 Å and a center near the minimum-energy crossing point, as detailed below. The branching plane was determined as the span of the vectors $g_{IJ}$ and $h_{IJ}$, where

$$g_{IJ} = \nabla_R (\omega_J - \omega_I)$$

and
FIG. 4. (a) Energies of the $1\Sigma^+_g$ and $3\Sigma^+_g$ states of He$_2$ computed at the FCI, CIS, and TDDFT/TDA levels, using the aug-cc-pVDZ basis set and the PBE0 functional for TDDFT. (b) Absolute value of the $z$-component of the derivative coupling between these two states, computed using the same methods.

is the energy difference gradient vector. The vector $\mathbf{h}_{IJ}$ is the NACV, which is given by $\mathbf{h}_{IJ}^{KS}$ Eq. (31) within the quadratic response approach. Within the pseudo-wavefunction approach,

$$\mathbf{h}_{IJ}^{PWA,TDFT} = \mathbf{d}_{IJ}^{PWA}(\omega_J - \omega_I),$$  

or upon invoking the TDA,

$$\mathbf{h}_{IJ}^{PWA,TDA} = \sum_{i,j,a,b} X^I_{ai} (\hat{\nabla}_R A_{ai,bj}) X^J_{bj}.$$  

Table I lists the geometric phases computed along two different circular loops, one that is centered at the minimum-energy crossing point ($\mathbf{R}_{\text{mex}}$) and another whose center is displaced from $\mathbf{R}_{\text{mex}}$ by 0.1 Å along a unit vector in the direction $\mathbf{g}_{IJ}$. For the loop centered at $\mathbf{R}_{\text{mex}}$, both $\mathbf{d}_{IJ}^{KS}$ and $\mathbf{d}_{IJ}^{PWA}$ afford the correct phase, $\phi = \pi$. For the loop whose center is displaced from $\mathbf{R}_{\text{mex}}$, however, both methods afford a phase $\phi = 0$, indicating that this displaced loop does not enclose a conical intersection. (This is expected, since we displaced the center of this path along one of the branching-plane degrees of freedom.) Overall, the geometric-phase behavior of both $\mathbf{d}_{IJ}^{KS}$ and $\mathbf{d}_{IJ}^{PWA}$ is correct.

C. Comparison with FCI derivative couplings

In the previous section, we showed that the QRA and PWA derivative couplings are almost identical in regions of the potential surface where energy gaps are small. In this section, we consider some systems with larger energy gaps, as well as some molecules with higher symmetry, in order to explore whether the agreement between the QRA and the PWA still holds. Specifically, we will examine the derivative couplings for H$_2$, He$_2$, Li$_2$, and linear H$_3$ using FCI, CIS, time-dependent Hartree-Fock (TDHF) theory, and TDDFT. For the TDDFT calculations, similar results are obtained using B3LYP$^{40,41}$ and PBE0$^{39}$ so only the PBE0 results are shown. As all of these examples are linear molecules, the $z$ axis is taken to be the molecular axis, only the $z$ components of the derivative couplings are non-zero.

1. H$_2$

Derivative couplings for the $2\Sigma^+_g$ and $3\Sigma^+_g$ states of H$_2$, computed in the aug-cc-pVDZ basis set with and without invoking the TDA, are shown in Figs. 2 and 3, respectively.
For bond lengths larger than 2.25 bohr, both states have strong double excitation character that cannot be captured by CIS, TDHF, or TDDFT, thus, our plots terminate at 2.25 bohr.

The CIS and TDHF-PWA methods are formally similar, in that both are wavefunction-based approaches that include only Hartree-Fock exchange, with excitation spaces that are truncated at single excitations. Derivative couplings computed using the these two methods are similar to one another, and agree quite well with FCI results for bond lengths ranging from 1.0–1.6 bohr [Figs. 2(b) and 3(b)]. For longer bond lengths, however, these methods fail to reproduce the increase in $d_I J, z$ as a function of bond length that is observed in the FCI results. This failure can be understood by examining the potential energy surfaces (PESs) of the $3 \Sigma^+_g$ state, whose curvature is incorrectly predicted by CIS and TDHF calculations for bond lengths larger than 1.6 bohr. The PBE0-PWA derivative couplings, both with or without the TDA, agree qualitatively with the CIS and TDHF-PWA results (and are thus incorrect for longer bond lengths), owing to the similar wavefunction ansatz that is used, and the fact that the PESs are very similar at the CIS, TDHF, and TDDFT levels of theory.

For this pair of states, TDHF-QRA affords almost identical results as compared to TDDHF-PWA, and PBE0-QRA agrees with FCI a little better than PBE0-PWA [see Fig. 3(b)].

This may indicate that the contribution from $|X_{I J}, Y_{I J}|$ is small for the current system, but full inclusion of the quadratic response contribution can still improve the derivative couplings calculated from the PWA. Note that the gap between the two states considered here is about 6 eV, and the PWA and QRA derivative couplings agree well with one another.

2. He$_2$

For He$_2$, we examine derivative couplings between the $3 \Sigma^+_u$ and $5 \Sigma^+_u$ states, both of which are basically single excitations in character. The $3 \Sigma^+_u$ state is dominated by a $1\sigma_u \rightarrow 2\sigma_u$ excitation and the $5 \Sigma^+_u$ state involves $1\sigma_u \rightarrow 3\sigma_u$ excitation. Potential energy curves and derivative couplings, with and without the TDA, are plotted in Figs. 4 and 5, respectively.

The CIS and TDHF methods reproduce all the features of the FCI potentials, even though the total energies are a bit shifted from the FCI values. This leads to good agreement among CIS, TDHF-PWA, and FCI derivative couplings. Interestingly, the TDHF-QRA derivative couplings are almost identical to the FCI derivative couplings at bond lengths larger than 2.0 bohr, whereas the PBE0-QRA results are less accurate, owing to qualitatively incorrect potential curves at the PBE0 level of theory [see Fig. 5(a)]. The success of TDHF-QRA for this
system is due to having a qualitatively correct wavefunction (leading to qualitatively correct potential curves), as well as full inclusion of the quadratic response contributions. Note that the current system has a large energy gap (>10 eV), and the QRA significantly improves upon the PW A derivative couplings.

At shorter bond lengths, however, the TDHF-QRA derivative coupling begins to diverge. As explained in the Appendix and in Ref. 21, this divergence is an intrinsic feature of the QRA approach in the presence of any excitation $|0\rangle \rightarrow |K\rangle$ whose excitation energy $\omega_K$ approaches degeneracy with the $IJ$ energy gap, $\omega_K \approx |E_I - E_J|$. In the present example, this divergence appears in the PBE0-QRA derivative coupling as well.

TDDFT potential curves fail to capture the state crossing between the $3^1\Sigma_u^+$ state and a higher-lying state at bond lengths of 3.5–4.0 bohr, and moreover, the double-well feature of the $5^1\Sigma_u^+$ state is also not correctly described in these calculations. Consequently, the TDDFT derivative couplings are qualitatively wrong at bond lengths larger than 4.0 bohr, although PBE0-QRA and PBE0-PWA results are equally good at shorter bond lengths. Moreover, the TDA makes almost no difference in the TDDFT derivative couplings for this particular system, in either the PWA or the QRA.

3. $Li_2$

For $Li_2$, we compute the derivative coupling between the two lowest $^1\Sigma_u^+$ states. These states have single-excitation character, involving $1\sigma_g \rightarrow 1\sigma_u$ and $1\sigma_g \rightarrow 2\sigma_u$ excitations, respectively. The calculated potential curves and derivative couplings are plotted, with and without the TDA, in Figs. 6 and 7, respectively. All methods afford similar potential curves for both states, and the derivative couplings obtained using the CIS, TDHF-PWA, and PBE0-PWA methods are about the same but are very different from the FCI derivative coupling.

In contrast, application of quadratic response theory shifts the TDHF and PBE0 derivative couplings closer to the FCI result. Note that the energy gap between the $1^1\Sigma_u^+$ and $2^1\Sigma_u^+$ states is <3 eV, yet $d_{K,\perp}^{PS}$ differs from $d_{PWA}^{PS}$ by about 50%. (This should be compared to the low-symmetry examples shown in Fig. 1, where the differences are <7% in all cases and <2% in most cases.) This example, along with that of $He_2$, may indicate the importance of the contributions from $|X_{IJ}, Y_{IJ}\rangle$ in some molecules with high symmetry, for reasons that are not clear to us.

4. $H_3$

For another high-symmetry example, we calculated the derivative couplings for $H_3$ in $D_{oh}$ symmetry. For HF and
density functional theory (DFT) calculations, the ground state in D_{och} symmetry has \( ^2\Sigma^+_u \) symmetry, and we compute derivative couplings between the \( ^1\Sigma^+_u \) and \( ^2\Sigma^+_u \) states.

In Figs. 8 and 9, we see that the potential curves obtained from CIS, TDHF, and TDDFT are very similar to the FCI potential curves, for both states. Furthermore, the derivative couplings computed at the CIS, TDHF, and TDDFT levels are qualitatively correct, and the differences between the QRA and PWA results are small. By observing the transition density matrix between FCI states, we found that the OO and VV blocks dominate, whereas the VO and OV blocks are negligible. The same is true for the TDDFT transition density matrix. This explains the similarity between the QRA and PWA derivative couplings. Note that the current \( \mathrm{H}_2 \) system has high symmetry and the energy gap between the two chosen states is large. Thus, in situations that the VO and OV blocks of the transition density matrix between TDDFT excited states are important is still unclear and needs further study.

IV. CONCLUSION

In the present study, we have presented a compact derivation of TDDFT derivative coupling between excited states that is entirely based on the quadratic response theory; we call this the quadratic response approach or QRA. Our previous formulation of the analytic derivative couplings for spin-flip TDDFT,\(^3\) which was based on direct differentiation of the Kohn-Sham determinant (what is herein called the PWA), is here shown to be rigorously correct. In other words, for spin-flip TDDFT, there is no distinction between the QRA and the PWA, whereas for traditional, spin-conserving TDDFT there is a distinction, which we have explored numerically here.

For small-gap systems, we find that the QRA and PWA can usually be used interchangeably, thus, the PWA may be a better choice considering its slightly lower computational cost. In certain cases, however (Li\(_2\) in the present work), the VO and OV blocks of the transition density matrix between excited states, which are neglected in the PWA, become important. In these cases, the QRA usually improves the accuracy of the derivative couplings, despite the latter method’s occasional divergence.

Overall, the quadratic response formalism for TDDFT derivative couplings is a potentially useful approach to calculate these quantities. Derivative couplings computed within the PWA tend to be qualitatively correct, provided that the potential energy surfaces for the states in question are qualitatively correct, and this approach is free of the divergent behavior encountered by quadratic response theory in the vicinity of excitations that are quasi-degenerate with excited-state energy gaps. Hence, for nonadiabatic dynamics simulations, where continuous derivative couplings are important, the PWA may be the method of choice.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (No. CHE-1300603) and calculations were performed at the Ohio Supercomputer Center under Project No. PAA-0003. J.M.H. is a Camille Dreyfus Teacher-Scholar.

APPENDIX: ADDITIONAL DETAILS OF THE DERIVATION

1. VO and OV blocks of the transition density matrix between TDDFT excited states

In Eq. (18), the matrix elements of \( R^{(\alpha \beta)} \) and \( S^{(\alpha \beta)} \) are

\[
R^{(\alpha \beta)}_{ai} = \Phi(\alpha, \beta) \left[ \sum_{bck} \left( \frac{\partial F_{ab}}{\partial P_{ck}} X_{bi}^{(\alpha)} Y_{ck}^{(\beta)} + \frac{\partial F_{ab}}{\partial P_{ck}} X_{bi}^{(\beta)} Y_{ck}^{(\alpha)} \right) - \sum_{jck} \left( X_{aj}^{(\beta)} \frac{\partial F_{ji}}{\partial P_{ck}} X_{bi}^{(\alpha)} Y_{ck}^{(\beta)} + X_{aj}^{(\alpha)} \frac{\partial F_{ji}}{\partial P_{ck}} X_{bi}^{(\beta)} Y_{ck}^{(\alpha)} \right) + \frac{1}{2} \sum_{b,j,b',j'} \frac{\partial^2 F_{ai}}{\partial P_{bj} \partial P_{b'j'}} (X_{aj}^{(\alpha)} Y_{b'j'}^{(\beta)} + Y_{aj}^{(\alpha)} X_{b'j'}^{(\beta)} + Y_{aj}^{(\beta)} X_{b'j'}^{(\alpha)}) \right] + \sum_{bcj} \left( X_{aj}^{(\beta)} Y_{bj}^{(\alpha)} + X_{aj}^{(\alpha)} Y_{bj}^{(\beta)} \right) - \sum_{bjk} \frac{\partial F_{ai}}{\partial P_{jk}} (X_{bk}^{(\alpha)} Y_{bj}^{(\beta)} + X_{bk}^{(\beta)} Y_{bj}^{(\alpha)}) \right] \quad (A1)
\]

and

\[
S^{(\alpha \beta)}_{ai} = \Phi(\alpha, \beta) \left[ -\sum_{jck} \left( \frac{\partial F_{ij}}{\partial P_{ck}} X_{ai}^{(\alpha)} Y_{ck}^{(\beta)} + \frac{\partial F_{ij}}{\partial P_{ck}} X_{ai}^{(\beta)} Y_{ck}^{(\alpha)} \right) + \sum_{bck} \left( X_{bi}^{(\beta)} \frac{\partial F_{ja}}{\partial P_{ck}} X_{ai}^{(\alpha)} Y_{ck}^{(\beta)} + Y_{bi}^{(\beta)} \frac{\partial F_{ja}}{\partial P_{ck}} Y_{ck}^{(\alpha)} \right) + \frac{1}{2} \sum_{b,j,b',j'} \frac{\partial^2 F_{ia}}{\partial P_{bij} \partial P_{b'j'}} (X_{ai}^{(\alpha)} Y_{bj}^{(\beta)} + Y_{ai}^{(\alpha)} X_{bj}^{(\beta)} + Y_{ai}^{(\beta)} X_{bj}^{(\alpha)}) \right] + \sum_{bcj} \left( X_{ai}^{(\beta)} Y_{bj}^{(\alpha)} + X_{ai}^{(\alpha)} Y_{bj}^{(\beta)} \right) - \sum_{bjk} \frac{\partial F_{ia}}{\partial P_{jk}} (X_{bk}^{(\alpha)} Y_{bj}^{(\beta)} + X_{bk}^{(\beta)} Y_{bj}^{(\alpha)}) \right] \quad (A2)
\]

where \( \Phi \) is the Fock matrix and \( P \) is the ground state density matrix. Derivatives of \( \Phi \) with respect to \( P \) have matrix elements

\[
\frac{\partial F_{pq}}{\partial P_{rs}} = \langle \phi_p | \phi_q | \psi_r | \psi_s \rangle - C_{rs} \langle \phi_p | \phi_q | \psi_r | \psi_s \rangle + \langle \phi_p | \phi_q | r^s | \psi_r \rangle \quad (A3)
\]

and

\[
\frac{\partial^2 F_{ai}}{\partial P_{bj} \partial P_{b'j'}} = \int \int \frac{d\mathbf{r} d\mathbf{r'} d\mathbf{r''}}{\delta(r-r') \delta(r-r'')} \frac{\partial^2 E_{xc}[\rho]}{\partial \rho(r) \partial \rho(r')} \phi_a(r) \phi_i(r) \phi_b(r') \phi_j(r'') \phi_b(r') \phi_j(r''), \quad (A4)
\]

From Eqs. (18), (26), (27), (A1), and (A2), it is straightforward to obtain

\[
\lim_{\omega_{\alpha} \rightarrow \omega_{\alpha}} \langle \omega_{\alpha} + \omega_{\alpha} \rangle \lim_{\omega_{\beta} \rightarrow \omega_{\beta} - \omega_{\beta}} \langle \omega_{\beta} - \omega_{\beta} \rangle \langle X^{(\alpha \beta)} | Y^{(\alpha \beta)} \rangle = [\Lambda - (\omega_{\beta} - \omega_{\alpha}) \Lambda^{-1}] R^{1J}, S^{1J}, \quad (A5)
\]
where the matrix elements of $R^{IJ}$ and $S^{IJ}$ are

$$R_{ai}^{IJ} = \sum_{bc} \left\{ \frac{\partial F_{aib}}{\partial P_{ck}} (Y_{ck}^I X_{bi}^J + X_{ck}^I Y_{bi}^J) + \frac{\partial F_{bic}}{\partial P_{ck}} (Y_{ck}^I X_{bi}^J + X_{ck}^I Y_{bi}^J) \right\}$$

and

$$S_{ai}^{IJ} = \sum_{bc} \left\{ \frac{\partial F_{bic}}{\partial P_{ck}} (Y_{ck}^I X_{bi}^J + X_{ck}^I Y_{bi}^J) + \frac{\partial F_{bic}}{\partial P_{ck}} (Y_{ck}^I X_{bi}^J + X_{ck}^I Y_{bi}^J) \right\}$$

By comparing to exact response theory, one discovers that the right side of Eq. (A5) constitutes the VO and OV blocks of the transition density matrix between two TDDFT excited states $|I\rangle$ and $|J\rangle$. We can define these blocks of the transition density matrix as

$$|X_{IJ}, Y_{IJ}\rangle = -[\Lambda - (\omega J - \omega I)\Lambda^{-1}]|R^{IJ}, S^{IJ}\rangle. \quad (A8)$$

Then, $|X_{IJ}, Y_{IJ}\rangle$ is the solution of the following linear equations:

$$[\Lambda - (\omega J - \omega I)\Lambda]|X_{IJ}, Y_{IJ}\rangle = -|R^{IJ}, S^{IJ}\rangle. \quad (A9)$$

Note that the quantity $[\Lambda - (\omega J - \omega I)\Lambda]^{-1}$ in Eq. (A8) has poles whenever $\omega_K = \omega_J - \omega_I$, where $\omega_K$ is the excitation energy for some excited state. As such, $|X_{IJ}, Y_{IJ}\rangle$ diverges when the energy difference between states $I$ and $J$ is close to the excitation energy of some other excited state. This divergence in $|X_{IJ}, Y_{IJ}\rangle$, with its concomitant divergence of the derivative couplings between states $I$ and $J$, has also been pointed out recently by Li et al.²¹

### 2. The last term in Eq. (28)

We can rewrite the last term in Eq. (28) as

$$\sum_{ai} (X_{ai}^{I'} \hat{\nabla} X_{ai}^{(\beta)} - Y_{ai}^{I'} \hat{\nabla} Y_{ai}^{(\beta)}) = \langle X_{I'}, Y_{I'} | \hat{\nabla} X_{ai}^{(\beta)} \rangle \hat{\nabla} Y_{ai}^{(\beta)} \rangle. \quad (A10)$$

Taking the nuclear derivative of Eq. (10), we obtain

$$\langle \Lambda - \omega \Lambda | \hat{\nabla} R^{(\beta)} \rangle \hat{\nabla} Y_{ai}^{(\beta)} \rangle = \langle \hat{\nabla} P^{(\beta)} \rangle \hat{\nabla} Q^{(\beta)} \rangle - \langle \hat{\nabla} R \Lambda | X_{ai}^{(\beta)} \rangle \hat{\nabla} Y_{ai}^{(\beta)} \rangle. \quad (A11)$$

Taking the inner product with $\langle X_{I'}, Y_{I'} |$ from the left, then we obtain

$$\langle X_{I'}, Y_{I'} | (\Lambda - \omega \Lambda) \hat{\nabla} X_{ai}^{(\beta)} \rangle \hat{\nabla} Y_{ai}^{(\beta)} \rangle = \langle \hat{\nabla} P^{(\beta)} \rangle \hat{\nabla} Q^{(\beta)} \rangle - \langle \hat{\nabla} R \Lambda | X_{ai}^{(\beta)} \rangle \hat{\nabla} Y_{ai}^{(\beta)} \rangle. \quad (A12)$$

From this equation, we know that

$$\langle X_{I'}, Y_{I'} | (\Lambda - \omega \Lambda) \hat{\nabla} X_{ai}^{(\beta)} \rangle \hat{\nabla} Y_{ai}^{(\beta)} \rangle = \langle \hat{\nabla} P^{(\beta)} \rangle \hat{\nabla} Q^{(\beta)} \rangle - \langle \hat{\nabla} R \Lambda | X_{ai}^{(\beta)} \rangle \hat{\nabla} Y_{ai}^{(\beta)} \rangle. \quad (A13)$$

After taking the residue of Eq. (A13), we obtain the final result

$$\lim_{\omega \to \omega_j} \langle \Lambda - \omega \Lambda | X_{ai}^{I'} \hat{\nabla} X_{ai}^{(\beta)} - Y_{ai}^{I'} \hat{\nabla} Y_{ai}^{(\beta)} \rangle = (\omega_j - \omega_i)^{-1} \langle X_{I'}, Y_{I'} | (\hat{\nabla} R \Lambda) X_{ai}^{(\beta)} \rangle Y_{ai}^{(\beta)} \rangle. \quad (A14)$$