AIP | The Journal of Chemical Physics

First order nonadiabatic coupling matrix elements between excited states: Implementation and application at the TD-DFT and pp-TDA levels

Zhendong Li, Bingbing Suo, and Wenjian Liu

Citation: The Journal of Chemical Physics **141**, 244105 (2014); doi: 10.1063/1.4903986
View online: http://dx.doi.org/10.1063/1.4903986
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/141/24?ver=pdfcov
Published by the AIP Publishing

Articles you may be interested in

First-order derivative couplings between excited states from adiabatic TDDFT response theory
J. Chem. Phys. **142**, 064114 (2015); 10.1063/1.4906941

Publisher's Note: "Derivative couplings between TDDFT excited states obtained by direct differentiation in the Tamm-Dancoff approximation" [J. Chem. Phys.141, 024114 (2014)]
J. Chem. Phys. **141**, 069903 (2014); 10.1063/1.4891539

Derivative couplings between TDDFT excited states obtained by direct differentiation in the Tamm-Dancoff approximation
J. Chem. Phys. **141**, 024114 (2014); 10.1063/1.4887256

First-order nonadiabatic coupling matrix elements between excited states: A Lagrangian formulation at the CIS, RPA, TD-HF, and TD-DFT levels
J. Chem. Phys. **141**, 014110 (2014); 10.1063/1.4885817

Multimode simulation of dimer absorption spectra from first principles calculations: Application to the 3,4,9,10-perylenetetracarboxylic diimide dimer
J. Chem. Phys. **131**, 154302 (2009); 10.1063/1.3245403

How can you REACH 100% of researchers at the Top 100 Physical Sciences Universities? (TIMES HIGHER EDUCATION RANKINGS, 2014)

With The Journal of Chemical Physics.

AIP | The Journal of Chemical Physics

THERE'S POWER IN NUMBERS. Reach the world with AIP Publishing.

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 134.129.120.3 On: Thu, 28 May 2015 13:34:55

THE JOURNAL OF CHEMICAL PHYSICS **141**, 244105 (2014)

First order nonadiabatic coupling matrix elements between excited states: Implementation and application at the TD-DFT and pp-TDA levels

Zhendong Li, Bingbing Suo, and Wenjian Liu[a)]

Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, State Key Laboratory of Rare Earth Materials Chemistry and Applications, College of Chemistry and Molecular Engineering, and Center for Computational Science and Engineering, Peking University, Beijing 100871, People's Republic of China

(Received 28 October 2014; accepted 1 December 2014; published online 23 December 2014)

The recently proposed rigorous yet abstract theory of first order nonadiabatic coupling matrix elements (fo-NACME) between electronically excited states [Z. Li and W. Liu, J. Chem. Phys. **141**, 014110 (2014)] is specified in detail for two widely used models: The time-dependent density functional theory and the particle-particle Tamm-Dancoff approximation. The actual implementation employs a Lagrangian formalism with atomic-orbital based direct algorithms, which makes the computation of fo-NACME very similar to that of excited-state gradients. Although the methods have great potential in investigating internal conversions and nonadiabatic dynamics between excited states of large molecules, only prototypical systems as a first pilot application are considered here to illustrate some conceptual aspects. © *2014 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4903986]

I. INTRODUCTION

While first-principles descriptions of fluorescence of large molecules have nowadays become more or less routine, largely thanks to the advent of time-dependent density functional theory (TD-DFT),[1,2] accurate and reliable descriptions of phosphorescence, intersystem crossings, internal conversions, as well as nonadiabatic dynamics of large molecules still remain a challenge to quantum chemistry. Among others, the spin-adapted open-shell TD-DFT,[3–5] combined with an efficient and accurate treatment of spin-orbit couplings,[6] provides an efficacious means for investigating phosphorescence and intersystem crossings of large molecules. To describe internal conversions and nonadiabatic dynamics, the so-called first order nonadiabatic coupling matrix elements (fo-NACME) between two adiabatic electronic states $|\Psi_I(x)\rangle$ and $|\Psi_J(x)\rangle$ at a given nuclear configuration $x \triangleq \{\vec{R}_A\}_{A=1}^{N_A}$,

$$g^{\xi}_{IJ} = \langle \Psi_I(x) | D_{\xi} | \Psi_J(x) \rangle, \quad D_{\xi} = \partial_{\xi}, \quad \xi \in x, \tag{1}$$

ought to be first calculated. The analytic formulation of fo-NACME has long been a nontrivial issue, especially for theoretical models without an explicit wavefunction. The most prominent example is the formulation of fo-NACME within TD-DFT. Although many interesting developments[7–21] have been made for the TD-DFT formulation of the fo-NACME g^{ξ}_{0I} between the ground and excited states, the corresponding TD-DFT formulation of the fo-NACME g^{ξ}_{IJ} between two excited states, which was once characterized as an open problem,[22–25] has only been achieved rather recently.[26] As a matter of fact, the particular formulation[26] is very general, in the sense that it encompasses all theoretical models that can be derived from either time-independent equation-of-motion (EOM)[27] or time-dependent response theory.[28–31] Briefly, for excited states with energies ω_I determined by a generalized eigenvalue equation,

$$\mathbf{E}\mathbf{t}_I = \omega_I \mathbf{S}\mathbf{t}_I, \tag{2}$$

the fo-NACME g^{ξ}_{0I} and g^{ξ}_{IJ} are given[26] simply by

$$g^{\xi}_{0I} = \sum_{pq} d^{\xi}_{pq} \gamma^{0I}_{pq}, \tag{3}$$

$$\begin{aligned} g^{\xi}_{IJ} &= \omega^{-1}_{JI} \mathbf{t}^{\dagger}_I \mathbf{E}^{\xi} \mathbf{t}_J + \sum_{pq} d^{\xi}_{pq} \gamma^{IJ}_{pq}, \\ \omega_{JI} &= E_J - E_I = \omega_J - \omega_I, \end{aligned} \tag{4}$$

where the transition density matrices γ^{0I}_{pq} and γ^{IJ}_{pq} are to be determined by a given theoretical model (e.g., TD-DFT), while the matrix elements d^{ξ}_{pq} are defined as

$$d^{\xi}_{pq} = \langle \psi_p | D_{\xi} | \psi_q \rangle = \langle \psi_p | \psi^{\xi}_q \rangle \tag{5}$$

in terms orthonormal molecular orbitals (MO) $\{\psi_p\}$, whether canonical or not. Note in passing that the matrix elements d^{ξ}_{pq} are antisymmetric, viz., $d^{\xi}_{pq} = -d^{\xi}_{qp}$, for real-valued MO. The generic expressions (3) and (4) cover, e.g., configuration interaction singles (CIS), particle-hole (ph) or particle-particle (pp) random phase approximation (RPA), time-dependent Hartree-Fock (TD-HF) and, in particular, adiabatic TD-DFT as special cases. For instance, the matrix $\mathbf{E}$ in Eq. (2) can be the Hamiltonian matrix $\langle m|H|n\rangle$ in the case of CIS, or the orbital Hessian in the case of TD-HF or TD-DFT. Note that, in the case of TD-HF or TD-DFT, Eq. (3) and the second term of Eq. (4) resemble the counterparts of standard linear and quadratic response theories[28–31] for the transition matrix elements ($\langle 0|V^a|I\rangle$ and $\langle I|V^a|J\rangle$) of an electronic operator V^a, respectively. By contrast, the first term of Eq. (4) originates[26] exclusively from the action of D_{ξ} on the amplitude $\mathbf{t}_J$.

In sum, the fo-NACME g^{ξ}_{0I} and g^{ξ}_{IJ} can be specified once the matrix $\mathbf{E}$ and the transition density matrices γ^{0I}_{pq} and γ^{IJ}_{pq}

[a)] Author to whom correspondence should be addressed. Electronic mail: liuwjbdf@gmail.com

0021-9606/2014/141(24)/244105/16/$30.00

are specified. To avoid the calculation of nuclear derivatives over the MO coefficients, $\mathbf{C}^{\xi}(x)$ in ψ_q^{ξ} (5), through the coupled-perturbed HF (CPHF) or Kohn-Sham (CPKS) equation for every nuclear coordinate ξ, the elegant Lagrangian formulation[32–36] can be employed. That is, by incorporating both the Brillouin condition $F_{ai\sigma}(x) = 0$ and the normalization condition $S_{pq}(x) = \delta_{pq}$, the Lagrangian for the fo-NACME can be constructed as[26]

$$L[x,\mathbf{C}(x),\mathbf{Z},\mathbf{W}] = g[x,\mathbf{C}(x)] + \sum_{ai} Z_{ai}F_{ai}(x) - \sum_{pq} W_{pq}(S_{pq}(x) - \delta_{pq}), \tag{6}$$

where Z_{ai} and $W_{pq} = W_{qp}$ are the Lagrangian multipliers, and the generating functions $g[x,\mathbf{C}(x)]$ are defined as

$$g_{0I}[x,\mathbf{C}(x)] = \sum_{pq} d_{pq}(x)\gamma_{pq}^{0I}, \tag{7}$$

$$g_{IJ}[x,\mathbf{C}(x)] = \omega_{JI}^{-1}\mathbf{t}_I^{\dagger}\mathbf{E}(x)\mathbf{t}_J + \sum_{pq} d_{pq}(x)\gamma_{pq}^{IJ}, \tag{8}$$

$$d_{pq}(x) = \langle\psi_p|\psi_q(x)\rangle = \sum_{\nu}\langle\psi_p|\chi_\nu(x)\rangle C_{\nu q}(x). \tag{9}$$

Once the multipliers Z_{ai} and W_{pq} are determined from the stationary condition

$$\left[\frac{\partial L}{\partial \mathbf{C}(x)}\right]_{x=x_0} = 0, \tag{10}$$

the fo-NACME are simply the *partial* derivatives $L^{(\xi)}$ of L at the reference point $x = x_0$, viz.,

$$\begin{aligned}
g^{\xi} &= L^{\xi} = L^{(\xi)} = g^{(\xi)} + \sum_{ai} Z_{ai}F_{ai}^{(\xi)} - \sum_{pq} W_{pq}S_{pq}^{(\xi)},\\
g_{0I}^{(\xi)} &= \sum_{pq} d_{pq}^{(\xi)}\gamma_{pq}^{0I},\\
g_{IJ}^{(\xi)} &= \omega_{JI}^{-1}\mathbf{t}_I^{\dagger}\mathbf{E}^{(\xi)}\mathbf{t}_J + \sum_{pq} d_{pq}^{(\xi)}\gamma_{pq}^{IJ},\\
d_{pq}^{(\xi)} &= \langle\psi_p|\psi_q^{(\xi)}\rangle = \sum_{\nu}\langle\psi_p|\chi_\nu^{(\xi)}\rangle C_{\nu q}.
\end{aligned} \tag{11}$$

Note in passing that the Lagrangian (6) applies also to the gradients[33,34] of excited states, where the generating function is just

$$g_I[x,\mathbf{C}(x)] = \mathbf{t}_I^{\dagger}\mathbf{E}(x)\mathbf{t}_I. \tag{12}$$

The first model considered here is TD-DFT. The spin-conserving TD-DFT variants of $\mathbf{E}$, γ_{pq}^{0I} and γ_{pq}^{IJ} required by g_{0I}^{ξ} and g_{IJ}^{ξ} are first derived in Sec. II A. Two points can already be made here. First, neither the use[13–16,19,20] of auxiliary wavefunctions nor the *ad hoc* replacement[21] of the CIS Hamiltonian with the TDA (Tamm-Dancoff approximation) matrix is necessary in the present formulation. Second, the extension to spin-flip TD-DFT can readily be made. Another model under concern here is the pp-TDA,[37,38] see Sec. II B. The Lagrangian technique can be used for both TD-DFT and pp-TDA, so as to make the computations of the fo-NACME and excited-state gradients rather similar. Some benchmark calculations are presented in Sec. III, while the conclusions are drawn in Sec. IV.

II. THEORY

Throughout the paper, the following convention for labeling the orbitals is to be used: $\{i, j, k, l, \ldots\}$ for occupied MO, $\{a, b, c, d, \ldots\}$ for virtual MO, $\{p, q, r, s, \ldots\}$ for unspecified MO, and $\{\mu, \nu, \kappa, \lambda, \ldots\}$ for atomic orbitals (AO). They are all assumed to be real valued since only spin-free Hamiltonians, nonrelativistic or scalar relativistic, are under concern here.

A. The fo-NACME at the TD-DFT level

1. *Defining quantities for g_{0I}^{ξ} and g_{IJ}^{ξ}*

The spin-conserving TD-DFT variant of the generalized eigenvalue problem (2) reads

$$\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B} & \mathbf{A}\end{bmatrix}\begin{bmatrix}\mathbf{X}_I\\ \mathbf{Y}_I\end{bmatrix} = \omega_I\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{X}_I\\ \mathbf{Y}_I\end{bmatrix}, \tag{13}$$

where

$$A_{ai\sigma,bj\tau} = \Delta_{ai\sigma,bj\tau} + K_{ai\sigma,bj\tau}, \tag{14}$$

$$B_{ai\sigma,bj\tau} = K_{ai\sigma,jb\tau}, \tag{15}$$

$$K_{ai\sigma,bj\tau} = g_{ai\sigma,jb\tau} + g_{ai\sigma,jb\tau}^{XC[2]}, \tag{16}$$

$$\Delta_{ai\sigma,bj\tau} = \delta_{\sigma\tau}(\delta_{ij}F_{ab\sigma} - \delta_{ab}F_{ji\tau}), \tag{17}$$

$$F_{pq\sigma} = H_{pq\sigma} + \sum_{rs\tau} g_{pq\sigma,rs\tau}D_{rs\tau} + v_{pq\sigma}^{XC}. \tag{18}$$

Here, $H_{pq\sigma}$ incorporates both the kinetic energy and the electron-nuclei attraction, $D_{pq\sigma} = \langle 0|E_{pq\sigma}|0\rangle$ represents the ground-state density matrix with $E_{pq\sigma} = a_{p\sigma}^{\dagger}a_{q\sigma}$ being the orbital replacement operator, and the two-electron integrals are written in the Mulliken notation

$$g_{pq\sigma,rs\tau} = j_{pq\sigma,rs\tau} - k_{pq\sigma,rs\tau} = g_{rs\tau,pq\sigma}, \tag{19}$$

$$j_{pq\sigma,rs\tau} = (p_\sigma q_\sigma|r_\tau s_\tau), \tag{20}$$

$$k_{pq\sigma,rs\tau} = \delta_{\sigma\tau}(p_\sigma s_\sigma|r_\sigma q_\sigma)_X. \tag{21}$$

The particular symbol $(p_\sigma s_\sigma|r_\sigma q_\sigma)_X$ represents a generalized exchange integral, which can also include a long-range interaction operator such as $r_{12}^{-1}\mathrm{erf}(r_{12})$. Assuming that the exchange correlation (XC) functional $E_{XC}[\rho]$ is parameterized as

$$E_{XC} = \int d^3\vec{r}\, e_{XC}[\{\rho_{\sigma,m}\}], \tag{22}$$

with m denoting the variable components such as the electron density $\rho_{\sigma,0} = \rho_\sigma$ and gradients $\rho_{\sigma,m} = \partial_m\rho_\sigma$ $(m = 1,2,3)$, the matrix elements of the XC potential can be written as

$$\begin{aligned}
v_{pq\sigma}^{XC} = g_{pq\sigma}^{XC[1]} &\triangleq \sum_m \int d^3\vec{r}\left(\frac{\partial e_{XC}}{\partial\rho_{\sigma,m}}\right)\Omega_{pq\sigma,m},\\
\Omega_{pq\sigma,m} &\triangleq \frac{\partial\rho_{\sigma,m}}{\partial D_{pq\sigma}}.
\end{aligned} \tag{23}$$

Accordingly, the matrix elements of the XC kernel in Eq. (16) read

$$g_{pq\sigma,rs\tau}^{XC[2]} = \sum_{mm'}\int d^3\vec{r}\left(\frac{\partial^2 e_{XC}}{\partial\rho_{\sigma,m}\partial\rho_{\tau,m'}}\right)\Omega_{pq\sigma,m}\Omega_{rs\tau,m'}. \tag{24}$$

Note that both $g_{pq\sigma,rs\tau}$ (19) and $g^{XC[2]}_{pq\sigma,rs\tau}$ (24) are invariant under the particle interchange, viz., $g_{pq\sigma,rs\tau} = g_{rs\tau,pq\sigma}$.

Formally, we can introduce the following excitation operators for the I-th excited state:

$$O_I^\dagger = \boldsymbol{\Lambda}^\dagger \mathbf{t}_I = -\mathbf{t}_{-I}^\dagger \boldsymbol{\Lambda}, \tag{25}$$

$$O_I = \mathbf{t}_I^\dagger \boldsymbol{\Lambda} = -\boldsymbol{\Lambda}^\dagger \mathbf{t}_{-I}, \tag{26}$$

$$\boldsymbol{\Lambda}^\dagger = (E_{\mathrm{VO}}, -E_{\mathrm{OV}}), \quad \mathbf{t}_I = (\mathbf{X}_I, \mathbf{Y}_I)^T, \tag{27}$$

where VO represents the virtual-occupied block labeled by the composite index $ai\sigma$. In terms of such excitation operators, the transition density matrices γ^{0I}_{pq} and γ^{IJ}_{pq} derived from the standard linear and quadratic response theories[28–31] can be expressed compactly as[26]

$$\gamma^{0I}_{pq\sigma} = \langle 0|[E_{pq\sigma}, O_I^\dagger]|0\rangle = \begin{bmatrix} 0 & X_{I,ai\sigma} \\ Y_{I,ai\sigma} & 0 \end{bmatrix}, \tag{28}$$

$$\begin{aligned}
\gamma^{IJ}_{pq\sigma} &= \gamma^{IJ}_{pq\sigma}(\mathrm{I}) + \gamma^{IJ}_{pq\sigma}(\mathrm{II}), \\
\gamma^{IJ}_{pq\sigma}(\mathrm{I}) &= \langle 0|[E_{pq\sigma}, \boldsymbol{\Lambda}^\dagger]|0\rangle \mathbf{t}_{IJ} = \begin{bmatrix} 0 & t_{IJ,ai\sigma} \\ t_{IJ,ia\sigma} & 0 \end{bmatrix}, \\
\gamma^{IJ}_{pq\sigma}(\mathrm{II}) &= \langle 0|[O_I, [E_{pq\sigma}, O_J^\dagger]]|0\rangle \\
&= \begin{bmatrix} -(X_J^T X_I + Y_I^T Y_J)_{ij\sigma} & 0 \\ 0 & (X_I X_J^T + Y_J Y_I^T)_{ab\sigma} \end{bmatrix},
\end{aligned} \tag{29}$$

where the VO (virtual-occupied) and OV (occupied-virtual) blocks of $\gamma^{IJ}_{pq}(\mathrm{I})$ are determined by[26]

$$(\mathbf{E} - \omega_{JI}\mathbf{S})\mathbf{t}_{IJ} = \mathbf{V}_{IJ}, \tag{30}$$

or more explicitly

$$\left(\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{bmatrix} - \omega_{JI} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix}\right) \begin{bmatrix} (\mathbf{t}_{IJ})_{\mathrm{VO}} \\ (\mathbf{t}_{IJ})_{\mathrm{OV}} \end{bmatrix} = \begin{bmatrix} (\mathbf{V}_{IJ})_{\mathrm{VO}} \\ (\mathbf{V}_{IJ})_{\mathrm{OV}} \end{bmatrix}, \tag{31}$$

$$\begin{aligned}
\mathbf{V}_{IJ} = \sum_{pq\sigma} \sum_{rs\tau} &(g_{pq\sigma,rs\tau} + g^{XC[2]}_{pq\sigma,rs\tau}) \\
&\times \Big(\langle 0|[\boldsymbol{\Lambda}, [O_I, E_{pq\sigma}]]|0\rangle \langle 0|[O_J^\dagger, E_{rs\tau}]|0\rangle \\
&+ \langle 0|[\boldsymbol{\Lambda}, [O_J^\dagger, E_{pq\sigma}]]|0\rangle \langle 0|[O_I, E_{rs\tau}]|0\rangle \\
&+ \langle 0|[\boldsymbol{\Lambda}, E_{pq\sigma}]|0\rangle \langle 0|[O_I, [O_J^\dagger, E_{rs\tau}]]|0\rangle\Big) \\
&+ \sum_{pq\sigma} \sum_{rs\tau} \sum_{r's'\tau'} g^{XC[3]}_{pq\sigma,rs\tau,r's'\tau'} \langle 0|[\boldsymbol{\Lambda}, E_{pq\sigma}]|0\rangle \\
&\times \langle 0|[O_I, E_{rs\tau}]|0\rangle \langle 0|[O_J^\dagger, E_{r's'\tau'}]|0\rangle,
\end{aligned} \tag{32}$$

$$\begin{aligned}
g^{XC[3]}_{pq\sigma,rs\tau,r's'\tau'} = \sum_{mm'm''} &\int d^3\vec{r} \left(\frac{\partial^3 e_{XC}}{\partial \rho_{\sigma,m} \partial \rho_{\tau,m'} \partial \rho_{\tau',m''}}\right) \\
&\times \Omega_{pq\sigma,m} \Omega_{rs\tau,m'} \Omega_{r's'\tau',m''}.
\end{aligned} \tag{33}$$

Although seemingly complicated, the construction of $\mathbf{V}_{IJ}$ (32) is actually rather straightforward: the "density matrices" $\langle 0|[O_J^\dagger, E_{rs\tau}]|0\rangle$, $\langle 0|[O_I, E_{rs\tau}]|0\rangle$, and $\langle 0|[O_I, [O_J^\dagger, E_{rs\tau}]]|0\rangle$ can first be evaluated and then contracted with the corresponding integrals $g_{pq\sigma,rs\tau}$, $g^{XC[2]}_{pq\sigma,rs\tau}$, and $g^{XC[3]}_{pq\sigma,rs\tau,r's'\tau'}$, so as to form effective one-electron operators $V_{eff} = \sum_{pq\sigma} V_{pq\sigma} E_{pq\sigma}$. The $\langle 0|[\boldsymbol{\Lambda}, V_{eff}]|0\rangle$-like terms can finally be assembled to form $\mathbf{V}_{IJ}$.

2. Notations

To make the expressions as compact as possible and meanwhile suitable for AO-based direct algorithms, we introduce the following notations:

(a) Full contractions of integrals of particle symmetry with density matrices are represented by

$$\begin{aligned}
&\langle O_n; \{D_1, D_2, \ldots, D_n\}\rangle \\
&\quad \triangleq \sum_{P_1 P_2 \cdots P_n} O_{P_1 P_2 \cdots P_n} (D_1)_{P_1} (D_2)_{P_2} \cdots (D_n)_{P_n},
\end{aligned} \tag{34}$$

where P_i represents the composite index $p_i p_i' \sigma_i$. The advantages of this notation are three folds: (1) a clear indication of the operators and densities, (2) particle symmetry can extensively be used within the bracket, e.g.,

$$\langle O_n; \{D_1, D_2, \ldots, D_n\}\rangle = \langle O_n; \{D_2, D_1, \ldots, D_n\}\rangle = \cdots, \tag{35}$$

(3) multiple linearity in both operators and densities, viz.,

$$\begin{aligned}
\langle O_n + O_n'; \{D_1, D_2, \ldots, D_n\}\rangle &= \langle O_n; \{D_1, D_2, \ldots, D_n\}\rangle \\
&\quad + \langle O_n'; \{D_1, D_2, \ldots, D_n\}\rangle, \\
\langle O_n; \{D_1, D_2 + D_2', \ldots, D_n\}\rangle &= \langle O_n; \{D_1, D_2, \ldots, D_n\}\rangle \\
&\quad + \langle O_n; \{D_1, D_2', \ldots, D_n\}\rangle.
\end{aligned} \tag{36}$$

With this notation, the full contraction of two-electron integrals (19) with two density matrices D_1 and D_2 can compactly be written as

$$\begin{aligned}
\langle g; \{D_1, D_2\}\rangle \triangleq \sum_{pq\sigma,rs\tau} &[(p_\sigma q_\sigma | r_\tau s_\tau) - \delta_{\sigma\tau}(p_\sigma s_\sigma | r_\sigma q_\sigma)_X] \\
&\times (D_1)_{pq\sigma} (D_2)_{rs\tau}.
\end{aligned} \tag{37}$$

(b) Partial contractions are presented by

$$\begin{aligned}
&O_n[\{D_1, D_2, \ldots, D_k\}] \\
&\quad \triangleq \sum_{P_1 P_2 \cdots P_k} O_{P_1 P_2 \cdots P_n} (D_1)_{P_1} (D_2)_{P_2} \cdots (D_k)_{P_k}, \quad k < n,
\end{aligned} \tag{38}$$

in terms of which the full contraction (37) can be reexpressed as

$$\langle g; \{D_1, D_2\}\rangle = \langle g[D_1]; D_2\rangle = \langle g[D_2]; D_1\rangle, \tag{39}$$

where the curly bracket in, e.g., $g[\{D_1\}]$, has been omitted since there is only one density in it.

(c) Quantities in the AO representation are denoted by boldface italic letters (e.g., $\boldsymbol{M}$), while those in the MO representation are denoted in plain italic letters (e.g., M). For instance, the full contraction of Δ (17) with $X_{I,ai\sigma}$ and $X_{J,bj\tau}$ can be written as

$$\begin{aligned}\langle\Delta;\{X_I,X_J\}\rangle &= \sum_{ai\sigma,bj\tau}\Delta_{ai\sigma,bj\tau}X_{I,ai\sigma}X_{J,bj\tau}\\ &= \sum_{ab\sigma}F_{ab\sigma}\sum_{i}X_{I,ai\sigma}X_{J,bi\sigma}\\ &\quad -\sum_{ij\sigma}F_{ji\sigma}\sum_{a}X_{I,ai\sigma}X_{J,bi\sigma}\\ &= \sum_{\mu\nu\sigma}F_{\mu\nu\sigma}\Bigg(\sum_{abi}C_{\mu a\sigma}X_{I,ai\sigma}X_{J,bi\sigma}C_{\nu b\sigma}\\ &\quad -\sum_{aij}C_{\mu j\sigma}X_{J,bj\sigma}X_{I,ai\sigma}C_{\nu i\sigma}\Bigg)\\ &= \langle\boldsymbol{F};\mathbf{C}_{\mathrm{V}}(X_IX_J^T)\mathbf{C}_{\mathrm{V}}^T-\mathbf{C}_{\mathrm{O}}(X_J^TX_I)\mathbf{C}_{\mathrm{O}}^T\rangle\\ &= \frac{1}{2}\langle\boldsymbol{F};\mathbf{C}_{\mathrm{V}}(X_IX_J^T+X_JX_I^T)\mathbf{C}_{\mathrm{V}}^T\\ &\quad -\mathbf{C}_{\mathrm{O}}(X_I^TX_J+X_J^TX_I)\mathbf{C}_{\mathrm{O}}^T\rangle,\end{aligned} \tag{40}$$

where the last equality is due to the symmetry of the AO Fock matrix. Here, $\mathbf{C}_{\mathrm{V}}$ and $\mathbf{C}_{\mathrm{O}}$ are the respective virtual and occupied MO coefficients such that density matrices like $\mathbf{C}_{\mathrm{V}}(X_IX_J^T)\mathbf{C}_{\mathrm{V}}^T$ are in the AO representation. In practice, each such full or partial contraction in the AO space can be calculated by the same subroutine.

(d) Total derivative with respect to ξ, partial derivative with respect to ξ, and single replacement of orbital p with p' in a quantity such as the Lagrangian, are denoted, respectively, as

$$\begin{aligned}L^{\xi}\triangleq\frac{\mathrm{d}L}{\mathrm{d}\xi},\quad L^{(\xi)}\triangleq\frac{\partial L}{\partial\xi},\\ L^{(pp'\sigma)}\triangleq\sum_{\mu\sigma}\frac{\partial L}{\partial C_{\mu p\sigma}(x)}C_{\mu p'\sigma}(x).\end{aligned} \tag{41}$$

(e) The contractions with $d_{pq}(x)$ in Eqs. (7) and (8) can be written as

$$\begin{aligned}\sum_{pq\sigma}\gamma_{pq\sigma}d_{pq\sigma}(x)=\langle d(x);\gamma\rangle=\langle\underline{\boldsymbol{d}};\underline{\boldsymbol{\gamma}}\rangle,\\ [\underline{\boldsymbol{d}}]_{\mu\nu}=\langle\chi_\mu|\chi_\nu(x)\rangle,\quad \underline{\boldsymbol{\gamma}}=\mathbf{C}\gamma[\mathbf{C}(x)]^T.\end{aligned} \tag{42}$$

Here, the underline of a quantity is to emphasize that the dependence on x occurs only to the ket of the quantity in the AO representation. An important relation at the reference point $x=x_0$ can be derived as

$$\begin{aligned}\langle\underline{\boldsymbol{d}};\underline{\boldsymbol{\gamma}}^{(pp'\sigma)}\rangle_{x=x_0} &= \sum_{\mu\nu}[d_{\mu\nu}(x)]_{x=x_0}\\ &\quad\times\sum_{\lambda}\left(\frac{\partial[\sum_{rs}C_{\mu r\sigma}\gamma_{rs\sigma}C_{\nu s\sigma}(x)]}{\partial C_{\lambda p\sigma}(x)}C_{\lambda p'}(x)\right)_{x=x_0}\\ &= \sum_{\mu\nu}S_{\mu\nu}\sum_{r}C_{\mu r\sigma}\gamma_{rp\sigma}C_{\nu p'}\\ &= \sum_{r}[\mathbf{C}^T\mathbf{S}\mathbf{C}]_{rp'}\gamma_{rp\sigma}\\ &= \sum_{r}\delta_{rp'}\gamma_{rp\sigma}=\gamma_{p'p\sigma}.\end{aligned} \tag{43}$$

Note that γ is not assumed to be symmetric here.

(f) To derive Eq. (10), we will make extensive use of the following relation:

$$\begin{aligned}\langle\boldsymbol{M};\boldsymbol{N}^{(pp'\sigma)}\rangle=2\eta(NM)_{pp'\sigma},\quad \boldsymbol{N}=\mathbf{C}N\mathbf{C}^{\boldsymbol{T}},\\ M=\mathbf{C}^T\boldsymbol{M}\mathbf{C},\end{aligned} \tag{44}$$

which can be obtained in the same way as Eq. (43). Here, the $\boldsymbol{M}$ and $\boldsymbol{N}$ matrices in the AO representation are either both *symmetric* ($\eta=1$) or both *antisymmetric* ($\eta=-1$). It follows that the AO-based quantity on the left hand side of Eq. (44) can be obtained by first going to the MO representation (i.e., M and N) and then picking up the corresponding elements of the matrix product followed by multiplying the factor 2η.

It will be shown later on that the notations (a)-(f) can simplify greatly the formulations of fo-NACME and excited-state gradients on one hand, and lead automatically to expressions suitable for AO-based direct algorithms on the other.

3. *Derivatives of the Lagrangians*

In terms of the notations introduced in Sec. II A 2, the generating function $g_I[x,\mathbf{C}(x)]$ (12) for excited-state gradients can be reexpressed as

$$\begin{aligned}g_I[x,\mathbf{C}(x)]=\langle\boldsymbol{F};\boldsymbol{T}_{II}\rangle+\langle\boldsymbol{g};\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}+\{\boldsymbol{L}_I^A,\boldsymbol{L}_I^A\}\rangle\\ +\langle\boldsymbol{g}^{XC[2]};\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}\rangle,\end{aligned} \tag{45}$$

where the intermediates are defined as

$$\boldsymbol{R}_I^S=\frac{1}{2}(\boldsymbol{R}_I+\boldsymbol{R}_I^T),\quad \boldsymbol{R}_I=\mathbf{C}_{\mathrm{V}}R_I\mathbf{C}_{\mathrm{O}}^T,\quad R_I=X_I+Y_I, \tag{46}$$

$$\boldsymbol{L}_I^A=\frac{1}{2}(\boldsymbol{L}_I-\boldsymbol{L}_I^T),\quad \boldsymbol{L}_I=\mathbf{C}_{\mathrm{V}}L_I\mathbf{C}_{\mathrm{O}}^T,\quad L_I=X_I-Y_I, \tag{47}$$

$$\boldsymbol{T}_{IJ}=\frac{1}{2}(\boldsymbol{\gamma}^{IJ}(\mathrm{II})+[\boldsymbol{\gamma}^{IJ}(\mathrm{II})]^T),\quad \boldsymbol{\gamma}^{IJ}(\mathrm{II})=\mathbf{C}\gamma^{IJ}(\mathrm{II})\mathbf{C}^T \tag{48}$$

with $\gamma^{IJ}(\mathrm{II})$ given in Eq. (29). Note that the symbol x has been omitted on the right hand side of Eq. (45) for simplicity. However, it should be kept in mind that the dependence of $g_I[x,\mathbf{C}(x)]$ on x arises from the integrals $\boldsymbol{g}$ and the MO coefficients $\mathbf{C}$ but not from the amplitudes X_I and Y_I.

By further introducing the following intermediates

$$\boldsymbol{P}_I=\boldsymbol{T}_{II}+\boldsymbol{Z}^S, \tag{49}$$

$$\boldsymbol{Z}^S=\frac{1}{2}(\boldsymbol{Z}+\boldsymbol{Z}^T),\quad \boldsymbol{Z}=\mathbf{C}_{\mathrm{V}}Z\mathbf{C}_{\mathrm{O}}^T, \tag{50}$$

$$\boldsymbol{\Gamma}_{IJ}=\{\boldsymbol{R}_I^S,\boldsymbol{R}_J^S\}+\{\boldsymbol{L}_I^A,\boldsymbol{L}_J^A\}, \tag{51}$$

where $\boldsymbol{T}_{II}$ (48) and $\boldsymbol{P}_I$ (49) are usually referred to as the respective unrelaxed and relaxed difference density matrices[33,34] for the I-th excited state, the Lagrangian (6) with g_I (45) can be rewritten as

$$\begin{aligned}L_I &= g_I+\langle\boldsymbol{F};\boldsymbol{Z}^S\rangle-\langle\boldsymbol{S};\boldsymbol{W}\rangle+\mathrm{tr}(W)\\ &= \langle\boldsymbol{F};\boldsymbol{P}_I\rangle+\langle\boldsymbol{g};\boldsymbol{\Gamma}_{II}\rangle+\langle\boldsymbol{g}^{XC[2]};\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}\rangle\\ &\quad -\langle\boldsymbol{S};\boldsymbol{W}\rangle+\mathrm{tr}(W)\\ &= \langle\boldsymbol{H};\boldsymbol{P}_I\rangle+\langle\boldsymbol{g};\{\boldsymbol{D},\boldsymbol{P}_I\}+\boldsymbol{\Gamma}_{II}\rangle\\ &\quad +\langle\boldsymbol{g}^{XC[1]};\boldsymbol{P}_I\rangle+\langle\boldsymbol{g}^{XC[2]};\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}\rangle-\langle\boldsymbol{S};\boldsymbol{W}\rangle+\mathrm{tr}(W).\end{aligned} \tag{52}$$

Use of the AO representation of F (18), viz., $\boldsymbol{F}=\boldsymbol{H}+\boldsymbol{g}[\boldsymbol{D}]+\boldsymbol{g}^{XC[1]}$, has been made here. The partial derivative of L_I (52)

at the reference point $x = x_0$ can then be expressed as

$$\begin{aligned} g_I^{\xi} = L_I^{(\xi)} = {} & \langle \boldsymbol{H}^{(\xi)};\boldsymbol{P}_I\rangle + \langle \boldsymbol{g}^{(\xi)};\{\boldsymbol{D},\boldsymbol{P}_I\} + \boldsymbol{\Gamma}_{II}\rangle \\ & + \langle \boldsymbol{g}^{XC[1](\xi)};\boldsymbol{P}_I\rangle + \langle \boldsymbol{g}^{XC[2](\xi)};\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}\rangle \\ & - \langle \boldsymbol{S}^{(\xi)};\boldsymbol{W}\rangle. \end{aligned} \tag{53}$$

Note that $\boldsymbol{g}^{XC[1](\xi)}$ and $\boldsymbol{g}^{XC[2](\xi)}$ contain nuclear derivatives of both the basis functions and the XC functional, the latter involving second and third order derivatives of the XC functional, respectively, e.g.,

$$\begin{aligned} g_{\mu\nu\sigma}^{XC[1](\xi)} = {} & \sum_m \int d^3\vec{r} \left(\frac{\partial e_{XC}}{\partial \rho_{\sigma,m}}\right) \Omega_{\mu\nu\sigma,m}^{(\xi)} \\ & + \sum_{mm'} \int d^3\vec{r} \left(\frac{\partial^2 e_{XC}}{\partial \rho_{\sigma,m}\partial \rho_{\tau,m'}}\right) \Omega_{\mu\nu\sigma,m} \rho_{\tau,m'}^{(\xi)}. \end{aligned} \tag{54}$$

Note in passing that by direct differentiating Eq. (53) with respect to another parameter $\zeta \in x$, the second order derivatives (Hessian) can be obtained as

$$\begin{aligned} g_I^{\xi\zeta} = L_I^{(\xi)\zeta} = {} & \langle \boldsymbol{F}^{(\xi)\zeta};\boldsymbol{P}_I\rangle + \langle \boldsymbol{F}^{(\xi)};\boldsymbol{P}_I^{[\zeta]}\rangle + \langle \boldsymbol{F}^{(\xi)};\boldsymbol{T}_{II}^{(\zeta)}\rangle \\ & + \langle \boldsymbol{g}^{(\xi\zeta)};\boldsymbol{\Gamma}_{II}\rangle + \langle \boldsymbol{g}^{(\xi)};\boldsymbol{\Gamma}_{II}^{\zeta}\rangle \\ & + \langle \boldsymbol{g}^{XC[2](\xi)\zeta};\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}\rangle \\ & + \langle \boldsymbol{g}^{XC[2](\xi)};\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}^{\zeta}\rangle \\ & - \langle \boldsymbol{S}^{(\xi\zeta)};\boldsymbol{W}\rangle - \langle \boldsymbol{S}^{(\xi)};\boldsymbol{W}^{[\zeta]}\rangle \\ & + \langle \boldsymbol{F}^{(\xi)};\boldsymbol{Z}^{S(\zeta)}\rangle - \langle \boldsymbol{S}^{(\xi)};\boldsymbol{W}^{(\zeta)}\rangle, \end{aligned} \tag{55}$$

where $\boldsymbol{M}^{\xi} = \boldsymbol{M}^{(\xi)} + \boldsymbol{M}^{[\xi]}$ with $\boldsymbol{M}^{(\xi)} = \mathbf{C}M^{\xi}\mathbf{C}^T$ and $\boldsymbol{M}^{[\xi]} = \mathbf{C}^{\xi}M\mathbf{C}^T + c.c.$ for $\boldsymbol{M} = \boldsymbol{T}_{IJ},\boldsymbol{R}_I,\boldsymbol{L}_I,\boldsymbol{Z},\boldsymbol{W}$. Two points can be made here. First, Eq. (55) can be transformed to an alternative form[39] that does not depend on Z^{ζ} and W^{ζ}. Second, at variance with the asymmetric form, a symmetric form of the Hessian is also possible. Each of the three forms of $g_I^{\xi\zeta}$ has its own merits.

Following the same procedure, the Lagrangian (6) with $g_{0I}[x,\mathbf{C}(x)]$ (7) can be written as

$$g_{0I} = \langle \underline{\boldsymbol{d}};\underline{\boldsymbol{\gamma}}^{0I}\rangle, \tag{56}$$

$$L_{0I} = g_{0I} + \langle \boldsymbol{F};\boldsymbol{Z}^S\rangle - \langle \boldsymbol{S};\boldsymbol{W}\rangle + \mathrm{tr}(W), \tag{57}$$

such that the fo-NACME g_{0I}^{ξ} as the partial first order derivative of L_{0I} at $x = x_0$ reads

$$\begin{aligned} g_{0I}^{\xi} = L_{0I}^{(\xi)} = {} & \langle \underline{\boldsymbol{d}}^{(\xi)};\underline{\boldsymbol{\gamma}}^{0I}\rangle_{x=x_0} + \langle \boldsymbol{F}^{(\xi)};\boldsymbol{Z}^S\rangle - \langle \boldsymbol{S}^{(\xi)};\boldsymbol{W}\rangle \\ = {} & \langle \boldsymbol{H}^{(\xi)};\boldsymbol{P}_{0I}\rangle + \langle \boldsymbol{g}^{(\xi)};\{\boldsymbol{D},\boldsymbol{P}_{0I}\}\rangle \\ & + \langle \boldsymbol{g}^{XC[1](\xi)};\boldsymbol{P}_{0I}\rangle - \langle \boldsymbol{S}^{(\xi)};\boldsymbol{W}\rangle + \langle \boldsymbol{d}^{(\xi)};\boldsymbol{\gamma}^{0I}\rangle. \end{aligned} \tag{58}$$

Use of the prescription $\boldsymbol{T}_{0I} = 0$ and hence, $\boldsymbol{P}_{0I} = \boldsymbol{Z}^S$ has been made here. Similarly, the Lagrangian (6) with $g_{IJ}[x,\mathbf{C}(x)]$ (8) can be written as

$$g_{IJ} = \omega_{JI}^{-1}Q_{IJ} + \langle \underline{\boldsymbol{d}};\underline{\boldsymbol{\gamma}}^{IJ}\rangle, \tag{59}$$

$$Q_{IJ} = \langle \boldsymbol{F};\boldsymbol{T}_{IJ}\rangle + \langle \boldsymbol{g};\boldsymbol{\Gamma}_{IJ}\rangle + \langle \boldsymbol{g}^{XC[2]};\{\boldsymbol{R}_I^S,\boldsymbol{R}_J^S\}\rangle, \tag{60}$$

$$L_{IJ} = g_{IJ} + \langle \boldsymbol{F};\boldsymbol{Z}^S\rangle - \langle \boldsymbol{S};\boldsymbol{W}\rangle + \mathrm{tr}(W). \tag{61}$$

By introducing the scaled quantities

$$\tilde{M} = \omega_{JI}M, \quad M = Z,W,\gamma^{IJ},L_{IJ}, \tag{62}$$

the fo-NACME g_{IJ}^{ξ} can conveniently be obtained as

$$g_{IJ}^{\xi} = L_{IJ}^{(\xi)} = \omega_{JI}^{-1}\tilde{L}_{IJ}^{(\xi)}, \tag{63}$$

$$\begin{aligned} \tilde{L}_{IJ}^{(\xi)} = {} & Q_{IJ}^{(\xi)} + \langle \boldsymbol{d}^{(\xi)};\tilde{\boldsymbol{\gamma}}^{IJ}\rangle + \langle \boldsymbol{F}^{(\xi)};\tilde{\boldsymbol{Z}}^S\rangle - \langle \boldsymbol{S}^{(\xi)};\tilde{\boldsymbol{W}}\rangle \\ = {} & \langle \boldsymbol{H}^{(\xi)};\boldsymbol{P}_{IJ}\rangle + \langle \boldsymbol{g}^{(\xi)};\{\boldsymbol{D},\boldsymbol{P}_{IJ}\} + \boldsymbol{\Gamma}_{IJ}\rangle \\ & + \langle \boldsymbol{g}^{XC[1](\xi)};\boldsymbol{P}_{IJ}\rangle + \langle \boldsymbol{g}^{XC[2](\xi)};\{\boldsymbol{R}_I^S,\boldsymbol{R}_J^S\}\rangle \\ & - \langle \boldsymbol{S}^{(\xi)};\tilde{\boldsymbol{W}}\rangle + \langle \boldsymbol{d}^{(\xi)};\tilde{\boldsymbol{\gamma}}^{IJ}\rangle, \end{aligned} \tag{64}$$

$$\boldsymbol{P}_{IJ} = \boldsymbol{T}_{IJ} + \tilde{\boldsymbol{Z}}^S. \tag{65}$$

It is clear that g_{IJ}^{ξ} Eq. (63) shares a similar mathematical form as the excited-state gradient g_I^{ξ} Eq. (53), with only one extra term $\langle \boldsymbol{d}^{(\xi)};\tilde{\boldsymbol{\gamma}}^{IJ}\rangle$ arising from the second term of Eq. (4).

It is evident that the derivatives of the Lagrangians can be evaluated efficiently with AO-based direct algorithms by using the back-transformed "density matrices." The remaining task is to determine the multipliers Z and W.

4. Equations for the multipliers Z and W

Before deriving the equations for the multipliers Z and W, we first present a theorem that is vital for the subsequent manipulations.

Theorem 1. *For a given set of MO coefficients* $\mathbf{C}(x)$ *(either canonical or noncanonical), if the generating function is invariant under unitary transformation* $\mathbf{U}$ *of an orbital manifold, viz.,*

$$g[x,\mathbf{C}(x)] = g[x,\mathbf{C}'(x)], \quad \mathbf{C}'(x) = \mathbf{C}(x)\mathbf{U}, \tag{66}$$

the following result then holds:

$$g^{(rs)} - g^{(sr)} = 0, \quad r > s. \tag{67}$$

Proof. The orbital rotation matrix $\mathbf{U}$ can generally be parameterized as

$$\mathbf{U} = \exp(\boldsymbol{\kappa}), \quad \boldsymbol{\kappa} = -\boldsymbol{\kappa}^T. \tag{68}$$

Eq. (66) then implies that

$$\begin{aligned} 0 = \frac{\partial g[x,\mathbf{C}'(x)]}{\partial \kappa_{rs}} & = \sum_{\mu p} \frac{\partial g[x,\mathbf{C}'(x)]}{\partial C'_{\mu p}(x)} \frac{\partial C'_{\mu p}(x)}{\partial \kappa_{rs}} \\ & = \sum_{\mu p} \frac{\partial g[x,\mathbf{C}'(x)]}{\partial C'_{\mu p}(x)} \sum_q C_{\mu q}(x) \frac{\partial U_{qp}(\kappa)}{\partial \kappa_{rs}}. \end{aligned} \tag{69}$$

Since at the current point $\mathbf{C}'(x) = \mathbf{C}(x)$, viz., $\boldsymbol{\kappa} = 0$, Eq. (69) becomes

$$\begin{aligned} 0 & = \sum_{\mu p} \left(\frac{\partial g[x,\mathbf{C}'(x)]}{\partial C'_{\mu p}(x)}\right)_{\boldsymbol{\kappa}=0} \sum_q C_{\mu q}(x) \left(\frac{\partial U_{qp}(\kappa)}{\partial \kappa_{rs}}\right)_{\boldsymbol{\kappa}=0} \\ & = \sum_{\mu p} \frac{\partial g[x,\mathbf{C}(x)]}{\partial C_{\mu p}(x)} \sum_q C_{\mu q}(x) \left(\frac{\partial U_{qp}(\kappa)}{\partial \kappa_{rs}}\right)_{\boldsymbol{\kappa}=0} \\ & = \sum_{pq} g^{(pq)}[x,\mathbf{C}(x)](\delta_{qr}\delta_{ps} - \delta_{qs}\delta_{pr}) = g^{(rs)} - g^{(sr)}. \end{aligned} \tag{70}$$

□

It follows that the matrix elements $g^{(pq)}$ are not all independent if the generating function g is invariant under certain type of orbital rotations. It can readily be verified that the generating functions of TD-DFT (13) are invariant under rotations among the occupied or virtual MO, viz., $\mathbf{U} = \mathbf{U}_{OO} \oplus \mathbf{U}_{VV}$.

Therefore, Eq. (67) implies in this case that

$$g^{(ij)}-g^{(ji)}=g^{(ab)}-g^{(ba)}=0. \tag{71}$$

Now the equations for the multipliers can be derived from the stationarity condition Eq. (10), which is equivalent to

$$\begin{aligned}
0 &= L^{(pp'\sigma)}\\
&= g^{(pp'\sigma)}+\langle \boldsymbol{F}^{(pp'\sigma)};\boldsymbol{Z}^S\rangle+\langle \boldsymbol{F};\boldsymbol{Z}^{S(pp'\sigma)}\rangle-\langle \boldsymbol{S};\boldsymbol{W}^{(pp'\sigma)}\rangle\\
&= g^{(pp'\sigma)}+\langle \boldsymbol{g}[\boldsymbol{D}^{(pp'\sigma)}]+\boldsymbol{g}^{XC[2]}[\boldsymbol{D}^{(pp'\sigma)}];\boldsymbol{Z}^S\rangle\\
&\quad+\langle \boldsymbol{F};\boldsymbol{Z}^{S(pp'\sigma)}\rangle-\langle \boldsymbol{S};\boldsymbol{W}^{(pp'\sigma)}\rangle\\
&= g^{(pp'\sigma)}+\langle \boldsymbol{g}[\boldsymbol{Z}^S]+\boldsymbol{g}^{XC[2]}[\boldsymbol{Z}^S];\boldsymbol{D}^{(pp'\sigma)}\rangle\\
&\quad+\langle \boldsymbol{F};\boldsymbol{Z}^{S(pp'\sigma)}\rangle-\langle \boldsymbol{S};\boldsymbol{W}^{(pp'\sigma)}\rangle\\
&= g^{(pp'\sigma)}+2\sum_{i\sigma}\delta_{pi\sigma}G_{ip'\sigma}[\boldsymbol{Z}^S]+\sum_{ai\sigma}\delta_{pa\sigma}Z_{ai\sigma}F_{ip'\sigma}\\
&\quad+\sum_{ib\sigma}\delta_{pi\sigma}F_{p'b\sigma}Z_{bi\sigma}-2W_{pp'\sigma},
\end{aligned} \tag{72}$$

where the last equality follows from Eq. (44) and the (symmetric) intermediate $G[\boldsymbol{Z}^S]$ is defined as

$$G[\boldsymbol{Z}^S]=\mathbf{C}^T\boldsymbol{G}[\boldsymbol{Z}^S]\mathbf{C},\quad \boldsymbol{G}[\boldsymbol{Z}^S]=\boldsymbol{g}[\boldsymbol{Z}^S]+\boldsymbol{g}^{XC[2]}[\boldsymbol{Z}^S]. \tag{73}$$

Eq. (72) shows that the multiplier W can be expressed as

$$\begin{aligned}
W_{pp'\sigma} &= \frac{1}{2}g^{(pp'\sigma)}+\sum_{i\sigma}\delta_{pi\sigma}G_{ip'\sigma}[\boldsymbol{Z}^S]\\
&\quad+\frac{1}{2}\Big(\sum_{ai\sigma}\delta_{pa\sigma}Z_{ai\sigma}F_{ip'\sigma}+\sum_{ib\sigma}\delta_{pi\sigma}F_{p'b\sigma}Z_{bi\sigma}\Big),
\end{aligned} \tag{74}$$

or more explicitly,

$$W_{ij\sigma} = \frac{1}{2}g^{(ij\sigma)}+G_{ij\sigma}[\boldsymbol{Z}^S], \tag{75}$$

$$W_{ab\sigma} = \frac{1}{2}g^{(ab\sigma)}, \tag{76}$$

$$W_{ia\sigma} = \frac{1}{2}g^{(ia\sigma)}+G_{ia\sigma}[\boldsymbol{Z}^S]+\frac{1}{2}\sum_{b\sigma}F_{ab\sigma}Z_{bi\sigma}, \tag{77}$$

$$W_{ai\sigma} = \frac{1}{2}g^{(ai\sigma)}+\frac{1}{2}\sum_{j\sigma}Z_{aj\sigma}F_{ji\sigma}. \tag{78}$$

The requirement of $W_{ij\sigma}=W_{ji\sigma}$ and $W_{ab\sigma}=W_{ba\sigma}$ is fulfilled automatically due to the property (71) and the fact that $G[\boldsymbol{Z}^S]$ is symmetric. In contrast, the requirement of $W_{ia\sigma}=W_{ai\sigma}$ dictates that the multipliers $Z_{ai\sigma}$ for the virtual-occupied block should satisfy the following condition:

$$\begin{aligned}
\sum_{b\sigma}F_{ab\sigma}Z_{bi\sigma}-\sum_{j\sigma}Z_{aj\sigma}F_{ji\sigma}+G_{ia\sigma}[2\boldsymbol{Z}^S]\\
=g^{(ai\sigma)}-g^{(ia\sigma)},
\end{aligned} \tag{79}$$

which is just the so-called Z-vector equation[33,34,40] and can be recast into a more compact form

$$(\mathbf{A}+\mathbf{B})\mathbf{Z}=g^{(\mathrm{VO})}-g^{(\mathrm{OV})} \tag{80}$$

in terms of the **A** (14) and **B** (15) matrices defined before.

The equations for W (74) and Z (80) can be applied to both the fo-NACME and excited-state gradients. The only difference lies in the generating function and hence $g^{(pp'\sigma)}$. For excited-state gradients g_I Eq. (45), $g_I^{(pp'\sigma)}$ reads

$$\begin{aligned}
g_I^{(pp'\sigma)} &= \langle \boldsymbol{F};\boldsymbol{T}_{II}\rangle^{(pp'\sigma)}+\langle \boldsymbol{g};\boldsymbol{\Gamma}_{II}\rangle^{(pp'\sigma)}\\
&\quad+\langle \boldsymbol{g}^{XC[2]};\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}\rangle^{(pp'\sigma)}\\
&= \langle \boldsymbol{F}^{(pp'\sigma)};\boldsymbol{T}_{II}\rangle+\langle \boldsymbol{F};\boldsymbol{T}_{II}^{(pp'\sigma)}\rangle+\langle \boldsymbol{g};\boldsymbol{\Gamma}_{II}^{(pp'\sigma)}\rangle\\
&\quad+\langle \boldsymbol{g}^{XC[3]};\{\boldsymbol{D}^{(pp'\sigma)},\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}\rangle\\
&\quad+\langle \boldsymbol{g}^{XC[2]};\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}^{(pp'\sigma)}\rangle\\
&= \langle \boldsymbol{g}[\boldsymbol{T}_{II}]+\boldsymbol{g}^{XC[2]}[\boldsymbol{T}_{II}]+\boldsymbol{g}^{XC[3]}[\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}];\boldsymbol{D}^{(pp'\sigma)}\rangle\\
&\quad+\langle \boldsymbol{F};\boldsymbol{T}_{II}^{(pp'\sigma)}\rangle\\
&\quad+2\langle \boldsymbol{g}[\boldsymbol{R}_I^S]+\boldsymbol{g}^{XC[2]}[\boldsymbol{R}_I^S];\boldsymbol{R}_I^{S(pp'\sigma)}\rangle\\
&\quad+2\langle \boldsymbol{g}[\boldsymbol{L}_I^A];\boldsymbol{L}_I^{A(pp'\sigma)}\rangle\\
&= 2\big[D\big(G[\boldsymbol{T}_{II}]+g^{XC[3]}[\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}]\big)\big]_{pp'\sigma}\\
&\quad+2[T_{II}F]_{pp'\sigma}\\
&\quad+4[R_I^SG[\boldsymbol{R}_I^S]]_{pp'\sigma}-4[L_I^Ag[\boldsymbol{L}_I^A]]_{pp'\sigma},
\end{aligned} \tag{81}$$

where the last equality follows from Eq. (44) and the fact that $\boldsymbol{L}_I^A$ and $\boldsymbol{g}[\boldsymbol{L}_I^A]$ are antisymmetric ($\eta=-1$) while the other matrices are all symmetric. To evaluate $g_I^{(pp'\sigma)}$ (81), the matrices $\boldsymbol{g}[\boldsymbol{T}_{II}]$, $\boldsymbol{g}[\boldsymbol{R}_I^S]$, $\boldsymbol{g}[\boldsymbol{L}_I^A]$, $\boldsymbol{g}^{XC[2]}[\boldsymbol{T}_{II}]$, $\boldsymbol{g}^{XC[2]}[\boldsymbol{R}_I^S]$, and $\boldsymbol{g}^{XC[3]}[\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}]$) can first be constructed using AO-based direct algorithms and then transformed to the MO space, followed by contracting with the corresponding density matrices

$$\begin{aligned}
D &= \begin{bmatrix} I & 0\\ 0 & 0\end{bmatrix},\\
T_{II} &= \begin{bmatrix} -\frac{1}{2}(R_I^TR_I+L_I^TL_I) & 0\\ 0 & \frac{1}{2}(R_IR_I^T+L_IL_I^T)\end{bmatrix},\\
R_I^S &= \begin{bmatrix} 0 & \frac{1}{2}R_I^T\\ \frac{1}{2}R_I & 0\end{bmatrix},\quad L_I^A=\begin{bmatrix} 0 & -\frac{1}{2}L_I^T\\ \frac{1}{2}L_I & 0\end{bmatrix}.
\end{aligned} \tag{82}$$

It is hence clear that compact and programable expressions for $g^{(pp'\sigma)}$ are obtained automatically in the present formulation.

In view of Eq. (43), the $g^{(pp'\sigma)}$ for g_{0I} (56) is just $\gamma_{p'p\sigma}^{0I}$ (28), such that the Z-vector equation (80) reads simply

$$(\mathbf{A}+\mathbf{B})\mathbf{Z}=\gamma_{\mathrm{OV}}^{0I}-\gamma_{\mathrm{VO}}^{0I}=\mathbf{X}_I-\mathbf{Y}_I=\mathbf{L}_I. \tag{83}$$

By comparing with the sum of the two equations in Eq. (13), viz.,

$$(\mathbf{A}+\mathbf{B})\mathbf{R}_I=\omega_I\mathbf{L}_I, \tag{84}$$

we obtain immediately

$$\mathbf{Z}=\omega_I^{-1}\mathbf{R}_I. \tag{85}$$

This accomplishes the derivation of g_{0I}^{ξ} for TD-DFT using the Lagrangian technique. The result agrees with that[17] obtained in a different way though.

Since the g_{IJ} (59) has a similar structure as the excited-state gradient g_I (45), the expression for $g_{IJ}^{(pp'\sigma)}$ can be obtained in the same way as done for $g_I^{(pp'\sigma)}$ (81), viz.,

$$g_{IJ}^{(pp'\sigma)} = \omega_{JI}^{-1}Q_{IJ}^{(pp'\sigma)}+\gamma_{p'p\sigma}^{IJ}, \tag{86}$$

$$\begin{aligned}
Q_{IJ}^{(pp'\sigma)} &= 2\big[D\big(G[\boldsymbol{T}_{IJ}]+g^{XC[3]}[\{\boldsymbol{R}_I^S,\boldsymbol{R}_J^S\}]\big)\big]_{pp'\sigma}\\
&\quad+2[T_{IJ}F]_{pp'\sigma}\\
&\quad+2[R_I^SG[\boldsymbol{R}_J^S]]_{pp'\sigma}-2[L_I^Ag[\boldsymbol{L}_J^A]]_{pp'\sigma}\\
&\quad+2[R_J^SG[\boldsymbol{R}_I^S]]_{pp'\sigma}-2[L_J^Ag[\boldsymbol{L}_I^A]]_{pp'\sigma},
\end{aligned} \tag{87}$$

such that the corresponding Z-vector equation reads

$$(\mathbf{A}+\mathbf{B})\mathbf{Z}=\omega_{JI}^{-1}(\mathbf{Q}_{IJ}^{\mathrm{VO}}-\mathbf{Q}_{IJ}^{\mathrm{OV}})+(\gamma_{\mathrm{OV}}^{IJ}-\gamma_{\mathrm{VO}}^{IJ}). \tag{88}$$

By further noting that the sum of the two equations in Eq. (31) can be written as

$$(\mathbf{A}+\mathbf{B})\omega_{JI}^{-1}[(\mathbf{t}_{IJ})_{\mathrm{VO}}+(\mathbf{t}_{IJ})_{\mathrm{OV}}]=\omega_{JI}^{-1}[(\mathbf{V}_{IJ})_{\mathrm{VO}}+(\mathbf{V}_{IJ})_{\mathrm{OV}}]+[(\mathbf{t}_{IJ})_{\mathrm{VO}}-(\mathbf{t}_{IJ})_{\mathrm{OV}}], \tag{89}$$

and that $\gamma_{\mathrm{OV}}^{IJ}-\gamma_{\mathrm{VO}}^{IJ}=(\mathbf{t}_{IJ})_{\mathrm{VO}}-(\mathbf{t}_{IJ})_{\mathrm{OV}}$ [cf. Eq. (29)] and $\mathbf{Q}_{IJ}^{\mathrm{VO}}-\mathbf{Q}_{IJ}^{\mathrm{OV}}=(\mathbf{V}_{IJ})_{\mathrm{VO}}+(\mathbf{V}_{IJ})_{\mathrm{OV}}$ [by comparing $Q_{IJ}^{(pp'\sigma)}$ (87) with $\mathbf{V}_{IJ}$ Eq. (32)], Eq. (88) can be solved immediately

$$\mathbf{Z}=\omega_{JI}^{-1}[(\mathbf{t}_{IJ})_{\mathrm{VO}}+(\mathbf{t}_{IJ})_{\mathrm{OV}}]. \tag{90}$$

As shown before,[26] Eq. (90) is essential for ensuring that, in the complete basis set limit, the g_{IJ}^{ξ} Eq. (63) coincides with the exact expression for the fo-NACME, $g_{IJ}^{\xi}=\omega_{JI}^{-1}\int d^3\vec{r}\rho_{IJ}(\vec{r})V_{en}^{\xi}(\vec{r})$, provided that the transition energy ω_{JI} and transition density $\rho_{IJ}(\vec{r})$ are both exact.

5. Discussion

Having derived the TD-DFT expressions for the excited-state gradients and the fo-NACME with a unified Lagrangian framework, it is of first interest to compare them from a computational aspect. It is easily figured out that the computation of g_{0I}^{ξ} (58) is cheaper than that of g_{I}^{ξ} (53), since the latter involves an additional XC term $\langle \boldsymbol{g}^{XC[2](\xi)};\{\boldsymbol{R}_I^S,\boldsymbol{R}_I^S\}\rangle$ and meanwhile requires the solution of the Z-vector equation (80), while the additional one-electron term $\boldsymbol{d}^{\xi}$ in the former is computationally very cheap. On the other hand, the computation of g_{IJ}^{ξ} (63) should be somewhat more expensive than that of g_{I}^{ξ} (53) for an obvious reason: given the same number of iterations, both sides of (30) for $\mathbf{t}_{IJ}$ [which leads directly to $\mathbf{Z}$ (90)] are operationally more expensive than those of the Z-vector equation (80) with $g_I^{(pp'\sigma)}$ (81) as input.

If the $\gamma^{IJ}(\mathrm{I})$ term of γ^{IJ} (29) is neglected, Eq. (90) will no longer hold. The Z-vector equation

$$(\mathbf{A}+\mathbf{B})\mathbf{Z}=\omega_{JI}^{-1}(\mathbf{Q}_{IJ}^{\mathrm{VO}}-\mathbf{Q}_{IJ}^{\mathrm{OV}}), \tag{91}$$

obtained by setting $\mathbf{t}_{IJ}=0$ in Eq. (88), should then be solved explicitly, in a similar effort as Eq. (30). To indicate such an approximation, TD-HF and TD-DFT will be denoted by TD-HF(n) and TD-DFT(n), respectively. The former is actually nothing but the RPA derived from the EOM.[26] Both TD-HF(n) and TD-DFT(n) are in line with those obtained by means of pseudowavefunctions along with intuitive approximations.[19,20]

Consider now the TDA for the fo-NACME, which amounts to setting $\mathbf{B}=0$ in Eq. (13), leading to

$$\mathbf{A}\mathbf{X}_I=\omega_I\mathbf{X}_I,\quad \mathbf{Y}_I=0. \tag{92}$$

Unlike TD-DFT, where the multiplier Z for g_{0I}^{ξ} is given simply by Eq. (85), the Z-vector equation

$$(\mathbf{A}+\mathbf{B})\mathbf{Z}=\mathbf{X}_I \tag{93}$$

must now be solved explicitly. Therefore, the computation of g_{0I}^{ξ} by TDA is actually more expensive than by TD-DFT. If the $\mathbf{B}$ term here is also neglected, we would obtain immediately $\mathbf{Z}=\omega_I^{-1}\mathbf{X}_I$. Since the TDA is not really defined for properties of excited states (due to the lack of wavefunction), there is no *a priori* reason to say whether this approximation is acceptable or not. However, there exists an indirect reason: the CIS shares formally the same defining equations, Eqs. (92) and (93), where the $\mathbf{B}$ term in the latter should not be neglected to make an already crude model even cruder. Therefore, the neglect of $\mathbf{B}$ in Eq. (93) is not to be considered. As for g_{IJ}^{ξ}, there are more possibilities for approximations. For instance, Eq. (30) for $\mathbf{t}_{IJ}$ can be solved with $\mathbf{V}_{IJ}$ yet constructed with the TDA/CIS Ref. 41 eigenvectors (92) Eq. (90) still holds in this case. Alternatively, if Eq. (30) is ignored completely, amounting to setting $\mathbf{t}_{IJ}=0$, the Z-vector equation (91) reduced from Eq. (88) can be solved with the $Q_{IJ}^{(pp'\sigma)}$ constructed with the TDA eigenvectors (92). To indicate this approximation, TDA and CIS[41] will be denoted by TDA(n) and CIS(n), respectively. The computational cost of both TDA and TDA(n) for g_{IJ}^{ξ} is rather similar to that of TD-DFT, since the cost for solving Eq. (91) in the former is similar to that for solving Eq. (30) in the latter.

As a final point, at variance with the above analytic formulation, a finite-difference formulation of the fo-NACME is also possible. Once the transition density matrices γ_{pq}^{0I} and γ_{pq}^{IJ} at the reference point $x=x_0$ are constructed analytically, a finite-difference scheme amounts simply to calculating the d_{pq}^{ξ} term with, e.g., a central difference form

$$d_{pq}^{\xi}=\langle\psi_p(x_0)|\psi_q^{\xi}(x_0)\rangle=\frac{1}{2h}[\langle\psi_p(x_0)|\psi_q(x_0+h)\rangle-\langle\psi_p(x_0)|\psi_q(x_0-h)\rangle]+O(h^2), \tag{94}$$

and the first term of Eq. (4) as

$$\begin{aligned}\omega_{JI}^{-1}\mathbf{t}_I^{\dagger}\mathbf{E}^{\xi}\mathbf{t}_J&=\omega_{JI}^{-1}\mathbf{t}_I^{\dagger}(\mathbf{E}-\omega_J\mathbf{S})^{\xi}\mathbf{t}_J=-\omega_{JI}^{-1}\mathbf{t}_I^{\dagger}(\mathbf{E}-\omega_J\mathbf{S})\mathbf{t}_J^{\xi}\\&=\mathbf{t}_I^{\dagger}(x_0)\mathbf{S}\mathbf{t}_J^{\xi}(x_0)\\&=\frac{1}{2h}[\mathbf{t}_I^{\dagger}(x_0)\mathbf{S}\mathbf{t}_J(x_0+h)-\mathbf{t}_I^{\dagger}(x_0)\mathbf{S}\mathbf{t}_J(x_0-h)]\\&\quad+O(h^2).\end{aligned} \tag{95}$$

However, such a finite-difference scheme should be used with great care, particularly, when there exist degenerate occupied or virtual orbitals (for which the $\mathbf{C}^{\xi}$ required by Eq. (94) cannot readily be determined by finite difference), needless to say that the TD-DFT/TDA eigenvalue problem has to be solved for each displaced nuclear configuration, the number of which is proportional to the number of atoms in the molecule.

B. The fo-NACME at the pp-TDA level

1. Defining quantities for g_{IJ}^{ξ}

The pp-RPA (Refs. 37 and 38) takes a state of $N\pm2$ electrons as the reference to target a number of N-electron states by adding or removing two electrons. The two processes get decoupled under the TDA. Specifically, the target N-electron state $|I\rangle$ is obtained in pp-TDA from the $(N-2)$-electron reference $|\tilde{0}\rangle$ via

$$|I\rangle = O_I^\dagger|\tilde{0}\rangle, \quad O_I^\dagger = \sum_{ab} X_{I,ab}a_a^\dagger a_b^\dagger, \tag{96}$$

which is effectively a two-electron-in-($N_V + m$)-active-orbital model, with N_V being the number of virtual orbitals and $m =$ 1 (closed shell) or 2 (open-shell). For comparison, the spin-flip TD-DFT is basically a two-electron-in-two-active-orbital model. For a closed-shell system, the spin-adapted pp-TDA equation reads[37]

$$\begin{aligned}
\mathbf{A}\mathbf{X}_I &= \omega_I\mathbf{X}_I,\\
A_{ab,cd} &= \Delta_{ab,cd} + K_{ab,cd},\\
\Delta_{ab,cd} &= \delta_{ac}\delta_{bd}(\epsilon_a + \epsilon_b),\\
K_{ab,cd} &= \begin{cases} \dfrac{1}{\sqrt{(1+\delta_{ab})(1+\delta_{cd})}}[(ac|bd)+(ad|bc)], & (s=0, a\geq b, c\geq d)\\ [(ac|bd)-(ad|bc)], & (s=1, a>b, c>d)\end{cases}.
\end{aligned} \tag{97}$$

Here, $s = 0$ and $s = 1$ refer to singlet and triplet states, respectively, and the orbitals are canonical MO (CMO) of an $(N-2)$-electron Hartree-Fock reference $|\tilde{0}\rangle$. Loosely speaking, Kohn-Sham orbitals and eigenvalues can also be used here, for pp-RPA can in a way be viewed as an extension of TD-DFT for a pairing field.[38] Since the reference $|\tilde{0}\rangle$ and the target states have different numbers of electrons, the quantities $\gamma_{pq}^{\tilde{0}I}$ and hence $g_{\tilde{0}I}^{\xi}$ are naturally zero. However, the fo-NACME g_{IJ}^{ξ} between two N-electron states are accessible by pp-TDA. A nice feature of pp-TDA/RPA lies in that the N-electron ground and excited states are treated on the equal, correlated footing. Using the excitation operators (96), the transition density matrix γ_{pq}^{IJ} can be calculated as

$$\begin{aligned}
\gamma_{pq}^{IJ} &= \langle\tilde{0}|[O_I,[E_{pq},O_J^\dagger]]|\tilde{0}\rangle\\
&= \begin{cases} 4[\bar{X}_I^S\bar{X}_J^S]_{pq}, & s=0\\ -4[X_I^AX_J^A]_{pq}, & s=1\end{cases}.
\end{aligned} \tag{98}$$

Here, the symmetrized ($\bar{X}_I^S$) and antisymmetrized (X_I^A) matrices are defined as

$$\bar{X}_I^S = \frac{1}{2}(\bar{X}_I + \bar{X}_I^T), \quad X_I^A = \frac{1}{2}(X_I - X_I^T), \tag{99}$$

with the scaled vector $\bar{X}_I$ being defined by the eigenvector X_I for $s = 0$,

$$\begin{aligned}
\bar{X}_{I,pq} &= \frac{1}{\sqrt{1+\delta_{pq}}}X_{I,pq},\\
&= \begin{cases} \dfrac{1}{\sqrt{1+\delta_{ab}}}X_{I,ab}, & a\geq b\\ 0, & \text{other cases}\end{cases}.
\end{aligned} \tag{100}$$

2. Derivatives of the Lagrangians

The pp-TDA equation (97) can be written explicitly as

$$(\epsilon_a + \epsilon_b)X_{I,ab} + \frac{2}{\sqrt{1+\delta_{ab}}}k_{ab}[\bar{X}_I^S] = \omega_I X_{I,ab}, \quad (s=0, a\geq b), \tag{101}$$

$$(\epsilon_a + \epsilon_b)X_{I,ab} + 2k_{ab}[X_I^A] = \omega_I X_{I,ab}, \quad (s=1, a>b), \tag{102}$$

with

$$\begin{aligned}
k_{ab}[\bar{X}_I^S] &= \frac{1}{2}\sum_{c\geq d}\frac{1}{\sqrt{1+\delta_{cd}}}[(ac|db)+(ad|cb)]X_{I,cd}\\
&= \sum_{\mu\nu}(a\mu|\nu b)\bar{\boldsymbol{X}}_{I,\mu\nu}^S,
\end{aligned} \tag{103}$$

$$\begin{aligned}
k_{ab}[X_I^A] &= \frac{1}{2}\sum_{c>d}[(ac|db)-(ad|cb)]X_{I,cd}\\
&= \sum_{\mu\nu}(a\mu|\nu b)\boldsymbol{X}_{I,\mu\nu}^A.
\end{aligned} \tag{104}$$

In terms of the notations in Sec. II A 2, the generating function g_{IJ} can then be written as

$$g_{IJ} = \omega_{JI}^{-1}Q_{IJ} + \langle\underline{\boldsymbol{d}};\underline{\boldsymbol{\gamma}}^{IJ}\rangle, \tag{105}$$

$$Q_{IJ} = \langle\boldsymbol{F};\boldsymbol{T}_{IJ}\rangle + \langle\boldsymbol{k};\boldsymbol{\Gamma}_{IJ}\rangle, \tag{106}$$

$$\boldsymbol{\Gamma}_{IJ} = \begin{cases} 2\{\bar{\boldsymbol{X}}_I^S,\bar{\boldsymbol{X}}_J^S\}, & s=0\\ 2\{\boldsymbol{X}_I^A,\boldsymbol{X}_J^A\}, & s=1\end{cases}. \tag{107}$$

Consequently, the pp-TDA fo-NACME g_{IJ}^{ξ} can be written in the same form as Eq. (63), viz.,

$$g_{IJ}^{\xi} = L_{IJ}^{(\xi)} = \omega_{JI}^{-1}\tilde{L}_{IJ}^{(\xi)}, \tag{108}$$

$$\begin{aligned}
\tilde{L}_{IJ}^{(\xi)} &= \langle\boldsymbol{H}^{(\xi)};\boldsymbol{P}_{IJ}\rangle + \langle\boldsymbol{g}^{(\xi)};\{\boldsymbol{D},\boldsymbol{P}_{IJ}\}\rangle + \langle\boldsymbol{k}^{(\xi)};\boldsymbol{\Gamma}_{IJ}\rangle\\
&\quad + \langle\boldsymbol{g}^{XC[1](\xi)};\boldsymbol{P}_{IJ}\rangle - \langle\boldsymbol{S}^{(\xi)};\tilde{\boldsymbol{W}}\rangle + \langle\boldsymbol{d}^{(\xi)};\tilde{\boldsymbol{\gamma}}^{IJ}\rangle,
\end{aligned} \tag{109}$$

with $\boldsymbol{P}_{IJ}$ defined in Eq. (65). As shown before, the expressions for g_I^{ξ} and g_{IJ}^{ξ} have a similar structure, such that the gradient for state $|I\rangle$ can, in view of Eq. (109), be obtained immediately as

$$\begin{aligned}
g_I^{\xi} = L_I^{(\xi)} &= \langle\boldsymbol{H}^{(\xi)};\boldsymbol{P}_I\rangle + \langle\boldsymbol{g}^{(\xi)};\{\boldsymbol{D},\boldsymbol{P}_I\}\rangle + \langle\boldsymbol{k}^{(\xi)};\boldsymbol{\Gamma}_{II}\rangle\\
&\quad + \langle\boldsymbol{g}^{XC[1](\xi)};\boldsymbol{P}_I\rangle - \langle\boldsymbol{S}^{(\xi)};\boldsymbol{W}\rangle,
\end{aligned} \tag{110}$$

with the relaxed difference density matrix $\boldsymbol{P}_I$ defined in Eq. (49).

3. Equations for the multipliers Z and W

Similar to the TD-DFT $g_{IJ}^{(pp')}$ (87), the pp-TDA $g_{IJ}^{(pp')}$ can be obtained for singlet states as

$$g_{IJ}^{(pp')} = \omega_{JI}^{-1} Q_{IJ}^{(pp')} + \gamma_{p'p}^{IJ}, \tag{111}$$

$$\begin{aligned} Q_{IJ}^{(pp')} &= \langle \boldsymbol{F}^{(pp')}; \boldsymbol{T}_{IJ}\rangle + \langle \boldsymbol{F}; \boldsymbol{T}_{IJ}^{(pp')}\rangle + 2\langle \boldsymbol{k}; \{\bar{\boldsymbol{X}}_I^S, \bar{\boldsymbol{X}}_J^S\}^{(pp')}\rangle \\ &= 2(DG[\boldsymbol{T}_{IJ}])_{pp'} + 2(T_{IJ}F)_{pp'} + 4(\bar{X}_J^S k[\bar{\boldsymbol{X}}_I^S])_{pp'} \\ &\quad + 4(\bar{X}_I^S k[\bar{\boldsymbol{X}}_J^S])_{pp'}, \end{aligned} \tag{112}$$

and for triplet states as

$$\begin{aligned} Q_{IJ}^{(pp')} &= \langle \boldsymbol{F}^{(pp')}; \boldsymbol{T}_{IJ}\rangle + \langle \boldsymbol{F}; \boldsymbol{T}_{IJ}^{(pp')}\rangle + 2\langle \boldsymbol{k}; \{\bar{\boldsymbol{X}}_I^A, \bar{\boldsymbol{X}}_J^A\}^{(pp')}\rangle \\ &= 2(DG[\boldsymbol{T}_{IJ}])_{pp'} + 2(T_{IJ}F)_{pp'} - 4(X_J^A k[\boldsymbol{X}_I^A])_{pp'} \\ &\quad - 4(X_I^A k[\boldsymbol{X}_J^A])_{pp'}. \end{aligned} \tag{113}$$

Likewise, the matrices $g_I^{(pp')}$ for the gradients of singlet and triplet states can be obtained simply as

$$\begin{aligned} g_I^{(pp')} &= Q_I^{(pp')} \\ &= \begin{cases} 2(DG[\boldsymbol{T}_I])_{pp'} + 2(T_I F)_{pp'} + 8(\bar{X}_I^S k[\bar{\boldsymbol{X}}_I^S])_{pp'}, & s = 0 \\ 2(DG[\boldsymbol{T}_I])_{pp'} + 2(T_I F)_{pp'} - 8(\bar{X}_I^A k[\bar{\boldsymbol{X}}_I^A])_{pp'}, & s = 1 \end{cases}. \end{aligned} \tag{114}$$

The quantity $\boldsymbol{T}_{IJ}$ appearing in the first term of Eq. (106) remains to be specified. Actually, it can be obtained in the same way as Eq. (40) by using the first term of Eq. (97), viz.,

$$\boldsymbol{T}_{IJ} = \mathbf{C}_V T_{IJ} \mathbf{C}_V^T,$$

$$T_{IJ} = \frac{1}{2}(X_I^T X_J + X_J^T X_I) + \frac{1}{2}(X_I X_J^T + X_J X_I^T). \tag{115}$$

However, the generating function g_{IJ} (105) with the so-obtained T_{IJ} (115) is not invariant with respect to rotations among the virtual orbitals, so as to violate property (67). Because of this, the Lagrangian (6) has to be augmented with the term $\sum_{a>b} Z_{ab} F_{ab}$. Following the same procedure in Sec. II A 4, the corresponding Z-vector equation can be obtained as

$$Z_{ab}(\epsilon_a - \epsilon_b) = g^{(ab)} - g^{(ba)}, \quad a > b. \tag{116}$$

At first glance, this equation will become ill-defined for accidentally degenerate virtual orbital pairs. Fortunately, such instability is spurious due to the underlying canonical representation. To see this, we spell out the expression

$$\begin{aligned} g_{IJ}^{(ab)} - g_{IJ}^{(ba)} &= \omega_{JI}^{-1}(Q_{IJ}^{(ab)} - Q_{IJ}^{(ba)}) + \gamma_{ba}^{IJ} - \gamma_{ab}^{IJ} \\ &= \omega_{JI}^{-1} 2(\epsilon_a - \epsilon_b)[2(\bar{X}_I^S \bar{X}_J^S + \bar{X}_J^S \bar{X}_I^S) - T^{IJ}]_{ab} \end{aligned} \tag{117}$$

for singlet states. It is hence clear that the factor $(\epsilon_a - \epsilon_b)$ in Eq. (116) can be cancelled out by that in Eq. (117), thereby leading to

$$Z_{ab} = \omega_{JI}^{-1} 2[2(\bar{X}_I^S \bar{X}_J^S + \bar{X}_J^S \bar{X}_I^S) - T_{IJ}]_{ab}, \quad a > b. \tag{118}$$

As a result, the T_{IJ} (115) term in $\boldsymbol{P}_{IJ}$ (65) is cancelled out precisely by that in $\tilde{Z}^S = \frac{1}{2}\omega_{JI}(Z + Z^T)$, leading finally to $\boldsymbol{P}_{IJ} = 2\mathbf{C}(\bar{X}_I^S \bar{X}_J^S + \bar{X}_J^S \bar{X}_I^S)\mathbf{C}^T$. Effectively, this is equivalent to choosing $Z_{ab} = 0$ and

$$T_{IJ} = 2(\bar{X}_I^S \bar{X}_J^S + \bar{X}_J^S \bar{X}_I^S). \tag{119}$$

Similarly, the following expression can be obtained for the T_{IJ} of triplet states:

$$T_{IJ} = -2(X_I^A X_J^A + X_J^A X_I^A). \tag{120}$$

With Eqs. (119) and (120) for T_{IJ}, the Z-vector equation (116) for Z_{ab} is no longer needed and the original Lagrangian (6) including only the virtual-occupied part Z_{ai} is sufficient also for pp-TDA. Actually, it can be verified that Eqs. (119) and (120) are nothing but the T_{IJ} of pp-TDA in a noncanonical representation. The situation is quite similar to the derivatives of Møller-Plesset correlation energy,[42] where the spurious singularity for accidentally degenerate orbitals appears only in the canonical representation but not in a noncanonical representation.

III. BENCHMARK CALCULATIONS

The implementation of the above formalisms is based on the existing DFT and TD-DFT modules (restricted, unrestricted spin-conserving/spin-flip, and spin-adapted open-shell TD-DFT[3–6,43]) in the BDF (Beijing density functional) package,[44–46] which are further modified for performing pp-TDA

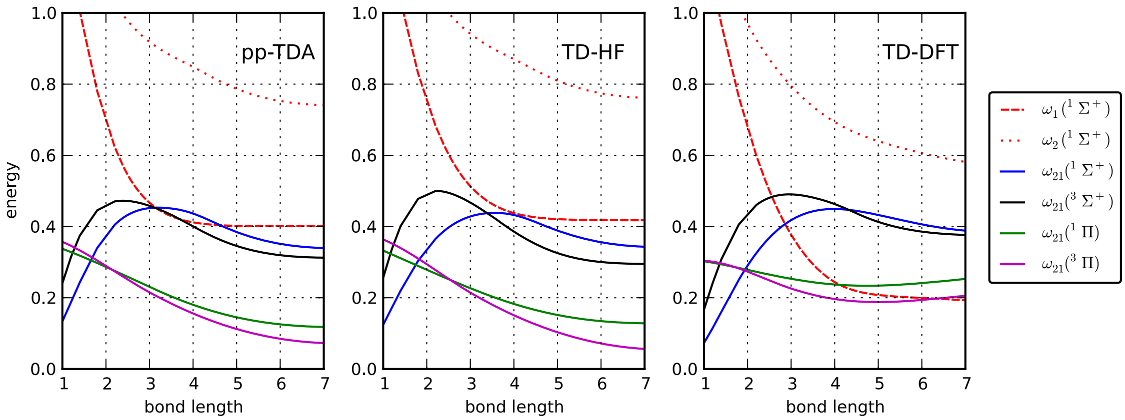

FIG. 1. Excitation energy curves of HeH$^+$. ω_1 and ω_2: excitation energies of the first two $^1\Sigma^+$ states; ω_{21}: energy difference between the lowest two excited states in each symmetry.

type of calculations. A new module RESP[47] is developed to calculate linear/quadratic spin-independent/spin-dependent response properties, excited-state gradients, and fo-NACME. For such properties of closed-shell systems, the spin adaptation is trivial by employing a spin-restricted HF/KS (RHF/RKS) reference. Yet, for open-shell systems, the spin adaption[3–5] based on a restricted high-spin reference remains to be developed. It also deserves to be mentioned that the computed fo-NACME are in general not translationally invariant, meaning that the sum of all g_{IJ}^{ξ} is not zero. Following the recipe in Ref. 18, the translationally invariant fo-NACME can be obtained by using only the symmetric part $[\boldsymbol{d}^{(\xi)}]^S = \frac{1}{2}(\boldsymbol{d}^{(\xi)}+\boldsymbol{d}^{(\xi)T}) = \frac{1}{2}\boldsymbol{S}^{(\xi)}$ of $\boldsymbol{d}^{(\xi)}$ in the AO representation, $(\boldsymbol{d}^{(\xi)})_{\mu\nu} = \langle \chi_\mu | \chi_\nu^{\xi} \rangle$. The derivatives of one- and two-electron integrals with respect to nuclear coordinates are evaluated with the GEN1INT package[48,49] and the ERI module in BDF, respectively. The third order derivatives of XC functionals are computed with the XCFUN library.[50,51] The implementation of the analytic formulations has been verified by comparing with results from the finite-difference calculations, see Eqs. (94) and (95).

In the following, we report the results obtained with TD-HF/CIS and TD-DFT/TDA using the VWN5 parametrization[52] of the local density approximation (LDA). For comparison, the fo-NACME were also computed by using CASSCF (complete active space self-consistent field) or FCI (full configuration interaction) with the MOLPRO (Ref. 53) program. The atomic units are used for all the quantities, including coordinates, energies, transition dipole moments, and fo-NACMEs.

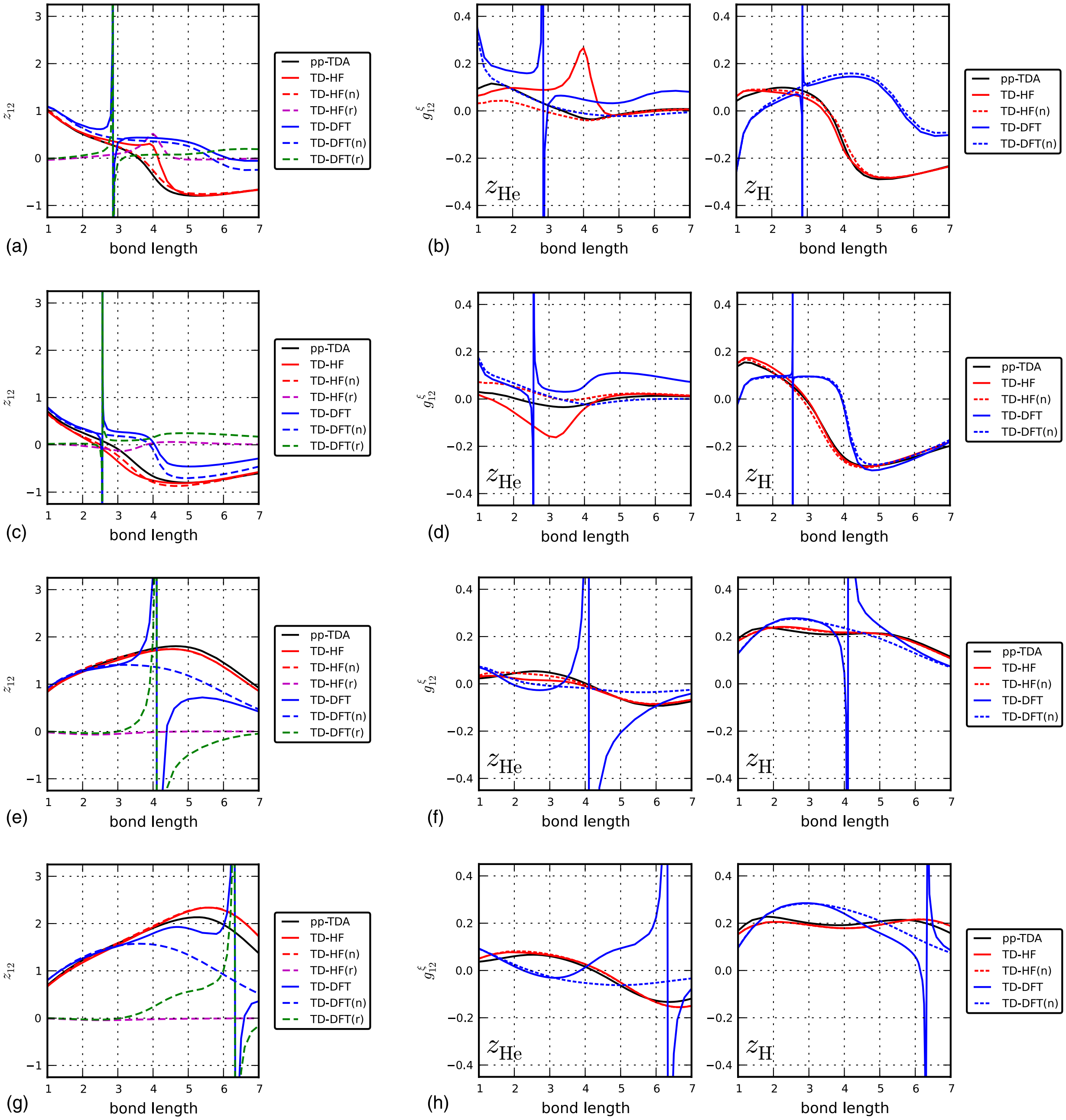

FIG. 2. Nonzero components of the TDM and fo-NACME between the lowest two excited states of HeH$^+$ in each symmetry.

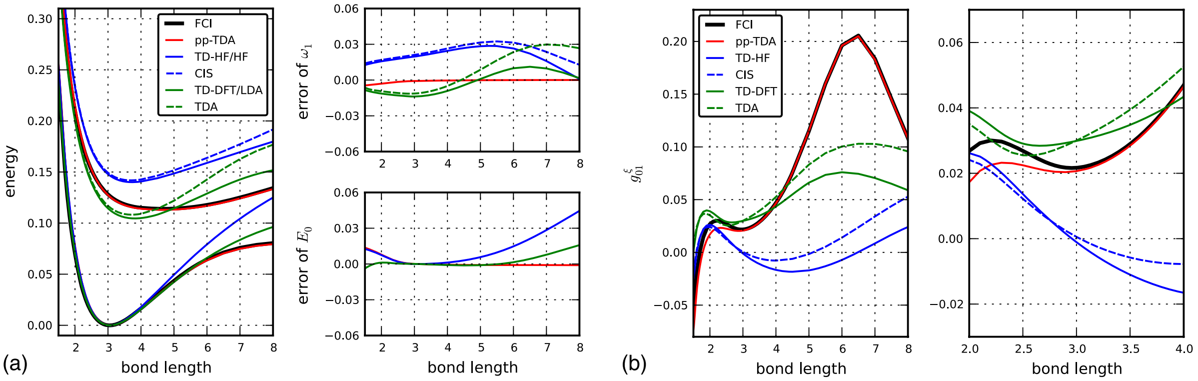

FIG. 3. (a) Potential energy curves for the $1^1\Sigma^+$ and $2^1\Sigma^+$ states of LiH. The energy minima of $1^1\Sigma^+$ by different methods are aligned to the same zero energy point. (b) The fo-NACME g_{01}^{ξ} between the two states.

A. fo-NACME g_{IJ}^{ξ} of HeH$^+$

For a better understanding of the above comprehensive formulation, we first take the simple two-electron system HeH$^+$ as an example, for which pp-TDA is identical with FCI. The lowest two excited states in each of the $^1\Sigma^+$, $^3\Sigma^+$, $^1\Pi$, and $^3\Pi$ symmetries were calculated with pp-TDA, TD-HF, and TD-DFT and the aug-cc-pVTZ basis set.[54,55] The excitation energy curves are depicted in Fig. 1. The nonzero components of the transition dipole moments (TDM) and fo-NACME between the two excited states in each symmetry are plotted in Fig. 2, where the fo-NACME obtained only with the transition density matrix γ_{pq}^{IJ}(II) (29) (designated by TD-HF(n) and TD-DFT(n)) are also shown. According to Eq. (32), for a nonvanishing $\mathbf{V}_{IJ}$ element, the irreducible representation (irrep) of the basis operator $\boldsymbol{\Lambda}$ must be contained in the direct product of the irreps of O_I and $O_J^{\dagger}$. Since both O_I and $O_J^{\dagger}$ belong to either $^{1,3}\Sigma^+$ or $^{1,3}\Pi$ for the present case, only $^1\Sigma^+$ type of $\boldsymbol{\Lambda}$ operator enters Eq. (30). It is then readily envisaged that, when the excitation energy ω_k of a $^1\Sigma^+$ state coincides at some geometry with $\omega_{JI} = \omega_J - \omega_I$, the response part γ_{pq}^{IJ}(I) of the transition density matrix γ_{pq}^{IJ} (29) will diverge as $1/(\omega_k - \omega_{JI})$. As can be seen from Fig. 1, the first $^1\Sigma^+$ excited state by TD-HF is very close to but does not intersect with the $\omega_{21}(^1\Sigma^+)$ curve at a distance around 4 a.u. However, the first $^1\Sigma^+$ excited state by TD-DFT intersects with the ω_{21} curves in each symmetry. This is because the energy of the first $^1\Sigma^+$ excited state, which has strong charge transfer character from He(1s) to H(1s) at long distances, is severely underestimated by the adiabatic LDA. Consequently, while the TD-HF fo-NACME g_{IJ}^{ξ} between the two $^1\Sigma^+$ states exhibits only a hump, the TD-DFT ones all diverge at the intersection points. The same situation can also be observed from the TDM, whose nonzero element for HeH$^+$, $\sum_{pq} z_{pq}\gamma_{pq}^{IJ}$, can be viewed as a simple measure of the transition density matrix. It is seen that, for all the cases, the TD-HF(n) and TD-DFT(n) TDM curves, calculated as $\sum_{pq} z_{pq}\gamma_{pq}^{IJ}$(II), are smooth and close to the corresponding pp-TDA/FCI results, while the TD-DFT(r) TDM curves, calculated as $\sum_{pq} z_{pq}\gamma_{pq}^{IJ}$(I), indeed diverge. It can also be seen that the overall accuracy of the TD-HF and TD-DFT fo-NACME follows closely that of the TDM, as compared with pp-TDA/FCI.

Given the simplicity of the HeH$^+$ system, it can be concluded that the response part γ_{pq}^{IJ}(I) of the transition density matrix γ_{pq}^{IJ} (29) should not be included in TD-HF and TD-DFT calculations of the fo-NACME as well as TDM. The (occasional) divergence of γ_{pq}^{IJ}(I) represents a deficiency of quadratic response functions of approximate models. For a more detailed analysis, see the Appendix.

B. g_{0I}^{ξ} of LiH and BH

The fo-NACME g_{0I}^{ξ} between the ground state $1^1\Sigma^+$ and first excited state $2^1\Sigma^+$ of LiH and BH are revisited here to compare with the previous TD-HF and TD-DFT calculations.[17] The potential energy curves and the fo-NACME (with respect

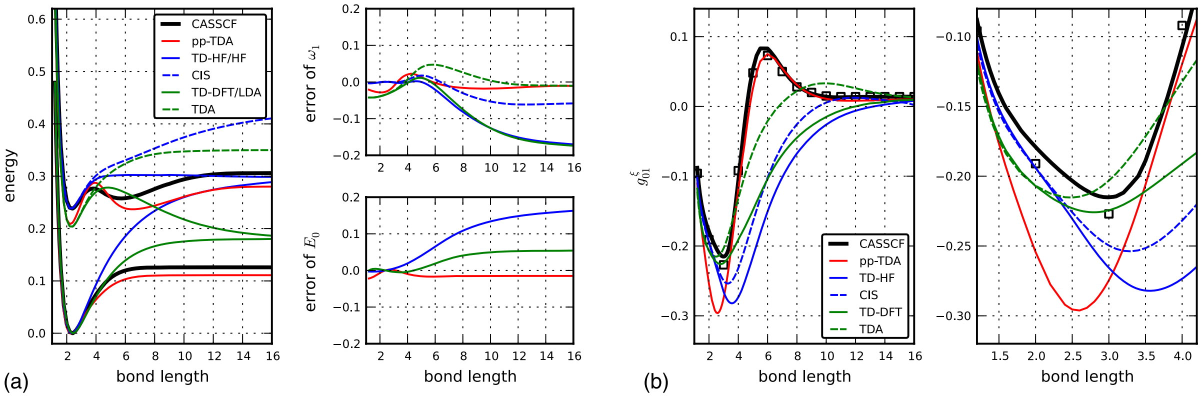

FIG. 4. (a) Potential energy curves for the $1^1\Sigma^+$ and $2^1\Sigma^+$ states of BH. The energy minima of $1^1\Sigma^+$ by different methods are aligned to the same zero energy point. (b) The fo-NACME g_{01}^{ξ} between the two states. Black squares designate the results of Ref. 58.

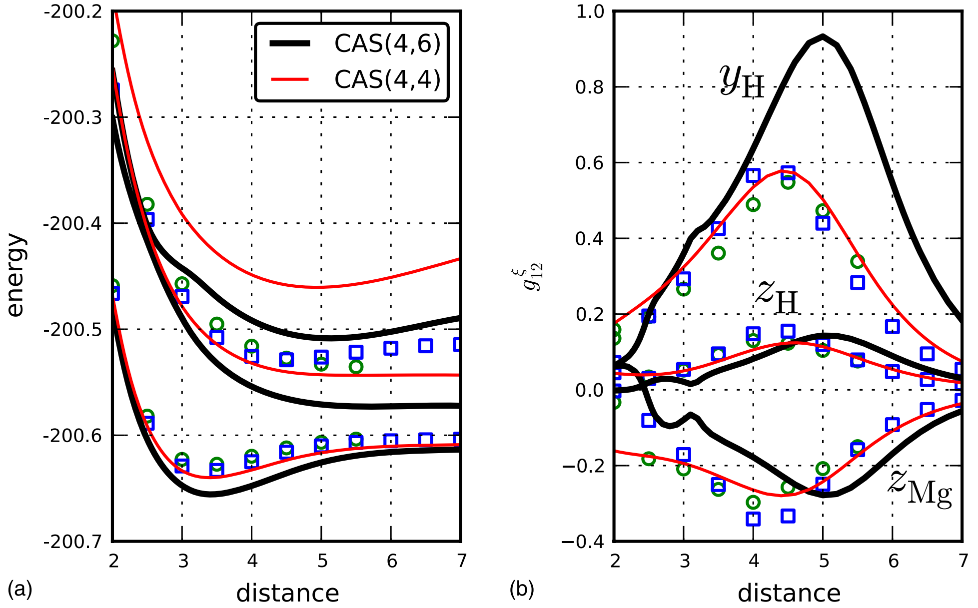

FIG. 5. CAS(4,4) and CAS(4,6) potential energy curves and fo-NACME for the lowest three $^3B_2$ states of MgH_2. Blue squares and green circles designate the results of Refs. 57 and 58, respectively.

to the interatomic distance) for the two $^1\Sigma^+$ states of LiH were calculated with FCI, pp-TDA, TD-HF, CIS, TD-DFT, and TDA and the cc-pVDZ basis,[56] see Fig. 3. It is first seen that pp-TDA is an excellent approximation to FCI throughout the whole distance. By contrast, both TD-HF and TD-DFT have large errors at distances beyond 4 a.u., simply because the underlying RHF/RKS reference does not dissociate correctly. In particular, the hump of g^{ξ}_{0I} due to the avoided crossing of the two states at the ionic-covalent transition point around 6.5 a.u. is totally missed by TD-HF and TD-DFT. The situation does not change much even if the unrestricted Hartree-Fock/Kohn-Sham determinant is taken as the reference.[17] It is also seen that, in this region, TDA (CIS) performs somewhat better than TD-DFT (TD-HF), whereas in the region around the ground-state equilibrium (between 2 and 4 a.u.), TDA (CIS) performs rather similarly as TD-DFT (TD-HF). Overall, TD-DFT (TDA) performs somewhat better than TD-HF (CIS).

The results for the first two $^1\Sigma^+$ states of BH are shown in Fig. 4. The aug-cc-pVDZ basis set[56] was employed here. The present CAS(4,18) (viz., 4 electrons in 18 active orbitals) results for g^{ξ}_{0I} are very similar to the previous CAS(4,14) ones[58] using a cc-pVDZ basis augmented with several nonstandard diffuse functions, see black squares in Fig. 4(b). Noticeably, the pp-TDA g^{ξ}_{0I} deviate significantly from the CAS(4,18) ones around 2 a.u. This can be traced back to the particular mechanism of pp-TDA for getting the two $^1\Sigma^+$ states: both are treated as excited states by adding two electrons to the double-cation reference $(BH)^{2+}$. Since the highest occupied MO (HOMO)-1 (B2*s*-like MO) is energetically close to the HOMO of BH, treating only the latter active but the former inactive will result in an unbalanced treatment of $1^1\Sigma^+$ and $2^1\Sigma^+$, viz., the former is less correlated. As a result, the pp-TDA gap between $2^1\Sigma^+$ and $1^1\Sigma^+$ is too small as compared with the CAS(4,18) one. Yet, more severe is that the pp-TDA $2^1\Sigma^+$ energy curve has in this

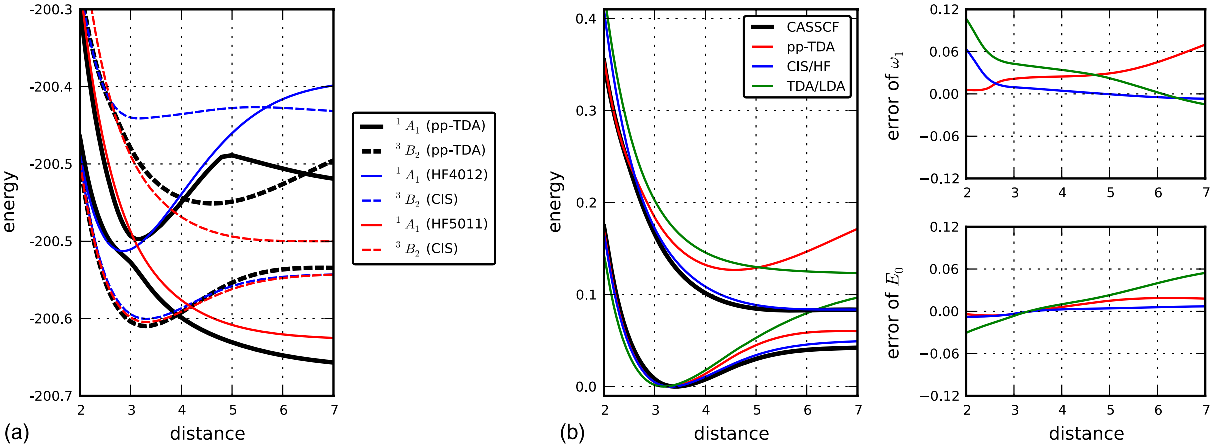

FIG. 6. (a) Potential energy curves for the first two $^3B_2$ states of MgH_2 calculated by pp-TDA and CIS. Two Hartree-Fock references of $^1A_1$ symmetry are used for CIS, with the respective occupations 4012 (in blue) and 5011 (in red) in the A_1, A_2, B_1, and B_2 irreps of C_{2V}. (b) Potential energy curves of the two $^3B_2$ states (left) and their deviations from CASSCF (right). Both CIS and TDA are here based on the 5011 reference.

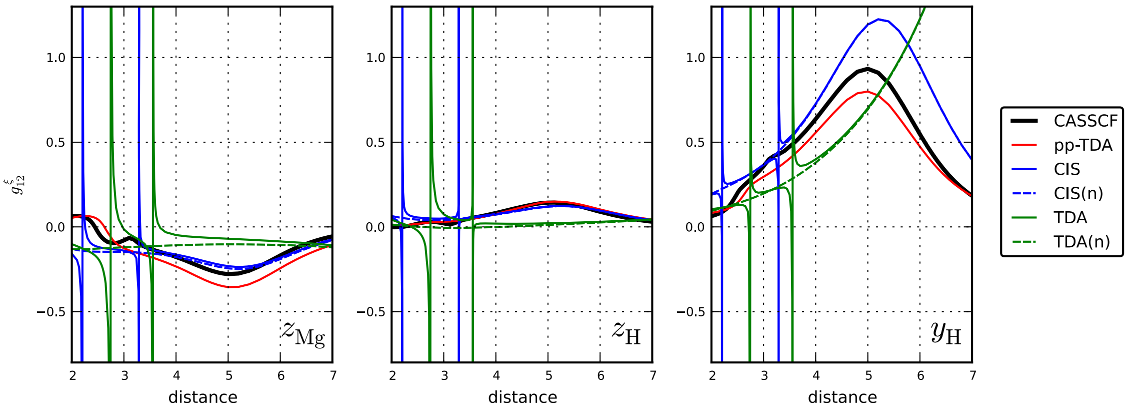

FIG. 7. Nonzero components of the fo-NACME between the first two $^3B_2$ states of MgH_2.

region a much larger curvature than the CAS(4,18) one. This argument is further confirmed by the good agreement between the pp-TDA and CAS(4,18) g_{0I}^{ξ} nearby the second minimum of $2^1\Sigma^+$ at ca. 6 a.u. (due to the avoided crossing between $2^1\Sigma^+$ and $3^1\Sigma^+$), where the error in the pp-TDA gap between $2^1\Sigma^+$ and $1^1\Sigma^+$ is very much the same as that nearby the first minimum, but the $2^1\Sigma^+$ curve has a roughly correct shape in this region. As a rule of thumb, pp-TDA performs only well if the HOMO of the N-electron system is well above the HOMO-1. Finally, the rather similar performance of the other approximate methods observed in LiH can also be found here.

C. g_{IJ}^{ξ} of MgH_2

Another example is the fo-NACME between the lowest two triplet states $^3B_2$ of T-shaped MgH_2, which was also investigated previously with state-averaged multi-configurational self-consistent field (MCSCF) (Ref. 57) and linear response MCSCF (Ref. 58). Here, we employed the same basis set but in the uncontracted form. In addition to the incomplete valence space CAS(4,4) used before,[57,58] the full valence space CAS(4,6) was also considered. The CASSCF potential energy curves along the distance between Mg and the center of H_2 (the H-H distance was fixed to 2.5 a.u.) are shown in Fig. 5(a), and the g_{IJ}^{ξ} for three coordinates (z_{Mg}, $z_{\mathrm{H}} = z_{\mathrm{H1}} + z_{\mathrm{H2}}$, and $y_{\mathrm{H}} = y_{\mathrm{H1}} - y_{\mathrm{H2}}$) are displayed in Fig. 5(b). It is clear that the present CAS(4,4) results almost coincide with those in Refs. 57 and 58 but are quite different from the CAS(4,6) ones. It is the latter that are employed for the purpose of benchmarking. Note in passing that, while pp-TDA can directly be applied to this example, there exists an ambiguity for TD-HF/CIS or TD-DFT/TDA, viz., two spin-restricted closed-shell references of $^1A_1$ symmetry, with the respective occupations 4012 and 5011 in the A_1, A_2, B_1, and B_2 irreps of C_{2v} symmetry, are possible. The potential energy curves for the two HF references and the corresponding four CIS $^3B_2$ states are shown in Fig. 6(a). It is clear that the linear combination of the two $^1A_1$ references will give rise to an avoided crossing around 3 a.u., as can be seen by the two pp-TDA $^1A_1$ states. While the first CIS $^3B_2$ state is almost independent of the references, the second CIS $^3B_2$ state based on the 5011 reference is significantly lower than that based on the 4012 reference. Therefore, the 5011 reference is always employed subsequently, although its energy is actually higher than the 4012 reference at distances shorter than 3 a.u. Since both TD-HF and TD-DFT yield imaginary excitation energies near the intersection point between 1^1A_1 and 1^3B_2, the corresponding results are not to be documented. Then, only the CIS and TDA results are meaningful, see Fig. 6(b) and Fig. 7 for the potential energy curves and fo-NACME, respectively. Again, the CIS and TDA fo-NACME become divergent when the response part $\gamma_{pq}^{IJ}(\mathrm{I})$ of the transition density matrix γ_{pq}^{IJ} (29) is taken into account, but behave well when it is ignored. It is seen from Fig. 7 that the y_{H}-component of the TDA(n) fo-NACME becomes increasingly too large as the two TDA(n) $^3B_2$ states get closer in energy (cf. Fig. 6(b)). In contrast, both CIS(n) and especially pp-TDA perform rather well throughout the distance.

IV. CONCLUSIONS AND OUTLOOK

Both the TD-DFT and pp-TDA formulations of the fo-NACME have been discussed in depth. They are specific realizations of the rigorous theory[26] rather than introduced in some *ad hoc* manner as done before. In particular, the Lagrangian technique combined with the well-defined notations leads to expressions that are very compact on one hand, and can directly be implemented using atomic orbital-based direct algorithms on the other. While more extensive applications are certainly necessary, some decisive conclusions can already be drawn based on the preliminary results for the prototypical systems considered here. It is generally true that TD-DFT can faithfully describe the fo-NACME as long as it can well describe the excited states. The latter can be monitored by the quality of the TD-DFT excitation energies and transition dipole moments. The TD-DFT(n) variant of TD-DFT, which ignores the response part of the transition density matrix, is particularly recommended for practical calculations of the fo-NACME, even around a conical intersection point.[20] Moreover, pp-TDA can be recommended as an alternative of spin-flip TD-DFT for describing excited-state energy surfaces and nonadiabatic couplings when near degeneracy and double excitations are encountered, given that the treatment of correlation needs to be further improved.

Many interesting developments can be done in near future. Obviously, the same Lagrangian technique applies also to the gradient difference $g_I^{\xi} - g_J^{\xi}$ with $g_I - g_J$ as the generating function. Further combined with the fo-NACME g_{IJ}^{ξ}, conical intersections can be located. The nuclear derivative of g_{IJ}^{ξ}, i.e., $g_{IJ}^{\xi\zeta} = D_{\zeta}\langle\Psi_I|D_{\xi}|\Psi_J\rangle$, which is required by the so-called group Born-Oppenheimer equation[59] for nonadiabatic dynamics, can also be formulated, similarly as the analytic Hessian $g_I^{\xi\zeta}$ (55) of the TD-DFT excitation energy. All these can be incorporated into the spin-adapted open-shell TD-DFT.[3–5] Work along these directions is being carried out at our laboratory.

ACKNOWLEDGMENTS

One of the authors (Z.L.) acknowledges helpful discussions with Dr. Yang Yang (Duke University) on pp-TDA, and Dr. Bin Gao (University of Tromsø) on the GEN1INT package. This work was supported by the NSFC (Project Nos. 21033001, 21273011, and 21290192).

APPENDIX: POSSIBLE DIVERGENCE OF g_{IJ}^{ξ} BY APPROXIMATE RESPONSE THEORY

The linear (LRF) and quadratic (QRF) response functions of TD-HF or adiabatic TD-DFT are compared with those of exact theory, so as to reveal the origin of divergence of g_{IJ}^{ξ} by TD-HF or adiabatic TD-DFT. The exact LRF and QRF in the frequency domain take the following sum-over-states (SOS) forms:[29–31]

$$A^b(\omega_k)=\mathcal{P}_{AB}\sum_{I>0}\frac{\langle 0|V^a(\omega_\sigma)|I\rangle\langle I|V^b(\omega_k)|0\rangle}{\omega_k-\omega_I}, \tag{A1}$$

$$A^{bc}(\omega_k,\omega_l)=-\mathcal{P}_{ABC}\times\sum_{I,J>0}\frac{\langle 0|V^a(\omega_\sigma)|I\rangle\langle I|\bar{V}^b(\omega_k)|J\rangle\langle J|V^c(\omega_l)|0\rangle}{(\omega_\sigma+\omega_I)(\omega_l-\omega_J)}, \tag{A2}$$

$$\bar{V}^b(\omega_k)=V^b(\omega_k)-\langle 0|V^b(\omega_k)|0\rangle, \tag{A3}$$

where the time-independent operator V^a is formally associated with a frequency ω_σ equal to the negative of the sum of the other frequencies in a response function, e.g., $\omega_\sigma = -\omega_k$ in Eq. (A1), and the permutation operators $\mathcal{P}_{AB} = 2!\mathcal{S}_{AB}$ and $\mathcal{P}_{ABC} = 3!\mathcal{S}_{ABC}$ symmetrize both the perturbation operators and the frequencies, e.g.,

$$\mathcal{S}_{AB}f(\{a,\omega_\sigma\},\{b,\omega_k\})=\frac{1}{2}[f(\{a,\omega_\sigma\},\{b,\omega_k\})+f(\{b,\omega_k\},\{a,\omega_\sigma\})]. \tag{A4}$$

To extract excitation energies and transition moments between the ground and excited states from the LRF of TD-HF/TD-DFT, it is important to note that, by using $O_I^\dagger$ (25) or $O_I^\dagger = |I\rangle\langle 0|$, the LRF of TD-HF/TD-DFT and exact theory can be written in the same mathematical form

$$A^b(\omega_k)=\sum_{I>0}\left(\frac{\langle 0|[V^a,O_I^\dagger]|0\rangle\langle 0|[O_I,V^b]|0\rangle}{\omega_k-\omega_I}-\frac{\langle 0|[V^b,O_I^\dagger]|I\rangle\langle 0|[O_I,V^a]|0\rangle}{\omega_k+\omega_I}\right), \tag{A5}$$

such that the TD-HF/TD-DFT excitation energies and transition moments can be identified from the respective poles and residues of LRF unambiguously. However, such termwise correspondence does not hold for the SOS QRF of TD-HF/TD-DFT and exact theory (A2). Instead, the QRF of TD-HF/TD-DFT and exact theory can both be written as[28–31]

$$A^{bc}(\omega_k,\omega_l)=-\boldsymbol{\kappa}^{a\dagger}(\omega_k,\omega_l)\mathbf{V}^{bc}(\omega_k,\omega_l)+\langle 0|[\kappa^b(\omega_k),[\kappa^c(\omega_l),V^a]]|0\rangle, \tag{A6}$$

with

$$\boldsymbol{\kappa}^{a\dagger}(\omega_k,\omega_l)[\mathbf{E}-(\omega_k+\omega_l)\mathbf{S}]=\mathbf{V}^{a\dagger}, \tag{A7}$$

$$[\mathbf{E}-\omega_k\mathbf{S}]\boldsymbol{\kappa}^b(\omega_k)=\mathbf{V}^b, \tag{A8}$$

$$[\mathbf{E}-\omega_l\mathbf{S}]\boldsymbol{\kappa}^c(\omega_l)=\mathbf{V}^c, \tag{A9}$$

where $\mathbf{V}^x = \langle 0|[\boldsymbol{\Lambda},V^x]|0\rangle$ $(x = a,b,c)$ and $\mathbf{V}^{bc}(\omega_k,\omega_l)$ is defined as[26]

$$\mathbf{V}^{bc}(\omega_k,\omega_l)=\mathcal{S}_{BC}\Big(\langle 0|[\boldsymbol{\Lambda},[\kappa^b(\omega_k),[\kappa^c(\omega_l),H^{(0)}]]]|0\rangle+\langle 0|[\boldsymbol{\Lambda},[\kappa^b(\omega_k),\omega_l\kappa^c(\omega_l)]]|0\rangle+2\langle 0|[\boldsymbol{\Lambda},[\kappa^b(\omega_k),H^c(\omega_l)+V^c(\omega_l)]]|0\rangle+\langle 0|[\boldsymbol{\Lambda},\bar{H}^{bc}(\omega_k,\omega_l)]|0\rangle\Big). \tag{A10}$$

Here, $H^b(\omega_k)$, $H^c(\omega_l)$, and $\bar{H}^{bc}(\omega_k,\omega_l)$ denote perturbed Hamiltonians depending on $\kappa^b(\omega_k)$ and $\kappa^c(\omega_l)$ (for details, see Ref. 26). By introducing the quantities $\mathbf{V}_J^b(\omega_k)$ and $\mathbf{V}_{IJ}$ via

$$\lim_{\omega_l\to\omega_J}(\omega_l-\omega_J)\mathbf{V}^{bc}(\omega_k,\omega_l)=\mathbf{V}_J^b(\omega_k)(\mathbf{t}_J^\dagger\mathbf{V}^c), \tag{A11}$$

$$\lim_{\omega_k\to-\omega_I}(\omega_k+\omega_I)\lim_{\omega_l\to\omega_J}(\omega_l-\omega_J)\mathbf{V}^{bc}(\omega_k,\omega_l)=(\mathbf{t}_{-I}^\dagger\mathbf{V}^b)\mathbf{V}_{IJ}(\mathbf{t}_J^\dagger\mathbf{V}^c), \tag{A12}$$

the single residue of the QRF (A6) can be obtained as

$$\lim_{\omega_l\to\omega_J}(\omega_l-\omega_J)A^{bc}(\omega_k)=[-\boldsymbol{\kappa}^{a\dagger}(\omega_k,\omega_J)\mathbf{V}_J^b(\omega_k)-\langle 0|[\kappa^b(\omega_k),[O_J^\dagger,V^a]]|0\rangle](\mathbf{t}_J^\dagger\mathbf{V}^c), \tag{A13}$$

which can be compared with that obtained from Eq. (A2), so as to identify the transition moment as

$$\mathcal{P}_{AB}\sum_{I>0}\frac{\langle 0|V^a(\omega_\sigma)|I\rangle\langle I|\bar{V}^b(\omega_k)|J\rangle}{\omega_\sigma+\omega_I}=\boldsymbol{\kappa}^{a\dagger}(\omega_k,\omega_J)\mathbf{V}_J^b+\langle 0|[\kappa^b(\omega_k),[O_J^\dagger,V^a]]|0\rangle. \tag{A14}$$

By using the spectral form of the resolvent

$$(\mathbf{E}-\omega\mathbf{S})^{-1}=\sum_{I>0}\left[\frac{\mathbf{t}_I\mathbf{t}_I^\dagger}{\omega_I-\omega}+\frac{\mathbf{t}_{-I}\mathbf{t}_{-I}^\dagger}{\omega_I+\omega}\right], \tag{A15}$$

the right hand side of Eq. (A14) can, in view of Eqs. (A7) and (A8), be further written in a SOS form

$$\boldsymbol{\kappa}^{a\dagger}(\omega_k,\omega_J)\mathbf{V}_J^b+\langle 0|[\kappa^b(\omega_k),[O_J^\dagger,V^a]]|0\rangle$$
$$=\sum_{I'>0}\left[\frac{(\mathbf{V}^{a\dagger}\mathbf{t}'_I)(\mathbf{t}_{I'}^\dagger,\mathbf{V}_J^b(\omega_k))}{\omega_{I'}-(\omega_k+\omega_J)}+\frac{(\mathbf{V}^{a\dagger}\mathbf{t}_{-I'})(\mathbf{t}_{-I'}^\dagger,\mathbf{V}_J^b(\omega_k))}{\omega_{I'}+(\omega_k+\omega_J)}\right]$$
$$+\sum_{I'>0}\left[\frac{(\langle 0|[\boldsymbol{\Lambda}^\dagger\mathbf{t}_{I'},[O_J^\dagger,V^a]]|0\rangle)(\mathbf{t}_{I'}^\dagger,\mathbf{V}^b)}{\omega_{I'}-\omega_k}\right.$$

$$+\frac{(\langle 0|[\boldsymbol{\Lambda}^{\dagger}\mathbf{t}_{-I'},[O_J^{\dagger},V^a]]|0\rangle)(\mathbf{t}_{-I'}^{\dagger},\mathbf{V}^b)}{\omega_{I'}+\omega_k}\Bigg]. \tag{A16}$$

Several important differences between Eqs. (A14) and (A16) can readily be identified.

First, compared with the exact expression on the left hand side of Eq. (A14), the pole structure of Eq. (A16) is more complicated, since apart from the common poles $\omega_k = \omega_{I'} - \omega_J$ and $\omega_k = -\omega_{I'}$, Eq. (A16) contains additional poles at $\omega_k = -(\omega_{I'} + \omega_J)$ and $\omega_k = \omega_{I'}$ if the numerators of the second and third terms of Eq. (A16) are nonvanishing. This is usually the case for TD-HF/TD-DFT. For instance, the $\langle 0|[\boldsymbol{\Lambda}^{\dagger}\mathbf{t}_{I'},[O_J^{\dagger},V^a]]|0\rangle = \langle 0|[O_{I'}^{\dagger},[O_J^{\dagger},V^a]]|0\rangle$ term in the numerator of the third term of Eq. (A16) involves double excitations $O_{I'}^{\dagger},O_J^{\dagger}|0\rangle$[28] and is not necessarily zero. In contrast, this term is zero for exact theory as can be verified by inserting the state transfer operator $O_I^{\dagger} = |I\rangle\langle 0|$. Similarly, in exact theory, we have $\mathbf{V}_J^b(\omega_k) = \langle 0|[\boldsymbol{\Lambda},[-O_J^{\dagger},V^b]]|0\rangle$ (see Eqs. (A10) and (A11)), such that $\mathbf{t}_{I'}^{\dagger},\mathbf{V}_J^b(\omega_k) = \langle 0|[O_{I'},[-O_J^{\dagger},V^b]]|0\rangle = \langle I'|\bar{V}^b|J\rangle$ in the first term, and $\mathbf{t}_{-I'}^{\dagger},\mathbf{V}_J^b(\omega_k) = 0$ in the second term of Eq. (A16). That is, in exact theory, the second term of Eq. (A16) vanishes and the first term therein has the correct form in view of the left hand side of Eq. (A14).

Second, it is important to note that the pole $\omega_k = -\omega_{I'}$ presents not only in the fourth term of Eq. (A16) but also in the first two terms through $\boldsymbol{\kappa}^b(\omega_k)$ in $\mathbf{V}_J^b(\omega_k)$, whose pole structure can be inferred from Eq. (A8). The dependence of $\mathbf{V}_J^b(\omega_k)$ on $\boldsymbol{\kappa}^b(\omega_k)$ in TD-HF/TD-DFT implies that the first term of Eq. (A16) contains terms with a pole structure like $1/[(-\omega_k - \omega_{JI'})(\omega_k \pm \omega_I)]$. The two first order poles $\omega_k = -\omega_{JI'}$ and $\omega_k = -\omega_I$ will merge into a second-order pole when $\omega_{JI'} = \omega_I$. This is exactly the origin of divergence observed in the calculations of TDM and fo-NACME between two excited states $|I\rangle$ and $|J\rangle$, since the corresponding formula for these two properties are obtained by taking the first order residue of expressions like Eq. (A16) at $\omega_k = -\omega_I$. To the best of our knowledge, this point has not been noticed in previous studies of transition properties between two excited states. Only by assuming $\omega_{JI'} \neq \omega_I$ for any I', we can identify

$$\begin{aligned}\langle I|\bar{V}^a|J\rangle &= \sum_{I'>0}\left[\frac{(\mathbf{V}^{a\dagger}\mathbf{t}_I')(\mathbf{t}_{I'}^{\dagger},\mathbf{V}_{IJ})}{\omega_{I'}-(-\omega_I+\omega_J)} + \frac{(\mathbf{V}^{a\dagger}\mathbf{t}_{-I'})(\mathbf{t}_{-I'}^{\dagger},\mathbf{V}_{IJ})}{\omega_{I'}+(-\omega_I+\omega_J)}\right]\\ &\quad + \langle 0|[O_I,[V^a,O_J^{\dagger}]]|0\rangle\\ &= \mathbf{V}^{a\dagger}[\mathbf{E}-\omega_{JI}\mathbf{S}]^{-1}\mathbf{V}_{IJ} + \langle 0|[O_I,[V^a,O_J^{\dagger}]]|0\rangle\\ &= \sum_{pq} V_{pq}^a \gamma_{pq}^{IJ},\end{aligned} \tag{A17}$$

which is just the formula used before for the TDM.

Therefore, the standard way for identifying transition properties from the residues of approximate QRF may be plagued by second order poles, which are otherwise not present in exact theory. A frequency-dependent kernel of TD-DFT may bring in some extra terms, so as to cancel the erroneous poles. However, such a frequency-dependent kernel is not yet available. As a practical solution, only those terms that have the same structure as in exact theory will be retained. Specifically, for the transition moment (A14), the terms in Eq. (A16) to be retained are

$$\begin{aligned}&\sum_{I'>0}\frac{(\mathbf{V}^{a\dagger}\mathbf{t}_I')(\mathbf{t}_{I'}^{\dagger},\langle 0|[\boldsymbol{\Lambda},[-O_J^{\dagger},V^b]]|0\rangle)}{\omega_{I'}-(\omega_k+\omega_J)}\\ &\quad + \sum_{I'>0}\frac{(\langle 0|[\boldsymbol{\Lambda}^{\dagger}\mathbf{t}_{-I'},[O_J^{\dagger},V^a]]|0\rangle)(\mathbf{t}_{-I'}^{\dagger},\mathbf{V}^b)}{\omega_{I'}+\omega_k}\\ &\quad = \sum_{I'>0}\frac{\langle 0|V^a|I\rangle\langle 0|[O_I,[V^b,O_J^{\dagger}]]|0\rangle}{\omega_{I'}-(\omega_k+\omega_J)}\\ &\quad + \sum_{I'>0}\frac{\langle 0|V^b|I\rangle\langle 0|[O_I,[V^a,O_J^{\dagger}]]|0\rangle}{\omega_{I'}+\omega_k},\end{aligned} \tag{A18}$$

which suggests immediately the transition moment between two excited states should be

$$\langle I|\bar{V}^a|J\rangle = \langle 0|[O_I,[V^a,O_J^{\dagger}]]|0\rangle. \tag{A19}$$

For TD-HF, this amounts just to neglecting $\gamma_{pq}^{IJ}(\mathrm{I})$ (29), thereby going back to the RPA derived from the EOM formalism.[26] One disadvantage of Eq. (A18) lies in that, for such transition moment, the implicit summation over all the excited states furnished by solving the response equation as in Eq. (A14) cannot be achieved. Finally, we mention that for the special case of $I = J$, Eq. (A17) is always well-behaved unless $\omega_{I'} = 0$.

[1] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).
[2] M. E. Casida, in *Recent Advances in Density Functional Methods*, edited by D. P. Chang (World Scientific, Singapore, 1995), Vol. 1.
[3] Z. Li and W. Liu, J. Chem. Phys. **133**, 064106 (2010).
[4] Z. Li, W. Liu, Y. Zhang, and B. Suo, J. Chem. Phys. **134**, 134101 (2011).
[5] Z. Li and W. Liu, J. Chem. Phys. **135**, 194106 (2011).
[6] Z. Li, B. Suo, Y. Zhang, Y. Xiao, and W. Liu, Mol. Phys. **111**, 3741 (2013).
[7] V. Chernyak and S. Mukamel, J. Chem. Phys. **112**, 3572 (2000).
[8] R. Baer, Chem. Phys. Lett. **364**, 75 (2002).
[9] C. Hu, H. Hirai, and O. Sugino, J. Chem. Phys. **127**, 064103 (2007).
[10] C. Hu, H. Hirai, and O. Sugino, J. Chem. Phys. **128**, 154111 (2008).
[11] C. Hu, O. Sugino, and Y. Tateyama, J. Chem. Phys. **131**, 114101 (2009).
[12] C. Hu, O. Sugino, H. Hirai, and Y. Tateyama, Phys. Rev. A **82**, 062508 (2010).
[13] E. Tapavicza, I. Tavernelli, and U. Röthlisberger, Phys. Rev. Lett. **98**, 023001 (2007).
[14] I. Tavernelli, E. Tapavicza, and U. Röthllisberger, J. Chem. Phys. **130**, 124107 (2009).
[15] I. Tavernelli, B. F. E. Curchod, and U. Röthlisberger, J. Chem. Phys. **131**, 196101 (2009).
[16] I. Tavernelli, B. F. E. Curchod, A. Laktionov, and U. Röthlisberger, J. Chem. Phys. **133**, 194104 (2010).
[17] R. Send and F. Furche, J. Chem. Phys. **132**, 044107 (2010).
[18] S. Fatehi, E. Alguire, Y. Shao, and J. E. Subotnik, J. Chem. Phys. **135**, 234105 (2011).
[19] E. Alguire, Q. Ou, and J. E. Subotnik, "Calculating derivative couplings between time-dependent Hartree–Fock excited states with pseudo-wavefunctions," J. Phys. Chem. B (published online).
[20] Q. Ou, E. Alguire, and J. E. Subotnik, "Derivative couplings between time-dependent density functional theory excited states in the random-phase approximation based on pseudo-wavefunctions: Behavior around conical intersections," J. Phys. Chem. B (published online).
[21] X. Zhang and J. M. Herbert, J. Chem. Phys. **141**, 064104 (2014).
[22] F. Cordova, L. J. Doriol, A. Ipatov, M. E. Casida, C. Filippi, and A. Vela, J. Chem. Phys. **127**, 164111 (2007).
[23] E. Tapavicza, I. Tavernelli, U. Rothlisberger, C. Filippi, and M. E. Casida, J. Chem. Phys. **129**, 124108 (2008).
[24] M. E. Casida, B. Natarajan, and T. Deutsch, *Fundamentals of Time-Dependent Density-Functional Theory*, Lecture Notes in Physics Vol. 837, edited by M. A. L. Marques, N. T. Maitra, F. M. S. Nogueira, E. K. U. Gross, and A. Rubio (Springer Verlag, 2011), p. 279.

[25]N. T. Maitra, J. Chem. Phys. **125**, 014110 (2006).
[26]Z. Li and W. Liu, J. Chem. Phys. **141**, 014110 (2014).
[27]D. J. Rowe, Rev. Mod. Phys. **40**, 153 (1968).
[28]E. Dalgaard, Phys. Rev. A **26**, 42 (1982).
[29]J. Olsen and P. Jørgensen, J. Chem. Phys. **82**, 3235 (1985).
[30]K. Sasagane, F. Aiga, and R. Itoh, J. Chem. Phys. **99**, 3738 (1993).
[31]O. Christiansen, P. Jørgensen, and C. Hättig, Int. J. Quantum Chem. **68**, 1 (1998).
[32]T. Helgaker and P. Jørgensen, Theor. Chim. Acta **75**, 111 (1989).
[33]F. Furche, J. Chem. Phys. **114**, 5982 (2001).
[34]F. Furche and R. Ahlrichs, J. Chem. Phys. **117**, 7433 (2002).
[35]Y. G. Khait, D. Theis, and M. R. Hoffmann, Mol. Phys. **108**, 2703 (2010).
[36]Y. G. Khait, D. Theis, and M. R. Hoffmann, Chem. Phys. **401**, 88 (2012).
[37]Y. Yang, H. van Aggelen, and W. Yang, J. Chem. Phys. **139**, 224105 (2013).
[38]D. Peng, H. van Aggelen, Y. Yang, and W. Yang, J. Chem. Phys. **140**, 18A522 (2014).
[39]J. Liu and W. Liang, J. Chem. Phys. **135**, 184111 (2011).
[40]N. C. Handy and H. F. Schaefer III, J. Chem. Phys. **81**, 5031 (1984).
[41]The abbreviation CIS used here for g_{IJ}^{ξ} refers to the TDA-like approximation to TD-HF resulting from time-dependent response theory, instead of the standard CIS, denoted by CIS(n), resulting from time-independent EOM.[26] The former contains contributions from $\mathbf{t}_{IJ}$ but the latter does not. That is, the present CIS(n) corresponds to the CIS in Ref. 18.
[42]T. Helgaker, P. Jørgensen, and N. C. Handy, Theor. Chim. Acta **76**, 227 (1989).
[43]Z. Li and W. Liu, J. Chem. Phys. **136**, 024107 (2012).
[44]W. Liu, G. Hong, D. Dai, L. Li, and M. Dolg, Theor. Chem. Acc. **96**, 75 (1997).
[45]W. Liu, F. Wang, and L. Li, J. Theor. Comput. Chem. **2**, 257 (2003).
[46]W. Liu, F. Wang, and L. Li, *Recent Advances in Relativistic Molecular Theory*, Recent Advances in Computational Chemistry Vol. 5, edited by K. Hirao and Y. Ishikawa (World Scientific, Singapore, 2004), p. 257.
[47]See http://bdfwiki.com/resp for the description of the RESP module in BDF.
[48]B. Gao, A. J. Thorvaldsen, and K. Ruud, Int. J. Quantum Chem. **111**, 858 (2011).
[49]B. Gao, *Gen1Int Version* 0.3.0, 2014, http://repo.ctcc.no/projects/gen1int.
[50]U. Ekström, L. Visscher, R. Bast, A. J. Thorvaldsen, and K. Ruud, J. Chem. Theory Comput. **6**, 1971 (2010).
[51]U. Ekström, *XCFun library*, 2010, http://www.admol.org/xcfun.
[52]S. J. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).
[53]H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz *et al.*, MOLPRO, version 2010.1, a package of *ab initio* programs, 2010, see http://www.molpro.net.
[54]T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[55]D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **100**, 2975 (1994).
[56]R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
[57]B. H. Lengsfield III, P. Saxe, and D. R. Yarkony, J. Chem. Phys. **81**, 4549 (1984).
[58]K. L. Bak, P. Jørgensen, H. J. A. Jensen, J. Olsen, and T. Helgaker, J. Chem. Phys. **97**, 7573 (1992).
[59]L. Cederbaum, in *Conical Intersections: Electronic Structure, Dynamics and Spectroscopy*, edited by W. Domcke, D. R. Yarkony, and H. Köppel (World Scientific, Singapore, 2004), p. 3.