**University of Texas Rio Grande Valley**

# ScholarWorks @ UTRGV

Chemistry Faculty Publications and Presentations

College of Sciences

12-19-2011

# Analytic derivative couplings between configuration-interaction-singles states with built-in electron-translation factors for translational invariance

Shervin Fatehi
*University of Texas Rio Grande Valley*, shervin.fatehi@utrgv.edu

Joseph Subotnik

Yihan Shao

Ethan Alguire

Follow this and additional works at: https://scholarworks.utrgv.edu/chem_fac

Part of the Chemistry Commons

## Recommended Citation

# Analytic derivative couplings between configuration-interaction-singles states with built-in electron-translation factors for translational invariance

Shervin Fatehi,[1] Ethan Alguire,[1] Yihan Shao,[2] and Joseph E. Subotnik[1,a)]

[1]*Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, Pennsylvania 19104, USA*
[2]*Q-Chem, Inc., The Design Center, Suite 690, 5001 Baum Boulevard, Pittsburgh, Pennsylvania 15213, USA*

We present a method for analytically calculating the derivative couplings between a pair of configuration-interaction-singles (CIS) excited states obtained in an atom-centered basis. Our theory is exact and has been derived using two completely independent approaches: one inspired by the Hellmann-Feynman theorem and the other following from direct differentiation. (The former is new, while the latter is in the spirit of existing approaches in the literature.) Our expression for the derivative couplings incorporates all Pulay effects associated with the use of an atom-centered basis, and the computational cost is minimal, roughly comparable to that of a single CIS energy gradient. We have validated our method against CIS finite-difference results and have applied it to the lowest lying excited states of naphthalene; we find that naphthalene derivative couplings include Pulay contributions sufficient to have a qualitative effect. Going beyond standard problems in analytic gradient theory, we have also constructed a correction, based on perturbative electron-translation factors, for including electronic momentum and eliminating spurious components of the derivative couplings that break translational symmetry. This correction is general and can be applied to any level of electronic structure theory. © *2011 American Institute of Physics*. [doi:10.1063/1.3665031]

## I. INTRODUCTION

### A. Derivative couplings

The Born-Oppenheimer approximation, ubiquitous in chemistry, is based on an adiabatic separation of electronic and nuclear motions; electronic states are treated as parametrically dependent on the nuclear configuration $\mathbf{R}$ and are assumed to follow the nuclei instantaneously. As a result, nuclear motion cannot induce an electronic transition within this approximation, and molecular dynamics unfolds on only one of the set of potential-energy surfaces $\{E_I(\mathbf{R})\}$ associated with the electronic states $|\Psi_I\rangle \equiv |\Psi_I(\mathbf{R})\rangle$. To treat nonadiabatic phenomena involving dynamical transitions between states, one must go beyond Born-Oppenheimer. To that end, it is straightforward to show that the leading-order nonadiabatic corrections to the Born-Oppenheimer approximation are the derivative couplings given by $\langle\Psi_I|\nabla_Q|\Psi_J\rangle$ for each nucleus $Q$.[1,2] These matrix elements couple different adiabatic potential surfaces, and they appear in the corrected Hamiltonian and electronic equations of motion as dot products with the nuclear velocities,

$$\sum_Q \langle\Psi_I|\nabla_Q|\Psi_J\rangle \cdot \frac{\mathbf{P}_Q}{m_Q}. \qquad (1)$$

Thus, derivative couplings switch on electronic transitions while conserving energy through compensatory changes in nuclear momentum.

Many schemes have been developed for the propagation of otherwise classical nuclei on multiple potential-

energy surfaces—the Meyer-Miller-Stock-Thoss approach,[3,4] surface hopping,[2,5] and *ab initio* multiple spawning,[6] to name a few—and all of these algorithms require as input the potential-energy surfaces and the couplings between them. Derivative couplings are also needed in some methods for finding conical intersections[7] and to compute rigorous diabatic states using the formalism of Baer.[8–10]

Historically, derivative couplings were calculated by Lengsfield, Saxe, Yarkony, Lischka, Shepard, and co-workers using multi-reference configuration-interaction (MR-CI) wavefunction theory.[1,7,11–13] MR-CI has dominated this field of study because, when used properly, the method correctly captures static correlation and generates ground and excited states in a balanced framework, capturing the topology of their curve crossings where simpler models of electronic structure fail.[14,15] Moreover, recent developments in MR-CI analytic gradient theory have led to quite general approaches for calculating nonadiabatic derivate couplings.[13,16–22] The downside of using MR-CI is that, although the relative cost of derivative couplings is not high, the *total* computational cost can be large. After all, MR-CI effectively requires a small active space if it is to be applicable to large molecules, and this requirement is demanding. As a result, state-of-the-art nonadiabatic dynamics simulations using MR-CI have been limited to roughly a dozen atoms.[23–25]

In the last decade, there has been considerable interest in computing derivative couplings more affordably, typically within time-dependent Hartree-Fock theory and time-dependent density-functional theory (TD-DFT). (Some work has also been done for equation-of-motion coupled-cluster theory.[26,27]) In this context, expressions for the derivative

couplings between the ground and excited electronic states were first obtained analytically by Chernyak, Mukamel, and co-workers.[28,29] These authors applied the logic of the Hellmann-Feynman theorem to express the derivative coupling as a matrix element between different electronic states of the force exerted on the nucleus. In a clever argument, they related that same matrix element to electronic response properties (specifically, the polarizability). The resulting closed expression for derivative couplings can be evaluated as a sum over states and can therefore be computed approximately in practice. An important component is missing from the Chernyak-Mukamel theory, however, at least in relation to common practice in quantum chemistry: The expressions do not include any Pulay terms—corrections arising from the use of an atom-centered basis.[30]

Since the publication of Ref. 28, several competing methods have been developed along similar lines,[31–40] only some of which have employed an atom-centered basis. The Gaussian-basis implementation of the method of Hu *et al.* includes Pulay terms associated with the one-electron integrals.[34] Send and Furche draw a rigorous connection between the derivative couplings of the physical system and its fictitious Kohn-Sham counterpart, accounting for all Pulay effects in the process.[40]

The methods described so far are primarily concerned with derivative couplings between ground and excited states. Given that single-reference methods do not provide a balanced description of the states being coupled,[14,15] a better focus might be derivative couplings between excited states, which *are* balanced in the context of TD-DFT or the configuration-interaction-singles (CIS) wavefunction theory used here. The Chernyak-Mukamel theory can be extended to derivative couplings between excited states,[41] but the Send-Furche derivation applies specifically to the case of ground-to-excited-state couplings. The auxiliary-wavefunction approach recently developed by Röthlisberger and co-workers for the purpose of calculating excited-state couplings bears strong similarities to the theory of Chernyak and Mukamel, but it is not equivalent.[39]

In light of the explosion of interest in derivative couplings and nonadiabatic phenomena, we now present a self-contained discussion of the problem of computing derivative couplings between excited states. Rather than focusing on TD-DFT, we work within the simpler CIS.[42] We show that our theory for CIS derivative couplings can be derived from two completely independent approaches, one based on a Hellmann-Feynman expression and the other based on direct differentiation. (The former derivation is novel, while the latter is a special case of that developed for MR-CI.[1,11–13,22]) Our method is formally exact, accounts rigorously for all Pulay effects, and entails computational cost comparable to that of a single CIS gradient. We perform proof-of-principle calculations of derivative couplings for a prototypical multi-electron diatom, lithium hydride, and find that they are in excellent agreement with finite-difference results. We also present an application to couplings between the lowest-lying excited states of naphthalene, for which Pulay effects are shown to be of qualitative significance.

Now, although the couplings computed using our theory are exact in the sense that no approximations have been made in obtaining them (and they match finite-difference calculations), these couplings exhibit unphysical translational-variance properties that stem from neglect of electronic motion. After all, Born-Oppenheimer eigenstates are real-valued in real space; if the states are to move with the nuclei, they should of course be complex. Thus, this problem has nothing to do with CIS, and it is common to MR-CI and TD-DFT derivative couplings also—it even appears in the analytical treatment of the hydrogen atom.[11,40,43,44] To correct this behavior, we draw on the scattering literature and apply electron-translation factors (ETFs), complex multiplicative factors that, when applied to the wavefunction in the time-dependent Schrödinger equation, account for electronic motion in tandem with the nuclei.[45–52] We show through a series of well-defined approximations that the effect of these ETFs is to form a universal operator that generates perturbative corrections to the derivative couplings. Evaluating the matrix elements of this correction operator between adiabatic states leads to a scheme for calculating derivative couplings with "built-in" ETFs. This ETF correction can be applied to CIS or MR-CI calculations, or to those involving any other type of adiabatic electronic state.

To summarize, we present in Eqs. (81)–(83) and (121) a method for computing derivative couplings between CIS states with built-in translational symmetry, making them suitable for use in nonadiabatic dynamics calculations.

## B. Notation

Before proceeding with the derivation, we summarize our notational conventions: Lowercase italics $\{i, j, k, \ell, m\}$ index occupied molecular spin-orbitals; $\{a, b, c, d\}$, virtual molecular spin-orbitals; and $\{p, q, r, s, t, u, w\}$, arbitrary (occupied *or* virtual) molecular spin-orbitals. Greek letters $\{\mu, \nu, \lambda, \sigma\}$ and $\{\alpha, \beta, \gamma, \delta\}$ are used interchangeably to indicate atomic orbitals. Two-electron integrals are given in antisymmetrized physicists' notation, $\Pi_{pqsr} = \langle pq||sr \rangle$.

Uppercase italic $E$ indicates the Hartree-Fock energy, unless otherwise subscripted; $I$ and $J$ label electronic states, usually CIS states; $N$ denotes the number of nuclei; and $V$ and $O$ are the numbers of virtual and occupied orbitals, respectively.

Lowercase $x$ is used as shorthand to indicate *any* of the Cartesian coordinates $\{x, y, z\}$ associated with a nucleus. $x_Q$ is a Cartesian coordinate belonging to nucleus $Q$. The total derivative with respect to $x$ is written as a bracketed superscript $^{[x]}$. We suppress the parametric $x$-dependence of the Hamiltonian, wavefunctions, etc. $e$ will be used as a subscript for the electronic mass and gradient (and, occasionally, as an electronic summation index). Vectors, matrices, and tensors (as opposed to their individual elements) are represented in bold type.

## II. ANALYTIC HELLMANN-FEYNMAN THEORY FOR CIS DERIVATIVE COUPLINGS

### A. Preliminary considerations

Consider two states $|\Psi_I\rangle$ and $|\Psi_J\rangle$ that are eigenstates of the electronic Hamiltonian $H$,

$$H = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{4} \sum_{pqrs} \Pi_{pqsr} a_p^\dagger a_q^\dagger a_r a_s \tag{2}$$

$$= \sum_{pq} \langle p|h|q\rangle a_p^\dagger a_q + \frac{1}{4} \sum_{pqrs} \langle pq||sr\rangle a_p^\dagger a_q^\dagger a_r a_s. \tag{3}$$

From the time-independent Schrödinger equation and the orthonormality condition,

$$H |\Psi_J\rangle = E_J |\Psi_J\rangle \tag{4a}$$

$$\langle \Psi_I|\Psi_J\rangle = \delta_{IJ}, \tag{4b}$$

we can derive the familiar expressions

$$\langle \Psi_I^{[x]}|\Psi_J\rangle = -\langle \Psi_I|\Psi_J^{[x]}\rangle \tag{5a}$$

$$\langle \Psi_I|\Psi_J^{[x]}\rangle = \frac{\langle \Psi_I|H^{[x]}|\Psi_J\rangle}{E_J - E_I}. \tag{5b}$$

Equation (5b) is the usual Hellmann-Feynman expression for the derivative coupling. (Note that this formula is inapplicable when $I = J$; analogous work yields the Hellmann-Feynman prescription for the energy gradient.) We would like to find a similar expression for derivative couplings between the CIS states,[42]

$$|\Psi_I\rangle = \sum_{ia} t_i^{Ia} |\Phi_i^a\rangle \tag{6}$$

$$= \sum_{ia} t_i^{Ia} a_a^\dagger a_i |\Phi_{\mathrm{HF}}\rangle, \tag{7}$$

where $\{|\Phi_i^a\rangle\}$ denotes the set of Slater determinants with occupied orbital $i$ promoted to virtual orbital $a$, the singles. The $V \times O$ amplitudes $\{t_i^{Ia}\}$ are determined by diagonalizing the Hamiltonian in the singles space (which defines the tensor **A**),

$$\langle \Psi_I|H|\Psi_J\rangle = \sum_{k\ell cd} \langle \Psi_I|\Phi_\ell^d\rangle\langle \Phi_\ell^d|H|\Phi_k^c\rangle\langle \Phi_k^c|\Psi_J\rangle \tag{8}$$

$$\equiv \sum_{k\ell cd} t_\ell^{Id} A_{\ell dkc} t_k^{Jc} \tag{9}$$

$$= \sum_{\ell d} t_\ell^{Id} E_J t_\ell^{Jd} \tag{10}$$

$$= E_J \delta_{IJ}. \tag{11}$$

Because **A** is diagonalized rather than the full Hamiltonian $H$, Eq. (4a) is not true for CIS states, and Eq. (5b) is invalid. We want to find a modified Hamiltonian $\widetilde{H}$ for which

Eq. (5b) holds, and in a sense we already have: If we define the projector

$$\mathcal{P} \equiv \sum_{kc} |\Phi_k^c\rangle\langle \Phi_k^c|, \tag{12}$$

which satisfies the relationship $\mathcal{P}^2 = \mathcal{P}$ and projects singles and CIS states into themselves, we see that Eqs. (8)–(11) imply

$$\mathcal{P} H \mathcal{P} |\Psi_J\rangle = E_J |\Psi_J\rangle. \tag{13}$$

We therefore conclude that the Hamiltonian compatible with the Hellmann-Feynman theorem for CIS states is

$$\widetilde{H} \equiv \mathcal{P} H \mathcal{P}. \tag{14}$$

### B. Evaluating the derivative coupling

We now replace $H$ with $\widetilde{H}$ in Eq. (5b),

$$\langle \Psi_I|\Psi_J^{[x]}\rangle = \frac{\langle \Psi_I|\widetilde{H}^{[x]}|\Psi_J\rangle}{E_J - E_I} \tag{15}$$

$$= \frac{\langle \Psi_I|\mathcal{P}^{[x]}H\mathcal{P}|\Psi_J\rangle}{E_J - E_I} + \frac{\langle \Psi_I|\mathcal{P}H^{[x]}\mathcal{P}|\Psi_J\rangle}{E_J - E_I}$$
$$+ \frac{\langle \Psi_I|\mathcal{P}H\mathcal{P}^{[x]}|\Psi_J\rangle}{E_J - E_I} \tag{16}$$

$$= \frac{\langle \Psi_I|H^{[x]}|\Psi_J\rangle}{E_J - E_I}$$
$$+ \sum_{kc} \frac{\langle \Psi_I|\Phi_k^{c[x]}\rangle\langle \Phi_k^c|H\mathcal{P}|\Psi_J\rangle}{E_J - E_I}$$
$$+ \sum_{kc} \frac{\langle \Psi_I|\Phi_k^c\rangle\langle \Phi_k^{c[x]}|H\mathcal{P}|\Psi_J\rangle}{E_J - E_I}$$
$$+ \sum_{kc} \frac{\langle \Psi_I|\mathcal{P}H|\Phi_k^{c[x]}\rangle\langle \Phi_k^c|\Psi_J\rangle}{E_J - E_I}$$
$$+ \sum_{kc} \frac{\langle \Psi_I|\mathcal{P}H|\Phi_k^c\rangle\langle \Phi_k^{c[x]}|\Psi_J\rangle}{E_J - E_I} \tag{17}$$

$$= \frac{\langle \Psi_I|H^{[x]}|\Psi_J\rangle}{E_J - E_I}$$
$$+ \sum_{kc} \frac{\langle \Psi_I|\Phi_k^{c[x]}\rangle}{E_J - E_I} E_J t_k^{Jc}$$
$$+ \sum_{kc} \frac{\langle \Phi_k^{c[x]}|H\mathcal{P}|\Psi_J\rangle}{E_J - E_I} t_k^{Ic}$$
$$+ \sum_{kc} \frac{\langle \Psi_I|\mathcal{P}H|\Phi_k^{c[x]}\rangle}{E_J - E_I} t_k^{Jc}$$
$$+ \sum_{kc} \frac{\langle \Phi_k^{c[x]}|\Psi_J\rangle}{E_J - E_I} E_I t_k^{Ic}. \tag{18}$$

To proceed further, we need the configuration-interaction resolution of the identity,

$$\mathbf{1} = |\Phi_{\text{HF}}\rangle\langle\Phi_{\text{HF}}| + \mathcal{P} + \frac{1}{4}\sum_{ijab}|\Phi_{ij}^{ab}\rangle\langle\Phi_{ij}^{ab}|$$

$$+ \frac{1}{36}\sum_{ij\ell abd}|\Phi_{ij\ell}^{abd}\rangle\langle\Phi_{ij\ell}^{abd}| + \dots \tag{19}$$

Noting that the Hamiltonian contains only one- and two-electron operators and does not couple CIS states to the Hartree-Fock determinant (Brillouin's theorem),[53] we retain singles, doubles, and triples in Eq. (19). Inserting the result between $H$ and the singles derivatives, we obtain

$$\langle\Psi_I|\Psi_J^{[x]}\rangle = \frac{\langle\Psi_I|H^{[x]}|\Psi_J\rangle}{E_J - E_I}$$

$$+ \sum_{kc}\frac{\langle\Psi_I|\Phi_k^{c[x]}\rangle}{E_J - E_I}E_J t_k^{Jc} + \sum_{kc}\frac{\langle\Phi_k^{c[x]}|\Psi_J\rangle}{E_J - E_I}E_J t_k^{Ic}$$

$$+ \sum_{kc}\frac{\langle\Psi_I|\Phi_k^{c[x]}\rangle}{E_J - E_I}E_I t_k^{Jc} + \sum_{kc}\frac{\langle\Phi_k^{c[x]}|\Psi_J\rangle}{E_J - E_I}E_I t_k^{Ic}$$

$$+ \frac{1}{4}\sum_{ijkabc}\frac{\langle\Phi_k^{c[x]}|\Phi_{ij}^{ab}\rangle\langle\Phi_{ij}^{ab}|H|\Psi_J\rangle}{E_J - E_I}t_k^{Ic}$$

$$+ \frac{1}{4}\sum_{ijkabc}\frac{\langle\Psi_I|H|\Phi_{ij}^{ab}\rangle\langle\Phi_{ij}^{ab}|\Phi_k^{c[x]}\rangle}{E_J - E_I}t_k^{Jc}$$

$$+ \frac{1}{36}\sum_{ijk\ell abcd}\frac{\langle\Phi_k^{c[x]}|\Phi_{ij\ell}^{abd}\rangle\langle\Phi_{ij\ell}^{abd}|H|\Psi_J\rangle}{E_J - E_I}t_k^{Ic}$$

$$+ \frac{1}{36}\sum_{ijk\ell abcd}\frac{\langle\Psi_I|H|\Phi_{ij\ell}^{abd}\rangle\langle\Phi_{ij\ell}^{abd}|\Phi_k^{c[x]}\rangle}{E_J - E_I}t_k^{Jc}. \tag{20}$$

All terms involving $\langle\Psi_I|\Phi_k^{c[x]}\rangle$ and $\langle\Phi_k^{c[x]}|\Psi_J\rangle$ are seen to cancel on expansion of $|\Psi_I\rangle$ and $|\Psi_J\rangle$, and we can use the hermiticity of $H$ to rewrite our expression as

$$\langle\Psi_I|\Psi_J^{[x]}\rangle = \frac{\langle\Psi_I|H^{[x]}|\Psi_J\rangle}{E_J - E_I} + \frac{1}{4}\sum_{ijkabc}\frac{\langle\Phi_{ij}^{ab}|\Phi_k^{c[x]}\rangle}{E_J - E_I}$$

$$\times \left[\langle\Phi_{ij}^{ab}|H|\Psi_I\rangle t_k^{Jc} + \langle\Phi_{ij}^{ab}|H|\Psi_J\rangle t_k^{Ic}\right]$$

$$+ \frac{1}{36}\sum_{ijk\ell abcd}\frac{\langle\Phi_{ij\ell}^{abd}|\Phi_k^{c[x]}\rangle}{E_J - E_I}$$

$$\times \left[\langle\Phi_{ij\ell}^{abd}|H|\Psi_I\rangle t_k^{Jc} + \langle\Phi_{ij\ell}^{abd}|H|\Psi_J\rangle t_k^{Ic}\right]. \tag{21}$$

### C. Derivative couplings between singles and higher excitations

We must now determine $\langle\Phi_{ij}^{ab}|\Phi_k^{c[x]}\rangle$ and $\langle\Phi_{ij\ell}^{abd}|\Phi_k^{c[x]}\rangle$. (For the sake of brevity, we discuss the doubles term in detail and treat triples by analogy.) We begin by rewriting the singles-doubles derivative coupling in terms of the Hartree-Fock determinant,

$$\langle\Phi_{ij}^{ab}|\Phi_k^{c[x]}\rangle = \langle\Phi_{\text{HF}}|a_i^\dagger a_j^\dagger a_b a_a (a_c^\dagger a_k |\Phi_{\text{HF}}\rangle)^{[x]} \tag{22}$$

$$= \langle\Phi_{\text{HF}}|a_i^\dagger a_j^\dagger a_b a_a a_c^{\dagger[x]} a_k|\Phi_{\text{HF}}\rangle$$

$$+ \langle\Phi_{\text{HF}}|a_i^\dagger a_j^\dagger a_b a_a a_c^\dagger a_k^{[x]}|\Phi_{\text{HF}}\rangle$$

$$+ \langle\Phi_{\text{HF}}|a_i^\dagger a_j^\dagger a_b a_a a_c^\dagger a_k|\Phi_{\text{HF}}^{[x]}\rangle. \tag{23}$$

It is certainly possible to evaluate these couplings by taking nuclear derivatives of the singles determinants explicitly. To evaluate the first two terms of Eq. (23), we instead derive formal expressions for nuclear derivatives of the annihilation and creation operators. These expressions, once obtained, may be used to develop other analytic gradient and derivative coupling theories for states referenced to Hartree-Fock. Further, they allow us to work in a fully second-quantized formalism, which encodes finicky issues of determinantal structure in straightforward operator algebra.[53,54]

The annihilation operator couples determinants differing only by the presence or absence of a single electron in a specific spin orbital; abbreviating the list of other orbitals as $L$,

$$\langle L|a_s|mL\rangle = \delta_{sm}. \tag{24}$$

We take the nuclear derivative of this standard formula to obtain

$$0 = \langle L^{[x]}|a_s|mL\rangle + \langle L|a_s^{[x]}|mL\rangle + \langle L|a_s|m^{[x]}L\rangle$$

$$+ \langle L|a_s|mL^{[x]}\rangle \tag{25}$$

$$= \delta_{sm}(\langle L^{[x]}|L\rangle + \langle L|L^{[x]}\rangle) + \langle sL|m^{[x]}L\rangle$$

$$+ \langle L|a_s^{[x]}|mL\rangle. \tag{26}$$

Noting that determinants are orthonormal and integrating over $L$ where possible, this expression reduces to

$$\langle L|a_s^{[x]}|mL\rangle = -\langle s|m^{[x]}\rangle \equiv -O_{sm}^{\text{R}[x]}, \tag{27}$$

where we have defined the "right" molecular-spin-orbital overlap derivative $\mathbf{O}^{\text{R}[x]}$; note that $O_{sm}^{\text{R}[x]} = -O_{ms}^{\text{R}[x]}$.

Equation (27) is equivalent to

$$a_s^{[x]} = -\sum_w O_{sw}^{\text{R}[x]} a_w, \tag{28}$$

where the sum runs over *all* spin orbitals; an exactly analogous expression holds for the creation operator $a_s^\dagger$. Equation (23) therefore becomes

$$\langle\Phi_{ij}^{ab}|\Phi_k^{c[x]}\rangle = -\sum_w\langle\Phi_{ij}^{ab}|\Phi_k^w\rangle O_{cw}^{\text{R}[x]} - \sum_w\langle\Phi_{ij}^{ab}|\Phi_w^c\rangle O_{kw}^{\text{R}[x]}$$

$$+ \langle\Phi_{\text{HF}}|a_i^\dagger a_j^\dagger a_b a_a a_c^\dagger a_k|\Phi_{\text{HF}}^{[x]}\rangle \tag{29}$$

$$= -\sum_d\langle\Phi_{ij}^{ab}|\Phi_k^d\rangle O_{cd}^{\text{R}[x]} - \sum_\ell\langle\Phi_{ij}^{ab}|\Phi_\ell^c\rangle O_{k\ell}^{\text{R}[x]}$$

$$+ \langle\Phi_{\text{HF}}|a_i^\dagger a_j^\dagger a_b a_a a_c^\dagger a_k|\Phi_{\text{HF}}^{[x]}\rangle, \tag{30}$$

where the second equality follows from Eq. (7) and the Pauli exclusion principle (no spin orbital may be doubly-occupied). Because they involve brackets of doubles and singles, which are orthogonal, the terms involving $\mathbf{O}^{R[x]}$ vanish. After rearranging creation and annihilation operators to satisfy normal ordering, we are therefore left with

$$\langle \Phi_{ij}^{ab} | \Phi_k^{c[x]} \rangle = \delta_{bc}\delta_{jk}\langle \Phi_i^a | \Phi_{HF}^{[x]} \rangle + \delta_{ac}\delta_{ik}\langle \Phi_j^b | \Phi_{HF}^{[x]} \rangle$$
$$- \delta_{bc}\delta_{ik}\langle \Phi_j^a | \Phi_{HF}^{[x]} \rangle - \delta_{ac}\delta_{jk}\langle \Phi_i^b | \Phi_{HF}^{[x]} \rangle. \quad (31)$$

The Hartree-Fock state contains occupied orbitals that we number consecutively from 1 to $O$. We can find its derivative in terms of the electronic vacuum state $|\rangle$, which contains no MOs and thus no dependence on the nuclear coordinate,

$$|\Phi_{HF}^{[x]}\rangle = \left(a_1^\dagger \cdots a_\ell^\dagger \cdots a_O^\dagger |\rangle\right)^{[x]} \quad (32)$$

$$= -\sum_{\ell w} a_1^\dagger \cdots a_w^\dagger \cdots a_O^\dagger |\rangle \, O_{\ell w}^{R[x]} \quad (33)$$

$$= \sum_{\ell d} |\Phi_\ell^d\rangle \, O_{d\ell}^{R[x]}. \quad (34)$$

Because the singles are orthonormal, we now obtain

$$\langle \Phi_{ij}^{ab} | \Phi_k^{c[x]} \rangle = \delta_{bc}\delta_{jk}O_{ai}^{R[x]} + \delta_{ac}\delta_{ik}O_{bj}^{R[x]}$$
$$- \delta_{bc}\delta_{ik}O_{aj}^{R[x]} - \delta_{ac}\delta_{jk}O_{bi}^{R[x]}. \quad (35)$$

By analogy to Eq. (31), $\langle \Phi_{ij\ell}^{abd} | \Phi_k^{c[x]} \rangle$ involves terms of the form $\langle \Phi_{ij}^{ab} | \Phi_{HF}^{[x]} \rangle$. Referring to Eq. (34) and noting that doubles and singles are orthogonal, we find that the triples terms vanish in Eq. (21). Inserting Eq. (35) yields

$$\langle \Psi_I | \Psi_J^{[x]} \rangle$$
$$= \frac{\langle \Psi_I | H^{[x]} | \Psi_J \rangle}{E_J - E_I}$$
$$+ \frac{1}{4}\sum_{ikac} \frac{O_{ai}^{R[x]}}{E_J - E_I}\left[\langle \Phi_{ik}^{ac}|H|\Psi_I\rangle t_k^{Jc} + \langle \Phi_{ik}^{ac}|H|\Psi_J\rangle t_k^{Ic}\right]$$
$$+ \frac{1}{4}\sum_{jkbc} \frac{O_{bj}^{R[x]}}{E_J - E_I}\left[\langle \Phi_{kj}^{cb}|H|\Psi_I\rangle t_k^{Jc} + \langle \Phi_{kj}^{cb}|H|\Psi_J\rangle t_k^{Ic}\right]$$
$$- \frac{1}{4}\sum_{jkac} \frac{O_{aj}^{R[x]}}{E_J - E_I}\left[\langle \Phi_{kj}^{ac}|H|\Psi_I\rangle t_k^{Jc} + \langle \Phi_{kj}^{ac}|H|\Psi_J\rangle t_k^{Ic}\right]$$
$$- \frac{1}{4}\sum_{ikbc} \frac{O_{bi}^{R[x]}}{E_J - E_I}\left[\langle \Phi_{ik}^{cb}|H|\Psi_I\rangle t_k^{Jc} + \langle \Phi_{ik}^{cb}|H|\Psi_J\rangle t_k^{Ic}\right]. \quad (36)$$

Permuting $c$ and $k$ in the remaining doubles brackets and noting that all other indices are dummies, we find that the latter four terms collapse,

$$\langle \Psi_I | \Psi_J^{[x]} \rangle = \frac{\langle \Psi_I | H^{[x]} | \Psi_J \rangle}{E_J - E_I} - \sum_{ikbc} \frac{O_{bi}^{R[x]}}{E_J - E_I}\left[\langle \Phi_{ik}^{cb}|H|\Psi_I\rangle t_k^{Jc}\right.$$
$$\left. + \langle \Phi_{ik}^{cb}|H|\Psi_J\rangle t_k^{Ic}\right] \quad (37)$$

$$= \frac{\langle \Psi_I | H^{[x]} | \Psi_J \rangle}{E_J - E_I} - \sum_{ik\ell bcd} \frac{\langle \Phi_{ik}^{cb}|H|\Phi_\ell^d\rangle}{E_J - E_I}\left[t_\ell^{Id}t_k^{Jc}\right.$$
$$\left. + t_k^{Ic}t_\ell^{Jd}\right]O_{bi}^{R[x]}. \quad (38)$$

## D. Singles-doubles coupling and the Hamiltonian derivative

The next term we must evaluate is the singles-doubles (Hamiltonian) coupling $\langle \Phi_{ik}^{cb}|H|\Phi_\ell^d\rangle$. Treating the one- and two-electron-operator contributions separately gives

$$\sum_{pq} h_{pq}\langle \Phi_{ik}^{cb}|a_p^\dagger a_q|\Phi_\ell^d\rangle$$
$$= \delta_{bd}\delta_{k\ell}h_{ci} - \delta_{bd}\delta_{i\ell}h_{ck} + \delta_{cd}\delta_{i\ell}h_{bk} - \delta_{cd}\delta_{k\ell}h_{bi} \quad (39a)$$

$$\frac{1}{4}\sum_{pqrs}\Pi_{pqsr}\langle\Phi_{ik}^{cb}|a_p^\dagger a_q^\dagger a_r a_s|\Phi_\ell^d\rangle$$
$$= \delta_{k\ell}\Pi_{bcdi} + \delta_{bd}\Pi_{\ell cik} - \delta_{cd}\Pi_{\ell bik} - \delta_{i\ell}\Pi_{bcdk}$$
$$+ \delta_{cd}\delta_{i\ell}\sum_m \Pi_{bmkm} - \delta_{cd}\delta_{k\ell}\sum_m \Pi_{bmim}$$
$$- \delta_{bd}\delta_{i\ell}\sum_m \Pi_{cmkm} + \delta_{bd}\delta_{k\ell}\sum_m \Pi_{cmim}. \quad (39b)$$

Plugging the singles-doubles coupling into Eq. (38), we find that the one-electron terms combine with the latter four two-electron terms to form off-diagonal elements of the Fock matrix ($F_{pq} = \varepsilon_p \delta_{pq}$, with $\varepsilon_p$ the orbital energy), leaving behind residual contributions from the two-electron integrals that we call **Y**,

$$-\sum_{ik\ell bcd} \frac{\langle \Phi_{ik}^{cb}|H|\Phi_\ell^d\rangle}{E_J - E_I}\left[t_\ell^{Id}t_k^{Jc} + t_k^{Ic}t_\ell^{Jd}\right]O_{bi}^{R[x]} \quad (40)$$

$$= -\sum_{ik\ell bcd} \frac{t_\ell^{Id}t_k^{Jc} + t_k^{Ic}t_\ell^{Jd}}{E_J - E_I}$$
$$\times [\delta_{k\ell}\Pi_{bcdi} + \delta_{bd}\Pi_{\ell cik} - \delta_{cd}\Pi_{\ell bik}$$
$$- \delta_{i\ell}\Pi_{bcdk}]O_{bi}^{R[x]} \quad (41)$$

$$= \frac{1}{E_J - E_I}\sum_{ikbc}\left[\begin{array}{l}\sum_d \left(t_k^{Id}t_k^{Jc} + t_k^{Ic}t_k^{Jd}\right)\Pi_{cbdi}\\[4pt] + \sum_\ell \left(t_\ell^{Ib}t_k^{Jc} + t_k^{Ic}t_\ell^{Jb}\right)\Pi_{\ell cki}\\[4pt] - \sum_\ell \left(t_\ell^{Ic}t_k^{Jc} + t_k^{Ic}t_\ell^{Jc}\right)\Pi_{\ell bki}\\[4pt] + \sum_d \left(t_i^{Id}t_k^{Jc} + t_k^{Ic}t_i^{Jd}\right)\Pi_{bcdk}\end{array}\right]O_{bi}^{R[x]}$$
$$(42)$$

$$\equiv \frac{1}{E_J - E_I}\sum_{ib}Y_{bi}O_{bi}^{R[x]}. \quad (43)$$

We now have a very simple expression for the CIS derivative coupling that shows it to be the combination of the usual

Hellmann-Feynman contribution and a correction term associated with nuclear derivatives of the molecular orbitals—precisely the structure found in the Chernyak-Mukamel formalism and also in MR-CI,[7,11–13,28,41]

$$\langle \Psi_I | \Psi_J^{[x]} \rangle = \frac{\langle \Psi_I | H^{[x]} | \Psi_J \rangle}{E_J - E_I} + \frac{1}{E_J - E_I} \sum_{ib} Y_{bi} O_{bi}^{\mathrm{R}[x]}. \quad (44)$$

In the standard literature,[1,7,11–13] MR-CI couplings consist of two contributions: a "CI" term associated with the deriva-

tive of the configuration-interaction coefficients (in this case, the singles amplitudes) and a "CSF" term containing derivatives of the configuration-state functions (the singles determinants). The Hellmann-Feynman piece of Eq. (44) is roughly the analogue of the CI term, and the correction is equivalent to the CSF term. That said, the analogy between our work and MR-CI theory is clearer in the separate derivation provided in the Appendix.

To complete the theory, we must evaluate the Hamiltonian derivative, which contains eight terms:

$$H^{[x]} = \sum_{pq} \left[ h_{pq}^{[x]} a_p^\dagger a_q + h_{pq} a_p^{\dagger[x]} a_q + h_{pq} a_p^\dagger a_q^{[x]} \right] + \frac{1}{4} \sum_{pqrs} \begin{bmatrix} \Pi_{pqsr}^{[x]} a_p^\dagger a_q^\dagger a_r a_s \\ + \Pi_{pqsr} a_p^{\dagger[x]} a_q^\dagger a_r a_s \\ + \Pi_{pqsr} a_p^\dagger a_q^{\dagger[x]} a_r a_s \\ + \Pi_{pqsr} a_p^\dagger a_q^\dagger a_r^{[x]} a_s \\ + \Pi_{pqsr} a_p^\dagger a_q^\dagger a_r a_s^{[x]} \end{bmatrix}. \quad (45)$$

Applying Eq. (28) throughout this expression yields

$$\langle \Psi_I | H^{[x]} | \Psi_J \rangle = \sum_{iab} t_i^{Ia} t_i^{Jb} \left[ h_{ab}^{[x]} - \sum_p \left( h_{pb} O_{pa}^{\mathrm{R}[x]} + h_{ap} O_{pb}^{\mathrm{R}[x]} \right) \right]$$

$$- \sum_{ija} t_i^{Ia} t_j^{Ja} \left[ h_{ij}^{[x]} - \sum_p \left( h_{pj} O_{pi}^{\mathrm{R}[x]} + h_{ip} O_{pj}^{\mathrm{R}[x]} \right) \right]$$

$$+ \sum_{ijab} t_i^{Ia} t_j^{Jb} \left[ \Pi_{ajib}^{[x]} - \sum_p \begin{pmatrix} \Pi_{pjib} O_{pa}^{\mathrm{R}[x]} + \Pi_{apib} O_{pj}^{\mathrm{R}[x]} \\ + \Pi_{ajpb} O_{pi}^{\mathrm{R}[x]} + \Pi_{ajip} O_{pb}^{\mathrm{R}[x]} \end{pmatrix} \right]$$

$$+ \sum_{imab} t_i^{Ia} t_i^{Jb} \left[ \Pi_{ambm}^{[x]} - \sum_p \begin{pmatrix} \Pi_{pmbm} O_{pa}^{\mathrm{R}[x]} + \Pi_{apbm} O_{pm}^{\mathrm{R}[x]} \\ + \Pi_{ampm} O_{pb}^{\mathrm{R}[x]} + \Pi_{ambp} O_{pm}^{\mathrm{R}[x]} \end{pmatrix} \right]$$

$$- \sum_{ijma} t_i^{Ia} t_j^{Ja} \left[ \Pi_{imjm}^{[x]} - \sum_p \begin{pmatrix} \Pi_{pmjm} O_{pi}^{\mathrm{R}[x]} + \Pi_{ipjm} O_{pm}^{\mathrm{R}[x]} \\ + \Pi_{impm} O_{pj}^{\mathrm{R}[x]} + \Pi_{imjp} O_{pm}^{\mathrm{R}[x]} \end{pmatrix} \right]. \quad (46)$$

## E. Conversion from molecular to atomic orbitals

While our work so far has been couched in terms of molecular orbitals (MOs), quantum chemistry codes are written to exploit the efficiency of integral evaluation in the atomic-orbital (AO) representation. The molecular orbitals are linear combinations of atomic orbitals ($|p\rangle = \sum_\mu C_{\mu p} |\mu\rangle$), such that the AO representation of $h_{pq}$ is

$$h_{pq} = \langle p | h | q \rangle \quad (47)$$

$$= \left( \sum_\mu C_{\mu p} \langle \mu | \right) h \left( \sum_\nu | \nu \rangle C_{\nu q} \right) \quad (48)$$

$$= \sum_{\mu\nu} C_{\mu p} h_{\mu\nu} C_{\nu q}, \quad (49)$$

and its derivative can be represented as

$$h_{pq}^{[x]} = \sum_{\mu\nu} C_{\mu p}^{[x]} h_{\mu\nu} C_{\nu q} + \sum_{\mu\nu} C_{\mu p} h_{\mu\nu}^{[x]} C_{\nu q} + \sum_{\mu\nu} C_{\mu p} h_{\mu\nu} C_{\nu q}^{[x]}. \quad (50)$$

Similarly, the two-electron integrals are

$$\Pi_{pqsr} = \sum_{\mu\nu\lambda\sigma} C_{\mu p} C_{\nu q} \Pi_{\mu\nu\lambda\sigma} C_{\lambda s} C_{\sigma r}, \quad (51)$$

and their derivative is

$$\Pi_{pqsr}^{[x]} = \sum_{\mu\nu\lambda\sigma} C_{\mu p}^{[x]} C_{\nu q} \Pi_{\mu\nu\lambda\sigma} C_{\lambda s} C_{\sigma r}$$

$$+ \sum_{\mu\nu\lambda\sigma} C_{\mu p} C_{\nu q}^{[x]} \Pi_{\mu\nu\lambda\sigma} C_{\lambda s} C_{\sigma r}$$

$$+ \sum_{\mu\nu\lambda\sigma} C_{\mu p} C_{\nu q} \Pi_{\mu\nu\lambda\sigma}^{[x]} C_{\lambda s} C_{\sigma r}$$

$$+ \sum_{\mu\nu\lambda\sigma} C_{\mu p} C_{\nu q} \Pi_{\mu\nu\lambda\sigma} C_{\lambda s}^{[x]} C_{\sigma r}$$

$$+ \sum_{\mu\nu\lambda\sigma} C_{\mu p} C_{\nu q} \Pi_{\mu\nu\lambda\sigma} C_{\lambda s} C_{\sigma r}^{[x]}. \tag{52}$$

In converting $O_{bi}^{R[x]}$ to the AO representation, we must keep in mind that the MO coefficients depend parametrically on $x$,

$$O_{bi}^{R[x]} = \langle b | i^{[x]} \rangle \tag{53}$$

$$= \left( \sum_{\mu} C_{\mu b} \langle \mu | \right) \left( \sum_{\nu} | \nu \rangle C_{\nu i}^{[x]} + \sum_{\nu} | \nu^{[x]} \rangle C_{\nu i} \right) \tag{54}$$

$$= \sum_{\mu\nu} C_{\mu b} S_{\mu\nu} C_{\nu i}^{[x]} + \sum_{\mu\nu} C_{\mu b} S_{\mu\nu}^{R[x]} C_{\nu i}, \tag{55}$$

where we have identified the atomic-orbital overlap $S_{\mu\nu} = \langle \mu | \nu \rangle$. We have also defined its left and right derivatives,

$$S_{\mu\nu}^{[x]} = \langle \mu^{[x]} | \nu \rangle + \langle \mu | \nu^{[x]} \rangle \tag{56}$$

$$\equiv S_{\mu\nu}^{L[x]} + S_{\mu\nu}^{R[x]}. \tag{57}$$

The overlap derivatives are not symmetric; they are transposes, such that

$$S_{\mu\nu}^{R[x]} = S_{\nu\mu}^{L[x]}, \tag{58}$$

and $(1/2)S^{[x]}$ can be understood as a symmetrized version of $S^{R[x]}$.

To complete the change of representation, we need a few more definitions:[55]

$$R_{\mu\nu}^{I} \equiv \sum_{ia} C_{\mu a} t_i^{Ia} C_{\nu i} \tag{59a}$$

$$B_{\mu\nu}^{IJ} \equiv \sum_{iab} C_{\mu a} t_i^{Ia} t_i^{Jb} C_{\nu b} - \sum_{ija} C_{\mu i} t_i^{Ia} t_j^{Ja} C_{\nu j} \tag{59b}$$

$$P_{\mu\nu} \equiv \sum_{m} C_{\mu m} C_{\nu m} \tag{59c}$$

$$\widetilde{P}_{\mu\nu} \equiv \sum_{p} C_{\mu p} C_{\nu p} = P_{\mu\nu} + \sum_{a} C_{\mu a} C_{\nu a}. \tag{59d}$$

In order, these expressions define the transition density for state $|\Psi_I\rangle$, a generalized difference density for states $|\Psi_I\rangle$ and $|\Psi_J\rangle$, the usual density matrix, and a symmetric matrix formally equivalent to $S^{-1}$. Combining the results of this section

with Eqs. (40)–(43) and (46), we obtain

$$\langle \Psi_I | H^{[x]} | \Psi_J \rangle = \sum_{\mu\nu} B_{\mu\nu}^{IJ} h_{\mu\nu}^{[x]} + \sum_{\mu\nu\lambda\sigma} \left( R_{\mu\lambda}^I R_{\sigma\nu}^J + B_{\mu\lambda}^{IJ} P_{\sigma\nu} \right) \Pi_{\mu\nu\lambda\sigma}^{[x]}$$

$$- \sum_{\mu\nu\alpha\beta} S_{\mu\nu}^{R[x]} \widetilde{P}_{\mu\alpha} \left( B_{\beta\nu}^{IJ} + B_{\nu\beta}^{IJ} \right) F_{\alpha\beta}$$

$$- \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{R[x]} \widetilde{P}_{\mu\alpha} \left( R_{\nu\gamma}^I R_{\delta\beta}^J + R_{\delta\beta}^I R_{\nu\gamma}^J \right) \Pi_{\alpha\beta\gamma\delta}$$

$$- \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{R[x]} \widetilde{P}_{\mu\alpha} \left( R_{\gamma\nu}^I R_{\beta\delta}^J + R_{\beta\delta}^I R_{\gamma\nu}^J \right) \Pi_{\alpha\beta\gamma\delta}$$

$$+ \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{R[x]} \widetilde{P}_{\mu\alpha} P_{\nu\delta} \left( B_{\gamma\beta}^{IJ} + B_{\beta\gamma}^{IJ} \right) \Pi_{\alpha\beta\gamma\delta} \tag{60a}$$

$$Y_{bi} = \sum_{d\mu\nu\lambda\sigma} C_{\mu b} C_{\lambda d} \left( R_{\nu\sigma}^I t_i^{Jd} + t_i^{Id} R_{\nu\sigma}^J \right) \Pi_{\mu\nu\lambda\sigma}$$

$$+ \sum_{\mu\nu\lambda\sigma} C_{\nu b} C_{\sigma i} \left( B_{\mu\lambda}^{IJ} + B_{\lambda\mu}^{IJ} \right) \Pi_{\mu\nu\lambda\sigma}$$

$$+ \sum_{\ell\mu\nu\lambda\sigma} C_{\mu\ell} C_{\sigma i} \left( R_{\nu\lambda}^I t_\ell^{Jb} + t_\ell^{Ib} R_{\nu\lambda}^J \right) \Pi_{\mu\nu\lambda\sigma}. \tag{60b}$$

### F. Derivatives of the molecular orbital coefficients

In order to evaluate $O_{bi}^{R[x]}$ in Eq. (55), we must determine the derivative of the canonical Hartree-Fock coefficients **C**; we follow the standard procedure summarized in Ref. 55. MO coefficients for theories referenced to Hartree-Fock are parameterized as

$$\mathbf{C}(\mathbf{C}^\circ, \mathbf{S}, \Theta) = \mathbf{C}^\circ \left( \mathbf{C}^{\circ\mathsf{T}} \mathbf{S} \mathbf{C}^\circ \right)^{-1/2} \exp \left( \begin{bmatrix} \mathbf{0} & \Theta^\mathsf{T} \\ -\Theta & \mathbf{0} \end{bmatrix} \right), \tag{61}$$

in terms of some initial guess $\mathbf{C}^\circ$, the AO overlap matrix $\mathbf{S}$, and the occupied-to-virtual rotation angles $\Theta$; the latter conventionally being written as the $V \times O$ matrix

$$\Theta = \begin{bmatrix} \Theta_{ai} & \Theta_{aj} & \cdots \\ \Theta_{bi} & \Theta_{bj} & \\ \vdots & & \ddots \end{bmatrix}. \tag{62}$$

As shown in Ref. 55, the derivatives of the MO coefficients with respect to $\mathbf{S}$ and $\Theta$ are

$$\frac{\partial C_{\mu p}}{\partial S_{\alpha\beta}} = -\frac{1}{2} \sum_{q} C_{\mu q} C_{\alpha q} C_{\beta p}, \tag{63a}$$

$$\frac{\partial C_{\mu p}}{\partial \Theta_{aj}} = C_{\mu j} \delta_{ap} - C_{\mu a} \delta_{jp}. \tag{63b}$$

Now we apply Eq. (63) to $O_{bi}^{R[x]}$,

$$O_{bi}^{R[x]} = \sum_{\mu\nu} C_{\mu b} S_{\mu\nu} C_{\nu i}^{[x]} + \sum_{\mu\nu} C_{\mu b} S_{\mu\nu}^{R[x]} C_{\nu i} \qquad (64)$$

$$= \sum_{\mu\nu} C_{\mu b} S_{\mu\nu} \sum_{\alpha\beta} \frac{\partial C_{\nu i}}{\partial S_{\alpha\beta}} S_{\alpha\beta}^{[x]}$$
$$+ \sum_{\mu\nu} C_{\mu b} S_{\mu\nu} \sum_{ja} \frac{\partial C_{\nu i}}{\partial \Theta_{aj}} \Theta_{aj}^{[x]} + \sum_{\mu\nu} C_{\mu b} S_{\mu\nu}^{R[x]} C_{\nu i} \qquad (65)$$

$$= \sum_{\mu\nu} C_{\mu b} S_{\mu\nu}^{R[x]} C_{\nu i} - \frac{1}{2} \sum_{q\mu\nu\alpha\beta} C_{\mu b} S_{\mu\nu} C_{\nu q} C_{\alpha q} S_{\alpha\beta}^{[x]} C_{\beta i}$$
$$+ \sum_{ja\mu\nu} C_{\mu b} S_{\mu\nu} C_{\nu j} \delta_{ia} \Theta_{aj}^{[x]} - \sum_{ja\mu\nu} C_{\mu b} S_{\mu\nu} C_{\nu a} \delta_{ij} \Theta_{aj}^{[x]} \qquad (66)$$

$$= \sum_{\mu\nu} C_{\mu b} S_{\mu\nu}^{R[x]} C_{\nu i} - \frac{1}{2} \sum_{q\alpha\beta} \delta_{bq} C_{\alpha q} S_{\alpha\beta}^{[x]} C_{\beta i} - \sum_{ia} \delta_{ba} \Theta_{ai}^{[x]} \qquad (67)$$

$$= \sum_{\mu\nu} C_{\mu b} \left( S_{\mu\nu}^{R[x]} - \frac{1}{2} S_{\mu\nu}^{[x]} \right) C_{\nu i} - \Theta_{bi}^{[x]} \qquad (68)$$

$$\equiv \sum_{\mu\nu} C_{\mu b} S_{\mu\nu}^{A[x]} C_{\nu i} - \Theta_{bi}^{[x]}, \qquad (69)$$

where we have used the fact that the MO coefficients orthogonalize the overlap ($\mathbf{C}^{\mathsf{T}}\mathbf{S}\mathbf{C} = \mathbf{1}$) and defined $\mathbf{S}^{A[x]}$, the antisymmetrized version of $\mathbf{S}^{R[x]}$ (cf. Eq. (57)),

$$S_{\mu\nu}^{A[x]} \equiv S_{\mu\nu}^{R[x]} - \frac{1}{2} S_{\mu\nu}^{[x]} \qquad (70)$$

$$= \frac{1}{2} \left( S_{\mu\nu}^{R[x]} - S_{\nu\mu}^{R[x]} \right). \qquad (71)$$

In Eq. (69), we see that $O_{bi}^{R[x]}$ includes a contribution from the orbital response $\Theta_{bi}^{[x]}$, just as one finds for a CIS gradient. In the language of MR-CI theory, we need only compute the response for "essential" rotations between the occupied and virtual subspace; we need not determine the response associated with "redundant" orbital rotations, which do not affect the CI energy.[13] Indeed, the trivially-redundant rotations appearing in our theory—those within the occupied and virtual Hartree-Fock subspaces—vanish identically. (Complex redundancies involving rotations among the configuration-state functions do appear in MR-CI.) Note also that the $\mathbf{S}^{A[x]}$ term in Eq. (69) is a direct analogue of the "orbital" portion of the CSF term in MR-CI theory.[13]

## G. Orbital response and the z vector

In practice, almost every term in Eqs. (44), (60), and (69) can be computed in memory using standard quantum chemistry or simple modifications to existing codes. The only

task remaining is to calculate $\Theta_{bi}^{[x]}$, which we do using standard coupled-perturbed Hartree-Fock (CPHF) theory.[56,57] For completeness, we now review the necessary formalism: From $(\partial E/\partial\Theta_{aj})^{[x]} = 0$, we obtain the CPHF equation,

$$\sum_{ib} \frac{\partial^2 E}{\partial \Theta_{aj} \partial \Theta_{bi}} \Theta_{bi}^{[x]} = -\sum_{\alpha\beta} \frac{\partial^2 E}{\partial \Theta_{aj} \partial S_{\alpha\beta}} S_{\alpha\beta}^{[x]}$$
$$- \sum_{\alpha\beta} \frac{\partial^2 E}{\partial \Theta_{aj} \partial h_{\alpha\beta}} h_{\alpha\beta}^{[x]}$$
$$- \sum_{\alpha\beta\gamma\delta} \frac{\partial^2 E}{\partial \Theta_{aj} \partial \Pi_{\alpha\beta\gamma\delta}} \Pi_{\alpha\beta\gamma\delta}^{[x]} \qquad (72)$$

$$\equiv -M_{aj}^{[x]}, \qquad (73)$$

where we have defined the mixed-derivative term $\mathbf{M}^{[x]}$. We can *formally* invert the tensor associated with the double angle-derivative at left to obtain

$$\Theta_{bi}^{[x]} = -\sum_{ja} \left( \frac{\partial^2 E}{\partial \Theta_{aj} \partial \Theta_{bi}} \right)^{-1} M_{aj}^{[x]}. \qquad (74)$$

Equations (44) and (69) indicate that we only ever need the dot product of $\mathbf{\Theta}^{[x]}$ and $\mathbf{Y}$. The celebrated **z**-vector method of Handy and Schaefer exploits this fact to avoid *actually* inverting the tensor by instead iteratively constructing the *aj* elements of its product with $\mathbf{Y}$, the **z**-vector Lagrangian:

$$-\sum_{ib} Y_{bi} \Theta_{bi}^{[x]} = \sum_{ijab} Y_{bi} \left( \frac{\partial^2 E}{\partial \Theta_{aj} \partial \Theta_{bi}} \right)^{-1} M_{aj}^{[x]} \qquad (75)$$

$$= \sum_{ja} \left[ \sum_{ib} Y_{bi} \left( \frac{\partial^2 E}{\partial \Theta_{aj} \partial \Theta_{bi}} \right)^{-1} \right] M_{aj}^{[x]} \qquad (76)$$

$$\equiv \sum_{ja} z_{aj} M_{aj}^{[x]}, \qquad (77)$$

where we have defined **z** according to the usual prescription.[58]

Finally, we construct $\mathbf{M}^{[x]}$. We quote the relevant results from Hartree-Fock theory:[59,60]

$$\frac{\partial^2 E}{\partial \Theta_{aj} \partial h_{\alpha\beta}} = -C_{\alpha a} C_{\beta j} - C_{\alpha j} C_{\beta a} \qquad (78a)$$

$$\frac{\partial^2 E}{\partial \Theta_{aj} \partial \Pi_{\alpha\beta\gamma\delta}} = -C_{\alpha a} P_{\beta\delta} C_{\gamma j} - C_{\alpha j} P_{\beta\delta} C_{\gamma a} \qquad (78b)$$

$$\frac{\partial^2 E}{\partial \Theta_{aj} \partial S_{\alpha\beta}} = \varepsilon_j C_{\alpha j} C_{\beta a} + \varepsilon_a C_{\alpha a} C_{\beta j}$$
$$+ \frac{1}{2} \sum_{\mu\nu\lambda\sigma} \Pi_{\mu\lambda\nu\sigma} (\widetilde{P}_{\lambda\alpha} P_{\sigma\beta} + P_{\lambda\beta} \widetilde{P}_{\sigma\alpha})$$
$$\times (C_{\mu a} C_{\nu j} + C_{\mu j} C_{\nu a}). \qquad (78c)$$

By taking dot products of Eq. (78) with **z**, we obtain

$$\sum_{ja} z_{aj} \frac{\partial^2 E}{\partial \Theta_{aj} \partial h_{\alpha\beta}} = -(z_{\alpha\beta} + z_{\beta\alpha}) \tag{79a}$$

$$\sum_{ja} z_{aj} \frac{\partial^2 E}{\partial \Theta_{aj} \partial \Pi_{\alpha\beta\gamma\delta}} = -(z_{\alpha\gamma} + z_{\gamma\alpha})P_{\delta\beta} \tag{79b}$$

$$\sum_{ja} z_{aj} \frac{\partial^2 E}{\partial \Theta_{aj} \partial S_{\alpha\beta}} = \sum_{\mu\nu} \widetilde{P}_{\mu\alpha} F_{\mu\nu}(z_{\beta\nu} + z_{\nu\beta})$$
$$+ \frac{1}{2} \sum_{\mu\nu\lambda\sigma} (\widetilde{P}_{\lambda\alpha} P_{\sigma\beta} + P_{\lambda\beta} \widetilde{P}_{\sigma\alpha})$$
$$\times (z_{\mu\nu} + z_{\nu\mu}) \Pi_{\mu\lambda\nu\sigma}, \tag{79c}$$

which expressions appear implicitly in Eq. (77). Taken together, these terms account for relaxation due to orbital response of the derivatives of the one- and two-electron integrals and the overlap matrix.

Although most professional quantum-chemistry codes have built-in **z**-vector solvers, we note for completeness the double-angle-derivative tensor used in solving the CPHF equations,

$$\frac{\partial^2 E}{\partial \Theta_{aj} \partial \Theta_{bi}} = 2\delta_{ij}\delta_{ab}(\varepsilon_b - \varepsilon_i)$$
$$+ \sum_{\mu\nu\lambda\sigma} \Pi_{\mu\lambda\nu\sigma}(C_{\lambda b}C_{\sigma i} + C_{\lambda i}C_{\sigma b})$$
$$\times (C_{\mu a}C_{\nu j} + C_{\mu j}C_{\nu a}). \tag{80}$$

## H. Exact CIS derivative couplings

Unifying the above results, we summarize the derivative coupling expression as

$$\langle \Psi_I | \Psi_J^{[x]} \rangle = \frac{\langle \Psi_I | H^{[x]} | \Psi_J \rangle}{E_J - E_I} + \frac{1}{E_J - E_I} \sum_{ja} z_{aj} M_{aj}^{[x]}$$
$$+ \frac{1}{E_J - E_I} \sum_{ib\mu\nu} Y_{bi} C_{\mu b} S_{\mu\nu}^{A[x]} C_{\nu i}. \tag{81}$$

The first term is the naïve Hellmann-Feynman contribution to the coupling; the second term corrects for orbital response; and the third term is a non-response correction. Their components are, in detail,

$$\langle \Psi_I | H^{[x]} | \Psi_J \rangle$$
$$= \sum_{\mu\nu} B_{\mu\nu}^{IJ} h_{\mu\nu}^{[x]} + \sum_{\mu\nu\lambda\sigma} \left( R_{\mu\lambda}^I R_{\sigma\nu}^J + B_{\mu\lambda}^{IJ} P_{\sigma\nu} \right) \Pi_{\mu\nu\lambda\sigma}^{[x]}$$
$$- \sum_{\mu\nu\alpha\beta} S_{\mu\nu}^{R[x]} \widetilde{P}_{\mu\alpha} \left( B_{\beta\nu}^{IJ} + B_{\nu\beta}^{IJ} \right) F_{\alpha\beta}$$
$$- \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{R[x]} \widetilde{P}_{\mu\alpha} \left( R_{\nu\gamma}^I R_{\delta\beta}^J + R_{\delta\beta}^I R_{\nu\gamma}^J \right) \Pi_{\alpha\beta\gamma\delta}$$

$$- \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{R[x]} \widetilde{P}_{\mu\alpha} \left( R_{\gamma\nu}^I R_{\beta\delta}^J + R_{\beta\delta}^I R_{\gamma\nu}^J \right) \Pi_{\alpha\beta\gamma\delta}$$
$$+ \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{R[x]} \widetilde{P}_{\mu\alpha} P_{\nu\delta} \left( B_{\gamma\beta}^{IJ} + B_{\beta\gamma}^{IJ} \right) \Pi_{\alpha\beta\gamma\delta} \tag{82a}$$

$$\sum_{ja} z_{aj} M_{aj}^{[x]} = -\sum_{\alpha\beta}(z_{\alpha\beta} + z_{\beta\alpha})h_{\alpha\beta}^{[x]}$$
$$- \sum_{\alpha\beta\gamma\delta}(z_{\alpha\gamma} + z_{\gamma\alpha})P_{\delta\beta}\Pi_{\alpha\beta\gamma\delta}^{[x]}$$
$$+ \sum_{\mu\nu\alpha\beta} S_{\alpha\beta}^{[x]} \widetilde{P}_{\mu\alpha} F_{\mu\nu}(z_{\beta\nu} + z_{\nu\beta})$$
$$+ \frac{1}{2} \sum_{\mu\nu\lambda\sigma\alpha\beta} S_{\alpha\beta}^{[x]} (\widetilde{P}_{\lambda\alpha} P_{\sigma\beta} + P_{\lambda\beta} \widetilde{P}_{\sigma\alpha})$$
$$\times (z_{\mu\nu} + z_{\nu\mu}) \Pi_{\mu\lambda\nu\sigma} \tag{82b}$$

$$Y_{bi} = \sum_{d\mu\nu\lambda\sigma} C_{\mu b} C_{\lambda d} \left( R_{\nu\sigma}^I t_i^{Jd} + t_i^{Id} R_{\nu\sigma}^J \right) \Pi_{\mu\nu\lambda\sigma}$$
$$+ \sum_{\mu\nu\lambda\sigma} C_{\nu b} C_{\sigma i} \left( B_{\mu\lambda}^{IJ} + B_{\lambda\mu}^{IJ} \right) \Pi_{\mu\nu\lambda\sigma}$$
$$+ \sum_{\ell\mu\nu\lambda\sigma} C_{\mu\ell} C_{\sigma i} \left( R_{\nu\lambda}^I t_\ell^{Jb} + t_\ell^{Ib} R_{\nu\lambda}^J \right) \Pi_{\mu\nu\lambda\sigma}, \tag{82c}$$

with the **z** vector determined by the relationship

$$\sum_{ja} \frac{\partial^2 E}{\partial \Theta_{aj} \partial \Theta_{bi}} z_{aj} = Y_{bi} \tag{83}$$

and the quantities $\mathbf{S}^{R[x]}$, $\mathbf{R}^I$, $\mathbf{B}^{IJ}$, $\mathbf{P}$, and $\widetilde{\mathbf{P}}$ defined in Eqs. (57) and (59). As usual, **C**, **h**, $\Pi$, and $\{t_i^{Ia}\}$ and $\{t_i^{Ja}\}$ are the MO coefficients, one-electron integrals, two-electron integrals, and CIS amplitudes for states $|\Psi_I\rangle$ and $|\Psi_J\rangle$.

We stress that derivative couplings given by Eqs. (81)–(83) are exact within CIS theory; no approximations have been made, so all Pulay effects are included. And because our method requires only matrix multiplications and a single **z**-vector call—as is true for MR-CI—we can compute all $3N$ derivative couplings at a cost comparable to that of a single CIS gradient.

If we compare our expressions with those for MR-CI derivative couplings in the case of a single-determinant reference, the final answers must agree.[13] The correspondence is not obvious, however. In particular, the Hellmann-Feynman term in Eq. (81) contains components proportional to $\mathbf{S}^{A[x]}$, which (by contrast) is cleanly isolated in the orbital-CSF term of MR-CI. We present an alternate derivation in the Appendix that emphasizes the close correspondence to MR-CI throughout.

## III. COMPARISON WITH FINITE DIFFERENCE

We have implemented Eqs. (81)–(83) in both restricted and unrestricted form within a development version of the

commercially available program Q-Chem.[61] Initial testing showed that Eq. (5a) was indeed satisfied. Couplings between singlet and triplet states vanished, as required by symmetry.

In order to validate our analytic derivative couplings explicitly, we made a direct comparison with finite-difference results. The standard central-difference formula yields the following expression for the derivative coupling:

$$\langle \Psi_I | \Psi_J^{[x]} \rangle \approx \frac{\langle \Psi_I(x) | \Psi_J(x + \Delta x) \rangle - \langle \Psi_I(x) | \Psi_J(x - \Delta x) \rangle}{2\Delta x}$$
$$+ \, O((\Delta x)^2). \tag{84}$$

Neither the singles determinants nor the underlying molecular orbitals are orthonormal when computed for different geometries; the same is therefore true of the CIS states. Using brackets with internal primes ($|\cdot'\rangle$) as shorthand for spin orbitals of the system in a distorted configuration, we can write

$$\langle \Psi_I | \Psi_J' \rangle = \sum_{ijab} t_i^{Ia} t_j^{Jb'} \langle \Phi_i^a | \Phi_j^{b'} \rangle. \tag{85}$$

$$= \sum_{ijab} t_i^{Ia} t_j^{Jb'} (\delta_{ij} \langle a|b'\rangle + \langle a|i'\rangle \langle j|b'\rangle) \prod_{k \neq i,j} \langle k|k'\rangle \tag{86}$$

$$= \sum_{iab} t_i^{Ia} t_i^{Jb'} \langle a|b'\rangle \prod_{k \neq i} \langle k|k'\rangle$$
$$+ \sum_{i \neq j,ab} t_i^{Ia} t_j^{Jb'} \langle a|i'\rangle \langle j|b'\rangle \prod_{k \neq i,j} \langle k|k'\rangle. \tag{87}$$

The overlaps in Eq. (87) depend implicitly on two sets of MO coefficients *and* the corresponding AO overlap matrix, $\langle \mu | \nu' \rangle$.

The test system we chose is lithium hydride (LiH), a prototypical heteronuclear diatom containing four electrons; we calculated derivative couplings at the HF/cc-pVDZ levels of theory. The geometry used in both calculations was that optimized for HF/cc-pVDZ; the LiH bond length is 1.618436 Å, aligned for convenience with the z axis. Although not reported here, we have made similar comparisons for $H_2$, $HeH^+$, $O_2$, and LiF; the agreement between our analytical theory and finite difference is confirmed for all of these systems.

Table I lists the CIS energies for the first six singlet excited states of LiH. Note in particular that the second/third and fifth/sixth CIS states form degenerate pairs. Theories of Hellmann-Feynman form will, of course, diverge when $|\Psi_I\rangle$ and $|\Psi_J\rangle$ are degenerate, so we tested our theory against finite difference for the coupling between the first and fourth CIS states.

TABLE I. Excitation energies for the first six singlet CIS states of HF/cc-pVDZ-optimized LiH. The ground state (VDZ) Hartree-Fock energy is $E = -7.983686 \, E_h$.

| State | $E_{CIS} - E$ (eV) |
|---|---|
| 1 | 4.0248 |
| 2 | 5.0651 |
| 3 | 5.0651 |
| 4 | 6.9219 |
| 5 | 7.8317 |
| 6 | 7.8317 |

TABLE II. Derivative couplings between the first and fourth CIS states of LiH at the HF/cc-pVDZ level of theory computed by finite difference (FD, Eq. (87)), analytical theory (CIS, Eqs. (81)–(83)), CIS with Pulay terms omitted (NP, Eq. (88)), and CIS corrected using built-in electron-translation factors (ETF, Eq. (121)). Note that FD and CIS agree to at least the expected order of errors ($\sim 10^{-4} \, a_0^{-1}$).

| Atom moved ($Q$) | $\langle \Psi_1 | \Psi_4^{[Q]} \rangle$ ($a_0^{-1}$) | | | |
| --- | --- | --- | --- | --- |
| | FD | CIS | NP | ETF |
| H | 0.0478 | 0.0479 | 0.0634 | 0.0177 |
| Li | −0.1466 | −0.1466 | −0.0634 | −0.0177 |

The analytically-vanishing couplings $\langle \Psi_I | \Psi_I^{[x]} \rangle$, which provide a direct measure of the finite-difference error, should scale quadratically with $\Delta x$. We observed that this scaling is broken for very small step sizes—numerical imprecisions in the wavefunction begin to overwhelm the finite-difference error. Consistent with this behavior, we chose a step size of $\Delta x = 10^{-4}$ Å, which lies at the low end of the acceptable range. We set the SCF and CIS energy convergence conditions, which also more loosely bound the precision of the MO coefficients and CIS amplitudes, respectively, to $10^{-8} \, E_h$. We therefore expect agreement between the analytical theory and finite difference to be no better than $\sim 10^{-4} \, a_0^{-1}$. The results in Table II ("CIS" column) show that we do, in fact, reach this level of agreement. When referring to Table II, note that while derivative couplings are signed quantities, only relative signs are locally physically meaningful; signs have been adjusted for consistency.

Though finite difference is a straightforward procedure, it becomes impractical for larger systems; obtaining the full set of derivative couplings requires the calculation of at least $3N + 1$ CIS energies. ($n$th-order finite-difference methods, such as the second-order central-difference formula of Eq. (84), require $n3N + 1$ energy calculations.) Our method is effectively fixed at a cost comparable to that of a single gradient (as is true in MR-CI theory).[13]

## IV. PULAY EFFECTS IN NAPHTHALENE DERIVATIVE COUPLINGS

As an interesting application of our theory, we have computed the derivative couplings between the two lowest-lying singlet excited states of naphthalene ($C_{10}H_8$, as shown in Figure 1) for all 54 of its nuclear degrees of freedom. Coupling between these states has long been appreciated to be important in electronic spectroscopy of naphthalene and its derivatives;[62–65] Table III shows that they are well-separated from the next-highest excitation. The coupling-vector magnitudes are given in Table IV; we depict the vectors superimposed on the HF/VDZ-optimized naphthalene structure in Figure 2. Note that the derivative-coupling vectors exhibit the same $D_{2h}$ symmetry as naphthalene itself.

Irrespective of the basis set used, we find that the largest derivative couplings are due to in-plane motion of the ring-fusion carbons ($C_A$) perpendicular to the bond between them; the next-largest couplings arise from ring deformations involving the outer carbons ($C_C$); and the motions of the middle
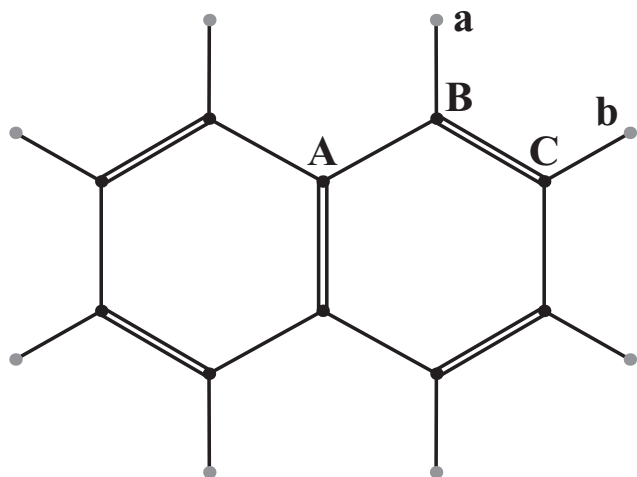
FIG. 1. The structure of naphthalene. Carbon labels {A, B, C} and hydrogen labels {a, b} are used in Table IV. Black and grey points represent carbons and hydrogens, respectively.

carbons ($C_B$) and hydrogens lead to couplings that, though non-zero, are significantly weaker. The cc-pVDZ couplings are always smaller than those for VDZ, which suggests that small basis sets may over-couple somewhat.

Referring to Table IV ("NP" column) shows that the above description of naphthalene derivative couplings does not apply if Pulay effects are omitted altogether—equivalent to making the replacements

$$\mathbf{S}^{[x]}, \mathbf{S}^{R[x]}, \mathbf{S}^{A[x]} \rightarrow 0 \qquad (88)$$

in Eqs. (81)–(83). Without the inclusion of Pulay effects, the observed couplings are qualitatively different; in particular, the strongest couplings are those associated with motion of the middle carbons. Hydrogen-induced couplings are also enhanced significantly. This qualitative change provides a stark example of the importance of Pulay effects in derivative-coupling calculations.

## V. TRANSLATIONAL VARIANCE AND ELECTRON-TRANSLATION FACTORS

### A. Translational variance in derivative couplings

It has long been recognized that MR-CI derivative couplings exhibit translational and rotational variance.[11,43] (We do not address rotational variance here.) In particular, derivative couplings computed using only internal nuclear coordinates exhibit origin dependence, while derivative couplings

TABLE III. Excitation energies for the first three singlet CIS states of HF/VDZ-optimized naphthalene. The ground state Hartree-Fock energy is $E = -383.384480\ E_h$.

| State | $E_{CIS} - E$ (eV) | |
|---|---|---|
|  | HF/VDZ | HF/cc-pVDZ |
| 1 | 5.4365 | 5.2436 |
| 2 | 5.5361 | 5.3782 |
| 3 | 7.2344 | 7.0382 |

TABLE IV. Magnitudes of the vector derivative couplings between the first and second CIS states of naphthalene, as computed using analytical theory (CIS, Eqs. (81)–(83)), CIS with Pulay terms omitted (NP, Eq. (88)), and CIS corrected using built-in electron-translation factors (ETF, Eq. (121)). Atom labels refer to Figure 1.

| | $\left\| \langle \Psi_1 \big| \Psi_2^{[Q]} \rangle \right\|\ (a_0^{-1})$ | | | | | |
|---|---|---|---|---|---|---|
| | HF/VDZ | | | HF/cc-pVDZ | | |
| Atom moved ($Q$) | CIS | NP | ETF | CIS | NP | ETF |
| $C_A$ | 1.1759 | 0.0333 | 1.2618 | 0.8317 | 0.0306 | 0.9270 |
| $C_B$ | 1.0807 | 1.4258 | 1.0644 | 0.7794 | 1.1132 | 0.7566 |
| $C_C$ | 0.1443 | 2.0371 | 0.1772 | 0.0907 | 1.5059 | 0.1268 |
| $H_a$ | 0.0196 | 0.3305 | 0.0196 | 0.0074 | 0.2608 | 0.0075 |
| $H_b$ | 0.0850 | 0.7771 | 0.0850 | 0.0753 | 0.4849 | 0.0756 |

computed using atomic coordinates fail to satisfy the sum rule

$$\sum_Q \left\langle \Psi_I \big| \Psi_J^{[x_Q]} \right\rangle = 0. \qquad (89)$$

A glance at the CIS column in Table II shows that our theory also generates translationally variant derivative couplings (as the analogy to MR-CI demands). Translational variance of the derivative couplings is not unique to CI-based theories; it appears in the context of TD-DFT as well.[40] Most strikingly, Eq. (89) fails for couplings obtained from an analytical treatment of the hydrogen atom—there are, for example, non-zero derivative couplings between the $|2s\rangle$ and $|2p_x\rangle$ states![40,44] From a dynamical point of view, such couplings are disastrous; they indicate that a hydrogen atom in its $|2p_x\rangle$ state undergoing straight-line, constant-velocity motion would transfer amplitude to $|2s\rangle$. The equivalent statement for a molecule is that constant-velocity motion of the center of mass would couple electronic states—which is equally troubling. Given that an inertial reference frame exists in which the system is stationary, this scenario is nonsensical.

The only source of translational variance in our theory is the antisymmetrized overlap $\mathbf{S}^{A[x]}$; hence the translational invariance of the NP results in Tables II and IV, which were obtained from applying Eq. (88) throughout Eqs. (81)–(83). To see that $\mathbf{S}^{A[x]}$ is the culprit, consider the simplest possible example: a one-electron system consisting of two orbitals on two centers, $|\mu\rangle$ on nucleus A and $|\nu\rangle$ on nucleus B, with A to the left of B on the x axis. A very small amount of work shows that $\mathbf{S}^{[x]}$ is translationally invariant,

$$S_{\mu\nu}^{[x_A]} + S_{\mu\nu}^{[x_B]} = \langle \mu^{[x_A]} | \nu \rangle + \langle \mu | \nu^{[x_B]} \rangle = 0. \qquad (90)$$

The first term brings the two centers infinitesimally closer together, while the second term pulls them infinitesimally further apart. The corresponding changes in the overlap are equal and opposite, and the assertion holds. (Similar arguments apply to $\mathbf{h}^{[x]}$ and $\mathbf{\Pi}^{[x]}$.) Recalling the definition in Eq. (70), we find that

$$S_{\mu\nu}^{A[x_A]} + S_{\mu\nu}^{A[x_B]} = \frac{1}{2}(\langle \mu^{[x_A]} | \nu \rangle - \langle \mu | \nu^{[x_B]} \rangle) \neq 0. \qquad (91)$$

As per the discussion in Sec. II H, this result is consistent with the fact that the CI contribution to the MR-CI derivative
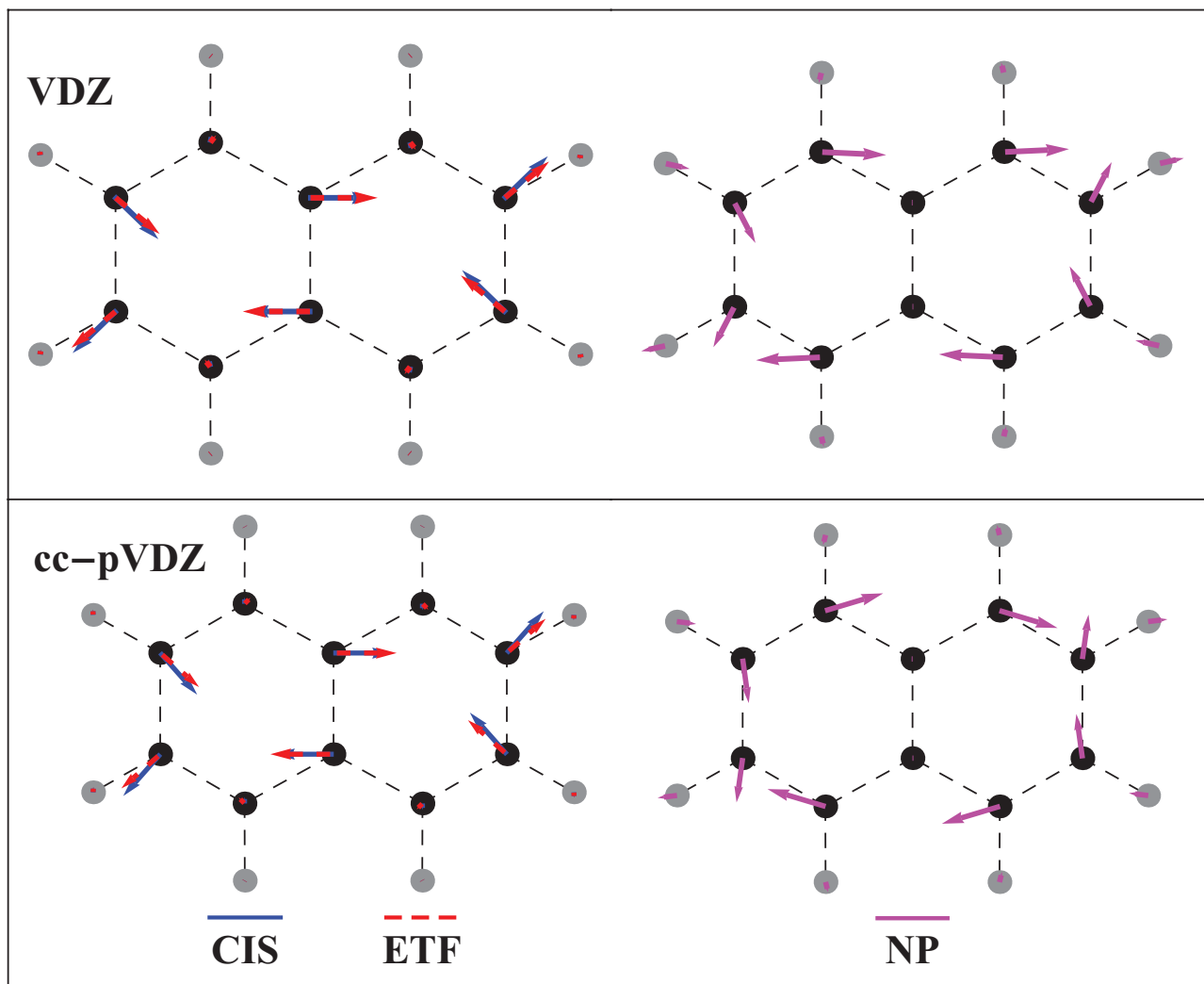
FIG. 2. Derivative coupling vectors for naphthalene computed using analytical theory (solid blue arrows, CIS, Eqs. (81)–(83)) and corrected using built-in electron-translation factors (dashed red arrows, ETF, Eq. (121)) at left; omitting all Pulay terms (solid magenta arrows, NP, Eq. (88)) at right. The qualitative difference between the panels indicates that Pulay effects cannot be neglected.

coupling is translationally and rotationally invariant, while the CSF contribution is not.[11,43]

## B. Electron-translation factors

If we wish to use our derivative couplings as input for dynamics simulations, then, we must find a way to correct this spurious behavior; the standard approach for small scattering problems is to use electron-translation factors.[45–52] The underlying concept is straightforward: The derivative coupling $\langle\Psi_I|\Psi_J^{[x]}\rangle$ arises from a combination of changes in state $|\Psi_J\rangle$ induced by the nuclear motion—such as "rotation, distortion, polarization, and change of character"[48]—*and* its translational displacement with the atom. This latter contribution leads to coupling only because Born-Oppenheimer states are real, carrying no momentum; electrons cannot simply move in tandem with the nuclei. ETFs are complex multiplicative terms that provide the necessary momentum and phase.

To put the conceptual discussion above on a more concrete footing, we consider the time evolution of an arbitrary electronic wavefunction expanded in a basis of Born-

Oppenheimer states,

$$|\Psi(t)\rangle = \sum_J c_J(t)\,|\Psi_J\rangle. \tag{92}$$

The nuclei follow classical trajectories $\{\mathbf{R}_Q(t), \mathbf{P}_Q(t)\}$, such that the electronic states are themselves implicitly time dependent. Inserting Eq. (92) into the time-dependent Schrödinger equation $i\hbar(\partial/\partial t)|\Psi(t)\rangle = H|\Psi(t)\rangle$ and projecting onto $|\Psi_I\rangle$ leads to the usual equations of motion for the expansion coefficients,[66]

$$i\hbar\dot{c}_I(t) = c_I(t)\,\langle\Psi_I|H|\Psi_J\rangle$$

$$- i\hbar\sum_J c_J(t)\sum_Q\langle\Psi_I|\nabla_Q|\Psi_J\rangle\cdot\mathbf{V}_Q. \tag{93}$$

In the case of the hydrogen atom, it is clear to see that the initial state $|\Psi(0)\rangle = |2p_x\rangle$ is not stationary—the second term on the right-hand side of Eq. (93) couples the $|2p_x\rangle$ and $|2s\rangle$ states. The solution to this problem is to add an ETF such that

the time-dependent wavefunction is of form

$$\langle \mathbf{r} | \Psi(t) \rangle = \psi_{2p_x}(\mathbf{r}) e^{-\frac{it}{\hbar} E_2} e^{\frac{i}{\hbar} m_e \mathbf{V}_H \cdot \mathbf{r}} e^{-\frac{it}{\hbar} \frac{m_e V_H^2}{2}} \equiv \langle \mathbf{r} | \widetilde{2p}_x(t) \rangle,$$
(94)

where $\mathbf{r}$ is the electronic coordinate, $E_2$ is the electronic energy of the $|2p_x\rangle$ state, $\mathbf{V}_H$ is the velocity of the hydrogen atom, and the tilde ($\sim$) indicates multiplication by an ETF. In other words, we modify the original state to include the momentum and phase associated with an electron traveling in parallel with the nucleus. (Note that this modified state is orthonormal to the other traveling bound states $\{|\widetilde{n\ell m}\rangle\}$.) Because the electronic Laplacian $\nabla_e^2$ that appears in the Hamiltonian can act on both the bound state and the translation factor, a new term of form $-i\hbar c_{2s}(t)\langle 2s | \nabla_e | 2p_x \rangle \cdot \mathbf{V}_H$ appears in the equations of motion and cancels the derivative coupling exactly. We then arrive at the correct conclusion: $|\widetilde{2p}_x(t)\rangle$, the traveling equivalent of the $|2p_x\rangle$ state, is indeed stationary.

## C. Dynamics in a traveling atomic orbital basis

The hydrogen example of Sec. V B is simplified considerably by the fact that the electron may be definitively associated with a single nucleus; there is no ambiguity as to what additional velocity the electron should carry. But the states of interest to us here—the CIS states—are formed from molecular orbitals, such that any given electron may be delocalized over the entire system. Moreover, it is not enough to account only for the motion of the nuclear center of mass; we need a detailed treatment of the instantaneous motion of each electron as it is pulled by nearby nuclei. These complications have inspired a host of ideas for assigning MOs to different atoms while also interpolating between the molecular- and atomic-orbital basis sets appropriate to various asymptotic configurations of a given system.[48,50] In light of these rather thorny issues, ETFs have most commonly been applied only to problems of atom-atom scattering.

We avoid the difficulties associated with applying ETFs at the level of the molecular orbitals by altering the underlying AOs to form a *traveling* atomic orbital (TAO) basis,[48,49,51] such that each AO carries the velocity of its nucleus. The corresponding change in the MOs is

$$\langle \mathbf{r} | n \rangle = \sum_n C_{\mu n} \chi_\mu(\mathbf{r})$$

$$\longrightarrow \quad \langle \mathbf{r} | \widetilde{n}(t) \rangle = \sum_\mu C_{\mu n} \chi_\mu(\mathbf{r}) e^{\frac{i}{\hbar} m_e \mathbf{V}_{Q_\mu} \cdot \mathbf{r}} e^{-\frac{it}{\hbar} \frac{m_e V_{Q_\mu}^2}{2}},$$
(95)

where $\chi_\mu(\mathbf{r})$ is the AO centered on nucleus $\mathbf{Q}_\mu$ at position $\mathbf{R}_{Q_\mu}$. The equations of motion for the modified wavefunction $|\widetilde{\Psi}(t)\rangle$ are

$$i\hbar \sum_J \dot{c}_J(t) \langle \widetilde{\Psi}_I(t) | \widetilde{\Psi}_J(t) \rangle$$

$$\approx \sum_J c_J(t) \langle \widetilde{\Psi}_I(t) | H | \widetilde{\Psi}_J(t) \rangle$$

$$-i\hbar \sum_J c_J(t) \sum_Q \langle \widetilde{\Psi}_I(t) | \nabla_Q | \widetilde{\Psi}_J(t) \rangle \cdot \mathbf{V}_Q$$

$$-i\hbar \sum_J c_J(t) \left\langle \widetilde{\Psi}_I(t) \left| \frac{\partial}{\partial t} \right|_{\{\mathbf{R}_Q\}} \right| \widetilde{\Psi}_J(t) \right\rangle.$$
(96)

This replacement is, of course, approximate: If we used TAOs from the very beginning of our calculations, the MO coefficients would necessarily change from instant to instant (and would differ from the original ones even at $t = 0$). It is only in the limit that the nuclear velocities $\{\mathbf{V}_Q\}$ are very small that we are justified in using this expression directly.

With this perturbative approximation in mind, we see immediately that we need only consider the effect of the TAO basis in the Hamiltonian-coupling term on the right-hand side of Eq. (96); the derivative-coupling term is already first-order, and all terms obtained from the remaining time derivative are at *least* second-order. The resulting set of equations of motion is somewhat simpler,

$$i\hbar \sum_J \dot{c}_J(t) \langle \widetilde{\Psi}_I(t) | \widetilde{\Psi}_J(t) \rangle$$

$$\approx \sum_J c_J(t) \langle \widetilde{\Psi}_I(t) | H | \widetilde{\Psi}_J(t) \rangle$$

$$-i\hbar \sum_J c_J(t) \sum_Q \langle \Psi_I | \nabla_Q | \Psi_J \rangle \cdot \mathbf{V}_Q.$$
(97)

Before moving on, we should note that approaches similar to the use of ETFs are common in obtaining molecular magnetic properties, which suffer from gauge-origin variance—spurious dependence on the origin to which the electromagnetic vector potential is referenced.[67] Standard perturbative treatments employ the London (or gauge-invariant) atomic orbitals, which are of the same form as TAOs but involve the vector potential at the nucleus instead of the electronic momentum.[67–75] This correspondence is intuitive, to some extent, since nuclear momenta and derivative couplings may also be understood as vector fields. It is not, however, perfect: Because translational variance in derivative couplings is a consequence of using real-valued adiabatic states, it would persist even if we had a complete AO basis and a complete set of electronic states. (The dynamics would, however, be exact; the exact wavefunction could be expressed as some linear combination of the states at all times.) By contrast, gauge variance in magnetic properties is simply a consequence of the use of a finite AO basis.[76]

We find that many aspects of the ETF approach developed below have been anticipated in the magnetic context: Pople applied identical atom-centering approximations and provided similar reasoning for using them,[69] and Helgaker and Jørgensen incorporated the perturbative effect of adopting a London-orbital basis into the Hamiltonian.[72] In general, the most interesting magnetic property of a molecule is its shielding tensor, which is typically evaluated at zero field. As such, the standard methodology of Wolinski *et al.* for calculating shielding tensors makes a zero-field approximation,[71] just as we assume $\{\mathbf{V}_Q\} \approx 0$. In the end, we will obtain a universal operator that corrects the derivative couplings independent of the nuclear velocities.

## D. The first-order "electron-travel" operator *W*

We want to evaluate the modified terms in Eq. (97) to first order in the atomic velocities $\{\mathbf{V}_Q\}$. We therefore define a one-electron operator $W(\mathbf{r})$ such that

$$W(\mathbf{r}_i)\phi_n(\mathbf{r}_i) = \frac{im_e}{\hbar}\sum_\mu C_{\mu n}\mathbf{V}_{Q_\mu}\cdot\mathbf{r}_i\chi_\mu(\mathbf{r}_i), \qquad (98)$$

where $\phi_n(\mathbf{r}_i) = \langle\mathbf{r}_i|n\rangle$. $W(\mathbf{r})$ is the first-order truncation of an "electron-travel" operator that adds the appropriate ETF to each atomic orbital in a given MO.

In order to obtain the full set of first-order terms in our modified electronic states, we must sum $W(\mathbf{r}_i)$ over all electrons; in second-quantized form,[53,54] the combined operator (simply called *W*) is

$$W = \sum_{pq}\langle p|W(\mathbf{r})|q\rangle\, a_p^\dagger a_q \qquad (99)$$

$$= \frac{im_e}{\hbar}\sum_{pq\mu} C_{\mu q}\mathbf{V}_{Q_\mu}\cdot\mathbf{r}_{p\mu}a_p^\dagger a_q. \qquad (100)$$

And with this definition in hand, we may rewrite our perturbative modification to the wavefunction as

$$|\widetilde{\Psi}_J(t)\rangle \approx (1 + W)\,|\Psi_J\rangle. \qquad (101)$$

Consequently, we see that the first-order state overlap and Hamiltonian matrix elements in Eq. (97) reduce to

$$\langle\widetilde{\Psi}_I(t)|\widetilde{\Psi}_J(t)\rangle \approx \delta_{IJ} + \langle\Psi_I|W + W^\dagger|\Psi_J\rangle \qquad (102)$$

$$\langle\widetilde{\Psi}_I(t)|H|\widetilde{\Psi}_J(t)\rangle \approx \langle\Psi_I|H|\Psi_J\rangle + \langle\Psi_I|HW + W^\dagger H|\Psi_J\rangle. \qquad (103)$$

The velocity factor $\mathbf{V}_{Q_\mu}$ that appears in Eq. (100) now seems quite troublesome: Because of its connection to the atomic orbital expansion of $|q\rangle$ (but *not* $|p\rangle$), $W$ is not anti-Hermitian ($W^\dagger \neq -W$). As a result, there will be modifications to both sides of the equations of motion. Furthermore, the products $HW$ and $W^\dagger H$ are neither Hermitian (like the Hamiltonian) nor anti-Hermitian (like the derivative couplings). Any complex matrix can be rewritten as the sum of a Hermitian and an anti-Hermitian part, indicating that the ETF correction would modify the derivative couplings *and* rescale the energies of the CIS states. We will invoke an additional approximation that removes these undesirable effects.

The simplest approximation is to insist that the molecular orbitals themselves be strictly atom-centered, or very nearly so. (The symmetrically-orthogonalized AOs are one such set of localized MOs.[77–80]) For molecular orbitals of this type, the velocity can be indexed by $Q_q$ instead of by $Q_\mu$. And, to a first approximation, the matrix element $\mathbf{r}_{pq}$ vanishes unless $|p\rangle$ and $|q\rangle$ are centered on the same atom. $W$ therefore becomes

$$W \approx \frac{im_e}{\hbar}\sum_{pq}\mathbf{V}_{pq}\cdot\mathbf{r}_{pq}a_p^\dagger a_q, \qquad (104)$$

where $\mathbf{V}_{pq} = \mathbf{V}_{Q_p} = \mathbf{V}_{Q_q}$. Because this approximate $W$ is anti-Hermitian, the second term on the right-hand side of

Eq. (102) disappears, and we may rewrite our equations of motion as

$$i\hbar\sum_J \dot{c}_J(t) \approx \sum_J c_J(t)\langle\Psi_I|H|\Psi_J\rangle$$
$$+ \sum_J c_J(t)\langle\Psi_I|[H, W]|\Psi_J\rangle$$
$$- i\hbar\sum_J c_J(t)\sum_Q\langle\Psi_I|\nabla_Q|\Psi_J\rangle\cdot\mathbf{V}_Q. \qquad (105)$$

## E. The correction operator [*H, W*]

We have now seen that it is possible to encode the effect of having introduced atomic ETFs in a single correction term, the details of which are determined by the correction operator $[H, W]$. Because we have not placed any restrictions on $H$, apart from its being an electronic Hamiltonian with parametric dependence on some set of nuclear positions, this operator can be evaluated for essentially any molecular system, including those with classical fields. In this section, we compute the correction operator for the simple case in which we have the bare electronic Hamiltonian of Eq. (2). (Linearity of the commutator ensures that any additional terms, such as would be required to treat the Stark effect or the spin-orbit coupling, may be evaluated separately and added to the result in Eq. (110).) In doing so, we treat the one- and two-electron portions of $[H, W]$ separately,

$$\text{one-electron part} \approx \frac{im_e}{\hbar}\sum_{pqst} h_{pq}\mathbf{V}_{st}\cdot\mathbf{r}_{st}\left[a_p^\dagger a_q, a_s^\dagger a_t\right] \qquad (106a)$$

$$\text{two-electron part} \approx \frac{im_e}{\hbar}\sum_{pqrstu}\Pi_{pqs}\mathbf{V}_{tu}\cdot\mathbf{r}_{tu}[a_p^\dagger a_q^\dagger a_r a_s, a_t^\dagger a_u]. \qquad (106b)$$

Evaluating the commutator and relabeling indices yields

$$\text{one-electron part} = \frac{im_e}{\hbar}\sum_{pst}(h_{ps}\mathbf{V}_{st}\cdot\mathbf{r}_{st} - \mathbf{V}_{ps}\cdot\mathbf{r}_{ps}h_{st})a_p^\dagger a_t \qquad (107a)$$

$$\text{two-electron part} = \frac{im_e}{\hbar}\sum_{pqrtu} a_p^\dagger a_q^\dagger a_r a_u$$
$$\times\left(\begin{array}{l}\Pi_{pqtr}\mathbf{V}_{tu}\cdot\mathbf{r}_{tu} - \mathbf{V}_{qt}\cdot\mathbf{r}_{qt}\Pi_{ptur}\\ +\Pi_{pqut}\mathbf{V}_{tr}\cdot\mathbf{r}_{tr} - \mathbf{V}_{pt}\cdot\mathbf{r}_{pt}\Pi_{tqur}\end{array}\right). \qquad (107b)$$

We stress once again that $\mathbf{r}_{pq}$ and $h_{pq}$ are matrix elements of one-electron operators that are local in real space, such that they should decay for well-separated orbitals $|p\rangle$ and $|q\rangle$. As such, the largest contributions will come from $|p\rangle$ and $|q\rangle$ on the same atom. Focusing on those orbital pairs, we see that our atom-centering approximation requires $\mathbf{V}_{st} = \mathbf{V}_{ps} = \mathbf{V}_{pt}$ in the one-electron part. We can use analogous equalities in

the two-electron part to obtain

$$\text{one-electron part} \approx \frac{im_e}{\hbar} \sum_{pst} \mathbf{V}_{pt} \cdot (h_{ps}\mathbf{r}_{st} - \mathbf{r}_{ps}h_{st}) \, a_p^\dagger a_t$$

(108a)

$$\text{two-electron part} \approx \frac{im_e}{\hbar} \sum_{pqrtu} a_p^\dagger a_q^\dagger a_r a_u$$

$$\times \left[ \begin{array}{c} \mathbf{V}_{qu} \cdot (\Pi_{pqtr}\mathbf{r}_{tu} - \mathbf{r}_{qt}\Pi_{ptur}) \\ +\mathbf{V}_{pr} \cdot (\Pi_{pqut}\mathbf{r}_{tr} - \mathbf{r}_{pt}\Pi_{tqur}) \end{array} \right].$$

(108b)

Finally, if our set of molecular orbitals is close enough to complete, we can approximate the commutator of projected operators as the projection of the exact commutator, yielding

$$\text{one-electron part} \approx \frac{im_e}{\hbar} \sum_{pt} \mathbf{V}_{pt} \cdot ([h, \mathbf{r}])_{pt} \, a_p^\dagger a_t \quad (109\text{a})$$

$$\text{two-electron part} \approx \frac{im_e}{\hbar} \sum_{pqru} a_p^\dagger a_q^\dagger a_r a_u (\mathbf{V}_{qu} + \mathbf{V}_{pr})$$

$$\times ([\Pi, \mathbf{r}])_{pqur}. \quad (109\text{b})$$

Because the two-electron integrals are purely coordinate-dependent, the commutator in the two-electron part vanishes. Inserting $[h, \mathbf{r}] = -(\hbar^2/m_e)\nabla_e$, we find that the correction operator for the bare electronic Hamiltonian is

$$[H, W] \approx -i\hbar \sum_{pt} \langle p|\nabla_e|t\rangle \cdot \mathbf{V}_{pt} a_p^\dagger a_t. \quad (110)$$

Note that—apart from the dot product with $\mathbf{V}_{pt}$—Eq. (110) is the *exact* second-quantized form of the sum of electronic momenta, $-i\hbar \sum_e \nabla_e$. Our result reflects the simple concept underpinning the use of ETFs, namely, that corrections to the derivative coupling arise from electronic motion in tandem with the nuclear motion.

  Next we expand the molecular orbitals and use the atom-centering approximation to write $\mathbf{V}_{pt} = \mathbf{V}_t = \mathbf{V}_{Q_v} = \mathbf{V}_{Q_\mu}$, meaning that the MOs $|p\rangle$ and $|t\rangle$ and the AOs $|v\rangle$ and $|\mu\rangle$ lie on the same atom:

$$[H, W] \approx -i\hbar \sum_{pt\mu v} \langle \mu|\nabla_e|v\rangle \cdot \mathbf{V}_{Q_v} C_{\mu p} C_{vt} a_p^\dagger a_t. \quad (111)$$

Because the atomic orbitals are functions of the distance between the nucleus and the electron, we may invoke the identity $\nabla_e = -\nabla_{Q_v}$, independent of any assumptions about atom centering. The atom-centering approximation further requires that $\nabla_{Q_v} = \nabla_{Q_\mu}$, such that we have a pair of equations that place a condition on the matrix element $\langle \mu|\nabla_e|v\rangle$,

$$\left\{ \begin{array}{l} \langle \mu|\nabla_e + \nabla_{Q_v}|v\rangle = 0 \\ \langle v|\nabla_e + \nabla_{Q_\mu}|\mu\rangle = 0 \end{array} \right\} \longrightarrow \left\{ \begin{array}{l} \langle \mu|\nabla_e|v\rangle + S_{\mu v}^{\mathrm{R}[Q_v]} = 0 \\ -\langle \mu|\nabla_e|v\rangle + S_{v\mu}^{\mathrm{R}[Q_v]} = 0 \end{array} \right\},$$

(112)

where we have used the above gradient identities and Eq. (57). (We have also implicitly defined vector forms of

$\mathbf{S}^{\mathrm{R}[x]}$ and the other overlap derivatives.) Adding and subtracting these equations and recalling that $\mathbf{S}^{\mathrm{A}[x]}$ is antisymmetric (cf. Eq. (70)) yields the equivalent conditions

$$S_{\mu v}^{[Q_v]} = 0 \quad (113\text{a})$$

$$\langle \mu|\nabla_e|v\rangle = S_{\mu v}^{\mathrm{A}[Q_v]}. \quad (113\text{b})$$

Equation (113a) makes sense: The overlap between AOs centered on the same atom cannot change if it moves.

  Now, $|v\rangle$ is associated with a specific nucleus, such that we always have

$$\nabla_{Q_v}|v\rangle \cdot \mathbf{V}_{Q_v} = \sum_Q \nabla_Q |v\rangle \cdot \mathbf{V}_Q. \quad (114)$$

Combining this identity with Eq. (113) and collecting terms in Eq. (111), we find

$$[H, W] \approx i\hbar \sum_Q \left[ \sum_{pt\mu v} S_{\mu v}^{\mathrm{A}[Q]} C_{\mu p} C_{vt} a_p^\dagger a_t \right] \cdot \mathbf{V}_Q, \quad (115)$$

such that the corrected derivative couplings for any given pair of wavefunctions may be computed as (cf. Eq. (105))

$$\langle \Psi_I|\Psi_J^{[Q]}\rangle_{\mathrm{ETF}} \equiv \langle \Psi_I|\Psi_J^{[Q]}\rangle - \sum_{pt\mu v} S_{\mu v}^{\mathrm{A}[Q]} C_{\mu p} C_{vt} \langle \Psi_I|a_p^\dagger a_t|\Psi_J\rangle.$$

(116)

Equation (116) is a general prescription for incorporating electron-translation factors into the derivative couplings. The correction is structured as a sum of nuclear terms, and there is no ambiguity as to whether and how the associated correction should be partitioned among the derivative couplings. Summing Eq. (116) over $Q$ and recalling the correspondence between the correction operator in Eq. (110) and the sum of electronic momenta, we recover the full-system center-of-mass identity,

$$\langle \Psi_I|\mathbf{P}_{\mathrm{CM}}|\Psi_J\rangle = -i\hbar \langle \Psi_I| \sum_Q \nabla_Q + \sum_e \nabla_e |\Psi_J\rangle = 0, \quad (117)$$

which *always* holds for isolated systems. The corrected derivative couplings are those associated with *effective* nuclear momenta obtained by partitioning the electronic momentum over $Q$. This appears to be a very "natural" way of imposing translational invariance on derivative couplings within the context of electronic-structure theory.

### F. Correcting CIS derivative couplings: Scheme and numerical results

  We now obtain the approximate correction specific to the CIS derivative couplings. To do so requires only that we evaluate the one-particle density-matrix element

$$\langle \Psi_I|a_p^\dagger a_t|\Psi_J\rangle = \sum_{ijab} t_i^{Ia} \langle \Phi_i^a|a_p^\dagger a_t|\Phi_j^b\rangle t_j^{Jb} \quad (118)$$

$$= \sum_{ijab} t_i^{Ia} t_j^{Jb} \left( \delta_{ap}\delta_{bt}\delta_{ij} - \delta_{ab}\delta_{it}\delta_{pj} + \delta_{ab}\delta_{ij}\delta_{pt}^{\mathrm{occ}} \right) \quad (119)$$

$$= \delta_{ap}\delta_{bt} \sum_{iab} t_i^{Ia} t_i^{Jb} - \delta_{it}\delta_{pj} \sum_{ija} t_i^{Ia} t_j^{Ja}. \quad (120)$$

The "occ" superscript indicates that $p$ and $t$ must be the same occupied orbital; we have used the orthogonality of the CIS amplitudes to obtain the last equality. Inserting this result into Eq. (116) and once again invoking the antisymmetry of $\mathbf{S}^{A[Q]}$, we obtain the corrected couplings for CIS,

$$
\langle \Psi_I | \Psi_J^{[Q]} \rangle_{\text{ETF}} = \langle \Psi_I | \Psi_J^{[Q]} \rangle
$$
$$
+ \sum_{\mu\nu} S_{\mu\nu}^{A[Q]} \left( \sum_{iab} C_{va} t_i^{Ia} t_i^{Jb} C_{\mu b} \right.
$$
$$
\left. + \sum_{ija} C_{vi} t_i^{Ia} t_j^{Ja} C_{\mu j} \right). \quad (121)
$$

If we consider the correction given by Eq. (116) for a single nuclear coordinate $x$, we see that the term proportional to $\mathbf{S}^{A[x]}$ *exactly* cancels one that appears in our derivative-coupling expressions (cf. the Appendix, Eq. (A25)). (In MR-CI language, it is the orbital-CSF term.) And, because $\mathbf{S}^{A[x]}$ is the only source of translational variance in our theory, including electron-translation factors restores translational invariance.

Because some of the $\mathbf{S}^{A[x]}$ terms in the Hellmann-Feynman theory of Eqs. (81)–(83) are folded into the Hamiltonian derivative, the cancellation of terms just mentioned may not be immediately clear. Thus, for concreteness, we emphasize that evaluating Eq. (121) is equivalent to making the replacement

$$
\mathbf{S}^{A[x]} \to 0 \quad (122)
$$

throughout Eqs. (81)–(83). Note the implication that $\mathbf{S}^{R[x]} \to (1/2)\mathbf{S}^{[x]}$.

Equation (121) defines CIS derivative couplings with built-in electron-translation factors. In terming the ETFs "built-in," we mean that an approximate treatment of electronic motion is included without any additional computational effort. Indeed, because the corrected couplings depend only on the standard overlap derivative $\mathbf{S}^{[x]}$, obviating the construction of $\mathbf{S}^{R[x]}$, they are slightly cheaper to compute!

To test the effect of our ETF correction, we have applied Eq. (121) to the lithium hydride and naphthalene couplings (Tables II and IV, "ETF" columns). We find that the correction is quite significant for LiH; for naphthalene, the effect is much smaller—perhaps because the electronic momentum is partitioned over a larger number of nuclei—and enhances the relative importance of ring-fusion-carbon motions. In any case, additional work will be needed to fully explore the impact of this ETF correction for molecules of various sizes (as well as for other electronic-structure theories).

## VI. CONCLUDING REMARKS

We have presented a formally exact analytical theory for the derivative couplings between CIS states, calculated in the spirit of the Hellmann-Feynman theorem and summarized in Equations (81)–(83). Our theory rigorously includes all Pulay terms associated with the use of atom-centered basis sets. In the Appendix, we present a second, independent derivation of this same result, structured instead as a direct evaluation of the derivative coupling; this derivation corresponds to well-established theory for MR-CI. Thus, Equations (81)–(83) are firmly established on theoretical grounds. Moreover, we have numerically validated the theory (as implemented in Q-Chem) against finite-difference calculations for a prototypical multi-electron diatom, lithium hydride.

The computational cost of our algorithm is low, comparable to that required to obtain a single CIS gradient, and this attribute hinges on our ability to avoid solving the coupled-perturbed Hartree-Fock equations for the orbital response associated with each degree of freedom. (The discussion in the Appendix shows that we have also avoided solving the coupled-perturbed CIS equations for the derivatives of the CIS amplitudes.) As a straightforward application of our theory, we have calculated the derivative couplings between the two lowest-lying excited singlet states of naphthalene. Our results show that these couplings are large, and that the inclusion of Pulay effects can lead to crucial qualitative changes regarding which of the atoms are involved in nonadiabatic transitions.

In order to rectify the unphysical translational variance of the derivative couplings that appears in any theory based on adiabatic electronic states, we have drawn on the scattering literature and developed a theory that modifies electronic states with approximate electron-translation factors. In doing so, we have obtained a universal operator $[H, W]$ that generates corrections to the derivative couplings for any desired Hamiltonian to first order in the nuclear momenta $\{\mathbf{V}_Q\}$; its form for the bare molecular Hamiltonian is given by Eq. (116). For isolated molecules within CIS, Eq. (121) shows that this ETF-based correction can be built into the derivative couplings, restoring translational invariance in the process. The associated corrections were smaller for naphthalene than for lithium hydride.

We have demonstrated the validity of our method, but many questions remain to be explored. What is the basis-set convergence behavior of the CIS derivative couplings, with and without built-in electron-translation factors? How well do they compare with derivative couplings from other theories? How do their translationally-variant components (and, thus, the associated corrections) scale with molecular size? What are the implications of our ETF correction when applied to MR-CI and other electronic-structure theories? And how can we best optimize the implementation of our method for use in *ab initio* molecular dynamics studies of interesting chemical systems? We hope to answer these and other questions in the future.

variance in the context of electron-translation factors. This work was supported in part by the U.S. Air Force Office of Scientific Research (USAFOSR) Grant No. FA9550-11-0092.

## APPENDIX: DIRECT DERIVATION OF THE CIS DERIVATIVE COUPLING

It is possible to obtain the derivative-coupling expression of Sec. II by means entirely distinct from the Hellmann-Feynman approach outlined there, which fact reinforces the rigor of our theory. Specifically, we can compute the coupling directly—strictly paralleling the MR-CI coupling derivation[13]—using the definition of the CIS state,

$$\langle \Psi_I | \Psi_J^{[x]} \rangle = \sum_{ijab} t_i^{Ia} \langle \Phi_i^a | \left( t_j^{Jb} | \Phi_j^b \rangle \right)^{[x]} \quad \text{(A1)}$$

$$= \sum_{ia} t_i^{Ia} t_i^{Ja[x]} + \sum_{ijab} t_i^{Ia} t_j^{Jb} \langle \Phi_i^a | \Phi_j^{b[x]} \rangle. \quad \text{(A2)}$$

The dot product of the CIS amplitudes with their derivatives is our version of the CI term; we set it aside for now. The derivative couplings between singles (the CSF term) can be evaluated according to the logic of Sec. II C, the main difference being that the terms obtained from derivatives of the annihilation and creation operators are the ones that contribute,

$$\langle \Phi_i^a | \Phi_j^{b[x]} \rangle = \langle \Phi_{HF} | a_i^\dagger a_a \left( a_b^\dagger a_j | \Phi_{HF} \rangle \right)^{[x]} \quad \text{(A3)}$$

$$= -\delta_{ij} O_{ba}^{R[x]} - \delta_{ab} O_{ji}^{R[x]}. \quad \text{(A4)}$$

Inserting this result into Eq. (A2) yields

$$\langle \Psi_I | \Psi_J^{[x]} \rangle = \sum_{ia} t_i^{Ia} t_i^{Ja[x]} - \sum_{iab} t_i^{Ia} t_i^{Jb} O_{ba}^{R[x]} - \sum_{ija} t_i^{Ia} t_j^{Ja} O_{ji}^{R[x]}. \quad \text{(A5)}$$

Equation (A5) is deceptively simple. To evaluate the CIS-amplitude derivatives directly requires the solution of the $3N$ coupled-perturbed CIS (CPCIS) equations,[55] a daunting task which we prefer to avoid. We also see that evaluating the latter

two terms requires an explicit treatment of intra-subspace orbital response, which is related to the redundant rotations of MR-CI theory. The corresponding rotation angles are unimportant in the context of Hartree-Fock, but there is no reason to believe that their derivatives should vanish. Fortunately, we may recast our expression so that we need not concern ourselves with CPCIS *or* these response terms.

### 1. Eliminating the amplitude derivatives

We begin by taking the derivative of Eqs. (8)–(11), obtaining

$$0 = \sum_{ijab} t_i^{Ia[x]} A_{iajb} t_j^{Jb} + \sum_{ijab} t_i^{Ia} A_{iajb}^{[x]} t_j^{Jb} + \sum_{ijab} t_i^{Ia} A_{iajb} t_j^{Jb[x]} \quad \text{(A6)}$$

$$= \sum_{ia} t_i^{Ia[x]} E_J t_i^{Ja} + \sum_{jb} t_j^{Ib} E_I t_j^{Jb[x]} + \sum_{ijab} t_i^{Ia} A_{iajb}^{[x]} t_j^{Jb} \quad \text{(A7)}$$

$$= (E_I - E_J) \sum_{ia} t_i^{Ia} t_i^{Ja[x]} + \sum_{ijab} t_i^{Ia} A_{iajb}^{[x]} t_j^{Jb}. \quad \text{(A8)}$$

We rearrange this result to isolate the amplitude-derivative dot product— simply rewriting the CI term—and plug it into Eq. (A5), yielding

$$\langle \Psi_I | \Psi_J^{[x]} \rangle = \frac{1}{E_J - E_I} \sum_{ijab} t_i^{Ia} A_{iajb}^{[x]} t_j^{Jb} - \sum_{iab} t_i^{Ia} t_i^{Jb} O_{ba}^{R[x]}$$
$$- \sum_{ija} t_i^{Ia} t_j^{Ja} O_{ji}^{R[x]}. \quad \text{(A9)}$$

Tedious (but straightforward) manipulations show that the tensor **A** and its derivative are[81]

$$A_{iajb} = \Pi_{ajib} + \delta_{ij} F_{ab} - \delta_{ab} F_{ij} + \delta_{ij} \delta_{ab} E \quad \text{(A10a)}$$

$$= \Pi_{ajib} + \delta_{ij} \delta_{ab} (\varepsilon_a - \varepsilon_i + E), \quad \text{(A10b)}$$

$$A_{iajb}^{[x]} = \Pi_{ajib}^{[x]} + \delta_{ij} F_{ab}^{[x]} - \delta_{ab} F_{ij}^{[x]} + \delta_{ij} \delta_{ab} E^{[x]}. \quad \text{(A10c)}$$

Inserting $A_{iajb}^{[x]}$ into Eq. (A9), we obtain

$$\langle \Psi_I | \Psi_J^{[x]} \rangle = \frac{1}{E_J - E_I} \sum_{ijab} t_i^{Ia} t_j^{Jb} \Pi_{ajib}^{[x]} + \frac{1}{E_J - E_I} \sum_{iab} t_i^{Ia} t_i^{Jb} F_{ab}^{[x]}$$

$$- \frac{1}{E_J - E_I} \sum_{ija} t_i^{Ia} t_j^{Ja} F_{ij}^{[x]} + \frac{1}{E_J - E_I} \sum_{ia} t_i^{Ia} t_i^{Ja} E^{[x]} - \sum_{iab} t_i^{Ia} t_i^{Jb} O_{ba}^{R[x]} - \sum_{ija} t_i^{Ia} t_j^{Ja} O_{ji}^{R[x]} \quad \text{(A11)}$$

$$= \frac{1}{E_J - E_I} \sum_{ijab} t_i^{Ia} t_j^{Jb} \Pi_{ajib}^{[x]} + \sum_{iab} t_i^{Ia} t_i^{Jb} \left( \frac{1}{E_J - E_I} F_{ab}^{[x]} - O_{ba}^{R[x]} \right) - \sum_{ija} t_i^{Ia} t_j^{Ja} \left( \frac{1}{E_J - E_I} F_{ij}^{[x]} + O_{ji}^{R[x]} \right). \quad \text{(A12)}$$

Note that we have begun to mix the separate CI and CSF terms of MR-CI theory! To proceed further, we must compute the derivatives of the arbitrary off-diagonal Fock matrix element $F_{pq}$ and the two-electron integral $\Pi_{ajib}$. We need not treat the Hartree-Fock gradient in Eq. (A10) explicitly, however; the orthogonality of CIS amplitudes ensures that its contribution to the derivative coupling vanishes.

Given the definition of the Fock matrix,

$$F_{pq} \equiv h_{pq} + \sum_m \Pi_{pmqm}, \tag{A13}$$

we can use Eqs. (47)–(52) to find the atomic-orbital representation of its derivative,

$$F_{pq}^{[x]} = \sum_{\mu\nu} C_{\mu p}^{[x]} h_{\mu\nu} C_{\nu q} + \sum_{\mu\nu} C_{\mu p} h_{\mu\nu}^{[x]} C_{\nu q} + \sum_{\mu\nu} C_{\mu p} h_{\mu\nu} C_{\nu q}^{[x]}$$

$$+ \sum_{m\mu\nu\lambda\sigma} \Pi_{\mu\nu\lambda\sigma} \begin{bmatrix} C_{\mu p}^{[x]} C_{\nu m} C_{\lambda q} C_{\sigma m} + C_{\mu p} C_{\nu m}^{[x]} C_{\lambda q} C_{\sigma m} \\ + C_{\mu p} C_{\nu m} C_{\lambda q}^{[x]} C_{\sigma m} + C_{\mu p} C_{\nu m} C_{\lambda q} C_{\sigma m}^{[x]} \end{bmatrix}$$

$$+ \sum_{m\mu\nu\lambda\sigma} C_{\mu p} C_{\nu m} \Pi_{\mu\nu\lambda\sigma}^{[x]} C_{\lambda q} C_{\sigma m}. \tag{A14}$$

Applying Eq. (63) and collecting terms yields the result

$$F_{pq}^{[x]} = \sum_{\mu\nu} C_{\mu p} h_{\mu\nu}^{[x]} C_{\nu q} + \sum_{m\mu\nu\lambda\sigma} C_{\mu p} C_{\nu m} \Pi_{\mu\nu\lambda\sigma}^{[x]} C_{\lambda q} C_{\sigma m}$$

$$-\frac{1}{2}\varepsilon_p \sum_{\alpha\beta} C_{\alpha p} S_{\alpha\beta}^{[x]} C_{\beta q} - \frac{1}{2}\varepsilon_q \sum_{\alpha\beta} C_{\alpha q} S_{\alpha\beta}^{[x]} C_{\beta p}$$

$$-\frac{1}{2} \sum_{mw\alpha\beta} C_{\alpha w} S_{\alpha\beta}^{[x]} C_{\beta m} (\Pi_{pwqm} + \Pi_{pmqw})$$

$$+(\varepsilon_q - \varepsilon_p)\Theta_{pq}^{[x]} - \sum_{mc} \Theta_{cm}^{[x]} (\Pi_{pcqm} + \Pi_{pmqc}). \tag{A15}$$

The same procedure applied to $\Pi_{ajib}$ leads to the expression

$$\Pi_{ajib}^{[x]} = \sum_{\mu\nu\lambda\sigma} C_{\mu a} C_{\nu j} \Pi_{\mu\nu\lambda\sigma}^{[x]} C_{\lambda i} C_{\sigma b}$$

$$-\frac{1}{2} \sum_{w\alpha\beta} C_{\alpha w} S_{\alpha\beta}^{[x]} [C_{\beta a} \Pi_{wjib} + C_{\beta j} \Pi_{awib} + C_{\beta i} \Pi_{ajwb} + C_{\beta b} \Pi_{ajiw}]$$

$$+ \sum_s \left[ \Pi_{sjib} \Theta_{as}^{[x]} + \Pi_{asib} \Theta_{js}^{[x]} + \Pi_{ajsb} \Theta_{is}^{[x]} + \Pi_{ajis} \Theta_{bs}^{[x]} \right]. \tag{A16}$$

## 2. Eliminating intra-subspace response

We now eliminate the intra-subspace (or, in the language of MR-CI, redundant) response terms that appear in the above expressions. If we isolate the terms involving mixing between virtuals (such as $O_{ba}^{R[x]}$) in Eq. (A12)—being careful to include the appropriate portions of the sums over $s$ from Eq. (A16)—we obtain

*VV* contribution

$$= \frac{1}{E_J - E_I} \sum_{ijabc} t_i^{Ia} t_j^{Jb} \left[ \Pi_{cjib} \Theta_{ac}^{[x]} + \Pi_{ajic} \Theta_{bc}^{[x]} \right]$$

$$+ \sum_{iab} t_i^{Ia} t_i^{Jb} \left[ \frac{1}{E_J - E_I} (\varepsilon_b - \varepsilon_a) \Theta_{ab}^{[x]} + \Theta_{ba}^{[x]} \right]. \tag{A17}$$

Combining Eqs. (8)–(11) with Eq. (A10), we find that

$$\sum_{jb} \Pi_{ajib} t_j^{Ib} = (E_I + \varepsilon_i - \varepsilon_a - E) t_i^{Ia}. \tag{A18}$$

Applying this relationship to Eq. (A17), noting that $\Theta_{pq}^{[x]} = -\Theta_{qp}^{[x]}$, and relabeling indices where appropriate, we obtain

*VV* contribution

$$= \frac{1}{E_J - E_I} \sum_{iac} t_i^{Ia} t_i^{Jc} (E_J + \varepsilon_i - \varepsilon_c - E) \Theta_{ac}^{[x]} \tag{A19}$$

$$+ \frac{1}{E_J - E_I} \sum_{jbc} t_j^{Ic} t_j^{Jb} (E_I + \varepsilon_j - \varepsilon_c - E) \Theta_{bc}^{[x]}$$

$$+ \sum_{iab} t_i^{Ia} t_i^{Jb} \left[ \frac{1}{E_J - E_I} (\varepsilon_b - \varepsilon_a) \Theta_{ab}^{[x]} + \Theta_{ba}^{[x]} \right],$$

$$= 0. \tag{A20}$$

The mixing within the occupied subspace vanishes in analogous fashion.

## 3. "Direct" CIS derivative couplings

The intra-subspace (redundant) response terms having been eliminated, the full "direct" CIS derivative-coupling expression in the MO representation is

$$
\langle\Psi_I|\Psi_J^{[x]}\rangle = \frac{1}{E_J - E_I}\sum_{ijab} t_i^{Ia} t_j^{Jb}\left\{ \sum_{\mu\nu\lambda\sigma} C_{\mu a} C_{\nu j}\Pi_{\mu\nu\lambda\sigma}^{[x]} C_{\lambda i} C_{\sigma b} - \frac{1}{2}\sum_{w\alpha\beta} C_{\alpha w} S_{\alpha\beta}^{[x]}\begin{bmatrix} C_{\beta a}\Pi_{wjib} + C_{\beta j}\Pi_{awib} \\[6pt] + C_{\beta i}\Pi_{ajwb} + C_{\beta b}\Pi_{ajiw} \end{bmatrix} \right.
$$

$$
\left. + \sum_k \Pi_{kjib}\Theta_{ak}^{[x]} + \sum_c \Pi_{acib}\Theta_{jc}^{[x]} + \sum_c \Pi_{ajcb}\Theta_{ic}^{[x]} + \sum_k \Pi_{ajik}\Theta_{bk}^{[x]} \right\}
$$

$$
+ \sum_{iab} t_i^{Ia} t_i^{Jb}\left\{ \frac{1}{E_J - E_I}\left[ \sum_{\mu\nu} C_{\mu a} h_{\mu\nu}^{[x]} C_{\nu b} + \sum_{m\mu\nu\lambda\sigma} C_{\mu a} C_{\nu m}\Pi_{\mu\nu\lambda\sigma}^{[x]} C_{\lambda b} C_{\sigma m} - \frac{1}{2}\varepsilon_a\sum_{\alpha\beta} C_{\alpha a} S_{\alpha\beta}^{[x]} C_{\beta b} \right.\right.
$$

$$
\left.\left. - \frac{1}{2}\varepsilon_b\sum_{\alpha\beta} C_{\alpha b} S_{\alpha\beta}^{[x]} C_{\beta a} - \frac{1}{2}\sum_{m w\alpha\beta} C_{\alpha w} S_{\alpha\beta}^{[x]} C_{\beta m}(\Pi_{awbm} + \Pi_{ambw}) - \sum_{mc}\Theta_{cm}^{[x]}(\Pi_{acbm} + \Pi_{ambc}) \right] - \sum_{\mu\nu} C_{\mu b} S_{\mu\nu}^{A[x]} C_{\nu a} \right\}
$$

$$
- \sum_{ija} t_i^{Ia} t_j^{Ja}\left\{ \frac{1}{E_J - E_I}\left[ \sum_{\mu\nu} C_{\mu i} h_{\mu\nu}^{[x]} C_{\nu j} + \sum_{m\mu\nu\lambda\sigma} C_{\mu i} C_{\nu m}\Pi_{\mu\nu\lambda\sigma}^{[x]} C_{\lambda j} C_{\sigma m} - \frac{1}{2}\varepsilon_i\sum_{\alpha\beta} C_{\alpha i} S_{\alpha\beta}^{[x]} C_{\beta j} \right.\right.
$$

$$
\left.\left. - \frac{1}{2}\varepsilon_j\sum_{\alpha\beta} C_{\alpha j} S_{\alpha\beta}^{[x]} C_{\beta i} - \frac{1}{2}\sum_{m w\alpha\beta} C_{\alpha w} S_{\alpha\beta}^{[x]} C_{\beta m}\left(\Pi_{iwjm} + \Pi_{imjw}\right) - \sum_{mc}\Theta_{cm}^{[x]}\left(\Pi_{icjm} + \Pi_{imjc}\right) \right] + \sum_{\mu\nu} C_{\mu j} S_{\mu\nu}^{A[x]} C_{\nu i} \right\}.
$$

$$(A21)$$

Converting the various two-electron integrals into the AO representation, combining terms wherever possible, relabeling indices, and applying the notation of Eq. (59), we arrive at the equivalent expression

$$
\langle\Psi_I|\Psi_J^{[x]}\rangle = \frac{1}{E_J - E_I}\left\{ \sum_{\mu\nu} B_{\mu\nu}^{IJ} h_{\mu\nu}^{[x]} + \sum_{\mu\nu\lambda\sigma}\left( R_{\mu\lambda}^I R_{\sigma\nu}^J + B_{\mu\lambda}^{IJ} P_{\sigma\nu} \right)\Pi_{\mu\nu\lambda\sigma}^{[x]} + \frac{1}{2}\sum_{ija\alpha\beta} t_i^{Ia} t_j^{Ja} S_{\alpha\beta}^{[x]}\left( \varepsilon_i C_{\alpha i} C_{\beta j} + \varepsilon_j C_{\alpha j} C_{\beta i} \right) \right.
$$

$$
- \frac{1}{2}\sum_{iab\alpha\beta} t_i^{Ia} t_i^{Jb} S_{\alpha\beta}^{[x]}(\varepsilon_a C_{\alpha a} C_{\beta b} + \varepsilon_b C_{\alpha b} C_{\beta a}) - \frac{1}{2}\sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[x]}\widetilde{P}_{\mu\alpha}\left( R_{\nu\gamma}^I R_{\delta\beta}^J + R_{\delta\beta}^I R_{\nu\gamma}^J \right)\Pi_{\alpha\beta\gamma\delta}
$$

$$
\left. - \frac{1}{2}\sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[x]}\widetilde{P}_{\mu\alpha}\left( R_{\gamma\nu}^I R_{\beta\delta}^J + R_{\beta\delta}^I R_{\gamma\nu}^J \right)\Pi_{\alpha\beta\gamma\delta} + \frac{1}{2}\sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[x]}\widetilde{P}_{\mu\alpha} P_{\nu\delta}\left( B_{\gamma\beta}^{IJ} + B_{\beta\gamma}^{IJ} \right)\Pi_{\alpha\beta\gamma\delta} \right\}
$$

$$
- \frac{1}{E_J - E_I}\sum_{bi} Y_{bi}\Theta_{bi}^{[x]} - \sum_{\mu\nu} S_{\mu\nu}^{A[x]}\left( \sum_{iab} C_{\nu a} t_i^{Ia} t_i^{Jb} C_{\mu b} + \sum_{ija} C_{\nu i} t_i^{Ia} t_j^{Ja} C_{\mu j} \right).
$$

$$(A22)$$

## 4. Reconciling the theories

Once the **z**-vector method for the orbital response is applied, Eq. (A22) is *almost* identical to the theory set forth in Eqs. (81)–(83). Because no approximations were made in either derivation; however, we know that they must be *completely* identical. We need to reconcile the two theories.

One ostensible difference between them that disappears under closer inspection is in the terms in Eq. (A22) involving the orbital energies $\varepsilon_p$. Substituting $\varepsilon_p = C_{\mu p} F_{\mu\nu} C_{\nu p}$ allows us to rewrite Eq. (A22) in a form almost identical to that of Eqs. (81)–(83),

$$\langle \Psi_I | \Psi_J^{[x]} \rangle = \frac{1}{E_J - E_I} \left\{ \sum_{\mu\nu} B_{\mu\nu}^{IJ} h_{\mu\nu}^{[x]} + \sum_{\mu\nu\lambda\sigma} \left( R_{\mu\lambda}^I R_{\sigma\nu}^J + B_{\mu\lambda}^{IJ} P_{\sigma\nu} \right) \Pi_{\mu\nu\lambda\sigma}^{[x]} - \frac{1}{2} \sum_{\mu\nu\alpha\beta} S_{\mu\nu}^{[x]} \widetilde{P}_{\mu\alpha} \left( B_{\beta\nu}^{IJ} + B_{\nu\beta}^{IJ} \right) F_{\alpha\beta} \right.$$

$$- \frac{1}{2} \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[x]} \widetilde{P}_{\mu\alpha} \left( R_{\nu\gamma}^I R_{\delta\beta}^J + R_{\delta\beta}^I R_{\nu\gamma}^J \right) \Pi_{\alpha\beta\gamma\delta} - \frac{1}{2} \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[x]} \widetilde{P}_{\mu\alpha} \left( R_{\gamma\nu}^I R_{\beta\delta}^J + R_{\beta\delta}^I R_{\gamma\nu}^J \right) \Pi_{\alpha\beta\gamma\delta}$$

$$\left. + \frac{1}{2} \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[x]} \widetilde{P}_{\mu\alpha} P_{\nu\delta} \left( B_{\gamma\beta}^{IJ} + B_{\beta\gamma}^{IJ} \right) \Pi_{\alpha\beta\gamma\delta} \right\} - \frac{1}{E_J - E_I} \sum_{bi} Y_{bi} \Theta_{bi}^{[x]}$$

$$- \sum_{\mu\nu} S_{\mu\nu}^{A[x]} \left( \sum_{iab} C_{\nu a} t_i^{Ia} t_i^{Jb} C_{\mu b} + \sum_{ija} C_{\nu i} t_i^{Ia} t_j^{Ja} C_{\mu j} \right). \tag{A23}$$

To show that the expression above is *totally* identical to that in Eqs. (81)–(83)—and to the MR-CI theory, now that there is a clean separation between the orbital-CSF terms containing $\mathbf{S}^{A[x]}$ and everything else—we decompose $\mathbf{S}^{R[x]}$ into its symmetric and antisymmetric components wherever it appears (using Eq. (70)). That accomplished, we find that the terms involving $\mathbf{S}^{[x]}$ are identical. Our task is therefore reduced to proving that the portions of Eqs. (81)–(83) containing $\mathbf{S}^{A[x]}$,

$$\frac{1}{E_J - E_I} \sum_{\mu\nu} S_{\mu\nu}^{A[x]} \left\{ - \sum_{\alpha\beta} \widetilde{P}_{\mu\alpha} \left( B_{\beta\nu}^{IJ} + B_{\nu\beta}^{IJ} \right) F_{\alpha\beta} \right.$$

$$- \sum_{\alpha\beta\gamma\delta} \widetilde{P}_{\mu\alpha} \left( R_{\nu\gamma}^I R_{\delta\beta}^J + R_{\delta\beta}^I R_{\nu\gamma}^J \right) \Pi_{\alpha\beta\gamma\delta}$$

$$- \sum_{\alpha\beta\gamma\delta} \widetilde{P}_{\mu\alpha} \left( R_{\gamma\nu}^I R_{\beta\delta}^J + R_{\beta\delta}^I R_{\gamma\nu}^J \right) \Pi_{\alpha\beta\gamma\delta}$$

$$+ \sum_{\alpha\beta\gamma\delta} \widetilde{P}_{\mu\alpha} P_{\nu\delta} \left( B_{\gamma\beta}^{IJ} + B_{\beta\gamma}^{IJ} \right) \Pi_{\alpha\beta\gamma\delta}$$

$$\left. + C_{\mu b} Y_{bi} C_{\nu i} \right\}, \tag{A24}$$

are equal to those from the direct method in Eq. (A23),

$$- \sum_{\mu\nu} S_{\mu\nu}^{A[x]} \left( \sum_{iab} C_{\nu a} t_i^{Ia} t_i^{Jb} C_{\mu b} + \sum_{ija} C_{\nu i} t_i^{Ia} t_j^{Ja} C_{\mu j} \right). \tag{A25}$$

Demonstrating this equivalence, which does indeed hold, requires a significant amount of algebra; we refer the interested reader to the supplemental material.[82] The general approach is to manipulate the MO representation of Eq. (A24), relabeling indices to identify terms that cancel; to make liberal use of Eq. (A18); and to exploit the antisymmetry of $\mathbf{S}^{A[x]}$ and its implications for the trace.

[1] B. H. Lengsfield III and D. R. Yarkony, Adv. Chem. Phys. **82**(2), 1 (1992).
[2] J. C. Tully, J. Chem. Phys. **93**, 1061 (1990).
[3] H.-D. Meyer and W. H. Miller, J. Chem. Phys. **70**, 3214 (1979).
[4] G. Stock and M. Thoss, Phys. Rev. Lett. **78**, 578 (1997).
[5] J. C. Tully, Faraday Discuss. **110**, 407 (1998).
[6] M. Ben-Nun, J. Quenneville, and T. J. Martínez, J. Phys. Chem. A **104**, 5161 (2000).
[7] D. R. Yarkony, in *Conical Intersections: Electronic Structure, Dynamics and Spectroscopy*, edited by W. Domcke, D. R. Yarkony, and H. Köppel, (World Scientific, Singapore, 2004), pp. 41–128.
[8] M. Baer, *Beyond Born-Oppenheimer: Nonadiabatic Coupling Terms and Conical Intersections* (Wiley, Hoboken, NJ, 2006).
[9] M. Baer, Chem. Phys. Lett. **35**, 112 (1975).
[10] G. J. Halász, Á. Vibók, and M. Baer, J. Chem. Phys. **127**, 144108 (2007).
[11] B. H. Lengsfield III, P. Saxe, and D. R. Yarkony, J. Chem. Phys. **81**, 4549 (1984).
[12] P. Saxe, B. H. Lengsfield III, and D. R. Yarkony, Chem. Phys. Lett. **113**, 159 (1985).
[13] H. Lischka, M. Dallos, P. G. Szalay, D. R. Yarkony, and R. Shepard, J. Chem. Phys. **120**, 7322 (2004).
[14] B. G. Levine, C. Ko, J. Quenneville, and T. J. Martínez, Mol. Phys. **104**, 1039 (2006).
[15] B. G. Levine, J. D. Coe, and T. J. Martínez, J. Phys. Chem. B **112**, 405 (2008).
[16] P. Jørgensen and J. Simons, J. Chem. Phys. **79**, 334 (1983).
[17] T. Helgaker and J. Almlöf, Int. J. Quantum Chem. **26**, 275 (1984).
[18] M. Page, P. Saxe, G. F. Adams, and B. H. Lengsfield III, J. Chem. Phys. **81**, 434 (1984).
[19] R. Shepard, Int. J. Quantum Chem. **31**, 33 (1987).
[20] R. Shepard, H. Lischka, P. G. Szalay, T. Kovar, and M. Ernzerhof, J. Chem. Phys. **96**, 2085 (1991).
[21] H. Lischka, M. Dallos, and R. Shepard, Mol. Phys. **100**, 1647 (2002).
[22] M. Dallos, H. Lischka, R. Shepard, D. R. Yarkony, and P. G. Szalay, J. Chem. Phys. **120**, 7330 (2004).
[23] M. Barbatti and H. Lischka, J. Am. Chem. Soc. **130**, 6831 (2008).
[24] J. J. Szymczak, M. Barbatti, J. T. Soo Hoo, J. A. Adkins, T. L. Windus, D. Nachtigallová, and H. Lischka, J. Phys. Chem. A **113**, 12686 (2009).
[25] H. Lischka, T. Müller, P. G. Szalay, I. Shavitt, R. M. Pitzer, and R. Shepard, WIREs Comput. Mol. Sci. **1**, 191 (2011).
[26] T. Ichino, J. Gauss, and J. F. Stanton, J. Chem. Phys. **130**, 174015 (2009).
[27] A. Tajti and P. G. Szalay, J. Chem. Phys. **131**, 124104 (2009).
[28] V. Chernyak and S. Mukamel, J. Chem. Phys. **112**, 3572 (2000).
[29] M. Tommasini, V. Chernyak, and S. Mukamel, Int. J. Quantum Chem. **85**, 225 (2001).
[30] P. Pulay, Mol. Phys. **17**, 197 (1969).
[31] R. Baer, Chem. Phys. Lett. **364**, 75 (2002).
[32] C. Hu, H. Hirai, and O. Sugino, J. Chem. Phys. **127**, 064103 (2007).
[33] C. Hu, H. Hirai, and O. Sugino, J. Chem. Phys. **128**, 154111 (2008).
[34] C. Hu, O. Sugino, and Y. Tateyama, J. Chem. Phys. **131**, 114101 (2009).
[35] C. Hu, O. Sugino, H. Hirai, and Y. Tateyama, Phys. Rev. A **82**, 062508 (2010).
[36] E. Tapavicza, I. Tavernelli, and U. Röthlisberger, Phys. Rev. Lett. **98**, 023001 (2007).
[37] I. Tavernelli, E. Tapavicza, and U. Röthllisberger, J. Chem. Phys. **130**, 124107 (2009).
[38] I. Tavernelli, B. F. E. Curchod, and U. Röthlisberger, J. Chem. Phys. **131**, 196101 (2009).
[39] I. Tavernelli, B. F. E. Curchod, A. Laktionov, and U. Röthlisberger, J. Chem. Phys. **133**, 194104 (2010).
[40] R. Send and F. Furche, J. Chem. Phys. **132**, 044107 (2010).

[41]T. Nelson, S. Fernandez-Alberti, V. Chernyak, A. E. Roitberg, and S. Tretiak, J. Phys. Chem. B **115**, 5402 (2011).

[42]A. Dreuw and M. Head-Gordon, Chem. Rev. **105**, 4009 (2005).

[43]D. R. Yarkony, J. Chem. Phys. **84**, 3206 (1986).

[44]W. Kutzelnigg, Mol. Phys. **105**, 2627 (2007).

[45]D. R. Bates and R. McCarroll, Proc. R. Soc. A **245**, 175 (1958).

[46]S. B. Schneiderman and A. Russek, Phys. Rev. **181**, 311 (1969).

[47]W. R. Thorson and J. B. Delos, Phys. Rev. A **18**, 117 (1978).

[48]J. B. Delos, Rev. Mod. Phys. **53**, 287 (1981).

[49]T. G. Winter, Phys. Rev. A **25**, 697 (1982).

[50]L. F. Errea, C. Harel, H. Jouin, L. Méndez, B. Pons, and A. Riera, J. Phys. B **27**, 3603 (1994).

[51]E. Deumens, A. Diz, R. Longo, and Y. Öhrn, Rev. Mod. Phys. **66**, 917 (1994).

[52]C. Illescas and A. Riera, Phys. Rev. Lett. **80**, 3029 (1998).

[53]A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover, New York, 1996).

[54]P. Jørgensen and J. Simons, *Second Quantization-Based Methods in Quantum Chemistry* (Academic, New York, 1981).

[55]D. Maurice, "Single electron theories of excited states," Ph.D. dissertation, (University of California, Berkeley, 1998).

[56]J. Gerratt and I. M. Mills, J. Chem. Phys. **49**, 1719 (1968).

[57]J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem., Symp. **13**, 225 (1979).

[58]N. C. Handy and H. F. Schaefer III, J. Chem. Phys. **81**, 5031 (1984).

[59]P. Pulay, in *Modern Electronic Structure Theory, Part II*, edited by D. R. Yarkony (World Scientific, Singapore, 1995), pp. 1191–1240.

[60]Y. Yamaguchi, Y. Osamura, J. D. Goddard, and H. F. Schaefer III, "A new dimension to quantum chemistry: analytical derivative methods," in *Ab-Initio Molecular Electronic Structure Theory* (Oxford University Press, New York, 1994).

[61]Y. Shao, L. Fusti-Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. Distasio, Jr., R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khaliullin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock III, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer III, J. Kong, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, Phys. Chem. Chem. Phys. **8**, 3172 (2006).

[62]M. Stockburger, H. Gattermann, and W. Klusmann, J. Chem. Phys. **63**, 4529 (1975).

[63]J. R. Reimers and N. S. Hush, Chem. Phys. **146**, 105 (1990).

[64]F. Negri and M. Z. Zgierski, J. Chem. Phys. **104**, 3486 (1996).

[65]R. Montero, A. Longarte, R. Martínez, M. N. Sánchez Rayo, and F. Castaño, Chem. Phys. Lett. **468**, 134 (2009).

[66]N. Shenvi, J. E. Subotnik, and W. Yang, J. Chem. Phys. **135**, 024101 (2011).

[67]R. Bast, U. Ekström, B. Gao, T. Helgaker, K. Ruud, and A. J. Thorvaldsen, Phys. Chem. Chem. Phys. **13**, 2627 (2011).

[68]F. London, J. Phys. Radium **8**, 397 (1937).

[69]J. A. Pople, J. Chem. Phys. **37**, 53 (1962).

[70]R. Ditchfield, Mol. Phys. **27**, 789 (1974).

[71]K. Wolinski, J. F. Hinton, and P. Pulay, J. Am. Chem. Soc. **112**, 8251 (1990).

[72]T. Helgaker and P. Jørgensen, J. Chem. Phys. **95**, 2595 (1991).

[73]K. Ruud, T. Helgaker, K. L. Bak, P. Jørgensen, and H. J. Aa. Jensen, J. Chem. Phys. **99**, 3847 (1993).

[74]A. M. Lee, N. C. Handy, and S. M. Colwell, J. Chem. Phys. **103**, 10095 (1995).

[75]J. Jusélius, D. Sundholm, and J. Gauss, J. Chem. Phys. **121**, 3952 (2004).

[76]P. Schmelcher and L. S. Cederbaum, Phys. Rev. A **37**, 672 (1988).

[77]P.-O. Löwdin, J. Chem. Phys. **18**, 365 (1950).

[78]P.-O. Löwdin, Adv. Phys. **5**, 1 (1956).

[79]B. C. Carlson and J. K. Keller, Phys. Rev. **105**, 102 (1957).

[80]I. Mayer, Int. J. Quantum Chem. **90**, 63 (2002).

[81]D. Maurice and M. Head-Gordon, Int. J. Quantum Chem., Symp. **29**, 361 (1995).

[82]See supplementary material at http://dx.doi.org/10.1063/1.3665031 for a full discussion of the transformations that complete the proof of the equivalence of Eqs. (A24) and (A25).