Derivative couplings and analytic gradients for diabatic states, with an implementation for Boys-localized configuration-interaction singles

Shervin Fatehi, Ethan Alguire, and Joseph E. Subotnik

Derivative couplings and analytic gradients for diabatic states, with an implementation for Boys-localized configuration-interaction singles

Shervin Fatehi, Ethan Alguire, and Joseph E. Subotnik[a)]
Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, Pennsylvania 19104, USA

We demonstrate that Boys-localized diabatic states do indeed exhibit small derivative couplings, as is required of quasidiabatic states. In doing so, we present a general formalism for calculating derivative couplings and analytic gradients for diabatic states. We then develop additional equations specific to the case of Boys-localized configuration-interaction singles (CIS)—in particular, the analytic gradient of the CIS dipole matrix—and we validate our implementation against finite-difference results. In a forthcoming paper, we will publish additional algorithmic and computational details and apply our method to the Closs energy-transfer systems as a further test of the validity of Boys-localized diabatic states. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4820485]

I. INTRODUCTION

A. Exact diabatic states for use in dynamical theories

In quantum dynamics, a vast and rich literature exists describing the flow of energy between electronic and nuclear degrees of freedom in model Hamiltonians with fixed diabatic states. The most famous dynamical models are undoubtedly Marcus theory[1] and Redfield theory.[2] On the one hand, Marcus theory describes a simple model Hamiltonian for a pair of fictitious, weakly-coupled diabatic electronic states, each of which is in turn coupled strongly and linearly to a bath of harmonic oscillators. On the other hand, Redfield theory is more general, capturing the electronic dynamics of a system of multiple diabatic electronic states, each coupled weakly to an arbitrary bath (usually chosen as harmonic). Marcus and Redfield theory both begin with a set of fixed and immutable diabatic states and then invoke Fermi's golden rule to determine the dynamical implications of their respective Hamiltonians.

From this perspective, the adiabatic representation is not privileged over a diabatic representation. Indeed, within Marcus theory, the adiabatic states are never observable and only serve as a distraction. Admittedly, there are some specialized semiclassical nonadiabatic-dynamics algorithms that require adiabatic representations to function in practice, e.g., surface-hopping dynamics.[3] However, a truly rigorous theory of nonadiabatic dynamics must be independent of representation, e.g., mixed quantum-classical Liouville dynamics,[4] fully-converged multiple spawning,[5] and the semiclassical implementation of the Meyer–Miller–Stock–Thoss method.[6] Moreover, dynamicists often prefer a diabatic basis as a convenience:[7] Diabatic potential-energy surfaces are as smooth as adiabatic surfaces, but diabatic couplings are much smoother than derivative couplings. In a best-case scenario, a true dynamicist would prefer to work with parameterized diabatic surfaces so as to focus computational resources

on capturing nuclear motion over a long timescale, rather than on computing expensive electronic energies and forces many times along short trajectories.

To sum up: Problems in nonadiabatic dynamics can often be posed and addressed most naturally using a simple model Hamiltonian with a fixed set of diabatic states, $\{|\Xi_A\rangle\}$, with derivative couplings vanishing by hypothesis,

$$\nabla_{\mathbf{R}}|\Xi_A\rangle = 0 \quad \Rightarrow \quad \mathbf{d}_{AB}^{[\mathbf{R}]} \equiv \langle \Xi_A|\nabla_{\mathbf{R}}|\Xi_B\rangle = 0. \quad (1)$$

There are enough difficulties inherent in simulating and analyzing dynamics that a practicing theorist does not want to worry about computing diabatic potential-energy surfaces and coupling parameters.

B. Generating adiabatic states for real molecules from quantum chemistry

The motivation above notwithstanding, diabatic states are not the fundamental starting point for modern quantum chemistry as practiced by electronic-structure theorists.[8] Rather, the basic mantra is the Born–Oppenheimer approximation, which assumes that all electronic motion occurs on a timescale over which the nuclei are effectively frozen. Thus, to generate potential-energy surfaces, the common approach is to diagonalize the electronic Hamiltonian at each given set of nuclear coordinates. The result is that one always computes adiabatic states *ab initio*.

From the perspective of quantum chemistry, the full Hamiltonian is composed of the nuclear kinetic energy added to an electronic Hamiltonian,

$$H(\mathbf{r}, \mathbf{R}) = T_{\mathrm{N}}(\mathbf{R}) + H_{\mathrm{el}}(\mathbf{r}; \mathbf{R}). \quad (2)$$

The adiabatic states are simply defined as those electronic states which diagonalize the electronic Hamiltonian and depend parametrically on the nuclear position $\mathbf{R}$,

$$H_{\mathrm{el}}(\mathbf{r}; \mathbf{R})|\Psi_I(\mathbf{R})\rangle = E_I(\mathbf{R})|\Psi_I(\mathbf{R})\rangle. \quad (3)$$

[a)]Electronic mail: subotnik@sas.upenn.edu

It is clear that the nature of the adiabatic states can change dramatically with the nuclear coordinate. To first order, electronic relaxation is mediated by so-called derivative (or nonadiabatic) couplings between adiabatic states; these couplings are induced by nuclear motion,

$$\mathbf{d}_{IJ}^{[\mathbf{R}]} = \langle \Psi_I(\mathbf{R})|\nabla_{\mathbf{R}}|\Psi_J(\mathbf{R})\rangle = \int_{-\infty}^{\infty} \Psi_I(\mathbf{r};\mathbf{R})\nabla_{\mathbf{R}}\Psi_J(\mathbf{r};\mathbf{R})\,d\mathbf{r}. \quad (4)$$

Following Tully's treatment of the time-dependent Schrödinger equation,[3] the electronic wavefunction can be propagated in time with the nuclear trajectory $\mathbf{R}(t)$ treated as a parameter. All electronic relaxation is reflected in a rate of transition proportional to the nuclear velocity $\mathbf{V}(t) \equiv \dot{\mathbf{R}}(t)$,

$$\frac{1}{\tau_{\text{hop}}^{I \to J}} \propto \mathbf{d}_{IJ}^{[\mathbf{R}(t)]} \cdot \mathbf{V}(t). \quad (5)$$

In the end, standard electronic-structure theory addresses nonadiabatic dynamics by computing accurate adiabatic potential-energy surfaces for the ground and excited states, $E_I(\mathbf{R})$, and the derivative couplings between those states, $\mathbf{d}_{IJ}^{[\mathbf{R}]}$.

C. Generating approximate diabatic states from real adiabatic states

Given the need for diabatic states in quantum dynamics and the ease of generating adiabatic states from electronic-structure theory, many approaches have been proposed over the last 50 years for constructing approximate diabatic states from adiabatic states. Early approaches by Baer[9] and Truhlar[10] focused on small molecules and sought to obtain diabatic states with rigorously small derivative couplings, $\mathbf{d}_{AB}^{[\mathbf{R}]} \approx 0$. More recently, Yarkony[11,12] has focused on methods to fit diabatic states with minimal derivative couplings for molecules with a handful of atoms.

As interest in larger molecules has increased over the last two decades, there has been an explosion of new transformations[13] for constructing diabatic states. Using chemical intuition, Köppel, Cederbaum, and co-workers introduced[14] block-diagonalization of chemically relevant reference states; later, Atchity and Ruedenberg introduced configurational uniformity;[15,16] and Nakamura and Truhlar developed the four-fold way.[17–19] Focusing on physical observables, Cave and Newton wrote seminal papers[20,21] proposing the Generalized Mulliken–Hush (GMH) method of generating diabatic states with a localized charge distribution for describing electron transfer (as did Meyer and Werner[22]). These notions of localized diabatization were later taken up by Voityuk[23] (fragment-charge difference) and Hsu[24,25] (fragment-energy difference). Our group showed recently that GMH is really a limiting case of Boys localization,[26] and we have introduced both Boys- and Edmiston–Ruedenberg-localized[27] diabatic states as generalizations of orbital localization in traditional quantum chemistry.[28,29] We further showed that localized diabatization can be justified on the basis of different assumptions about system-solvent interactions.[27,30]

In many applications, the essential aim of localized-diabatization techniques is to generate the Hamiltonian matrix element between diabatic states, H_{AB}, also called the diabatic coupling. H_{AB} is the only inter-state matrix element in an exactly diabatic basis, and the Marcus rate of electron or energy transfer is proportional to $|H_{AB}|^2$. But localized diabatic states are *not* exactly diabatic states, and their derivative couplings are not exactly zero. Thus, localized diabatic states are coupled together by diabatic *and* derivative couplings, the latter being presumed small. To our knowledge, however, derivative couplings between localized diabatic states have never been computed, and only a few reports of estimates exist in the literature.[31,32] That such derivative couplings are small has never actually been verified.

With these facts in mind, the goal of this paper is to develop a rigorous method for computing derivative couplings between localized diabatic states. We will focus on the case of Boys-localized diabatic states in configuration-interaction-singles (CIS) theory, but our approach is quite general. Having computed the couplings, we will be able to verify or dismiss the validity and usefulness of localized diabatic representations. We will also show that our formalism can be used to obtain the analytic first derivative of any quantity in the diabatic representation, e.g., the gradient of the diabatic coupling $H_{AB}^{[Q]}$, which provides an estimate of non-Condon effects. In this paper, we elaborate the necessary theory. In an upcoming publication,[33] we will apply our methods to the Closs energy-transfer molecules, which constitute a canonical test set with experimentally-measured properties largely consistent with the predictions of Marcus theory.[34–38] In particular, we will assess for each molecule whether the diabatic derivative couplings are actually small along the reaction coordinate and whether the Condon approximation is violated, which bears directly on the accuracy of the Marcus energy-transfer rates previously computed by our group.[38]

D. Outline of the present work

The remainder of this paper is organized as follows: We conclude this section with an overview of our notational conventions and symbols. In Sec. II, we develop the general theory for derivative couplings between diabatic states and provide the relevant expressions for the Boys diabatization; we note that the same formalism can be used to obtain essentially any first-order analytic gradient involving diabatic states. Because gradients of the adiabatic dipole matrix are required in computing Boys-diabatic derivative couplings, we present in Sec. III an analytic expression for their calculation in the context of CIS. We validate our theory against finite-difference calculations for lithium hydride and make a simple application to p-benzoquinone in Sec. IV. Finally, we recapitulate our findings in Sec. V, concluding with the most immediate questions raised by our work. Respectively, Appendices A–E discuss the coupled-perturbed CIS (CPCIS) and coupled-perturbed Hartree–Fock (CPHF) equations for calculating amplitude- and orbital-response contributions to the adiabatic dipole-matrix gradient; provide expressions for the various constituents of the mixed-derivative matrix appearing in the

CPCIS equations; detail the manipulations leading to our final analytic equation for the dipole-matrix gradient; present the exact formula for overlap between nonorthogonal, singly-excited Slater determinants used in the finite-difference calculations of Sec. IV; and list Cartesian coordinates for the distorted p-benzoquinone system.

E. Notation and symbols

In the text to follow, we will adhere to a variety of notational conventions and necessarily introduce quite a few symbols. This section explains the basic choices we have made in this regard, while Table I lists symbols that will appear, together with brief descriptions and an index to associated defining equations.

Diabatic states will be represented by Ξ, with labels $\{A, B, C, D, \dots\}$. They are obtained by mixing the adiabatic states, which are given by Ψ and the labels $\{I, J, K, L, \dots\}$. (We will sometimes refer to these states as "diabats" and "adiabats.") Elements of the diabatization matrix $\mathbf{U}$ will be written with the diabatic-state index first, such that mixing of state I into state A is due to the element U_{AI}. M of the adiabatic states are selected for diabatization.

The system contains N_A atoms, corresponding to $3N_A$ degrees of freedom in Cartesian space. $\mathbf{R}$ is the full set of nuclear positions and also, by convention, the CIS transition density; context will make this distinction clear. $\mathbf{R}_Q$ is a specific nuclear position, and ∇_Q is the gradient with respect to this position. The superscript $[Q]$ is shorthand for the nuclear gradient of a quantity or a label for the derivative coupling induced by atom Q. We will use a superscript $[\mathbf{R}]$ to refer to the gradient over the entire configuration space.

When the adiabats are given explicit form, Slater determinants will be represented by Φ. Excitations will be labeled to indicate which occupied MOs have been promoted to virtuals, e.g., Φ_i^a for a singles determinant; the associated amplitudes are, e.g., t_i^a. The constituent molecular orbitals are represented by ϕ. We use spin orbitals everywhere, such that spin variables remain implicit.

The Hartree–Fock molecular-orbital basis comprises N_o spin orbitals, O occupied and V virtual. Occupied MOs will be labeled by $\{i, j, k, \ell, m, \dots\}$; virtual MOs by $\{a, b, c, d, \dots\}$; and generic (occupied *or* virtual) MOs by $\{p, q, s, t, \dots\}$. Atomic orbitals, represented by φ, will be indexed interchangeably with either $\{\mu, \nu, \lambda, \sigma\}$ or $\{\alpha, \beta, \gamma, \delta\}$. The coefficients characterizing a specific molecular orbital will be given by $C_{\mu p}$, with the AO index always listed first. We will use $\Pi_{\mu\nu\lambda\sigma} \equiv \langle\mu\nu||\lambda\sigma\rangle$ to represent antisymmetrized two-electron integrals in physicists' notation.

Atomic-orbital quantities, once introduced, will be used almost exclusively. In doing so, we will follow certain indexing conventions in order to keep our expressions as compact as possible. $\mathbf{P}'$ (formally equivalent to $\mathbf{S}^{-1}$) will always be indexed by $\mu\alpha$. Gradients of AO quantities will always be labeled from the set $\{\mu, \nu, \lambda, \sigma\}$; the orbital and amplitude response gradients $\boldsymbol{\Theta}^{[Q]}$ and $\mathbf{t}^{[Q]}$, which are naturally

TABLE I. Symbols appearing in this paper, their defining equations, and their meanings.

Eq.	Symbol	Meaning	Eq.	Symbol	Meaning
…	N_A	Number of atoms	…	ε	Molecular-orbital energy
…	N_o	Total number of spin orbitals	…	$a^\dagger/a$	Creation/annihilation operators
…	O	Number of occupied spin orbitals	(41)	$\boldsymbol{\mu}$	Dipole operator or matrix
…	V	Number of virtual spin orbitals	…	$\mathcal{O}$	Arbitrary operator
…	M	Number of states in diabatization	(31)	$\mathbf{A}$	CIS supermatrix
(30)	Ψ	Generic (or CIS) adiabatic state	(36)	ω	CIS excitation energy
(7)	Ξ	Generic (quasi-)diabatic state	(30)	$\mathbf{t}$	CIS excitation amplitudes
(1)	$\mathbf{d}^{[Q]}$	Derivative coupling due to atom Q	(37)	$\mathbf{R}$	CIS transition density
(13)	f	Boys diabatization functional	(37)	$\mathbf{B}$	CIS difference density
(7)	$\mathbf{U}$	Diabatization (rotation) matrix	(39)	$\mathbf{D}$	General (or inter-CIS) density
(10)	θ	Diabatic mixing angle	(A4)	$\mathcal{D}$	CPCIS supermatrix
(16)	$\mathcal{A}$	$\mathbf{U}^{[Q]}$ unitarity supermatrix	(A5)	$\overline{\mathbf{L}}$	CPCIS Lagrangian for $\boldsymbol{\mu}^{[Q]}$
(18)	$\mathcal{B}$	$\mathbf{U}^{[Q]}$ constraint supermatrix	(A4)	$\overline{\mathbf{M}}^{[Q]}$	CPCIS mixed-derivative matrix
(17)	$\mathcal{C}^{[Q]}$	$\mathbf{U}^{[Q]}$ constraint matrix	(A9)	$\overline{\mathbf{z}}$	CPCIS z-vector for $\boldsymbol{\mu}^{[Q]}$
…	Φ	Slater determinant	(C3)	η	CPCIS scaling vector
…	ϕ	Molecular spin orbital	(C3)	ζ	CPCIS relaxer of attachment
…	φ	Atomic orbital	(C3)	ζ'	CPCIS relaxer of detachment
…	$\mathbf{H}$	Electronic Hamiltonian	(C8)	$\overline{\mathbf{t}}$	CPCIS-relaxed CIS amplitude
…	$\mathbf{h}$	One-electron integrals	(C8)	$\overline{\mathbf{R}}$	CPCIS-relaxed transition density
…	$\boldsymbol{\Pi}$	Two-electron integrals	(C8)	$\overline{\mathbf{B}}$	CPCIS-relaxed difference density
…	$\mathbf{F}$	Fock matrix	(A17)	$\frac{\partial^2 E}{\partial\boldsymbol{\Theta}^2}$	CPHF supermatrix
…	$\mathbf{C}$	Molecular-orbital coefficients	(A17)	$\mathbf{M}^{[Q]}$	CPHF mixed-derivative matrix
…	$\mathbf{S}$	Overlap matrix	(B8)	$\mathbf{Y}$	CPHF Lagrangian for $\omega^{[Q]}$ and $\mathbf{d}^{[Q]}$
(37)	$\mathbf{P}$	Hartree–Fock density	(A14)	$\widetilde{\mathbf{L}}$	CPHF Lagrangian for $\boldsymbol{\mu}^{[Q]}$
(43)	$\mathbf{P}'$	Formal equivalent to $\mathbf{S}^{-1}$	…	$\mathbf{z}$	CPHF z-vector for $\omega^{[Q]}$ and $\mathbf{d}^{[Q]}$
…	E	Energy of an electronic state	(A18)	$\widetilde{\mathbf{z}}$	CPHF z-vector for $\boldsymbol{\mu}^{[Q]}$
…	$\boldsymbol{\Theta}$	Hartree–Fock orbital-rotation angle	(A21)	$\widetilde{\mathbf{B}}$	Fully-relaxed difference density

represented in the MO basis, will be indexed with ck and $\begin{smallmatrix}a\\i\end{smallmatrix}$, respectively. We will reserve the index pairs $\{\mu\lambda, \sigma\nu\}$ and $\{\alpha\gamma, \delta\beta\}$ for quantities that will be contracted with $\mathbf{\Pi}$ or $\mathbf{\Pi}^{[Q]}$.

Matrices will be italicized when in coordinate form (e.g., U_{AI}) and set in boldface when referred to in their entirety (e.g., $\mathbf{U}$). Because the dipole operator $\boldsymbol{\mu}$ is associated with a 3-vector in Cartesian space, it will *always* be set in boldface, as will vector quantities derived from it. These quantities include the Lagrangians and $\mathbf{z}$-vectors appearing in the coupled-perturbed CIS and coupled-perturbed Hartree–Fock equations, as well as the corresponding relaxed quantities.

Any asymmetric matrix can be partitioned into symmetric and antisymmetric components according to the relationships

$$\begin{cases} D_{\mu\nu} = D_{\mu\nu}^{\mathrm{S}} + D_{\mu\nu}^{\mathrm{A}} \\ D_{\mu\nu}^{\mathrm{S}} \equiv \frac{1}{2}\left(D_{\mu\nu} + D_{\nu\mu}\right) \\ D_{\mu\nu}^{\mathrm{A}} \equiv \frac{1}{2}\left(D_{\mu\nu} - D_{\nu\mu}\right) \end{cases} \Leftrightarrow \begin{cases} \mathbf{D} = \mathbf{D}^{\mathrm{S}} + \mathbf{D}^{\mathrm{A}} \\ \mathbf{D}^{\mathrm{S}} \equiv \frac{1}{2}\left(\mathbf{D} + \mathbf{D}^{\mathrm{T}}\right) \\ \mathbf{D}^{\mathrm{A}} \equiv \frac{1}{2}\left(\mathbf{D} - \mathbf{D}^{\mathrm{T}}\right) \end{cases}.$$
(6)

As such, we know that the trace of $\mathbf{D}$ against some other symmetric (antisymmetric) matrix is the same as the trace of $\mathbf{D}^{\mathrm{S}}$ ($\mathbf{D}^{\mathrm{A}}$) against that matrix. We will make frequent use of this fact, generally without comment.

II. DERIVATIVE COUPLINGS BETWEEN DIABATIC STATES

In this section, we develop the formalism necessary for computing derivative couplings between diabatic states; our approach is general, insofar as our expressions are initially agnostic with regard to both the diabatization method *and* the initial set of adiabatic states. We will select a diabatization and an adiabatic basis when necessary to further our discussion, but our reasoning remains valid independent of these choices.

A. General theory of diabatic-state derivative couplings

We begin with a set of M orthonormal adiabatic states, $\{|\Psi_I\rangle\}$. Applying an arbitrary diabatization method mixes these states together to return M diabatic states, $\{|\Xi_A\rangle\}$, defined by

$$|\Xi_A\rangle \equiv \sum_I |\Psi_I\rangle U_{AI},$$
(7)

where $\mathbf{U}$ is the $M \times M$ diabatization matrix. The overlap between any two of the resulting diabats is given by

$$\langle\Xi_A|\Xi_B\rangle = \sum_{IJ} U_{AI}\langle\Psi_I|\Psi_J\rangle U_{BJ} = \sum_I U_{AI}U_{BI}.$$
(8)

We will require that the diabats maintain orthonormality, which is equivalent to imposing the unitary condition on $\mathbf{U}$,

$$\sum_I U_{AI}U_{BI} = \delta_{AB} \quad\Leftrightarrow\quad \mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I} \quad\Leftrightarrow\quad \mathbf{U}^{\mathrm{T}} = \mathbf{U}^{-1}.$$
(9)

Consequently, the diabatization matrix has the form of a rotation matrix; for a two-state diabatization, it can be written as

$$\mathbf{U}(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix},$$
(10)

where we have introduced the mixing angle θ. Equal mixtures of the adiabatic states correspond to $\theta = (\pi/4)$.

The derivative couplings between the diabats can be obtained by inserting Eq. (7) into the definition in Eq. (1),

$$\mathbf{d}_{AB}^{[Q]} = \left(\sum_I U_{AI}\langle\Psi_I|\right)\nabla_Q\left(\sum_J |\Psi_J\rangle U_{BJ}\right)$$
$$= \sum_{IJ} U_{AI}\mathbf{d}_{IJ}^{[Q]}U_{BJ} + \sum_I U_{AI}U_{BI}^{[Q]}.$$
(11)

We see immediately that the derivative couplings between diabatic states naturally partition into two discrete contributions. The first term in Eq. (11) simply conjugates the adiabatic derivative couplings into the new basis; the second term accounts for couplings induced by changes in the composition of the diabats, as encoded in $\mathbf{U}^{[Q]}$.

Assuming that adiabatic derivative couplings are available for some desirable model chemistry, the problem at hand reduces to computing the gradient of the diabatization matrix. Obviously, the unitary condition of Eq. (9) does not provide enough information to fix its elements; taking the gradient yields

$$\sum_I U_{AI}^{[Q]}U_{BI} + \sum_I U_{AI}U_{BI}^{[Q]} = 0$$
(12)

for each of the $[M(M+1)/2]$ distinct (matched or unmatched) AB pairs. Additional constraints on $\mathbf{U}^{[Q]}$ must be supplied by specifying a diabatization method; for the sake of concreteness, we will select the Boys diabatization.

B. Convergence condition for Boys-diabatic states

Boys-diabatic states are defined as maximizing the functional of $\{|\Xi_A\rangle\}$ (function of $\mathbf{U}$)

$$f(\mathbf{U}) = \sum_{AB} |\boldsymbol{\mu}_{AA} - \boldsymbol{\mu}_{BB}|^2,$$
(13)

where $\boldsymbol{\mu}$ is the electronic dipole operator and $|\cdot|$ denotes the vector norm. In other words, Boys diabats are orthonormal linear combinations of the adiabatic states chosen to have maximally dissimilar charge distributions; they are well-suited to describing charge transfer within and between molecules.

In practice, the Boys functional is maximized using the iterative method of Jacobi sweeps. At convergence, the first variation of f in the diabats must vanish; the associated condition is[39]

$$\boldsymbol{\mu}_{AB} \cdot (\boldsymbol{\mu}_{AA} - \boldsymbol{\mu}_{BB}) = 0$$
(14)

for each of the $[M(M-1)/2]$ distinct (unmatched) AB pairs. This result, when combined with the unitary conditions of Eq. (9), yields M^2 equations in M^2 unknowns, uniquely determining the diabatization matrix. By extension, Eq. (12) can

be combined with the gradient of Eq. (14),

$$0 = \sum_{IJ} \boldsymbol{\mu}_{IJ}^{[Q]} \cdot (\boldsymbol{\mu}_{AA} - \boldsymbol{\mu}_{BB}) \, U_{AI} U_{BJ}$$

$$+ \sum_{IJ} \boldsymbol{\mu}_{IJ} \cdot (\boldsymbol{\mu}_{AA} - \boldsymbol{\mu}_{BB}) \left(U_{AI}^{[Q]} U_{BJ} + U_{AI} U_{BJ}^{[Q]} \right)$$

$$+ \sum_{KL} \boldsymbol{\mu}_{AB} \cdot \boldsymbol{\mu}_{KL}^{[Q]} (U_{AK} U_{AL} - U_{BK} U_{BL})$$

$$+ \sum_{KL} 2\boldsymbol{\mu}_{AB} \cdot \boldsymbol{\mu}_{KL} \left(U_{AK}^{[Q]} U_{AL} - U_{BK}^{[Q]} U_{BL} \right), \quad (15)$$

in order to fix the elements of $\mathbf{U}^{[Q]}$.

C. Supermatrix equations for U[Q]

We have enough information to solve for the diabatization-matrix gradient, but the conditions of Eqs. (12) and (15) are not practical for the purpose—at least, not as written. We emphasize that these equations form a linear system by rewriting them in terms of a single matrix element, $U_{CK}^{[Q]}$. Reformulating Eq. (12) in this way is straightforward:

$$0 = \sum_{KC} (\delta_{AC} U_{BK} + \delta_{BC} U_{AK}) \, U_{CK}^{[Q]}$$

$$\equiv \sum_{KC} \mathcal{A}_{ABKC} U_{CK}^{[Q]} \quad \forall \, A \leq B, \quad (16)$$

where we define the unitarity supermatrix $\mathcal{A}$. We can similarly reexpress Eq. (15) by separating out the terms

that are independent of $\mathbf{U}^{[Q]}$,

$$\mathcal{C}_{AB}^{\text{Boys}[Q]} \equiv \sum_{IJKL} \left[\boldsymbol{\mu}_{IJ}^{[Q]} \cdot \boldsymbol{\mu}_{KL} + \boldsymbol{\mu}_{IJ} \cdot \boldsymbol{\mu}_{KL}^{[Q]} \right]$$

$$\times U_{AI} U_{BJ} (U_{AK} U_{AL} - U_{BK} U_{BL}). \quad (17)$$

Simple manipulations then yield

$$-\mathcal{C}_{AB}^{\text{Boys}[Q]}$$

$$= \sum_{KC} \left\{ \begin{array}{l} \delta_{AC} \left[2\boldsymbol{\mu}_{AB} \cdot \boldsymbol{\mu}_{KA} + (\boldsymbol{\mu}_{AA} - \boldsymbol{\mu}_{BB}) \cdot \boldsymbol{\mu}_{KB} \right] \\ + \delta_{BC} \left[(\boldsymbol{\mu}_{AA} - \boldsymbol{\mu}_{BB}) \cdot \boldsymbol{\mu}_{KA} - 2\boldsymbol{\mu}_{AB} \cdot \boldsymbol{\mu}_{KB} \right] \end{array} \right\} U_{CK}^{[Q]}$$

$$(18)$$

$$\equiv \sum_{KC} \mathcal{B}_{ABKC}^{\text{Boys}} U_{CK}^{[Q]} \quad \forall \, A < B. \quad (19)$$

In light of these results, we see that obtaining $\mathbf{U}^{[Q]}$ requires us to solve $3N_A$ supermatrix equations

$$\sum_{KC} \begin{bmatrix} \mathcal{A}_{ABKC} \\ \mathcal{B}_{ABKC}^{\text{Boys}} \end{bmatrix} U_{CK}^{[Q]} = - \begin{bmatrix} \mathbf{0} \\ \mathcal{C}_{AB}^{\text{Boys}[Q]} \end{bmatrix}, \quad (20)$$

with both universal ($\mathcal{A}$) and diabatization-specific ($\mathcal{B}$, $\mathcal{C}^{[Q]}$) components.

To solve Eq. (20), we combine $\mathcal{A}$ and $\mathcal{B}$ to form a square constraint matrix of size $M^2 \times M^2$, as shown in partial detail below (Eq. (21)). Given that diabatic states are usually formed from a small subset of the adiabats—typically $M = 2$—the constraint matrix can be inverted directly. Of course, there can in theory be specific systems and geometries for which the diabatization itself is ill-behaved, giving rise to a vanishing row—i.e., singularity—in the constraint matrix.

$$(21)$$

D. Analytic gradients for diabatic states: $E_A^{[Q]}$, $H_{AB}^{[Q]}$, and $\mu_{AB}^{[Q]}$

Although the primary aim of this paper is to compute diabatic-state derivative couplings, we note that we can re-purpose $\mathbf{U}^{[Q]}$ to calculate analytic gradients of essentially *any* property of interest. We briefly outline the necessary equations here, deferring more detailed discussion to future publications.

Suppose we have some arbitrary operator $\mathcal{O}$ and that its matrix elements and their analytic gradients, $\mathcal{O}_{IJ}$ and $\mathcal{O}_{IJ}^{[Q]}$, are available. The corresponding diabatic quantities are

$$\mathcal{O}_{AB} = \sum_{IJ} U_{AI}\mathcal{O}_{IJ}U_{BJ}, \tag{22}$$

$$\mathcal{O}_{AB}^{[Q]} = \sum_{IJ} U_{AI}^{[Q]}\mathcal{O}_{IJ}U_{BJ} + \sum_{IJ} U_{AI}\mathcal{O}_{IJ}^{[Q]}U_{BJ}$$
$$+ \sum_{IJ} U_{AI}\mathcal{O}_{IJ}U_{BJ}^{[Q]}, \tag{23}$$

which can be computed straightforwardly, given $\mathbf{U}^{[Q]}$. For example, diabatic energies and gradients may be obtained from

$$E_A = \sum_I U_{AI}E_I U_{AI}, \tag{24}$$

$$E_A^{[Q]} = \sum_I U_{AI}E_I^{[Q]}U_{AI} + 2\sum_I U_{AI}E_I U_{AI}^{[Q]} \tag{25}$$

for use in optimizing geometries on the diabatic surface; diabatic couplings and coupling gradients from

$$H_{AB} = \sum_I U_{AI}E_I U_{BI}, \tag{26}$$

$$H_{AB}^{[Q]} = \sum_I U_{AI}^{[Q]}E_I U_{BI} + \sum_I U_{AI}E_I^{[Q]}U_{BI}$$
$$+ \sum_I U_{AI}E_I U_{BI}^{[Q]} \tag{27}$$

for use in evaluating the validity of the Condon approximation ($H_{AB}^{[Q]} \approx 0$); and diabatic dipole matrices and dipole gradients from

$$\boldsymbol{\mu}_{AB} = \sum_{IJ} U_{AI}\boldsymbol{\mu}_{IJ}U_{BJ}, \tag{28}$$

$$\boldsymbol{\mu}_{AB}^{[Q]} = \sum_{IJ} U_{AI}^{[Q]}\boldsymbol{\mu}_{IJ}U_{BJ} + \sum_{IJ} U_{AI}\boldsymbol{\mu}_{IJ}^{[Q]}U_{BJ}$$
$$+ \sum_{IJ} U_{AI}\boldsymbol{\mu}_{IJ}U_{BJ}^{[Q]} \tag{29}$$

for use in constructing spectra incorporating Herzberg–Teller effects, such as vibronic intensity borrowing.[40] Of course, CIS spectra should be qualitatively accurate at best, irrespective of whether they are obtained from the adiabatic or diabatic representation; CIS transition dipole moments do not even satisfy the Thomas–Reiche–Kuhn sum rule for oscillator strengths.[41]

III. DERIVATIVE COUPLINGS BETWEEN BOYS-DIABATIC CIS STATES: ANALYTIC THEORY FOR $\mu_{IJ}^{[Q]}$

To obtain $\mathbf{U}^{[Q]}$ in practice, we must make a specific choice of adiabatic states, then compute the adiabatic dipole matrix $\boldsymbol{\mu}_{IJ}$ and its gradient $\boldsymbol{\mu}_{IJ}^{[Q]}$ in order to construct $\mathcal{B}$ and $\mathcal{C}^{[Q]}$. Here we will diabatize and couple CIS states, for two reasons: First, CI derivative couplings[42–48] and gradients[49–53] are well-known and readily available to us from our previous work.[54] (Please refer to the latter part of Appendix B for the equations.) Second, and most immediately relevant, it is not unreasonably difficult to write down a complete analytic-gradient theory for $\boldsymbol{\mu}_{IJ}^{[Q]}$ within CIS by invoking basic definitions and applying the chain rule. Similar calculations in time-dependent Hartree–Fock or time-dependent density-functional theory require a more sophisticated approach based on linear response.[55]

A. CIS states

CIS is perhaps the simplest theory of excited states.[41] Each adiabatic excited state has the form

$$|\Psi_I\rangle = \sum_{ia} t_i^{Ia}|\Phi_i^a\rangle. \tag{30}$$

The CIS amplitudes $\mathbf{t}^I$ are uniquely specified by diagonalizing the CIS supermatrix

$$A_{\ell dkc} \equiv \langle\Phi_\ell^d|H|\Phi_k^c\rangle \tag{31}$$

$$= \Pi_{dk\ell c} + \delta_{k\ell}F_{dc} - \delta_{cd}F_{\ell k} + \delta_{k\ell}\delta_{cd}E \tag{32}$$

$$= \Pi_{dk\ell c} + \delta_{kl}\delta_{cd}(E + \varepsilon_c - \varepsilon_k), \tag{33}$$

where we have introduced the orbital energies (diagonal Fock-matrix elements) $F_{pq} = \varepsilon_p\delta_{pq}$ and the Hartree–Fock energy E. Because singles determinants are orthogonal, $\langle\Phi_i^a|\Phi_j^b\rangle = \delta_{ij}\delta_{ab}$, the CIS normalization condition can be written solely in terms of the amplitudes,

$$\langle\Psi_I|\Psi_J\rangle = \sum_{ijab} t_i^{Ia}\langle\Phi_i^a|\Phi_j^b\rangle t_j^{Jb} = \sum_{ia} t_i^{Ia}t_i^{Ja} = \delta_{IJ}. \tag{34}$$

The amplitudes also satisfy the CIS energy equation

$$\sum_{kc} A_{\ell dkc}t_k^{Ic} = E_I t_\ell^{Id}, \tag{35}$$

with energies given by

$$E_I = E + \sum_{\mu\nu} h_{\mu\nu}B_{\mu\nu}^{II} + \sum_{\mu\nu\lambda\sigma}\left(R_{\mu\lambda}^I R_{\sigma\nu}^I + B_{\mu\lambda}^{II}P_{\sigma\nu}\right)\Pi_{\mu\nu\lambda\sigma}$$

$$\equiv E + \omega_I. \tag{36}$$

In Eq. (36), we have defined the CIS excitation energy ω_I and atomic-orbital representations of the Hartree–Fock density $\mathbf{P}$, CIS transition density $\mathbf{R}^I$, and (generalized) CIS difference density $\mathbf{B}^{IJ}$ (in terms of the molecular-orbital coefficients $\mathbf{C}$),

$$P_{\mu\nu} \equiv \sum_m C_{\mu m}C_{\nu m}, \tag{37a}$$

$$R^I_{\mu\nu} \equiv \sum_{ia} C_{\mu a} t^{Ia}_i C_{\nu i}, \tag{37b}$$

$$B^{IJ}_{\mu\nu} \equiv \sum_{iab} C_{\mu a} t^{Ia}_i t^{Jb}_i C_{\nu b} - \sum_{ija} C_{\mu i} t^{Ia}_i t^{Ja}_j C_{\nu j}. \tag{37c}$$

B. Adiabatic dipole matrix $\boldsymbol{\mu}_{IJ}$

The elements of the adiabatic dipole matrix are given in second quantization by

$$\boldsymbol{\mu}_{IJ} = \sum_{pq} \boldsymbol{\mu}_{pq} \langle \Psi_I | a^\dagger_p a_q | \Psi_J \rangle \equiv \sum_{pq} \boldsymbol{\mu}_{pq} D^{IJ}_{pq} = \sum_{\mu\nu} \boldsymbol{\mu}_{\mu\nu} D^{IJ}_{\mu\nu}, \tag{38}$$

where $\mathbf{D}^{IJ}$ is the inter-CIS transition density between states I and J,

$$D^{IJ}_{pq} = \sum_{iab} t^{Ia}_i t^{Jb}_i \delta_{ap}\delta_{bq} - \sum_{ija} t^{Ia}_i t^{Ja}_j \delta_{jp}\delta_{iq} + \delta_{IJ}\delta^{\mathrm{occ}}_{pq}. \tag{39}$$

We can immediately write down $\boldsymbol{\mu}_{IJ}$ in MO and AO representations,

$$\boldsymbol{\mu}_{IJ} = \sum_{iab} t^{Ia}_i \boldsymbol{\mu}_{ab} t^{Jb}_i - \sum_{ija} t^{Ia}_i \boldsymbol{\mu}_{ij} t^{Ja}_j + \delta_{IJ} \sum_m \boldsymbol{\mu}_{mm} \tag{40}$$

$$= \sum_{\mu\nu} \boldsymbol{\mu}_{\mu\nu} B^{IJ}_{\mu\nu} + \delta_{IJ} \boldsymbol{\mu}_{\mathrm{HF}}, \tag{41}$$

where we have identified the Hartree–Fock dipole moment. This HF contribution will cancel from Eqs. (13) and (14) and cannot have any effect on our diabatization. We may therefore omit this term when constructing $\mathcal{B}$ and $\mathcal{C}^{[Q]}$, although we retain it in our expressions.

C. Adiabatic dipole-matrix gradient $\boldsymbol{\mu}^{[Q]}_{IJ}$

The gradient of Eq. (41) can be obtained by noting the dependence of the difference density on the CIS amplitudes and the MO coefficients, which depend in turn on the overlap matrix and the Hartree–Fock orbital-rotation angles, i.e., $\mathbf{B}^{IJ} \equiv \mathbf{B}^{IJ}[\mathbf{t}^I, \mathbf{t}^J, \mathbf{C}(\mathbf{S}, \boldsymbol{\Theta})]$. (The implicit dependence of the density on the underlying atomic-orbital basis—and the associated overlap matrix $\mathbf{S}$—is the source of Pulay forces in analytic gradients;[56] we must of course account for these terms explicitly.) Applying the chain rule, we obtain

$$\boldsymbol{\mu}^{[Q]}_{IJ} = \delta_{IJ} \boldsymbol{\mu}^{[Q]}_{\mathrm{HF}} + \sum_{\mu\nu} \boldsymbol{\mu}_{\mu\nu} B^{IJ}_{\mu\nu}$$

$$+ \sum_{\mu\nu\alpha\beta} \boldsymbol{\mu}_{\mu\nu} \left(\frac{\partial B^{IJ}_{\mu\nu}}{\partial S_{\alpha\beta}} \right)_{\mathbf{t}^I, \mathbf{t}^J, \boldsymbol{\Theta}} S^{[Q]}_{\alpha\beta}$$

$$+ \sum_{kc\mu\nu} \boldsymbol{\mu}_{\mu\nu} \left(\frac{\partial B^{IJ}_{\mu\nu}}{\partial \Theta_{ck}} \right)_{\mathbf{t}^I, \mathbf{t}^J, \mathbf{S}} \Theta^{[Q]}_{ck}$$

$$+ \sum_{iab\mu\nu} \boldsymbol{\mu}_{\mu\nu} C_{\mu a} \left(t^{Ia[Q]}_i t^{Jb}_i + t^{Ia}_i t^{Jb[Q]}_i \right) C_{\nu b}$$

$$- \sum_{ija\mu\nu} \boldsymbol{\mu}_{\mu\nu} C_{\mu i} \left(t^{Ia[Q]}_i t^{Ja}_j + t^{Ia}_i t^{Ja[Q]}_j \right) C_{\nu j}. \tag{42}$$

Invoking the standard relations[57]

$$\frac{\partial C_{\mu p}}{\partial S_{\alpha\beta}} = -\frac{1}{2} \sum_q C_{\mu q} C_{\alpha q} C_{\beta p} \equiv -\frac{1}{2} P'_{\mu\alpha} C_{\beta p}, \tag{43}$$

$$\frac{\partial C_{\mu p}}{\partial \Theta_{ck}} = \delta_{cp} C_{\mu k} - \delta_{kp} C_{\mu c} \tag{44}$$

leads in short order to the partial derivatives

$$\left(\frac{\partial B^{IJ}_{\mu\nu}}{\partial S_{\alpha\beta}} \right)_{\mathbf{t}^I, \mathbf{t}^J, \boldsymbol{\Theta}} = -\frac{1}{2} \left(P'_{\mu\alpha} B^{IJ}_{\beta\nu} + P'_{\nu\alpha} B^{IJ}_{\mu\beta} \right), \tag{45}$$

$$\left(\frac{\partial B^{IJ}_{\mu\nu}}{\partial \Theta_{ck}} \right)_{\mathbf{t}^I, \mathbf{t}^J, \mathbf{S}} = \sum_{ia} t^{Ic}_i t^{Ja}_i C_{\nu a} C_{\mu k} + \sum_{ia} t^{Jc}_i t^{Ia}_i C_{\mu a} C_{\nu k}$$

$$+ \sum_{ia} C_{\mu c} C_{\nu i} t^{Ja}_i t^{Ia}_k + \sum_{ia} C_{\nu c} C_{\mu i} t^{Ia}_i t^{Ja}_k. \tag{46}$$

Inserting these results into Eq. (42) and collecting terms leads to the expression

$$\boldsymbol{\mu}^{[Q]}_{IJ} = \delta_{IJ} \boldsymbol{\mu}^{[Q]}_{\mathrm{HF}} + \sum_{\mu\nu} \boldsymbol{\mu}^{[Q]}_{\mu\nu} B^{IJ\mathrm{S}}_{\mu\nu} - \sum_{\mu\nu\alpha\beta} S^{[Q]}_{\mu\nu} P'_{\mu\alpha} \boldsymbol{\mu}_{\alpha\beta} B^{IJ\mathrm{S}}_{\nu\beta}$$

$$+ \sum_{kc} \left[\begin{array}{l} \sum_{ia\mu\nu} \left(t^{Ic}_i t^{Ja}_i + t^{Jc}_i t^{Ia}_i \right) C_{\mu a} \boldsymbol{\mu}_{\mu\nu} C_{\nu k} \\ + \sum_{ia\mu\nu} C_{\mu c} \boldsymbol{\mu}_{\mu\nu} C_{\nu i} \left(t^{Ia}_i t^{Ja}_k + t^{Ia}_i t^{Ja}_k \right) \end{array} \right] \Theta^{[Q]}_{ck}$$

$$+ \sum_{ia} \left[\sum_{b\mu\nu} C_{\mu a} \boldsymbol{\mu}_{\mu\nu} C_{\nu b} t^{Jb}_i - \sum_{j\mu\nu} t^{Ja}_j C_{\mu j} \boldsymbol{\mu}_{\mu\nu} C_{\nu i} \right] t^{Ia[Q]}_i$$

$$+ \sum_{ia} \left[\sum_{b\mu\nu} C_{\mu a} \boldsymbol{\mu}_{\mu\nu} C_{\nu b} t^{Ib}_i - \sum_{j\mu\nu} t^{Ia}_j C_{\mu j} \boldsymbol{\mu}_{\mu\nu} C_{\nu i} \right] t^{Ja[Q]}_i, \tag{47}$$

which is manifestly symmetric in I and J, if one considers that $\left(\mathbf{B}^{IJ} \right)^{\mathsf{T}} = \mathbf{B}^{JI}$. For completeness, we note that the HF dipole gradient is given by

$$\boldsymbol{\mu}^{[Q]}_{\mathrm{HF}} = \sum_{\mu\nu} \boldsymbol{\mu}^{[Q]}_{\mu\nu} P_{\mu\nu} - \sum_{\mu\nu\alpha\beta} S^{[Q]}_{\mu\nu} P'_{\mu\alpha} \boldsymbol{\mu}_{\alpha\beta} P_{\nu\beta}$$

$$- 2 \sum_{kc\mu\nu} C_{\mu c} \boldsymbol{\mu}_{\mu\nu} C_{\nu k} \Theta^{[Q]}_{ck}. \tag{48}$$

Details of the amplitude- and orbital-response calculations necessary for evaluating Eq. (47) are provided in Appendix A. The resulting analytic expression for the CIS

dipole-matrix gradient is

$$
\begin{aligned}
\boldsymbol{\mu}_{IJ}^{[Q]} = {} & \delta_{IJ}\boldsymbol{\mu}_{\mathrm{HF}}^{[Q]} + \sum_{\mu\nu}\boldsymbol{\mu}_{\mu\nu}^{[Q]}B_{\mu\nu}^{IJS} - \sum_{\mu\nu\alpha\beta}S_{\mu\nu}^{[Q]}P'_{\mu\alpha}\boldsymbol{\mu}_{\alpha\beta}B_{\nu\beta}^{IJS} \\
& + \sum_{\mu\nu}h_{\mu\nu}^{[Q]}\left(\widetilde{\mathbf{B}}_{\mu\nu}^{IJ} + \widetilde{\mathbf{B}}_{\mu\nu}^{JI}\right) \\
& + \sum_{\mu\nu\lambda\sigma}\left[\overline{\mathbf{R}}_{\mu\lambda}^{IJ}R_{\sigma\nu}^{J} + \overline{\mathbf{R}}_{\mu\lambda}^{JI}R_{\sigma\nu}^{I} + \left(\widetilde{\mathbf{B}}_{\mu\lambda}^{IJ} + \widetilde{\mathbf{B}}_{\mu\lambda}^{JI}\right)P_{\sigma\nu}\right]\Pi_{\mu\nu\lambda\sigma}^{[Q]} \\
& - \sum_{\mu\nu\alpha\beta}S_{\mu\nu}^{[Q]}P'_{\mu\alpha}F_{\alpha\beta}\left(\widetilde{\mathbf{B}}_{\nu\beta}^{IJ} + \widetilde{\mathbf{B}}_{\nu\beta}^{JI}\right) \\
& - \frac{1}{2}\sum_{\mu\nu\alpha\beta\gamma\delta}S_{\mu\nu}^{[Q]}P'_{\mu\alpha}
\begin{bmatrix}
\overline{\mathbf{R}}_{\nu\gamma}^{IJ}R_{\delta\beta}^{J} + R_{\nu\gamma}^{J}\overline{\mathbf{R}}_{\delta\beta}^{IJ} \\
+\overline{\mathbf{R}}_{\gamma\nu}^{IJ}R_{\beta\delta}^{J} + R_{\gamma\nu}^{J}\overline{\mathbf{R}}_{\beta\delta}^{IJ} \\
+\overline{\mathbf{R}}_{\nu\gamma}^{JI}R_{\delta\beta}^{I} + R_{\nu\gamma}^{I}\overline{\mathbf{R}}_{\delta\beta}^{JI} \\
+\overline{\mathbf{R}}_{\gamma\nu}^{JI}R_{\beta\delta}^{I} + R_{\gamma\nu}^{I}\overline{\mathbf{R}}_{\beta\delta}^{JI}
\end{bmatrix}
\Pi_{\alpha\beta\gamma\delta} \\
& - \sum_{\mu\nu\alpha\beta\gamma\delta}S_{\mu\nu}^{[Q]}P'_{\mu\alpha}P_{\nu\gamma}\Pi_{\alpha\beta\gamma\delta}\left(\widetilde{\mathbf{B}}_{\delta\beta}^{IJ} + \widetilde{\mathbf{B}}_{\delta\beta}^{JI}\right).
\end{aligned}
\tag{49}
$$

The various quantities appearing in this equation are named in Table I and/or defined in Eqs. (37) (**P**, **R**, **B**), (43) (**P**′), (48) ($\boldsymbol{\mu}_{\mathrm{HF}}^{[Q]}$), (A21) ($\widetilde{\mathbf{B}}$), and (C8) ($\overline{\mathbf{R}}$). Further useful definitions are given in Eqs. (A5), (A9), (A14), (A18), and (C3).

D. Estimated computational costs of our method

Having obtained a closed form for $\boldsymbol{\mu}_{IJ}^{[Q]}$, we take a moment to consider the cost of calculating the diabatization-matrix gradient. The most computationally costly and memory-intensive operation required to obtain the elements of the dipole-matrix gradient is the contraction of form $\mathbf{D}\cdot\boldsymbol{\Pi}^{[Q]}\cdot\mathbf{D}'$, which occurs twice in calculating the CIS gradient or (adiabatic) derivative coupling. Referring to Eq. (49) and noting that the relaxed densities $\overline{\mathbf{R}}$ and $\widetilde{\mathbf{B}}$ are Cartesian 3-vectors, we see that each diagonal element of the dipole-matrix gradient requires 6 contractions with $\boldsymbol{\Pi}^{[Q]}$, while each off-diagonal element involves 9. In other words, we may write

$$
\mathrm{cost}\left(\omega_I^{[Q]}\right) \approx \mathrm{cost}\left(\mathbf{d}_{IJ}^{[Q]}\right) \approx \frac{1}{3}\mathrm{cost}\left(\boldsymbol{\mu}_I^{[Q]}\right) \approx \frac{2}{9}\mathrm{cost}\left(\boldsymbol{\mu}_{IJ}^{[Q]}\right).
\tag{50}
$$

For a Boys diabatization involving M states, the relevant part of $\boldsymbol{\mu}_{IJ}^{[Q]}$ for computing $\mathbf{U}^{[Q]}$ includes M diagonal and $[M(M-1)/2]$ off-diagonal elements:

$$
\begin{aligned}
\mathrm{cost}(\mathbf{U}^{[Q]}) & \approx M\,\mathrm{cost}\left(\boldsymbol{\mu}_I^{[Q]}\right) + \frac{M(M-1)}{2}\mathrm{cost}\left(\boldsymbol{\mu}_{IJ}^{[Q]}\right) \\
& \approx \frac{3M(3M+1)}{4}\mathrm{cost}\left(\omega_I^{[Q]}\right).
\end{aligned}
\tag{51}
$$

We therefore estimate that the cost of our method for a two-state Boys diabatization will be 10 times that of a CIS gradient.

IV. NUMERICAL RESULTS

We have implemented restricted and unrestricted versions of Eqs. (11), (20), and (49) within a development version of the Q-CHEM software package.[58,59] To verify our

derivations and implementation, we validated our method against finite difference for a variety of diabatic quantities and a range of systems, including H_2, HeH^+, LiF, H_2O, and NHFCl. Results for LiH are provided below; a simple application of our method to *p*-benzoquinone follows.

A. Validation by finite difference: LiH Boys-diabatic derivative couplings

A stringent test of our analytic derivative couplings is to compare them for a small system against the corresponding central-difference formula for each Cartesian degree of freedom x,

$$
\begin{aligned}
\mathbf{d}_{IJ}^{[x]} \approx {} & \frac{\langle\Psi_I(x)|\Psi_J(x+\Delta x)\rangle - \langle\Psi_I(x)|\Psi_J(x-\Delta x)\rangle}{2\Delta x} \\
& + O[(\Delta x)^2],
\end{aligned}
\tag{52a}
$$

$$
\begin{aligned}
\mathbf{d}_{AB}^{[x]} \approx {} & \frac{\langle\Xi_A(x)|\Xi_B(x+\Delta x)\rangle - \langle\Xi_A(x)|\Xi_B(x-\Delta x)\rangle}{2\Delta x} \\
& + O[(\Delta x)^2].
\end{aligned}
\tag{52b}
$$

Slater determinants associated with different molecular geometries need not be orthonormal, so we must write down an exact expression for their overlap. Distinguishing determinants with different underlying geometries by a prime, we write

$$
\langle\Psi_I|\Psi_J'\rangle = \sum_{ijab}t_i^{Ia}\left\langle\Phi_i^a\big|\Phi_j^{b\prime}\right\rangle t_j^{\prime Jb},
\tag{53a}
$$

$$
\langle\Xi_A|\Xi_B'\rangle = \sum_{ijab}U_{AI}t_i^{Ia}\left\langle\Phi_i^a\big|\Phi_j^{b\prime}\right\rangle t_j^{\prime Jb}U'_{BJ}.
\tag{53b}
$$

We then apply a theorem due to Löwdin[60] to evaluate the singles bracket; the result can be found in Appendix D.

We performed calculations for LiH at the HF/cc-pVDZ minimum characterized by Li–H = 1.618436 Å, with the molecule aligned to the z axis. Table II lists the lowest-lying CIS/cc-pVDZ singlet states at this geometry; we focus our attention on the non-degenerate S_1/S_3 state pair. Consistent with our previous analysis[54] of the scaling of finite-difference error in LiH derivative couplings, we used a step size of $\Delta x = 10^{-4}$ Å in our calculations; when combined with HF and CIS energy-convergence conditions of 10^{-8} E_h, this choice leads to an expected agreement between finite difference and analytical theory of $\sim 10^{-4}$ a_0^{-1}. As shown in Table III, we more than satisfy this expectation—in other words, our derivation

TABLE II. Symmetries, excitation energies ω, and dipole moments μ^z of the lowest-lying CIS/cc-pVDZ singlet states of LiH at the Hartree–Fock minimum. The HF energy is $E = -7.983686$ E_h.

State	Symmetry	ω (eV)	μ^z (D)
S_1	Σ^+	4.0248	−6.7308
S_2	Π	5.0651	−1.1415
S_3	Σ^+	6.9219	6.2950
S_4	Π	7.8317	−1.0054

TABLE III. Derivative couplings induced by motion of atom Q between the adiabatic state pair S_1/S_3 and the corresponding Boys diabats of LiH, as computed for CIS/cc-pVDZ by finite difference (FD: Eqs. (52), (53), and (D1)), analytic theory for the adiabats (CIS: Eqs. (B5) and (B7)–(B9)), and analytic theory for the diabats (Boys: Eqs. (11), (20), and (49)). The agreement is far better than the expected order of errors ($\sim 10^{-4}\,a_0^{-1}$). Note that the couplings have vanishing components except along the z axis.

Atom (Q)	$\mathbf{d}_{IJ}^{[Q]}\ \left(a_0^{-1}\right)$		$\mathbf{d}_{AB}^{[Q]}\ \left(a_0^{-1}\right)$	
	FD	CIS	FD	Boys
H	0.047933	0.047931	0.079044	0.079043
Li	−0.146642	−0.146641	−0.177753	−0.177753

of the Boys-diabatic analytic-gradient theory is correct, and our implementation is sound.

We emphasize that, in diabatizing S_1 and S_3, we intend *only* to validate our implementation of Eqs. (11), (20), and (49). Obviously, S_1 and S_3 are not usefully mixed at this geometry; the energy gap between them is significant.[61] For a more physical example, we turn to *p*-benzoquinone.

B. Boys diabatization of distorted *p*-benzoquinone

We now make a simple demonstration of the fact that Boys diabats have near-vanishing derivative couplings when applied to adiabatic states close to a crossing point. Our test case is *p*-benzoquinone (Figure 1), which is of biological significance as a building block of photosynthetic electron acceptors, Vitamin K, and chemical defenses in insects.[62,63] *p*-benzoquinone and its derivatives have been studied extensively, with a significant literature discussing their electronic structure and spectroscopy,[64–68] normal modes and vibrational spectra,[63,69–71] and photoacceptance/-detachment properties.[72–75] Our group has previously studied an electron-transfer system consisting of benzene and several *p*-benzoquinone molecules.[26]

Instead of the Hartree–Fock-minimum structure of *p*-benzoquinone, we study a geometry with stretched bonds and distorted angles that nevertheless maintains approximate D_{2h} symmetry. Average geometric parameters are provided in Table IV; the associated Cartesian coordinates are listed in
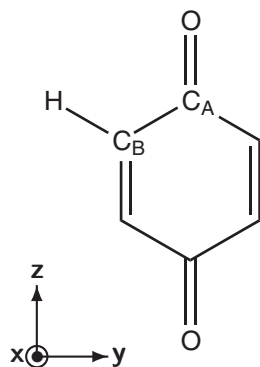

FIG. 1. *p*-benzoquinone, with labels to distinguish carbon types and one hydrogen shown explicitly. The coordinate system used in assigning state symmetries and reporting dipole moments is depicted at bottom left.

TABLE IV. Average geometric parameters of the distorted *p*-benzoquinone studied in this paper. Bond types indicated are meant only to provide a correspondence with the sketch in Figure 1; we did not compute bond orders.

Bond type	Bond length (Å)	Angle type	Angle (°)
C=O	1.428	∠ C–C–C	157.8
C=C	1.064	∠ C–C=C	101.1
C–C	1.659	∠ C=C–H	143.4
C–H	1.155	∠ C–C–H	115.3

Appendix E. As shown in Table V, this geometry has several closely-spaced CIS/6-31G** singlet states; indeed, S_2, S_3, and S_4 are separated by less than the thermal energy at room temperature.

The mixing angle for Boys diabatization grows larger with the transition dipole moment between states; as shown in Refs. 20, 21, and 26, it is given by

$$\tan 2\theta = \frac{2\boldsymbol{\mu}_{IJ}}{|\boldsymbol{\mu}_I - \boldsymbol{\mu}_J|} \quad (54)$$

when the moments are aligned. Thus, we might expect diabatization to be most important for state pairs I/J with a large transition dipole. The symmetry products of our candidate states are

$$S_2/S_3:\quad B_{1u} \otimes B_{3g} = B_{2u} \quad y\text{-active}:\quad \boldsymbol{\mu}_{23} \to \mu_{23}^y, \quad (55a)$$

$$S_2/S_4:\quad B_{1u} \otimes B_{1g} = A_u \quad \text{inactive}:\quad \boldsymbol{\mu}_{24} = 0, \quad (55b)$$

$$S_3/S_4:\quad B_{3g} \otimes B_{1g} = B_{2g} \quad \text{inactive}:\quad \boldsymbol{\mu}_{34} = 0. \quad (55c)$$

Following this intuition, we diabatize the S_2/S_3 pair, which does indeed have a comparatively large transition dipole aligned with the y axis, $\mu_{23}^y = -2.444$ D.

We find that the Boys diabatization mixes S_2 and S_3 almost equally, $\theta \approx (\pi/4)$. The effect of the diabatization is conveniently visualized by imaging the attachment/detachment densities of the adiabatic and diabatic states together[76]—that is to say, by separately plotting isosurfaces of the two terms of $\mathbf{B}^{IJ}$ (cf. Eq. (37c)) and $\mathbf{B}^{AB}$, as in Figure 2. We see that the adiabatic attachment densities strongly resemble each other—and, therefore, the corresponding diabatic attachment densities, which remain delocalized. At the same time, the diabatic detachment densities are clearly charge-localized on the left- or right-hand side of the molecule, in stark contrast to the adiabatic densities.

TABLE V. Approximate symmetries, excitation energies ω, and dipole moments μ^y—other moments are negligible—of the lowest-lying CIS/6-31G** singlet states of *p*-benzoquinone in the distorted geometry of Appendix E. The HF energy is $E = -378.417577\ E_h$.

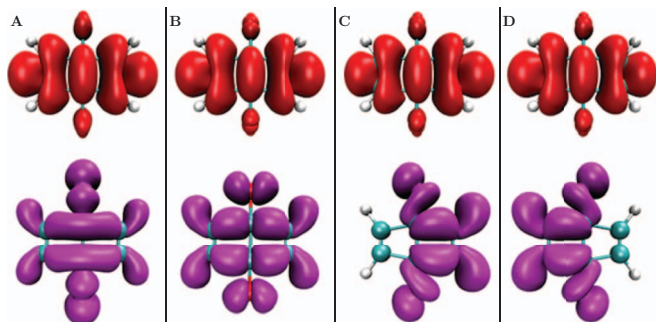| State	Approx. symmetry	ω (eV)	μ^y (D)
S_1	A_u	2.4012	0.0001
S_2	B_{1u}	2.8532	0.0201
S_3	B_{3g}	2.8562	−0.0201
S_4	B_{1g}	2.8586	0.0000
S_5	B_{2g}	2.9195	0.0000
S_6	B_{3u}	3.9543	0.0000

FIG. 2. Attachment/detachment (red/purple) densities for S_2 (Panel A) and S_3 (Panel B) and the corresponding Boys diabats (Panels C and D). All of these densities are symmetric in the plane of the molecule; the attachment densities are nearly indistinguishable. Because excited electrons are transferred *to* areas covered by the attachment density *from* those covered by the detachment density, Panel C depicts the left-polarized state with dipole $\mu^y = -2.444$ D; Panel D, the right-polarized state with dipole $\mu^y = 2.444$ D.
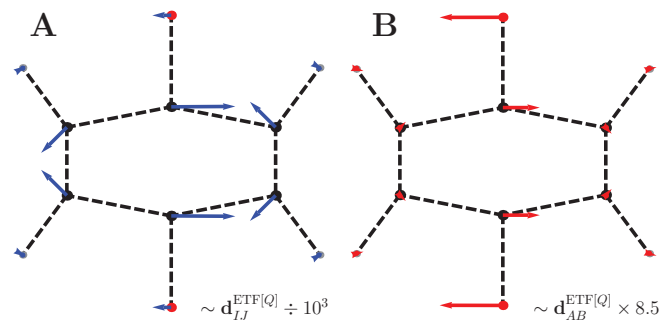

FIG. 3. The ETF-corrected derivative couplings (Eq. (B10)) between adiabatic states S_2 and S_3 (Panel A) compared with those between the corresponding Boys-diabatic states (Panel B). The vectors in each panel are scaled by the norm of the largest coupling, per Table VI. Note that the derivative couplings correspond to a b_{2u} collective motion, consistent with the direct product of state symmetries. Normal modes of the same symmetry include carbon-carbon bond stretches, carbon-hydrogen bond stretches and bends, and the carbonyl bend.[63, 70]

Having mixed S_2 and S_3 to form the diabats, we next computed the derivative couplings, the magnitudes of which are listed in Table VI; the vectors are plotted against the *p*-benzoquinone bond edges in Figure 3. (To be precise, we computed the derivative couplings with and without perturbative electron-translation factors. Results presented for *p*-benzoquinone include this correction, as per Eq. (B10), and are superscripted accordingly with *ETF*; for more information on this point, refer to the latter part of Appendix B and Refs. 54 and 77.) The adiabatic derivative couplings are extremely large—as we would expect, owing to the energy denominator in $\mathbf{d}_{IJ}^{[Q]}$ (cf. Eq. (B5)). By contrast, the derivative couplings between the diabats are reduced almost to zero; the effective scaling associated with the diabatization is no less than ≈ 2600 and as large as $\approx 37\,000$ (cf. Table VI). These data confirm that Boys diabatization yields a nearly perfect pair of diabats for distorted *p*-benzoquinone.

C. CPCIS contributions to $\mu_{IJ}^{[Q]}$ are crucial

It is reasonable to wonder whether the numerical effort involved in computing $\mu_{IJ}^{[Q]}$ can be avoided by instead using its unrelaxed equivalent

$$\text{unrelaxed } \boldsymbol{\mu}_{IJ}^{[Q]} = \delta_{IJ} \boldsymbol{\mu}_{\text{HF}}^{[Q]} + \sum_{\mu\nu} \boldsymbol{\mu}_{\mu\nu}^{[Q]} B_{\mu\nu}^{IJS}$$
$$- \sum_{\mu\nu\alpha\beta} S_{\mu\nu}^{[Q]} P'_{\mu\alpha} \boldsymbol{\mu}_{\alpha\beta} B_{\nu\beta}^{IJS}, \quad (56)$$

which is essentially free to compute.

General considerations based on the structure of the CP-CIS supermatrix (cf. Appendix A 1) show that this approximation would fail abjectly at a conical intersection between states *I* and *J*: Because the corresponding CIS energies are degenerate, the bare supermatrix $(\mathbf{A} - E_I)$ has vanishing eigenvalues for both $\mathbf{t}^I$ and $\mathbf{t}^J$. While the normalization condition can be incorporated into the supermatrix to shift the eigenvalue of $\mathbf{t}^I$, as in Eq. (A2), no similar modification can be made for $\mathbf{t}^J$. Thus, the full supermatrix $\mathcal{D}^I$ also becomes singular, and the amplitude response $\mathbf{t}^{I[Q]}$ will necessarily diverge, with exquisite sensitivity to the precise values of $\mathbf{t}^J$. (Similar arguments hold for $\mathbf{t}^{J[Q]}$ and its sensitivity to $\mathbf{t}^I$.) This divergence is carried from $\boldsymbol{\mu}_{IJ}^{[Q]}$ into $\mathbf{U}^{[Q]}$ and the unitarity-gradient term in Eq. (11), where it ultimately cancels the divergence of the adiabatic derivative couplings. To neglect relaxation of the dipole-matrix gradient at or near a conical intersection would be to omit crucial information.

As a concrete example of this behavior, we have partitioned the unique elements of $\boldsymbol{\mu}_{IJ}^{[Q]}$ for our distorted *p*-benzoquinone system into an unrelaxed contribution (Eq. (56)), a CPCIS contribution (from the terms in Eq. (47) proportional to $\mathbf{t}^{I[Q]}$ and $\mathbf{t}^{J[Q]}$), and a CPHF contribution (from the term proportional to $\Theta^{[Q]}$). As shown in Table VII, CPCIS accounts for nearly all of the transition-dipole gradient, and the same is true for the unitarity-gradient term.

TABLE VI. Magnitudes of the ETF-corrected derivative couplings (Eq. (B10)). induced by the atoms of *p*-benzoquinone for the adiabatic-state pair S_2/S_3 and the diabatic-state pair formed on Boys localization. The values presented are averaged over atom type; their ratio provides an effective scaling associated with the Boys diabatization.

Atom (Q)	$\|\mathbf{d}_{IJ}^{\text{ETF}[Q]}\|$ (a_0^{-1})	$\|\mathbf{d}_{AB}^{\text{ETF}[Q]}\|$ (a_0^{-1})	$\|\mathbf{d}_{AB}^{\text{ETF}[Q]}\|/\|\mathbf{d}_{IJ}^{\text{ETF}[Q]}\|$
C_A	1 041.418	0.069	15 000
C_B	589.622	0.016	37 000
O	307.772	0.118	2 610
H	60.235	0.014	4 300

TABLE VII. Magnitudes of contributions to the full transition-dipole gradient $\boldsymbol{\mu}_{IJ}^{[\mathbf{R}]}$ and unitarity-gradient term $(\sum_I U_{AI} U_{BI}^{[\mathbf{R}]})$ for the S_2/S_3 state pair of the distorted *p*-benzoquinone molecule. (See the text of Sec. IV C for more details.) The conjugated ETF-corrected derivative couplings for this system are of size $|\sum_{IJ} \mathbf{U}_{AI} \mathbf{d}_{IJ}^{\text{ETF}[\mathbf{R}]} \mathbf{U}_{BJ}| = 1,940.022\ a_0^{-1}$, and they cancel with the unitarity gradient to yield $|\mathbf{d}_{AB}^{\text{ETF}[\mathbf{R}]}| = 0.197\ a_0^{-1}$.

Quantity (units)	Unrelaxed	CPCIS	CPHF	Total
$\|\boldsymbol{\mu}_{22}^{[\mathbf{R}]}\|$ (D/Å)	6.180	17 911.827	7.604	17 914.844
$\|\boldsymbol{\mu}_{23}^{[\mathbf{R}]}\|$ (D/Å)	2.256	147.665	4.115	147.569
$\|\boldsymbol{\mu}_{33}^{[\mathbf{R}]}\|$ (D/Å)	5.393	17 936.132	5.346	17 934.350
$\|\sum_I U_{AI} U_{BI}^{[\mathbf{R}]}\|$ (a_0^{-1})	0.066	1 940.049	0.061	1 940.116

TABLE VIII. Magnitudes of contributions to the full transition-dipole gradient $\boldsymbol{\mu}_{IJ}^{[\mathbf{R}]}$ and unitarity-gradient term ($\sum_I U_{AI} U_{BI}^{[\mathbf{R}]}$) for the S_1/S_3 state pair of LiH. (See the text of Sec. IV C for more details.) The conjugated adiabatic derivative couplings for this system are of size $|\sum_{IJ} \mathbf{U}_{AI} \mathbf{d}_{IJ}^{[\mathbf{R}]} \mathbf{U}_{BJ}|$ $= 0.154\ a_0^{-1}$, and they combine with the unitarity gradient to yield $|\mathbf{d}_{AB}^{[\mathbf{R}]}|$ $= 0.195\ a_0^{-1}$.

Quantity (units)	Unrelaxed	CPCIS	CPHF	Total
$\|\boldsymbol{\mu}_{11}^{[\mathbf{R}]}\|$ (D/Å)	6.300	7.355	4.288	16.406
$\|\boldsymbol{\mu}_{13}^{[\mathbf{R}]}\|$ (D/Å)	0.842	1.669	0.364	1.804
$\|\boldsymbol{\mu}_{33}^{[\mathbf{R}]}\|$ (D/Å)	7.086	10.393	4.411	8.865
$\|\sum_I U_{AI} U_{BI}^{[\mathbf{R}]}\|$ (a_0^{-1})	0.032	0.003	0.015	0.044

To establish whether $\boldsymbol{\mu}_{IJ}^{[Q]}$ can be safely approximated by Eq. (56) in regions of configuration space far from a crossing point would require a systematic survey over a range of molecules. For LiH, the unrelaxed, CPCIS, and CPHF components of $\boldsymbol{\mu}_{IJ}^{[Q]}$ are basically comparable (cf. Table VIII), suggesting that such an approximation may always be ill-advised. It is clear, in any case, that the CPCIS terms *must* be accounted for in regions near a conical intersection, where the diabatization itself will be most important.

V. CONCLUDING REMARKS

Our work has addressed a long-standing question—are the derivative couplings between localized diabatic states actually small? For the canonical electron acceptor *p*-benzoquinone, the answer is a resounding yes. We have computed analytical derivative couplings between Boys-diabatic CIS states and shown that the diabatic derivative couplings are reduced by factors of greater than 10^3.

To arrive at this conclusion, we have formulated a comprehensive, completely general analytic-gradient theory for (quasi-)diabatic states. Furthermore, we have realized our theory for Boys diabatization (Eq. (13)) by taking the derivative of the Boys convergence condition (Eq. (14)), thus determining the Boys-specific constraint matrices of Eqs. (17) and (18). The latter matrices depend in turn on the adiabatic dipole-matrix gradient; we therefore developed the necessary analytic theory for CIS states, summarized in the boxed Eq. (49). A simple estimate suggests that the combined method of Eqs. (11), (20), and (49) incurs a computational cost comparable to 10 CIS gradients for a two-state Boys diabatization; while not insignificant, this cost is not unreasonable, either.

We have validated both our derivations and our implementation rigorously by comparison with finite-difference calculations, including those for LiH presented in Table III. We have also made a first step toward studying large systems with *p*-benzoquinone (Table VI and Figure 3). Further steps in this direction will be presented in a forthcoming paper,[33] which will detail our algorithms as implemented and apply them to the Closs energy-transfer molecules, which contain tens of atoms, as a further test of the validity of Boys-diabatic states.

Answers to many questions now seem within reach, the most immediate being: Given that the Boys diabatization appears to be nearly rigorous, what of other approaches, such as the BoysOV variant[26] or the Edmiston–Ruedenberg methods?[29, 38, 61] How small are the diabatic derivative couplings in the neighborhood of avoided crossings? Can we accurately determine the derivative couplings at conical intersections? We intend to investigate these and other intriguing avenues of research in short order.

ACKNOWLEDGMENTS

APPENDIX A: AMPLITUDE- AND ORBITAL-RESPONSE CALCULATIONS

1. Amplitude response from coupled-perturbed CIS

In order to evaluate Eq. (47), we must account for the effects of amplitude and orbital response associated with $\mathbf{t}^{[Q]}$ and $\boldsymbol{\Theta}^{[Q]}$. As such, we now detail our approach to the CPCIS and CPHF equations. We will soon see that relaxation of the CIS amplitudes will introduce additional terms into the prefactor of $\boldsymbol{\Theta}^{[Q]}$, such that we should address CPCIS first, then move on to CPHF.

The CPCIS equations can be developed completely analogously to those for CPHF;[49, 78–81] namely, we take the gradient of the variational condition for the amplitudes and solve for the amplitude response, $\mathbf{t}^{[Q]}$. In the case of CIS, the variational condition and the energy equation are identical; that is, we evaluate

$$\nabla_Q \left[\sum_{kc} (A_{\ell dkc} - \delta_{k\ell}\delta_{cd}E_I)\, t_k^{Ic} = 0 \right]. \quad (A1)$$

There is a subtle problem in Eq. (A1): The secular matrix $(\mathbf{A} - E_I)$ has a vanishing eigenvalue associated with $\mathbf{t}^I$ and is therefore not invertible. In order to solve for the CIS amplitude gradients, we must make some modification that moves the eigenvalue away from zero without altering the amplitudes themselves. Following Schaefer and co-workers,[82] we add and subtract the normalization condition of Eq. (34),

$$\nabla_Q \left\{ \sum_{kc} \left[A_{\ell dkc} - \delta_{k\ell}\delta_{cd}E_I - \delta_{k\ell}\delta_{cd}\left(1 - \sum_{ia}\left(t_i^{Ia}\right)^2\right) \right] t_k^{Ic} = 0 \right\}. \quad (A2)$$

Evaluating and rearranging the gradient, and canceling contributions to E_I from Hartree–Fock, we obtain the $3N_A$ CPCIS

equations,

$$\sum_{kc} \left[\Pi_{dk\ell c} + 2t_k^{Ic} t_\ell^{Id} + \delta_{k\ell}\delta_{cd}(\varepsilon_c - \varepsilon_k - \omega_I) \right] t_k^{Ic[Q]}$$

$$= -\sum_{ia} \left[\Pi_{di\ell a}^{[Q]} + \delta_{i\ell} F_{da}^{[Q]} - \delta_{ad} F_{\ell i}^{[Q]} - \delta_{i\ell}\delta_{ad}\omega_I^{[Q]} \right] t_i^{Ia}. \tag{A3}$$

By introducing the supermatrix $\mathcal{D}^I$ and the mixed-derivative matrix $\overline{\mathbf{M}}^{I[Q]}$, we write these equations compactly as

$$\sum_{kc} \mathcal{D}_{\ell dkc}^I t_k^{Ic[Q]} = -\overline{M}_{d\ell}^{I[Q]}. \tag{A4}$$

The various terms appearing in the mixed-derivative matrix will necessarily include contributions from orbital response, justifying our assertion that the CPCIS equations should be solved before CPHF. For a given state I, finding this solution requires that we invert the supermatrix $\mathcal{D}^I$, which has $(OV)^2$ elements. Storing an object of this size quickly becomes inconvenient in larger basis sets and for larger molecules, and inverting such an enormous matrix becomes unappealingly costly. We will instead solve Eq. (A4) using a matrix turnover (or **z**-vector) approach.[81]

We begin by defining a state-dependent, 3-vector Lagrangian,

$$\overline{\mathbf{L}}_{ai}^J \equiv \sum_{b\mu\nu} C_{\mu a}\boldsymbol{\mu}_{\mu\nu} C_{\nu b} t_i^{Jb} - \sum_{j\mu\nu} t_j^{Ja} C_{\mu j}\boldsymbol{\mu}_{\mu\nu} C_{\nu i}, \tag{A5}$$

where the overline indicates that this quantity is associated with CPCIS. The relevant portions of Eq. (47) then become

$$\text{CPCIS part of } \boldsymbol{\mu}_{IJ}^{[Q]} = \sum_{ia} \overline{\mathbf{L}}_{ai}^J t_i^{Ia[Q]} + \sum_{ia} \overline{\mathbf{L}}_{ai}^I t_i^{Ja[Q]}. \tag{A6}$$

Consider the partition of the (amplitude-dependent) Lagrangian $\overline{\mathbf{L}}^J$ into components parallel and perpendicular to $\mathbf{t}^I$,

$$\overline{\mathbf{L}}^J = \overline{\mathbf{L}}^{J\parallel} + \overline{\mathbf{L}}^{J\perp}, \tag{A7a}$$

$$\overline{\mathbf{L}}^{J\parallel} = (\mathbf{t}^I \cdot \overline{\mathbf{L}}^J)\mathbf{t}^I, \tag{A7b}$$

$$\overline{\mathbf{L}}^{J\perp} = \overline{\mathbf{L}}^J - \overline{\mathbf{L}}^{J\parallel}. \tag{A7c}$$

The gradient of the CIS normalization equation (Eq. (34)) is simply $\mathbf{t}^I \cdot \mathbf{t}^{I[Q]} = 0$, which implies $\overline{\mathbf{L}}^{J\parallel} \cdot \mathbf{t}^{I[Q]} = 0$. In other words, the parallel component of the Lagrangian cannot contribute to the final value of $\boldsymbol{\mu}_{IJ}^{[Q]}$; it can, however, destabilize the **z**-vector solution of the CPCIS equations. From this point on, all CPCIS Lagrangians appearing in our equations have implicitly been projected against the appropriate CIS amplitudes.

Continuing on, we formally invert Eq. (A4) and insert the result into Eq. (A6), thereby obtaining

$$\text{CPCIS part of } \boldsymbol{\mu}_{IJ}^{[Q]} = -\sum_{i\ell ad} \overline{\mathbf{L}}_{ai}^J \left(\mathcal{D}_{\ell dia}^I\right)^{-1} \overline{M}_{d\ell}^{I[Q]}$$

$$- \sum_{i\ell ad} \overline{\mathbf{L}}_{ai}^I \left(\mathcal{D}_{\ell dia}^J\right)^{-1} \overline{M}_{d\ell}^{J[Q]}. \tag{A8}$$

We may now define the CPCIS $\overline{\mathbf{z}}$-vector according to the usual prescription,

$$\sum_{\ell d} \mathcal{D}_{\ell dia}^I \overline{\mathbf{z}}_{d\ell}^{JI} \equiv \overline{\mathbf{L}}_{ai}^J, \tag{A9}$$

such that we have

$$\text{CPCIS part of } \boldsymbol{\mu}_{IJ}^{[Q]} = -\sum_{\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} \overline{M}_{d\ell}^{I[Q]} - \sum_{\ell d} \overline{\mathbf{z}}_{d\ell}^{IJ} \overline{M}_{d\ell}^{J[Q]}. \tag{A10}$$

In this way, we have converted the problem of inverting the CPCIS supermatrix into one of constructing the $\overline{\mathbf{z}}$-vector.

Combining Eqs. (47) and (A10) yields an intermediate expression for the dipole-matrix gradient,

$$\boldsymbol{\mu}_{IJ}^{[Q]} = \delta_{IJ}\boldsymbol{\mu}_{\text{HF}}^{[Q]} + \sum_{\mu\nu} \boldsymbol{\mu}_{\mu\nu}^{[Q]} B_{\mu\nu}^{IJS} - \sum_{\mu\nu\alpha\beta} S_{\mu\nu}^{[Q]} P_{\mu\alpha}' \boldsymbol{\mu}_{\alpha\beta} B_{\nu\beta}^{IJS}$$

$$+ \sum_{kc} \left[\begin{array}{c} \sum_{ia\mu\nu}(t_i^{Ic}t_i^{Ja} + t_i^{Jc}t_i^{Ia}) C_{\mu a}\boldsymbol{\mu}_{\mu\nu}C_{\nu k} \\ + \sum_{ia\mu\nu} C_{\mu c}\boldsymbol{\mu}_{\mu\nu}C_{\nu i}(t_i^{Ia}t_k^{Ja} + t_i^{Ja}t_k^{Ia}) \end{array} \right] \Theta_{ck}^{[Q]}$$

$$- \sum_{\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} \overline{M}_{d\ell}^{I[Q]} - \sum_{\ell d} \overline{\mathbf{z}}_{d\ell}^{IJ} \overline{M}_{d\ell}^{J[Q]}. \tag{A11}$$

We must now evaluate $\text{Tr } \overline{\mathbf{z}}\overline{\mathbf{M}}^{[Q]}$. The results are transferable between sets of state indices, so we begin by distributing the $\overline{\mathbf{z}}^{JI}$-vector among the terms of $\overline{\mathbf{M}}^{I[Q]}$,

$$\sum_{\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} \overline{M}_{d\ell}^{I[Q]} = \sum_{i\ell ad} \overline{\mathbf{z}}_{d\ell}^{JI} \Pi_{di\ell a}^{[Q]} t_i^{Ia} + \sum_{\ell ad} \overline{\mathbf{z}}_{d\ell}^{JI} F_{da}^{[Q]} t_\ell^{Ia}$$

$$- \sum_{i\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} t_i^{Id} F_{\ell i}^{[Q]} - \sum_{\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} t_\ell^{Id} \omega_I^{[Q]}. \tag{A12}$$

The somewhat complicated manipulations that follow have been consigned to Appendix C; the result is

$$\sum_{\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} \overline{M}_{d\ell}^{I[Q]}$$

$$= -\sum_{\mu\nu} h_{\mu\nu}^{[Q]} \overline{\mathbf{B}}_{\mu\nu}^{JIS} - \sum_{\mu\nu\lambda\sigma} \left(\overline{\mathbf{R}}_{\mu\lambda}^{JI} R_{\sigma\nu}^I + \overline{\mathbf{B}}_{\mu\lambda}^{JIS} P_{\sigma\nu} \right) \Pi_{\mu\nu\lambda\sigma}^{[Q]}$$

$$+ \sum_{\mu\nu\alpha\beta} S_{\mu\nu}^{[Q]} P_{\mu\alpha}' F_{\alpha\beta} \overline{\mathbf{B}}_{\nu\beta}^{JIS}$$

$$+ \frac{1}{2} \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[Q]} P_{\mu\alpha}' \left[\begin{array}{c} \overline{\mathbf{R}}_{\nu\gamma}^{JI} R_{\delta\beta}^I + R_{\nu\gamma}^I \overline{\mathbf{R}}_{\delta\beta}^{JI} \\ + \overline{\mathbf{R}}_{\gamma\nu}^{JI} R_{\beta\delta}^I + R_{\gamma\nu}^I \overline{\mathbf{R}}_{\beta\delta}^{JI} \end{array} \right] \Pi_{\alpha\beta\gamma\delta}$$

$$+ \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[Q]} P_{\mu\alpha}' P_{\nu\gamma} \Pi_{\alpha\beta\gamma\delta} \overline{\mathbf{B}}_{\delta\beta}^{JIS}$$

$$+ \sum_{kc} \left\{ \begin{array}{l} 2\sum_{\mu\nu\lambda\sigma} C_{\mu c} C_{\lambda k} \Pi_{\mu\nu\lambda\sigma} \overline{\mathbf{B}}_{\sigma\nu}^{JIS} \\ + \sum_{\mu\nu\lambda\sigma} \left[\begin{array}{c} \sum_a C_{\mu c} C_{\lambda a} \overline{\mathbf{t}}_k^{JIa} \\ -\sum_i \overline{\mathbf{t}}_i^{JIc} C_{\mu i} C_{\lambda k} \end{array} \right] \Pi_{\mu\nu\lambda\sigma} R_{\nu\sigma}^I \\ + \sum_{\mu\nu\lambda\sigma} \left[\begin{array}{c} \sum_a C_{\mu c} C_{\lambda a} t_k^{Ia} \\ -\sum_i t_i^{Ic} C_{\mu i} C_{\lambda k} \end{array} \right] \Pi_{\mu\nu\lambda\sigma} \overline{\mathbf{R}}_{\nu\sigma}^{JI} \end{array} \right\} \Theta_{ck}^{[Q]}, \tag{A13}$$

where we have introduced CPCIS-relaxed versions—to repeat, as indicated by the overline—of the CIS amplitudes, transition density, and generalized difference density.

2. Orbital response from coupled-perturbed Hartree–Fock

Inserting Eq. (A13) into Eq. (A11) gives us an expression that fully accounts for amplitude response:

$$
\boldsymbol{\mu}_{IJ}^{[Q]} = \delta_{IJ}\boldsymbol{\mu}_{\mathrm{HF}}^{[Q]} + \sum_{\mu\nu}\boldsymbol{\mu}_{\mu\nu}^{[Q]}B_{\mu\nu}^{IJS} - \sum_{\mu\nu\alpha\beta}S_{\mu\nu}^{[Q]}P'_{\mu\alpha}\boldsymbol{\mu}_{\alpha\beta}B_{\nu\beta}^{IJS}
$$

$$
+ \sum_{\mu\nu}h_{\mu\nu}^{[Q]}\left(\overline{\mathbf{B}}_{\mu\nu}^{IJS}+\overline{\mathbf{B}}_{\mu\nu}^{JIS}\right)
$$

$$
+ \sum_{\mu\nu\lambda\sigma}\left[\overline{\mathbf{R}}_{\mu\lambda}^{IJ}R_{\sigma\nu}^{J}+\overline{\mathbf{R}}_{\mu\lambda}^{JI}R_{\sigma\nu}^{I}+\left(\overline{\mathbf{B}}_{\mu\lambda}^{IJS}+\overline{\mathbf{B}}_{\mu\lambda}^{JIS}\right)P_{\sigma\nu}\right]\Pi_{\mu\nu\lambda\sigma}^{[Q]}
$$

$$
- \sum_{\mu\nu\alpha\beta}S_{\mu\nu}^{[Q]}P'_{\mu\alpha}F_{\alpha\beta}\left(\overline{\mathbf{B}}_{\nu\beta}^{IJS}+\overline{\mathbf{B}}_{\nu\beta}^{JIS}\right)
$$

$$
- \frac{1}{2}\sum_{\mu\nu\alpha\beta\gamma\delta}S_{\mu\nu}^{[Q]}P'_{\mu\alpha}
\begin{bmatrix}
\overline{\mathbf{R}}_{\nu\gamma}^{IJ}R_{\delta\beta}^{J}+R_{\nu\gamma}^{J}\overline{\mathbf{R}}_{\delta\beta}^{IJ}\\
+\overline{\mathbf{R}}_{\gamma\nu}^{IJ}R_{\beta\delta}^{J}+R_{\gamma\nu}^{J}\overline{\mathbf{R}}_{\beta\delta}^{IJ}\\
+\overline{\mathbf{R}}_{\nu\gamma}^{JI}R_{\delta\beta}^{I}+R_{\nu\gamma}^{I}\overline{\mathbf{R}}_{\delta\beta}^{JI}\\
+\overline{\mathbf{R}}_{\gamma\nu}^{JI}R_{\beta\delta}^{I}+R_{\gamma\nu}^{I}\overline{\mathbf{R}}_{\beta\delta}^{JI}
\end{bmatrix}\Pi_{\alpha\beta\gamma\delta}
$$

$$
- \sum_{\mu\nu\alpha\beta\gamma\delta}S_{\mu\nu}^{[Q]}P'_{\mu\alpha}P_{\nu\gamma}\Pi_{\alpha\beta\gamma\delta}\left(\overline{\mathbf{B}}_{\delta\beta}^{IJS}+\overline{\mathbf{B}}_{\delta\beta}^{JIS}\right)
$$

$$
- \sum_{kc}\underbrace{\left\{
\begin{array}{l}
2\sum_{\mu\nu\lambda\sigma}C_{\mu c}C_{\lambda k}\Pi_{\mu\nu\lambda\sigma}\left(\overline{\mathbf{B}}_{\sigma\nu}^{IJS}+\overline{\mathbf{B}}_{\sigma\nu}^{JIS}\right)\\[6pt]
+\sum_{\mu\nu\lambda\sigma}\left[\begin{array}{l}\sum_a C_{\mu c}C_{\lambda a}\bar{\mathbf{t}}_k^{JIa}\\-\sum_i \bar{\mathbf{t}}_i^{JIc}C_{\mu i}C_{\lambda k}\end{array}\right]\Pi_{\mu\nu\lambda\sigma}R_{\nu\sigma}^{I}\\[12pt]
+\sum_{\mu\nu\lambda\sigma}\left[\begin{array}{l}\sum_a C_{\mu c}C_{\lambda a}\bar{\mathbf{t}}_k^{IJa}\\-\sum_i \bar{\mathbf{t}}_i^{IJc}C_{\mu i}C_{\lambda k}\end{array}\right]\Pi_{\mu\nu\lambda\sigma}R_{\nu\sigma}^{J}\\[12pt]
+\sum_{\mu\nu\lambda\sigma}\left[\begin{array}{l}\sum_a C_{\mu c}C_{\lambda a}t_k^{Ja}\\-\sum_i t_i^{Jc}C_{\mu i}C_{\lambda k}\end{array}\right]\Pi_{\mu\nu\lambda\sigma}\overline{\mathbf{R}}_{\nu\sigma}^{IJ}\\[12pt]
+\sum_{\mu\nu\lambda\sigma}\left[\begin{array}{l}\sum_a C_{\mu c}C_{\lambda a}t_k^{Ia}\\-\sum_i t_i^{Ic}C_{\mu i}C_{\lambda k}\end{array}\right]\Pi_{\mu\nu\lambda\sigma}\overline{\mathbf{R}}_{\nu\sigma}^{JI}\\[12pt]
-\sum_{ia\mu\nu}\left(t_i^{Ic}t_i^{Ja}+t_i^{Jc}t_i^{Ia}\right)C_{\mu a}\boldsymbol{\mu}_{\mu\nu}C_{\nu k}\\[6pt]
-\sum_{ia\mu\nu}C_{\mu c}\boldsymbol{\mu}_{\mu\nu}C_{\nu i}\left(t_i^{Ia}t_k^{Ja}+t_i^{Ja}t_k^{Ia}\right)
\end{array}
\right\}}_{\widetilde{\mathbf{L}}_{ck}^{IJ}}\Theta_{ck}^{[Q]}.
$$

(A14)

As indicated by the presence of $\Theta^{[Q]}$ and our implicit definition in the underbrace of the CPHF Lagrangian—denoted with a wide tilde—we must still incorporate the effects of orbital response. $\widetilde{\mathbf{L}}^{JI}=\widetilde{\mathbf{L}}^{IJ}$, such that we can use them interchangeably; it is also true that $\widetilde{\mathbf{z}}^{JI}=\widetilde{\mathbf{z}}^{IJ}$. Note, however, that $\widetilde{\mathbf{z}}_{\mu\nu}^{JI}\neq\widetilde{\mathbf{z}}_{\nu\mu}^{JI}$.

The CPHF equations are derived from the gradient of the Hartree–Fock variational condition,[78,79]

$$
\nabla_Q\left(\frac{\partial E}{\partial \Theta_{ai}}=-F_{ai}-F_{ia}=0\right).
$$

(A15)

Inserting the Fock-matrix derivative of Eq. (B2), collecting terms, and permuting indices, we obtain the $3N_A$ CPHF equations,

$$
\sum_{kc}2\left[\Pi_{caki}+\Pi_{cika}+\delta_{ik}\delta_{ac}\left(\varepsilon_c-\varepsilon_k\right)\right]\Theta_{ck}^{[Q]}
$$

$$
= -\sum_{\mu\nu}h_{\mu\nu}^{[Q]}(C_{\mu a}C_{\nu i}+C_{\mu i}C_{\nu a})
$$

$$
- \sum_{\mu\nu\lambda\sigma}(C_{\mu a}C_{\lambda i}+C_{\mu i}C_{\lambda a})\Pi_{\mu\nu\lambda\sigma}^{[Q]}P_{\sigma\nu}
$$

$$
+ \sum_{\mu\nu\alpha\beta}S_{\mu\nu}^{[Q]}P'_{\mu\alpha}F_{\alpha\beta}(C_{\beta a}C_{\nu i}+C_{\beta i}C_{\nu a})
$$

$$
+ \sum_{\mu\nu\alpha\beta\gamma\delta}S_{\mu\nu}^{[Q]}P'_{\mu\alpha}P_{\nu\gamma}\Pi_{\alpha\beta\gamma\delta}(C_{\delta a}C_{\nu i}+C_{\delta i}C_{\nu a}).
$$

(A16)

This equation can be rewritten to obtain the standard form in terms of the usual CPHF supermatrix ($\partial^2 E/\partial\Theta^2$) and the corresponding mixed-derivative matrix $\mathbf{M}^{[Q]}$,

$$
\sum_{kc}\left(\frac{\partial^2 E}{\partial\Theta_{ai}\partial\Theta_{ck}}\right)\Theta_{ck}^{[Q]}=-M_{ai}^{[Q]}.
$$

(A17)

As in the CPCIS case, we invoke the formal inversion of Eq. (A17); define the $\widetilde{\mathbf{z}}$-vector in terms of a CPHF Lagrangian,

$$
\sum_{ia}\left(\frac{\partial^2 E}{\partial\Theta_{ai}\partial\Theta_{ck}}\right)\widetilde{\mathbf{z}}_{ai}^{IJ}\equiv\widetilde{\mathbf{L}}_{ck}^{IJ};
$$

(A18)

and make the identification

$$
\text{CPHF part of } \boldsymbol{\mu}_{IJ}^{[Q]}=-\sum_{kc}\widetilde{\mathbf{L}}_{ck}^{IJ}\Theta_{ck}^{[Q]}=\sum_{ia}\widetilde{\mathbf{z}}_{ai}^{IJ}M_{ai}^{[Q]}.
$$

(A19)

Simple manipulations involving the symmetry rules of Eq. (6) quickly lead us to an explicit expression for the CPHF contribution to the dipole-matrix gradient,

$$
\text{CPHF part of } \boldsymbol{\mu}_{IJ}^{[Q]}
$$

$$
= -2\sum_{\mu\nu}h_{\mu\nu}^{[Q]}\widetilde{\mathbf{z}}_{\mu\nu}^{IJS}-2\sum_{\mu\nu\lambda\sigma}\widetilde{\mathbf{z}}_{\mu\lambda}^{IJS}\Pi_{\mu\nu\lambda\sigma}^{[Q]}P_{\sigma\nu}
$$

$$
+ 2\sum_{\mu\nu\alpha\beta}S_{\mu\nu}^{[Q]}P'_{\mu\alpha}F_{\alpha\beta}\widetilde{\mathbf{z}}_{\nu\beta}^{IJS}
$$

$$
+ 2\sum_{\mu\nu\alpha\beta\gamma\delta}S_{\mu\nu}^{[Q]}P'_{\mu\alpha}P_{\nu\gamma}\Pi_{\alpha\beta\gamma\delta}\widetilde{\mathbf{z}}_{\delta\beta}^{IJS}.
$$

(A20)

Referring to Eq. (A14), we find that we subtract $\widetilde{\mathbf{z}}^{IJS}=\widetilde{\mathbf{z}}^{JIS}$ from each instance of $\overline{\mathbf{B}}^S$. We therefore define a fully relaxed (and totally symmetric) version of the difference density,

$$
\widetilde{\mathbf{B}}_{\mu\nu}^{JI}=\overline{\mathbf{B}}_{\mu\nu}^{JIS}-\widetilde{\mathbf{z}}_{\mu\nu}^{JIS}.
$$

(A21)

Applying Eqs. (A19)–(A21) gives us the final analytic expression for the CIS dipole-matrix gradient, namely, Eq. (49).

APPENDIX B: CONSTITUENTS OF THE CPCIS MIXED-DERIVATIVE MATRIX

The CPCIS mixed-derivative matrix (initially defined in Eq. (A4)) is given by

$$\overline{M}_{d\ell}^{I[Q]} = \sum_{ia} \left[\Pi_{di\ell a}^{[Q]} + \delta_{i\ell} F_{da}^{[Q]} - \delta_{ad} F_{\ell i}^{[Q]} - \delta_{i\ell}\delta_{ad}\omega_I^{[Q]} \right] t_i^{Ia}. \tag{B1}$$

In this Appendix, we provide expressions for the bracketed gradients, as well as for the adiabatic CIS derivative couplings. We conclude with a brief discussion of translational-symmetry breaking in the derivative couplings and how it can be corrected.

We begin by quoting $\Pi_{di\ell a}^{[Q]}$ and $F_{pq}^{[Q]}$ directly from Eqs. (A15) and (A16) of our previous work,[54] rewriting them as necessary to fit our current notational conventions:

$$F_{pq}^{[Q]} = \sum_{\mu\nu} C_{\mu p} h_{\mu\nu}^{[Q]} C_{\nu q} + \sum_{\mu\nu\lambda\sigma} C_{\mu p} C_{\lambda q} \Pi_{\mu\nu\lambda\sigma}^{[Q]} P_{\sigma\nu}$$
$$- \frac{1}{2}\sum_{\mu\nu\alpha\beta} S_{\mu\nu}^{[Q]} P_{\mu\alpha}' F_{\alpha\beta}(C_{\beta p} C_{\nu q} + C_{\beta q} C_{\nu p})$$
$$- \frac{1}{2}\sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[Q]} P_{\mu\alpha}' P_{\nu\gamma} \Pi_{\alpha\beta\gamma\delta}(C_{\delta p} C_{\beta q} + C_{\delta q} C_{\beta p})$$
$$- \sum_{kc}(\Pi_{pcqk} + \Pi_{qcpk})\Theta_{ck}^{[Q]} - (\varepsilon_p - \varepsilon_q)\Theta_{pq}^{[Q]}, \tag{B2}$$

where we discard the last term if p and q are drawn from the same subspace, and

$$\Pi_{di\ell a}^{[Q]}$$
$$= \sum_{\mu\nu\lambda\sigma} C_{\mu d} C_{\lambda\ell} \Pi_{\mu\nu\lambda\sigma}^{[Q]} C_{\sigma a} C_{\nu i}$$
$$- \frac{1}{2}\sum_{s\mu\nu} C_{\mu s} S_{\mu\nu}^{[Q]}[C_{\nu d}\Pi_{si\ell a} + C_{\nu i}\Pi_{ds\ell a}$$
$$+ C_{\nu\ell}\Pi_{disa} + C_{\nu a}\Pi_{di\ell s}]$$
$$+ \sum_{kc}[\delta_{cd}\Pi_{ki\ell a} - \delta_{ik}\Pi_{dc\ell a} - \delta_{k\ell}\Pi_{dica} + \delta_{ac}\Pi_{di\ell k}]\Theta_{ck}^{[Q]}. \tag{B3}$$

Reference 54 introduced a new derivation of the (adiabatic) CIS derivative couplings. Starting from the premise that the CIS states satisfy the time-independent Schrödinger equation for the singles-projected Hamiltonian $H' = \sum_{k\ell cd} |\Phi_\ell^d\rangle A_{\ell dkc}\langle\Phi_k^c|$,

$$H'|\Psi_I\rangle = E_I|\Psi_I\rangle, \tag{B4}$$

the Hellmann–Feynman theorem was invoked to obtain an expression of form

$$\mathbf{d}_{IJ}^{[Q]} = \frac{1}{E_J - E_I}\left[\langle\Psi_I|H^{[Q]}|\Psi_J\rangle - \sum_{kc} Y_{ck}^{IJ}\Theta_{ck}^{[Q]}\right]$$
$$+ \sum_{\mu\nu} S_{\mu\nu}^{A[Q]} D_{\mu\nu}^{IJ}. \tag{B5}$$

Exactly the same reasoning—apart from the fact that the Schrödinger equation is now written as $(H' - E)|\Psi_I\rangle = \omega_I|\Psi_I\rangle$—may be used to write down the CIS excitation-energy gradient,

$$\omega_I^{[Q]} = \langle\Psi_I|H^{[Q]}|\Psi_I\rangle - \sum_{kc} Y_{ck}^{II}\Theta_{ck}^{[Q]}. \tag{B6}$$

In Eqs. (B5) and (B6), the (bare) Hamiltonian derivative is

$$\langle\Psi_I|H^{[Q]}|\Psi_I\rangle$$
$$= \sum_{\mu\nu} h_{\mu\nu}^{[Q]} B_{\mu\nu}^{II} + \sum_{\mu\nu\lambda\sigma}\left(R_{\mu\lambda}^I R_{\sigma\nu}^I + B_{\mu\lambda}^{II} P_{\sigma\nu}\right)\Pi_{\mu\nu\lambda\sigma}^{[Q]}$$
$$- \sum_{\mu\nu\alpha\beta} S_{\mu\nu}^{[Q]} P_{\mu\alpha}' F_{\alpha\beta} B_{\nu\beta}^{II}$$
$$- \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[Q]} P_{\mu\alpha}'\left(R_{\nu\gamma}^I R_{\delta\beta}^I + R_{\gamma\nu}^I R_{\beta\delta}^I\right)\Pi_{\alpha\beta\gamma\delta}$$
$$- \sum_{\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[Q]} P_{\mu\alpha}' P_{\nu\gamma}\Pi_{\alpha\beta\gamma\delta} B_{\delta\beta}^{II}, \tag{B7}$$

and the Lagrangian for the orbital response is

$$Y_{ck}^{IJ} = 2\sum_{\mu\nu\lambda\sigma} C_{\mu c} C_{\lambda k}\Pi_{\mu\nu\lambda\sigma} B_{\sigma\nu}^{IJS}$$
$$+ \sum_{\mu\nu\lambda\sigma}\left[\sum_a C_{\mu c} C_{\lambda a} t_k^{Ja} - \sum_i t_i^{Jc} C_{\mu i} C_{\lambda k}\right]\Pi_{\mu\nu\lambda\sigma} R_{\nu\sigma}^I$$
$$+ \sum_{\mu\nu\lambda\sigma}\left[\sum_a C_{\mu c} C_{\lambda a} t_k^{Ia} - \sum_i t_i^{Ic} C_{\mu i} C_{\lambda k}\right]\Pi_{\mu\nu\lambda\sigma} R_{\nu\sigma}^J. \tag{B8}$$

The antisymmetric form of the overlap matrix appearing only in Eq. (B5) is defined as

$$S_{\mu\nu}^{A[Q]} \equiv \frac{1}{2}\left(\langle\varphi_\mu|\varphi_\nu^{[Q]}\rangle - \langle\varphi_\nu|\varphi_\mu^{[Q]}\rangle\right). \tag{B9}$$

As discussed in our previous papers,[54,77] the associated term arises from the failure of adiabatic eigenstates to account for tandem motion of the electrons and nuclei, and it breaks the translational symmetry and momentum conservation required of the corresponding dynamics. (The unitarity-gradient term in Eq. (11), by contrast, respects translational symmetry.) Our method of perturbative electron-translation factors (ETFs) rectifies this unphysical behavior at no additional computational expense; the operational prescription for obtaining ETF-corrected derivative couplings is simply to discard the term containing $S^{A[Q]}$ from Eq. (B5). In other words,

$$\mathbf{d}_{IJ}^{ETF[Q]} = \mathbf{d}_{IJ}^{[Q]} - \sum_{\mu\nu} S_{\mu\nu}^{A[Q]} D_{\mu\nu}^{IJ}. \tag{B10}$$

In validating our implementation against LiH finite-difference results in Section IV, we naturally used the analytical derivative couplings given by Eq. (B5). As a general rule, however, we recommend the use of the ETF-corrected couplings, which ensure physical behavior in the associated dynamics. For the distorted p-benzoquinone system of Section IV, the ETF-corrected adiabatic derivative couplings are identical to

their uncorrected counterparts within fractions of a percent. But because the associated diabatic derivative couplings are themselves quite small, the effect of the correction is qualitatively apparent. Hence, we label the *p*-benzoquinone derivative couplings with a superscript *ETF*.

APPENDIX C: DERIVATION OF Tr $\overline{\mathbf{z}}\overline{\mathbf{M}}^{[Q]}$ FOR $\mu_{IJ}^{[Q]}$

In this Appendix, we evaluate the trace of the CPCIS $\overline{\mathbf{z}}$-vector against the mixed-derivative matrix $\overline{\mathbf{M}}^{[Q]}$, starting from Eq. (A12) (which we repeat here):

$$\sum_{\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} \overline{M}_{d\ell}^{I[Q]} = \sum_{i\ell ad} \overline{\mathbf{z}}_{d\ell}^{JI} \Pi_{di\ell a}^{[Q]} t_i^{Ia} + \sum_{\ell ad} \overline{\mathbf{z}}_{d\ell}^{JI} F_{da}^{[Q]} t_\ell^{Ia}$$
$$- \sum_{i\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} t_i^{Id} F_{\ell i}^{[Q]} - \sum_{\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} t_\ell^{Id} \omega_I^{[Q]}. \quad \text{(C1)}$$

In doing so, we will draw expressions for the various constituent pieces of $\overline{\mathbf{M}}^{[Q]}$ from Appendix B. First, we plug in the Fock-matrix and two-electron-integral derivatives (Eqs. (B2) and (B3)) to yield a rather complicated equation,

$$\sum_{\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} \overline{M}_{d\ell}^{I[Q]} = \sum_{i\ell ad\mu\nu\lambda\sigma} C_{\mu d} \overline{\mathbf{z}}_{d\ell}^{JI} C_{\lambda \ell} \Pi_{\mu\nu\lambda\sigma}^{[Q]} C_{\sigma a} t_i^{Ia} C_{\nu i}$$

$$- \frac{1}{2} \sum_{i\ell sad\mu\nu} C_{\mu s} S_{\mu\nu}^{[Q]} \overline{\mathbf{z}}_{d\ell}^{JI} \left[\begin{matrix} C_{\nu d} \Pi_{si\ell a} + C_{\nu i} \Pi_{ds\ell a} \\ + C_{\nu \ell} \Pi_{disa} + C_{\nu a} \Pi_{di\ell s} \end{matrix} \right] t_i^{Ia}$$

$$+ \sum_{kc} \left[\begin{matrix} \sum_{i\ell a} \overline{\mathbf{z}}_{c\ell}^{JI} \Pi_{ki\ell a} t_i^{Ia} - \sum_{\ell ad} \overline{\mathbf{z}}_{d\ell}^{JI} \Pi_{dc\ell a} t_k^{Ia} \\ - \sum_{iad} \overline{\mathbf{z}}_{dk}^{JI} \Pi_{dica} t_i^{Ia} + \sum_{i\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} \Pi_{di\ell k} t_i^{Ic} \end{matrix} \right] \Theta_{ck}^{[Q]}$$

$$+ \sum_{\ell ad\mu\nu} h_{\mu\nu}^{[Q]} C_{\mu d} \overline{\mathbf{z}}_{d\ell}^{JI} t_\ell^{Ia} C_{\nu a} + \sum_{\ell ad\mu\nu\lambda\sigma} C_{\mu d} \overline{\mathbf{z}}_{d\ell}^{JI} t_\ell^{Ia} C_{\lambda a} \Pi_{\mu\nu\lambda\sigma}^{[Q]} P_{\sigma\nu}$$

$$- \frac{1}{2} \sum_{\ell ad\mu\nu\alpha\beta} S_{\mu\nu}^{[Q]} P'_{\mu\alpha} F_{\alpha\beta} \overline{\mathbf{z}}_{d\ell}^{JI} \left(C_{\beta d} C_{\nu a} + C_{\beta a} C_{\nu d} \right) t_\ell^{Ia}$$

$$- \frac{1}{2} \sum_{\ell ad\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[Q]} P'_{\mu\alpha} P_{\nu\gamma} \Pi_{\alpha\beta\gamma\delta} \overline{\mathbf{z}}_{d\ell}^{JI} \left(C_{\delta d} C_{\beta a} + C_{\delta a} C_{\beta d} \right) t_\ell^{Ia}$$

$$- \sum_{kc} \left(\sum_{\ell ad} \overline{\mathbf{z}}_{d\ell}^{JI} \Pi_{dcak} t_\ell^{Ia} + \sum_{\ell ad} \overline{\mathbf{z}}_{d\ell}^{JI} \Pi_{acdk} t_\ell^{Ia} \right) \Theta_{ck}^{[Q]}$$

$$- \sum_{i\ell d\mu\nu} h_{\mu\nu}^{[Q]} C_{\mu\ell} \overline{\mathbf{z}}_{d\ell}^{JI} t_i^{Id} C_{\nu i} - \sum_{i\ell d\mu\nu\lambda\sigma} C_{\mu\ell} \overline{\mathbf{z}}_{d\ell}^{JI} t_i^{Id} C_{\lambda i} \Pi_{\mu\nu\lambda\sigma}^{[Q]} P_{\sigma\nu}$$

$$+ \frac{1}{2} \sum_{i\ell d\mu\nu\alpha\beta} S_{\mu\nu}^{[Q]} P'_{\mu\alpha} F_{\alpha\beta} \overline{\mathbf{z}}_{d\ell}^{JI} \left(C_{\beta\ell} C_{\nu i} + C_{\beta i} C_{\nu\ell} \right) t_i^{Id}$$

$$+ \frac{1}{2} \sum_{i\ell d\mu\nu\alpha\beta\gamma\delta} S_{\mu\nu}^{[Q]} P'_{\mu\alpha} P_{\nu\gamma} \Pi_{\alpha\beta\gamma\delta} \overline{\mathbf{z}}_{d\ell}^{JI} \left(C_{\delta\ell} C_{\beta i} + C_{\delta i} C_{\beta\ell} \right) t_i^{Id}$$

$$+ \sum_{kc} \left(\sum_{i\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} \Pi_{\ell cik} t_i^{Id} + \sum_{i\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} \Pi_{ic\ell k} t_i^{Id} \right) \Theta_{ck}^{[Q]}$$

$$- \sum_{\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} t_\ell^{Id} \omega_I^{[Q]}. \quad \text{(C2)}$$

We can sift through its various terms in order to identify a few helpful definitions, all of which involve the $\overline{\mathbf{z}}$-vector in combination with CIS amplitudes and/or MO coefficients:

$$\overline{\mathbf{z}}_{\mu\nu}^{JI} \equiv \sum_{\ell d} C_{\mu d} \overline{\mathbf{z}}_{d\ell}^{JI} C_{\nu\ell}, \quad \text{(C3a)}$$

$$\zeta_{\mu\nu}^{JI} \equiv \sum_{\ell ad} C_{\mu d} \overline{\mathbf{z}}_{d\ell}^{JI} t_\ell^{Ia} C_{\nu a}, \quad \text{(C3b)}$$

$$\zeta_{\mu\nu}'^{JI} \equiv \sum_{i\ell d} C_{\mu\ell} \overline{\mathbf{z}}_{d\ell}^{JI} t_i^{Id} C_{\nu i}, \quad \text{(C3c)}$$

$$\eta^{JI} \equiv \sum_{\ell d} \overline{\mathbf{z}}_{d\ell}^{JI} t_\ell^{Id}. \quad \text{(C3d)}$$

A point that deserves emphasis is that $\boldsymbol{\eta}^{JI}$ is a *spinless* 3-vector in Cartesian space—we must trace the spin blocks of $\overline{\mathbf{z}}^{JI}$ and $\mathbf{t}^I$ separately and add the results to obtain $\boldsymbol{\eta}^{JI}$.

We now insert as many definitions as we can from Eq. (C3) in order to simplify Eq. (C2), collecting terms and permuting indices so that we obtain as compact an expression

as possible. In particular, we aim to have $P'_{\mu\alpha}$ and $C_{\mu c}C_{\lambda k}$ (or the equivalent combination of $\mathbf{C}$, $\bar{\mathbf{z}}^{JI}$, and $\mathbf{t}^I$) throughout. The result is

$$\sum_{\ell d} \bar{\mathbf{z}}^{JI}_{d\ell} \overline{M}^{I[Q]}_{d\ell}$$

$$= -\sum_{\mu\nu} h^{[Q]}_{\mu\nu} \left(-\zeta^{JI}_{\mu\nu} + \zeta'^{JI}_{\mu\nu} \right)$$

$$- \sum_{\mu\nu\lambda\sigma} \left[-\bar{\mathbf{z}}^{JI}_{\mu\lambda} R^I_{\sigma\nu} + \left(-\zeta^{JI}_{\mu\lambda} + \zeta'^{JI}_{\mu\lambda} \right) P_{\sigma\nu} \right] \Pi^{[Q]}_{\mu\nu\lambda\sigma}$$

$$+ \sum_{\mu\nu\alpha\beta} S^{[Q]}_{\mu\nu} P'_{\mu\alpha} F_{\alpha\beta} \left(-\zeta^{JIS}_{\nu\beta} + \zeta'^{JIS}_{\nu\beta} \right)$$

$$- \frac{1}{2} \sum_{i\ell sad\mu\nu} C_{\mu s} S^{[Q]}_{\mu\nu} \bar{\mathbf{z}}^{JI}_{d\ell} \left[\begin{array}{c} C_{\nu d}\Pi_{si\ell a} + C_{\nu i}\Pi_{sda\ell} \\ + C_{\nu\ell}\Pi_{sadi} + C_{\nu a}\Pi_{s\ell id} \end{array} \right] t^{Ia}_i$$

$$+ \sum_{\mu\nu\alpha\beta\gamma\delta} S^{[Q]}_{\mu\nu} P'_{\mu\alpha} P_{\nu\gamma} \Pi_{\alpha\beta\gamma\delta} \left(-\zeta^{JIS}_{\delta\beta} + \zeta'^{JIS}_{\delta\beta} \right)$$

$$+ \sum_{kc} \left[\begin{array}{c} \sum_{i\ell a}\bar{\mathbf{z}}^{JI}_{c\ell}\Pi_{\ell aki}t^{Ia}_i - \sum_{\ell ad}t^{Ia}_k\Pi_{cda\ell}\bar{\mathbf{z}}^{JI}_{d\ell} \\ -\sum_{\ell ad}\Pi_{cdka}\bar{\mathbf{z}}^{JI}_{d\ell}t^{Ia}_\ell - \sum_{\ell ad}\Pi_{cakd}t^{Ia}_\ell\bar{\mathbf{z}}^{JI}_{d\ell} \\ -\sum_{iad}\bar{\mathbf{z}}^{JI}_{dk}\Pi_{cadi}t^{Ia}_i + \sum_{i\ell d}t^{Ic}_i\Pi_{idk\ell}\bar{\mathbf{z}}^{JI}_{d\ell} \\ +\sum_{i\ell d}\Pi_{c\ell ki}\bar{\mathbf{z}}^{JI}_{d\ell}t^{Id}_i + \sum_{i\ell d}\Pi_{cik\ell}t^{Id}_i\bar{\mathbf{z}}^{JI}_{d\ell} \end{array} \right] \Theta^{[Q]}_{ck}$$

$$- \eta^{JI}\omega^{[Q]}_I. \tag{C4}$$

We must now find the atomic-orbital representation of the various contractions appearing in Eq. (C4). After applying the identity

$$\Pi_{pqst} = \sum_{\mu\nu\lambda\sigma} C_{\mu p} C_{\nu q} \Pi_{\mu\nu\lambda\sigma} C_{\lambda s} C_{\sigma t}, \tag{C5}$$

we once again invoke definitions from Eq. (C3), collect terms, and permute indices, yielding

$$\sum_{\ell d} \bar{\mathbf{z}}^{JI}_{d\ell} \overline{M}^{I[Q]}_{d\ell} = -\sum_{\mu\nu} h^{[Q]}_{\mu\nu} \left(-\zeta^{JI}_{\mu\nu} + \zeta'^{JI}_{\mu\nu} \right)$$

$$- \sum_{\mu\nu\lambda\sigma} \left[-\bar{\mathbf{z}}^{JI}_{\mu\lambda} R^I_{\sigma\nu} + \left(-\zeta^{JI}_{\mu\lambda} + \zeta'^{JI}_{\mu\lambda} \right) P_{\sigma\nu} \right] \Pi^{[Q]}_{\mu\nu\lambda\sigma}$$

$$+ \sum_{\mu\nu\alpha\beta} S^{[Q]}_{\mu\nu} P'_{\mu\alpha} F_{\alpha\beta} \left(-\zeta^{JIS}_{\nu\beta} + \zeta'^{JIS}_{\nu\beta} \right)$$

$$- \frac{1}{2} \sum_{\mu\nu\alpha\beta\gamma\delta} S^{[Q]}_{\mu\nu} P'_{\mu\alpha} \left[\begin{array}{c} \bar{\mathbf{z}}^{JI}_{\nu\gamma} R^I_{\delta\beta} + R^I_{\nu\gamma}\bar{\mathbf{z}}^{JI}_{\delta\beta} \\ +\bar{\mathbf{z}}^{JI}_{\gamma\nu} R^I_{\beta\delta} + R^I_{\gamma\nu}\bar{\mathbf{z}}^{JI}_{\beta\delta} \end{array} \right] \Pi_{\alpha\beta\gamma\delta}$$

$$+ \sum_{\mu\nu\alpha\beta\gamma\delta} S^{[Q]}_{\mu\nu} P'_{\mu\alpha} P_{\nu\gamma} \Pi_{\alpha\beta\gamma\delta} \left(-\zeta^{JIS}_{\delta\beta} + \zeta'^{JIS}_{\delta\beta} \right)$$

$$+ \sum_{kc} \left\{ \begin{array}{c} 2\sum_{\mu\nu\lambda\sigma} C_{\mu c}C_{\lambda k}\Pi_{\mu\nu\lambda\sigma}\left(-\zeta^{JIS}_{\sigma\nu} + \zeta'^{JIS}_{\sigma\nu} \right) \\ -\sum_{\mu\nu\lambda\sigma}\left[\begin{array}{c} \sum_a C_{\mu c}C_{\lambda a}\bar{\mathbf{z}}^{JI}_{ak} \\ -\sum_i \bar{\mathbf{z}}^{JI}_{ci}C_{\mu i}C_{\lambda k} \end{array} \right]\Pi_{\mu\nu\lambda\sigma}R^I_{\nu\sigma} \\ -\sum_{\mu\nu\lambda\sigma}\left[\begin{array}{c} \sum_a C_{\mu c}C_{\lambda a}t^{Ia}_k \\ -\sum_i t^{Ic}_i C_{\mu i}C_{\lambda k} \end{array} \right]\Pi_{\mu\nu\lambda\sigma}\bar{\mathbf{z}}^{JI}_{\nu\sigma} \end{array} \right\} \Theta^{[Q]}_{ck}$$

$$- \eta^{JI}\omega^{[Q]}_I. \tag{C6}$$

We *could* consider Eq. (C6) final and use it as the basis for our working expressions; after all, quantum-chemistry packages are typically capable of efficiently calculating the CIS excitation-energy gradient $\omega^{[Q]}_I$. What would be gained in simple expressions would be lost, however, in unnecessary contractions—especially the costly contractions involving $\mathbf{\Pi}^{[Q]}$. As such, we press on and substitute Eqs. (B6)–(B8) for the excitation-energy gradient into Eq. (C6), then collect terms once again. We thereby obtain a complete expression,

$$\sum_{\ell d} \bar{\mathbf{z}}^{JI}_{d\ell} \overline{M}^{I[Q]}_{d\ell}$$

$$= -\sum_{\mu\nu} h^{[Q]}_{\mu\nu} \left(\eta^{JI} B^{II}_{\mu\nu} - \zeta^{JI}_{\mu\nu} + \zeta'^{JI}_{\mu\nu} \right)$$

$$- \sum_{\mu\nu\lambda\sigma} \left[\begin{array}{c} \left(\eta^{JI} R^I_{\mu\lambda} - \bar{\mathbf{z}}^{JI}_{\mu\lambda} \right) R^I_{\sigma\nu} \\ + \left(\eta^{JI} B^{II}_{\mu\lambda} - \zeta^{JI}_{\mu\lambda} + \zeta'^{JI}_{\mu\lambda} \right) P_{\sigma\nu} \end{array} \right] \Pi^{[Q]}_{\mu\nu\lambda\sigma}$$

$$+ \sum_{\mu\nu\alpha\beta} S^{[Q]}_{\mu\nu} P'_{\mu\alpha} F_{\alpha\beta} \left(\eta^{JI} B^{II}_{\nu\beta} - \zeta^{JIS}_{\nu\beta} + \zeta'^{JIS}_{\nu\beta} \right)$$

$$+ \frac{1}{2} \sum_{\mu\nu\alpha\beta\gamma\delta} S^{[Q]}_{\mu\nu} P'_{\mu\alpha} \left[\begin{array}{c} 2\eta^{JI} R^I_{\nu\gamma} R^I_{\delta\beta} - \bar{\mathbf{z}}^{JI}_{\nu\gamma} R^I_{\delta\beta} - R^I_{\nu\gamma}\bar{\mathbf{z}}^{JI}_{\delta\beta} \\ +2\eta^{JI} R^I_{\gamma\nu} R^I_{\beta\delta} - \bar{\mathbf{z}}^{JI}_{\gamma\nu} R^I_{\beta\delta} - R^I_{\gamma\nu}\bar{\mathbf{z}}^{JI}_{\beta\delta} \end{array} \right] \Pi_{\alpha\beta\gamma\delta}$$

$$+ \sum_{\mu\nu\alpha\beta\gamma\delta} S^{[Q]}_{\mu\nu} P'_{\mu\alpha} P_{\nu\gamma} \Pi_{\alpha\beta\gamma\delta} \left(\eta^{JI} B^{II}_{\delta\beta} - \zeta^{JIS}_{\delta\beta} + \zeta'^{JIS}_{\delta\beta} \right)$$

$$+ \sum_{kc} \left\{ \begin{array}{c} 2\sum_{\mu\nu\lambda\sigma} C_{\mu c}C_{\lambda k}\Pi_{\mu\nu\lambda\sigma}\left(\eta^{JI} B^{II}_{\sigma\nu} - \zeta^{JIS}_{\sigma\nu} + \zeta'^{JIS}_{\sigma\nu} \right) \\ +2\eta^{JI}\sum_{\mu\nu\lambda\sigma}\left[\begin{array}{c} \sum_a C_{\mu c}C_{\lambda a}t^{Ia}_k \\ -\sum_i t^{Ic}_i C_{\mu i}C_{\lambda k} \end{array} \right]\Pi_{\mu\nu\lambda\sigma}R^I_{\nu\sigma} \\ -\sum_{\mu\nu\lambda\sigma}\left[\begin{array}{c} \sum_a C_{\mu c}C_{\lambda a}\bar{\mathbf{z}}^{JI}_{ak} \\ -\sum_i \bar{\mathbf{z}}^{JI}_{ci}C_{\mu i}C_{\lambda k} \end{array} \right]\Pi_{\mu\nu\lambda\sigma}R^I_{\nu\sigma} \\ -\sum_{\mu\nu\lambda\sigma}\left[\begin{array}{c} \sum_a C_{\mu c}C_{\lambda a}t^{Ia}_k \\ -\sum_i t^{Ic}_i C_{\mu i}C_{\lambda k} \end{array} \right]\Pi_{\mu\nu\lambda\sigma}\bar{\mathbf{z}}^{JI}_{\nu\sigma} \end{array} \right\} \Theta^{[Q]}_{ck}. $$

$$\tag{C7}$$

While the above expression is complete, it is not especially intuitive. To reveal more of its physical content, we define three new quantities—the CPCIS-relaxed amplitudes, transition density, and difference density:

$$\bar{\mathbf{t}}^{JIc}_k \equiv \eta^{JI} t^{Ic}_k - \bar{\mathbf{z}}^{JI}_{ck}, \tag{C8a}$$

$$\overline{\mathbf{R}}^{JI}_{\mu\nu} \equiv \eta^{JI} R^I_{\mu\nu} - \bar{\mathbf{z}}^{JI}_{\mu\nu}, \tag{C8b}$$

$$\overline{\mathbf{B}}^{JI}_{\mu\nu} \equiv \eta^{JI} B^{II}_{\mu\nu} - \zeta^{JI}_{\mu\nu} + \zeta'^{JI}_{\mu\nu}. \tag{C8c}$$

In this context, it becomes apparent that ζ^{JI} and ζ'^{JI} can be understood as separately relaxing the rescaled attachment and detachment densities.

There are two other simplifications that we can identify: First, $\mathbf{h}^{[Q]}$ is symmetric, such that we can replace $\overline{\mathbf{B}}^{JI}$ with $\overline{\mathbf{B}}^{JIS}$ in that term. Second, the symmetry of $\mathbf{P}$ allows us to make the same replacement in the $\overline{\mathbf{B}}^{JI} \cdot \Pi^{[Q]} \cdot \mathbf{P}$ term. Inserting Eq. (C8) and exploiting these symmetries, we arrive at the final expression resulting from the introduction of the CPCIS $\bar{\mathbf{z}}$-vector, i.e., Eq. (A13).

APPENDIX D: OVERLAP BETWEEN SINGLES DETERMINANTS AT DIFFERENT GEOMETRIES

As shown by Löwdin,[60] the overlap between Slater determinants with nonorthogonal molecular orbitals can be written as a determinant of overlap integrals between individual MOs from each set. For nonorthogonal singles determinants, we obtain

$$
\langle \Phi_i^a | \Phi_j^{b\prime} \rangle =
\begin{cases}
\begin{vmatrix}
\langle \phi_1 | \phi_1' \rangle & \cdots & \langle \phi_1 | \phi_b' \rangle & \cdots & \langle \phi_1 | \phi_{N_o}' \rangle \\
\vdots & & \vdots & & \vdots \\
\langle \phi_a | \phi_1' \rangle & \cdots & \langle \phi_a | \phi_b' \rangle & \cdots & \langle \phi_a | \phi_{N_o}' \rangle \\
\vdots & & \vdots & & \vdots \\
\langle \phi_{N_o} | \phi_1' \rangle & \cdots & \langle \phi_{N_o} | \phi_b' \rangle & \cdots & \langle \phi_{N_o} | \phi_{N_o}' \rangle
\end{vmatrix} & i = j \\[4em]
\begin{vmatrix}
\langle \phi_1 | \phi_1' \rangle & \cdots & \langle \phi_1 | \phi_i' \rangle & \cdots & \langle \phi_1 | \phi_b' \rangle & \cdots & \langle \phi_1 | \phi_{N_o}' \rangle \\
\vdots & & \vdots & & \vdots & & \vdots \\
\langle \phi_a | \phi_1' \rangle & \cdots & \langle \phi_a | \phi_i' \rangle & \cdots & \langle \phi_a | \phi_b' \rangle & \cdots & \langle \phi_a | \phi_{N_o}' \rangle \\
\vdots & & \vdots & & \vdots & & \vdots \\
\langle \phi_j | \phi_1' \rangle & \cdots & \langle \phi_j | \phi_i' \rangle & \cdots & \langle \phi_j | \phi_b' \rangle & \cdots & \langle \phi_j | \phi_{N_o}' \rangle \\
\vdots & & \vdots & & \vdots & & \vdots \\
\langle \phi_{N_o} | \phi_1' \rangle & \cdots & \langle \phi_{N_o} | \phi_i' \rangle & \cdots & \langle \phi_{N_o} | \phi_b' \rangle & \cdots & \langle \phi_{N_o} | \phi_{N_o}' \rangle
\end{vmatrix} & i \neq j,
\end{cases}
\tag{D1}
$$

Eqs. (86) and (87) of Ref. 54 correspond to retaining only the diagonal parts of these determinants.

APPENDIX E: CARTESIAN COORDINATES OF DISTORTED *p*-BENZOQUINONE

See Table IX.

TABLE IX. Cartesian coordinates of distorted *p*-benzoquinone; due to the sensitivity of the coupled-perturbed CIS equations to the values of the amplitudes near a conical intersection, all digits provided are necessary to replicate the results presented in Sections IV B and IV C. All *x*-coordinates are zero.

atom	y (Å)	z (Å)
C	−1.6282	0.531824
C	0.000000215445	0.850617
C	1.6282	0.531826
C	1.6282	−0.531828
C	−0.0000014475	−0.850618
C	−1.62821	−0.531829
H	−2.31725	−1.45939
O	0.000000107777	−2.2785
H	2.31725	−1.4594
H	2.31724	1.4594
O	−0.000000107777	2.2785
H	−2.31725	1.45939

[1] R. A. Marcus, J. Phys. Chem. **67**, 853 (1963).
[2] A. Nitzan, *Chemical Dynamics in Condensed Phases* (Oxford, New York, 2006).
[3] J. C. Tully, J. Chem. Phys. **93**, 1061 (1990).
[4] R. Kapral and G. Ciccotti, J. Chem. Phys. **110**, 8919 (1999).
[5] M. Ben-Nun and T. J. Martínez, J. Chem. Phys. **112**, 6113 (2000).
[6] N. Ananth, C. Venkataraman, and W. H. Miller, J. Chem. Phys. **127**, 084114 (2007).
[7] W. H. Miller, J. Phys. Chem. A **113**, 1405 (2009).
[8] A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover, Mineola, NY, 1996).
[9] M. Baer, Chem. Phys. Lett. **35**, 112 (1975).
[10] C. A. Mead and D. G. Truhlar, J. Chem. Phys. **77**, 6090 (1982).
[11] R. G. Sadygov and D. R. Yarkony, J. Chem. Phys. **109**, 20 (1998).
[12] X. Zhu and D. R. Yarkony, J. Chem. Phys. **132**, 104101 (2010).
[13] C. Hsu, Acc. Chem. Res. **42**, 509 (2009).
[14] T. Pacher, L. S. Cederbaum, and H. Köppel, J. Chem. Phys. **89**, 7367 (1988).
[15] K. Ruedenberg and G. J. Atchity, J. Chem. Phys. **99**, 3799 (1993).
[16] G. J. Atchity and K. Ruedenberg, Theor. Chim. Acta **97**, 47 (1997).
[17] H. Nakamura and D. G. Truhlar, J. Chem. Phys. **115**, 10353 (2001).
[18] H. Nakamura and D. G. Truhlar, J. Chem. Phys. **117**, 5576 (2002).
[19] H. Nakamura and D. G. Truhlar, J. Chem. Phys. **118**, 6816 (2003).
[20] R. J. Cave and M. D. Newton, Chem. Phys. Lett. **249**, 15 (1996).
[21] R. J. Cave and M. D. Newton, J. Chem. Phys. **106**, 9213 (1997).
[22] H. J. Werner and W. Meyer, J. Chem. Phys. **74**, 5802 (1981).
[23] A. A. Voityuk and N. Rosch, J. Chem. Phys. **117**, 5607 (2002).
[24] C. Hsu, Z. Q. You, and H. C. Chen, J. Phys. Chem. C **112**, 1204 (2008).
[25] H. C. Chen, Z. Q. You, and C. Hsu, J. Chem. Phys. **129**, 084708 (2008).
[26] J. E. Subotnik, S. Yeganeh, R. J. Cave, and M. A. Ratner, J. Chem. Phys. **129**, 244101 (2008).

[27]J. E. Subotnik, R. J. Cave, R. P. Steele, and N. Shenvi, J. Chem. Phys. **130**, 234102 (2009).

[28]S. F. Boys, Rev. Mod. Phys. **32**, 296 (1960).

[29]C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. **35**, 457 (1963).

[30]M. Pavanello and J. Neugebauer, J. Chem. Phys. **135**, 134113 (2011).

[31]E. S. Kryachko, J. Phys. Chem. A **103**, 4368 (1999).

[32]E. S. Kryachko, Adv. Quantum Chem. **44**, 119 (2003).

[33]E. Alguire, S. Fatehi, Y. Shao, and J. E. Subotnik (unpublished).

[34]G. L. Closs, P. Piotrowiak, J. M. MacInnis, and G. R. Fleming, J. Am. Chem. Soc. **110**, 2652 (1988).

[35]G. L. Closs, M. Johnson, J. R. Miller, and P. Piotrowiak, J. Am. Chem. Soc. **111**, 3751 (1989).

[36]N. Koga, K. Sameshima, and K. Morokuma, J. Phys. Chem. **97**, 13117 (1993).

[37]M. Braga and S. Larsson, J. Phys. Chem. **97**, 8929 (1993).

[38]J. E. Subotnik, J. Vura-Weis, A. Sodt, and M. A. Ratner, J. Phys. Chem. A **114**, 8665 (2010).

[39]D. A. Kleier, T. A. Halgren, J. H. Hall, and W. N. Lipscomb, J. Chem. Phys. **61**, 3905 (1974).

[40]G. Orlandi and W. Siebrand, J. Chem. Phys. **58**, 4513 (1973).

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

[42]B. H. Lengsfield III and D. R. Yarkony, Adv. Chem. Phys. **82**(2), 1 (1992).

[43]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. 129–174.

[44]B. H. Lengsfield III, P. Saxe, and D. R. Yarkony, J. Chem. Phys. **81**, 4549 (1984).

[45]P. Saxe, B. H. Lengsfield III, and D. R. Yarkony, Chem. Phys. Lett. **113**, 159 (1985).

[46]H. Lischka, M. Dallos, and R. Shepard, Mol. Phys. **100**, 1647 (2002).

[47]H. Lischka, M. Dallos, P. G. Szalay, D. R. Yarkony, and R. Shepard, J. Chem. Phys. **120**, 7322 (2004).

[48]M. Dallos, H. Lischka, R. Shepard, D. R. Yarkony, and P. G. Szalay, J. Chem. Phys. **120**, 7330 (2004).

[49]P. Jørgensen and J. Simons, J. Chem. Phys. **79**, 334 (1983).

[50]T. Helgaker and J. Almlöf, Int. J. Quantum Chem. **26**, 275 (1984).

[51]M. Page, P. Saxe, G. F. Adams, and B. H. Lengsfield III, J. Chem. Phys. **81**, 434 (1984).

[52]R. Shepard, Int. J. Quantum Chem. **31**, 33 (1987).

[53]R. Shepard, H. Lischka, P. G. Szalay, T. Kovar, and M. Ernzerhof, J. Chem. Phys. **96**, 2085 (1992).

[54]S. Fatehi, E. Alguire, Y. Shao, and J. E. Subotnik, J. Chem. Phys. **135**, 234105 (2011).

[55]S. Coriani, T. Kjærgaard, P. Jørgensen, K. Ruud, J. Huh, and R. Berger, J. Chem. Theory Comput. **6**, 1028 (2010).

[56]P. Pulay, Mol. Phys. **17**, 197 (1969).

[57]D. Maurice, *Single Electron Theories of Excited States* (University of California, Berkeley, 1998).

[58]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. Hsu, G. Kedziora, R. Z. Khalliulin, 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).

[59]A. I. Krylov and P. M. W. Gill, WIREs Comput. Mol. Sci. **3**, 317 (2013).

[60]P. Löwdin, Phys. Rev. **97**, 1474 (1955).

[61]E. Alguire and J. E. Subotnik, J. Chem. Phys. **137**, 194108 (2012).

[62]I. Abraham, R. Joshi, P. Pardasani, and R. T. Pardasani, J. Braz. Chem. Soc. **22**, 385 (2011).

[63]A. K. Grafton, S. E. Boesch, and R. A. Wheeler, J. Mol. Struct.: THEOCHEM **392**, 1 (1997).

[64]V. Gold, Trans. Faraday Soc. **46**, 109 (1950).

[65]H. P. Trommsdorff, J. Chem. Phys. **56**, 5358 (1972).

[66]R. W. Bigelow, J. Chem. Phys. **68**, 5086 (1978).

[67]R. Pou-Amérigo, M. Merchán, and E. Ortí, J. Chem. Phys. **110**, 9536 (1999).

[68]Y. Honda, M. Hada, M. Ehara, and H. Nakatsuji, J. Phys. Chem. A **106**, 3838 (2002).

[69]D. M. Chipman and M. F. Prebenda, J. Phys. Chem. **90**, 5557 (1986).

[70]R. Liu, X. Zhou, and P. Pulay, J. Phys. Chem. **96**, 4255 (1992).

[71]M. Nonella, J. Phys. Chem. B **101**, 1235 (1997).

[72]S. E. Boesch, A. K. Grafton, and R. A. Wheeler, J. Phys. Chem. **100**, 10083 (1996).

[73]L. A. Eriksson, F. Himo, P. E. M. Siegbahn, and G. T. Babcock, J. Phys. Chem. A **101**, 9496 (1997).

[74]J. Schiedt and R. Weinkauf, J. Chem. Phys. **110**, 304 (1999).

[75]Q. Fu, J. Yang, and X. Wang, J. Phys. Chem. A **115**, 3201 (2011).

[76]M. Head-Gordon, A. M. Grana, D. Maurice, and C. A. White, J. Phys. Chem. **99**, 14261 (1995).

[77]S. Fatehi and J. E. Subotnik, J. Phys. Chem. Lett. **3**, 2039 (2012).

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

[79]J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem. **16**, 225 (1979).

[80]J. Simons, P. Jørgensen, and T. U. Helgaker, Chem. Phys. **86**, 413 (1984).

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

[82]Y. Osamura, Y. Yamaguchi, and H. F. Schaefer III, Theor. Chim. Acta **72**, 71 (1987).