

Analytic Nuclear Gradients for Complete Active Space Linearized Pair-Density Functional Theory

Matthew R. Hennefarth,¹ Matthew R. Hermes,¹ Donald G. Truhlar,^{2, a)} and Laura Gagliardi^{3, 4, b)}

¹⁾*Department of Chemistry and Chicago Center for Theoretical Chemistry, University of Chicago, Chicago, IL 60637, USA*

²⁾*Department of Chemistry, Chemical Theory Center, and Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, MN 55455-0431, USA*

³⁾*Department of Chemistry, Pritzker School of Molecular Engineering, The James Franck Institute, and Chicago Center for Theoretical Chemistry, University of Chicago, Chicago, IL 60637, USA*

⁴⁾*Argonne National Laboratory, 9700 S. Cass Avenue, Lemont, IL 60439, USA*

Accurately modeling photochemical reactions is difficult due to the presence of conical intersections and locally avoided crossings as well as the inherently multiconfigurational character of excited states. As such, one needs a multi-state method that incorporates state interaction in order to accurately model the potential energy surface at all nuclear coordinates. The recently developed linearized pair-density functional theory (L-PDFT) is a multi-state extension of multiconfiguration PDFT, and it has been shown to be a cost-effective post-MCSCF method (as compared to more traditional and expensive multireference many-body perturbation methods or multireference configuration interaction methods) that can accurately model potential energy surfaces in regions of strong nuclear-electronic coupling in addition to accurately predicting Franck-Condon vertical excitations. In this paper, we report the derivation of analytic gradients for L-PDFT and their implementation in the PYSCF-FORGE software, and we illustrate the utility of these gradients for predicting ground- and excited-state equilibrium geometries and adiabatic excitation energies for formaldehyde, *s-trans*-butadiene, phenol, and cytosine.

I. INTRODUCTION

Accurate characterization and modeling of electronically excited states is important for a variety of chemical and biochemical processes including light-harvesting,^{1–5} photochemistry,^{6,7} photocatalysis,^{8,9} vision,^{10,11} and UV damage to DNA.^{12–14} However, since excited states are typically strongly multiconfigurational, a multireference method is necessary for quantitative and qualitative accuracy. Additionally, since conical intersections and locally avoided crossings are common topological features of excited-state potential energy surfaces, the method should also be able to properly incorporate state interaction so that states of the same symmetry do not unphysically cross or cross on surfaces of the wrong dimensionality.

While state-averaged complete active space self-consistent field (CASSCF) theory^{15,16} is a multireference method that can properly account for static correlation, it does not include dynamic correlation outside of the active space. Multireference many-body perturbation theories such as CAS second-order perturbation theory (CASPT2)¹⁷ or *n*-electron valence state second-order perturbation theory (NEVPT2)¹⁸ are able to recover external dynamic correlation; however, they are computationally very expensive. Multiconfiguration pair-density functional theory (MC-PDFT)^{19–21} is an alternative post-MCSCF approach that combines wave function

theory and density functional theory where the final electronic energy of a multiconfigurational wave function is computed using an on-top energy functional for the non-classical component of the energy. MC-PDFT recovers the missing correlation energy with an accuracy similar to CASPT2²² and NEVPT2²³ but with a significantly reduced computational cost.

MC-PDFT is inaccurate in regions of strong nuclear-electronic coupling because it is a single-state method.^{24–26} Linearized PDFT²⁷ is a recently developed multi-state extension of MC-PDFT that incorporates state interaction by defining an effective Hamiltonian that is a functional of a set of densities. It has been shown to be as accurate as extended multi-state CASPT2 (MS-CASPT2)²⁸ in modeling potential energy surfaces near conical intersections and locally avoided crossings,²⁷ and it is as accurate as NEVPT2 for predicting Franck-Condon vertical excitations.²⁹ L-PDFT also has the benefit of being computationally faster than MC-PDFT since its computational cost is independent of the number of states included in the model space, making it an excellent method to study photochemical reactions and dynamics.

Nuclear gradients are used to optimize molecular geometries to study vertical and adiabatic excitations and to perform molecular dynamics simulations. While it is always possible to calculate gradients numerically, it is often very slow to converge the gradients with respect to the stepsize. Here we report the derivation and implementation of analytic gradients for L-PDFT based on a SA-CASSCF reference wave function. Since the L-PDFT energy is not fully variational with respect to all wave function parameters, we use a Lagrangian-based approach similar to SA-CASSCF gradients,³⁰ SA-MC-

^{a)}corresponding author: truhlar@umn.edu

^{b)}corresponding author: lgagliardi@uchicago.edu

PDFT gradients,³¹ state-specific MC-PDFT (SS-MC-PDFT) gradients,³² and MC-PDFT gradients with density fitting.³³ We then validate our implementation by comparing the analytic gradients to numerical gradients for the diatomic systems HeH⁺ and LiH. Finally, we show the utility of L-PDFT gradients in predicting both ground- and excited-state geometries, as well as vertical and adiabatic excitation energies for formaldehyde, phenol, *s-trans*-butadiene, and cytosine.

II. THEORY

Throughout this manuscript, lowercase roman letters $p, q, r, s, t, u, v, w, x, y$ indicate general spatial molecular orbitals (MO), and lowercase Greek letters τ, μ, ν, ξ indicate atomic orbital (AO) basis functions. I, J refer to CASSCF eigenstates in the state-averaged space (which is taken to be the same as the model space); Λ, Γ label L-PDFT eigenstates (within the model space); and M, N label CASSCF eigenstates within the complementary part of the state-averaged space (outside the model space). κ is used for MO rotations, λ for general nuclear coordinates, and P for the state-transfer operator. Bold-faced variables are tensors (vectors, matrices, etc.). Einstein summation notation is used throughout (repeated indices are summed implicitly). An efficient implementation of the following equations should make use of the partitioning of orbitals into inactive, active, and virtual, and our code does this; however, our derivation presented in this manuscript does not account for such partitioning for simplicity.

A. L-PDFT

The L-PDFT energy²⁷ of a given state $|\Gamma\rangle$ is defined as the first-order Taylor expansion of the MC-PDFT energy expression¹⁹ in the one- and two-reduced density matrix (RDM) elements around the zero-order density ($\tilde{\gamma}$).

$$E_{\Gamma}^{\text{L-PDFT}} = E^{\text{PDFT}}[\tilde{\gamma}] + \left. \frac{\partial E^{\text{PDFT}}}{\partial \gamma_q^p} \right|_{\tilde{\gamma}} \Delta_q^p + \left. \frac{\partial E^{\text{PDFT}}}{\partial \gamma_{qs}^{pr}} \right|_{\tilde{\gamma}} \Delta_{qs}^{pr} \quad (1)$$

$$\Delta_q^p = \gamma_q^p - \tilde{\gamma}_q^p \quad (2a)$$

$$\Delta_{qs}^{pr} = \gamma_{qs}^{pr} - \tilde{\gamma}_{qs}^{pr} \quad (2b)$$

Here, γ_q^p and γ_{qs}^{pr} are the one- and two-RDM elements of state $|\Gamma\rangle$, $\tilde{\gamma}_q^p$ and $\tilde{\gamma}_{qs}^{pr}$ are the one- and two-RDM elements of the zero-order density, Δ represents the difference between in RDM elements between the state $|\Gamma\rangle$ and zero-order density, and $E^{\text{PDFT}}[\tilde{\gamma}]$ is the MC-PDFT energy expression evaluated with the zero-order RDM elements.

$$E^{\text{PDFT}}[\tilde{\gamma}] = h_p^q \tilde{\gamma}_q^p + \frac{1}{2} \mathcal{J}_p^q[\tilde{\gamma}] \tilde{\gamma}_q^p + E^{\text{ot}}[\tilde{\rho}_{\tilde{\gamma}}] + V^{\text{nuc}} \quad (3)$$

$$\mathcal{J}_p^q[\tilde{\gamma}] = g_{pr}^{qs} \tilde{\gamma}_s^r \quad (4)$$

h_p^q and g_{pr}^{qs} are the normal one- and two-electron integrals respectively, V^{nuc} is the nuclear-nuclear repulsion, $\mathcal{J}_p^q[\tilde{\gamma}]$ is the Coulomb interaction of the density $\tilde{\gamma}$, and $E^{\text{ot}}[\tilde{\rho}_{\tilde{\gamma}}]$ is an on-top energy functional that depends on the collective density variables $\tilde{\rho}_{\tilde{\gamma}}$.

$$\tilde{\rho}_{\tilde{\gamma}}^{\top} = [\rho_{\tilde{\gamma}} \quad \Pi_{\tilde{\gamma}} \quad \rho'_{\tilde{\gamma}} \quad \Pi'_{\tilde{\gamma}}] \quad (5)$$

The density (ρ), on-top pair density (Π), and their gradients are generated through the one- and two-RDM elements as

$$\rho_{\tilde{\gamma}} = \phi_p \tilde{\gamma}_q^p \phi^q \quad (6)$$

$$\Pi_{\tilde{\gamma}} = \frac{1}{2} \phi_p \phi_r \tilde{\gamma}_{qs}^{pr} \phi^q \phi^s \quad (7)$$

$$\rho'_{\tilde{\gamma}} = (\phi'_p \phi^q + \phi_p \phi'^q) \tilde{\gamma}_q^p \quad (8)$$

$$\begin{aligned} \Pi'_{\tilde{\gamma}} = & \frac{1}{2} (\phi'_p \phi_r \phi^q \phi^s + \phi_p \phi'_r \phi^q \phi^s \\ & + \phi_p \phi_r \phi'^q \phi^s + \phi_p \phi_r \phi^q \phi'^s) \tilde{\gamma}_{qs}^{pr} \end{aligned} \quad (9)$$

Note that ρ , Π , and ϕ_p are all functions of one three-dimensional variable \mathbf{r} ; and ϕ_p is the p 'th MO. Appendix A describes how the current generation of on-top functionals are evaluated using existing Kohn-Sham (KS) functionals via translated and fully-translated schemes.

The L-PDFT energy is expressible as²⁷

$$\begin{aligned} E_{\Gamma}^{\text{L-PDFT}} = & h_p^q \gamma_q^p + (\mathcal{J}_p^q[\tilde{\gamma}] + V_p^q[\tilde{\rho}_{\tilde{\gamma}}]) \Delta_q^p + \frac{1}{2} v_{pr}^{qs}[\tilde{\rho}_{\tilde{\gamma}}] \Delta_{qs}^{pr} \\ & + E^{\text{ot}}[\tilde{\rho}_{\tilde{\gamma}}] + \frac{1}{2} \mathcal{J}_p^q[\tilde{\gamma}] \tilde{\gamma}_q^p + V^{\text{nuc}} \end{aligned} \quad (10)$$

$$V_p^q = V_p^q[\tilde{\rho}_{\tilde{\gamma}}] = \left. \frac{\partial E^{\text{ot}}}{\partial \gamma_q^p} \right|_{\tilde{\rho}_{\tilde{\gamma}}} \quad (11a)$$

$$v_{pr}^{qs} = v_{pr}^{qs}[\tilde{\rho}_{\tilde{\gamma}}] = 2 \left. \frac{\partial E^{\text{ot}}}{\partial \gamma_{qs}^{pr}} \right|_{\tilde{\rho}_{\tilde{\gamma}}} \quad (11b)$$

The L-PDFT energy contains derivatives of the on-top functional with respect to the one- (V_p^q) and two-RDM (v_{qs}^{pr}) elements (which we call the one- and two-electron on-top potential terms^{31,32}).

In L-PDFT, all functional terms (E^{ot} , V_p^q , v_{pr}^{qs}) are evaluated only at the zero-order density. In practice, the zero-order density is taken to be the weighted average of densities within the state-averaged manifold.

$$\tilde{\gamma}_q^p = \omega_I \langle I | \hat{E}_q^p | I \rangle \quad (12a)$$

$$\tilde{\gamma}_{qs}^{pr} = \omega_I \langle I | \hat{e}_{qs}^{pr} | I \rangle \quad (12b)$$

\hat{E}_q^p and \hat{e}_{qs}^{pr} the one- and two-electron excitation operators respectively. We take ω_I to be the same weight as in the underlying SA-CASSCF or state-averaged complete active space configuration interaction calculation. For the analytic L-PDFT gradients, we require equal weights ($\omega_I = \omega_J$) so that the zero-order density is invariant to rotation among states within the model space, as is done in SA-CASSCF.³⁰

Because the L-PDFT energy depends linearly on the one- and two-RDM elements of the state, it is possible to express it as the expectation value of a Hermitian operator which we call the L-PDFT Hamiltonian ($\hat{H}^{\text{L-PDFT}}$).

$$E_{\Gamma}^{\text{L-PDFT}} = \langle \Gamma | \hat{H}^{\text{L-PDFT}} | \Gamma \rangle \quad (13)$$

$$\hat{H}^{\text{L-PDFT}} = (h_p^q + \mathcal{J}_p^q[\tilde{\gamma}] + V_p^q) \hat{E}_q^p + \frac{1}{2} v_{pr}^{qs} \hat{e}_{qs}^{pr} + h^{\text{const}} \quad (14)$$

Here h^{const} is a constant term that only depends on $\tilde{\gamma}$.

$$h^{\text{const}} = V^{\text{nuc}} + E^{\text{ot}} - \left(\frac{1}{2} \mathcal{J}_p^q[\tilde{\gamma}] + V_p^q \right) \tilde{\gamma}_q^p - \frac{1}{2} v_{pr}^{qs} \tilde{\gamma}_{qs}^{pr} \quad (15)$$

The final L-PDFT energies and states are the solutions to the eigenvalue equation of the operator $\hat{H}^{\text{L-PDFT}}$ defined within the model space spanned by the eigenvectors of the underlying SA-CASSCF calculation.

$$|\Gamma\rangle \langle \Gamma | \hat{H}^{\text{L-PDFT}} | \Lambda \rangle = \delta_{\Gamma}^{\Lambda} E_{\Lambda}^{\text{L-PDFT}} |\Gamma\rangle \quad (16)$$

Note that $|\Gamma\rangle$ and $|I\rangle$ span the same model space, but the state $|I\rangle$ are chosen such that they diagonalize the normal electronic Hamiltonian projected within that space.

B. The L-PDFT Energy Lagrangian

The Hellmann-Feynman theorem³⁴⁻³⁶ implies that if an energy E is stationary with respect to all parameters defining the wave function ($|\Psi\rangle$), then

$$\frac{dE}{d\lambda} = \langle \Psi | \frac{d\hat{H}}{d\lambda} | \Psi \rangle \quad (17)$$

As such, one does not have to account for the response of the wave function to a change in nuclear coordinate (λ).

The L-PDFT energy is not stationary with respect to MO rotations and state rotations out of the model space; hence, the Hellmann-Feynman theory cannot be applied. However, one can avoid calculating the response of the wave function with respect to a nuclear displacement by using Lagrange's method of undetermined multipliers. This requires us to enumerate the variables defining the wave function and the systems of equations that set them to their particular values.

Similarly to SA-CASSCF, the final L-PDFT eigenstates can be parameterized as

$$|\Gamma\rangle = e^{\hat{P}^{\Gamma}} e^{\hat{\kappa}} |0\rangle \quad (18)$$

where $\hat{\kappa}$ is the orbital rotation operator

$$\hat{\kappa} = \sum_{p < q} \kappa_p^q \left[\hat{\mathbf{E}} - \hat{\mathbf{E}}^{\dagger} \right]_q^p \quad (19)$$

and \hat{P}^{Γ} is the state transfer operator for state $|\Gamma\rangle$. \hat{P}^{Γ} can be decomposed into rotations within the model space ($\hat{P}_{\parallel}^{\Gamma}$) and those outside the model space (\hat{P}_{\perp}^{Γ}).

$$\hat{P}^{\Gamma} = \hat{P}_{\perp}^{\Gamma} + \hat{P}_{\parallel}^{\Gamma} \quad (20)$$

$$\hat{P}_{\perp}^{\Gamma} = P_M^{\Gamma} (|M\rangle\langle\Gamma| - |\Gamma\rangle\langle M|) \quad (21)$$

$$\hat{P}_{\parallel}^{\Gamma} = P_{\Lambda}^{\Gamma} (|\Lambda\rangle\langle\Gamma| - |\Gamma\rangle\langle\Lambda|) \quad (22)$$

Since our reference wave function comes from a SA-CASSCF calculation, the parameters κ_p^q and P_M^{Γ} are optimized with respect to the SA-CASSCF energy (E^{SA}).

$$\frac{\partial E^{\text{SA}}}{\partial \kappa_p^q} = 0 \quad (23)$$

$$\frac{\partial E^{\text{SA}}}{\partial P_M^{\Gamma}} = 0 \quad (24)$$

The P_{Λ}^{Γ} parameters are determined by diagonalizing $\hat{H}^{\text{L-PDFT}}$, which is equivalent to making the energies stationary with respect to interstate rotations.

$$\frac{\partial E_{\Gamma}^{\text{L-PDFT}}}{\partial P_{\Lambda}^{\Gamma}} = 0 \quad (25)$$

For equal weights, E^{SA} is invariant to rotation within the model space, and hence

$$\frac{d}{dx} \left(\frac{\partial E^{\text{SA}}}{\partial P_{\Lambda}^{\Gamma}} \right) = 0 \quad (26)$$

for any parameter x .³⁰

Our constraints are given by Eqs. 23 and 24; therefore, our Lagrangian for state $|\Gamma\rangle$ takes the form

$$\mathcal{L}_{\Gamma} = E_{\Gamma}^{\text{L-PDFT}} + \bar{\kappa} \cdot \nabla_{\kappa} E^{\text{SA}} + \bar{\mathbf{P}}^{\perp} \cdot \nabla_{\mathbf{P}^{\perp}} E^{\text{SA}} \quad (27)$$

As a reminder $\bar{\kappa}$ and $\bar{\mathbf{P}}^{\perp}$ are the associated Lagrange multipliers for the orbital rotations and state rotations out of the model space respectively. Both $\bar{\kappa}$ and $\bar{\mathbf{P}}^{\perp}$ are determined by making the Lagrangian stationary with respect to κ_p^q and P_M^{Λ} .

$$\frac{\partial \mathcal{L}_{\Gamma}}{\partial \kappa_p^q} = 0 = \frac{\partial \mathcal{L}_{\Gamma}}{\partial P_M^{\Lambda}} \quad (28)$$

As indicated by Eqs. 25 and 26, \mathcal{L}_{Γ} is already invariant to rotations within the model space, and therefore does

not need to be accounted for. Substituting Eq. 27 into Eq. 28 yields a system of coupled linear equations.

$$\begin{bmatrix} \nabla_{\kappa} E_{\Gamma}^{\text{L-PDFT}} \\ \nabla_{\mathbf{P}^{\perp}} E_{\Gamma}^{\text{L-PDFT}} \end{bmatrix} = - \begin{bmatrix} \mathbf{H}_{\kappa\kappa}^{E^{\text{SA}}} & \mathbf{H}_{\kappa\mathbf{P}^{\perp}}^{E^{\text{SA}}} \\ \mathbf{H}_{\mathbf{P}^{\perp}\kappa}^{E^{\text{SA}}} & \mathbf{H}_{\mathbf{P}^{\perp}\mathbf{P}^{\perp}}^{E^{\text{SA}}} \end{bmatrix} \begin{bmatrix} \bar{\kappa} \\ \bar{\mathbf{P}}^{\perp} \end{bmatrix} \quad (29)$$

The left-hand side is the energy response of the L-PDFT energy, and the first factor on the right is the Hessian of E^{SA} with respect to κ_p^q and P_M^{Λ} ; these terms will be discussed in Section II C. Equations 27 and 29 are almost identical to those presented in the SA-CASSCF analytic gradients,³⁰ with the difference being that the wave function energy response terms on the left side have been replaced with the L-PDFT energy response terms. As such, Eq. 29 can be solved using the standard SA-CASSCF preconditioned conjugate gradient iterative solver.^{37,38}

C. Energy Response

The SA-CASSCF Hessian that appears in Eq. 29 is well established^{30,39} and unchanged in this derivation. In practice, we evaluate $\mathbf{H}^{E^{\text{SA}}}$ within the L-PDFT eigenstate basis, rather than the SA-CASSCF eigenstate basis, which slightly modifies the form of the equation as is described in Appendix B.

As seen on the left side of Eq. 29, we need the response of the L-PDFT energy with respect to orbital and CI rotations. However, it is important to realize that $E_{\Gamma}^{\text{L-PDFT}}$ depends on the model space, and changing either the orbitals or CI parameters for any state within the model space changes the zero-order density. Hence, we break each component into an explicit and implicit part; the explicit part represents the response of only the state, whereas the implicit part accounts for changes due to the zero-order density.

1. Explicit Dependence

The derivatives of γ with respect to κ_p^q are given by

$$\frac{\partial \gamma_p^q}{\partial \kappa_x^y} = \delta_p^x \gamma_y^q - \delta_p^y \gamma_x^q - \delta_y^q \gamma_p^x + \delta_x^q \gamma_p^y \quad (30a)$$

$$\begin{aligned} \frac{\partial \gamma_{pr}^{qs}}{\partial \kappa_x^y} &= \delta_p^x \gamma_{yr}^{qs} + \delta_x^q \gamma_{pr}^{ys} - \delta_p^y \gamma_{xr}^{qs} - \delta_y^q \gamma_{pr}^{xs} \\ &+ \delta_r^x \gamma_{py}^{qs} + \delta_x^s \gamma_{pr}^{qy} - \delta_r^y \gamma_{px}^{qs} - \delta_y^s \gamma_{pr}^{qx} \end{aligned} \quad (30b)$$

Therefore,

$$\frac{\partial E_{\Gamma}^{\text{L-PDFT}}}{\partial \gamma} \cdot \frac{\partial \gamma}{\partial \kappa_x^y} = 2[\mathcal{F}_{\text{ex}} - \mathcal{F}_{\text{ex}}^{\dagger}]_y^x \quad (31)$$

where \mathcal{F}_{ex} is the explicit part of the L-PDFT generalized Fock matrix.

$$[\mathcal{F}_{\text{ex}}]_y^x = (h_y^q + \mathcal{J}_y^q[\tilde{\gamma}] + V_y^q) \gamma_q^x + v_{yr}^{qs} \gamma_{qs}^{xr} \quad (32)$$

The derivative of γ with respect to P_M^{Γ} is given by

$$\frac{\partial \gamma_q^p}{\partial P_M^{\Lambda}} = 2\delta_{\Lambda}^{\Gamma} \gamma_{\Lambda q}^{Mp} \quad (33a)$$

$$\frac{\partial \gamma_{qs}^{pr}}{\partial P_M^{\Lambda}} = 2\delta_{\Lambda}^{\Gamma} \gamma_{\Lambda qs}^{Mpr} \quad (33b)$$

where $\gamma_{\Lambda q}^{Mp}$ and $\gamma_{\Lambda qs}^{Mpr}$ are the transition density matrix elements from $|M\rangle$ to $|\Lambda\rangle$.

$$\gamma_{\Lambda q}^{Mp} = \langle \Lambda | \hat{E}_q^p | M \rangle \quad (34a)$$

$$\gamma_{\Lambda qs}^{Mpr} = \langle \Lambda | \hat{e}_{qs}^{pr} | M \rangle \quad (34b)$$

Therefore, the explicit response of the L-PDFT energy to state rotations out of the model space is given by

$$\frac{\partial E_{\Gamma}^{\text{L-PDFT}}}{\partial \gamma} \cdot \frac{\partial \gamma}{\partial P_M^{\Lambda}} = 2\delta_{\Lambda}^{\Gamma} \langle \Lambda | \hat{H}^{\text{L-PDFT}} | M \rangle \quad (35)$$

In general, these expressions are similar to the SA-CASSCF response equations³⁰ but with modified one- and two-electron integrals.

2. Implicit Dependence

Since the L-PDFT energy directly depends on the first derivatives of E^{ot} with respect to γ , the energy response involves explicit second derivatives. We define the elements of the Hessian of E^{ot} with respect to the RDM elements as

$$F_{p,r}^{q,s} = \left. \frac{\partial^2 E^{\text{ot}}}{\partial \gamma_q^p \partial \gamma_r^s} \right|_{\bar{\rho}_\gamma} \quad (36a)$$

$$F_{p,rt}^{q,su} = F_{rt,p}^{su,q} = 2 \left. \frac{\partial^2 E^{\text{ot}}}{\partial \gamma_q^p \partial \gamma_{su}^{rt}} \right|_{\bar{\rho}_\gamma} \quad (36b)$$

$$F_{pr,tv}^{qs,uw} = 4 \left. \frac{\partial^2 E^{\text{ot}}}{\partial \gamma_{qs}^{pr} \partial \gamma_{uv}^{tv}} \right|_{\bar{\rho}_\gamma} \quad (36c)$$

Since E^{ot} is the integral of the on-top kernel (ϵ^{ot}) over all space,

$$E^{\text{ot}}[\bar{\rho}] = \int \epsilon^{\text{ot}}[\bar{\rho}(\mathbf{r})] d\mathbf{r} \quad (37)$$

we can move the derivatives of Eq. 36 inside the integral. Applying the chain rule twice gives

$$\begin{aligned} \mathbf{F} &= \begin{bmatrix} \{F_{p,t}^{q,u}\} & \{F_{pr,t}^{qs,u}\} \\ \{F_{t,pr}^{u,qs}\} & \{F_{pr,tv}^{qs,uw}\} \end{bmatrix} \\ &= \int d\mathbf{r} \begin{bmatrix} \left\{ \frac{\partial \bar{\rho}}{\partial \gamma_q^p} \right\} \\ 2 \left\{ \frac{\partial \bar{\rho}}{\partial \gamma_{qs}^{pr}} \right\} \end{bmatrix} \cdot \mathbf{f}^{\text{ot}} \cdot \left[\left\{ \frac{\partial \bar{\rho}}{\partial \gamma_u^t} \right\} \right] 2 \left\{ \frac{\partial \bar{\rho}}{\partial \gamma_{uv}^{tv}} \right\} \end{aligned} \quad (38)$$

$$\frac{\partial \vec{\rho}}{\partial \gamma_q^p} = \begin{bmatrix} \phi_p \phi^q \\ 0 \\ 2\phi'_p \phi^q \\ 0 \end{bmatrix} \quad (39a)$$

$$\frac{\partial \vec{\rho}}{\partial \gamma_{qs}^{pr}} = \frac{1}{2} \begin{bmatrix} 0 \\ \phi_p \phi_r \phi^q \phi^s \\ 0 \\ 4\phi'_p \phi_r \phi^q \phi^s \end{bmatrix} \quad (39b)$$

where \mathbf{f}^{ot} is the Hessian of the on-top potential kernel with elements

$$\mathbf{f}^{\text{ot}} = \begin{bmatrix} \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \rho^2} & \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \rho \partial \Pi} & \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \rho \partial \rho'} & \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \rho \partial \Pi'} \\ \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \rho \partial \Pi} & \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \Pi^2} & \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \Pi \partial \rho'} & \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \Pi \partial \Pi'} \\ \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \rho \partial \rho'} & \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \Pi \partial \rho'} & \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \rho'^2} & \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \rho' \partial \Pi'} \\ \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \rho \partial \Pi'} & \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \Pi \partial \Pi'} & \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \rho' \partial \Pi'} & \frac{\partial^2 \epsilon^{\text{ot}}}{\partial \Pi'^2} \end{bmatrix} \quad (40)$$

See Appendix D for a description of how \mathbf{f}^{ot} is evaluated.

As we will see, \mathbf{F} is always contracted with the Δ density elements. We define the Hessian-vector product $\Delta \cdot \mathbf{F}$ to be the on-top gradient response.

$$[\Delta \cdot \mathbf{F}]_t^u = F_{p,t}^{q,u} \Delta_q^p + \frac{1}{2} F_{pr,t}^{qs,u} \Delta_{qs}^{pr} \quad (41a)$$

$$[\Delta \cdot \mathbf{F}]_{tv}^{uw} = F_{p,tv}^{q,uw} \Delta_q^p + \frac{1}{2} F_{pr,tv}^{qs,uw} \Delta_{qs}^{pr} \quad (41b)$$

We directly compute the elements of $\Delta \cdot \mathbf{F}$ on the grid by moving the contraction with Δ within the integral, thereby avoiding constructing tensors of up to rank 8. We see this by first noting that the density variables $\vec{\rho}$ are linear with respect to the RDM elements

$$\vec{\rho}_\Delta = \frac{\partial \vec{\rho}}{\partial \gamma_q^p} \Delta_q^p + \frac{\partial \vec{\rho}}{\partial \gamma_{qs}^{pr}} \Delta_{qs}^{pr} = \frac{\partial \vec{\rho}}{\partial \gamma} \cdot \Delta \quad (42)$$

Thus, contracting Eq. 38 with the elements of the Δ , we find that the elements of $\Delta \cdot \mathbf{F}$ are given by

$$[\Delta \cdot \mathbf{F}]_p^q = \int \text{dr} \vec{\rho}_\Delta^\top \cdot \mathbf{f}^{\text{ot}} \cdot \frac{\partial \vec{\rho}}{\partial \gamma_p^q} \quad (43a)$$

$$[\Delta \cdot \mathbf{F}]_{pr}^{qs} = 2 \int \text{dr} \vec{\rho}_\Delta^\top \cdot \mathbf{f}^{\text{ot}} \cdot \frac{\partial \vec{\rho}}{\partial \gamma_{qs}^{pr}} \quad (43b)$$

The derivative of the $E_\Gamma^{\text{L-PDFT}}$ with respect to the zero-order density matrix elements is given by

$$\frac{\partial E_\Gamma^{\text{L-PDFT}}}{\partial \tilde{\gamma}_t^u} = \mathcal{J}_t^u[\Delta] + [\Delta \cdot \mathbf{F}]_t^u \quad (44a)$$

$$\frac{\partial E_\Gamma^{\text{L-PDFT}}}{\partial \tilde{\gamma}_{tw}^{uv}} = \frac{1}{2} [\Delta \cdot \mathbf{F}]_{tw}^{uv} \quad (44b)$$

Using Eq. 30, we see that the L-PDFT energy implicit response to MO rotations is

$$\frac{\partial E_\Gamma^{\text{L-PDFT}}}{\partial \tilde{\gamma}} \cdot \frac{\partial \tilde{\gamma}}{\partial \kappa_x^y} = 2 \left[\mathcal{F}_{\text{impl}} - \mathcal{F}_{\text{impl}}^\dagger \right]_y^x \quad (45)$$

where $\mathcal{F}_{\text{impl}}$ is the implicit contribution to the L-PDFT generalized Fock matrix.

$$[\mathcal{F}_{\text{impl}}]_y^x = \left(\mathcal{J}_y^q[\Delta] + [\Delta \cdot \mathbf{F}]_y^q \right) \tilde{\gamma}_q^x + [\Delta \cdot \mathbf{F}]_{yr}^{qs} \tilde{\gamma}_{qs}^{xr} \quad (46)$$

Additionally, if we consider Eq. 33, we find that

$$\frac{\partial \tilde{\gamma}_q^p}{\partial P_M^\Lambda} = 2\omega_\Lambda \gamma_{\Lambda q}^{Mp} \quad (47a)$$

$$\frac{\partial \tilde{\gamma}_{qs}^{pr}}{\partial P_M^\Lambda} = 2\omega_\Lambda \gamma_{\Lambda qs}^{Mpr} \quad (47b)$$

Therefore, we have that the implicit response of $E_\Gamma^{\text{L-PDFT}}$ to state rotations out of the model space is given by

$$\frac{\partial E_\Gamma^{\text{L-PDFT}}}{\partial \tilde{\gamma}} \cdot \frac{\partial \tilde{\gamma}}{\partial P_M^\Lambda} = 2\omega_\Lambda G_\Lambda^M \quad (48)$$

where we have defined G_Λ^M as

$$G_\Lambda^M = \langle M | \hat{G} | \Lambda \rangle \quad (49)$$

$$\hat{G} = \left(\mathcal{J}_p^q[\Delta] + [\Delta \cdot \mathbf{F}]_p^q \right) \hat{E}_q^p + \frac{1}{2} [\Delta \cdot \mathbf{F}]_{pr}^{qs} \hat{e}_{qs}^{pr} \quad (50)$$

3. Response Summary

Given the above decompositions, we have the L-PDFT energy response to orbitals rotations is given by

$$\frac{\partial E_\Gamma^{\text{L-PDFT}}}{\partial \kappa_x^y} = 2 \left[\mathcal{F} - \mathcal{F}^\dagger \right]_y^x \quad (51)$$

where \mathcal{F} is the full L-PDFT generalized Fock matrix.

$$\begin{aligned} \mathcal{F}_y^x &= [\mathcal{F}_{\text{ex}} + \mathcal{F}_{\text{impl}}]_y^x \\ &= (h_y^q + \mathcal{J}_y^q[\tilde{\gamma}] + V_y^q) \gamma_q^x + v_{yr}^{qs} \gamma_{qs}^{xr} \\ &\quad + \left(\mathcal{J}_y^q[\Delta] + [\Delta \cdot \mathbf{F}]_y^q \right) \tilde{\gamma}_q^x + [\Delta \cdot \mathbf{F}]_{yr}^{qs} \tilde{\gamma}_{qs}^{xr} \end{aligned} \quad (52)$$

Additionally, we have that the L-PDFT energy response to state rotations out of the model space is given by

$$\frac{\partial E_\Gamma^{\text{L-PDFT}}}{\partial P_M^\Lambda} = 2 \left(\delta_\Lambda^\Gamma \left[\hat{H}^{\text{L-PDFT}} \right]_\Lambda^M + \omega_\Lambda G_\Lambda^M \right) \quad (53)$$

Note that when there is only one state within the model space, $\tilde{\gamma} \rightarrow \gamma$ and the L-PDFT energy is exactly the same as the SS-MC-PDFT energy. Correspondingly, one can see that all of the implicit terms would vanish from Eqs. 52 and 53, which would result in the same response equations as derived for the SS-MC-PDFT gradients.³¹

D. Derivative of the Lagrangian

Having solved for the Lagrange multipliers in Eq. 27 via Eq. 29, the gradient of \mathcal{L}_Γ yields the same gradient as fully differentiating $E_\Gamma^{\text{L-PDFT}}$. Differentiating Eq. 27 with respect to nuclear displacements yields

$$\begin{aligned} \frac{d\mathcal{L}_\Gamma}{d\lambda} &= \frac{dE_\Gamma^{\text{L-PDFT}}}{d\lambda} \\ &= \langle \Gamma | \frac{d\hat{H}^{\text{L-PDFT}}}{d\lambda} | \Gamma \rangle + \bar{\boldsymbol{\kappa}} \cdot \nabla_{\boldsymbol{\kappa}} \left(\omega_\Lambda \langle \Lambda | \frac{d\hat{H}^{\text{el}}}{d\lambda} | \Lambda \rangle \right) \\ &\quad + \bar{\mathbf{P}}^\perp \cdot \nabla_{\mathbf{P}^\perp} \left(\omega_\Lambda \langle \Lambda | \frac{d\hat{H}^{\text{el}}}{d\lambda} | \Lambda \rangle \right) \end{aligned} \quad (54)$$

Both the electronic Hamiltonian and the L-PDFT Hamiltonian in Eq. 54 are expressed within second quantization, which requires that the MOs always be orthonormal. A change in nuclear coordinates can cause the underlying MOs to no longer be orthogonal. In particular, if we consider the following AO to MO transformation at a fixed geometry (λ_0),

$$|\phi^p; \lambda_0\rangle = C_\mu^p |\phi^\mu; \lambda_0\rangle \quad (55)$$

then for any slight nuclear displacement δ , the overlap between the perturbed MOs at the this perturbed geometry is no longer orthonormal.

$$\langle \phi^p; \lambda_0 + \delta | \phi^q; \lambda_0 + \delta \rangle \neq \delta_p^q \quad (56)$$

To ensure that our MOs are orthogonal regardless of the geometry, we introduce the orthogonal molecular orbital (OMO) picture⁴⁰ (using the Löwdin orthonormalization⁴¹),

$$|\psi^{\bar{p}}; \lambda\rangle = \left[\mathbf{S}^{-\frac{1}{2}}(\lambda) \right]_{\bar{p}}^{\bar{p}} C_\mu^p |\phi^\mu; \lambda\rangle \quad (57)$$

where $\mathbf{S}(\lambda)$ is the overlap matrix in the MO basis

$$S_q^p(\lambda) = \langle \phi_q; \lambda | \phi^p; \lambda \rangle \quad (58)$$

At the reference geometry ($\lambda = \lambda_0$), the OMOs are the same as the MOs since $\mathbf{S}(\lambda_0)$ is the identity matrix. Equation 57 also highlights that the transformation from AOs to OMOs also depends on the nuclear coordinates. The derivatives of the one- and two-electron integrals in the OMO basis are given by

$$\frac{dh_{\bar{p}}^{\bar{q}}}{d\lambda} = h_{p,\lambda}^q - \frac{1}{2} \left(S_{p,\lambda}^x h_x^q + S_{x,\lambda}^q h_p^x \right) \quad (59)$$

$$\frac{dg_{\bar{p}\bar{r}}^{\bar{q}\bar{s}}}{d\lambda} = g_{pr,\lambda}^{qs} - \frac{1}{2} \left(S_{p,\lambda}^x g_{qs}^{xr} + S_{x,\lambda}^r g_{qs}^{px} + S_{q,\lambda}^x g_{xs}^{pr} + S_{s,\lambda}^x g_{qx}^{pr} \right) \quad (60)$$

where $h_{p,\lambda}^q$ and $g_{pr,\lambda}^{qs}$ are the derivatives of the one- and two-electron integrals in the MO basis, and $S_{q,\lambda}^p$ is the

derivative of the overlap matrix element. The term involving the derivative of the overlap elements is often called the ‘‘connection’’ or ‘‘renormalization’’ contribution. The one- and two-electron derivative integrals and the overlap matrix elements are obtained in the AO basis as

$$h_{\tau,\lambda}^\mu = \langle \phi_{\tau,\lambda} | \hat{h} | \phi^\mu \rangle + \left\langle \phi_\tau \left| \hat{h} \right| \phi_{,\lambda}^\mu \right\rangle + \langle \phi_\tau | \frac{\partial \hat{h}}{\partial \lambda} | \phi^\mu \rangle \quad (61)$$

$$\begin{aligned} g_{\tau\nu,\lambda}^{\mu\xi} &= \langle \phi_{\tau,\lambda} \phi_\nu | \hat{g} | \phi^\mu \phi^\xi \rangle + \langle \phi_\tau \phi_{\nu,\lambda} | \hat{g} | \phi^\mu \phi^\xi \rangle \\ &\quad + \left\langle \phi_\tau \phi_\nu \left| \hat{g} \right| \phi_{,\lambda}^\mu \phi^\xi \right\rangle + \left\langle \phi_\tau \phi_\nu \left| \hat{g} \right| \phi^\mu \phi_{,\lambda}^\xi \right\rangle \end{aligned} \quad (62)$$

$$S_{\tau,\lambda}^\mu = \langle \phi_{\tau,\lambda} | \phi^\mu \rangle + \left\langle \phi_\tau \left| \phi_{,\lambda}^\mu \right\rangle \quad (63)$$

where $\phi_{\tau,\lambda}$ represents the partial derivative of ϕ_τ with respect to the nuclear coordinate.

$$\phi_{\tau,\lambda} = \frac{\partial \phi_\tau}{\partial \lambda} \quad (64)$$

Within the L-PDFT Hamiltonian, there are no explicit two-electron integrals, but only Coulomb integrals. The Coulomb derivative integrals, generated from a density γ , are given by

$$\mathcal{J}_{p,\lambda}^q[\gamma] = \left(\langle \phi_{p,\lambda} \phi_r | \hat{g} | \phi^q \phi^s \rangle + \langle \phi_p \phi_r | \hat{g} | \phi^{q,\lambda} \phi^s \rangle \right) \gamma_s^r \quad (65)$$

The derivative of the Coulomb contribution can be rewritten in terms of the Coulomb derivative integrals. For example,

$$\begin{aligned} \frac{\partial \mathcal{J}_{\bar{p}}^q[\tilde{\gamma}]}{\partial \lambda} \Delta_q^p &= g_{pr,\lambda}^{qs} \tilde{\gamma}_s^r \Delta_q^p \\ &= \mathcal{J}_{p,\lambda}^q[\tilde{\gamma}] \Delta_q^p + \mathcal{J}_{p,\lambda}^q[\Delta] \tilde{\gamma}_q^p \end{aligned} \quad (66)$$

Since the on-top energy and on-top potential terms depend on the collective density variables (which are constructed from the OMOs), they will also give a renormalization contributions.^{31,32}

$$\frac{dE_{,\lambda}^{\text{ot}}}{d\lambda} = E_{,\lambda}^{\text{ot}} - S_{x,\lambda}^y \left(V_y^p \tilde{\gamma}_p^x + v_{yr}^{qs} \tilde{\gamma}_{qs}^{xr} \right) \quad (67)$$

$$\frac{dV_{\bar{p}}^{\bar{q}}}{d\lambda} \Delta_{\bar{q}}^{\bar{p}} = V_{\bar{p},\lambda}^q \Delta_q^p - S_{x,\lambda}^y \left(V_y^p \Delta_x^p + [\Delta \cdot \mathbf{F}]_y^p \tilde{\gamma}_p^x \right) \quad (68a)$$

$$\frac{dv_{\bar{p}\bar{r}}^{\bar{q}\bar{s}}}{d\lambda} \Delta_{\bar{q}\bar{s}}^{\bar{p}\bar{r}} = v_{pr,\lambda}^{qs} \Delta_{qs}^{pr} - 2S_{x,\lambda}^y \left(v_{yr}^{qs} \Delta_{qs}^{xr} + [\Delta \cdot \mathbf{F}]_{yr}^{qs} \tilde{\gamma}_{qs}^{xr} \right) \quad (68b)$$

The explicit derivatives of the on-top energy ($E_{,\lambda}^{\text{ot}}$) and one- and two-electron on-top potentials ($V_{p,\lambda}^q$ and $v_{pr,\lambda}^{qs}$) can be evaluated in various forms since the tensor contractions can be performed in different orders. Here we present the derivatives as they are implemented in our

PYSCF-FORGE implementation. The nuclear derivative of the collective density variables can be written as

$$\frac{\partial \bar{\rho}_\gamma^\top}{\partial \lambda} = \left[\frac{\partial \rho_\gamma}{\partial \lambda} \quad \frac{\partial \Pi_\gamma}{\partial \lambda} \quad \frac{\partial \rho'_\gamma}{\partial \lambda} \quad \frac{\partial \Pi'_\gamma}{\partial \lambda} \right] + \delta_\lambda(\mathbf{r}) \nabla \bar{\rho}_\gamma \cdot \bar{\mathbf{n}}_\lambda \quad (69)$$

$$\frac{\partial \rho_\gamma}{\partial \lambda} = 2\phi_{p,\lambda} \gamma_q^p \phi_\lambda^q \quad (70)$$

$$\frac{\partial \Pi_\gamma}{\partial \lambda} = 2\phi_{p,\lambda} \phi_r \gamma_{qs}^{pr} \phi^q \phi^s \quad (71)$$

$$\frac{\partial \rho'_\gamma}{\partial \lambda} = 2 \left(\phi'_{p,\lambda} \phi^q + \phi'_p \phi_{,\lambda}^q \right) \gamma_q^p \quad (72)$$

$$\frac{\partial \Pi'_\gamma}{\partial \lambda} = 2 \left(\phi'_{p,\lambda} \phi_r \phi^q \phi^s + 3\phi'_p \phi_{r,\lambda} \phi^q \phi^s \right) \gamma_{qs}^{pr} \quad (73)$$

where $\delta_\lambda(\mathbf{r})$ is 1 if \mathbf{r} is evaluated at a grid point associated with the atom of the coordinate λ and is 0 otherwise, and $\bar{\mathbf{n}}_\lambda$ is the Cartesian unit vector for the coordinate direction λ . The explicit derivative of the on-top energy with respect to nuclear coordinates is given as

$$E_{,\lambda}^{\text{ot}} = \int d\mathbf{r} \left(\mathbf{v}^{\text{ot}} \cdot \frac{\partial \bar{\rho}_\gamma}{\partial \lambda} \right) + \{ \epsilon^{\text{ot}} \}_{\mathbf{r} \in \mathcal{G}} \cdot \frac{\partial \bar{w}}{\partial \lambda} \quad (74)$$

where \mathcal{G} is the set of all grid points, \bar{w} is the corresponding set of quadrature weights, \mathbf{v}^{ot} is the derivative of the on-top kernel with respect to the density variables,

$$\mathbf{v}^{\text{ot}} = \nabla_{\bar{\rho}} \epsilon^{\text{ot}} = \left[\frac{\partial \epsilon^{\text{ot}}}{\partial \rho} \quad \frac{\partial \epsilon^{\text{ot}}}{\partial \Pi} \quad \frac{\partial \epsilon^{\text{ot}}}{\partial \rho'} \quad \frac{\partial \epsilon^{\text{ot}}}{\partial \Pi'} \right] \quad (75)$$

and

$$\{ \epsilon^{\text{ot}} \}_{\mathbf{r} \in \mathcal{G}}^\top = [\epsilon^{\text{ot}}(\mathbf{r}_1) \quad \epsilon^{\text{ot}}(\mathbf{r}_2) \quad \dots \quad \epsilon^{\text{ot}}(\mathbf{r}_n)] \quad (76)$$

is the on-top kernel evaluated at every grid point (as a vector). See Appendix C for how \mathbf{v}^{ot} is evaluated. The nuclear derivative of the on-top potential is given as

$$V_{p,\lambda}^q = \int d\mathbf{r} \left(\frac{\partial \bar{\rho}^\top}{\partial \gamma_q^p} \cdot \mathbf{f}^{\text{ot}} \cdot \frac{\partial \bar{\rho}_\gamma}{\partial \lambda} + \mathbf{v}^{\text{ot}} \cdot \frac{\partial^2 \bar{\rho}}{\partial \gamma_q^p \partial \lambda} \right) \quad (77a)$$

$$+ \left\{ \mathbf{v}^{\text{ot}} \cdot \frac{\partial \bar{\rho}}{\partial \gamma_q^p} \right\}_{\mathbf{r} \in \mathcal{G}} \cdot \frac{\partial \bar{w}}{\partial \lambda}$$

$$v_{pr,\lambda}^{qs} = 2 \int d\mathbf{r} \left(\frac{\partial \bar{\rho}^\top}{\partial \gamma_{qs}^{pr}} \cdot \mathbf{f}^{\text{ot}} \cdot \frac{\partial \bar{\rho}_\gamma}{\partial \lambda} + \mathbf{v}^{\text{ot}} \cdot \frac{\partial^2 \bar{\rho}}{\partial \gamma_{qs}^{pr} \partial \lambda} \right)$$

$$+ 2 \left\{ \mathbf{v}^{\text{ot}} \cdot \frac{\partial \bar{\rho}}{\partial \gamma_{qs}^{pr}} \right\}_{\mathbf{r} \in \mathcal{G}} \cdot \frac{\partial \bar{w}}{\partial \lambda} \quad (77b)$$

However, because the density variables are linear with respect to the RDM elements (Eq. 42), we can move the contraction with Δ in Eq. 68 within the integral (just like in Eq. 43). Additionally, to reduce quadrature loops, we evaluate the on-top energy derivative at the same time

such that the nuclear derivative due to all on-top terms is given as

$$\begin{aligned} E_{,\lambda}^{\text{ot}} + V_{p,\lambda}^q \Delta_q^p + \frac{1}{2} v_{pr,\lambda}^{qs} \Delta_{qs}^{pr} \\ = \int d\mathbf{r} \left(\bar{\rho}_\Delta^\top \cdot \mathbf{f}^{\text{ot}} \cdot \frac{\partial \bar{\rho}_\gamma}{\partial \lambda} + \mathbf{v}^{\text{ot}} \cdot \frac{\partial \bar{\rho}_\gamma}{\partial \lambda} \right) \quad (78) \\ + \{ \epsilon^{\text{ot}} + \mathbf{v}^{\text{ot}} \cdot \bar{\rho}_\Delta \}_{\mathbf{r} \in \mathcal{G}} \cdot \frac{\partial \bar{w}}{\partial \lambda} \end{aligned}$$

We can therefore write the full gradient of the L-PDFT energy for state $|\Gamma\rangle$ as

$$\begin{aligned} \frac{dE_\Gamma^{\text{L-PDFT}}}{d\lambda} = \langle \Gamma | \frac{d\hat{H}^{\text{L-PDFT}}}{d\lambda} | \Gamma \rangle \quad (79) \\ + h_{p,\lambda}^q \check{\gamma}_q^p + \frac{1}{2} g_{pr,\lambda}^{qs} \check{\gamma}_{qs}^{pr} - S_{x,\lambda}^y \check{\mathcal{F}}_y^x \end{aligned}$$

where $\check{\gamma}$ are the effective RDM elements, which contain the Lagrange multiplier terms as

$$\check{\gamma}_q^p = (\check{\gamma}_q^x \bar{\kappa}_x^p - \check{\gamma}_x^p \bar{\kappa}_q^x) + \omega^\Lambda \bar{P}_M^\Lambda \left[\hat{\mathbf{E}} + \hat{\mathbf{E}}^\dagger \right]_{\Lambda q}^{Mp} \quad (80a)$$

$$\begin{aligned} \check{\gamma}_{qs}^{pr} = (\check{\gamma}_{qs}^{xr} \bar{\kappa}_x^p - \check{\gamma}_{xs}^{pr} \bar{\kappa}_q^x + \check{\gamma}_{qs}^{px} \bar{\kappa}_r^x - \check{\gamma}_{qx}^{pr} \bar{\kappa}_s^x) \\ + \omega^\Lambda \bar{P}_M^\Lambda \left[\hat{\mathbf{e}} + \hat{\mathbf{e}}^\dagger \right]_{\Lambda qs}^{Mpr} \quad (80b) \end{aligned}$$

and $\check{\mathcal{F}}$ is the effective Fock matrix,

$$\check{\mathcal{F}}_\tau^\mu = h_\tau^\nu \check{\gamma}_\nu^\mu + g_{\tau\eta}^{\nu\xi} \check{\gamma}_{\nu\xi}^{\mu\eta} \quad (81)$$

The L-PDFT Hellmann-Feynman contribution is given as

$$\begin{aligned} \langle \Gamma | \frac{d\hat{H}^{\text{L-PDFT}}}{d\lambda} | \Gamma \rangle = h_{p,\lambda}^q \gamma_q^p + \mathcal{J}_{p,\lambda}^q [\check{\gamma}] \Delta_q^p + \mathcal{J}_{p,\lambda}^q [\gamma] \check{\gamma}_q^p \\ + V_{,\lambda}^{\text{nuc}} + E_{,\lambda}^{\text{ot}} + V_{p,\lambda}^q \Delta_q^p + \frac{1}{2} v_{pr,\lambda}^{qs} \Delta_{qs}^{pr} \\ - S_{x,\lambda}^y \mathcal{F}_y^x \quad (82) \end{aligned}$$

where \mathcal{F} is the generalized L-PDFT Fock matrix (defined in Eq. 52). Equations 80 and 81 are essentially the same as those that appear in other Lagrange-based analytic gradient approaches.^{30–32,52}

III. COMPUTATIONAL METHODS

All calculations used PYSCF^{53,54} (Version 2.3, commit v1.1-8104-g6c1ea86eb) compiled with the LIBXC^{55,56} (Version 6.1.0) and LIBCINT⁵⁷ (Version 6.0.0) libraries, MRH⁵⁸ (commit SHA-1 b3185fe), and PYSCF-FORGE⁵⁹ (commit SHA-1 c503f41). Geometry optimizations were performed with the GEOMETRIC⁶⁰ package (version 1.0) within PYSCF. All PDFT calculations used a numerical quadrature grid size of 6 (80/120 radial points and 770/974 angular points for atoms of periods 1/2 respectively). All L-PDFT calculations used the model

TABLE I. Systems studied along with the symmetry, basis set, number of states (N), number of active electrons (n_e), active space orbitals, and the on-top functional used.

System	Sym. ^a	Basis Set	N	n_e	Active Orbitals	Functional
HeH ⁺	C_1	cc-pVDZ ⁴²	2	2	σ, σ^*	tPBE ^{19,43}
LiH	C_1	aug-cc-pVTZ ^{42,48}	2	2	σ, σ^*	ftSVWN3 ^b tPBE ^{19,43} ftPBE ^{43,44}
formaldehyde	C_1	jun-cc-pVTZ ^{42,48-51}	2	12	full valence	tPBE ^{19,43}
<i>s-trans</i> -butadiene	C_{2h}	jul-cc-pVTZ ^{42,48-51}	2 (1A_g)	4	$2(\pi, \pi^*)$	tPBE ^{19,43}
phenol	C_1	jul-cc-pVDZ ^{42,48-51}	3	12	$3(\pi, \pi^*), p_z$ of O	tPBE ^{19,43}
cytosine	C_s	jul-cc-pVTZ ^{42,48-51}	3 ($^1A'$)	14	$\sigma_{OH}, \sigma_{OH}^*, \sigma_{CO}, \sigma_{CO}^*$ $5\pi, 3\pi^*$, and 2 lone-pairs	tPBE ^{19,43}

^a Point group symmetry

^b ftSVWN3 is a fully-translated⁴⁴ local-spin-density approximation with the exchange functional being the one called Slater exchange^{45,46} in LIBXC and the correlation functional being correlation functional number 3 of Vosko, Wilk, and Nusair⁴⁷.

space spanned by the SA-CASSCF eigenvectors to construct \hat{H}^{L-PDFT} and equal weights ($\omega_I = \omega_J$). System-specific computational details including symmetry, basis set, number of states in the model space, active space, and on-top functional used are summarized in Table I.

Numerical gradients were computed using the central difference method. Because there is a dependence on the step size (δ), we calculated numerical gradients with differing δ and extrapolated to the $\delta \rightarrow 0$ limit. Specifically, a linear regression of the numerical gradient versus δ^2 is performed for a subset of data where the correlation coefficient R^2 is greater than 0.9, and the y -intercept is taken to be the extrapolated numerical gradient.

Both numerical and analytic gradients suffer from numerical error since they rely on a wave function that is only converged to finite precision. For all comparisons between analytic and numerical gradients, we use an energy convergence threshold of 10^{-12} hartree and an orbital and CI rotation gradient threshold of 10^{-6} . For convenience, we will refer to the numerical gradient as the reference for the rest of this manuscript (as we have done previously^{32,61}). We also define the unsigned error (UE) (in units of hartree/bohr) and relative error (RE) (unitless) as

$$UE = |\text{Analytic} - \text{Numerical}| \quad (83)$$

$$RE = \left| \frac{\text{Analytic} - \text{Numerical}}{\text{Numerical}} \right| \quad (84)$$

IV. RESULTS AND DISCUSSION

A. Validation of Analytic Gradients Using Diatomic Molecules

We first test our analytic gradient implementation with translated and fully-translated functionals at a variety of points on the potential energy curves of two diatomic systems: HeH⁺⁶²⁻⁶⁶ and LiH.⁶⁷⁻⁷¹ Analytic and numerical

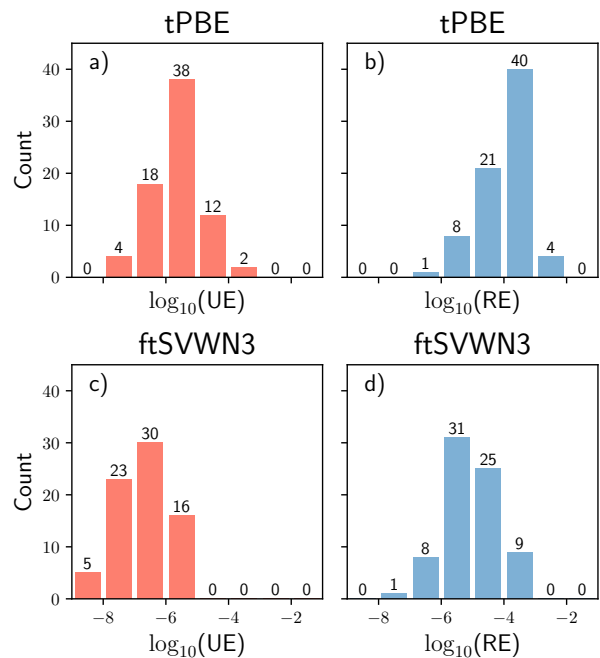


FIG. 1. Distribution of the common log of the unsigned (a,c) and relative (b,d) error of analytic gradients relative to numerical gradients for all states of HeH⁺ at various internuclear distances. The top row (a,b) is using the tPBE functional and bottom row (c,d) is using the ftSVWN3 functional.

gradients were computed for interatomic distances in the range 0.4 Å to 4.0 Å with a step size of 0.1 Å. As both the common log of UE and RE occur in roughly standard distributions, we present the errors as histograms.

HeH⁺ is the simplest possible system for which many terms in the programmable equations for L-PDFT analytic gradients do not vanish due to symmetry. The $\log_{10}(UE)$ and $\log_{10}(RE)$ distributions for both the tPBE and ftSVWN3 functional are shown in Fig. 1. For the tPBE functional, the majority of the UEs are below

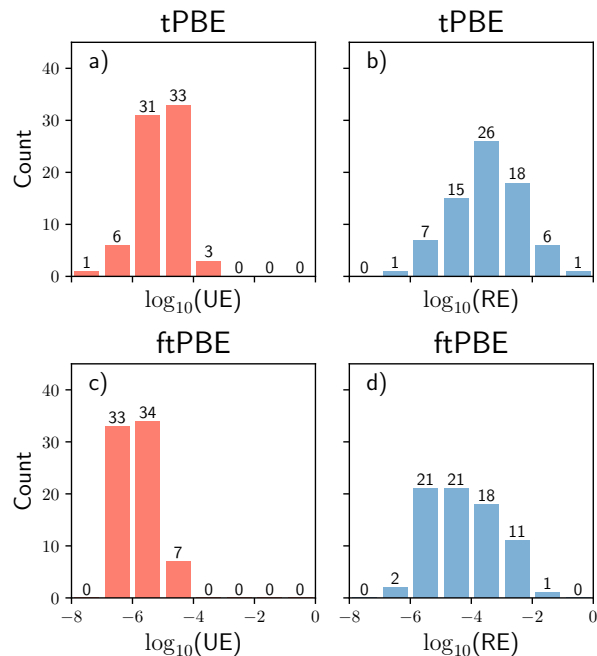


FIG. 2. Distribution of the common log of the unsigned (a,c) and relative (b,d) error of analytic gradients relative to numerical gradients for all states of LiH at various internuclear distances. The top row (a,b) is for the tPBE functional, and the bottom row (c,d) is for the ftPBE functional.

1×10^{-5} hartree/bohr whereas for the ftSVWN3 functional they are below 1×10^{-6} hartree/bohr. The RE is fairly constant throughout the potential energy curve for both functionals (Fig. S1 and S2) with the RE for tPBE being an order of magnitude greater than for ftSVWN3.

The $\log_{10}(\text{UE})$ and $\log_{10}(\text{RE})$ distributions using both the tPBE and ftPBE functionals for the lowest two $^1\Sigma^+$ states of LiH are summarized in Fig. 2. We see that the majority of UEs are below 1×10^{-4} hartree/bohr for tPBE and below 1×10^{-5} hartree/bohr for ftPBE. There are slightly larger REs for both functionals in the dissociation region of the potential energy curve (Figs. S3 and S4) which are likely due to the flatness of the potential. For example, the largest REs for tPBE and ftPBE occur on the upper $^1\Sigma^+$ state at 2.7 Å and 2.8 Å with analytic gradients of 1.7532×10^{-3} hartree/bohr and 9.0825×10^{-5} hartree/bohr respectively.

Table II summarizes the statistical agreement between the analytic and numerical gradients for both systems and all functionals. Overall, tPBE has mean unsigned error (MUE) of 1.4×10^{-5} hartree/bohr and 2.5×10^{-5} hartree/bohr for HeH⁺ and LiH respectively; these are consistent with our previous implementations of MC-PDFT gradients. There, we saw that the analytic and numerical gradients for LiH computed with tPBE and CMS-tPBE had MUEs of 4×10^{-5} hartree/bohr³² and 4.9×10^{-5} hartree/bohr⁶¹ respectively.

The fully-translated functionals have better agreement

TABLE II. Mean signed error (MSE), mean unsigned error (MUE), and root-mean-squared error (RMSE) of the analytic gradients relative to the numerical gradients. All values are in hartree/bohr.

Functional	HeH ⁺		LiH	
	tPBE	ftSVWN3	tPBE	ftPBE
MSE	2.2×10^{-6}	-1.2×10^{-7}	-3.8×10^{-6}	5.3×10^{-7}
MUE	1.4×10^{-5}	9.0×10^{-7}	2.5×10^{-5}	4.4×10^{-6}
RMSE	5.1×10^{-5}	1.9×10^{-6}	5.1×10^{-5}	9.1×10^{-6}

with the numerical gradients for both HeH⁺ and LiH (Table II). This can likely be attributed to the fact that translated functionals have a discontinuity in the first derivative of the translation scheme (where values of R in Eq. A9 equal 1); and since L-PDFT gradients require the second derivative, this discontinuity leads to larger numerical instabilities. The fully-translated functionals, on the other hand, use a polynomial interpolation to avoid the discontinuity (Appendix A), and this likely leads to more accurate gradients.

Overall, we find good agreement between the analytic and numerical gradients at all geometries considered for both HeH⁺ and LiH using both translated and fully-translated functionals.

B. Formaldehyde

Here we consider calculations with a full-valence active space of the optimized ground state (S_0) and the first excited state of $n \rightarrow \pi^*$ character (S_1) of formaldehyde. In the ground state, formaldehyde is a planar molecule with C_{2v} symmetry. For the first excited state of formaldehyde, it is known that the C=O double bond is elongated, and the molecule is no longer planar. We define η to be the angle between the H-C-H plane and the C=O bond, which is a measure of the nonplanarity of the structure. We will take as our reference the experimental ground- and excited-state geometries reported by Duncan⁷² and Jensen and Bunker⁷³ respectively.

The L-PDFT ground- and excited-state structural parameters are summarized in Table III, where they are compared to results obtained by other methods including MC-PDFT³² and CASPT2.⁷⁴ Both the MC-PDFT and CASPT2 calculations utilize a full-valence active space, with the MC-PDFT calculation also including two additional oxygen lone-pair orbitals. Additionally, there are results⁶¹ from MC-PDFT and compressed multi-state PDFT (CMS-PDFT)²⁶ using a smaller (6,5) active space. This smaller active space was chosen using the ABC2 automatic active-space selection scheme⁷⁵ by setting the parameters A , B , and C to 3, 2, and 0 respectively. We also include results from high-level single-reference methods computed by Budzák, Scalmani, and Jacquemin⁷⁴ including second-order algebraic diagrammatic construction (ADC(2)),⁷⁶ second-order coupled cluster (CC2),⁷⁷

TABLE III. L-PDFT ground and first excited state bond lengths, bond angles, and out-of-plane dihedral (η) for formaldehyde compared with various methods. All bond lengths are in Å and all angles are in degrees. Experimental uncertainty shown in parentheses.

State	Method	Basis Set	r_{CO}	r_{CH}	θ_{HCH}	η
S_0	L-PDFT(12,10)	jun-cc-pVTZ	1.210	1.115	116.4	0
	SA-CASSCF(12,12) ³²	aug-cc-pVTZ	1.214	1.103	117.3	0
	MC-PDFT(12,12) ³²	aug-cc-pVTZ	1.210	1.114	116.1	0
	MC-PDFT(6,5) ⁶¹	jun-cc-pVTZ	1.202	1.112	115.8	0
	CMS-PDFT(6,5) ⁶¹	jun-cc-pVTZ	1.203	1.112	115.8	0
	CASPT2(12,10) ⁷⁴	aug-cc-pVTZ	1.209	1.102	116.1	0
	ADC(2) ⁷⁴	aug-cc-pVTZ	1.209	1.096	116.5	0
	CCSD ⁷⁴	aug-cc-pVTZ	1.201	1.097	116.4	0
	CCSDR(3) ⁷⁴	aug-cc-pVTZ	1.207	1.100	116.4	0
	CC2 ⁷⁴	aug-cc-pVTZ	1.217	1.098	116.4	0
	CC3 ⁷⁴	aug-cc-pVTZ	1.208	1.100	116.2	0
	expt. ⁷²		1.207(1)	1.117(1)	116.2(1)	0
	S_1	L-PDFT(12,10)	jun-cc-pVTZ	1.328	1.100	118.1
SA-CASSCF(12,12) ³²		aug-cc-pVTZ	1.356	1.079	118.1	32
MC-PDFT(12,12) ³²		aug-cc-pVTZ	1.323	1.102	117.6	28
MC-PDFT(6,5) ⁶¹		jun-cc-pVTZ	1.333	1.095	119.9	30
CMS-PDFT(6,5) ⁶¹		jun-cc-pVTZ	1.333	1.095	119.9	30
CASPT2(12,10) ⁷⁴		aug-cc-pVTZ	1.326	1.090	118.1	38
ADC(2) ⁷⁴		aug-cc-pVTZ	1.380	1.081	123.8	19
CCSD ⁷⁴		aug-cc-pVTZ	1.300	1.087	118.9	30.9
CCSDR(3) ⁷⁴		aug-cc-pVTZ	1.320	1.089	118.2	37
CC2 ⁷⁴		aug-cc-pVTZ	1.353	1.085	121.3	29.5
CC3 ⁷⁴		aug-cc-pVTZ	1.326	1.089	118.3	37
expt. ⁷³			1.323(3)	1.103(1)	118.1(1)	34

third-order coupled cluster (CC3),^{78,79} and coupled cluster response method with single and double excitations and noniterative connected triple excitations from CC3 (CCSDR(3)).⁸⁰

Relative to the experimental geometry, the L-PDFT ground-state C=O bond length differs by 0.003 Å, the C–H bond length differs by 0.002 Å, and the H–C–H bond angle differs by 0.2° (Table III). For the excited state, the L-PDFT structure has a deviation of 0.005 Å for the C=O bond length, 0.003 Å for the C–H bond length, and 1° for the out-of-plane dihedral (η) relative to the experimental geometry (Table III). The L-PDFT H–C–H bond angle agrees with the experimental value to within 0.1°. L-PDFT has the most accurate η and H–C–H bond angle values of any of the methods presented.

Overall, for predicting the ground- and excited-state structures of formaldehyde, L-PDFT performs similarly to MC-PDFT with the slightly larger (12,12) active space and also similarly to the much more expensive CASPT2 method with the same active space.

Table IV summarizes the adiabatic and vertical excitation energies calculated by L-PDFT and compares them with results from the literature. All values exclude the vibrational ZPE. We take the experimental vertical excitation energy measured by electron-impact spectroscopy as our reference.⁸² It is not possible to compare the adia-

TABLE IV. L-PDFT adiabatic and vertical excitations in eV (not including vibration ZPE) for the first excited state of formaldehyde compared to reported values in the literature.

Method	Basis Set	Adiabatic	Vertical
L-PDFT(12,10)	jun-cc-pVTZ	3.61	3.98
SA-CASSCF(12,12) ³²	aug-cc-pVTZ	3.56	4.04
MC-PDFT(12,12) ³²	aug-cc-pVTZ	3.58	3.92
MC-PDFT(6,5) ⁶¹	jun-cc-pVTZ	3.65	4.07
CMS-PDFT(6,5) ⁶¹	jun-cc-pVTZ	3.65	4.07
CASPT2(12,10) ⁷⁴	aug-cc-pVTZ	3.53	3.92
ADC(2) ⁷⁴	aug-cc-pVTZ		3.92
CCSDR(3) ⁷⁴	aug-cc-pVTZ		3.97
CC3 ^{74,81}	aug-cc-pVTZ	3.58	3.96
expt. ⁸²			3.79

batic excitation energy to experiments without the vibrational ZPE. Instead, since the CC3 method is known to get within 0.03 eV of the extrapolated full configuration interaction for a variety of molecules,⁸¹ we take CC3 to be our reference for the adiabatic excitation energy.

Like MC-PDFT, L-PDFT overestimates the vertical excitation energy with a difference of 0.19 eV relative to the experimental value.⁸² However, this should be considered in the context that all methods presented in Table IV deviate by more than 0.1 eV from the experimen-

tal vertical excitation energy, which may itself have some uncertainty. The L-PDFT predicted vertical excitation energy only differs from the CC3 result by 0.02 eV and the much more expensive CASPT2 by 0.06 eV. Additionally, L-PDFT almost exactly reproduces the CC3 adiabatic excitation energy, with a slightly higher excitation energy as compared to CASPT2. Overall, L-PDFT performs similarly to CC3, MC-PDFT, and CASPT2 in predicting both the adiabatic and vertical excitation energies of formaldehyde.

C. *s-trans*-butadiene

Unlike formaldehyde, *s-trans*-butadiene has been shown to have a strong multireference character even in its 1A_g ground state.⁸⁴ Table V summarizes the selected optimized structural parameters of the 1 and 2 1A_g states of *s-trans*-butadiene as calculated by L-PDFT and other multireference methods.³² All active spaces are comprised of four electrons in two π orbitals and two π^* orbitals. Our reference for the ground-state structure is the experimental geometry from Haugen *et al.*⁸³, whereas for the excited-state structure we take our prior results calculated at the CASPT2(4,4) level of theory as our reference.³²

In general, L-PDFT performs similarly to both MC-PDFT and CASPT2 at predicting the equilibrium structures. For the ground state, L-PDFT deviates slightly from MC-PDFT for the C–C bond length (difference of 0.007 Å), but still does better than CASPT2. Both L-PDFT and MC-PDFT deviate the most from the experimental C=C bond length. For the excited state, L-PDFT and MC-PDFT agree on nearly every parameter except for the C–C bond length where they differ by only 0.002 Å. Overall, L-PDFT performs similarly to both MC-PDFT and CASPT2 at predicting the ground- and excited-state structures of the challenging *s-trans*-butadiene molecule.

The theoretical best estimate (TBE) of the vertical excitation energy of *s-trans*-butadiene to the 2 1A_g state is 6.39 eV.⁸⁵ Multi-state CASPT2⁸⁶ using a (4,4) active space at the equilibrium geometry predicts a vertical excitation of 6.69 eV,⁸⁴ 0.3 eV above the TBE. Table VI summarizes the adiabatic and vertical excitation energies of L-PDFT as compared to the previously computed SA-CASSCF, MC-PDFT, and CASPT2 energies using a two-state model space with a (4,4) active space.³² Although the L-PDFT vertical excitation energy differs by only 0.1 eV from the CASPT2 predicted adiabatic excitation energy and by only 0.01 eV from MC-PDFT, it overestimates the vertical excitation energy by more than 0.5 eV as compared to the best available estimate.

D. Phenol

The photochemistry of phenol has been extensively studied as it is a prototype of the $^1\pi\sigma^*$ motif which is common in a variety of biomolecules and aromatic compounds.^{12,87–94} In the original L-PDFT paper, we studied the O–H photodissociation potential energy surface and found that L-PDFT was able to correctly model the potential energy surface near the conical intersection, whereas MC-PDFT surfaces unphysically crossed.²⁷ Our active space in this paper is the same as in our prior studies of phenol^{26,27,61} consisting of $3(\pi, \pi^*)$, the p_z of O, and the C–O and O–H σ and σ^* orbitals.

Selected L-PDFT optimized ground- and first excited-state internal coordinates of phenol are presented in Table VII and are compared with results from other, similar methods. Our reference for the ground- and excited-state geometries are the experimental structures determined by Larsen⁹⁵ and Spangenberg, Imhof, and Kleiner⁹⁶ respectively. All of the PDFT and CASSCF methods in Table VII use the same (12,11) active space. We also include results from a high-level semiempirical fit that was designed to replicate the ZPE-inclusive experimental adiabatic excitation energy.⁹² All of the methods in the table predict relatively similar ground-state geometries, in line with the experimentally determined geometry.⁹⁵ Of interest is that the excited-state geometry optimized with MC-PDFT is nonplanar, with a substantial C–C–O–H dihedral of 14.3°.⁶¹ It was noted that CMS-PDFT does not suffer from this incorrect nonplanarity because it correctly incorporates the state interaction between S_0 and S_1 . L-PDFT also correctly predicts a planar excited-state geometry, in agreement with CMS-PDFT⁶¹ and experimental results.⁹⁶ This confirms that L-PDFT accounts for the state interaction as well as CMS-PDFT does. Overall, L-PDFT performs better than MC-PDFT in accurately predicting the first excited-state geometry of phenol, and the results are similar to those for CMS-PDFT and other high-level methods.

Table VIII contains a summary of the vertical and adiabatic excitation energies of phenol as obtained by L-PDFT and other methods. As in our previous work,⁶¹ we take as our reference the high-level semiempirical fit by Zhu, Malbon, and Yarkony⁹² for both excitation energies. We also include results from CC2⁹⁹ (for which the ground-state geometries are optimized at the second-order Møller-Plesset perturbation theory (MP2)^{104,105}), internally contracted multireference configuration interaction (MRCI)¹⁰⁶ based on a CASSCF(10,9) wave function,¹⁰⁰ and CASPT2 using a (8,8)⁹⁷ and (10,10)⁹⁸ active space. Both the MRCI and CASPT2 results were computed at their respective reference CASSCF optimized geometries.

L-PDFT performs similarly to the previously reported MC-PDFT results⁶¹ and overestimates the vertical and adiabatic excitation energy relative to the reference by about 0.2 eV. Comparatively, CASPT2 using both an (8,8)⁹⁷ and (10,10)⁹⁸ active space underestimates the

TABLE V. Selected L-PDFT optimized internal coordinates for the ground and excited 1A_g states of *s-trans*-butadiene as compared to results from other methods. All bond lengths are in Å and all bond angles are in degrees. Experimental uncertainty shown in parentheses.

State	Method	Basis Set	$r_{C=C}$	r_{CC}	θ_{CCC}
$1 {}^1A_g$	L-PDFT(4,4)	jul-cc-pVTZ	1.335	1.460	124.1
	SA-CASSCF(4,4) ³²	aug-cc-pVTZ	1.345	1.456	124.3
	MC-PDFT(4,4) ³²	aug-cc-pVTZ	1.336	1.470	124.1
	CASPT2(4,4) ³²	aug-cc-pVTZ	1.342	1.454	123.6
	expt. ⁸³		1.343(1)	1.467(1)	122.8(5)
$2 {}^1A_g$	L-PDFT(4,4)	jul-cc-pVTZ	1.496	1.399	124.1
	SA-CASSCF(4,4) ³²	aug-cc-pVTZ	1.489	1.413	123.2
	MC-PDFT(4,4) ³²	aug-cc-pVTZ	1.496	1.397	124.1
	CASPT2(4,4) ³²	aug-cc-pVTZ	1.488	1.394	122.1

TABLE VI. L-PDFT adiabatic and vertical excitations in eV (not including vibrational ZPE) for the $2 {}^1A_g$ state of *s-trans*-butadiene compared to reported values in the literature. The experimental uncertainty is shown in parentheses.

Method	Basis Set	Adiabatic	Vertical
L-PDFT(4,4)	jul-cc-pVTZ	5.78	6.92
SA-CASSCF(4,4) ³²	aug-cc-pVTZ	5.42	6.57
MC-PDFT(4,4) ³²	aug-cc-pVTZ	5.77	6.91
CASPT2(4,4) ³²	aug-cc-pVTZ	5.68	6.68
CCSD ^{a84}	6-31G**		7.69
MS-CASPT2(4,4) ^{a84}	6-31G**		6.69
TBE ^{b85}			6.39(7)

^a Excitation energy calculated at experimental equilibrium geometry.

^b Theoretical best estimate.

vertical and adiabatic excitation energy by more than 0.2 eV. The relative difference between the L-PDFT vertical and adiabatic excitation energies is 0.18 eV which is similar to the relative difference of 0.17 eV predicted by the semiempirical fit.⁹²

E. Cytosine

Finally, we report the optimized ground- and excited-state geometries and the adiabatic and vertical excitation energies of the nucleobase cytosine (Fig. 3). Previous studies have shown that the (14,10) active space composed of five π , two lone-pair, and three π^* orbitals is sufficient for studying the low-lying excited states of cytosine.^{107–109}

For the ground state, we take the experimental structure determined by Barker and Marsh¹⁰² as our reference. Tables IX and X contain selected optimized bond lengths and angles respectively obtained from L-PDFT and other methods. All methods presented in Table IX perform similarly with a mean unsigned deviation (MUD) of 0.02 Å relative to the experimental bond

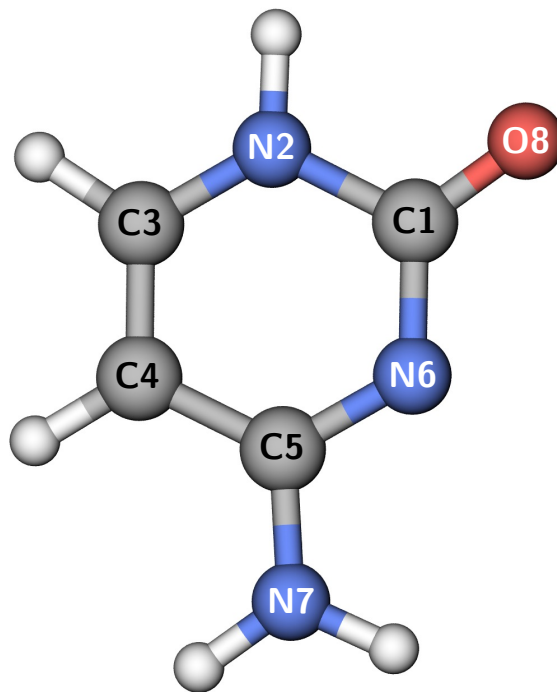


FIG. 3. Ground-state geometry of cytosine optimized with L-PDFT.

lengths. L-PDFT performs identically to MC-PDFT for determining the cytosine bond angles with an MUD of 1.6° for both methods.

Due to the lack of experimental data for the $2 {}^1A'$ relaxed geometry, we take the MS-CASPT2 with an (8,7) active space from the study by Nakayama, Yamazaki, and Taketsugu¹⁰³ as our reference. Tables XI and XII compares the same selected $2 {}^1A'$ optimized bond lengths and bond angles for L-PDFT and other methods. All methods give similar optimized bond lengths for the excited state with L-PDFT and MC-PDFT both having a MUD of 0.02 Å. CASPT2 and SA-CASSCF have relatively larger MUDs of 0.04 Å and 0.05 Å respectively.

TABLE VII. L-PDFT selected internal coordinates for the ground- and first excited-state of phenol compared with various methods. All bond lengths are in Å and all angles are in degrees. The experimental uncertainty is shown in parentheses.

State	Method	Basis Set	Avg. r_{CC}	r_{CO}	r_{OH}	θ_{COH}	τ_{CCOH}
S_0	L-PDFT(12,11)	jul-cc-pVDZ	1.400	1.369	0.962	109.4	0.0
	SA-CASSCF(12,11) ⁶¹	jul-cc-pVDZ	1.399	1.384	0.966	109.3	0.0
	MC-PDFT(12,11) ⁶¹	jul-cc-pVDZ	1.401	1.370	0.964	109.2	0.0
	CMS-PDFT(12,11) ⁶¹	jul-cc-pVDZ	1.398	1.367	0.966	109.2	0.0
	semiemp. fit ⁹²		1.395	1.382	0.965	108.5	0.0
	expt. ⁹⁵		1.393	1.375(5)	0.957(6)	108.8(4)	0.0
S_1	L-PDFT(12,11)	jul-cc-pVDZ	1.431	1.346	0.964	109.4	0.0
	SA-CASSCF(12,11) ⁶¹	jul-cc-pVDZ	1.434	1.379	0.960	109.3	0.0
	MC-PDFT(12,11) ⁶¹	jul-cc-pVDZ	1.429	1.337	0.975	108.2	14.3
	CMS-PDFT(12,11) ⁶¹	jul-cc-pVDZ	1.435	1.362	0.962	109.2	0.0
	semiemp. fit ⁹²		1.427	1.367	0.963	108.7	0.0
	expt. ⁹⁶		1.423	1.356	0.992	108.8	0.0

TABLE VIII. L-PDFT adiabatic and vertical excitations in eV (not including vibration ZPE) for the first excited state of phenol compared to reported values in the literature.

Method	Basis Set	Adiabatic	Vertical
L-PDFT(12,11)	jul-cc-pVDZ	4.85	5.03
SA-CASSCF(12,11) ⁶¹	jul-cc-pVDZ	4.73	4.93
MC-PDFT(12,11) ⁶¹	jul-cc-pVDZ	4.83	5.03
CMS-PDFT(12,11) ⁶¹	jul-cc-pVDZ	4.72	4.93
CASPT2(8,8) ^{a97}	cc-pVDZ	4.36	4.64
CASPT2(10,10) ^{a98}	aug(O)-AVTZ ^b	4.37	4.52
CC2 ^{c99}	aug-cc-pVDZ	4.67	4.86
MRCI(10,9) ^{ad100}	aug-cc-pVDZ	4.82	4.75
semiemp. fit ⁹²		4.66	4.83

^a Excitations calculated at the CASSCF optimized geometry.

^b Modified aug-cc-pVTZ basis set with extra even tempered sets of s and p diffuse functions on the oxygen atom.

^c Ground state optimized with MP2 and excited state optimized with CC2.

^d It is possible for the vertical excitation to be lower than the adiabatic excitation when the excitations are computed at geometries optimized at a different level of theory.

L-PDFT and MC-PDFT predict similar bond angles of the excited state, with L-PDFT being slightly closer to the MS-CASPT2 results with a MUD of 1.1° . CASPT2 has the widest difference in the bond angles, and it overestimates the 7-5-6 bond angle by about 7° as compared to all of the other methods.

Table XIII summarizes the adiabatic and vertical excitation energies for the $2^1A'$ state of cytosine computed at various levels of theory. We take the experimental vertical excitation from Abouaf *et al.*¹¹¹, and the MS-CASPT2(12,9) result (ground-state geometry optimized with MP2 and excited-state geometry optimized with MS-CASPT2(8,7)) as the reference for the adiabatic excitation energy Nakayama, Yamazaki, and Taketsugu¹⁰³. All methods, except CC2, perform similarly with L-PDFT underestimating the reference vertical ex-

citation energy the most. CC2 gets the closest to the experimental vertical excitation energy, differing only by 0.11 eV. Both L-PDFT and MC-PDFT get within 0.1 eV of the MS-CASPT2 adiabatic excitation energy. Overall, L-PDFT differs from MC-PDFT by only 0.03 eV for both the adiabatic and vertical excitation energies.

V. CONCLUSION

We presented the derivation and implementation of analytic nuclear gradients for L-PDFT calculations based on SA-CASSCF wave functions. Because the L-PDFT wave function is not fully variational with respect to all its parameters, we used a Lagrangian method similar to that used previously for SA-CASSCF and MC-PDFT analytic gradients. As in SA-CASSCF, we assumed equal weights to exclude the model states from the response equations. We then implemented the gradients in PYSCF-FORGE, which is a library of PYSCF extensions, and we showed that they agree with numerical gradients for both HeH^+ and LiH using both translated and fully-translated functionals. We showed the utility of the L-PDFT analytic gradients by optimizing the ground and first excited singlet states of formaldehyde, *s-trans*-butadiene, phenol, and the nucleobase cytosine. Whereas MC-PDFT predicts a nonplanar first excited state of phenol, we showed that L-PDFT correctly predicts a planar structure. Additionally, we computed the vertical and adiabatic excitation energy for each molecule and saw that L-PDFT performs similarly for excitation energies to MC-PDFT and other high-level multireference methods like CASPT2.

These results are consistent with our prior study and benchmarking of L-PDFT for calculating vertical excitation energies and for modeling potential energy surfaces.^{27,29} The results are especially encouraging because of the low cost of L-PDFT relative to CASPT2 and MRCI. We conclude that L-PDFT is promising new tool

TABLE IX. Selected L-PDFT ground state cytosine bond lengths (in Å) compared with similar methods and experimental quantities. Atoms are labeled according to Fig. 3.

Method	Basis Set	C1-N2	N2-C3	C3-C4	C4-C5	C5-N7	C5-N6	N6-C1	C1-O8	MUD ^a
L-PDFT(14,10)	jul-cc-pVTZ	1.438	1.344	1.361	1.434	1.352	1.327	1.363	1.218	0.02
SA-CASSCF(14,10) ³²	aug-cc-pVTZ	1.391	1.354	1.346	1.446	1.350	1.291	1.391	1.296	0.02
MC-PDFT(14,10) ³²	aug-cc-pVTZ	1.440	1.340	1.358	1.436	1.351	1.326	1.364	1.215	0.02
CCSD ¹⁰¹	TZP	1.416	1.360	1.353	1.446	1.357	1.313	1.379	1.214	0.02
CASPT2(14,10) ³²	6-311G+(2df)	1.420	1.362	1.357	1.443	1.360	1.320	1.378	1.222	0.02
expt. ¹⁰²		1.374	1.357	1.342	1.424	1.330	1.337	1.364	1.234	

^a Mean unsigned deviation from experiment.

TABLE X. Selected L-PDFT ground state cytosine bond angles (in degrees) compared with similar methods and experimental quantities. Atoms are labeled according to Fig. 3.

Method ^a	θ_{6-1-2}	θ_{5-6-1}	θ_{4-5-6}	θ_{3-2-1}	θ_{3-4-5}	θ_{4-3-2}	θ_{8-1-2}	θ_{8-1-6}	θ_{7-5-6}	θ_{7-5-4}	MUD ^b
L-PDFT(14,10)	116.0	120.4	124.2	123.2	115.8	120.4	117.3	126.6	116.2	119.6	1.6
SA-CASSCF(14,10) ³²	119.3	119.3	120.6	122.3	118.4	120.3	120.0	120.8	118.0	121.7	0.9
MC-PDFT(14,10) ³²	115.8	120.5	124.0	123.4	115.7	119.6	117.5	117.7	116.4	119.5	1.6
CCSD ¹⁰¹	119.7	120.1	119.5	121.9	118.8	120.2	120.9	119.6	119.2	121.5	1.3
CASPT2(14,10) ³²	119.9	120.1	119.3	121.7	118.9	120.3	120.9	119.4	119.5	121.4	1.4
expt. ¹⁰²	119.1	119.9	122.0	122.7	117.3	120.1	119.8	122.2	118.2	119.9	

^a Methods are the same as in Table IX.

^b Mean unsigned deviation from experiment.

for studying excited-state geometries, both vertical and adiabatic energies, photochemical reactions, and electronically nonadiabatic dynamics.

SUPPLEMENTARY MATERIAL

See the supplementary material associated with this article for the analytic and numerical gradients of HeH⁺ and LiH at each geometry and L-PDFT optimized structures with their corresponding energies.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under Grant No. CHE-2054723. M.R. Hennefarth acknowledges support by the National Science Foundation Graduate Research Fellowship under Grant No. 2140001. We also acknowledge the University of Chicago's Research Computing Center for their support of this work. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Appendix A: The Translated and Fully-Translated On-Top Functionals

Current generation on-top functionals include translated¹⁹ or fully-translated⁴⁴ KS local-spin density approximations or generalized gradient approximations functionals. These on-top functionals are defined such that

$$E^{\text{ot}}[\vec{\rho}] = E^{\text{xc}}[\vec{\rho}] \quad (\text{A1})$$

where E^{xc} is a KS exchange-correlation functional and $\vec{\rho}$ are the collective translated (or fully-translated) spin-density variables and their gradients.

$$\vec{\rho}^{\text{T}} = [\rho_{\uparrow} \ \rho_{\downarrow} \ \sigma_{\uparrow\uparrow} \ \sigma_{\uparrow\downarrow} \ \sigma_{\downarrow\downarrow}] \quad (\text{A2})$$

Here, ρ_{\uparrow} and ρ_{\downarrow} are effective spin densities, with primes denoting the gradient with respect to electron coordinate; and $\sigma_{\uparrow\uparrow}$, $\sigma_{\uparrow\downarrow}$, and $\sigma_{\downarrow\downarrow}$ being the inner product of the effective spin density gradients.

$$\sigma_{\uparrow\uparrow} = \rho'_{\uparrow} \cdot \rho'_{\uparrow} \quad (\text{A3a})$$

$$\sigma_{\uparrow\downarrow} = \rho'_{\uparrow} \cdot \rho'_{\downarrow} \quad (\text{A3b})$$

$$\sigma_{\downarrow\downarrow} = \rho'_{\downarrow} \cdot \rho'_{\downarrow} \quad (\text{A3c})$$

For translated functionals, the following mapping is used to generate the effective spin densities and their gradients from the wave function's density and on-top pair density:

$$\rho_{\uparrow} = \frac{\rho}{2}(1 + \zeta_t) \quad (\text{A4})$$

TABLE XI. Selected L-PDFT $2^1A'$ excited state cytosine bond lengths (in Å) compared with similar methods and experimental quantities. Atoms are labeled according to Fig. 3.

Method	Basis Set	C1-N2	N2-C3	C3-C4	C4-C5	C5-N7	C5-N6	N6-C1	C1-O8	MUD ^a
L-PDFT(14,10)	jul-cc-pVTZ	1.398	1.390	1.415	1.376	1.361	1.406	1.313	1.271	0.02
SA-CASSCF(14,10) ³²	aug-cc-pVTZ	1.358	1.381	1.424	1.365	1.372	1.417	1.274	1.328	0.04
MC-PDFT(14,10) ³²	aug-cc-pVTZ	1.429	1.382	1.416	1.385	1.369	1.416	1.342	1.236	0.02
CASPT2(14,10) ³²	6-311G+(2df)	1.453	1.301	1.394	1.444	1.350	1.320	1.374	1.210	0.05
MS-CASPT2(8,7) ¹⁰³	Sapporo-DZP	1.410	1.376	1.430	1.391	1.397	1.351	1.325	1.276	

^a Mean unsigned deviation from MS-CASPT2.

TABLE XII. Selected L-PDFT $2^1A'$ excited state cytosine bond angles (in degrees) compared with similar methods and experimental quantities. Atoms are labeled according to Fig. 3.

Method ^a	θ_{6-1-2}	θ_{5-6-1}	θ_{4-5-6}	θ_{3-2-1}	θ_{3-4-5}	θ_{4-3-2}	θ_{8-1-2}	θ_{8-1-6}	θ_{7-5-6}	θ_{7-5-4}	MUD ^b
L-PDFT(14,10)	122.7	117.1	122.6	121.7	119.9	115.9	114.1	123.2	111.9	125.5	1.1
SA-CASSCF(14,10) ³²	126.6	116.1	121.6	119.8	119.6	116.4	112.1	121.2	112.7	125.8	1.6
MC-PDFT(14,10) ³²	121.4	118.1	121.7	121.1	120.4	117.4	115.7	122.9	112.2	126.2	1.9
CASPT2(14,10) ³²	118.1	119.1	122.8	123.4	118.4	118.3	116.7	125.2	119.0	118.3	3.0
MS-CASPT2(8,7) ¹⁰³	123.0	115.8	123.4	122.7	119.6	115.4	111.8	125.2	112.8	123.7	

^a Methods are the same as in Table XI.

^b Mean unsigned deviation from MS-CASPT2.

TABLE XIII. L-PDFT adiabatic and vertical excitations in eV (not including vibration ZPE) for the $2^1A'$ state of cytosine compared to reported values in the literature.

Method	Basis Set	Adiabatic	Vertical
L-PDFT(14,10)	jul-cc-pVTZ	4.05	4.42
MC-PDFT(14,10) ³²	aug-cc-pVTZ	4.08	4.45
CC2 ¹¹⁰	aug-cc-pVDZ		4.54
MS-CASPT2(12,9) ^{a103}	Sapporo-DZP	3.98	4.48
expt. ¹¹¹			4.65

^a Ground-state geometry optimized with MP2 and excited-state geometry optimized with MS-CASPT2(8,7).

$$\rho_{\downarrow} = \frac{\rho}{2}(1 - \zeta_t) \quad (\text{A5})$$

$$\rho'_{\uparrow} = \frac{\rho'}{2}(1 + \zeta_t) \quad (\text{A6})$$

$$\rho'_{\downarrow} = \frac{\rho'}{2}(1 - \zeta_t) \quad (\text{A7})$$

where ζ_t is given by

$$\zeta_t = \begin{cases} \sqrt{1-R} & R \leq 1 \\ 0 & R > 1 \end{cases} \quad (\text{A8})$$

and R is proportional to the ratio of the on-top density to the density.

$$R = \frac{4\Pi}{\rho^2} \quad (\text{A9})$$

Both ζ_t and R are functions of \mathbf{r} . Functionals translated by this scheme are simply known as ‘translated’ functionals and are given the prefix ‘t’.¹⁹

The above translation scheme has a discontinuity in its first derivative at $R = 1$. The fully-translated scheme⁴⁴ fixes this by using a polynomial interpolation to smooth out the discontinuity in the region of R close to 1 as

$$\rho_{\uparrow} = \frac{\rho}{2}(1 + \zeta_{ft}) \quad (\text{A10})$$

$$\rho_{\downarrow} = \frac{\rho}{2}(1 - \zeta_{ft}) \quad (\text{A11})$$

$$\rho'_{\uparrow} = \frac{\rho'}{2}(1 + \zeta_{ft}) + \frac{\rho}{2}\zeta'_{ft} \quad (\text{A12})$$

$$\rho'_{\downarrow} = \frac{\rho'}{2}(1 - \zeta_{ft}) - \frac{\rho}{2}\zeta'_{ft} \quad (\text{A13})$$

where ζ_{ft} is

$$\zeta_{ft} = \begin{cases} \sqrt{1-R} & R < R_0 \\ P(R) & R_0 \leq R \leq R_1 \\ 0 & R > R_1 \end{cases} \quad (\text{A14})$$

and P is the polynomial interpolation from R_0 to R_1 .

$$P = P(R) = A(R - R_1)^5 + B(R - R_1)^4 + C(R - R_1)^3 \quad (\text{A15})$$

The gradient of ζ_{ft} is given by

$$\zeta'_{ft} = \begin{cases} -\frac{R'}{2\zeta_{ft}} & R < R_0 \\ R'P^{(1)} & R_0 \leq R \leq R_1 \\ 0 & R > R_1 \end{cases} \quad (\text{A16})$$

where \mathbf{v}^{xc} is the first derivative of ϵ^{xc} with respect to $\vec{\rho}$ (Eq. C2), \mathbf{f}^{xc} is defined as the symmetric matrix (here we show only the lower triangle) of the Hessian of ϵ^{xc} with respect to the translated effective spin-density variables and their gradients, $\mathbf{J}_{\vec{\rho}}$ is the translation Jacobian (Eq. C3),^{31,32} and $\mathbf{H}_{\vec{\rho}}$ is the Hessian of the translation. Both \mathbf{v}^{xc} and \mathbf{f}^{xc} are evaluated using standard KS density functional theory techniques. Specifically, in the PYSCF implementation, they are evaluated using LIBXC. Throughout the rest of this section, we will only show the lower triangular portion of all Hessian matrices since they are symmetric.

Due to the complexity of these equations, we will instead derive the necessary equations through a series of transformation with much more manageable Jacobians and Hessians. We will denote each set of coordinates (except the first and last corresponding to $\vec{\rho}$ and $\vec{\rho}$) as (a, b) which corresponds to the variables $a, b, \sigma_{aa}, \sigma_{ab}, \sigma_{bb}$. Correspondingly, we have that σ denotes the inner product between the gradient of the two variables as

$$\sigma_{aa} = a' \cdot a' \quad (\text{D3a})$$

$$\sigma_{ab} = a' \cdot b' \quad (\text{D3b})$$

$$\sigma_{bb} = b' \cdot b' \quad (\text{D3c})$$

Furthermore, we have that the gradient and Hessian of ϵ^{xc} with respect to these variables are denoted as \mathbf{v}^{ab} and \mathbf{f}^{ab} respectively.

$$\begin{aligned} \mathbf{v}^{ab} &= [\epsilon_a^{\text{xc}} \quad \epsilon_b^{\text{xc}} \quad \epsilon_{\sigma_{aa}}^{\text{xc}} \quad \epsilon_{\sigma_{ab}}^{\text{xc}} \quad \epsilon_{\sigma_{bb}}^{\text{xc}}] \\ &= \left[\frac{\partial \epsilon^{\text{xc}}}{\partial a} \quad \frac{\partial \epsilon^{\text{xc}}}{\partial b} \quad \frac{\partial \epsilon^{\text{xc}}}{\partial \sigma_{aa}} \quad \frac{\partial \epsilon^{\text{xc}}}{\partial \sigma_{ab}} \quad \frac{\partial \epsilon^{\text{xc}}}{\partial \sigma_{bb}} \right] \end{aligned} \quad (\text{D4})$$

$$\mathbf{f}^{ab} = \begin{bmatrix} \frac{\partial^2 \epsilon^{\text{xc}}}{\partial a^2} & & & & \\ \frac{\partial^2 \epsilon^{\text{xc}}}{\partial a \partial b} & \frac{\partial^2 \epsilon^{\text{xc}}}{\partial b^2} & & & \\ \frac{\partial^2 \epsilon^{\text{xc}}}{\partial a \partial \sigma_{aa}} & \frac{\partial^2 \epsilon^{\text{xc}}}{\partial b \partial \sigma_{aa}} & \frac{\partial^2 \epsilon^{\text{xc}}}{\partial \sigma_{aa}^2} & & \\ \frac{\partial^2 \epsilon^{\text{xc}}}{\partial a \partial \sigma_{ab}} & \frac{\partial^2 \epsilon^{\text{xc}}}{\partial b \partial \sigma_{ab}} & \frac{\partial^2 \epsilon^{\text{xc}}}{\partial \sigma_{aa} \partial \sigma_{ab}} & \frac{\partial^2 \epsilon^{\text{xc}}}{\partial \sigma_{ab}^2} & \\ \frac{\partial^2 \epsilon^{\text{xc}}}{\partial a \partial \sigma_{bb}} & \frac{\partial^2 \epsilon^{\text{xc}}}{\partial b \partial \sigma_{bb}} & \frac{\partial^2 \epsilon^{\text{xc}}}{\partial \sigma_{aa} \partial \sigma_{bb}} & \frac{\partial^2 \epsilon^{\text{xc}}}{\partial \sigma_{ab} \partial \sigma_{bb}} & \frac{\partial^2 \epsilon^{\text{xc}}}{\partial \sigma_{bb}^2} \end{bmatrix} \quad (\text{D5})$$

Furthermore, for the transformation of (a, b) to (c, d) , we will let \mathbf{J}_{cd}^{ab} and \mathbf{H}_{cd}^{ab} be the corresponding Jacobian and Hessian of the transformation respectively.

The first transformation step involves going from spin-separated electron density and its derivatives to charge density (ρ) and spin density (m) and their derivatives. In this sense, $(a, b) = (\rho, m)$ so that we are translating as follows:

$$\begin{bmatrix} \rho \\ m \\ \sigma_{\rho\rho} \\ \sigma_{\rho m} \\ \sigma_{mm} \end{bmatrix} \rightarrow \begin{bmatrix} \rho_{\uparrow} \\ \rho_{\downarrow} \\ \sigma_{\uparrow\uparrow} \\ \sigma_{\uparrow\downarrow} \\ \sigma_{\downarrow\downarrow} \end{bmatrix} \quad (\text{D6})$$

The coordinates are related by the following linear trans-

formation:

$$\begin{bmatrix} \rho_{\uparrow} \\ \rho_{\downarrow} \\ \sigma_{\uparrow\uparrow} \\ \sigma_{\uparrow\downarrow} \\ \sigma_{\downarrow\downarrow} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(\rho + m) \\ \frac{1}{2}(\rho - m) \\ \frac{1}{4}(\sigma_{\rho\rho} + \sigma_{mm}) + \frac{1}{2}\sigma_{\rho m} \\ \frac{1}{4}(\sigma_{\rho\rho} - \sigma_{mm}) \\ \frac{1}{4}(\sigma_{\rho\rho} + \sigma_{mm}) - \frac{1}{2}\sigma_{\rho m} \end{bmatrix} \quad (\text{D7})$$

Since this is a strictly linear transformation, we can see that \mathbf{f}^{xc} can be related to the Hessian of ϵ^{xc} with respect to (ρ, m) and their gradients as

$$\mathbf{f}^{\rho m} = \left(\mathbf{J}_{\rho m}^{\uparrow\downarrow} \right)^{\top} \cdot \mathbf{f}^{\text{xc}} \cdot \mathbf{J}_{\rho m}^{\uparrow\downarrow} \quad (\text{D8})$$

$$\mathbf{J}_{\rho m}^{\uparrow\downarrow} = \frac{1}{4} \begin{bmatrix} 2 & 2 & 0 & 0 & 0 \\ 2 & -2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -2 & 1 \end{bmatrix} \quad (\text{D9})$$

Most subsequent intermediate translation steps to the coordinates ρ and Π will differ depending on whether the functional is translated or fully-translated. We first start with the simpler translated case and then go to the fully-translated case. Generally speaking though, we will undergo the following change of variables:

1. (ρ, m) to (ρ, ζ) .
2. (ρ, ζ) to (ρ, R) .
3. (ρ, R) to (ρ, Π) .
4. (ρ, Π) to $\vec{\rho}$.

The Hessians are related to one another by

$$\mathbf{f}^{\rho\zeta} = \left(\mathbf{J}_{\rho\zeta}^{\rho m} \right)^{\top} \cdot \mathbf{f}^{\rho m} \cdot \mathbf{J}_{\rho\zeta}^{\rho m} + \mathbf{v}^{\rho m} \cdot \mathbf{H}_{\rho\zeta}^{\rho m} \quad (\text{D10})$$

$$\mathbf{f}^{\rho R} = \left(\mathbf{J}_{\rho R}^{\rho\zeta} \right)^{\top} \cdot \mathbf{f}^{\rho\zeta} \cdot \mathbf{J}_{\rho R}^{\rho\zeta} + \mathbf{v}^{\rho\zeta} \cdot \mathbf{H}_{\rho R}^{\rho\zeta} \quad (\text{D11})$$

$$\mathbf{f}^{\rho\Pi} = \left(\mathbf{J}_{\rho\Pi}^{\rho R} \right)^{\top} \cdot \mathbf{f}^{\rho R} \cdot \mathbf{J}_{\rho\Pi}^{\rho R} + \mathbf{v}^{\rho R} \cdot \mathbf{H}_{\rho\Pi}^{\rho R} \quad (\text{D12})$$

$$\mathbf{f}^{\text{ot}} = \left(\mathbf{J}_{\vec{\rho}}^{\rho\Pi} \right)^{\top} \cdot \mathbf{f}^{\rho\Pi} \cdot \mathbf{J}_{\vec{\rho}}^{\rho\Pi} + \mathbf{v}^{\rho\Pi} \cdot \mathbf{H}_{\vec{\rho}}^{\rho\Pi} \quad (\text{D13})$$

and the gradients are related by

$$\mathbf{v}^{\rho m} = \mathbf{v}^{\text{xc}} \cdot \mathbf{J}_{\rho m}^{\uparrow\downarrow} \quad (\text{D14})$$

$$\mathbf{v}^{\rho\zeta} = \mathbf{v}^{\rho m} \cdot \mathbf{J}_{\rho\zeta}^{\rho m} \quad (\text{D15})$$

$$\mathbf{v}^{\rho R} = \mathbf{v}^{\rho\zeta} \cdot \mathbf{J}_{\rho R}^{\rho\zeta} \quad (\text{D16})$$

$$\mathbf{v}^{\rho\Pi} = \mathbf{v}^{\rho\Pi} \cdot \mathbf{J}_{\rho\Pi}^{\rho R} \quad (\text{D17})$$

All Jacobians and Hessians used for the translated functionals will be prefixed with a ‘t’ (for example, ${}^t\mathbf{J}$ and ${}^t\mathbf{H}$), and Jacobians and Hessians used for fully-translated functionals will be prefixed with an ‘ft’ (for example, ${}^{\text{ft}}\mathbf{J}$ and ${}^{\text{ft}}\mathbf{H}$). Once we arrive at the coordinates of (ρ, Π) , we can then change the variables to $\vec{\rho}$, which will be the same for translated and fully-translated functionals.

1. Translated On-Top Hessian

We now go to the variables (ρ, ζ) by

$$\begin{bmatrix} \rho \\ m \\ \sigma_{\rho\rho} \\ \sigma_{\rho m} \\ \sigma_{mm} \end{bmatrix} = \begin{bmatrix} \rho \\ \rho\zeta \\ \sigma_{\rho\rho} \\ \zeta\sigma_{\rho\rho} \\ \zeta^2\sigma_{\rho\rho} \end{bmatrix} \quad (\text{D18})$$

Then

$${}^t\mathbf{J}_{\rho\zeta}^{\rho m} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \zeta & \rho & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \sigma_{\rho\rho} & \zeta & 0 & 0 \\ 0 & 2\zeta\sigma_{\rho\rho} & \zeta^2 & 0 & 0 \end{bmatrix} \quad (\text{D19})$$

$$\mathbf{v}^{\rho m} \cdot {}^t\mathbf{H}_{\rho\zeta}^{\rho m} = \begin{bmatrix} 0 \\ \epsilon_m^{\text{ot}} & 2\epsilon_{\sigma_{mm}}^{\text{ot}} \sigma_{\rho\rho} \\ 0 & \epsilon_{\sigma_{\rho m}}^{\text{ot}} + 2\epsilon_{\sigma_{mm}}^{\text{ot}} \zeta & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{D20})$$

Our next transformation is to (ρ, R) using

$$\begin{bmatrix} \rho \\ \zeta \\ \sigma_{\rho\rho} \\ \sigma_{\rho\zeta} \\ \sigma_{\zeta\zeta} \end{bmatrix} = \begin{bmatrix} \rho \\ f(R) \\ \sigma_{\rho\rho} \\ 0 \\ 0 \end{bmatrix} \quad (\text{D21})$$

Note that in the translated case, there is no dependence on ζ' ; therefore, the $\sigma_{\rho\zeta}$ and $\sigma_{\zeta\zeta}$ components do not contribute. Here, we are treating $f(R) = f$ and a general function of R where $f' = \frac{df}{dR}$. Our Jacobian and Hessian for this step are

$${}^t\mathbf{J}_{\rho R}^{\rho\zeta} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & f' & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{D22})$$

$$\mathbf{v}^{\rho\zeta} \cdot {}^t\mathbf{H}_{\rho R}^{\rho\zeta} = \begin{bmatrix} 0 \\ 0 & \epsilon_{\zeta}^{\text{ot}} f'' \\ 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{D23})$$

where f'' is the second derivative of ζ with respect to R .

Next we go to the (ρ, Π) variables by the following transformation:

$$\begin{bmatrix} \rho \\ R \\ \sigma_{\rho\rho} \\ \sigma_{\rho R} \\ \sigma_{RR} \end{bmatrix} = \begin{bmatrix} \rho \\ \frac{4\Pi}{\rho^2} \\ \sigma_{\rho\rho} \\ 0 \\ 0 \end{bmatrix} \quad (\text{D24})$$

Again, we can omit the $\sigma_{\rho R}$ and σ_{RR} variables since translated functionals do not depend on R' . This transformation results in

$${}^t\mathbf{J}_{\rho\Pi}^{\rho R} = \frac{4}{\rho^2} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{-2\Pi}{\rho} & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{D25})$$

$$\mathbf{v}^{\rho R} \cdot {}^t\mathbf{H}_{\rho\Pi}^{\rho R} = \frac{2\epsilon_{\rho R}^{\text{ot}}}{\rho^2} \begin{bmatrix} 3R \\ \frac{-4}{\rho} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{D26})$$

2. Fully-Translated On-Top Hessian

For the fully-translated case, going to the (ρ, m) variables modifies $\sigma_{\rho m}$ and σ_{mm} in Eq. D18 such that

$$\begin{bmatrix} \rho \\ m \\ \sigma_{\rho\rho} \\ \sigma_{\rho m} \\ \sigma_{mm} \end{bmatrix} = \begin{bmatrix} \rho \\ \rho\zeta \\ \sigma_{\rho\rho} \\ \zeta\sigma_{\rho\rho} + \rho\sigma_{\rho\zeta} \\ \zeta^2\sigma_{\rho\rho} + 2\rho\zeta\sigma_{\rho\zeta} + \rho^2\sigma_{\zeta\zeta} \end{bmatrix} \quad (\text{D27})$$

The fully-translated Jacobian and Hessian for this translation step are slight modifications of the translated matrices.

$${}^{\text{ft}}\mathbf{J}_{\rho\zeta}^{\rho m} = {}^t\mathbf{J}_{\rho\zeta}^{\rho m} + \tilde{\mathbf{J}}_{\rho\zeta}^{\rho m} \quad (\text{D28})$$

$$\tilde{\mathbf{J}}_{\rho\zeta}^{\rho m} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \sigma_{\rho\zeta} & 0 & 0 & \rho & 0 \\ 2(\rho\sigma_{\zeta\zeta} + \zeta\sigma_{\rho\zeta}) & 2\rho\sigma_{\rho\zeta} & 0 & 2\rho\zeta & \rho^2 \end{bmatrix} \quad (\text{D29})$$

$$\mathbf{v}^{\rho m} \cdot {}^{\text{ft}}\mathbf{H}_{\rho\zeta}^{\rho m} = \mathbf{v}^{\rho m} \cdot {}^t\mathbf{H}_{\rho\zeta}^{\rho m} + \mathbf{v}^{\rho m} \cdot \tilde{\mathbf{H}}_{\rho\zeta}^{\rho m} \quad (\text{D30})$$

$$\mathbf{v}^{\rho m} \cdot \tilde{\mathbf{H}}_{\rho\zeta}^{\rho m} = \begin{bmatrix} 2\epsilon_{\sigma_{mm}}^{\text{ot}} \sigma_{\zeta\zeta} & 0 \\ 2\epsilon_{\sigma_{mm}}^{\text{ot}} \sigma_{\rho\zeta} & 0 & 0 \\ 0 & 0 & 0 \\ \epsilon_{\sigma_{\rho m}}^{\text{ot}} + 2\epsilon_{\sigma_{mm}}^{\text{ot}} \zeta & 2\epsilon_{\sigma_{mm}}^{\text{ot}} \rho & 0 & 0 \\ 2\epsilon_{\sigma_{mm}}^{\text{ot}} \rho & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{D31})$$

For the next transformation step to (ρ, R) , we must include the transformations for $\sigma_{\rho\zeta}$ and $\sigma_{\zeta\zeta}$ so that Eq. D21 is modified to be

$$\begin{bmatrix} \rho \\ \zeta \\ \sigma_{\rho\rho} \\ \sigma_{\rho\zeta} \\ \sigma_{\zeta\zeta} \end{bmatrix} = \begin{bmatrix} \rho \\ f(R) \\ \sigma_{\rho\rho} \\ f' \sigma_{\rho R} \\ (f')^2 \sigma_{RR} \end{bmatrix} \quad (\text{D32})$$

This leads to the modified Jacobian and Hessian as

$${}^{\text{ft}}\mathbf{J}_{\rho R}^{\rho\zeta} = {}^{\text{t}}\mathbf{J}_{\rho R}^{\rho\zeta} + \tilde{\mathbf{J}}_{\rho R}^{\rho\zeta} \quad (\text{D33})$$

$$\tilde{\mathbf{J}}_{\rho R}^{\rho\zeta} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{\rho R} f'' & 0 & f' & 0 \\ 0 & 2\sigma_{RR} f'' f' & 0 & 0 & (f')^2 \end{bmatrix} \quad (\text{D34})$$

$$\mathbf{v}^{\rho\zeta} \cdot {}^{\text{ft}}\mathbf{H}_{\rho R}^{\rho\zeta} = \mathbf{v}^{\rho\zeta} \cdot {}^{\text{t}}\mathbf{H}_{\rho R}^{\rho\zeta} + \mathbf{v}^{\rho\zeta} \cdot \tilde{\mathbf{H}}_{\rho R}^{\rho\zeta} \quad (\text{D35})$$

$$\mathbf{v}^{\rho\zeta} \cdot \tilde{\mathbf{H}}_{\rho R}^{\rho\zeta} = \epsilon_{\sigma_{\rho\zeta}}^{\text{ot}} \mathbf{H}_{\rho R}^{\sigma_{\rho\zeta}} + \epsilon_{\sigma_{\zeta\zeta}}^{\text{ot}} \mathbf{H}_{\rho R}^{\sigma_{\zeta\zeta}} \quad (\text{D36})$$

where $\mathbf{H}_{\rho R}^{\sigma_{\rho\zeta}}$ and $\mathbf{H}_{\rho R}^{\sigma_{\zeta\zeta}}$ are the Hessians of $\sigma_{\rho\zeta}$ and $\sigma_{\zeta\zeta}$

with respect to the (ρ, R) variables.

$$\mathbf{H}_{\rho R}^{\sigma_{\rho\zeta}} = \begin{bmatrix} 0 \\ 0 & \sigma_{\rho R} f''' \\ 0 & 0 & 0 \\ 0 & f'' & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{D37})$$

$$\mathbf{H}_{\rho R}^{\sigma_{\zeta\zeta}} = \begin{bmatrix} 0 \\ 0 & 2\sigma_{RR} (f''' f' + (f'')^2) \\ 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 2f'' f' & 0 & 0 & 0 \end{bmatrix} \quad (\text{D38})$$

The final transformation we must consider separately for the fully-translated functionals is to the (ρ, Π) coordinate. Here, we must include the transformation of the $\sigma_{\rho R}$ and σ_{RR} , which are not included in the translated case. The modified form of Eq. D24 for the fully-translated case is thus

$$\begin{bmatrix} \rho \\ R \\ \sigma_{\rho\rho} \\ \sigma_{\rho R} \\ \sigma_{RR} \end{bmatrix} = \begin{bmatrix} \frac{\rho}{4\Pi} \\ \frac{\rho^2}{\rho^2} \\ \sigma_{\rho\rho} \\ \frac{4}{\rho^2} \left(\sigma_{\rho\Pi} - \frac{2\Pi}{\rho} \sigma_{\rho\rho} \right) \\ \frac{16}{\rho^4} \left(\frac{4\Pi^2}{\rho^2} \sigma_{\rho\rho} - \frac{4\Pi}{\rho} \sigma_{\rho\Pi} + \sigma_{\Pi\Pi} \right) \end{bmatrix} \quad (\text{D39})$$

The fully-translated Jacobian is given by

$${}^{\text{ft}}\mathbf{J}_{\rho\Pi}^{\rho R} = {}^{\text{t}}\mathbf{J}_{\rho\Pi}^{\rho R} + \tilde{\mathbf{J}}_{\rho\Pi}^{\rho R} \quad (\text{D40})$$

$$\tilde{\mathbf{J}}_{\rho\Pi}^{\rho R} = \frac{4}{\rho^2} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{2}{\rho} \left(\frac{3\Pi\sigma_{\rho\rho}}{\rho} - \sigma_{\rho\Pi} \right) & \frac{-2\sigma_{\rho\rho}}{\rho} & \frac{-2\Pi}{\rho} & 1 & 0 \\ \frac{16}{\rho^3} \left(\frac{5\Pi\sigma_{\rho\Pi}}{\rho} - \frac{6\Pi^2\sigma_{\rho\rho}}{\rho^2} - \sigma_{\Pi\Pi} \right) & \frac{16}{\rho^3} \left(\frac{2\Pi\sigma_{\rho\rho}}{\rho} - \sigma_{\rho\Pi} \right) & \frac{16\Pi^2}{\rho^4} & \frac{-16\Pi}{\rho^3} & \frac{4}{\rho^2} \end{bmatrix} \quad (\text{D41})$$

And the fully-translated Hessian term is given by

$$\mathbf{v}^{\rho R} \cdot {}^{\text{ft}}\mathbf{H}_{\rho\Pi}^{\rho R} = \mathbf{v}^{\rho R} \cdot {}^{\text{t}}\mathbf{H}_{\rho\Pi}^{\rho R} + \mathbf{v}^{\rho R} \cdot \tilde{\mathbf{H}}_{\rho\Pi}^{\rho R} \quad (\text{D42})$$

$$\mathbf{v}^{\rho R} \cdot \tilde{\mathbf{H}}_{\rho\Pi}^{\rho R} = \epsilon_{\sigma_{\rho R}}^{\text{ot}} \mathbf{H}_{\rho\Pi}^{\sigma_{\rho R}} + \epsilon_{\sigma_{RR}}^{\text{ot}} \mathbf{H}_{\rho\Pi}^{\sigma_{RR}} \quad (\text{D43})$$

with $\mathbf{H}_{\rho\Pi}^{\sigma_{\rho R}}$ and $\mathbf{H}_{\rho\Pi}^{\sigma_{RR}}$ the Hessian of $\sigma_{\rho R}$ and σ_{RR} with

respect to the (ρ, Π) variables given by

$$\mathbf{H}_{\rho\Pi}^{\sigma_{\rho R}} = \begin{bmatrix} \frac{24}{\rho^3} (\sigma_{\rho\Pi} - R\sigma_{\rho\rho}) \\ \frac{24\sigma_{\rho\rho}}{\rho^4} & 0 \\ \frac{6R}{\rho^2} & \frac{-8}{\rho^3} & 0 \\ \frac{-8}{\rho^3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{D44})$$

- ²⁵J. J. Bao, C. Zhou, Z. Varga, S. Kanchanakungwankul, L. Gagliardi, and D. G. Truhlar, “Multi-state pair-density functional theory,” *Faraday Discuss.* **224**, 348–372 (2020).
- ²⁶J. J. Bao, C. Zhou, and D. G. Truhlar, “Compressed-state multi-state pair-density functional theory,” *J. Chem. Theory Comput.* **16**, 7444–7452 (2020).
- ²⁷M. R. Hennefarth, M. R. Hermes, D. G. Truhlar, and L. Gagliardi, “Linearized pair-density functional theory,” *J. Chem. Theory Comput.* **19**, 3172–3183 (2023).
- ²⁸A. A. Granovsky, “Extended multi-configuration quasi-degenerate perturbation theory: The new approach to multi-state multi-reference perturbation theory,” *J. Chem. Phys.* **134**, 214113 (2011).
- ²⁹M. R. Hennefarth, D. S. King, and L. Gagliardi, “Linearized pair-density functional theory for vertical excitation energies,” *J. Chem. Theory Comput.* (2023), 10.1021/acs.jctc.3c00863.
- ³⁰J. Stålring, A. Bernhardsson, and R. Lindh, “Analytical gradients of a state average MCSCF state and a state average diagnostic,” *Mol. Phys.* **99**, 103–114 (2001).
- ³¹A. M. Sand, C. E. Hoyer, K. Sharkas, K. M. Kidder, R. Lindh, D. G. Truhlar, and L. Gagliardi, “Analytic gradients for complete active space pair-density functional theory,” *J. Chem. Theory Comput.* **14**, 126–138 (2018).
- ³²T. R. Scott, M. R. Hermes, A. M. Sand, M. S. Oakley, D. G. Truhlar, and L. Gagliardi, “Analytic gradients for state-averaged multiconfiguration pair-density functional theory,” *J. Chem. Phys.* **153**, 014106 (2020).
- ³³T. R. Scott, M. S. Oakley, M. R. Hermes, A. M. Sand, R. Lindh, D. G. Truhlar, and L. Gagliardi, “Analytic gradients for multiconfiguration pair-density functional theory with density fitting: Development and application to geometry optimization in the ground and excited states,” *J. Chem. Phys.* **154**, 074108 (2021).
- ³⁴H. Hellmann, “Zur rolle der kinetischen elektronenergie für die zwischenatomaren kräfte,” *Z. Physik* **85**, 180–190 (1933).
- ³⁵H. Hellmann, *Einführung in Die Quantenchemie* (Franz Deuticke, Leipzig, Germany, 1937).
- ³⁶R. P. Feynman, “Forces in molecules,” *Phys. Rev.* **56**, 340–343 (1939).
- ³⁷W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, 2nd ed. (Cambridge: Cambridge University Press, Cambridge, United Kingdom, 1992).
- ³⁸A. Bernhardsson, R. Lindh, J. Olsen, and M. Fulscher, “A direct implementation of the second-order derivatives of multiconfigurational SCF energies and an analysis of the preconditioning in the associated response equation,” *Mol. Phys.* **96**, 617–628 (1999).
- ³⁹T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, Ltd, West Sussex, England, 2000).
- ⁴⁰T. U. Helgaker and J. Almlöf, “A second-quantization approach to the analytical evaluation of response properties for perturbation-dependent basis sets,” *Int. J. Quantum Chem.* **26**, 275–291 (1984).
- ⁴¹P.-O. Löwdin, “On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals,” *J. Chem. Phys.* **18**, 365–375 (1950).
- ⁴²T. H. Dunning, “Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen,” *J. Chem. Phys.* **90**, 1007–1023 (1989).
- ⁴³J. P. Perdew, K. Burke, and M. Ernzerhof, “Generalized gradient approximation made simple,” *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- ⁴⁴R. K. Carlson, D. G. Truhlar, and L. Gagliardi, “Multiconfiguration pair-density functional theory: A fully translated gradient approximation and its performance for transition metal dimers and the spectroscopy of $\text{Re}_2\text{Cl}_8^{2-}$,” *J. Chem. Theory Comput.* **11**, 4077–4085 (2015).
- ⁴⁵F. Bloch, “Bemerkung zur elektronentheorie des ferromagnetismus und der elektrischen leitfähigkeit,” *Z. Physik* **57**, 545–555 (1929).
- ⁴⁶P. A. M. Dirac, “Note on exchange phenomena in the Thomas atom,” *Math. Proc. Camb. Philos. Soc.* **26**, 376–385 (1930).
- ⁴⁷S. H. Vosko, L. Wilk, and M. Nusair, “Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis,” *Can. J. Phys.* **58**, 1200–1211 (1980).
- ⁴⁸R. A. Kendall, T. H. Dunning, and R. J. Harrison, “Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions,” *J. Chem. Phys.* **96**, 6796–6806 (1992).
- ⁴⁹D. Feller, “The role of databases in support of computational chemistry calculations,” *J. Comput. Chem.* **17**, 1571–1586 (1996).
- ⁵⁰K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurmoothi, J. Chase, J. Li, and T. L. Windus, “Basis set exchange: A community database for computational sciences,” *J. Chem. Inf. Model.* **47**, 1045–1052 (2007).
- ⁵¹E. Papajak and D. G. Truhlar, “Convergent partially augmented basis sets for post-hartree-fock calculations of molecular properties and reaction barrier heights,” *J. Chem. Theory Comput.* **7**, 10–18 (2011).
- ⁵²P. Celani and H.-J. Werner, “Analytical energy gradients for internally contracted second-order multireference perturbation theory,” *J. Chem. Phys.* **119**, 5044–5057 (2003).
- ⁵³Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, “PySCF: The Python-based simulations of chemistry framework,” *WIREs Comput. Mol. Sci.* **8**, e1340 (2018).
- ⁵⁴Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purvanto, P. J. Robinson, E. Ronca, E. R. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun, S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Y. Sokolov, and G. K.-L. Chan, “Recent developments in the PySCF program package,” *J. Chem. Phys.* **153**, 024109 (2020).
- ⁵⁵M. A. L. Marques, M. J. T. Oliveira, and T. Burnus, “Libxc: A library of exchange and correlation functionals for density functional theory,” *Comput. Phys. Commun.* **183**, 2272–2281 (2012).
- ⁵⁶S. Lehtola, C. Steigemann, M. J. T. Oliveira, and M. A. L. Marques, “Recent developments in libxc — A comprehensive library of functionals for density functional theory,” *SoftwareX* **7**, 1–5 (2018).
- ⁵⁷Q. Sun, “Libcint: An efficient general integral library for Gaussian basis functions,” *J. Comput. Chem.* **36**, 1664–1671 (2015).
- ⁵⁸M. R. Hermes, “MRH,” <https://github.com/MatthewRHermes/mrh> (2024).
- ⁵⁹“PySCF-FORGE,” <https://github.com/pyscf/pyscf-forge> (2024).
- ⁶⁰L.-P. Wang and C. Song, “Geometry optimization made simple with translation and rotation coordinates,” *J. Chem. Phys.* **144**, 214108 (2016).
- ⁶¹J. J. Bao, M. R. Hermes, T. R. Scott, A. M. Sand, R. Lindh, L. Gagliardi, and D. G. Truhlar, “Analytic gradients for compressed multistate pair-density functional theory,” *Mol. Phys.* **120**, e2110534 (2022).
- ⁶²G. Glockler and D. L. Fuller, “Helium hydride ion,” *J. Chem. Phys.* **1**, 886–887 (1933).
- ⁶³D. M. Bishop and L. M. Cheung, “A theoretical investigation of HeH^+ ,” *J. Mol. Spectrosc.* **75**, 462–473 (1979).
- ⁶⁴S. Peyerimhoff, “Hartree–Fock–Roothaan wavefunctions, potential curves, and charge-density contours for the HeH^+ ($X^1\Sigma^+$) and NeH^+ ($X^1\Sigma^+$) molecule ions,” *J. Chem. Phys.* **43**, 998–1010 (2004).

- ⁶⁵R. Güsten, H. Wiesemeyer, D. Neufeld, K. M. Menten, U. U. Graf, K. Jacobs, B. Klein, O. Ricken, C. Risacher, and J. Stutzki, "Astrophysical detection of the helium hydride ion HeH⁺," *Nature* **568**, 357–359 (2019).
- ⁶⁶O. Novotný, P. Wilhelm, D. Paul, Á. Kálosi, S. Saurabh, A. Becker, K. Blaum, S. George, J. Göck, M. Grieser, F. Grussie, R. von Hahn, C. Krantz, H. Kreckel, C. Meyer, P. M. Mishra, D. Muell, F. Nuesslein, D. A. Orlov, M. Rimmler, V. C. Schmidt, A. Shornikov, A. S. Terekhov, S. Vogel, D. Zajfman, and A. Wolf, "Quantum-state-selective electron recombination studies suggest enhanced abundance of primordial HeH⁺," *Science* **365**, 676–679 (2019).
- ⁶⁷R. J. Fallon, J. T. Vanderslice, and E. A. Mason, "Potential energy curves for lithium hydride," *J. Chem. Phys.* **32**, 1453–1455 (1960).
- ⁶⁸K. C. Li and W. C. Stwalley, "The A¹Σ⁺ → X¹Σ⁺ bands of the isotopic lithium hydrides," *J. Mol. Spectrosc.* **69**, 294–318 (1978).
- ⁶⁹A. Pardo, J. J. Camacho, and J. M. L. Poyato, "The padé-approximant method and its applications in the construction of potential-energy curves for the lithium hydride molecule," *Chem. Phys. Lett.* **131**, 490–495 (1986).
- ⁷⁰W. C. Stwalley and W. T. Zemke, "Spectroscopy and structure of the lithium hydride diatomic molecules and ions," *J. Phys. Chem. Ref. Data* **22**, 87–112 (1993).
- ⁷¹W.-C. Tung, M. Pavanello, and L. Adamowicz, "Very accurate potential energy curve of the LiH molecule," *J. Chem. Phys.* **134**, 064117 (2011).
- ⁷²J. Duncan, "The ground-state average and equilibrium structures of formaldehyde and ethylene," *Mol. Phys.* **28**, 1177–1191 (1974).
- ⁷³P. Jensen and P. R. Bunker, "The geometry and the inversion potential function of formaldehyde in the \tilde{A}^1A_2 and \tilde{a}^3A_2 electronic states," *J. Mol. Spectrosc.* **94**, 114–125 (1982).
- ⁷⁴Š. Budzák, G. Scalmani, and D. Jacquemin, "Accurate excited-state geometries: A CASPT2 and coupled-cluster reference database for small molecules," *J. Chem. Theory Comput.* **13**, 6237–6252 (2017).
- ⁷⁵J. J. Bao and D. G. Truhlar, "Automatic active space selection for calculating electronic excitation energies based on high-spin unrestricted hartree-fock orbitals," *J. Chem. Theory Comput.* **15**, 5308–5318 (2019).
- ⁷⁶A. Dreuw and M. Wormit, "The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states," *WIREs Comput. Mol. Sci.* **5**, 82–95 (2015).
- ⁷⁷O. Christiansen, H. Koch, and P. Jørgensen, "The second-order approximate coupled cluster singles and doubles model CC2," *Chemical Physics Letters* **243**, 409–418 (1995).
- ⁷⁸O. Christiansen, H. Koch, and P. Jørgensen, "Response functions in the CC3 iterative triple excitation model," *J. Chem. Phys.* **103**, 7429–7441 (1995).
- ⁷⁹H. Koch, O. Christiansen, P. Jørgensen, A. M. Sanchez de Merás, and T. Helgaker, "The CC3 model: An iterative coupled cluster approach including connected triples," *J. Chem. Phys.* **106**, 1808–1818 (1997).
- ⁸⁰O. Christiansen, H. Koch, and P. Jørgensen, "Perturbative triple excitation corrections to coupled cluster singles and doubles excitation energies," *J. Chem. Phys.* **105**, 1451–1459 (1996).
- ⁸¹P.-F. Loos, A. Scemama, A. Blondel, Y. Garniron, M. Caffarel, and D. Jacquemin, "A mountaineering strategy to excited states: Highly accurate reference energies and benchmarks," *J. Chem. Theory Comput.* **14**, 4360–4379 (2018).
- ⁸²K. N. Walzl, C. F. Koerting, and A. Kuppermann, "Electron-impact spectroscopy of acetaldehyde," *J. Chem. Phys.* **87**, 3796–3803 (1987).
- ⁸³W. Haugen, M. Trætteberg, F. Kaufmann, K. Motzfeldt, D. H. Williams, E. Bunnenberg, C. Djerassi, and R. Records, "The molecular structure of 1,3-butadiene and 1,3,5-trans-hexatriene." *Acta Chem. Scand.* **20**, 1726–1728 (1966).
- ⁸⁴Y. Shu and D. G. Truhlar, "Doubly excited character or static correlation of the reference state in the controversial 2¹A_g state of trans-butadiene?" *J. Am. Chem. Soc.* **139**, 13770–13778 (2017).
- ⁸⁵M. A. Watson and G. K.-L. Chan, "Excited states of butadiene to chemical accuracy: Reconciling theory and experiment," *J. Chem. Theory Comput.* **8**, 4013–4018 (2012).
- ⁸⁶J. Finley, P.-Å. Malmqvist, B. O. Roos, and L. Serrano-Andrés, "The multi-state CASPT2 method," *Chem. Phys. Lett.* **288**, 299–306 (1998).
- ⁸⁷M. N. R. Ashfold, B. Cronin, A. L. Devine, R. N. Dixon, and M. G. D. Nix, "The role of Πσ* excited states in the photodissociation of heteroaromatic molecules," *Science* **312**, 1637–1640 (2006).
- ⁸⁸A. L. Devine, M. G. D. Nix, R. N. Dixon, and M. N. R. Ashfold, "Near-ultraviolet photodissociation of thiophenol," *J. Phys. Chem. A* **112**, 9563–9574 (2008).
- ⁸⁹M. N. R. Ashfold, A. L. Devine, R. N. Dixon, G. A. King, M. G. D. Nix, and T. A. A. Oliver, "Exploring nuclear motion through conical intersections in the UV photodissociation of phenols and thiophenol," *Proc. Natl. Acad. Sci.* **105**, 12701–12706 (2008).
- ⁹⁰J. S. Lim, H. Choi, I. S. Lim, S. B. Park, Y. S. Lee, and S. K. Kim, "Photodissociation dynamics of thiophenol-d₁: The nature of excited electronic states along the S-D bond dissociation coordinate," *J. Phys. Chem. A* **113**, 10410–10416 (2009).
- ⁹¹X. Xu, K. R. Yang, and D. G. Truhlar, "Diabatic molecular orbitals, potential energies, and potential energy surface couplings by the 4-fold way for photodissociation of phenol," *J. Chem. Theory Comput.* **9**, 3612–3625 (2013).
- ⁹²X. Zhu, C. L. Malbon, and D. R. Yarkony, "An improved quasi-diabatic representation of the 1, 2, 3¹A coupled adiabatic potential energy surfaces of phenol in the full 33 internal coordinates," *J. Chem. Phys.* **144**, 124312 (2016).
- ⁹³L. Zhang, D. G. Truhlar, and S. Sun, "Electronic spectrum and characterization of diabatic potential energy surfaces for thiophenol," *Phys. Chem. Chem. Phys.* **20**, 28144–28154 (2018).
- ⁹⁴L. Zhang, D. G. Truhlar, and S. Sun, "Full-dimensional three-state potential energy surfaces and state couplings for photodissociation of thiophenol," *J. Chem. Phys.* **151**, 154306 (2019).
- ⁹⁵N. W. Larsen, "Microwave spectra of the six mono-¹³C-substituted phenols and of some monodeuterated species of phenol. Complete substitution structure and absolute dipole moment," *J. Mol. Struct.* **51**, 175–190 (1979).
- ⁹⁶D. Spangenberg, P. Imhof, and K. Kleinermanns, "The S₁ state geometry of phenol determined by simultaneous Franck-Condon and rotational constants fits," *Phys. Chem. Chem. Phys.* **5**, 2505–2514 (2003).
- ⁹⁷G. Granucci, J. T. Hynes, P. Millié, and T.-H. Tran-Thi, "A theoretical investigation of excited-state acidity of phenol and cyanophenols," *J. Am. Chem. Soc.* **122**, 12243–12253 (2000).
- ⁹⁸R. N. Dixon, T. A. A. Oliver, and M. N. R. Ashfold, "Tunnelling under a conical intersection: Application to the product vibrational state distributions in the UV photodissociation of phenols," *J. Chem. Phys.* **134**, 194303 (2011).
- ⁹⁹G. A. Pino, A. N. Oldani, E. Marceca, M. Fujii, S.-I. Ishiuchi, M. Miyazaki, M. Broquier, C. Dedonder, and C. Jovet, "Excited state hydrogen transfer dynamics in substituted phenols and their complexes with ammonia: Ππ* -Πσ* energy gap propensity and ortho-substitution effect," *J. Chem. Phys.* **133**, 124313 (2010).
- ¹⁰⁰O. P. J. Vieuxmaire, Z. Lan, A. L. Sobolewski, and W. Domcke, "Ab initio characterization of the conical intersections involved in the photochemistry of phenol," *J. Chem. Phys.* **129**, 224307 (2008).
- ¹⁰¹G. Fogarasi, "Relative stabilities of three low-energy tautomers of cytosine: A coupled cluster electron correlation study," *J. Phys. Chem. A* **106**, 1381–1390 (2002).
- ¹⁰²D. L. Barker and R. E. Marsh, "The crystal structure of cytosine," *Acta Cryst* **17**, 1581–1587 (1964).

- ¹⁰³A. Nakayama, S. Yamazaki, and T. Taketsugu, “Quantum chemical investigations on the nonradiative deactivation pathways of cytosine derivatives,” *J. Phys. Chem. A* **118**, 9429–9437 (2014).
- ¹⁰⁴Chr. Møller and M. S. Plesset, “Note on an approximation treatment for many-electron systems,” *Phys. Rev.* **46**, 618–622 (1934).
- ¹⁰⁵M. Head-Gordon, J. A. Pople, and M. J. Frisch, “MP2 energy evaluation by direct methods,” *Chem. Phys. Lett.* **153**, 503–506 (1988).
- ¹⁰⁶P. J. Knowles and H.-J. Werner, “Internally contracted multiconfiguration-reference configuration interaction calculations for excited states,” *Theoret. Chim. Acta* **84**, 95–103 (1992).
- ¹⁰⁷M. Merchán, R. González-Luque, T. Climent, L. Serrano-Andrés, E. Rodríguez, M. Reguero, and D. Peláez, “Unified model for the ultrafast decay of pyrimidine nucleobases,” *J. Phys. Chem. B* **110**, 26471–26476 (2006).
- ¹⁰⁸J. González-Vázquez and L. González, “A time-dependent picture of the ultrafast deactivation of keto-cytosine including three-state conical intersections,” *ChemPhysChem* **11**, 3617–3624 (2010).
- ¹⁰⁹A. Nakayama, Y. Harabuchi, S. Yamazaki, and T. Taketsugu, “Photophysics of cytosine tautomers: New insights into the non-radiative decay mechanisms from MS-CASPT2 potential energy calculations and excited-state molecular dynamics simulations,” *Phys. Chem. Chem. Phys.* **15**, 12322–12339 (2013).
- ¹¹⁰T. D. Cherneva, M. M. Todorova, R. I. Bakalska, I. G. Shterev, E. Horkel, and V. B. Delchev, “Experimental and theoretical

study of the cytosine tautomerism through excited states,” *J. Mol. Model.* **29**, 303 (2023).

- ¹¹¹R. Abouaf, J. Pommier, H. Dunet, P. Quan, P.-C. Nam, and M. T. Nguyen, “The triplet state of cytosine and its derivatives: Electron impact and quantum chemical study,” *J. Chem. Phys.* **121**, 11668–11674 (2004).

TOC GRAPHIC

