

Construction of Quasi-diabatic Hamiltonians That Accurately Represent *ab Initio* Determined Adiabatic Electronic States Coupled by Conical Intersections for Systems on the Order of 15 Atoms. Application to Cyclopentoxide Photoelectron Detachment in the Full 39 Degrees of Freedom

Published as part of The Journal of Physical Chemistry virtual special issue "Machine Learning in Physical Chemistry".

Yifan Shen* and David R. Yarkony*



Cite This: *J. Phys. Chem. A* 2020, 124, 4539–4548



Read Online

ACCESS |



Metrics & More

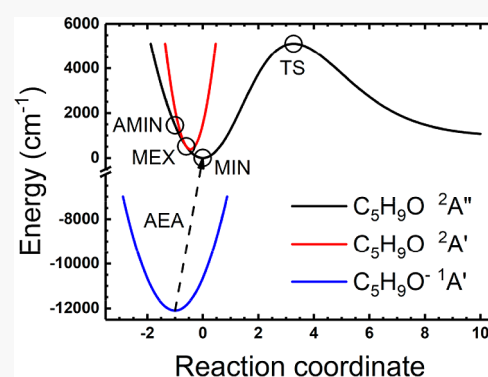


Article Recommendations



Supporting Information

ABSTRACT: We present, for systems of moderate dimension, a fitting framework to construct quasi-diabatic Hamiltonians that accurately represent *ab initio* adiabatic electronic structure data including the effects of conical intersections. The framework introduced here minimizes the difference between the fit prediction and the *ab initio* data obtained in the adiabatic representation, which is singular at a conical intersection seam. We define a general and flexible merit function to allow arbitrary representations and propose a representation to measure the fit–*ab initio* difference at geometries near electronic degeneracies. A fit Hamiltonian may behave poorly in insufficiently sampled regions, in which case a machine learning theory analysis of the fit representation suggests a regularization to address the deficiency. Our fitting framework including the regularization is used to construct the full 39-dimensional coupled diabatic potential energy surfaces for cyclopentoxide relevant to cyclopentoxide photoelectron detachment.



I. INTRODUCTION

When conical intersections^{1–4} are involved in a nonadiabatic process,^{5–8} a quasi-diabatic^{9,10} Hamiltonian (H^d) can greatly facilitate a reliable simulation. A variety of diabatizations have been reported in the literature, based on smooth molecular properties,^{11–16} ansatz of diabatic states,^{17–20} configuration uniformity,^{21–27} and derivative couplings.^{28–31} When accuracy is the primary concern, it would be safest for diabatizations to employ multireference electronic structure methods which provide derivative coupling information explicitly, such as multiconfigurational self-consistent field^{32–35} (MCSCF), extended multistate multireference second order perturbation theory^{36,37} (XMS-MRPT2), quasi-degenerate N -electron valence state second order perturbation theory³⁸ (QD-NEVPT2), and multireference configuration interaction^{39,40} (MRCI).

The focus of this work is a well-established derivative-coupling-based method which accurately fits energies, energy gradients, and the energy difference scaled derivative couplings obtained from *ab initio* electronic wave functions.^{41–49} This framework directly compares the fit prediction in the adiabatic representation to the corresponding *ab initio* quantities. Into this framework a partial diagonalization technique⁴⁴ has been

incorporated to yield a unique and continuous representation along a seam of conical intersections, where the degeneracy introduces ambiguity. However, when this approach is applied to adiabatic states that are only quasi-degenerate (energy gap up to 2000 cm^{-1} in ref 44 or 0.05 eV in ref 48), the *ab initio* electronic Hamiltonian contains nonzero off-diagonal elements in the partially diagonalized representation, which must be treated. In order to incorporate the off-diagonal contribution, here we introduce an alternative treatment by defining the merit function to be the difference between H^d in its eigenbasis and the *ab initio* Hamiltonian (H^{ab}) already in its eigenbasis, each transformed (by $U^{(n)}$) to a common but arbitrary representation. Our merit function is general and flexible, allowing any convenient representation to measure the fitting error, but it

Received: March 28, 2020

Revised: May 6, 2020

Published: May 6, 2020



reduces to a familiar one if the adiabatic representation is chosen ($U^{(1)} = 1$) or an accurate fit is obtained. Here we propose an alternative representation (transformation by $U^{(2)}$) for geometries with quasi-degenerate electronic states, which is similar to the partially diagonalized representation and easy to implement.

A fitted Hamiltonian may behave poorly in an insufficiently sampled region. The common practice is to generate an initial fit from chemically important geometries, then fill the holes (large negative values due to ill-defined polynomials) in the fitted potential energy surfaces with additional geometries. However, this solution may fail in a moderate-dimensional case because the volume of a space grows exponentially over its dimensionality, making it hopeless to saturate the entire space with finite number of geometries. If we reconsider the poor behavior in a more abstract way, e.g., from a machine learning point of view,⁵⁰ then it can be regarded as an overfitting problem, which can be treated with regularization. Compared to the hole filling technique, regularization is appealing for the following reasons:

- (1) Negligible extra fitting cost is introduced. The regularization term is a norm of the fitting parameters, while the hole filling technique requires fit-*ab initio* difference evaluation at every hole filling geometry, which can become prohibitively expensive in a moderate-dimensional case.
- (2) A limited number of additional geometries are required. Standard regularization techniques require no additional geometries at all; although additional geometries serving as a development set may be helpful to tune the regularization hyperparameter. This is advantageous, even for small molecules, when *ab initio* calculations are expensive, which is almost always the case in accuracy-demanding problems.

We will discuss this issue further in section II.B.

This paper is organized as follows. Section II.A elaborates the definition of our merit function. Section II.B proposes a representation to measure the fit-*ab initio* difference at geometries with quasi-degenerate adiabatic states and suggests a regularization for surface fitting. Our algorithm is summarized in section II.C. With an eye to a future spectral simulation, our fitting framework is applied to construct full 39-dimensional coupled potential energy surfaces for cyclopentoxide photoelectron detachment. The results are presented in section III. Section IV summarizes and discusses directions for future work.

II. THEORY

II.A. Merit Function. To determine a model quasi-diabatic Hamiltonian $H^d(Q; c)$ representing N^{state} electronic states by fitting a training set, a merit function is required to estimate the performance of $H^d(Q; c)$, where Q is the nuclear coordinate vector, c is the fitting parameter vector. For the approach of interest here the training set is based on energies, energy gradients and energy difference scaled derivative couplings obtained from *ab initio* determined adiabatic electronic wave functions. The existing fitting schemes based on this approach^{41–49} share a common adiabatic merit function $\epsilon_{old}(c)$, that is

$$\epsilon_{old}(c) = \sum_{i=1}^M w_i [\rho^2 \|E^d(Q_i; c) - E^{ab}(Q_i)\|_F^2 + \|\nabla H_a^d(Q_i; c) - \nabla H_a^{ab}(Q_i)\|_F^2] \quad (1)$$

where E is the N^{state} vector of energies, ρ is a scaling factor accounting for the difference in the units between E and ∇H , w is the fitting weight (there are different ways to pose w ; for simplicity here, we use one weight per geometry), M is the number of geometries contained in the training set, superscripts d and ab indicate coming from H^d or *ab initio*, subscript a is short for adiabatic, and F denotes the Frobenius norm. Here

$$H^x(q, Q)|i_a^x(q, Q)\rangle = E_i^x(Q)|i_a^x(q, Q)\rangle \quad (2a)$$

$$[\nabla H_a^x]_{ij}(Q) = \langle i_a^x(q, Q)|\nabla H(q, Q)|j_a^x(q, Q)\rangle_q \quad (2b)$$

where q denotes the electronic coordinates. The Frobenius norm squared of a $D_1 \times D_2 \times \dots \times D_n$ n th-order tensor is

$$\|A\|_F^2 = \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2} \dots \sum_{i_n=1}^{D_n} A_{i_1 i_2 \dots i_n}^2 \quad (3)$$

The adiabatic representation is singular along a conical intersection seam. Any unitary transformation of the degenerate states is still a legal adiabatic representation, and usually the rotation defining the orthogonal⁵¹ branching space⁵² vectors are chosen. However, nothing conceptually forbids using an alternative representation. Here we introduce an alternative merit function $\epsilon(c)$

$$\epsilon(c) = \sum_{i=1}^M w_i [\rho^2 \|H_r^d(Q_i; c) - H_r^{ab}(Q_i)\|_F^2 + \|\nabla H_r^d(Q_i; c) - \nabla H_r^{ab}(Q_i)\|_F^2] \quad (4)$$

where subscript r denotes the matrix representation of the operator. Comparing eq 4 to 1, H substitutes for E , and r replaces a . $\epsilon(c)$ reduces to $\epsilon_{old}(c)$ when $r = a$ choosing the adiabatic representation or an accurate fit is obtained. If H^d is accurate enough, then the fitting error would always be negligible no matter which representation it is defined in.

II.B. Composite Representation. In principle, any non-singular representation is compatible with $\epsilon(c)$. In practice, when a geometry (Q_i) is far away from a conical intersection seam, the adiabatic representation works fine, so we simply leave it; when a Q_i has quasi-degenerate energy levels, however, an alternative representation must be adopted. A desirable alternative should be able to be

- (1) Determined based solely on single point information. If not, then it is a nonlocal problem, which is as hard as our goal, constructing a global H^d , rather than a simple subproblem.
- (2) Uniquely transformed from quasi-diabatic (for H^d) or adiabatic (for H^{ab}) representation up to only trivial changes (e.g., the multiple solutions to eqs A7 or A10, see Appendix A). If not, then it is as ill-posed as the quasi-degenerate adiabatic representation.

Bearing these limitations in mind, we propose an alternative representation obtained by diagonalizing $\nabla H \cdot \nabla H$, which is defined as

$$[(\nabla H \cdot \nabla H)_r^x]_{ij} = \sum_{k=1}^{N^{state}} [\nabla H_r^x]_{ik} \cdot [\nabla H_r^x]_{kj} \quad (5a)$$

We will suppress the superscript x and subscript r when no confusion will result.

The $\nabla H \cdot \nabla H$ composite representation is desirable in that, first, it is a local or single point transformation and second, it is

unique except at degenerate points of $\nabla\mathbf{H}\cdot\nabla\mathbf{H}$. Since it is unlikely to have $\mathbf{H}$ and $\nabla\mathbf{H}\cdot\nabla\mathbf{H}$ both degenerate at a same geometry, this composite representation is almost always qualified as an alternative. Moreover, the composite representation is as valid as a previously reported partially diagonalized representation⁴⁴ and easy to implement. In the $N^{\text{state}} = 2$ case eq 5a leads to the requirement

$$\mathbf{s}_{ij}\cdot\mathbf{h}_{ij} = 0 \quad (\text{5b})$$

which is to be compared with the condition for the partially diagonalized representation

$$\mathbf{g}_{ij}\cdot\mathbf{h}_{ij} = 0 \quad (\text{5c})$$

Here $\mathbf{s}_{ij}$, $\mathbf{g}_{ij}$, and $\mathbf{h}_{ij}$ are the average energy, energy difference and interstate coupling gradients⁵¹ of adiabatic states i, j . A detailed comparison of the proposed representation to the partially diagonalized representation is presented in Appendix A.

To construct coupled potential energy surfaces of moderate dimension, the training set is highly limited: it is not possible to saturate an exponentially growing space with data at finite number of geometries. Consequently, a fitted $\mathbf{H}^d$ of moderate dimension may poorly reproduce $\mathbf{H}^{ab}$ in insufficiently sampled regions, due to incomplete knowledge of the performance of $\mathbf{H}^d$ outside the training set. This is the situation encountered in this work as explained in section III.A. The possible poor behavior of the fitted $\mathbf{H}^d$ is similar to overfitting (the inaccurate behavior) in machine learning.⁵⁰

- (1) Overfitting happens in insufficiently sampled regions and adding additional data points (geometries) helps.
- (2) The hypothesis class is rich (the fitting expansion is large) enough to fit the chemically important region accurately.

These similarities indicate the applicability of regularization, a well-established cure to overfitting in machine learning, to $\mathbf{H}^d$ construction. The use of regularization techniques to construct the surface fitting algorithm is described in Appendix B.

II.C. Algorithm. Our machine learning motivated algorithm to determine $\mathbf{H}^d(\mathbf{Q}; \mathbf{c})$, which is derived in Appendix B, minimizes a loss function

$$L(\mathbf{c}) = \varepsilon(\mathbf{c}) + \mu \|\mathbf{c} - \mathbf{c}_p\|_2^2 \quad (\text{6})$$

where $\varepsilon(\mathbf{c})$ has been defined in eq 4, and μ and $\mathbf{c}_p$ (p stands for prior) are introduced in eq B6. For $\mathbf{Q}_i$, if $E_{j+1}^{ab}(\mathbf{Q}_i) - E_j^{ab}(\mathbf{Q}_i) \geq \delta$ for all $j \in [1, N^{\text{state}}]$, then every operator is in the adiabatic representation; otherwise, the composite representation must be used.

The minimization of $L(\mathbf{c})$ is performed using the trust region⁵³ continued with the Dai–Yuan conjugate gradient.⁵⁴

III. RESULTS AND DISCUSSION

Alkoxy radicals are important participants in atmospheric and combustion processes, and photoelectron spectroscopy is a powerful experimental technique with which to study them. Here we apply our fitting framework to cyclopentoxo ($\text{C}_5\text{H}_9\text{O}$) to construct an $\mathbf{H}^d$, which will be employed to determine the photoelectron spectrum of cyclopentoxide ($\text{C}_5\text{H}_9\text{O}^-$) in future work. The experiment⁵⁵ has been done by Alconcel and Continetti. Figure 1 pictures the $\text{C}_5\text{H}_9\text{O}$ radical and indicates the atom labeling used in this section.

In the Condon approximation, the photodetachment of an electron from the negative ion $\text{C}_5\text{H}_9\text{O}^-$ (the precursor) produces neutral $\text{C}_5\text{H}_9\text{O}$ (the target or residual) in a vibronic

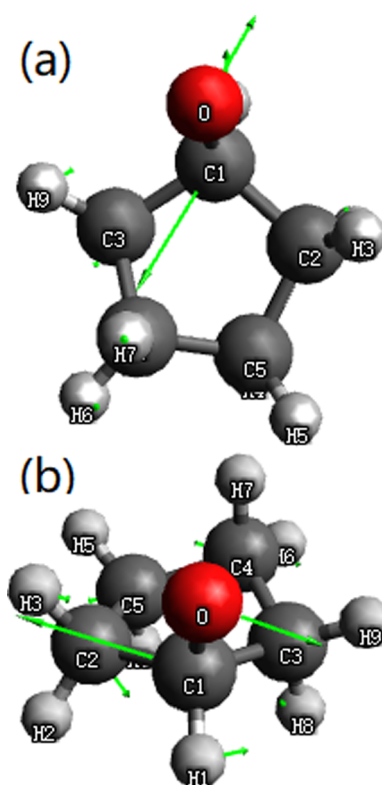


Figure 1. Cyclopentoxo $1,2^2\text{A}$ minimum energy crossing structure, along with (a) normalized $\mathbf{g}$, the energy difference gradient vector, or (b) normalized $\mathbf{h}$, the interstate coupling vector. The principle component of $\mathbf{g}$ is the scissoring of O–C1–H1, while the rocking of O–C1–H1 dominates $\mathbf{h}$. If we borrow the language of the C_s point group, then $\mathbf{g}$ is almost A' , while $\mathbf{h}$ is almost A'' . The atom numbering used in the text is also indicated.

state. For a low temperature experiment, the precursor resides in its ground vibronic state, which can be adequately described by a quadratic approximation obtained from an *ab initio* nuclear Hessian. Therefore, only for the neutral residual is an $\mathbf{H}^d$, comprised of 2 electronic states and 39 nuclear internal coordinates, required.

The two-state description is necessitated by a low-lying conical intersection which produces a Jahn–Teller type interaction that couples the lowest two potential energy surfaces of $\text{C}_5\text{H}_9\text{O}$. The precursor equilibrium geometry (anion minimum, AMIN) has a higher symmetry (C_s point group) than the low energy region (C_1 point group). The ground state minimum (MIN) and the Jahn–Teller saddle point (SP) are largely contained within the branching space of the $1,2^2\text{A}$ minimum energy crossing (MEX). A 1-dimensional sketch of the involved potential energy surfaces is provided in Figure 2.

Section III.A outlines the construction of $\mathbf{H}^d$. Section III.B assesses the quality of $\mathbf{H}^d$.

III.A. Construction of $\mathbf{H}^d$. The electronic structure description adopted here for $\text{C}_5\text{H}_9\text{O}$ is a previously reported⁵⁶ MRCI with single and double excitations (MRCISD), with 6 frozen core orbitals, 15 doubly occupied orbitals, and a 5-electron 3-orbital active space. The atomic orbital basis set is cc-pVTZ⁵⁷ for the oxygen and the carbons, and cc-pVDZ⁵⁸ for the hydrogens. This MRCISD is comprised of 19 302 445 configuration state functions. All electronic structure computations are performed using the COLUMBUS^{39,40,59–65} suite of programs. The zero of potential energy is taken as that of MIN.

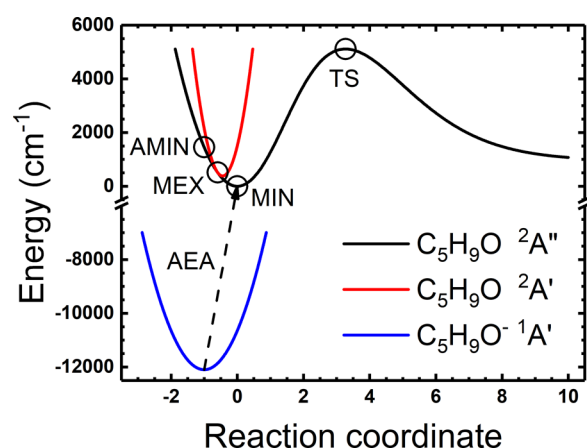


Figure 2. Sketch of the potential energy surfaces involved in cyclopentoxide ($\text{C}_5\text{H}_9\text{O}^-$) photoelectron detachment. The crude adiabatic states are labeled with the irreducible representations of C_s point group. $\text{C}_5\text{H}_9\text{O}^-$ ground state minimum (anion minimum, AMIN, 1452 cm^{-1}), cyclopentoxy ($\text{C}_5\text{H}_9\text{O}$) ground state minimum (MIN, 0 cm^{-1}), $1,2^2\text{A}$ minimum energy crossing (MEX, 515 cm^{-1}), and ring-opening transition state (TS, 5056 cm^{-1}) are circled. Experimental adiabatic electron affinity (AEA, 1.5 eV) is also marked.

$\text{C}_5\text{H}_9\text{O}$ may experience a ring opening, which is unrelated to the low energy photoelectron spectrum reflecting a classical barrier of 5056 cm^{-1} and a dramatic geometry change from the precursor. However, the ring-opening channel complicates the determination of vibronic states. The vertical excitation of the precursor is only 1452 cm^{-1} , after accounting for adiabatic electron affinity (see Figure 2), which supports the validity of approximating the coupled $\text{C}_5\text{H}_9\text{O}$ potential energy surfaces as bound.

For bound systems, the electronic Hamiltonian matrix elements can be well described by polynomials based on internal coordinates. The polynomial origin is set to MEX. There is non-negligible anharmonicity along the linear synchronous transit path from MEX to AMIN, so we cannot stop at a quadratic approximation. Owing to the dimensionality, a simple quartic expansion over full 39 degrees of freedom (123410 terms) is unaffordable, so to the quadratic approximation we add a quartic expansion for only eight important coordinates listed in the left most column in Table 1, resulting in 820 terms through quadratic order + 450 third and fourth order terms = 1270 terms in total. Since there are three independent elements in a 2×2 real symmetric $\mathbf{H}^d$, the total number of the fitting parameters is 3810. The training set contains 356 geometries (equivalent to 42364 independent fitting equations),

of which 257 are within Jahn–Teller region as follows: MIN and 78 geometries around it, MEX and 78 geometries around it, SP and 78 geometries around it, 10 along the linear synchronous transit path from MEX to MIN, and 10 along the linear synchronous transit path from MEX to SP. Others are around the precursor equilibrium geometry: AMIN and 78 geometries around it and 20 along the linear synchronous transit path from MEX to AMIN.

This is the (limited) training set. As suggested above it is not possible to obtain an ample training set for the 39-dimensional $\text{C}_5\text{H}_9\text{O}$. Without regularization the fitting algorithm, minimizing the first term on the right-hand side of eq 6, produces potential energy surfaces with holes. Filling the holes in the fitted surfaces with additional points merely pushed the holes further from the chemically important region, but the spectral simulation remains compromised. This failure necessitated the introduction of regularization in the surface fitting, which is including the second term in eq 6. We constructed the prior $\mathbf{c}_p$ for the regularization as follows:

- (1) Generate a harmonic fit using the 257 Jahn–Teller region geometries.
- (2) Determine MRCISD wave functions on an extra 9 (30) geometries for each of the 2 (4) bond lengths (angles) described at fourth order, then generate 6 independent quartic 1-dimensional fits from these geometries.
- (3) Determine MRCISD wave functions on an extra 106 geometries for the 2 dihedral angles described at fourth order, then generate a quartic 2-dimensional fit from these data.
- (4) Fill the remaining elements of $\mathbf{c}_p$ with 0.

This $\mathbf{c}_p$ showed qualitatively correct behavior, further from the chemically important region. So not unexpectedly a regularization based on it yielded a well-behaved fitted $\mathbf{H}^d$. Another factor involved in the regularization is the hyperparameter μ . Our choice is $\mu = 0.1$ considering the balance between accuracy in and out of the training set.

Before executing our fitting algorithm, the final missing parts of eq 6, w and ρ must be defined. Various choices are allowed. Here the construction used

$$w_i = \begin{cases} 1, & \text{if } E_1^{ab}(\mathbf{Q}_i) \leq E_{\text{thresh}} \\ \left(\frac{E_{\text{thresh}} - E_{\text{ref}}}{E_1^{ab}(\mathbf{Q}_i) - E_{\text{ref}}} \right)^2, & \text{else} \end{cases} \quad (7a)$$

Table 1. Representative Bond Lengths in Å, Bond Angles and Dihedral Angles in Degree, for Cyclopentoxy Ground State Minimum (MIN), $1,2^2\text{A}$ Minimum Energy Crossing (MEX), and Jahn–Teller Saddle Point (SP), Computed from MRCISD (*ab*) and the Quasi-diabatic Hamiltonian (*d*)

	MEX ^{ab}	MEX ^d	MIN ^{ab}	MIN ^d	SP ^{ab}	SP ^d
C1–O	1.3827	1.3829	1.3795	1.3796	1.3803	1.3803
C1–H1	1.0914	1.0913	1.0898	1.0898	1.0918	1.0918
O–C1–C2	112.2	112.2	105.4	105.4	113.6	113.5
O–C1–C3	109.8	109.9	111.1	111.0	109.6	109.6
H1–C1–C2	112.6	112.6	111.8	111.9	112.9	112.9
H1–C1–C3	111.6	111.4	113.2	113.2	110.7	110.7
C5–C4–C3–C1	−43.0	−42.7	−43.0	−43.1	−41.8	−41.9
C3–C4–C5–C2	37.6	38.7	37.6	36.9	39.9	39.7

$$\rho = \frac{\max\{\|\nabla H_a^{ab}(\mathbf{Q}_i)\|_M\}}{\max\{\|H_a^{ab}(\mathbf{Q}_i)\|_M\}} \quad (7b)$$

where E_{ref} is chosen as the $1,2^2A$ minimum energy crossing (515 cm^{-1}), E_{thresh} is set to 0.003 hartree (660 cm^{-1}), $\|A\|_M$ is the largest absolute value among all entries of A , and $\max\{\cdot\}$ denotes the maximum among the training set. This ρ is simply a convenient choice. This w comes from a straightforward idea that, the low energy region (lower than E_{thresh}) is of the most interest, while the importance of high energy region fades. We set δ to 0.0001 hartree (22 cm^{-1}). Then the algorithm described in section II.C is used to construct H^d .

III.B. Quality of H^d . As described in section III.A there are 356 points in the training set of which 71 are considered quasi-degenerate. Of those points, 257 are included in the regularization, to which are added 244 points. The root-mean-square deviations (RMSD) for the nondegenerate data points are

$$\text{RMSD}(E) = 9.419843 \text{ cm}^{-1}$$

$$\text{RMSD}(\nabla H) = 3.643409 \times 10^{-4} \text{ au}$$

and for the quasi-degenerate (energy difference $< 22 \text{ cm}^{-1}$) data points we find

$$\text{RMSD}(H) = 4.497788 \text{ cm}^{-1}$$

$$\text{RMSD}(\nabla H) = 2.819638 \times 10^{-5} \text{ au}$$

These results are quite encouraging but a clearer picture emerges from the tables and figures which follow.

Table 1 reports the H^d and MRCISD determined key internal coordinates for MIN, MEX, and SP. The fit and *ab initio* determined structures are in excellent agreement, with the largest error in a bond distance (angle) [dihedral angle] being $0.0002 \text{ \AA}$ (0.2°) [1.1°]. Table 2 extends the comparison to harmonic frequencies for MIN and SP. Most normal modes are well reproduced to within a several wavenumbers.

Parts a and b of Figure 3 assess the quality of H^d in the vicinity of the MEX, reporting the energy and the derivative coupling along paths in the $\mathbf{g}$ and $\mathbf{h}$ directions, with the MEX as the origin. As shown in Figure 1, the principle component of $\mathbf{g}$ is the scissoring of O–C1–H1, while the rocking of O–C1–H1 dominates $\mathbf{h}$. H^d reproduces the MRCISD results quantitatively. Parts c and d of Figure 3 report the $\mathbf{g}$ – $\mathbf{h}$ plane cross sections of the ground state and the pair of states, respectively. These plots are generated purely from H^d as they would be too expensive for MRCISD. From Figure 3c, we can see that MIN and SP are largely contained within the branching space of the MEX.

To understand the behavior of the coupled potential energy surfaces perpendicular to the branching plane of MEX we stretch the C–O bond, which is one of the key internal coordinates and whose direction is almost perpendicular to $\mathbf{g}$ and $\mathbf{h}$. Figure 4a presents the energetically accessible region (recall from Figure 2 that AMIN is only 1452 cm^{-1} higher than MIN), within which 1^2A and 2^2A barely separate. If we zoom in around MEX as shown in Figure 4b, we find that the $1,2^2A$ degeneracy breaks slowly but linearly, as a result of the small but nonzero projection of the C–O bond stretching in the $\mathbf{g}$ – $\mathbf{h}$ plane. It is worth noting that H^d agrees with MRCISD quantitatively in both parts a and b of Figure 4.

Potential energy surfaces can be rather flat along torsional coordinates, so the resulting fits can easily produce holes outside the training set. Although only the chemically important regions

Table 2. Harmonic Frequencies in cm^{-1} for Cyclopentoxo Ground State Minimum (MIN) and Jahn–Teller Saddle Point (SP), Computed from MRCISD (*ab*) and Quasi-adiabatic Hamiltonian (*d*)

mode	MIN ^{ab}	MIN ^d	SP ^{ab}	SP ^d
1	60.73	54.02	−1369.43	−1427.25
2	251.98	238.97	30.38	56.86
3	374.91	357.39	260.96	252.26
4	408.05	416.06	446.43	484.22
5	583.42	583.21	527.48	530.80
6	703.91	701.64	589.55	589.95
7	843.06	834.73	775.07	779.92
8	885.08	888.70	853.04	845.15
9	906.83	909.44	872.08	874.49
10	939.63	939.26	913.46	908.65
11	961.96	961.52	958.01	955.36
12	979.65	966.07	962.13	962.26
13	1052.88	1061.19	1008.48	1006.25
14	1075.83	1079.33	1061.35	1068.27
15	1097.25	1090.95	1083.83	1084.62
16	1120.44	1135.01	1108.56	1107.32
17	1241.42	1248.99	1165.83	1174.62
18	1261.25	1264.30	1251.03	1254.35
19	1301.38	1296.51	1270.73	1275.41
20	1310.86	1307.05	1304.14	1298.46
21	1344.95	1345.99	1313.55	1309.49
22	1376.05	1368.86	1372.88	1373.87
23	1402.90	1400.57	1400.18	1401.48
24	1413.23	1408.98	1412.05	1419.17
25	1423.55	1423.71	1418.88	1421.50
26	1453.64	1464.22	1430.43	1438.85
27	1539.77	1538.93	1540.34	1539.84
28	1542.22	1540.66	1542.46	1542.53
29	1551.95	1551.68	1553.00	1553.17
30	1574.06	1573.44	1573.78	1573.62
31	3118.15	3119.53	3109.16	3115.57
32	3134.24	3133.59	3126.38	3130.56
33	3134.99	3139.82	3130.50	3131.98
34	3144.70	3141.96	3141.34	3138.41
35	3158.14	3157.48	3153.70	3154.81
36	3188.75	3189.26	3186.39	3186.60
37	3191.68	3194.80	3191.99	3192.69
38	3199.76	3203.95	3204.42	3203.21
39	3212.06	3210.20	3212.67	3208.30

have to be fitted quantitatively, holes may appear due to inexact polynomial cancellation. To validate the globally bound behavior of H^d , we twist the dihedral angle C4–C5–C2–C1. Figure 5a reports the coupled potential energy surfaces $\pm 60^\circ$ from the $1,2^2A$ MEX, which includes points a dramatically long way from the training set. H^d always behaves reasonably. If we zoom in around the MEX as shown in Figure 5b, we find that H^d is best between -30° and 0° , a much broader range than the training set which extends, $\pm 5^\circ$, from the $1,2^2A$ MEX.

IV. CONCLUSIONS

We have presented a fitting framework to construct quasi-adiabatic Hamiltonians which accurately represent *ab initio* adiabatic electronic structure data and include the effects of conical intersections for systems of moderate dimension. This framework consists of a general and flexible merit function along with an alternative representation to avoid problematic quasi

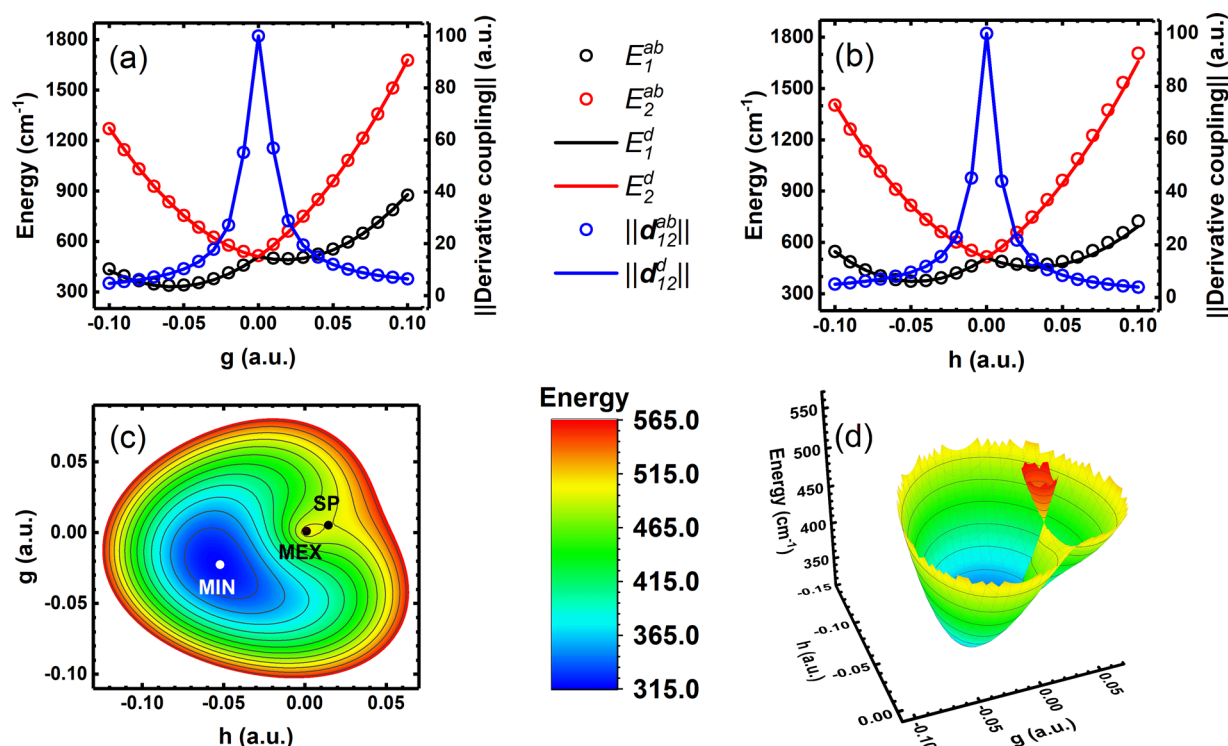


Figure 3. From cyclopentoxyl 1,2 ²A minimum energy crossing (MEX), scan by MRCISD (*ab*) and quasi-diabatic Hamiltonian (*d*) along (a) normalized *g*, the energy difference gradient vector; (b) normalized *h*, the interstate coupling vector. The infinite derivative coupling is arbitrarily plotted as 100. Within the branching plane spanned by *g* and *h*, the fitted (c) lower state surface, with MEX and the approximate locations of the ground state minimum (MIN) and the Jahn–Teller saddle point (SP) marked; (d) coupled surfaces illustrating double cone topography.

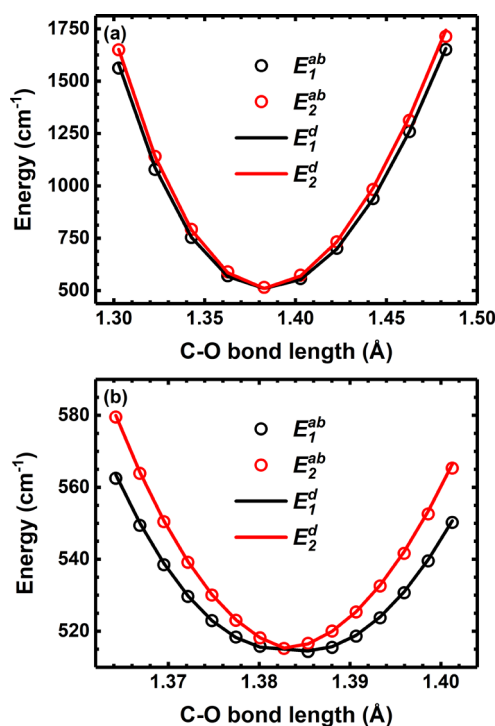


Figure 4. From cyclopentoxyl 1,2 ²A minimum energy crossing (MEX), scan by MRCISD (*ab*) and quasi-diabatic Hamiltonian (*d*) along the C–O bond stretching. Note that the C–O bond length at MEX is 1.3827 Å. Part b is the zoomed in version of part a around MEX.

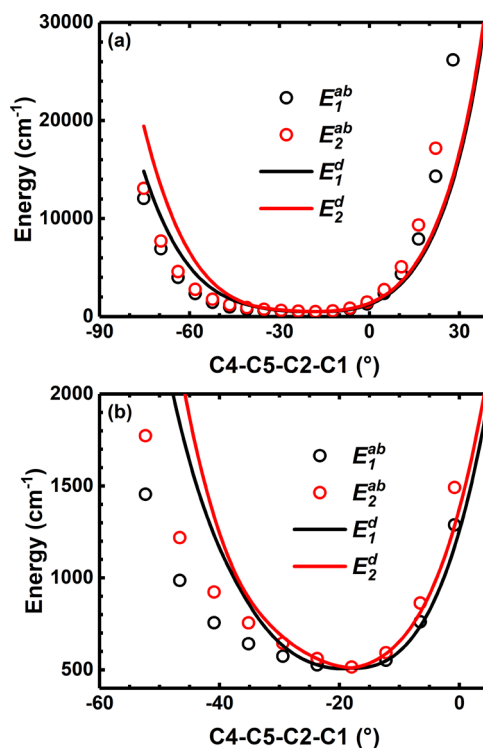


Figure 5. From cyclopentoxyl 1,2 ²A minimum energy crossing (MEX), scan by MRCISD (*ab*) and quasi-diabatic Hamiltonian (*d*) along the C4–C5–C2–C1 torsion angle. Note that the C4–C5–C2–C1 dihedral angle at MEX is –18.0°. (b) Zoomed in version of part a around MEX.

degeneracy, and a regularization technique to prevent pathological behavior of the fitted Hamiltonian. When the

regularization term is included, this fitting framework successfully constructs coupled 1,2²A potential energy surfaces for cyclopentoxo, of sufficient accuracy to describe cyclopentoxide photoelectron detachment. The photoelectron spectrum of cyclopentoxide will be determined employing the constructed H^d in future work.

■ APPENDIX A

Derivation of Equations 5b and 5c: The Composite Representation and Comparison to the Partial Diagonalization Approach

As noted in the main text, partial diagonalization is a unique unitary transformation of the eigenstate representation of electronic Hamiltonian H that avoids fitting the energy difference scaled derivative coupling data near a conical intersection. In this appendix we compare this approach with the composite representation reported in section II.B.

Let $U = \exp(\Theta)$, where Θ is a skew symmetric matrix, be the N^{state} transformation that transforms H , that is

$$H\psi_i = E_i\psi_i \quad (\text{A1})$$

We focus on the 2-state case where

$$\begin{bmatrix} \psi_i^z \\ \psi_j^z \end{bmatrix} = \begin{bmatrix} \cos \theta_{ij}^z & -\sin \theta_{ij}^z \\ \sin \theta_{ij}^z & \cos \theta_{ij}^z \end{bmatrix} \begin{bmatrix} \psi_i \\ \psi_j \end{bmatrix} \quad (\text{A2})$$

where superscript z can be pd or cr , denoting the partially diagonalized representation or the composite representation. The transformed gradients are

$$\begin{aligned} \nabla H^z &= \exp(-\Theta) \begin{bmatrix} \psi_i \\ \psi_j \end{bmatrix} \nabla H[\psi_i, \psi_j] \exp(\Theta) \\ &= \exp(-\Theta) \begin{bmatrix} \nabla H_{ii} & \nabla H_{ij} \\ \nabla H_{ji} & \nabla H_{jj} \end{bmatrix} \exp(\Theta) \end{aligned} \quad (\text{A3})$$

Define

$$\mathbf{s}_{ij} = \frac{1}{2}(\nabla H_{ii} + \nabla H_{jj}) \quad (\text{A4a})$$

$$\mathbf{g}_{ij} = \frac{1}{2}(\nabla H_{jj} - \nabla H_{ii}) \quad (\text{A4b})$$

$$\mathbf{h}_{ij} = \nabla H_{ij} \quad (\text{A4c})$$

Using eq A2 in eq A4 gives

$$\mathbf{s}_{ij}^z = \mathbf{s}_{ij} \quad (\text{A5a})$$

$$\mathbf{g}_{ij}^z = \cos 2\theta_{ij}^z \mathbf{g}_{ij} + \sin 2\theta_{ij}^z \mathbf{h}_{ij} \quad (\text{A5b})$$

$$\mathbf{h}_{ij}^z = \cos 2\theta_{ij}^z \mathbf{h}_{ij} - \sin 2\theta_{ij}^z \mathbf{g}_{ij} \quad (\text{A5c})$$

A well-established choice to determine θ_{ij}^{pd} is to require⁵¹

$$\mathbf{g}_{ij}^{pd} \cdot \mathbf{h}_{ij}^{pd} = 0 \quad (\text{A6})$$

The partial diagonalization is the extension of eq A6 to an arbitrary number of states, letting all state pairs satisfy $\mathbf{g}_{ij}^{pd} \perp \mathbf{h}_{ij}^{pd}$ provided the energy separation is less than a preassigned threshold. eq A6 leads to

$$(\|\mathbf{h}_{ij}\|_2^2 - \|\mathbf{g}_{ij}\|_2^2) \sin 4\theta_{ij}^{pd} = 2\mathbf{g}_{ij} \cdot \mathbf{h}_{ij} \cos 4\theta_{ij}^{pd} \quad (\text{A7})$$

which would always yield unique θ_{ij}^{pd} (up to a several $\pi/4$, in total 8 different transforms), unless $\|\mathbf{h}_{ij}\|_2^2 - \|\mathbf{g}_{ij}\|_2^2 = \mathbf{g}_{ij} \cdot \mathbf{h}_{ij} = 0$.

Alternatively using eq 5

$$\nabla \mathbf{H} \cdot \nabla \mathbf{H} = \begin{bmatrix} \|\nabla H_{ii}\|_2^2 + \|\nabla H_{jj}\|_2^2 & (\nabla H_{ii} + \nabla H_{jj}) \cdot \nabla H_{ij} \\ (\nabla H_{ii} + \nabla H_{jj}) \cdot \nabla H_{ij} & \|\nabla H_{jj}\|_2^2 + \|\nabla H_{ii}\|_2^2 \end{bmatrix} \quad (\text{A8})$$

and noting that

$$(\nabla H_{ii} + \nabla H_{jj}) \cdot \nabla H_{ij} = 2\mathbf{s}_{ij} \cdot \mathbf{h}_{ij} \quad (\text{A9})$$

We may choose a θ_{ij}^{cr} so that $(\nabla \mathbf{H} \cdot \nabla \mathbf{H})^{cr}$ is diagonal, i.e. $\mathbf{s}_{ij}^{cr} \perp \mathbf{h}_{ij}^{cr}$, that is

$$\mathbf{s}_{ij} \cdot \mathbf{h}_{ij} \cos 2\theta_{ij}^{cr} = \mathbf{s}_{ij} \cdot \mathbf{g}_{ij} \sin 2\theta_{ij}^{cr} \quad (\text{A10})$$

Like eq A7, eq A10 would always yield unique θ_{ij}^{cr} (up to a several $\pi/2$, so in total four different transforms), unless $\mathbf{s}_{ij} \cdot \mathbf{h}_{ij} = \mathbf{s}_{ij} \cdot \mathbf{g}_{ij} = 0$. From the properties of eqs A7 and A10, we can tell that composite representation is as valid as the partial diagonalization in 2-state case. The more general case of an arbitrary number of states^{66–70} will be the subject of a future work.

It is worthwhile to note that the composite representation is very easy to implement. As a Hermitian matrix, $\nabla \mathbf{H} \cdot \nabla \mathbf{H}$ can be diagonalized using standard routines. The derivative of the eigenbasis of $\nabla \mathbf{H} \cdot \nabla \mathbf{H}$ with respect to a parameter c (e.g. the fitting parameter) can also be conveniently evaluated in a standard way. Let

$$(\nabla \mathbf{H} \cdot \nabla \mathbf{H})\varphi_i = \lambda_i \varphi_i \quad (\text{A11})$$

The derivative of φ over some parameter c is

$$\left\langle \varphi_i \left| \frac{\partial}{\partial c} \right| \varphi_j \right\rangle = \frac{\left\langle \varphi_i \left| \frac{\partial}{\partial c} (\nabla \mathbf{H} \cdot \nabla \mathbf{H}) \right| \varphi_j \right\rangle}{\lambda_j - \lambda_i} \quad (\text{A12})$$

■ APPENDIX B

Derivation of Equation 6: The Regularization

To select an appropriate regularization for chemical use,⁷¹ we start from the origin of regularization: Bayesian statistics. From a Bayesian perspective,⁵⁰ fitting is estimating a probabilistic model from observed data by maximizing $P(\mathbf{c})$, the posterior probability of the fitting parameter vector $\mathbf{c}$

$$\begin{aligned} P(\mathbf{c}) &= \prod_{i=1}^M p(\mathbf{c} | \mathbf{H}_r^{ab}(\mathbf{Q}_i), \nabla \mathbf{H}_r^{ab}(\mathbf{Q}_i), \mathbf{Q}_i) \\ &= \prod_{i=1}^M p(\mathbf{H}_r^{ab}(\mathbf{Q}_i), \nabla \mathbf{H}_r^{ab}(\mathbf{Q}_i) | \mathbf{Q}_i, \mathbf{c}) p(\mathbf{c}) \end{aligned} \quad (\text{B1})$$

where $p(ab)$ is the conditional probability of a given b , $p(\mathbf{c})$ is the prior probability of $\mathbf{c}$. It will be more convenient to use log form

$$\begin{aligned} \underset{\mathbf{c}}{\operatorname{argmax}} P(\mathbf{c}) &= \underset{\mathbf{c}}{\operatorname{argmax}} \left[\sum_{i=1}^M \log p(\mathbf{H}_r^{ab}(\mathbf{Q}_i), \nabla \mathbf{H}_r^{ab}(\mathbf{Q}_i) | \mathbf{Q}_i, \mathbf{c}) \right. \\ &\quad \left. + M \log p(\mathbf{c}) \right] \end{aligned} \quad (\text{B2})$$

where the final term will provide the regularization, provided we have some prior knowledge from $p(\mathbf{c})$. If we have no prior

knowledge concerning $\mathbf{c}$, then we are equivalently applying uniform $p(\mathbf{c})$ to eq B2

$$\operatorname{argmax}_{\mathbf{c}} P(\mathbf{c}) = \operatorname{argmax}_{\mathbf{c}} \sum_{i=1}^M \log p(\mathbf{H}_r^{ab}(\mathbf{Q}_i), \nabla \mathbf{H}_r^{ab}(\mathbf{Q}_i) | \mathbf{Q}_i, \mathbf{c}) \quad (\text{B3})$$

In order to reduce Bayesian parameter estimation to familiar least square fit, a Gaussian model is adopted for $\mathbf{H}_r^{ab}(\mathbf{Q}_i)$ and $\nabla \mathbf{H}_r^{ab}(\mathbf{Q}_i)$

$$p(\mathbf{H}_r^{ab}(\mathbf{Q}_i), \nabla \mathbf{H}_r^{ab}(\mathbf{Q}_i) | \mathbf{Q}_i, \mathbf{c}) = a e^{-\sigma w [\rho^2 \|\mathbf{H}_r^d(\mathbf{Q}_i; \mathbf{c}) - \mathbf{H}_r^{ab}(\mathbf{Q}_i)\|_F^2 + \|\nabla \mathbf{H}_r^d(\mathbf{Q}_i; \mathbf{c}) - \nabla \mathbf{H}_r^{ab}(\mathbf{Q}_i)\|_F^2]} \quad (\text{B4})$$

where a is a renormalization factor, σ is a width constant, then we arrive at

$$\operatorname{argmax}_{\mathbf{c}} P(\mathbf{c}) = \operatorname{argmin}_{\mathbf{c}} \varepsilon(\mathbf{c}) \quad (\text{B5})$$

where $\varepsilon(\mathbf{c})$ has been defined in eq 4.

In common cases, we do have some understanding of $\mathbf{c}$. As shown in section III, we can then come up with a prior estimation $\mathbf{c}_p$ via various approaches, e.g.

- (1) Use a local harmonic approximation.
- (2) Fit only a few selected degrees of freedoms.
- (3) Simply let $\mathbf{c}_p = 0$, as what is usually done in L_1 and L_2 regularization.⁷²

It is not new to utilize chemical experience in a prior probability manner. The information theory^{73–75} of Levine and Bernstein provides a well-known example. To make use of $\mathbf{c}_p$ in surface fitting, a nonuniform $p(\mathbf{c})$ should be placed in eq B2. The common forms of $p(\mathbf{c})$ are Laplace distribution, Gaussian distribution, etc. Here we have used Gaussian forms owing to the simple quadratic regularization term it will produce

$$p(\mathbf{c}) = A e^{-\mu \sigma \|\mathbf{c} - \mathbf{c}_p\|_2^2} \quad (\text{B6})$$

where A is a renormalization factor, μ reflects our confidence in $\mathbf{c}_p$ over the observed training set. Plugging eqs B4 and (B6) into (B2), we obtain

$$\begin{aligned} \operatorname{argmax}_{\mathbf{c}} P(\mathbf{c}) &= \operatorname{argmin}_{\mathbf{c}} [\varepsilon(\mathbf{c}) + \mu \|\mathbf{c} - \mathbf{c}_p\|_2^2] \\ &\equiv \operatorname{argmin}_{\mathbf{c}} L(\mathbf{c}) \end{aligned} \quad (\text{B7})$$

where $L(\mathbf{c}) \equiv \varepsilon(\mathbf{c}) + \mu \|\mathbf{c} - \mathbf{c}_p\|_2^2$. As a hyperparameter adjusting the strength of regularization, we suggest tuning μ on a development set which

- (1) Does not overlap with the training set.
- (2) Consists of geometries whose ground state energies are as high as the total available energy in dynamics.
- (3) Has a reasonable size. Although it needs not to be large, at least two geometries along positive and negative directions of each coordinate should be ensured.

It is worthwhile to note the relationship between regularization and constraint. Comparing to constrained optimization, minimizing $L(\mathbf{c})$ allows a slight fluctuation⁵³ of $\mathbf{c}$ around $\mathbf{c}_p$

$$\operatorname{argmin}_{\mathbf{c}} L(\mathbf{c}) - \mathbf{c}_p \approx -\frac{\lambda}{\mu} \quad (\text{B8})$$

where λ is the Lagrange multiplier to enforce $\mathbf{c} = \mathbf{c}_p$. $\mathbf{c}_p$ reflects chemical experience, so on one hand $\mathbf{H}^d(\mathbf{Q}; \mathbf{c}_p)$ ought to be qualitatively correct, which means $\mathbf{c}$ should not deviate from $\mathbf{c}_p$ too much, yet on the other hand, $\mathbf{H}^d(\mathbf{Q}; \mathbf{c}_p)$ would not necessarily reproduce $\mathbf{H}^{ab}(\mathbf{Q})$ quantitatively; i.e., it would be inappropriate to enforce $\mathbf{c} = \mathbf{c}_p$. In conclusion, it is more reasonable to utilize chemical experience ($\mathbf{c}_p$) in a regularization manner, rather than an exact constraint to be enforced.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpca.0c02763>.

Documentation of the program SurGenBound and technical details of the cyclopentoxo diabatic Hamiltonian (PDF)

SurGenBound file with coupled diabatic potential energy surfaces constructed in this work, H^d , with details of the fit, the computer program, and sample inputs (ZIP)

■ AUTHOR INFORMATION

Corresponding Authors

Yifan Shen – Department of Chemistry, Johns Hopkins University, Baltimore, Maryland 21218, United States;
Email: yifanshensz@jhu.edu

David R. Yarkony – Department of Chemistry, Johns Hopkins University, Baltimore, Maryland 21218, United States;

orcid.org/0000-0002-5446-1350; Email: yarkony@jhu.edu

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acs.jpca.0c02763>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by National Science Foundation Grant CHE-1663692 to D.R.Y. The authors are grateful for a grant of computer time on the Maryland Advanced Research Computing Center (MARCC) at Johns Hopkins University.

■ REFERENCES

- (1) Celani, P.; Bernardi, F.; Olivucci, M.; Robb, M. A. Conical Intersection Mechanism for Photochemical Ring Opening in Benzospiropyran Compounds. *J. Am. Chem. Soc.* **1997**, *119*, 10815–10820.
- (2) Migani, A.; Robb, M. A.; Olivucci, M. Relationship between Photoisomerization Path and Intersection Space in a Retinal Chromophore Model. *J. Am. Chem. Soc.* **2003**, *125*, 2804–2808.
- (3) Domcke, W.; Yarkony, D. R. Role of Conical Intersections in Molecular Spectroscopy and Photoinduced Chemical Dynamics. *Annu. Rev. Phys. Chem.* **2012**, *63*, 325–352.
- (4) Yarkony, D. R. Nonadiabatic Quantum Chemistry—Past, Present, and Future. *Chem. Rev.* **2012**, *112*, 481–498.
- (5) Köuppel, H.; Domcke, W.; Cederbaum, L. S. Multimode Molecular Dynamics Beyond the Born-Oppenheimer Approximation. *Adv. Chem. Phys.* **2007**, *57*, 59–246.
- (6) Ben-Nun, M.; Quenneville, J.; Martínez, T. J. Ab Initio Multiple Spawning: Photochemistry from First Principles Quantum Molecular Dynamics. *J. Phys. Chem. A* **2000**, *104*, 5161–5175.
- (7) Jasper, A. W.; Nangia, S.; Zhu, C.; Truhlar, D. G. Non-Born–Oppenheimer Molecular Dynamics. *Acc. Chem. Res.* **2006**, *39*, 101–108.

- (8) Worth, G. A.; Cederbaum, L. S. Beyond Born-Oppenheimer: Molecular Dynamics Through a Conical Intersection. *Annu. Rev. Phys. Chem.* **2004**, *55*, 127–158.
- (9) Baer, M. Adiabatic and Diabatic Representations for Atom-Molecule Collisions: Treatment of the Collinear Arrangement. *Chem. Phys. Lett.* **1975**, *35*, 112–118.
- (10) Mead, C. A.; Truhlar, D. G. Conditions for the Definition of a Strictly Diabatic Electronic Basis for Molecular Systems. *J. Chem. Phys.* **1982**, *77*, 6090–6098.
- (11) Viel, A.; Eisfeld, W. Effects of Higher Order Jahn-Teller Coupling on the Nuclear Dynamics. *J. Chem. Phys.* **2004**, *120*, 4603–4613.
- (12) Werner, H.; Meyer, W. MCSCF Study of the Avoided Curve Crossing of the Two Lowest $^1\Sigma^+$ States of LiF. *J. Chem. Phys.* **1981**, *74*, 5802–5807.
- (13) Perić, M.; Buenker, R. J.; Peyerimhoff, S. D. Ab Initio Investigation of the Vibronic Structure of the C_2H Spectrum. *Mol. Phys.* **1990**, *71*, 673–691.
- (14) Petrongolo, C.; Hirsch, G.; Buenker, R. J. Diabatic Representation of the $\tilde{A}^2A_1/\tilde{B}^2B_2$ Conical Intersection in NH_2 . *Mol. Phys.* **1990**, *70*, 825–834.
- (15) Subotnik, J. E.; Yeganeh, S.; Cave, R. J.; Ratner, M. A. Constructing Diabatic States from Adiabatic States: Extending Generalized Mulliken–Hush to Multiple Charge Centers with Boys Localization. *J. Chem. Phys.* **2008**, *129*, 244101.
- (16) Hoyer, C. E.; Parker, K.; Gagliardi, L.; Truhlar, D. G. The DQ and DQ Φ Electronic Structure Diabatization Methods: Validation for General Applications. *J. Chem. Phys.* **2016**, *144*, 194101.
- (17) Cave, R. J.; Newton, M. D. Generalization of the Mulliken–Hush Treatment for the Calculation of Electron Transfer Matrix Elements. *Chem. Phys. Lett.* **1996**, *249*, 15–19.
- (18) Cave, R. J.; Newton, M. D. Calculation of Electronic Coupling Matrix Elements for Ground and Excited State Electron Transfer Reactions: Comparison of the Generalized Mulliken–Hush and Block Diagonalization Methods. *J. Chem. Phys.* **1997**, *106*, 9213–9226.
- (19) Joubert-Doriol, L.; Lasorne, B.; Lauvergnat, D.; Meyer, H.-D.; Gatti, F. A Generalised Vibronic-Coupling Hamiltonian Model for Benzopyran. *J. Chem. Phys.* **2014**, *140*, 044301.
- (20) Lenzen, T.; Manthe, U. Neural Network Based Coupled Diabatic Potential Energy Surfaces for Reactive Scattering. *J. Chem. Phys.* **2017**, *147*, 084105.
- (21) Ruedenberg, K.; Atchity, G. J. A Quantum Chemical Determination of Diabatic States. *J. Chem. Phys.* **1993**, *99*, 3799–3803.
- (22) Atchity, G. J.; Ruedenberg, K. Determination of Diabatic States through Enforcement of Configurational Uniformity. *Theor. Chem. Acc.* **1997**, *97*, 47–58.
- (23) Nakamura, H.; Truhlar, D. G. The Direct Calculation of Diabatic States Based on Configurational Uniformity. *J. Chem. Phys.* **2001**, *115*, 10353.
- (24) Nakamura, H.; Truhlar, D. G. Direct Diabatization of Electronic States by the Fourfold Way. II. Dynamical Correlation and Rearrangement Processes. *J. Chem. Phys.* **2002**, *117*, 5576–5593.
- (25) Nakamura, H.; Truhlar, D. G. Extension of the Fourfold Way for Calculation of Global Diabatic Potential Energy Surfaces of Complex, Multiarrangement, Non-Born-Oppenheimer Systems: Application to $HNCO(S_0, S_1)$. *J. Chem. Phys.* **2003**, *118*, 6816–6829.
- (26) Yang, K. R.; Xu, X.; Truhlar, D. G. Direct Diabatization of Electronic States by the Fourfold-Way: Including Dynamical Correlation by Multi-Configuration Quasidegenerate Perturbation Theory with Complete Active Space Self-Consistent-Field Diabatic Molecular Orbitals. *Chem. Phys. Lett.* **2013**, *573*, 84–89.
- (27) Li, S. L.; Truhlar, D. G.; Schmidt, M. W.; Gordon, M. S. Model Space Diabatization for Quantum Photochemistry. *J. Chem. Phys.* **2015**, *142*, 064106.
- (28) Evenhuis, C. R.; Collins, M. A. Interpolation of Diabatic Potential Energy Surfaces. *J. Chem. Phys.* **2004**, *121*, 2515–2527.
- (29) Godsi, O.; Evenhuis, C. R.; Collins, M. A. Interpolation of Multidimensional Diabatic Potential Energy Matrices. *J. Chem. Phys.* **2006**, *125*, 104105.
- (30) Abrol, R.; Kuppermann, A. An Optimal Adiabatic-to-Diabatic Transformation of the $1^2A'$ and $2^2A'$ States of H_3 . *J. Chem. Phys.* **2002**, *116*, 1035–1062.
- (31) Evenhuis, C.; Martínez, T. J. A Scheme to Interpolate Potential Energy Surfaces and Derivative Coupling Vectors without Performing a Global Diabatization. *J. Chem. Phys.* **2011**, *135*, 224110.
- (32) Lengsfeld, B. H.; Saxe, P.; Yarkony, D. R. On the Evaluation of Nonadiabatic Coupling Matrix Elements Using SA-MCSCF/CI Wave Functions and Analytic Gradient Methods. I. *J. Chem. Phys.* **1984**, *81*, 4549–4553.
- (33) Ågren, H.; Flores-Riveros, A.; Jensen, H. J. A. Evaluation of First- and Second-Order Nonadiabatic Coupling Elements from Large Multiconfigurational Self-Consistent-Field Wave Functions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1986**, *34*, 4606–4614.
- (34) Lengsfeld, B. H.; Yarkony, D. R. On the Evaluation of Nonadiabatic Coupling Matrix Elements for MCSCF/CI Wave Functions Using Analytic Derivative Methods. III. Second Derivative Terms. *J. Chem. Phys.* **1986**, *84*, 348–353.
- (35) Saxe, P.; Yarkony, D. R. On the Evaluation of Nonadiabatic Coupling Matrix Elements for MCSCF/CI Wave Functions. IV. Second Derivative Terms Using Analytic Gradient Methods. *J. Chem. Phys.* **1987**, *86*, 321–328.
- (36) Park, J. W.; Shiozaki, T. On-the-Fly CASPT2 Surface-Hopping Dynamics. *J. Chem. Theory Comput.* **2017**, *13*, 3676–3683.
- (37) Park, J. W.; Shiozaki, T. Analytical Derivative Coupling for Multistate CASPT2 Theory. *J. Chem. Theory Comput.* **2017**, *13*, 2561–2570.
- (38) Park, J. W. Analytical Gradient Theory for Quasidegenerate N-Electron Valence State Perturbation Theory (QD-NEVPT2). *J. Chem. Theory Comput.* **2020**, *16*, 326–339.
- (39) Lischka, H.; Dallos, M.; Szalay, P. G.; Yarkony, D. R.; Shepard, R. Analytic Evaluation of Nonadiabatic Coupling Terms at the MR-CI Level. I. Formalism. *J. Chem. Phys.* **2004**, *120*, 7322–7329.
- (40) Dallos, M.; Lischka, H.; Shepard, R.; Yarkony, D. R.; Szalay, P. G. Analytic Evaluation of Nonadiabatic Coupling Terms at the MR-CI Level. II. Minima on the Crossing Seam: Formaldehyde and the Photodimerization of Ethylene. *J. Chem. Phys.* **2004**, *120*, 7330–7339.
- (41) Zhu, X.; Yarkony, D. R. Toward Eliminating the Electronic Structure Bottleneck in Nonadiabatic Dynamics on the Fly: An Algorithm to Fit Nonlocal, Quasidiabatic, Coupled Electronic State Hamiltonians Based on Ab Initio Electronic Structure Data. *J. Chem. Phys.* **2010**, *132*, 104101.
- (42) Zhu, X.; Ma, J.; Yarkony, D. R.; Guo, H. Computational Determination of the $\tilde{A}$ State Absorption Spectrum of NH_3 and of ND_3 Using a New Quasi-Diabatic Representation of the $\tilde{X}$ and $\tilde{A}$ States and Full Six-Dimensional Quantum Dynamics. *J. Chem. Phys.* **2012**, *136*, 234301.
- (43) Zhu, X.; Yarkony, D. R. Quasi-Diabatic Representations of Adiabatic Potential Energy Surfaces Coupled by Conical Intersections Including Bond Breaking: A More General Construction Procedure and an Analysis of the Diabatic Representation. *J. Chem. Phys.* **2012**, *137*, 22A511.
- (44) Zhu, X.; Yarkony, D. R. Fitting Coupled Potential Energy Surfaces for Large Systems: Method and Construction of a 3-State Representation for Phenol Photodissociation in the Full 33 Internal Degrees of Freedom Using Multireference Configuration Interaction Determined Data. *J. Chem. Phys.* **2014**, *140*, 024112.
- (45) Zhu, X.; Yarkony, D. R. On the Elimination of the Electronic Structure Bottleneck in on the Fly Nonadiabatic Dynamics for Small to Moderate Sized (10–15 Atom) Molecules Using Fit Diabatic Representations Based Solely on Ab Initio Electronic Structure Data: The Photodissociation. *J. Chem. Phys.* **2016**, *144*, 024105.
- (46) Guan, Y.; Fu, B.; Zhang, D. H. Construction of Diabatic Energy Surfaces for LiFH with Artificial Neural Networks. *J. Chem. Phys.* **2017**, *147*, 224307.
- (47) Guan, Y.; Guo, H.; Yarkony, D. R. Neural Network Based Quasi-Diabatic Hamiltonians with Symmetry Adaptation and a Correct Description of Conical Intersections. *J. Chem. Phys.* **2019**, *150*, 214101.

- (48) Guan, Y.; Zhang, D. H.; Guo, H.; Yarkony, D. R. Representation of Coupled Adiabatic Potential Energy Surfaces Using Neural Network Based Quasi-Diabatic Hamiltonians: $1,2\ ^2A'$ States of LiFH. *Phys. Chem. Chem. Phys.* **2019**, *21*, 14205–14213.
- (49) Guan, Y.; Guo, H.; Yarkony, D. R. Extending the Representation of Multistate Coupled Potential Energy Surfaces To Include Properties Operators Using Neural Networks: Application to the $1,2\ ^1A$ States of Ammonia. *J. Chem. Theory Comput.* **2020**, *16*, 302–313.
- (50) Bishop, C. M. *Pattern Recognition and Machine Learning*; Jordan, M., Kleinberg, J., Schölkopf, B., Eds.; Springer: New York, 2006.
- (51) Yarkony, D. R. On the Adiabatic to Diabatic States Transformation near Intersections of Conical Intersections. *J. Chem. Phys.* **2000**, *112*, 2111–2120.
- (52) Atchity, G. J.; Xantheas, S. S.; Ruedenberg, K. Potential Energy Surfaces near Intersections. *J. Chem. Phys.* **1991**, *95*, 1862–1876.
- (53) Nocedal, J.; Wright, S. J. *Numerical Optimization*, 2nd ed.; Mikosch, T. V., Resnick, S. I., Robinson, S. M., Eds.; Springer: New York, 2006.
- (54) Dai, Y. H.; Yuan, Y. A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property. *SIAM J. Optim.* **1999**, *10*, 177–182.
- (55) Alconcel, L. S.; Continetti, R. E. Dissociation Dynamics and Stability of Cyclopentoxo and Cyclopentoxide. *Chem. Phys. Lett.* **2002**, *366*, 642–649.
- (56) Malbon, C. L.; Yarkony, D. R.; Zhu, X. On the Electronic Structure of the Ground State of Cyclopentoxo. The Case for a Two Coupled State Description. *J. Mol. Spectrosc.* **2015**, *311*, 36–41.
- (57) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (58) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (59) Lischka, H.; Müller, T.; Szalay, P. G.; Shavitt, I.; Pitzer, R. M.; Shepard, R. Columbus—a Program System for Advanced Multi-reference Theory Calculations. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 191–199.
- (60) Shepard, R. Geometrical Energy Derivative Evaluation with MRCI Wave Functions. *Int. J. Quantum Chem.* **1987**, *31*, 33–44.
- (61) Shepard, R. *Modern Electronic Structure Theory Part I*; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995.
- (62) Lischka, H.; Shepard, R.; Pitzer, R. M.; Shavitt, I.; Dallos, M.; Müller, T.; Szalay, P. G.; Seth, M.; Kedziora, G. S.; Yabushita, S.; et al. High-Level Multireference Methods in the Quantum-Chemistry Program System COLUMBUS: Analytic MR-CISD and MR-AQCC Gradients and MR-AQCC-LRT for Excited States, GUGA Spin–Orbit CI and Parallel CI Density. *Phys. Chem. Chem. Phys.* **2001**, *3*, 664–673.
- (63) Lischka, H.; Shepard, R.; Shavitt, I.; Pitzer, R. M.; Dallos, M.; Müller, T.; Szalay, P. G.; Brown, F. B.; Ahlrichs, R.; Böhm, H. J.; et al. COLUMBUS, an Ab Initio Electronic Structure Program, Release 7.0; 2017.
- (64) Lischka, H.; Dallos, M.; Shepard, R. Analytic MRCI Gradient for Excited States: Formalism and Application to the $n-\pi^*$ Valence- and $n-(3s,3p)$ Rydberg States of Formaldehyde. *Mol. Phys.* **2002**, *100*, 1647–1658.
- (65) Shepard, R.; Lischka, H.; Szalay, P. G.; Kovar, T.; Ernzerhof, M. A General Multireference Configuration Interaction Gradient Program. *J. Chem. Phys.* **1992**, *96*, 2085–2098.
- (66) Coe, J. D.; Martínez, T. J. Competitive Decay at Two- and Three-State Conical Intersections in Excited-State Intramolecular Proton Transfer. *J. Am. Chem. Soc.* **2005**, *127*, 4560–4561.
- (67) Coe, J. D.; Ong, M. T.; Levine, B. G.; Martínez, T. J. On the Extent and Connectivity of Conical Intersection Seams and the Effects of Three-State Intersections. *J. Phys. Chem. A* **2008**, *112*, 12559–12567.
- (68) Matsika, S. Three-State Conical Intersections in Nucleic Acid Bases. *J. Phys. Chem. A* **2005**, *109*, 7538–7545.
- (69) Kistler, K. A.; Matsika, S. Three-State Conical Intersections in Cytosine and Pyrimidinone Bases. *J. Chem. Phys.* **2008**, *128*, 215102.
- (70) Blancafort, L.; Robb, M. A. Key Role of a Threefold State Crossing in the Ultrafast Decay of Electronically Excited Cytosine. *J. Phys. Chem. A* **2004**, *108*, 10609–10614.
- (71) Köppel, H.; Schubert, B. The Concept of Regularized Diabatic States for a General Conical Intersection. *Mol. Phys.* **2006**, *104*, 1069–1079.
- (72) Ng, A. Y. Feature Selection, L_1 vs. L_2 Regularization, and Rotational Invariance. *Twenty-first international conference on Machine learning - ICML '04* **2004**, 78.
- (73) Bernstein, R. B.; Levine, R. D. Entropy and Chemical Change. I. Characterization of Product (and Reactant) Energy Distributions in Reactive Molecular Collisions: Information and Entropy Deficiency. *J. Chem. Phys.* **1972**, *57*, 434–449.
- (74) Ben-Shaul, A.; Levine, R. D.; Bernstein, R. B. Entropy and Chemical Change. II. Analysis of Product Energy Distributions: Temperature and Entropy Deficiency. *J. Chem. Phys.* **1972**, *57*, 5427–5447.
- (75) Ben-Shaul, A.; Levine, R. D.; Bernstein, R. B. Prior-expectation Distribution Functions for Energy Disposal and Energy Consumption in Reactive Molecular Collisions. *J. Chem. Phys.* **1974**, *61*, 4937–4938.