Machine Learning Seams of Conical Intersection: A Characteristic Polynomial Approach

Tzu Yu Wang,[†] Simon P. Neville,[*,‡] and Michael S. Schuurman[*,‡,†]

[†]*Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, Canada*

[‡]*National Research Council Canada, 100 Sussex Dr., Ottawa, Canada, K1A 0R6*

E-mail: simon.neville@nrc-cnrc.gc.ca; michael.schuurman@uottawa.ca

Abstract

The machine learning of potential energy surfaces (PESs) has undergone rapid progress in recent years. The vast majority of this work, however, has been focused on the learning of ground state PESs. To reliably extend machine learning protocols to excited state PESs, the occurrence of seams of conical intersections between adiabatic electronic states must be correctly accounted for. This introduces a serious problem, for at such points the adiabatic potentials are not differentiable to any order, complicating the application of standard machine learning methods. We show that this issue may be overcome by instead learning the coordinate-dependent coefficients of the characteristic polynomial of a simple decomposition of the potential matrix. We demonstrate that, through this approach, quantitatively accurate machine learning models of seams of conical intersection may be constructed.

Introduction

Potential energy surfaces (PESs) are of central importance in understanding chemical processes, forming a crucial ingredient for the simulation and rationalization of the spectroscopic and dynamic properties of molecular systems. Accordingly, there exists great interest in the construction of accurate PES models fitted to *ab initio* energies. Here, recent advances in machine learning (ML) promise to be transformative,[1–5] with accurate, high-dimensional ML models PESs now being routinely constructed. However, there exists an important caveat to this success: the majority of applications to date have been concerned with the learning of isolated ground state PESs. If one is interested in excited state PESs, then an additional complication must be accounted for. Namely, that, in general, excited state PESs exhibit extended seams of conical intersection (CIs).[6,7] The existence of a CI seam between two or more electronic states means that the corresponding adiabatic PESs belong to the C^0 differentiabilty class over any domain containing it. This lack of differentiability, or non-smoothness, of the PESs on the locus of seam points is highly problematic for many standard ML approaches, including kernel methods, and neural networks employing gradient-based optimization. Such models will struggle to correctly describe CI seams; in terms of topography and, more importantly, the dimensionality of the branching space. That is, the subspace in which the degeneracy between the intersecting PESs is lifted to first-order, a defining property of seams of CI. Given the central role played by CIs in photochemistry and photophysics, this poses a serious problem that previous works have not directly addressed.[8–15]

In this Letter, we demonstrate that accurate ML models of CI seams, including a correct description of the branching space, can in fact be constructed by forgoing the direct learning of PESs. Instead, we advocate for the learning of the nuclear coordinate-dependent coefficients of the characteristic polynomial (CP) of a simple decomposition of the potential matrix. Unlike the adiabatic PESs, these form smooth surfaces, even at points of CI. Furthermore, there exists a simple mapping between the CP coefficients and the adiabatic PESs, making the recovery of the latter a trivial exercise. The efficacy of the proposed

approach is demonstrated via the construction of kernel ridge regression (KRR) models of conical intersections for a number of small molecules, although the conclusions drawn should be transferable to other ML methods, e.g., the construction of neural network potentials.

Let $\boldsymbol{R}$ denote the vector of 3N nuclear coordinates, and $\boldsymbol{V}(\boldsymbol{R})$ the $n \times n$ nuclear coordinate-dependent adiabatic potential matrix with on-diagonal elements $V_{ii}(\boldsymbol{R})$ corresponding to the PESs $E_i(\boldsymbol{R})$ of interest. We begin with the decomposition of the potential matrix into an average energy and splitting matrix,

$$\boldsymbol{V}(\boldsymbol{R}) = \omega(\boldsymbol{R})\mathbf{1}_n + \boldsymbol{Z}(\boldsymbol{R}), \tag{1}$$

where $\omega(\boldsymbol{R}) = \mathrm{Tr}\boldsymbol{V}(\boldsymbol{R})/n$ denotes the average adiabatic energy, and the diagonal splitting matrix $\boldsymbol{Z}(\boldsymbol{R})$ has elements $Z_{ij}(\boldsymbol{R}) = [E_i(\boldsymbol{R}) - \omega(\boldsymbol{R})]\delta_{ij}$. The average energy $\omega(\boldsymbol{R})$ is a smooth function of the nuclear coordinates, even at points of CI. To arrive at a smooth representation of the splitting contribution $\boldsymbol{Z}(\boldsymbol{R})$, we follow Opalka and Domcke[16] and consider its characteristic polynomial

$$\begin{aligned} p^Z(\lambda) &= \det\left[\lambda\mathbf{1}_n - \boldsymbol{Z}(\boldsymbol{R})\right] \\ &= \prod_{i=1}^{n}\left[\lambda - Z_{ii}(\boldsymbol{R})\right] \\ &= \sum_{i=0}^{n-1} c_i^Z(\boldsymbol{R})\lambda^i, \end{aligned} \tag{2}$$

with CP coefficients $c_i^Z(\boldsymbol{R})$ given by

$$
\begin{aligned}
c_0^Z(\boldsymbol{R}) &= (-1)^n \prod_i^n Z_{ii}(\boldsymbol{R}) \\
&\vdots \\
c_{n-2}^Z(\boldsymbol{R}) &= \sum_{i<j}^n Z_{ii}(\boldsymbol{R}) Z_{jj}(\boldsymbol{R}) \\
c_{n-1}^Z(\boldsymbol{R}) &= -\sum_i^n Z_{ii}(\boldsymbol{R}) = 0 \\
c_n^Z(\boldsymbol{R}) &= 1
\end{aligned}
\tag{3}
$$

The CP coefficients are also smooth functions of the nuclear coordinates over any domain, including one that contains a CI. Given the CP coefficients, $c_i^Z(\boldsymbol{R})$, and average energy, $\omega(\boldsymbol{R})$, the adiabatic potentials may easily be recovered as the eigenvalues of the following companion matrix:[16]

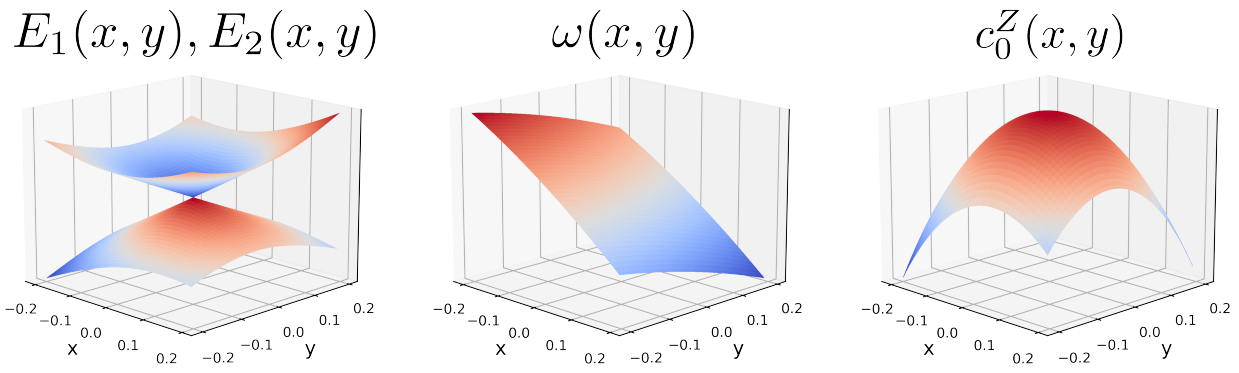

Figure 1: Adiabatic potential energy surfaces, $E_1(x,y)$ and $E_2(x,y)$, average energy, $\omega(x,y)$, and the single splitting matrix CP coefficient, $c_0^Z(x,y)$, for a representative first-order CI model (Equation 7) with parameters $\tilde{g} = \tilde{h} = 0.5$, $s_x = 0.25$, $s_y = 0.1$. Unlike the adiabatic potential energy surfaces, $\omega(x,y)$ and $c_0^Z(x,y)$ are smooth, differentiable functions, even at a point of CI.

$$\boldsymbol{C}(\boldsymbol{R}) = \begin{bmatrix} \omega(\boldsymbol{R}) & 0 & \cdots & 0 & 0 & -c_0^z(\boldsymbol{R}) \\ 1 & \omega(\boldsymbol{R}) & \cdots & 0 & 0 & -c_1^z(\boldsymbol{R}) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \omega(\boldsymbol{R}) & -c_{n-2}^z(\boldsymbol{R}) \\ 0 & 0 & \cdots & 0 & 1 & \omega(\boldsymbol{R}) \end{bmatrix} \tag{4}$$

The advantages of learning the set of functions $\{\omega(\boldsymbol{R}), c_i^Z(\boldsymbol{R}) : i = 0, \ldots, n-2\}$ instead of the PESs $E_i(\boldsymbol{R})$ are two-fold. Firstly, these functions are extremely smooth, as illustrated in Figure 1 for a simple first-order model of a two-state CI. This, in turn, renders them highly amenable to machine learning, unlike the underlying adiabatic PESs, which exhibit discontinuous derivatives at a point of CI. Secondly, and somewhat remarkably, all the branching space information is contained in the single CP coefficient $c_{n-2}^Z(\boldsymbol{R})$, irrespective of the number of intersecting states. To see this, we first note that the following identity holds:

$$\begin{aligned} c_{n-2}^Z(\boldsymbol{R}) &= -\frac{1}{2n}\sum_{i<j}^{n}[E_i(\boldsymbol{R}) - E_j(\boldsymbol{R})]^2 \\ &= -\frac{1}{2n}\sum_{i<j}^{n}\Delta E_{ij}^2. \end{aligned} \tag{5}$$

The proof of this is somewhat non-obvious and is given in the Supplementary Information. Now, since the squared energy differences $\Delta E_{ij}^2(\boldsymbol{R})$ are lifted to second-order at a point of CI with respect to the branching space coordinates, and to fourth-order with respect to the remaining seam space coordinates, the Hessian

$$\mathcal{H}_{\alpha\beta}(\boldsymbol{R}') = \left.\frac{\partial^2 c_{n-2}^Z}{\partial R_\alpha \partial R_\beta}\right|_{\boldsymbol{R}'}, \tag{6}$$

when evaluated at a point of CI, $\boldsymbol{R}_{CI}$, has a column space corresponding to the branching space, and a null space corresponding to the seam space. Thus, the orthogonalized branching (seam space) coordinates may be computed as the eigenvectors of $\boldsymbol{\mathcal{H}}(\boldsymbol{R}_{CI})$ with non-zero

(zero) eigenvalues.[17–19] Further, as detailed below, the non-zero eigenpairs can be used to characterize the conical intersection topography. Thus, in addition to being smooth, slowly-varying functions, the CP coefficients provide a simple, yet fundamental, encapsulation of the branching space information, irrespective of the number of intersecting states.

To show the advantages of constructing ML models of seams of CIs using the above-described CP formalism, KRR models were constructed for three prototypical CIs: (1) the symmetry-required $(E \otimes e)$ two-state CI between the two components of the D_1 state of NH_3^+; (2) the accidental 'twisted-pyramidalized' (TwPy) CI between the S_0 and S_1 states in ethylene, and; (3) the symmetry-required $(T_2 \otimes (e \oplus t_2))$ three-state intersection beteen the components of the D_0 state of CH_4^+. In addition, in order to highlight the shortcomings of the traditional, direct approach, KRR models of the adiabatic PESs themselves were computed for all three systems. In the following, we shall refer to these two sets of models as the 'ω-CP' and 'direct-energy' models.

In all cases, radial basis function (RBF) kernel was used in conjunction with the Smoothed Overlap of Atomic Positions (SOAP) descriptor.[20,21] In brief, the SOAP descriptor transforms the Cartesian nuclear coordinates into a power spectrum through a linear combination of Gaussian type orbital function and spherical harmonics centered at each atom site, which satisfies the desired symmetries and bijection of a molecular descriptor. More importantly, the SOAP descriptor allows for scaling to large molecules due to being independent of the molecule size.[22] Training sets were constructed at the multireference configuration interaction (MRCI) level of theory via latin hypercube sampling about each minimum energy CI (MECI) geometry. The parameters of the SOAP descriptor were separately optimised for each system using a genetic algorithm approach. The details of these calculations are given in the Supplementary Information.

We first consider the quality of the ω-CP and direct-energy KRR models as judged by the mean absolute errors (MAEs) in the vicinity of the MECI geometries as a function of training set size. These are shown in Figure 2. In all cases, sub-chemical accuracy in the learned

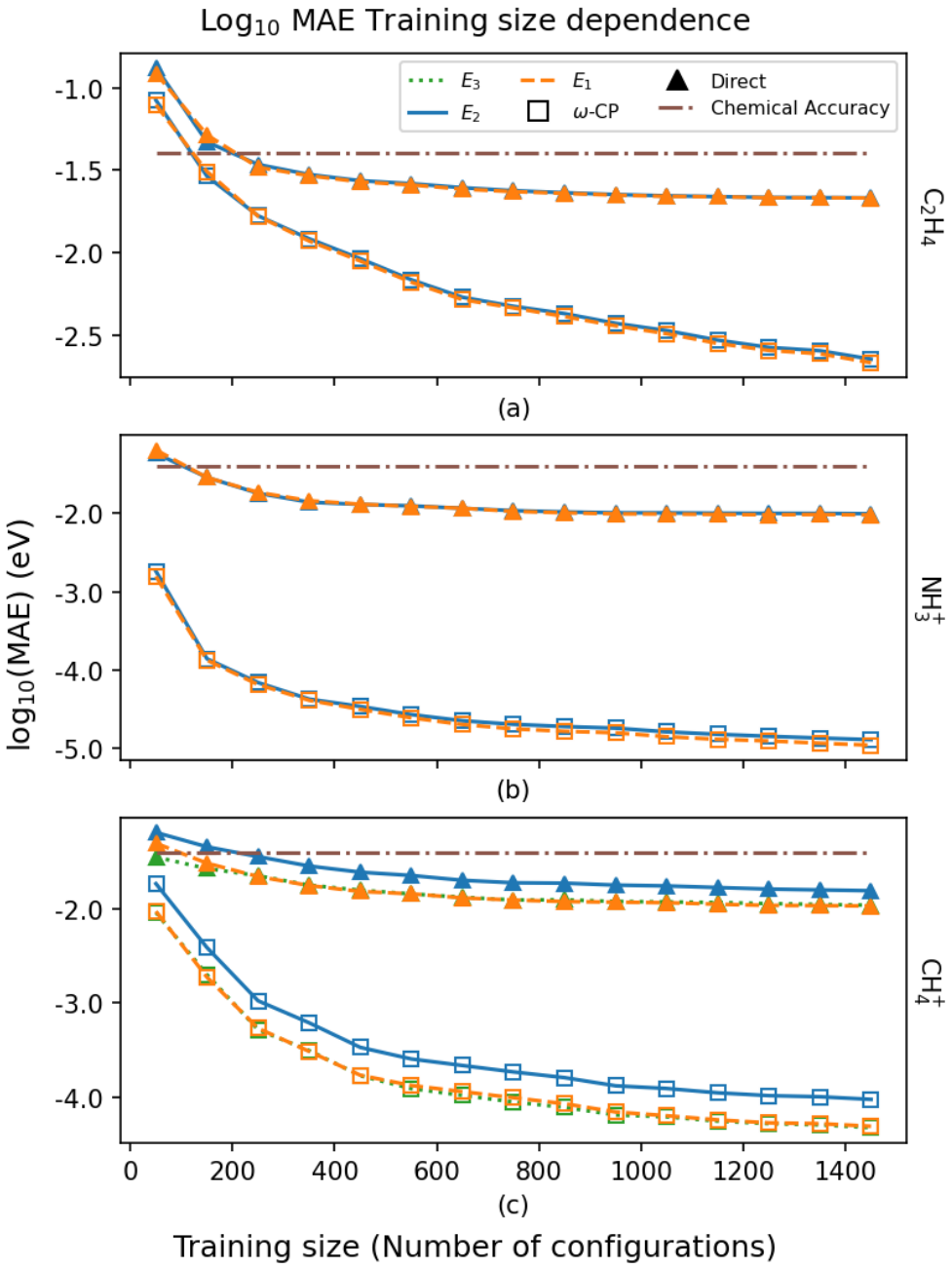

Figure 2: log of Mean absolute error as a function of training set size, in terms of number of nuclear configurations, when fitting the characteristic polynomial parameters (ω-CP) or the adiabatic energies (Direct).

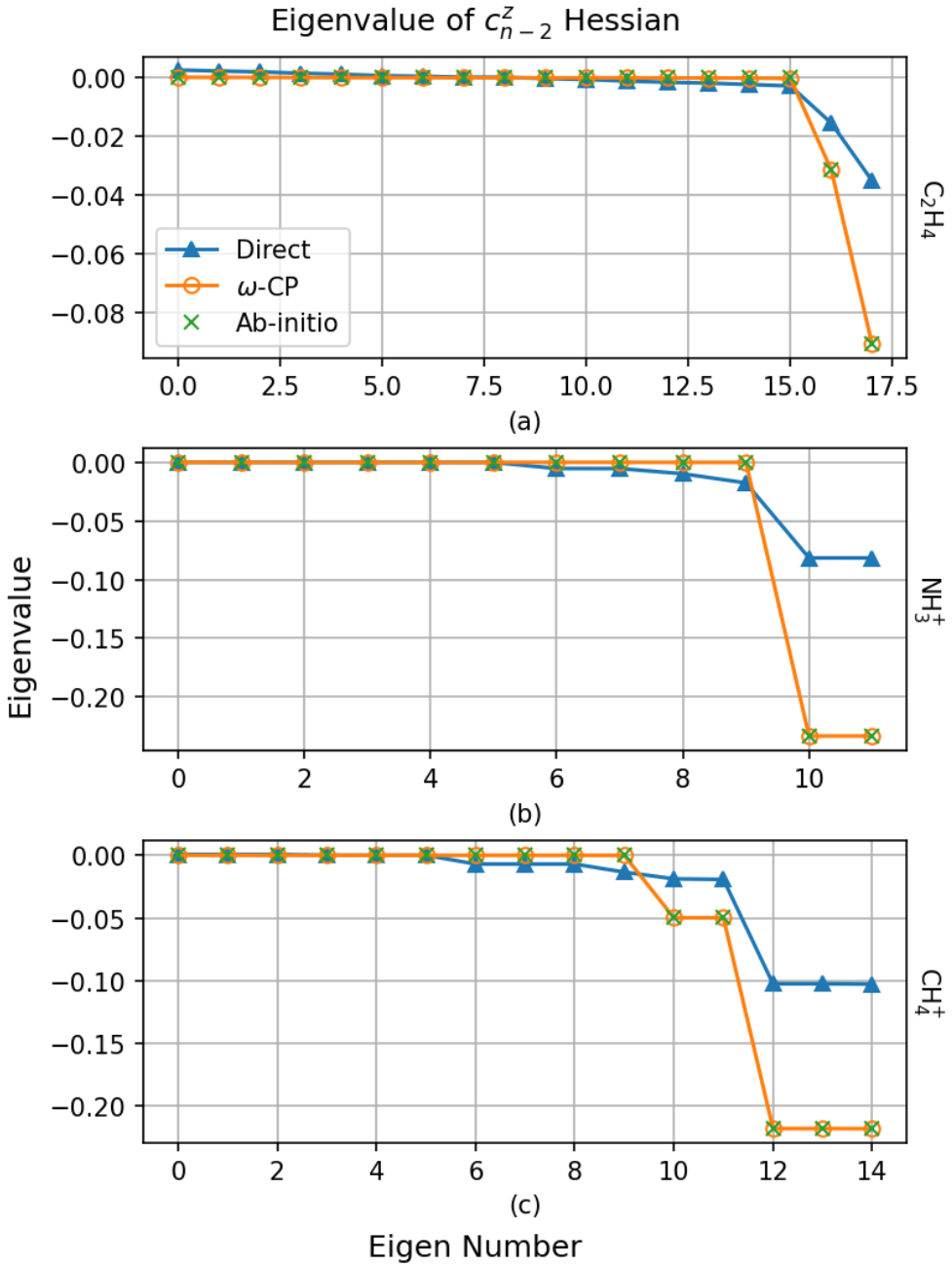

Figure 3: Eigenvalues of the c^z_{n-2} Hessian matrix as computed from *ab initio* data and using the ω-CP and direct energy fit surrogate potentials. The number of non-zero eigenvalues correspond to the dimension of the branching space.

PESs is attained with fewer than 300 training points. We note, however, that the PESs obtained from the ω-CP KRR models are consistently more accurate for a given training set size than the corresponding direct-energy models. The ability of the direct-energy models to accurately describe on average the adiabatic PESs in a subspace containing a CI, however, belies a larger failing. Namely, an inability to correctly describe the branching spaces, both in terms of direction and dimensionality. To see this, we consider the eigenvalues of the $c^Z_{n-2}(\boldsymbol{R})$ Hessian $\boldsymbol{\mathcal{H}}$ (Equation 6) evaluated at the MECI points using both the ω-CP and direct-energy models. These are shown in Figure 3 alongside the *ab initio* MRCI values. The number of non-zero eigenvalues of $\boldsymbol{\mathcal{H}}$ should equal the dimensionality of the branching space: two for a two-state CI and five for a three-state CI. In all cases the ω-CP models quantitatively reproduce the *ab initio* eigenvalues. On the other hand, the direct-energy models fail rather badly: for the NH_3^+ and CH_4^+ models, the dimensionality of the predicted branching space is too high. For the C_2H_4 model, only two significantly-non-zero eigenvalues are furnished by the direct-energy model. However, the values of these are significantly underestimated which, as we discuss below, is related to an incorrect description of the CI topography.

Having established the ability of the ω-CP models to correctly recover the dimensionality of the branching space, we now consider their ability to describe the conical intersection topography, that is, the parameters describing the tilt and pitch of the cone. Here, we restrict ourselves to the two-state CI case, which is described to first-order within the branching space by following model potential:[23,24]

$$\boldsymbol{W}^{(1)}(x,y) = (s_x x + s_y y)\,\mathbf{1}_2 + \begin{bmatrix} -\tilde{g}x & \tilde{h}y \\ \tilde{h}y & \tilde{g}x \end{bmatrix}. \tag{7}$$

Here, $\tilde{g}$ and $\tilde{h}$ are the norms of the orthogonalized gradient difference and derivative coupling vectors, respectively,

$$\tilde{\boldsymbol{g}} = -\cos\beta\boldsymbol{g} + \sin\beta\boldsymbol{h}, \tag{8}$$

$$\tilde{\boldsymbol{h}} = \sin\beta\boldsymbol{g} + \cos\beta\boldsymbol{h}, \tag{9}$$

$$\beta = \frac{2\boldsymbol{g}^T\boldsymbol{h}}{\boldsymbol{h}^T\boldsymbol{h} - \boldsymbol{g}^T\boldsymbol{g}}, \tag{10}$$

where $\boldsymbol{g}$ and $\boldsymbol{h}$ denote the nascent gradient difference and derivative coupling vectors. Together, the parameters $\tilde{g}$ and $\tilde{h}$ describe the asymmetry in the pitch of the cone. The parameters

$$s_x = \frac{\boldsymbol{s}^T\tilde{\boldsymbol{g}}}{\tilde{g}}, \tag{11}$$

and

$$s_y = \frac{\boldsymbol{s}^T\tilde{\boldsymbol{h}}}{\tilde{h}}, \tag{12}$$

$$\boldsymbol{s} = \frac{1}{2}\frac{\partial(E_1 + E_2)}{\partial\boldsymbol{R}}\Big|_{\boldsymbol{R}_{CI}}, \tag{13}$$

describe the tilt of the principal axis of the CI cone. For the ω-CP models, all first-order parameters are trivial to compute. The branching space coordinates x and y correspond to the column space of the Hessian $\boldsymbol{\mathcal{H}}$ of $c^Z_{n-2}(\boldsymbol{R})$. The corresponding non-zero eigenvalues take values of $-2\tilde{g}^2$ and $-2\tilde{h}^2$,[17–19] and the tilt parameters s_x and s_y are obtained using the relation

$$\boldsymbol{s} = \frac{\partial\omega}{\partial\boldsymbol{R}}\Big|_{\boldsymbol{R}_{CI}}. \tag{14}$$

The values of the first-order parameters computed using the ω-CP and direct-energy

models are shown in Table 1 for both two-state models. In addition, we also give the angles $\theta_{x/y}$ between the model and *ab initio* branching space vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. The ω-CP models are found to quantitatively reproduce the CI topography in both cases, with branching vector angles $\theta_{x/y}$ of less than 0.5° and parameters $\tilde{g}$, $\tilde{h}$ and $s_{x/y}$ in almost perfect agreement with the *ab initio* values. The direct-energy models, on the other hand, fare less well, with maximum branching vector angles $\theta_{x/y}$ angles of 8.0° and parameters $\tilde{g}$, $\tilde{h}$ and $s_{x/y}$ that fail to correctly describe the CI topography.

Table 1: Comparison of the branching space parameters determined from *ab initio* data and using the surrogate potentials. The parameters $\tilde{g}$ and $\tilde{h}$ are the norms of the orthogonalized gradient difference and non-adiabatic coupling vectors, respectively, and the θ_i, $i = x, y$ are the angles between the *ab initio* and ω-CP and direct-energy KRR model branching space vector vectors x and y. Angles are given in units of degrees. All other values are given in units of $\mathrm{E}_h/\mathring{A}$.

		θ_x	θ_y	$\tilde{g}$	$\tilde{h}$	s_x	s_y
C_2H_4	*ab initio*	0.0	0.0	0.213	0.125	-0.130	-0.046
	ω-CP	0.0	0.3	0.214	0.126	-0.130	-0.046
	direct-energy	3.7	8.0	0.130	0.084	0.128	-0.050
NH_3^+	*ab initio*	0.0	0.0	0.342	0.342	0.000	0.000
	ω-CP	0.3	0.3	0.342	0.342	0.000	0.000
	direct-energy	5.3	5.3	0.202	0.202	0.000	0.001

Next, we show in Figure 4 both sets of model potentials for the two-state systems along the two *ab initio* branching space coordinates, x and y. Two important aspects are immediately obvious. Firstly, the direct-energy models actually yield avoided crossings, not CIs. On the other hand, the ω-CP models correctly reproduce the intersection of the PESs. Secondly, and perhaps more surprisingly, the ω-CP models are able to accurately extrapolate the PESs to geometries far away from the training sets used to construct them. To see this, we first refer to Figure 5, where the model branching space cuts are extended to large displacements. Both ω-CP models remain accurate out to displacements of x=0.4 and y=0.4, which correspond to geometries outside the span of the training sets. This can be discerned from Figures S4 and S5, in which we show, respectively, a superposition of the training set geometries and the geometries corresponding to $x = 0.4$. Somewhat remarkably, for ethylene, this geometry

corresponds to ethylidene, a different structural isomer to the geometries present in the training set. Similarly, the ammonia geometry at x=0.4 corresponds to a near-dissociated N-H bond which, again, is not represented in the training set. The ability of the ω-CP models to extrapolate is a direct result of the significantly longer length scales on which the surfaces $\omega(\boldsymbol{R})$ and $c_i^Z(\boldsymbol{R})$ vary compared to the PESs $E_i(\boldsymbol{R})$. This, in turn results in large kernel length scales and an increased distance from which knowledge may be transferred from the training to prediction points. On the other hand, the rapidly-varying adiabatic PESs result in direct-energy KRR models that rapidly lose all predictive power when moving away from elements of the training set.

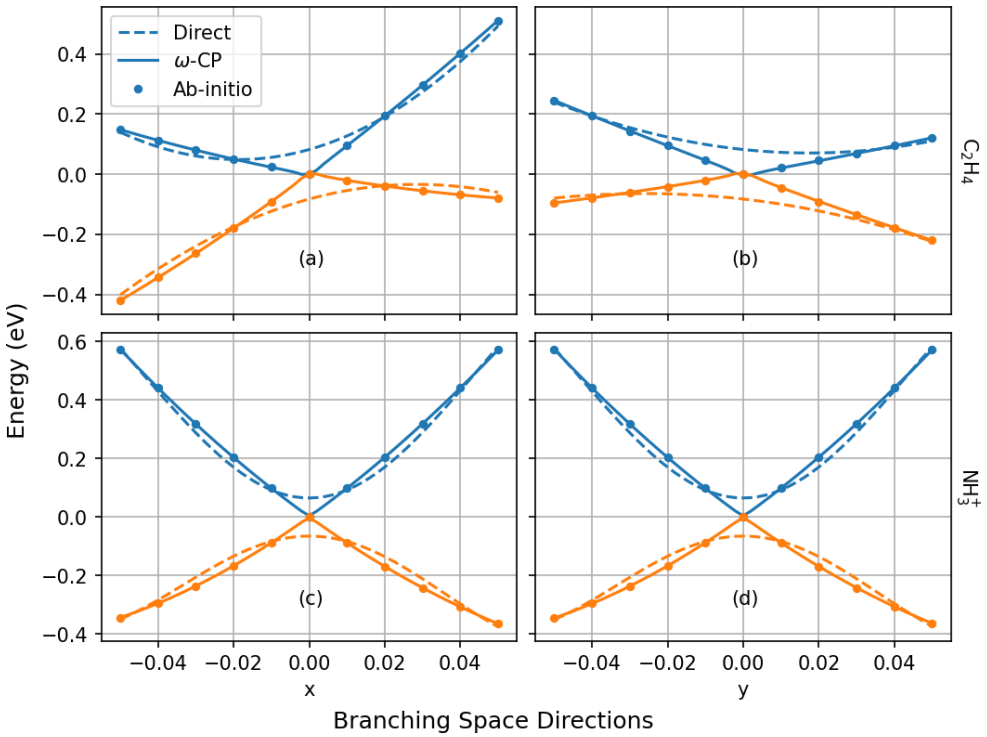

Figure 4: Adiabatic potentials in the immediate vicinity of a minimum energy conical intersection as determined from *ab initio* computations (isolated points) and the ω-CP and direct energy surrogate potentials. The latter fail to capture the degeneracy and instead show an avoided crossing.

It remains to comment on the choice of kernel. The RBF kernel used here is a simple 'default' choice in many kernel regression based methods. Belonging to the C^∞ differentiability class, it performs well for the smooth $\omega(\boldsymbol{R})$ and $c_i^Z(\boldsymbol{R})$ functions, the main focus of

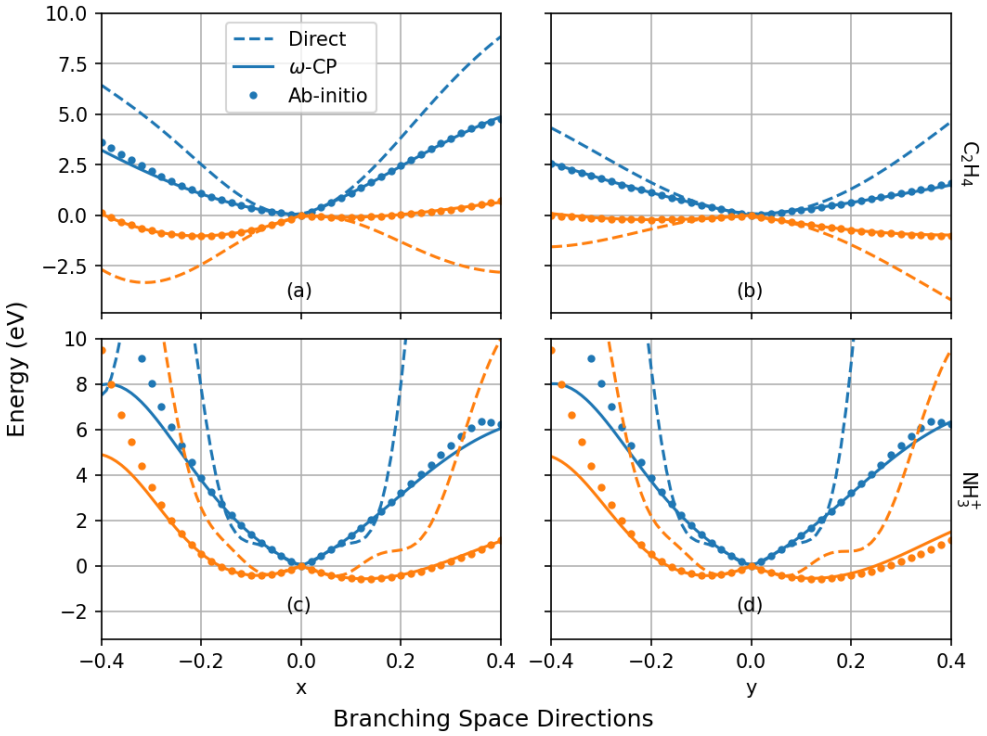

Figure 5: Adiabatic potentials over an extended region of coordinate space along the (rectilinear) branching space directions, x and y, for the two-state examples. While the points used to fit the ω-CP and direct-energy surrogate potentials were within 1.5 eV of the MECI, the ω-CP models exhibit impressive fidelity to the *ab initio* surfaces over a much larger energy range.

this work. However, for the direct learning of the PESs $E_i(\boldsymbol{R})$, other kernel choices might give improved results. An obvious choice here is the Matérn-1/2 kernel, which belongs to the same differentiability class (C^0) as a set of intersecting adiabatic PESs. However, even using such a kernel, poor accuracy of the direct-energy KRR models was attained. This is demonstrated in Figures S2 and S3, where the direct-energy Matérn-1/2 kernel model PESs are shown plotted along the branching space coordinates for ethylene. Even with this non-differentiable kernel choice, the model PESs form an avoided crossing instead of a CI. Yet more complicated, and potentially better performing, kernels could be conceived of, e.g., additive kernels designed to capture strong sub-dimensional interactions.[25] However, we believe that a strength of our proposed approach is that the complex intersecting PES structure can be transformed into functions that can be well described by a simple kernel, as exemplified by the RBF kernel.

Finally, we compare the proposed CP-based approach to learning seams of CI to alternative approaches. An obvious alternative would be to learn quasi-diabatic potential matrices,[26–28] the elements of which are also smooth functions of the nuclear coordinates. Indeed, this may be an advantageous approach if non-adiabatic couplings are also needed. If this is not the case, however, then the ω-CP approach seems advantageous for three main reasons. Firstly, quasi-diabatic potentials are somewhat ill-defined, a result of the non-existence of strictly diabatic states for polyatomic molecules.[29] The functions $\omega(\boldsymbol{R})$ and $c_i^Z(\boldsymbol{R})$, on the other hand, form a fundamental representation of the potential matrix, being as they are invariant to unitary transformations of the choice of electronic basis. Secondly, and related to the prior point, the $\omega(\boldsymbol{R})$ and $c_i^Z(\boldsymbol{R})$ can be trivially computed from the adiabatic potentials $E_i(\boldsymbol{R})$, which are directly furnished from quantum chemistry calculations. This stands in contrast to quasi-diabatic potentials, for which a subsequent adiabatic-to-diabatic transformation must be applied, typically requiring either the additional calculation of derivative couplings or electronic wave function overlaps. Lastly, the set $\{\omega(\boldsymbol{R}), c_i^Z(\boldsymbol{R}) : i = 0, \ldots, n-2\}$ contains only n quantities to be learned, whilst the

quasi-diabatic potential matrix is composed of $n(n+1)/2$ symmetry-unique elements. Thus, the ω-CP formalism provides a more compact representation of a set of n adiabatic PESs.

To conclude, we have demonstrated the ability to construct quantitatively accurate ML models of seams of CI via an indirect learning of the involved adiabatic PESs based on a CP formalism. The advocated ω-CP approach yields a correct description of the branching and seam spaces, a feat that seems out of reach for models based on the direct learning of adiabatic PESs. Furthermore, it is found that a single CP coefficient, $c^{Z}_{n-2}(\boldsymbol{R})$, contains all branching space information, irrespective of the number of intersecting states. In terms of practical application, the ω-CP formalism will be of use in a number of situations. For example, in MECI optimisation using quantum chemistry methods for which analytical gradients are not available; in such cases ω-CP based surrogate potentials may straightforwardly be utilized. Additionally, the use of ML adiabatic PESs and non-adiabatic couplings in excited-state dynamics simulations is starting to gain traction.[30,31] Here, there seems to be no reason to continue using directly-learned adiabatic PESs, given the unambiguous advantages of the indirect ω-CP approach. We thus anticipate that the results presented here shall be of great use in directing future developments in this nascent, yet important, field of work.

Acknowledgement

M.S.S. thanks the Natural Sciences and Engineering Research Council (NSERC) Discovery grant program for financial support, and all authors thank Isaac Tamblyn for helpful discussions.

Supporting Information Available

Additional information on the generation of the training sets, KRR hyperparameter optimization, choice of kernel, proof of Equation 5, and the ability of the ω-CP models to extrapolate are given in the Supplementary Information.

References

(1) Dral, P. O. Quantum Chemistry in the Age of Machine Learning. *The Journal of Physical Chemistry Letters* **2020**, *11*, 2336–2347, PMID: 32125858.

(2) Behler, J. Perspective: Machine learning potentials for atomistic simulations. *The Journal of Chemical Physics* **2016**, *145*, 170901.

(3) Wang, H.; Zhang, L.; Han, J.; E, W. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Computer Physics Communications* **2018**, *228*, 178–184.

(4) Schütt, K. T.; Arbabzadah, F.; Chmiela, S.; Müller, K. R.; Tkatchenko, A. Quantum-chemical insights from deep tensor neural networks. *Nature Communications* **2017**, *8*, 13890.

(5) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. SchNet – A deep learning architecture for molecules and materials. *The Journal of Chemical Physics* **2018**, *148*, 241722.

(6) Yarkony, D. R. Diabolical conical intersections. *Rev. Mod. Phys.* **1996**, *68*, 985–1013.

(7) Domcke, W.; Yarkony, D. R.; Köppel, H. *Conical Intersections*; WORLD SCIENTIFIC, 2004.

(8) Westermayr, J.; Gastegger, M.; Marquetand, P. Combining SchNet and SHARC: The SchNarc Machine Learning Approach for Excited-State Dynamics. *The Journal of Physical Chemistry Letters* **2020**, *11*, 3828–3834, PMID: 32311258.

(9) Westermayr, J.; Marquetand, P. Machine learning and excited-state molecular dynamics. *Machine Learning: Science and Technology* **2020**, *1*, 043001.

(10) Westermayr, J.; Marquetand, P. Machine Learning for Electronically Excited States of Molecules. *Chemical Reviews* **2021**, *121*, 9873–9926, PMID: 33211478.

(11) Tang, D.; Jia, L.; Shen, L.; Fang, W.-H. Fewest-Switches Surface Hopping with Long Short-Term Memory Networks. *The Journal of Physical Chemistry Letters* **2022**, *13*, 10377–10387, PMID: 36317657.

(12) Li, J.; Reiser, P.; Boswell, B. R.; Eberhard, A.; Burns, N. Z.; Friederich, P.; Lopez, S. A. Automatic discovery of photoisomerization mechanisms with nanosecond machine learning photodynamics simulations. *Chem. Sci.* **2021**, *12*, 5302–5314.

(13) Li, J.; Stein, R.; Adrion, D. M.; Lopez, S. A. Machine-Learning Photodynamics Simulations Uncover the Role of Substituent Effects on the Photochemical Formation of Cubanes. *Journal of the American Chemical Society* **2021**, *143*, 20166–20175, PMID: 34787403.

(14) Li, J.; Lopez, S. A. Excited-State Distortions Promote the Photochemical 4pi Electrocyclizations of Fluorobenzenes via Machine Learning Accelerated Photodynamics Simulations. *Chemistry – A European Journal* **2022**, *28*, e202200651.

(15) Axelrod, S.; Shakhnovich, E.; Gómez-Bombarelli, R. Excited state non-adiabatic dynamics of large photoswitchable molecules using a chemically transferable machine learning potential. *Nature Communications* **2022**, *13*, 3440.

(16) Opalka, D.; Domcke, W. Interpolation of multi-sheeted multi-dimensional potential-energy surfaces via a linear optimization procedure. *The Journal of Chemical Physics* **2013**, *138*, 224103.

(17) Köppel, H.; Schubert, B. The concept of regularized diabatic states for a general conical intersection. *Molecular Physics* **2006**, *104*, 1069–1079.

(18) Kammeraad, J. A.; Zimmerman, P. M. Estimating the Derivative Coupling Vector Using Gradients. *The Journal of Physical Chemistry Letters* **2016**, *7*, 5074–5079, PMID: 27973885.

(19) Gonon, B.; Perveaux, A.; Gatti, F.; Lauvergnat, D.; Lasorne, B. On the applicability of a wavefunction-free, energy-based procedure for generating first-order non-adiabatic couplings around conical intersections. *The Journal of Chemical Physics* **2017**, *147*, 114114.

(20) Bartók, A. P.; Kondor, R.; Csányi, G. On representing chemical environments. *Phys. Rev. B* **2013**, *87*, 184115.

(21) Himanen, L.; Jäger, M. O.; Morooka, E. V.; Federici Canova, F.; Ranawat, Y. S.; Gao, D. Z.; Rinke, P.; Foster, A. S. DScribe: Library of descriptors for machine learning in materials science. *Computer Physics Communications* **2020**, *247*, 106949.

(22) Barnard, T.; Tseng, S.; Darby, J. P.; Bartók, A. P.; Broo, A.; Sosso, G. C. Leveraging genetic algorithms to maximise the predictive capabilities of the SOAP descriptor. *Mol. Syst. Des. Eng.* **2023**, *8*, 300–315.

(23) Atchity, G. J.; Xantheas, S. S.; Ruedenberg, K. Potential energy surfaces near intersections. *The Journal of Chemical Physics* **1991**, *95*, 1862–1876.

(24) Yarkony, D. R. Energies and Derivative Couplings in the Vicinity of a Conical Intersection Using Degenerate Perturbation Theory and Analytic Gradient Techniques. 1. *The Journal of Physical Chemistry A* **1997**, *101*, 4263–4270.

(25) Hiroki Sugisawa, T. I.; Krems, R. V. Gaussian process model of 51-dimensional potential energy surface for protonated imidazole dimer. *The Journal of Chemical Physics* **2020**, *153*, 114101.

(26) Richings, G. W.; Habershon, S. MCTDH on-the-fly: Efficient grid-based quantum dynamics without pre-computed potential energy surfaces. *The Journal of Chemical Physics* **2018**, *148*, 134116.

(27) Richings, G. W.; Habershon, S. Predicting Molecular Photochemistry Using Machine-Learning-Enhanced Quantum Dynamics Simulations. *Accounts of Chemical Research* **2022**, *55*, 209–220, PMID: 34982533.

(28) Fdez. Galván, I.; Lindh, R. Smooth Things Come in Threes: A Diabatic Surrogate Model for Conical Intersection Optimization. *Journal of Chemical Theory and Computation* **2023**, *19*, 3418–3427, PMID: 37192531.

(29) Mead, C. A.; Truhlar, D. G. Conditions for the definition of a strictly diabatic electronic basis for molecular systems. *The Journal of Chemical Physics* **1982**, *77*, 6090–6098.

(30) Westermayr, J.; Gastegger, M.; Menger, M. F. S. J.; Mai, S.; González, L.; Marquetand, P. Machine learning enables long time scale molecular photodynamics simulations. *Chem. Sci.* **2019**, *10*, 8100–8107.

(31) Westermayr, J.; Gastegger, M.; Vörös, D.; Panzenboeck, L.; Joerg, F.; González, L.; Marquetand, P. Deep learning study of tyrosine reveals that roaming can lead to photodamage. *Nature Chemistry* **2022**, *14*, 914–919.