



HAL
open science

Diabatic neural network potentials for accurate vibronic quantum dynamics -The test case of planar NO₃

David M. G. Williams, Wolfgang Eisfeld, Alexandra Viel

► To cite this version:

David M. G. Williams, Wolfgang Eisfeld, Alexandra Viel. Diabatic neural network potentials for accurate vibronic quantum dynamics -The test case of planar NO₃. *Journal of Chemical Physics*, 2019, 151 (16), pp.164118. 10.1063/1.5125851 . hal-02317939

HAL Id: hal-02317939

<https://hal.science/hal-02317939>

Submitted on 29 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Diabatic neural network potentials for accurate vibronic quantum dynamics—The test case of planar NO_3

Cite as: J. Chem. Phys. **151**, 164118 (2019); <https://doi.org/10.1063/1.5125851>

Submitted: 27 August 2019 . Accepted: 11 October 2019 . Published Online: 28 October 2019

David M. G. Williams , Alexandra Viel , and Wolfgang Eisfeld 



View Online



Export Citation



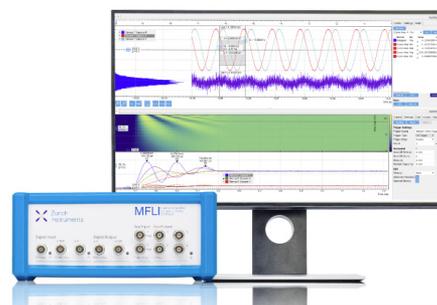
CrossMark

Challenge us.

What are your needs for periodic
signal detection?



Zurich
Instruments



Diabatic neural network potentials for accurate vibronic quantum dynamics—The test case of planar NO₃

Cite as: J. Chem. Phys. 151, 164118 (2019); doi: 10.1063/1.5125851

Submitted: 27 August 2019 • Accepted: 11 October 2019 •

Published Online: 28 October 2019



View Online



Export Citation



CrossMark

David M. G. Williams,^{1,a)}  Alexandra Viel,^{2,b)}  and Wolfgang Eisfeld^{1,c)} 

AFFILIATIONS

¹Theoretische Chemie, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany

²Univ Rennes, CNRS, IPR (Institut de Physique de Rennes) - UMR 6251, F-35000 Rennes, France

^{a)}Electronic mail: d.williams@uni-bielefeld.de

^{b)}Electronic mail: alexandra.viel@univ-rennes1.fr

^{c)}Electronic mail: wolfgang.eisfeld@uni-bielefeld.de

ABSTRACT

A recently developed scheme to produce high-dimensional coupled diabatic potential energy surfaces (PESs) based on artificial neural networks (ANNs) [D. M. G. Williams and W. Eisfeld, J. Chem. Phys. 149, 204106 (2019)] is tested for its viability for quantum dynamics applications. The method, capable of reproducing high-quality *ab initio* data with excellent accuracy, utilizes simple coupling matrices to produce a basic low-order diabatic potential matrix as an underlying backbone for the model. This crude model is then refined by making its expansion coefficients geometry-dependent by the output neurons of the ANN. This structure, strongly guided by a straightforward physical picture behind nonadiabatic coupling, combines structural simplicity with high accuracy, reproducing *ab initio* data without introducing unphysical artifacts to the surface, even for systems with complicated electronic structure. The properties of diabatic potentials obtained by this method are tested thoroughly in the present study. Vibrational/vibronic eigenstates are computed on the \tilde{X} and \tilde{A} states of NO₃, a notoriously difficult Jahn-Teller system featuring strong nonadiabatic couplings and complex spectra. The method is investigated in terms of how consistently it produces dynamics results for PESs of similar (fitting) quality and how the results depend on the ANN size and ANN topography. A central aspect of this work is to understand the convergence properties of the new method in order to evaluate its predictive power. A previously developed, high-quality model utilizing a purely (high-order) polynomial *ansatz* is used as a reference to showcase improvements of the overall quality which can be obtained by the new method.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5125851>

I. INTRODUCTION

The development of accurate high-dimensional potential energy surfaces (PESs) is of fundamental interest for the theoretical treatment of molecular systems, providing a foundation for quantum-dynamical studies. Several strategies have been established for developing such PESs in the case of a single adiabatic electronic state, such as local interpolation techniques,^{1–5} or least-squares fitting approaches utilizing invariant polynomials encoding complete nuclear permutation-inversion (CNPI) symmetry of indistinguishable nuclei.⁶ However, extending such approaches to PESs including excited states proves difficult as state-state interactions enter the

picture. Attempts to tackle this issue include the use of the modified Shepard interpolation^{7–9} or more recently invariant polynomials and the CNPI symmetry.^{10,11}

A fundamental concern that arises when accounting for state-state interactions is the appropriate representation of PESs and their couplings as the Born-Oppenheimer approximation will not hold for certain regions in nuclear configuration space (NCS) where interactions among electronic states become significant. In such a case, a *quasidiabatic* representation of the coupled electronic states has been found to be of great advantage.^{12–25} For the sake of brevity, we will refer to quasidiabatic representations as “diabatic” representations from here on. Such a representation requires the state

basis to represent the electronic Hamiltonian and to preserve the character of the electronic states as much as possible throughout NCS such that derivative coupling becomes small enough that it can be neglected safely. As a consequence, the electronic Hamiltonian is represented by a nondiagonal matrix encoding all state energies and couplings. Accordingly, the matrix elements of the electronic Hamiltonian become well-behaved functions of the nuclear coordinates, decreasing the complexity required to express them analytically. A commonly used, straightforward approach utilizing diabatic representations is the multimode linear vibronic coupling method by Köppel, Domcke, and Cederbaum.²⁶ Analogous methods have been developed also by Yarkony and co-workers.^{27–30} Such methods, while well-suited for ultrafast nonadiabatic processes, do not provide the flexibility needed to express extended regions of the coupled PESs needed for more complicated dynamical processes. Consequently, various approaches have been developed to properly account for extended regions of the PES, such as extensions of vibronic coupling models to higher orders^{31–36} and methods utilizing properties of the adiabatic electronic wave functions for a diabaticization.^{18,21,22,24,37–42} Members of the latter group of diabaticization approaches involving electronic wave function information usually do not require a model diabatic matrix, thus not yielding a set of PESs in closed form. Hence, these (pointwise) diabaticization methods require a second step in which an external model is used to construct (fit) the diabatic matrix elements.

Overall, constructing accurate diabatic models in closed form beyond simple models such as linear vibronic coupling remains a significant problem for systems beyond triatomics.^{43–46} Attempts to do so include the aforementioned extension of vibronic coupling models in a variety of ways,^{31–36} using the modified Shepard interpolation^{7–9} or invariant polynomials and the CNPI symmetry¹⁰ as well as by choosing elaborate functional forms for the diabatic matrix elements.^{43–45} Nonetheless, instances of such diabatic surfaces remain scarce in the literature.

An alternative, more recent approach to represent PESs involves using artificial neural networks (ANNs),^{47–64} which are fitted (“trained”) against *ab initio* or some reference data. ANNs, as has been shown mathematically, are capable of uniformly representing any continuous real function of n dimensions up to arbitrary accuracy.⁶⁵ This makes them in principle suitable for representing arbitrary PESs up to the limitations of the underlying data. Another prominent feature of ANNs is that, once trained, they can be evaluated very efficiently as their evaluation consists mostly of matrix-vector multiplications, making them particularly suitable for quantum dynamics methods such as multi-configuration time-dependent Hartree (MCTDH),^{66,67} especially when using the correlated discrete variable representation (CDVR)⁶⁸ scheme for which the evaluation of the PES is the most time demanding part. Without using the CDVR approach, a sum of products form of the PES model can be accounted for by special formulations of the ANN.^{56,64,69} Despite their mathematical properties, the principal capabilities of ANNs are of course limited by practical concerns such as the acquisition of data, limitations of the training algorithm, and an issue commonly referred to as “overfitting.” However, ANNs have been used already with impressive results to represent a single PES based on high-level *ab initio* data.^{54,63,64} Very recently, the first attempts to extend the use of ANNs to diabatic PES models were documented in the literature.^{70–73}

Our recently developed approach presented in Ref. 73 combines the basic, simple structure of a low-order vibronic coupling *ansatz* with the ANN introducing the flexibility needed to construct highly accurate PESs. Planar (5D) NO₃ was chosen as the first test case because it is certainly a tough, nontrivial problem and, second, plenty of data and experience are available. The new approach has proven to perform better in terms of overall accuracy than the previous approach which was based on inclusion of high-order polynomial expansions.^{31,32,36,74} More than one parameterization of the ANNs as defined in Ref. 73 leads to very similar rms vs the electronic *ab initio* energies. Selecting one of them as the “most physically relevant” is thus not straightforward. In the present study, we investigate the performance of these ANN based coupled PESs by computing the bound states they support. Similar to the previous paper,⁷³ planar NO₃ serves as a benchmark system. The low-lying electronic states of NO₃ are a notoriously difficult system with strong Jahn-Teller (JT) coupling.^{35,36,75–78} In this work, only planar geometries are considered, mostly because of the computational cost reduction in dealing with a 5-dimensional instead of a 6-dimensional system. This planarity constraint also allows us to determine the vibronic energies of the first excited state of NO₃ as it decouples from the other electronic states when NO₃ stays planar. The first excited state presents a strong JT effect. We investigate the method’s performance by converged accurate quantum dynamics computations of the vibrational states supported by the ²A₂’ adiabatic electronic ground and of the vibronic levels supported by the ²E’’ first excited state. Energies are compared with the ones supported by the potential energy surfaces previously published^{36,74} and based on a high-order polynomial diabatic model. A direct comparison is made possible due to the use of the same underlying *ab initio* data set. The results are also compared to the available experimental data for the two ²A₂’ and ²E’’ electronic states.

II. DIABATIC POTENTIAL MODEL

In this work, the properties of the recently developed restricted-dimensional (5D) model of the lowest PESs corresponding to the \tilde{X}^2A_2' , \tilde{A}^2E'' , and \tilde{B}^2E' state of the planar NO₃ radical are tested thoroughly. A brief summary of the crucial aspects of the presented model and diabaticization method is given below, all details regarding both being available in Ref. 73. The core idea of this method is to combine the advantages and general structure of a low-order vibronic coupling model with the accuracy of ANNs. To achieve this, we construct a minimal, symmetry-adapted polynomial model and fit its coefficients by a nonlinear least squares procedure yielding a qualitatively correct but not very accurate initial model that follows the general shape of the final surfaces. Then, in a second step, most of the obtained parameters are modified by an ANN making them coordinate dependent. This way, the polynomial model dictates the overall structure of the coupling model, while the neural network applies *corrections* to individual terms yielding significantly improved (and in fact excellent) accuracy.

In the present case, our underlying *ansatz* for the diabatic matrix is expressed as a sum of diagonal and coupling matrices,^{12,32,73}

$$\mathbf{W}^d(\mathbf{Q}) = \mathbf{W}_{\text{diag}}^d(\mathbf{Q}) + \begin{pmatrix} \lambda_1 & 0^T & 0^T \\ 0 & \lambda_2 \mathbf{1} & 0 \\ 0 & 0 & \lambda_3 \mathbf{1} \end{pmatrix} + \begin{pmatrix} 0 & 0^T & 0^T \\ 0 & \lambda_4 \boldsymbol{\varepsilon}_s + \lambda_5 \boldsymbol{\varepsilon}_b & 0 \\ 0 & 0 & \lambda_6 \boldsymbol{\varepsilon}_s + \lambda_7 \boldsymbol{\varepsilon}_b \end{pmatrix} + \begin{pmatrix} 0 & 0^T & \lambda_8 \boldsymbol{\rho}_s^T + \lambda_9 \boldsymbol{\rho}_b^T \\ 0 & 0 & 0 \\ \lambda_8 \boldsymbol{\rho}_s + \lambda_9 \boldsymbol{\rho}_b & 0 & 0 \end{pmatrix}. \quad (1)$$

The matrices $\boldsymbol{\varepsilon}_{s,b}$ are the first-order Jahn-Teller coupling blocks, and the vectors $\boldsymbol{\rho}_{s,b}$ are the pseudo-Jahn-Teller coupling blocks. These couplings are given by

$$\boldsymbol{\varepsilon}_{s,b} = \begin{pmatrix} x_{s,b} & y_{s,b} \\ y_{s,b} & -x_{s,b} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\rho}_{s,b} = \begin{pmatrix} x_{s,b} \\ -y_{s,b} \end{pmatrix}, \quad (2)$$

where symmetry-adapted coordinates corresponding to the degenerate asymmetric bending and stretching modes $x_{s,b}$, $y_{s,b}$ have been used. These coordinates, together with a totally symmetric stretch coordinate a , are collected in the nuclear coordinate vector \mathbf{Q} . They have been presented in previous work^{33,36} and are given in the Appendix for the ease of the reader. They are constructed from a set of primitive valence coordinates specifically chosen to account for the basic asymptotic behavior in the underlying low-order model. In Eq. (1), $\mathbf{W}_{\text{diag}}^d(\mathbf{Q})$ is a diagonal matrix which contains first and second order terms as detailed in previous work^{33,36,73,74} and reproduced in the Appendix. The 12 parameters appearing in this matrix are not modified by the ANN procedure. The three subsequent 5×5 matrices of Eq. (1) refer to the usual zero-order (or constant), linear Jahn-Teller and linear pseudo-Jahn-Teller coupling matrices. When the parameters $\lambda_j(\mathbf{Q})$ are kept constant (i.e., independent of \mathbf{Q}), the model resumes to the usual linear vibronic coupling model in symmetry-adapted nuclear coordinates with λ_j^0 , $j = 1, 9$ being the usual parameters of the expansion. The ANN step modifies this picture by allowing the 9 parameters to vary with \mathbf{Q} . These 9 coordinate-dependent λ_j are given by

$$\lambda_j(\mathbf{Q}) = \lambda_j^0 \cdot \left(1 + c_j \cdot \eta_j^{(f)}(\mathbf{Q})\right), \quad (3)$$

where $\eta_j^{(f)}$ are the neural network outputs and the λ_j^0 are obtained from a nonlinear least squares fit. Thus, the neural network output of a single ANN depending on the five input coordinates only provides corrections to the low-order polynomial model as a function of the coordinates, going beyond constant coefficients in the expansions. Additional scaling factors c_j are introduced to allow for further flexibility if particular λ_j^0 (read: the reference model terms) reside in a different order of magnitude than other terms and hence require a different treatment. The coupling blocks, together with their corresponding λ_j , account for ${}^2E'$ and ${}^2E''$ Jahn-Teller coupling as well as pseudo-Jahn-Teller coupling between ${}^2A'_2$ and ${}^2E'$, providing 6 independent coupling terms in total. Additionally, the constants defining the energy differences between the electronic states at the D_{3h} reference geometry are modified by the ANN (λ_{1-3}).

All neural networks tested belong to the broad category of feed-forward neural networks, that is, a function taking a vector $\boldsymbol{\eta}^{(1)}$ as input (layer) and processing it via intermediate results $\boldsymbol{\eta}^{(k)}$, the so-called hidden layers, to a final output vector $\boldsymbol{\eta}^{(f)}$ called the output layer. The vector elements $\eta_j^{(k)}$ of the k th layer are the neurons (perceptrons). Each intermediate $\boldsymbol{\eta}^{(k)}$ depends solely on the previous layer $\boldsymbol{\eta}^{(k-1)}$ by

$$\eta_j^{(k)} = f^{(k)}\left(\underbrace{\beta_j^{(k)} + \sum_l \omega_{jl}^{(k)} \eta_l^{(k-1)}}_{f^{(k)}(x_j^{(k)})}\right). \quad (4)$$

Here, $f^{(k)}$ is a function of one variable $x_j^{(k)}$ called the activation function, $x_j^{(k)}$ being a weighted sum of the values of the neurons $\eta_l^{(k-1)}$ of the previous layer with an added bias term $\beta_j^{(k)}$. In the present case, $f^{(k)}$ is chosen to be tanh for all hidden layers and as the identity Id for the output, respectively. The network architecture ultimately to turn out as optimal in our present case features a single hidden layer with 75 neurons, corresponding to a function of roughly 1100 formal parameters. Both neural networks with more or less neurons, respectively, as well as “deeper” architectures featuring more hidden layers have also been investigated. All networks have been trained using batches of 100 randomly generated initial guesses for weights and biases which are then optimized utilizing a specialized modification to a standard ANN Marquardt-Levenberg method we developed alongside the present diabaticization scheme.⁷³ The inclusion of more than one initial guess arises from the nonlinear nature of the fit. The reference model resulting from the fit of the constants $\lambda_j^{(0)}$ yields a root mean square (rms) error of 1730 cm^{-1} which is about two orders of magnitude higher than the final error after the ANN training of 38.4 cm^{-1} . More specifically, this model reproduced *ab initio* energies of the 5 adiabatic electronic states with an excellent accuracy with unweighted rms values of 22.9, 39.4, 29.3, 30.2, and 23.1 cm^{-1} , respectively, for energy ranges up to 1 eV above the energies of each state for the reference geometry point. The 10th best network of the same batch shows very similar rms values. These (and all other later considered) fits used the same data set containing roughly 90 000 adiabatic energies in total, of which 15% were withheld from the neural network fit for the sake of external validation. The necessary energy data points for performing the fit are taken from previous work⁷⁴ and were computed at the Multiconfiguration Reference Singles and Doubles Configuration Interaction (MR-SDCI) level of theory based on Complete Active Space Self-Consistent Field (CASSCF) reference wave functions using a slightly adapted correlation consistent aug-cc-pVTZ standard basis.⁷⁹ For further details regarding the *ab initio* calculations, see Refs. 36, 74, and 79–81.

At first glance, the remaining rms errors are slightly larger than what has been achieved for single, uncoupled, adiabatic PESs so far. However, the errors are similar to previous ANN representations of diabatic potential models based on *ab initio* reference data.^{70,72} Smaller rms errors were obtained for a diabatic ANN representation based on model reference data.⁷¹ One reason for this observation most probably is the complicated topography of the coupled adiabatic PESs to be reproduced. Very accurate ANN

PESs can be obtained for spectroscopic purposes in the case of a single adiabatic state single-well problem. This is not the case in the present work, and it seems indicative that the second adiabatic PES sheet, corresponding to a very pronounced triple-well PES with low dissociation threshold, shows a significantly higher rms error than the single-well ground state PES. This lower adiabatic sheet of the $\tilde{A} \ ^2E''$ excited state dominates the vibronic energy levels of the \tilde{A} state manifold (see below). A second reason most likely lies in the reference data and explains why generally smaller rms errors are obtained when analytic model PESs are used as reference data. It is a very demanding task to compute the electronic structure of excited states for which usually multiconfiguration reference methods have to be applied. In the present case, the internally contracted MR-SDCI method by Werner and Knowles has been used, which still is the state of the art.^{82,83} However, due to the internal contraction scheme, the number of contractions (variational parameters) is not constant throughout extended regions of the nuclear configuration space. Wherever the number of contractions changes significantly, small jumps in energy are observed, which can be up to about 30 cm^{-1} in our experience. This sets a natural limit to the rms error that can be achieved by the ANN diabatization without introducing artifacts by overfitting. A second aspect worth mentioning is the required number of reference data. For a 5D problem it might appear unnecessary to use 90 000 reference energies. However, since five adiabatic states are computed, this corresponds only to about 18 000 different geometries. Furthermore, nine independent matrix elements must be determined accurately using these reference data. In this light, the number of used reference data is certainly not too large but rather necessary to obtain robust results.

Finally, a short note on choosing a 5D rather than the full-dimensional 6D model is provided. First of all, NO_3 is predominantly planar and many experimental observations can be simulated without including the out-of-plane mode. Furthermore, the \tilde{A} state is decoupled for planar geometries and thus, the vibronic eigenstates can be obtained easily, which would be much more difficult in the full-dimensional and fully coupled case. Second, the performance of the ANN diabatization is tested by a large number of subsequent quantum dynamics calculations. Such calculations can be performed in 6D without problems, but doing literally hundreds to thousands of such calculations as was necessary for the present study would be prohibitively expensive without gaining more insight than by the restriction to planar 5D NO_3 . However, a full 6D ANN model will be published shortly.

III. COMPUTATIONAL DETAILS

Vibrational/vibronic energy levels have been calculated on the adiabatic ground state as well as the (for planar geometries separable) first excited state in order to benchmark the novel ANN diabatic PES model. The second excited state of $^2E'$ symmetry, however, is coupled to the $^2A_2'$ ground state by *pseudo*-JT coupling, and thus, vibronic eigenstates would not be accessible easily and are not computed in the present study. For the computation of the vibrational/vibronic energy levels, a time-independent Hermite discrete variable representation (Hermite-DVR) was used⁸⁴ in which the corresponding Hamiltonian has been diagonalized

by an exact short iterative Lanczos method. Normal coordinates from MRCI calculations at the D_{3h} point have been used, with the out-of-plane umbrella bending mode (ν_2, a_2'') being excluded for these 5D calculations. The remaining coordinates correspond to the symmetric stretch (ν_1, a_1') as well as the asymmetric stretch ($\nu_{3,x,3,y}, e'$) and asymmetric bend ($\nu_{4,x,4,y}, e'$). The kinetic energy operator is transformed into the DVR grid point basis, neglecting vibrational angular momenta, as previously established in Ref. 74.

The associated number of DVR grid points (i.e., basis functions) for each mode is 15, 17, 17, 17, and 17 for the ground state and 19, 25, 25, 25, and 25 for the computations on the two coupled surfaces of the first excited state. This yields total energies converged to better than 10^{-1} cm^{-1} , degeneracies being reproduced better than 10^{-2} cm^{-1} for all calculated levels. The obtained vibrational and vibronic eigenstates are analyzed in the same manner as in our previous studies.^{36,74,78}

Our previously developed high-order polynomial based 5×5 diabatic model⁷⁴ has been used already to compute the vibrational energy levels supported by ground electronic state in full dimensionality. The vibronic levels supported by the E'' state have also been determined using an earlier 2×2 model in Ref. 36. In order to generate usable reference data, we recomputed these data using the older PES models in the reduced 5D dimensionality considered here so that a comparison with the ANN models is possible. The MCTDH approach,^{66,67,85} together with the state average and block diagonalization scheme as described in Ref. 86, is employed to this end. As in the previous studies, the six internal curvilinear coordinates as proposed in Ref. 87 are utilized, that is $\rho^{(\text{cu})}$, $\vartheta^{(\text{cu})}$, $\phi^{(\text{cu})}$, $\theta^{(\text{cu})}$, $\phi^{(\text{cu})}$, $\chi^{(\text{cu})}$, keeping $\theta^{(\text{cu})} = \pi/2$ to impose planarity. With this choice of the coordinate system, the kinetic term is exact for D_{3h} and close to exact for the other geometries. As proposed in Ref. 87 and used in Refs. 74 and 78, the fourth order Taylor expansion of the only term not in the sum-of-product form is used here. The basis set definitions used for the MCTDH approach are given in Table I. The potential term is evaluated using the correlated discrete variable representation (CDVR).^{33,68}

TABLE I. Wave function representations given by the number of single particle functions (n), the number of Fourier points (N), and the range of the underlying box (in a.u.) for the six curvilinear coordinates constructed using mass weighted Cartesian coordinates. The two columns provide the details of the MCTDH basis for the $^2A_2'$ ground state surface⁷⁴ and for the $^2E''$ first excited state coupled surfaces.³⁶

Coord.	$^2A_2'$			$^2E''$		
	n	N	Range	n	N	Range
$\rho^{(\text{cu})}$	5	32	[628:694]	7	32	[640:800]
$\vartheta^{(\text{cu})}$	7	32	[0.845:1.055]	10	32	[0.805:1.105]
$\phi^{(\text{cu})}$	7	32	[0.655:0.915]	11	32	[0.615:0.955]
$\theta^{(\text{cu})}$			$\pi/2$			$\pi/2$
$\phi^{(\text{cu})}$	9	32	[0.777:1.377]	9	32	[0.860:1.235]
$\chi^{(\text{cu})}$	10	32	[2.641:3.541]	9	32	[2.829:3.454]
Electronic	1	1		2	2	

IV. RESULTS AND DISCUSSION

The various ANN parameterizations fitted and tested in Ref. 73 are studied systematically below. The results of that previous study suggested that an ANN model with a single hidden layer of 75 neurons would provide an optimal balance between fitting accuracy and robustness against overfitting. However, that assessment was based on fitting errors only, while the present study will focus on quantum dynamics results obtained with those ANN models. The 75 neuron model is chosen as a reference case and the other ANN models will be compared with respect to these reference results. The dynamics results of the various ANN models will also be compared with the results obtained with our previously developed purely polynomial models^{36,74,78} for both ground and ${}^2E''$ excited states. Our earlier polynomial PES models reproduce the reference data with similar rms errors but are much harder to fit. The performance of the new approach finally will be assessed by its capability of reproducing the available experimental data.^{88,89}

A. Impact of the hidden layer size

Since both the size of the input layer (coordinates) and the size of the output layer (modified coefficients) are dictated by the general model, the number of neurons in the hidden layer is the primary convergence parameter left for fitting a given *ab initio* data set. As such, it will be the primary focus of our investigations. If an ANN is viewed as a simple, parameterized function, the number of hidden layers as well as each hidden layer size determines the neural network's number of formal parameters. More specifically, in the case of a single hidden layer, each neuron exactly corresponds to a set of 15 formal parameters to be fitted additionally (9 + 5 weights, 1 bias). While the number of formal parameters increases in a more involved manner for multiple hidden layers, we will focus solely on the single layer case in this section. The central question here is whether or to which degree increasing/decreasing the number of hidden layer neurons (thus, fitting parameters) significantly influences computed energy levels. Therefore, in this section, we

are not particularly interested in the energy levels themselves, but rather ask how much equivalent levels deviate from the reference model.

In order to properly quantify what a *significant influence* on the computed energy levels is, let us first consider otherwise identical neural networks that differ only in the explicit values of their weights and biases and marginally in terms of fitting error. Due to the nonlinearity of ANNs, many fits with different initial guesses are performed and ranked according to the resulting fitting error. For the NO_3 system, we found that the fitting errors of the best 10 obtained parameterizations hardly differ (41 cm^{-1} vs 38 cm^{-1}). Therefore, we compare the ten best out of 100 fitted ANNs including the 75 neuron reference network. They are hereafter referenced to by their "rank," rank n meaning the n th best model when considering the fitting error. The computed vibrational excitation energies on the adiabatic electronic ground state PES are presented in Table II for the best ANN.

The computed vibrational excitation energies of the competing networks (rank 2–10) are provided in terms of deviations from the best network. The two last columns of Table II provide the average excitation energies over the 10 ANNs, as well as the standard deviation. Considering this set of competing networks, one finds that the produced energies differ within a few cm^{-1} , the standard deviation over all states being roughly 2 cm^{-1} . The values of Table II are reproduced in Fig. 1 in order to better visualize how the results of the different networks are spread around the average energies. It is observed that, considering that these 10 network parameterizations are of very similar quality in terms of fitting error, the deviations among the vibrational level energies within the given spread are intrinsic to the model. This means that they are produced by the flexibility of the ANN approach to reproduce the *ab initio* reference data. In this sense, we consider all changes to the model that only produce deviations within this range to be insignificant as they become *a priori* indistinguishable from competing networks of the same architecture as our reference. Thus, from here on, we will denote the excitation energies with a $\pm\sigma_i$ to

TABLE II. Comparison of computed transition energies for the \bar{X}^2A_2' state of ${}^{14}\text{NO}_3$ (in cm^{-1}) for the 10 best (in terms of fitting error) ANNs from a set of 100 fitted neural networks, including the reference ANN. Standard deviation is provided with respect to average energy levels μ_i .

State	Best	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	μ_i	σ_i
	$E_i - E_0$	Deviations from best network									Mean	Std. dev.
1/2	366.8	3.9	1.8	4.1	3.1	1.5	4.6	3.7	1.8	3.7	364.0	1.5
3	756.3	3.9	0.5	4.3	-0.6	-0.8	3.9	1.2	-3.2	3.5	755.1	2.5
4/5	774.6	6.1	2.4	3.8	3.6	1.2	6.1	5.6	0.8	2.9	771.3	2.2
6/7	1039.7	3.4	-2.5	-1.3	-0.8	1.8	3.3	-2.7	3.0	0.2	1039.3	2.3
8	1054.9	0.2	-0.2	-0.7	0.0	-0.4	0.6	-1.1	1.0	-1.0	1055.1	0.7
9/10	1179.3	3.7	1.4	3.6	-2.6	-2.3	3.0	-0.7	-5.0	-0.4	1179.2	2.9
11	1187.7	5.6	2.1	2.4	1.8	-0.7	5.6	4.4	-4.7	2.2	1185.8	3.1
12	1215.1	8.7	4.4	4.2	5.6	1.7	8.5	8.0	0.8	0.7	1210.9	3.4
13	1344.9	3.7	-3.2	0.6	-1.3	1.4	5.0	-3.1	-0.3	0.4	1344.5	2.6
14/15	1417.7	2.9	0.5	2.1	1.9	0.5	4.1	1.0	1.7	1.8	1416.1	1.2
16/17	1490.0	4.1	-1.0	1.0	1.9	1.7	4.3	-2.0	0.2	1.3	1488.9	2.0
18	1498.2	8.7	0.5	-0.1	5.3	3.2	9.8	0.9	7.5	1.9	1494.4	3.8

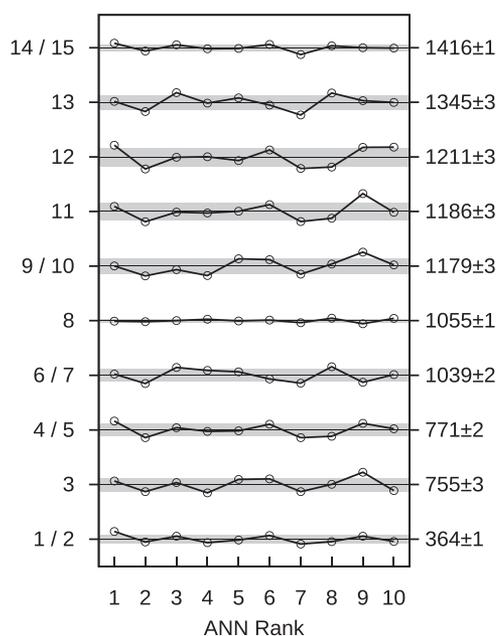


FIG. 1. Visual representation of the first 15 states given in Table II, showing deviations of the 10 best networks from respective average state energies. Level distances, not to scale, correspond to a deviation of 20 cm^{-1} from average. Gray: Standard deviations σ_i around average energies. Average energies are provided on the right-hand side together with the deviation σ_i in cm^{-1} .

emphasize this point. While individual networks deviate more from the average than others (e.g., network 7), the overall spread achieved remains around or below 10 cm^{-1} . This indicates that the method does not produce a single exceptional network among the number

of tested reasonable parameterizations. Instead, the method appears to produce several network parameterizations yielding consistent results under present conditions. Thus, the method appears to work robustly.

Deviations in excitation energies for a representative subset of hidden layer sizes are summarized in Table III, which can now be interpreted within the limits of our estimated intrinsic errors. The 40 neuron network, having almost half the number of formal parameters, deviates significantly for individual excitations (in particular 4/5, 11, and 12), causing energy levels to be significantly spread around the reference computation. This is intuitively understandable as the significant loss of flexibility compared to the reference model eventually limits the quality of the surface in areas relevant for the quantum dynamics. The 60 neuron network shows similar features although they are less pronounced as the network is significantly more flexible. In general, states 1/2, 8, and 14/15, which already proved to be extremely robust with respect to the network choice within the same set, show very low deviation among reasonable network sizes. Increasing or decreasing network size by 5 neurons (that is 70 and 80 neurons) largely leaves the overall quality of the model unchanged, with the ANN being marginally more susceptible to a decrease in parameters than to an increase. Similarly, increasing the number of neurons to 100 (over 30% more parameters) yields no energy shift that is distinguishable from noise caused by choosing a different 75 neuron network.

The vibronic levels supported by the ${}^2E''$ first excited state of the radical are an even more stringent test of the robustness of the proposed ANN parameterizations of coupled surfaces as the coupling terms are directly used in the dynamics computations. Vibronic energy levels for the reference coupled surfaces (single hidden layer with 75 neurons) as well as deviations obtained among the ten best networks of this architecture are summarized in Table IV in analogy to Table II. The data are reproduced, furthermore, in Fig. 2

TABLE III. Comparison of computed transition energies for the $\bar{X}^2A'_2$ state of ${}^{14}\text{NO}_3$ (in cm^{-1}) for different hidden layer sizes. The vibrational excitation energies of the reference (75) surface are reported together with the deviation σ_i as defined before (see text). For the five hidden layer sizes tested, $\delta(E_i - E_0)$ provides level shifts compared to the reference network (75). The dotted vertical line materializes the separation between increased and decreased hidden layer sizes with respect to the reference.

	75	40	60	70	80	100	
State	$E_i - E_0 \pm \sigma_i$	$\delta(E_i - E_0)$					
1/2	366.8 ± 1.5	6.2	7.8	1.2	∴	3.0	1.5
3	756.3 ± 2.5	5.2	5.8	-5.1	∴	3.0	2.6
4/5	774.6 ± 2.2	13.6	10.7	3.1	∴	4.9	2.9
6/7	1039.7 ± 2.3	-0.6	0.7	3.9	∴	-4.1	1.6
8	1054.9 ± 0.7	0.6	0.2	0.0	∴	-0.4	0.0
9/10	1179.3 ± 2.9	9.0	5.9	-6.1	∴	2.8	3.3
11	1187.7 ± 3.1	15.3	7.1	-1.0	∴	3.6	3.0
12	1215.1 ± 3.4	21.9	13.3	7.6	∴	7.8	3.1
13	1344.9 ± 2.6	7.0	5.3	-1.7	∴	-4.3	0.1
14/15	1417.7 ± 1.2	4.7	6.2	-0.8	∴	0.8	1.0
16/17	1490.0 ± 2.0	-0.1	0.8	1.4	∴	-0.4	2.6
18	1498.2 ± 3.8	6.0	4.4	9.6	∴	1.7	6.0

TABLE IV. Comparison of computed transition energies for $^{14}\text{NO}_3$ (in cm^{-1}) on the \tilde{A} state for the 10 best (in terms of fitting error) ANNs in analogy to Table II. All states are given labels in terms of excitations in the breathing mode (a) as well as radial and tangential excitations in the degenerate stretch and bend modes ($r_{s,b}$, $t_{s,b}$). Assignments marked with (*) display visible contamination of other r_s excitation(s). States to which no meaningful label could be assigned are marked with ?

State	Assignment	Best $E_i - E_0$	Deviations from best network										μ_i Mean	σ_i Std. dev.
			2nd	3rd	4th	5th	6th	7th	8th	9th	10th			
2	$0 a_1''$	56.6	-2.5	3.2	2.9	3.4	0.6	0.9	3.5	6.3	-1.8	55.0	2.7	
3/4	$r_s^1 e''$	516.2	-15.3	26.0	4.3	1.2	-14.0	-6.5	0.8	18.8	2.2	514.5	12.9	
5	$t_b^1 a_2''$	539.3	-5.6	-4.0	-3.2	-7.2	-9.7	-5.9	-0.2	-8.0	-2.2	543.9	3.3	
6	$r_s^1 a_1''$	576.4	-18.4	22.8	-0.4	-2.0	-11.8	-6.5	-2.9	17.1	-0.4	576.7	12.2	
7/8	$t_b^1 e''$	604.7	-14.5	2.8	0.9	-3.3	-10.5	-5.3	2.6	0.1	-2.7	607.7	5.7	
9/10	$r_b^1 e''$	790.0	-1.6	16.8	-3.4	14.2	10.5	-4.3	7.4	11.1	-10.6	786.0	9.2	
11	$r_b^1 (*) a_1''$	833.9	-17.8	39.1	2.6	16.1	-3.7	-3.7	4.6	19.0	-10.6	829.3	16.4	
12/13	$r_s^2 e''$	1013.4	-38.3	59.3	1.7	-3.2	-30.6	-7.1	-5.5	35.5	-6.4	1012.9	28.5	
14	$r_s^1 t_b^1 a_2''$	1024.8	-19.5	21.5	2.8	-2.1	-18.8	-10.1	2.3	15.9	4.8	1025.2	13.3	
15	$r_s^2 a_1''$	1084.4	-35.9	48.8	-2.3	-3.9	-20.2	-10.5	-3.9	31.2	-8.3	1084.9	24.2	
16/17	$t_b^2 (*) e''$	1100.1	-27.4	15.9	-2.8	-11.4	-21.5	-8.7	-3.4	5.4	-4.5	1105.9	12.5	
18/19	$r_s^1 t_b^1 (*) e''$	1151.2	-29.7	7.9	0.1	-8.5	-18.3	-8.5	0.9	1.4	-1.0	1156.8	11.1	
20	$t_b^2 a_1''$	1161.8	-24.4	0.7	0.1	-9.5	-16.8	-6.3	2.9	-6.0	-3.6	1168.1	8.6	
21/22	$a^1 e''$	1179.8	6.3	8.6	5.0	1.2	-2.6	4.3	0.5	4.1	4.0	1176.7	3.3	
23	$t_s^1 a_2''$	1276.1	-0.6	0.2	-7.6	-4.3	-4.3	-11.2	-3.6	-5.0	1.4	1279.6	3.9	
24	$a^1 (*) a_1''$	1291.4	-0.1	51.6	16.5	17.5	-1.3	16.0	5.9	21.9	-4.2	1279.1	16.7	
25/26	$r_b^1 r_s^1 e''$	1302.8	-25.3	54.3	0.0	19.1	-4.0	-0.2	5.6	22.1	-7.6	1296.4	21.4	
27/28	? e''	1357.4	-19.4	25.2	-1.7	4.1	-5.1	-7.0	4.6	9.8	-11.4	1357.5	12.3	

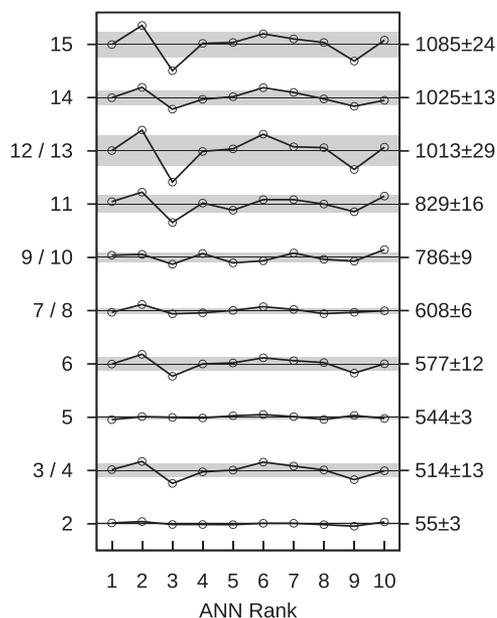


FIG. 2. Visual representation of the first 15 states of Table IV, showing deviations of the 10 best networks from respective average state energies. Level distances, not to scale, correspond to a deviation of 100 cm^{-1} from average. Gray: Standard deviation σ_i around average energies. Excitation energies together with the deviation σ_i are given at the right-hand side.

for better visualization. In addition to the energies, Table IV also provides D_{3h} labels and state assignments based on symmetry information and projected wave function densities. State assignments in the degenerate stretch and bend coordinates are given in terms of radial $r_{s,b}^n$ and tangential excitations $t_{s,b}^n$,⁹⁰ while excitations in the totally symmetric mode are denoted with a^n . In contrast to ground state calculations, the “spread” of individual energy levels is between 1 and 30 cm^{-1} , the standard deviation over all vibronic states being roughly 13 cm^{-1} , with few states showing individual deviations up to around 50 cm^{-1} . At first glance, such an increase in error may be fully attributed to the fact that the lower adiabatic sheet has been found to be a very dominant contribution to the fitting error, making it the most difficult part for the ANN to fit. In addition to this factor, the deviations are not evenly distributed among the vibronic levels but well-structured. More specifically, the ground state tunneling first excitation of 57 cm^{-1} is robustly reproduced by the 10 parameterizations. Similarly, levels without noticeable contributions of r_s excitations show lower deviations than the ones with considerable r_s excitations. This means that, while the triple well structure is overall well-described for dynamics, the accuracy of the surface diminishes for larger displacements in the asymmetric stretch. The reason for this observation is not clear but could be due to the sampling of the reference data.

In analogy to the vibrational levels supported by the ground electronic states (see Table III), the effect of the number of hidden neurons used in the ANN is tested for the vibronic levels of the $^2E''$ coupled surfaces and the results are summarized in Table V.

TABLE V. Comparison of computed transition energies for $^{14}\text{NO}_3$ (in cm^{-1}) on the \bar{A} state for different hidden layer sizes. $\delta(E_i - E_0)$ provides level shifts compared to the reference network (75). Assignments are identical to that of Table IV. The dotted vertical line materializes the separation between increased and decreased hidden layer sizes with respect to the reference.

State	Assignment	75	40	60	70	80	100
		$E_i - E_0$	$\delta(E_i - E_0)$				
2	$0 a_1''$	56.6 ± 2.7	-0.1	-1.2	-0.9	1.8	5.6
3/4	$r_s^1 e''$	516.2 ± 12.9	-16.4	-15.6	-10.9	7.4	30.1
5	$t_b^1 a_2''$	539.3 ± 3.3	-8.8	-4.3	-8.9	4.1	-9.2
6	$r_s^1 a_1''$	576.4 ± 12.2	-12.1	-13.3	-11.6	4.1	25.7
7/8	$t_b^1 e''$	604.7 ± 5.7	-8.3	-11.5	-13.7	6.1	-0.2
9/10	$r_b^1 e''$	790.0 ± 9.2	-27.7	10.0	9.0	25.7	22.1
11	$r_b^1 (*) a_1''$	833.9 ± 16.4	-55.6	-6.2	-2.9	31.4	42.4
12/13	$r_s^2 e''$	1013.4 ± 28.5	-47.1	-31.8	-27.2	17.1	60.6
14	$r_s^1 t_b^1 a_2''$	1024.8 ± 13.3	-21.4	-17.6	-16.4	14.6	23.4
15	$r_s^2 a_1''$	1084.4 ± 24.2	-32.9	-23.9	-24.2	13.4	50.1
16/17	$t_b^2 (*) e''$	1100.1 ± 12.5	-26.0	-17.7	-22.1	10.4	13.1
18/19	$r_s^1 t_b^1 (*) e''$	1151.2 ± 11.1	-13.5	-20.2	-22.6	12.1	4.0
20	$t_b^2 a_1''$	1161.8 ± 8.6	-13.1	-15.8	-21.6	11.7	-4.3
21/22	$a^1 e''$	1179.8 ± 3.3	-3.9	1.3	1.4	7.7	6.4
23	$t_s^1 a_2''$	1276.1 ± 3.9	-4.5	2.2	1.5	4.1	0.0
24	$a^1 (*) a_1''$	1291.4 ± 16.7	-44.9	2.3	7.0	42.8	44.2
25/26	$r_b^1 r_s^1 e''$	1302.8 ± 21.4	-62.2	-5.2	-3.6	47.3	52.5
27/28	? e''	1357.4 ± 12.3	-39.0	-4.6	-6.8	29.6	27.2

In contrast to ground state calculations, we find the 40 neuron network to just marginally deviate for most r_s^0 states such as 2, 5, and 7/8, while producing a significant difference for the r_b fundamental. Overall, agreement increases for increased network size up to the 80 and 100 neuron network, where, for example, state 9/10 suddenly shows an *increase* in deviation. The latter may actually be an overfitting effect as both the dynamics on the excited state and the fitting error are dominated by the lower adiabatic sheet and both begin to show discrepancies past the 75 neuron mark (one in the form of yielding different energy levels and the other in the form of being less consistent with the validation set).

In conclusion, our findings regarding the behavior of the model with respect to the hidden layer size coincide with previous conclusions drawn from fitting performance alone, that is, while reducing the number of hidden layer neurons has a significant influence on ground state dynamics, increasing the network size has an unclear effect. Furthermore, the dynamics on the first excited electronic state produce results analogous to previous findings, suggesting that networks significantly greater than the reference reproduce the *ab initio* data less consistently. As a consequence, the 75 neuron model remains the ideal candidate for further testing. In addition, the vibronic level analysis also suggests that this instability may result from lacking data in certain regions of the NCS.

B. Comparison of different ANN (and fitting) setups

Up to this point, we have only investigated the impact of the network size (i.e., the number of formal parameters of the model) with a single hidden layer architecture on the ground and excited state dynamics of the resulting surfaces. We have found that our initial reference of 75 neurons produces ground state excitations reliably and consistently across different parameterizations arising from different initial guesses. In the following, we give a brief summary of the impact of two other factors independent of the number of formal parameters of the fitting function. These tests are in analogy to the study discussed in Ref. 73 based on fitting performance. For the following two cases, we will focus on ground state dynamics as they give easier insight into the surfaces' overall quality without sampling effects.

*Early stopping*⁹¹ as used in the present work is a means of avoiding overfitting by imposing an additional convergence criterion. For this, we reserved about 15% of all *ab initio* data for an external validation set for which an independent fitting error is computed at each iteration. If this validation error does not improve along with the fitting error for three times in a row, the last parameter set to improve the validation error is considered converged. The validation set itself is not used for the training of the ANN. In order to investigate potential downsides of this method, the identical set of 100 initial guesses that produced the 75 neuron reference ANN is

refitted with early stopping disabled and a maximum number of 1000 Marquardt-Levenberg iterations. The deviations produced by this network, referred to as network 75* for brevity, are listed in Table VI alongside another networks discussed below. Deviations found remain indistinguishable from differences between competing networks within the same fitting set, meaning that while early stopping (in our present case) shows no clear advantage beyond serving as an algorithmic safeguard and speeding up the training, no discernible disadvantages of using it have been found.

One further aspect of relevance besides the number of free parameters is the actual ANN architecture. As we have noted previously, each *neuron* depends on a small set of formal parameters to fit, if we consider ANNs merely as fitting functions of a particular analytical form. This notion becomes marginally more involved when considering networks with more than one hidden layer. This is due to the fact that the number of formal parameters is roughly equal to the number of *connections* between individual nodes in the network. Hence, an ANN with 50 hidden layer neurons, one input and one output, will have a different number of formal parameters depending on whether it is one hidden layer with 50 neurons (100 connections, 50 in and 50 out) or two hidden layers with 25 each (675 connections, 25 in, 625 intermediate, 25 out). This makes comparing different architectures particularly difficult. Hence, for our practical purposes, we have chosen to take our best working model (75 neurons) and produce multilayered networks with close to the same number of formal parameters. For the sake of simplicity, we further chose the number of neurons to be the same for each hidden layer, keeping all other technical details of the fit the same. Of all network topographies considered, the only one that even remotely compared in terms of fitting error to the single layer case was a 5–20–20–20–9 network (see Ref. 73). Other tested topographies include two and four hidden layers. Considering again deviations of ground state excitation energies, Table VI shows that the

TABLE VI. Comparison of computed transition energies for $^{14}\text{NO}_3$ (in cm^{-1}) on the \tilde{X} state for various networks (see Sec. IV B). $\delta(E_i - E_0)$ provides level shifts compared to reference network (75). Networks fitted without early stopping are denoted with an asterisk (*). Networks fitted with more than one hidden layer list neuron numbers separated by –.

State	75	75*	5–20–20–20–9
	$E_i - E_0$	$\delta(E_i - E_0)$	$\delta(E_i - E_0)$
1/2	366.8 ± 1.5	0.6	2.6
3	756.3 ± 2.5	0.1	2.9
4/5	774.6 ± 2.2	2.2	5.6
6/7	1039.7 ± 2.3	0.2	2.7
8	1054.9 ± 0.7	0.2	–1.0
9/10	1179.3 ± 2.9	1.0	5.9
11	1187.7 ± 3.1	1.8	2.3
12	1215.1 ± 3.4	3.4	9.4
13	1344.9 ± 2.6	–1.4	0.5
14/15	1417.7 ± 1.2	0.0	0.1
16/17	1490.0 ± 2.0	–0.1	–0.1
18	1498.2 ± 3.8	3.2	7.3

5–20–20–20–9 network, while “noisier” than, for example, the 75* network discussed above, shows deviations of the same order of magnitude as the scattering of values among the best ten parameterizations of the 75 neuron reference model. While a more in-depth investigation could provide further insight into the intricacies of the effects more sophisticated network topographies could have, our current findings suggest that the added value might be marginal, while the sheer combinatorial effort required would be immense. We thus conclude that a single-layer topography might be the most robust and recommendable choice.

C. Comparison with experimental and previous theoretical data

So far we investigated the impact of the ANN design parameters on the quantum dynamics results by comparison with a reference ANN model. Special attention has been paid to how the number of hidden layer neurons influences the quality of the resulting surface. These comparisons aim at estimating the intrinsic error of the model with respect to differences in excitation energies produced by ANNs of (in terms of fitting error) indistinguishable quality. In the following, the focus will be on the quality of the surfaces based on external consistency with experimental data, considering \tilde{X} and \tilde{A} state transition energies.

1. Vibrational eigenstates in the \tilde{X}^2A_2' electronic state

The results of the vibrational eigenstates of the \tilde{X} state are given in Table VII in which comparisons between experimental frequencies and computed transitions for both the novel ANN model and the previous purely polynomial *ansatz* are listed. The MCTDH computations with the earlier PES models⁷⁴ are repeated in reduced (5D) dimensionality excluding the umbrella mode in order to evaluate the effect of the reduced dimensionality. As apparent from the third and fourth column of Table VII, the differences between the original 6D and the new 5D calculation (with the previous PES model) is rather small with a maximum absolute deviation of 3.5 cm^{-1} . In previous work, we also ensured that the deviations between the DVR method and the MCTDH calculations are very small. This was confirmed again computing the vibrational and vibronic states on the reference ANN PESs with both MCTDH and DVR. Thus, the 5D ANN results of the DVR calculations, also used to check for internal consistency (see above), are estimated to be within about 5 cm^{-1} of corresponding 6D MCTDH calculations once the full 6D ANN model is available. In the following, the focus will be on the comparison of the 5D data with available experiments and the improvements over the previous polynomial model (Ref. 74).

We report a striking decrease in the overall deviation from experimental data by about an order of magnitude for almost all states compared to Ref. 74, which is easily seen from the last two columns in Table VII. All assignments remain the same as for the corresponding state(s) in the polynomial model calculation. As in the case of the polynomial model,⁷⁴ the long-debated 3^1 fundamental can be assigned confidently to a state computed at 1039.7 cm^{-1} (previously 1021.8 cm^{-1} in 6D and 1025.3 cm^{-1} in 5D) with significantly improved agreement to experiment. This assignment, now at higher energy, agrees well with other theoretical treatments^{92–95} and recent experimental assignments.^{88,96,97} The second state of

TABLE VII. Summary of the comparison between experimental and computed transition energies for $^{14}\text{NO}_3$ (in cm^{-1}). The “assignment” column provides the corresponding D_{3h} labels as well as state assignments including dominant state contributions. $E_{\text{theo}}^{6\text{D},5\text{D}}$ provide excitation energies using the previous (polynomial) model⁷⁴ and $E_{\text{theo}}^{\text{ANN}}$ the ones obtained for the present ANN based model.⁷³ The remaining two columns provide respective differences from experimental data as listed under E_{exp} reported from Ref. 88. See text for discussion.

Assignment	E_{exp}	$E_{\text{theo}}^{6\text{D}}$	$E_{\text{theo}}^{5\text{D}}$	$E_{\text{theo}}^{\text{ANN}}$	$\delta E_{\text{exp}}^{5\text{D}}$	$\delta E_{\text{exp}}^{\text{ANN}}$
		Polynomial model ⁷⁴				
$4^1 e'$	365.5	361.1	361.1	366.8	4.3	−1.3
$4^2 a'_1$	752.4	711.0	711.1	756.3	41.3	−3.9
$4^2 e'$	771.8	742.2	742.3	774.6	29.5	−2.8
$1^1 a'_1$	1051.2	1038.6	1040.2	1054.9	11.0	−3.7
$3^1 e'$	1055.3	1021.8	1025.3	1039.7	30.0	15.6
$4^3 e'$	1173.6	1082.5	1082.7	1179.3	90.9	−5.7
$4^3 a'_1$	1214	1139.7	1140.1	1215.1	73.9	−1.1
$1^1 4^1 e'$	1413.6	1388.1	1390.0	1417.7	23.6	−4.1
$3^1 4^1 a'_2$	1491	1302.4	1305.9	1344.9	185.1	146.1
$3^1 4^1 e'$	1492.4	1438.6	1441.9	1490.0	50.5	2.4
$3^1 4^1 a'_1$	1499.8	1498.2	52.2	1.6

particular interest is the combination mode $3^1 4^1$, giving rise to three sublevels of e' , a'_1 , and a'_2 symmetry, respectively. Much like the fundamental mode, the $3^1 4^1 e'$ transition now computed at 1490.0 cm^{-1} displays significantly improved agreement with the experimental value at 1492.4 cm^{-1} with a deviation of only 2.4 cm^{-1} . This gives further strong evidence that this energy level does *not* correspond to the 3^1 fundamental. The effect of the 5D approximation to the 6D dynamics using the polynomial model on this transition is found to be less than 4 cm^{-1} (1438.6 cm^{-1} vs 1442.0 cm^{-1}). Similarly, the $3^1 4^1 a'_1$ state, to which no state could be assigned using the polynomial model, now can be assigned with certainty to an energy level at 1498.2 cm^{-1} . Indeed, the harmonic oscillator basis function of the DVR calculations corresponding to the $3^1 4^1 a'_1$ state accounts for over 70% of the total vibrational wave function. Finally, the $3^1 4^1 a'_2$ state, computed at an energy level of 1344.9 cm^{-1} , shows significant deviation from the experimental assignment to a level at 1491 cm^{-1} . This disagreement of more than 150 cm^{-1} is one order of magnitude larger than the second largest deviation (about 16 cm^{-1}) and over 25 times greater than the third largest deviation (about 6 cm^{-1}). Considering, however, that excitations even higher in energy, namely, $3^1 4^1 e'$ and a'_1 , do not deviate more than a few cm^{-1} from the experimental values suggests that this deviation is more likely to be due to an incorrect experimental assignment. It is worth noting that deviations between the new ANN results and the experimental assignments for most states reside well within deviations between different ANNs from the same fitting set as described in Sec. IV A. The impact of the ANN size on the quantum dynamics results is put into perspective with respect to experimental values in Table VIII. We find larger hidden layer sizes to have no significant effect on the agreement with experimental data when compared with the competing reference network of 75 hidden neurons. The same is observed when comparing with the 75^* ANN results.

The agreement of the present results with experimental data shows that the PES model must be very accurate and certainly must be qualitatively correct. This allows us to shine some light on another controversy in the literature. The equilibrium geometry of NO_3 in its electronic ground state has been subject to debate since the late 1970s (see Ref. 80 for an extensive discussion). Electronic structure calculations mostly yield a C_{2v} equilibrium geometry, but it was shown by one of the authors that this is an artifact due to artificial symmetry breaking of the electronic wave function caused by the use of single reference wave functions.⁸⁰ The adiabatic ground state PES resulting from the present diabatic model clearly has a single PES minimum of D_{3h} symmetry, thus supporting a D_{3h} equilibrium geometry. With this strong evidence from the current dynamics calculations using this PES model and the excellent agreement obtained with respect to experimental spectroscopy data, it is nearly unimaginable that a distorted C_{2v} equilibrium structure is more than an artifact.

2. Vibronic eigenstates in the $\tilde{A}^2 E''$ electronic state

For the electronically excited $^2 E''$ state, there are only a total of 3 experimentally measured transitions to compare with as all other observed transitions involve umbrella excitations and cannot be computed with the present ANN model. The corresponding results are gathered in Table IX together with their assignments. Due to the strength of the Jahn-Teller coupling in this system and the resulting triple-well structure on the lower diabatic sheet, the assignment of the D_{3h} labels based on normal mode excitations is not as straightforward compared to the \tilde{X} state labeling. The assignments used in Sec. IV A in terms of radial and tangential excitations provide deeper insights regarding the limits of accuracy of the model as states corresponding to r_s excitations have been found to be more limited in terms of accuracy. Further explanation as to how tunneling and geometric phase effects influence both the complexity of the

TABLE VIII. Comparison of absolute differences between experimental and computed (ground state) transition energies for $^{14}\text{NO}_3$ (in cm^{-1}) for different hidden layer sizes. Assignments are identical to that of Table VII. $|\delta E_{\text{exp}}|$ corresponds to $\delta E_{\text{exp}}^{\text{ANN}}$ from the same table. $\delta(|\delta E_{\text{exp}}|)$ provides decreases (negative) or increases (positive) in deviation from experiment compared to the first column. The third column provides data for the 10th percentile instead of the (in terms of fitting error) best fitted ANN from a set of 100 fitted neural networks, including the reference ANN. Bottom rows: average increase/decrease in deviation over all states and root mean square variations around the reference surface, disregarding the outlier state $3^1 4^1 a_2'$; see Sec. IV C 1.

	75	75 (10th best)	40	70	80	100
Assignment	$ \delta E_{\text{exp}} $	$\delta(\delta E_{\text{exp}})$				
$4^1 e'$	1.3	1.0	3.6	-1.2	0.3	-1.2
$4^2 a_1'$	3.9	-3.5	-2.6	5.1	-3.0	-2.6
$4^2 e'$	2.8	-2.7	8.0	-2.6	-0.8	-2.7
$1^1 a_1'$	15.6	0.2	-0.6	3.9	-4.1	1.6
$3^1 e'$	3.7	1.0	-0.6	0.0	0.4	0.0
$4^3 e'$	5.7	0.4	-2.4	6.1	-2.8	-3.3
$4^3 a_1'$	1.1	-0.7	19.6	5.4	5.5	0.8
$1^1 4^1 e'$	4.1	-1.8	-3.5	0.8	-0.8	-1.0
$3^1 4^1 a_2'$	146.1	0.4	7.0	-1.7	-4.3	0.1
$3^1 4^1 e'$	2.4	1.3	-0.1	1.4	-0.4	2.6
$3^1 4^1 a_1'$	1.6	1.9	6.0	9.6	1.7	6.0
Mean increase		-0.3	2.7	2.9	-0.4	0.0
Variance		1.7	7.2	4.4	2.8	2.6

resulting spectra and the structure of the vibronic eigenstates can be found in greater detail in Refs. 78 and 90. The deviations between experimental assignments and both the new ANN model and the prior, purely polynomial *ansatz* are given in Table IX. Of the three states for which experimental data are available, all show significant improvement in terms of reproducing experimental data over the previous polynomial model. Both computed purely tangential excitations reproduce experimental data up to a few cm^{-1} , while the r_s excited state shows a larger deviation. This means that the relative improvement of the $^2E''$ model is consistent with that of the ground state, with dissociative motions being a limiting factor on the excited state surface.

3. Numerical stability and potential artifacts

In Secs. IV C 1 and IV C 2, we outlined how our present ANN scheme significantly outperforms the previous purely polynomial-based *ansatz*. One of the reasons for *why* the neural network performs better than the polynomial model certainly is the more flexible functional form of the ANN compared to the previously used polynomial approach. To achieve higher flexibility with the polynomial *ansatz*, one would have to increase the polynomial order considerably. However, it is a fundamental property of polynomials that as the order increases, the function starts oscillating rapidly or getting unbound in areas not well sampled. This unphysical behavior

TABLE IX. Summary of the comparison between experimental and computed transition energies on the \tilde{A}^2E'' state of NO_3 (in cm^{-1}). Assignments in the first column correspond to those of Table VI of Ref. 89, with assignments from Table V given in parentheses. $E_{\text{theo}}^{6\text{D},5\text{D}}$ provide excitation energies using the previous (polynomial) model,^{36,78} the $E_{\text{theo}}^{6\text{D}}$ being reproduced from Ref. 78 and $E_{\text{theo}}^{\text{ANN}}$ the ones obtained for the present ANN based model.⁷³ The remaining two columns provide respective differences from experimental data.

Assignment	E_{exp}	$E_{\text{theo}}^{6\text{D}}$	$E_{\text{theo}}^{5\text{D}}$	$E_{\text{theo}}^{\text{ANN}}$	$\delta E_{\text{exp}}^{5\text{D}}$	$\delta E_{\text{exp}}^{\text{ANN}}$
		Polynomial model ^{36,78}				
$4_0^1 (t_b^1)$	539.5	549.1	544.9	539.3	-5.4	0.2
$4_0^2 (r_s^1 t_b^1)$	1056.6	1136.9	1142.2	1024.8	-85.6	31.8
$3_0^1 (t_s^1)$	1270.5	1300.1	1291.5	1276.1	-21.0	-5.6

becomes a problem for dynamics treatment even for physically irrelevant nuclear configurations when the oscillation tends to negative infinity. Such regions cause the wave packet to get trapped in these unphysical regions. However, since these artifacts only occur where little to no data are sampled, they only become apparent when the wave packet dynamics on the PESs are investigated, meaning that there is no structured, *a priori* method to predict or detect such regions in a high dimensional nuclear configuration space. Therefore, the development of high-dimensional coupled PESs, especially when using some sort of polynomial model, is usually plagued by the occurrence of such artifacts.

However, the general structure of the ANN approach tested in this work shows a crucial difference. Unlike polynomial terms of increasing order, the functional shape of hidden layer neurons is *bounded* as they are composed of sigmoid functions depending on weighted sums of coordinates. Furthermore, all hidden layer neurons are *homogeneous* in the sense that increasing the number of neurons always adds a term of the same functional form as the previous ones, merely differently parameterized, in stark contrast to adding a higher-order polynomial term to a function. As a consequence, none of the investigated ANNs produced the artifacts discussed above in any of the dynamics calculations performed here. This suggests that all ANN surfaces produced by our current method, even if insufficiently parameterized to describe the PESs well (e.g., the 40 neuron network) or potentially overfitted (e.g., the 100 neuron network) are well-behaved and free of severe oscillations or unbound areas comparable to those commonly occurring in polynomial fits. This is a property of the new ANN approach of invaluable importance for future applications.

V. CONCLUSIONS AND OUTLOOK

In the present study, we investigate the viability of our recently developed diabaticization method as a means to produce coupled PESs for reliable high-dimensional quantum dynamics calculations. This method, combining vibronic coupling models with artificial neural networks (ANNs), has been tested using the NO_3 radical as a benchmark system of exceptional complexity. For the purpose of this work, the molecule has been restricted to planar geometries, which will be extended to a full 6D space in future work. Using a time-independent DVR method, the behavior of vibrational eigenstates on the ${}^2A'_2$ ground state as well as vibronic eigenstates on the first excited ${}^2E''$ state has been investigated depending on the number of hidden layer neurons (corresponding to the number of formal parameters) and to a lesser extent depending on the number of hidden layers (conserving the number of formal parameters). The various resulting surfaces, exclusively differing in the underlying ANN, have been compared among one another as well as with previous efforts using a purely polynomial *ansatz*. To this end, the MCTDH approach has been used to evaluate the effect of the planar geometry restriction. The ANN results are also compared with available experimental data. The investigated approach utilizes the basic structure (and simplicity) of low-order vibronic coupling models and achieves the high accuracy needed for reliable dynamics simulations by the neural network. Despite the highly nonlinear nature of the fitting procedure, the presented method produces PESs of consistent quality for a large number of initial guesses, with deviations

of fitting errors as well as excitation frequencies of a few cm^{-1} for the best 10 out of 100 fitted networks for a fixed but reasonably large number of hidden layer neurons. Similarly, despite being restricted to planar geometries, the model reproduces measured transition energies (excluding umbrella mode excitations) with unprecedented accuracy, with deviations of only a few cm^{-1} . It therefore is particularly noteworthy that for the vast majority of measured excitations, deviations between best ANN results and experiment are of the same order of magnitude as deviations among ANN results of similar fitting quality but different parameterizations (different initial guesses). While deviations on the excited ${}^2E''$ state can be significantly larger (10 cm^{-1} range), they remain well-structured. The larger deviations correlate with radial excitations in the asymmetric stretching mode (r_s), which might indicate insufficient data in the dissociative motions. Overall, when compared to experimental data, the surface model based on our new approach produces vibrational/vibronic excitation frequencies that are an order of magnitude better than our previous polynomial model. This allows us to assign debated ground-state transitions with great confidence. The numerical results indicate that our new, ANN-based model is thus capable of providing reliable interpretations of experimental data. In this context, we report a resolution of the disputed assignment of the $3^1 e'$ state, now confirmed at 1055.5 cm^{-1} (computed at 1039.7 cm^{-1}) and the measured excitation at 1492.4 cm^{-1} (computed at 1490.0 cm^{-1}) corresponding to the $3^1 4^1 e'$ state. Given the accuracy reached, we can also confirm the measured transition corresponding to the $3^1 4^1 a'_2$ state to be most certainly an incorrect experimental assignment, as previously suggested by our polynomial model. This is supported by the fact that this state is the only one deviating from computed excitations by about 150 cm^{-1} across all investigated ANN models, an order of magnitude higher than any other (experimental or computational) deviation. Apart from a significant increase in accuracy, it is also found that throughout all investigated ANN-based PESs, there is not a single instance of oscillations producing deep, unphysical minima in scarcely sampled regions, a common problem with polynomial models. Such artifacts, if close enough to physically relevant areas in NCS, can cause wave packets to be trapped producing unphysical results. Consequently, these results suggest that the present model not only produces PESs of much higher accuracy than possible before but also better adapted to dynamics calculations.

The influence of ANN size (in terms of free parameter number) and ANN design (number of hidden layers) on the dynamics results is also studied. The same set of ANNs is investigated, which previously was assessed purely on the basis of the fitting error. The findings of the previous study confirmed (within the small scope of this study) that multilayered network architectures are not superior to single-layer ANNs for our diabaticization scheme. Similarly, no significant disadvantage is found to be caused by omitting *early stopping* from the fitting procedure as PESs produced remained indistinguishable from those of the unmodified fit. The previously reported indicator for overfitting is in agreement with the current finding that hidden layer sizes beyond that point either do not improve the dynamics results or even can lead to worse dynamics results.

In conclusion, our present findings (based on quantum dynamics calculations and experimental data) demonstrate that PESs produced by our novel ANN diabaticization method, despite its simple

setup, are not only very accurate but also robust against typical artifacts. Furthermore, the obtained PES models are very suitable for quantum dynamics calculations and yield results of high accuracy. The spread of quantum dynamics results among different parameterizations yielding a similar fitting error is rather small and gives a good idea of the size of errors to be expected in general. Remaining goals for future investigations include extending the model to a full-dimensional description of NO₃, including the umbrella motion, as well as corresponding couplings in the diabatic matrix, and applying the method to different kinds of coupling (e.g., relativistic coupling). Corresponding work is in progress.

ACKNOWLEDGMENTS

The authors are grateful for financial support via the PHC/DAAD, Grant No. PROCOPE 40442PD. Part of this work was also generously supported by the Deutsche Forschungsgemeinschaft (DFG). Calculations leading to the results presented here were performed on resources provided by the Paderborn Center for Parallel Computing.

APPENDIX: SUPPLEMENTARY EQUATIONS

1. Coordinates

The symmetry-adapted coordinates are constructed from a set of primitive valence coordinates as already described in previous work^{33,36} and account for the basic asymptotic behavior in the underlying low-order model. The primitive coordinates comprising the three N–O distances r_i and a set of O–N–O angles α_i are first transformed nonlinearly as

$$m_i = 1 - \exp(-\gamma(r_i - r_0)), \quad (\text{A1})$$

$$\alpha'_i = \frac{\alpha_i - \alpha_0}{r_j r_k}, \quad i \neq j \neq k, \quad (\text{A2})$$

where r_0 and α_0 are the respective distances and angles at the reference point and γ is a chosen Morse-parameter. The out-of-plane umbrella coordinate is omitted in the present study. These primitive coordinates then are linearly transformed to yield the symmetry-adapted coordinates a (breathing mode) and the degenerate asymmetric modes x_s, y_s (stretching) and x_b, y_b (bending),

$$a = \sqrt{\frac{1}{3}}(m_1 + m_2 + m_3), \quad (\text{A3})$$

$$x_s = \sqrt{\frac{1}{6}}(2m_1 - m_2 - m_3), \quad (\text{A4})$$

$$y_s = \sqrt{\frac{1}{2}}(m_2 - m_3), \quad (\text{A5})$$

$$x_b = \sqrt{\frac{1}{6}}(2\alpha'_1 - \alpha'_2 - \alpha'_3), \quad (\text{A6})$$

$$y_b = \sqrt{\frac{1}{2}}(\alpha'_2 - \alpha'_3). \quad (\text{A7})$$

2. Diagonal model terms

For the sake of simplicity, let $r_{s,b}^2$ be given as

$$r_{s,b}^2 = x_{s,b}^2 + y_{s,b}^2. \quad (\text{A8})$$

The totally symmetric diagonal contributions of the reference model are expressed here in terms of three independent scalar functions $V_i(\mathbf{Q})$,

$$\mathbf{W}_{\text{diag}}^d(\mathbf{Q}) = \begin{pmatrix} V_1(\mathbf{Q}) & 0^T & 0^T \\ 0 & V_2(\mathbf{Q})\mathbf{1} & 0 \\ 0 & 0 & V_3(\mathbf{Q})\mathbf{1} \end{pmatrix}. \quad (\text{A9})$$

Apart from the constant terms, referring to the vertical excitation energies at the reference point, an expansion of each $V_i(\mathbf{Q})$, $i = 1, 2, 3$ up to second order yields four (constant) coefficients μ_k^i and corresponding polynomial terms,

$$V_i(\mathbf{Q}) = \mu_1^i \cdot a + \mu_2^i \cdot a^2 + \mu_3^i \cdot r_s^2 + \mu_4^i \cdot r_b^2. \quad (\text{A10})$$

REFERENCES

- M. A. Collins and D. F. Parsons, *J. Chem. Phys.* **99**, 6756 (1993).
- J. Ischtwan and M. A. Collins, *J. Chem. Phys.* **100**, 8080 (1994).
- T. S. Ho and H. Rabitz, *J. Chem. Phys.* **104**, 2584 (1996).
- G. G. Maisuradze, D. L. Thompson, A. F. Wagner, and M. Minkoff, *J. Chem. Phys.* **119**, 10002 (2003).
- G. G. Maisuradze, A. Kawano, D. L. Thompson, A. F. Wagner, and M. Minkoff, *J. Chem. Phys.* **121**, 10329 (2004).
- B. J. Braams and J. M. Bowman, *Int. Rev. Phys. Chem.* **28**, 577 (2009).
- C. R. Evenhuis and M. A. Collins, *J. Chem. Phys.* **121**, 2515 (2004).
- C. R. Evenhuis, X. Lin, D. H. Zhang, D. Yarkony, and M. A. Collins, *J. Chem. Phys.* **123**, 134110 (2005).
- O. Godsi, C. R. Evenhuis, and M. A. Collins, *J. Chem. Phys.* **125**, 104105 (2006).
- X. Zhu, J. Y. Ma, D. R. Yarkony, and H. Guo, *J. Chem. Phys.* **136**, 234301 (2012).
- C. Xie, J. Ma, X. Zhu, D. H. Zhang, D. R. Yarkony, D. Xie, and H. Guo, *J. Phys. Chem. Lett.* **5**, 1055 (2014).
- Conical Intersections: Electronic Structure, Dynamics and Spectroscopy*, edited by W. Domcke, D. R. Yarkony, and H. Köppel (World Scientific, Singapore, 2004).
- H. C. Longuet-Higgins, *Adv. Spectrosc.* **2**, 429 (1961).
- W. Lichten, *Phys. Rev.* **131**, 229 (1963).
- W. Lichten, *Phys. Rev.* **164**, 131 (1967).
- F. T. Smith, *Phys. Rev.* **179**, 111 (1969).
- M. Baer, *Chem. Phys.* **15**, 49 (1976).
- H. Werner and W. Meyer, *J. Chem. Phys.* **74**, 5802 (1981).
- C. A. Mead and D. G. Truhlar, *J. Chem. Phys.* **77**, 6090 (1982).
- C. A. Mead, *J. Chem. Phys.* **78**, 807 (1983).
- H.-J. Werner, B. Follmeg, and M. H. Alexander, *J. Chem. Phys.* **89**, 3139 (1988).
- T. Pacher, L. S. Cederbaum, and H. Köppel, *J. Chem. Phys.* **89**, 7367 (1988).
- T. Pacher, C. A. Mead, L. S. Cederbaum, and H. Köppel, *J. Chem. Phys.* **91**, 7057 (1989).
- T. Pacher, H. Köppel, and L. S. Cederbaum, *J. Chem. Phys.* **95**, 6668 (1991).
- T. Pacher, L. S. Cederbaum, and H. Köppel, *Adv. Chem. Phys.* **84**, 293 (1993).
- H. Köppel, W. Domcke, and L. S. Cederbaum, *Adv. Chem. Phys.* **57**, 59 (1984).
- M. S. Schuurman and D. R. Yarkony, *J. Chem. Phys.* **127**, 094104 (2007).
- B. N. Papas, M. S. Schuurman, and D. R. Yarkony, *J. Chem. Phys.* **129**, 124104 (2008).
- X. Zhu and D. R. Yarkony, *J. Chem. Phys.* **130**, 234108 (2009).
- X. Zhu and D. R. Yarkony, *J. Chem. Phys.* **132**, 104101 (2010).
- A. Viel and W. Eisfeld, *J. Chem. Phys.* **120**, 4603 (2004).

- ³²W. Eisfeld and A. Viel, *J. Chem. Phys.* **122**, 204317 (2005).
- ³³A. Viel, W. Eisfeld, S. Neumann, W. Domcke, and U. Manthe, *J. Chem. Phys.* **124**, 214306 (2006).
- ³⁴A. Viel, W. Eisfeld, C. R. Evenhuis, and U. Manthe, *Chem. Phys.* **347**, 331 (2008).
- ³⁵S. Faraji, H. Köppel, W. Eisfeld, and S. Mahapatra, *Chem. Phys.* **347**, 110 (2008).
- ³⁶W. Eisfeld, O. Vieuxmaire, and A. Viel, *J. Chem. Phys.* **140**, 224109 (2014).
- ³⁷R. Cimiraglia, J. P. Malrieu, M. Persico, and F. Spiegelmann, *J. Phys. B: At. Mol. Phys.* **18**, 3073 (1985).
- ³⁸W. Domcke and C. Woywod, *Chem. Phys. Lett.* **216**, 362 (1993).
- ³⁹G. J. Atchity and K. Ruedenberg, *Theor. Chem. Acc.* **97**, 47 (1997).
- ⁴⁰H. Nakamura and D. G. Truhlar, *J. Chem. Phys.* **115**, 10353 (2001).
- ⁴¹H. Nakamura and D. G. Truhlar, *J. Chem. Phys.* **117**, 5576 (2002).
- ⁴²H. Nakamura and D. G. Truhlar, *J. Chem. Phys.* **118**, 6816 (2003).
- ⁴³P. Cattaneo and M. Persico, *Theor. Chem. Acc.* **103**, 390 (2000).
- ⁴⁴S. Nangia and D. G. Truhlar, *J. Chem. Phys.* **124**, 124309 (2006).
- ⁴⁵Z. H. Li, R. Valero, and D. G. Truhlar, *Theor. Chem. Acc.* **118**, 9 (2007).
- ⁴⁶X. Zhu and D. R. Yarkony, *J. Chem. Phys.* **140**, 024112 (2014).
- ⁴⁷T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, *J. Chem. Phys.* **103**, 4129 (1995).
- ⁴⁸D. F. R. Brown, M. N. Gibbs, and D. C. Clary, *J. Chem. Phys.* **105**, 7597 (1996).
- ⁴⁹K. T. No, B. H. Chang, S. Y. Kim, M. S. Jhon, and H. A. Scheraga, *Chem. Phys. Lett.* **271**, 152 (1997).
- ⁵⁰F. V. Prudente, P. H. Acioli, and J. J. S. Neto, *J. Chem. Phys.* **109**, 8801 (1998).
- ⁵¹S. Lorenz, A. Gross, and M. Scheffler, *Chem. Phys. Lett.* **395**, 210 (2004).
- ⁵²L. M. Raff, M. Malshe, M. Hagan, D. I. Doughan, M. G. Rockley, and R. Komanduri, *J. Chem. Phys.* **122**, 084104 (2005).
- ⁵³S. Lorenz, M. Scheffler, and A. Gross, *Phys. Rev. B* **73**, 115431 (2006).
- ⁵⁴S. Manzhos, X. G. Wang, R. Dawes, and T. Carrington, *J. Phys. Chem. A* **110**, 5295 (2006).
- ⁵⁵S. Manzhos and T. Carrington, Jr., *J. Chem. Phys.* **125**, 084109 (2006).
- ⁵⁶S. Manzhos and T. Carrington, Jr., *J. Chem. Phys.* **125**, 194105 (2006).
- ⁵⁷J. Behler and M. Parrinello, *Phys. Rev. Lett.* **98**, 146401 (2007).
- ⁵⁸S. Manzhos and T. Carrington, Jr., *J. Chem. Phys.* **127**, 014103 (2007).
- ⁵⁹M. Malshe, R. Narulkar, L. M. Raff, M. Hagan, S. Bukkapatnam, and R. Komanduri, *J. Chem. Phys.* **129**, 044111 (2008).
- ⁶⁰S. Manzhos and T. Carrington, Jr., *J. Chem. Phys.* **129**, 224104 (2008).
- ⁶¹J. Behler, *J. Chem. Phys.* **134**, 074106 (2011).
- ⁶²H. T. T. Nguyen and H. M. Le, *J. Phys. Chem. A* **116**, 4629 (2012).
- ⁶³B. Jiang and H. Guo, *J. Chem. Phys.* **139**, 054112 (2013).
- ⁶⁴W. Koch and D. H. Zhang, *J. Chem. Phys.* **141**, 021101 (2014).
- ⁶⁵G. Cybenko, *Math. Control Signals Syst.* **2**, 303 (1989).
- ⁶⁶H. D. Meyer, U. Manthe, and L. S. Cederbaum, *Chem. Phys. Lett.* **165**, 73 (1990).
- ⁶⁷U. Manthe, H. D. Meyer, and L. S. Cederbaum, *J. Chem. Phys.* **97**, 3199 (1992).
- ⁶⁸U. Manthe, *J. Chem. Phys.* **105**, 6989 (1996).
- ⁶⁹A. Brown and E. Pradhan, *J. Theor. Comput. Chem.* **16**, 1730001 (2017).
- ⁷⁰Y. Guan, B. Fu, and D. H. Zhang, *J. Chem. Phys.* **147**, 224307 (2017).
- ⁷¹T. Lenzen and U. Manthe, *J. Chem. Phys.* **147**, 084105 (2017).
- ⁷²C. Xie, X. Zhu, D. R. Yarkony, and H. Guo, *J. Chem. Phys.* **149**, 144107 (2018).
- ⁷³D. M. G. Williams and W. Eisfeld, *J. Chem. Phys.* **149**, 204106 (2018).
- ⁷⁴A. Viel and W. Eisfeld, *Chem. Phys.* **509**, 81 (2018).
- ⁷⁵M. Mayer, L. S. Cederbaum, and H. Köppel, *J. Chem. Phys.* **100**, 899 (1994).
- ⁷⁶M. Okumura, J. Stanton, A. Deev, and J. Sommar, *Phys. Scr.* **73**, C64 (2006).
- ⁷⁷S. Mahapatra, W. Eisfeld, and H. Köppel, *Chem. Phys. Lett.* **441**, 7 (2007).
- ⁷⁸W. Eisfeld and A. Viel, *J. Chem. Phys.* **146**, 034303 (2017).
- ⁷⁹W. Eisfeld and K. Morokuma, *J. Chem. Phys.* **114**, 9430 (2001).
- ⁸⁰W. Eisfeld and K. Morokuma, *J. Chem. Phys.* **113**, 5587 (2000).
- ⁸¹W. Eisfeld, *J. Chem. Phys.* **134**, 054303 (2011).
- ⁸²P. J. Knowles and H.-J. Werner, *Chem. Phys. Lett.* **145**, 514 (1988).
- ⁸³H.-J. Werner and P. J. Knowles, *J. Chem. Phys.* **89**, 5803 (1988).
- ⁸⁴J. C. Light and T. Carrington, *Adv. Chem. Phys.* **114**, 263 (2000).
- ⁸⁵U. Manthe, *J. Phys.: Condens. Matter* **29**, 253001 (2017).
- ⁸⁶U. Manthe, *J. Chem. Phys.* **128**, 064108 (2008).
- ⁸⁷C. Evenhuis, G. Nyman, and U. Manthe, *J. Chem. Phys.* **127**, 144302 (2007).
- ⁸⁸K. Kawaguchi, T. Narahara, R. Fujimori, J. Tang, and T. Ishiwata, *J. Mol. Spectrosc.* **334**, 10 (2017).
- ⁸⁹T. Codd, M.-W. Chen, M. Roudjane, J. F. Stanton, and T. A. Miller, *J. Chem. Phys.* **142**, 184305 (2015).
- ⁹⁰T. Weihe, D. M. G. Williams, A. Viel, and W. Eisfeld, *J. Chem. Phys.* **151**, 074302 (2019).
- ⁹¹M. Hagan, H. Demuth, M. Beale, and O. De Jesús, in *Neural Network Design*, 2nd ed., edited by M. Hagan (2014), ISBN: 9780971732117.
- ⁹²J. F. Stanton, *J. Chem. Phys.* **126**, 134309 (2007).
- ⁹³J. F. Stanton, *Mol. Phys.* **107**, 1059 (2009).
- ⁹⁴C. S. Simmons, T. Ichino, and J. F. Stanton, *J. Phys. Chem. Lett.* **3**, 1946 (2012).
- ⁹⁵Z. Homayoon and J. M. Bowman, *J. Chem. Phys.* **141**, 161104 (2014).
- ⁹⁶M. E. Jacox and W. E. Thompson, *J. Chem. Phys.* **129**, 204306 (2008).
- ⁹⁷K. Kawaguchi, R. Fujimori, J. Tang, and T. Ishiwata, *J. Mol. Spectrosc.* **314**, 73 (2015).