



HAL
open science

Conquering coupled diabatic potential energy surfaces with artificial neural networks for nonadiabatic dynamics

David M. G. Williams

► **To cite this version:**

David M. G. Williams. Conquering coupled diabatic potential energy surfaces with artificial neural networks for nonadiabatic dynamics. Chemical Sciences. University of Bielefeld, 2020. English. NNT : . tel-03917774

HAL Id: tel-03917774

<https://hal.science/tel-03917774v1>

Submitted on 6 Jan 2023

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.

PhD thesis manuscript of **David M. G. Williams** defended on January the 28th of 2021 in front of the jury members:

- Torben Lübke
- Uwe Manthe
- Michael Schwake
- Wolfgang Eisfeld (supervisor)

in order to obtain the title of **Doktors der Naturwissenschaften** of the **Universität Bielefeld**.

Manuscrit du doctorat de thèse de **David M. G. Williams** soutenue le 28 janvier 2021 devant le jury:

- Torben Lübke
- Uwe Manthe
- Michael Schwake
- Wolfgang Eisfeld (directeur de thèse)

en vue de l'obtention du grade de **Docteur** en sciences de l'**université de Bielefeld**.

Conquering coupled diabatic potential energy surfaces with artificial neural networks for nonadiabatic dynamics

David M. G. Williams

December 18, 2020

Contents

Acknowledgment	v
1 General Summary	1
1.1 State of the Art	1
1.2 Detailed Summaries of Presented Articles	4
Articles included in the present thesis	9
Bibliography	11
Appendix	15

Acknowledgment

First I would like to thank Prof. Dr. Wolfgang Eisfeld for giving me the opportunity working on such a diverse and fascinating project. I would also like to express my deepest gratitude for his continued guidance, advice and support.

Furthermore I would like to thank Prof. Dr. Uwe Manthe and the members of the AGTC, present and past, for the many enlightening discussions we shared over the years. In particular I would like to thank Thomas Weike and Florian Venghaus for their time and patience to discuss even the strangest of musings with me. Without them, I do not know whether my pursuits would have ended up quite as fruitful.

I also want to thank the *Free Software Foundation*, whose invaluable contributions to free software I gladly rely on with every keystroke. Last but not least I want to thank my family and friends for their endless love and encouragement.

1 General Summary

1.1 State of the Art

The Born-Oppenheimer (BO) approximation is a cornerstone of the theoretical treatment of molecular processes. It separates nuclear and electronic motion, splitting up an immensely challenging many-body problem into two highly involved but well-understood ones. The electronic problem, solved for fixed nuclear geometries, yields eigenvalues of the electronic Hamiltonian as a function of the nuclear coordinates. Each eigenvalue (as a function of the nuclear coordinates) forms a potential energy surface (PES) for a given electronic eigenstate. The set of these electronic eigenstates forms the so-called *adiabatic basis* throughout nuclear configuration space (NCS). The nuclear problem is then represented in this adiabatic basis in a second step usually. The BO approximation itself is applied in this second step, by discarding derivatives of the adiabatic basis with respect to nuclear coordinates arising from the nuclear kinetic energy operator. As a consequence, this approximation neglects coupling terms between electronic states. It generally holds for cases where the energies of electronic states are sufficiently separated for the dynamically relevant domain of the NCS, making it perfectly viable for a large range of applications. Conversely, the BO approximation fails when these *nonadiabatic coupling* terms become singular or very large for conical intersections (CI) or avoided crossings [1]. Jahn-Teller (JT) systems are a prototypical example where (symmetry-induced) CIs play a crucial role, as the presence of conical intersections typically induces the so-called geometric phase effect (GPE) [1, 2]. While the underlying fundamental theory of the GPE is well understood, its influence on molecular spectroscopy proves to be non-trivial even for $E \otimes e C_{3v}$ JT systems [2–10]. Publication [E1], discussed in greater detail below, provides a systematic study of how (and under what conditions) the GPE influences the vibronic level structure of a general $E \otimes e$ JT C_{nv} system. This in turn provides a way to experimentally observe the GPE by spectroscopic means for such general systems, connecting theory and experiment.

Comparing theoretical findings with experimental data necessitates theoretical computations of sufficiently high accuracy. Hence, the construction of highly accurate PES models is of fundamental interest. For JT systems, where the BO approximation breaks down at the usually dynamically relevant symmetry point, it proves more convenient to use a different representation. Such representations, called *diabatic representations*, are constructed such that derivative couplings are sufficiently small, and there is no GPE causing problems. In return, the electronic Hamiltonian can no longer be described by a diagonal matrix of PESs. As diabatic electronic states are no longer eigenfunctions of the electronic Hamiltonian throughout the NCS, the diabatic representation becomes a fully

1 General Summary

occupied matrix. However, diabatic matrix elements representing the electronic Hamiltonian become smooth, well-behaved functions, simplifying the construction of PESs. It may be important to note that while a true diabatic basis may not exist for realistic models beyond diatomics [11, 12], an approximate *quasi-diabatic* basis can be chosen in practice. Such basis sets will hence be referred to as “diabatic” for the sake of brevity. These representations allow for constructing a representation of a so-called *diabatic (potential) matrix* representing the electronic Hamiltonian with smooth, well-defined analytic functions. However, diabatic representations cannot be computed directly and thus methods for the diabatization of adiabatic states are required.

Common diabatization methods can be split roughly into two categories. The first category utilizes electronic wave function information. These approaches are usually characterized by different criteria enforced in the wave function [13–30]. The primary drawback of these schemes is that they do not yield analytic PES models on their own, meaning they need to be combined with an appropriate model in a second step to ensure efficient evaluation usually required by nuclear quantum dynamics calculations. In contrast to the above, so-called diabatization by *ansatz* [31–39] schemes exclusively utilize adiabatic energies obtained from *ab initio* calculations. The central idea of diabatization by *ansatz* exploits the simple structure of diabatic coupling elements with the necessity of the diagonalized form of the model reproducing adiabatic energies for each point in NCS. Linear vibronic coupling (LVC) [40] and similar models [25–28], apart from being immensely popular, are prototypical for how such a diabatization method is commonly implemented. Related models [34–39] going beyond LVC usually expand the necessary coupling terms as polynomials in suitable coordinates, up to some (usually low) order, with LVC approaches expanding coupling terms in first or *linear* order only. LVC models have been extremely successful in describing processes dominated by short-term dynamics with a compact and straightforward model. However, this simplicity comes at the cost of introducing arbitrariness, as the diabatic matrix elements are not specified beyond reproducing *ab initio* data once diagonalized. Additionally, increasing accuracy for systems where short-term dynamics are insufficient proves far from straightforward, to the point that the construction of highly accurate diabatic PESs in closed form continues to be a major challenge beyond triatomic systems [41–44]. Various attempts have been made to overcome this problem. Such approaches include the modified Shepard interpolation [45–47], the inclusion of invariant polynomials and CNPI symmetry [48], and manually tailoring the functional form of diabatic matrix elements to satisfy the requirements of the given system [41–43]. Finally, a new diabatization approach was developed recently in this research group called *hybrid diabatization* [49]. This diabatization technique combines the advantages of diabatization by *ansatz* with techniques based on wave function information, thus reducing the arbitrariness of the model and retaining more *ab initio* data per nuclear geometry. The hybrid diabatization method is the starting point of the work presented here. Nonetheless, accurate high-dimensional diabatic models remain scarce in the literature. One seemingly universally shared limitation of polynomial-based models is that increasing polynomial orders introduces numerical instabilities to the fitting process and increase the chance of unphysical artifacts. These artifacts can range from minor oscillations to pathological regions where the PES is no

longer bound. The latter case, if not treated, can render dynamics calculations unusable, as the wave packet tunnels into a normally dynamically less relevant domain. For this reason, alternative functional forms with more desirable properties are of fundamental interest.

One such functional form is provided by artificial neural networks (ANNs), which have found a recent surge in popularity. While initially used for the construction of adiabatic PESs [50–68], first attempts have been made towards diabatic ANN models [69–78]. To understand their advantages over polynomial structures, one must first take a closer look at their mathematical structure. The umbrella term “artificial neural network” describes a very broad class of computational systems, yet the primary development focus usually lies on the much more narrowly defined subclass of *feed-forward neural networks* or *multilayer perceptrons*, which can be understood as a computational process involving the successive computation of vectors (*layers*), beginning with a given *input layer* of coordinates and returning a final *output layer*. If thought of as a function, input and output layer serve as function arguments and values, respectively. Each layer is multiplied by a rectangular “weight matrix” and incremented by a “bias vector”. Each element of the resulting vector is then evaluated by a layer-dependent, nonlinear “activation function” of a single variable. The result defines the elements of the next layer. A given ANN may have an arbitrary amount of intermediate (*hidden*) layers which only serve a computational purpose and introduce further flexibility to the system. In such a setup, the weights and biases are the only fitting parameters, with the (usually sigmoid) activation function being fixed. ANNs with a single hidden layer were shown mathematically to be capable of reproducing any continuous, real function of n dimensions up to arbitrary accuracy [79]. Hence, increasing an ANN’s number of hidden layer neurons is (in terms of model flexibility) equivalent to increasing the order in a polynomial expansion. However, this means that ANNs extend rather naturally, as each “term” of a neural network is composed of bound sigmoid functions. This is in stark contrast to polynomial expansions, where increasing the flexibility involves including rapidly growing terms of increasing steepness. This structural difference translates into an increased stability during fitting, being less prone to causing unphysical artifacts in sparsely sampled regions of NCS. Most importantly, well-posed ANN models do not seem to produce pathological regions in sparsely sampled regions of the NCS as polynomials do. However, this advantage comes at the cost of losing the hierarchical structure of series expansions, making the neural network output hard to analyze and interpret. In the work presented here, a new scheme is developed, which unifies the flexibility of neural networks with the commonly known and well-understood structure of LVC. This scheme shows great potential to produce highly accurate diabatic models of exceptional simplicity, both computational *and* conceptual, reliability and extensibility, making it a promising tool to enable accurate nonadiabatic quantum dynamics studies in the future.

1.2 Detailed Summaries of Presented Articles

The Jahn-Teller effect and its influence on vibronic energy levels of doubly degenerate (E) electronic states is a long-standing subject of fundamental theoretical interest. A study of how the geometric phase effect influences general C_{nv} systems is presented in publication [E1] utilizing simple analytical and numerical models for the cases of n from 3–8, including the different combinations of $E_\alpha \otimes e_\beta$ Jahn-Teller systems for point groups with more than one irreducible E -representation. It is shown that findings in C_{3v} systems do not trivially generalize to arbitrary $E_\alpha \otimes e_\beta$ systems, as induced geometric phases strongly depend on the particular choice of n , α and β . It is first shown that restricting a model to a single polynomial coupling term is incapable of describing a system's multi-well structure, being itself of higher ($C_{\infty v}$) symmetry. A general derivation scheme is provided to construct necessary (single-mode) coupling terms up to arbitrary order. A minimal 2×2 model system is constructed for each case, composed of the two lowest coupling orders. The inclusion of more than one coupling term induces additional conical intersections away from the C_{nv} reference point, separating the NCS into an *inner* and *outer* region delimited by a regular polygon of conical intersections. Apart from analytical arguments, numerical dynamics calculations are performed to further confirm theoretical findings. The presence or absence of a GPE in conclusion is found to depend both on the leading polynomial coupling order of the model as well as the predominant localization of the wave function either within the inner or outer region. This work also presents a simple Hückel-type model capable of predicting qualitatively the structure of vibronic energy levels for systems with sufficiently deep multiple wells. In such cases the vibronic level structure is composed of multiplets corresponding to symmetrized local vibrational excitations in each single (harmonic) well. Since barriers between wells are finite, these tunneling multiplets split with a characteristic level ordering. The Hückel model is capable of predicting the expected level ordering and the irreducible representations of the vibronic eigenstates based on the distribution of excitation quanta in the local well picture. Consequently, this work provides a way to infer spectroscopically the presence (or absence) of the GPE based on the observed vibronic level structure. The prediction of the symmetry of the excited vibrational modes helps simplifying spectroscopic assignments.

(This author contributed to the development of the theoretical components of this work. He also contributed to the writing of the corresponding body of text of the manuscript.)

The construction of highly accurate, diabatic models for many electronic states for use in nuclear dynamics is an active and essential field of interest. The present work aims to propose a new method of constructing such models, utilizing state of the art artificial neural network (ANN) methodology. The newly developed ANN-based diabatization scheme is first established in publication [E2] in the literature. It utilizes the structure of a linear vibronic coupling *ansatz* which is then extended by modifying the polynomial expansion coefficients using a feed-forward neural network. The model is hence generated

from appropriately chosen symmetry-adapted coordinates in two steps. In a first step, a minimal, bound vibronic coupling model is fitted against available *ab initio* data. In this paper, the model is composed of all diagonal terms up to second order as well as all coupling terms up to linear order. The ANN model is then generated by introducing a simple multiplicative term for every (constant) coupling coefficient, as well as one diagonal term for every electronic state. Each multiplicative term is composed of a respective entry of the neural network’s output layer, meaning the number of so-called output neurons of the ANN equals the number of coupling matrices (plus the number of different diagonal elements). Consequently, increasing the number of polynomial coupling matrices no longer serves as the primary source for increasing model flexibility. Instead, the expansion coefficients are modulated as a function of nuclear coordinates. In a purely polynomial context, this corresponds to “pulling out” common scalar prefactors in individual matrix terms and contracting all prefactors with a shared (simplified) basis matrix to a function. A simple example of such a simplification can be found in Eq. 16 of [E1], where part of the infinite expansion is absorbed into “coefficient” $f_{\pm k}^{n\alpha}$. This approach is developed to combine the straightforwardness of a vibronic coupling approach with the accuracy and stability of an ANN. The modified model is fitted against adiabatic energies by means of diagonalization using an appropriately adapted Marquardt-Levenberg (ML) method. Since the resulting fit is highly non-linear, every tested ANN architecture is fitted using a set of 100 randomly generated initial guesses. A small part of the data set is withheld from the fit itself to serve as a validation set used in a relaxed form of the *early stopping* technique [80]. The root mean square error (RMSE) of the validation set serves in this context as a test of convergence, and is a key indicator for overfitting. The new scheme has been applied to a 5-state model of NO_3 restricted to planar geometries as a difficult benchmark system. A large range of hidden layer sizes has been studied, together with potential benefits of “deep” neural networks with more than one hidden layer. Such deep networks (within the limited confines of the study) are not found to improve fitting performance when compared to a representative single-layer network with an approximately equivalent number of fitting parameters. Fitting error, validation set error, and state-resolved RMSEs are utilized as primary metrics for network performance. The discrepancy between fitting and validation RMSE serves as a measure for risk of overfitting. Additionally, the tenth percentile of fitted neural networks is taken into consideration as a simple measure for the stability of the fitting result. The best fitted 75 hidden layer neuron ANN is deemed optimal in terms of fitting error versus risk of overfitting by all accounts. Overall, the approach is found to yield models of excellent accuracy in terms of fitting RMS.

(The candidate both developed and implemented the modified ML scheme and co-developed the *ansatz* itself. He also performed the necessary ANN fits using a specialized program he co-developed, and contributed significant portions to the manuscript’s body of text.)

1 General Summary

A rigorous and systematic test of the performance of the newly developed ANN diabatic PES model in accurate quantum dynamics calculations is presented in [E3]. It serves as a continued study of the scheme proposed in [E2] with a focus on the central point of the potential model: nuclear dynamics application. For this purpose, time-independent DVR calculations are carried out for all previously considered network sizes on the ${}^2A'_2$ ground state as well as the ${}^2E''$ excited state of the NO_3 radical. Vibrational/vibronic states also are computed for tenth percentile networks and multi-layer networks investigated previously, to ensure consistency with previous findings. First, computed levels are compared among different networks, with the tenth percentile networks of the 75 neuron network serving as a simple measure for how computed levels spread among (in terms of fitting error) equivalent models. This measure then is utilized to put differences between various network sizes and architectures into perspective. While reducing the number of neurons (compared to the reference 75 network) is found to have a significant impact on dynamics well beyond differences among competing networks, increasing the network size is observed to have a lesser overall impact, further suggesting that the reference network provides sufficient flexibility. The model also is compared to a state-of-the-art polynomial model from Ref. [81] as well as experimental data [82, 83]. To ensure a rigorous comparison, the 6D polynomial model is restricted to 5D, using the multi-configuration time-dependent Hartree (MCTDH) [84, 85] approach, with the potential being evaluated using the correlated discrete variable representation (CDVR) [86] method. Analogous computations also are carried out for the ANN model, to ensure dynamics results are not meaningfully affected by the particular choice of dynamics calculation. The reference network is found to not only reproduce experimental data optimally among the different network sizes tested, but also significantly improve agreement with experimental data when compared to the reference polynomial model.

(The candidate performed the necessary DVR calculations (including state assignments) and implemented the model for the nuclear dynamics application. He also developed an appropriate estimate for the expected accuracy of computed energy levels purely based on *a priori* information. Furthermore, he authored the manuscript together with the supervisor and another co-author.)

The subject of the fourth publication [E4] is the refinement of the diabaticization method previously established in [E2,E3] by ensuring the ANN model is intrinsically symmetrized without the need of additional symmetrization. To this end, CNPI invariants are used as coordinates. This approach again is tested using a combination of methods utilized in [E2,E3]. However, instead of comparing the modified model to the purely polynomial model described in Ref. [81], the best model from [E3] is used as a reference. The modified approach is found to improve previous efforts in every considered aspect, including fitting error, compactness, and agreement with experiment. The best network setup again is determined by analyzing the fitting performance w.r.t. hidden layer size [E2] and further confirmed based on dynamics calculations in analogy to [E3]. The optimal network size for the modified approach turns out to be 30 instead of 75 hidden layers, cutting

both fitting error and number of fitted parameters roughly in half. As a result, the remaining error is deemed low enough to approach the expected limit of accuracy for the available *ab initio* data. Ground state dynamics results resemble closely those produced by the previous 75 neuron model. Excited state (${}^2E_2''$) results are more significantly affected, coinciding with the model having significantly improved in reproducing the intricate triple well structure of the lower \tilde{A} sheet when compared to [E2]. This is likely caused by the modified input coordinates already encoding the required general structure (symmetry), thus greatly removing strain from the network. Similarly, agreement with experimental data is not changed significantly for the ground state, while it is vastly improved for the \tilde{A} state.

(The candidate developed and implemented the CNPI-modified ANN scheme, and carried out necessary ANN fits. The candidate also performed the necessary DVR computations, and wrote the manuscript together with the supervisor.)

Up to this point, the benchmark system NO_3 was restricted to planar geometries to ensure consistency among studies [E2–4]. This restriction is lifted in [E5], where the previous 5D CNPI-ANN model is embedded into a complete 6D model. For this purpose, a natural extension of the diabatization scheme was developed that allows additional ANNs to be included hierarchically. This way, a model can be fitted consecutively against growing sub-domains of the NCS in such a way that new contributions vanish for previously “covered” parts of the NCS. In this sense, the second neural network serves as a correction upon a coefficient that is already modified by another ANN. This gives way to a more general correction-upon-correction approach which may be used to avoid a duplication of effort. In the present case, this scheme is used to create a fully occupied PES matrix, taking into account umbrella motion. The new, full-dimensional model was used then to analyze the effects of temperature and vibronic coupling in the low-energy regime of the photodetachment spectrum of NO_3^- . Dynamics calculations were performed using MCTDH for wavepacket propagation. Spectra are obtained from partial spectra corresponding to the vibronically coupled \tilde{X} and \tilde{B} state manifolds and various vibrationally excited initial states to account for temperature effects. Furthermore, near threshold effects are accounted for in the simulations of the cryo-SEVI spectra by use of Wigner’s threshold law. Computed partial spectra, when combined, are found to be in excellent agreement with experimental data [87, 88]. This unprecedented accuracy, combined with a detailed simulation of temperature and near threshold effects allows to settle long-standing disputes regarding particular assignments, and made it possible to fully distinguish thermal and vibronic contributions to the spectrum.

(The candidate developed, generalized and implemented the necessary modifications to the CNPI ANN model. He also carried out necessary ANN fits and contributed to the writing of the manuscript.)

Articles included in the present thesis

- [E1] T. Weike, D. M. G. Williams, A. Viel, and W. Eisfeld,
Quantum dynamics and geometric phase in $E \otimes e$ Jahn-Teller systems with general C_{nv} symmetry, J. Chem. Phys. **151**, 074302 (2019)
- [E2] D. M. G. Williams and W. Eisfeld,
Neural network diabaticization: A new ansatz for accurate high-dimensional coupled potential energy surfaces, J. Chem. Phys. **149**, 204106 (2018)
- [E3] D. M. G. Williams, A. Viel, and W. Eisfeld,
Diabatic neural network potentials for accurate vibronic quantum dynamics—The test case of planar NO_3 , J. Chem. Phys. **151**, 164118 (2019)
- [E4] D. M. G. Williams and W. Eisfeld,
Complete nuclear permutation inversion invariant artificial neural network (CNPI-ANN) diabaticization for the accurate treatment of vibronic coupling problems, J. Phys. Chem. A, **124**, 7608-7621 (2020)
- [E5] A. Viel, D. M. G. Williams, and W. Eisfeld,
Accurate quantum dynamics simulation of the photodetachment spectrum of the nitrate anion (NO_3^-) based on an artificial neural network diabatic potential model, J. Chem. Phys., submitted (2020).

Bibliography

- [1] W. Domcke, D. R. Yarkony, and H. Köppel, eds., *Conical Intersections: Electronic Structure, Dynamics and Spectroscopy* (World Scientific, Singapore, 2004).
- [2] C. A. Mead, *Rev. Mod. Phys.* **64**, 51 (1992).
- [3] H. Koizumi and I. B. Bersuker, *Phys. Rev. Lett.* **83**, 3009 (1999).
- [4] K. Pae and V. Hizhnyakov, *J. Chem. Phys.* **147**, 084107 (2017).
- [5] J. W. Zwanziger and E. R. Grant, *J. Chem. Phys.* **87**, 2954 (1987).
- [6] M. C. M. O'Brien, *Proc. R. Soc. A* **281**, 323 (1964).
- [7] C. A. Mead, *Chem. Phys.* **49**, 23 (1980).
- [8] F. S. Ham, *Phys. Rev. Lett.* **58**, 725 (1987).
- [9] J. W. Zwanziger, M. Koenig, and A. Pines, *ann. Rev. Phys. Chem.* **41**, 601 (1990).
- [10] J. Schön and H. Köppel, *J. Chem. Phys.* **103**, 9292 (1995).
- [11] C. A. Mead and D. G. Truhlar, *J. Chem. Phys.* **77**, 6090 (1982).
- [12] T. Pacher, L. S. Cederbaum, and H. Köppel, *Adv. Chem. Phys.* **84**, 293 (1993).
- [13] T. Pacher, L. S. Cederbaum, and H. Köppel, *J. Chem. Phys.* **89**, 7367 (1988).
- [14] H. Werner and W. Meyer, *J. Chem. Phys.* **74**, 5802 (1981).
- [15] G. Hirsch, R. J. Buenker, and C. Petrongolo, *Mol. Phys.* **70**, 835 (1990).
- [16] R. Cimiraglia, J. P. Malrieu, M. Persico, and F. Spiegelmann, *J. Phys. B-At. Mol. Opt.* **18**, 3073 (1985).
- [17] W. Domcke and C. Woywod, *Chem. Phys. Lett.* **216**, 362 (1993).
- [18] W. Domcke, C. Woywod, and M. Stengle, *Chem. Phys. Lett.* **226**, 257 (1994).
- [19] G. J. Atchity and K. Ruedenberg, *Theo. Chem. Acc.* **97**, 47 (1997).
- [20] R. Sadygov and D. Yarkony, *J. Chem. Phys.* **109**, 20 (1998).
- [21] H. Nakamura and D. G. Truhlar, *J. Chem. Phys.* **115**, 10353 (2001).

Bibliography

- [22] H. Nakamura and D. G. Truhlar, *J. Chem. Phys.* **117**, 5576 (2002).
- [23] R. Abrol and A. Kuppermann, *J. Chem. Phys.* **116**, 1035 (2002).
- [24] H. Nakamura and D. G. Truhlar, *J. Chem. Phys.* **118**, 6816 (2003).
- [25] M. S. Schuurman and D. R. Yarkony, *J. Chem. Phys.* **127**, 094104 (2007).
- [26] B. N. Papas, M. S. Schuurman, and D. R. Yarkony, *J. Chem. Phys.* **129**, 124104 (2008).
- [27] X. Zhu and D. R. Yarkony, *J. Chem. Phys.* **130**, 234108 (2009).
- [28] X. Zhu and D. R. Yarkony, *J. Chem. Phys.* **132**, 104101 (2010).
- [29] X. Zhu and D. R. Yarkony, *J. Chem. Phys.* **136**, 174110 (2012).
- [30] H. Lischka, M. Dallos, P. G. Szalay, D. R. Yarkony, and R. Shepard, *J. Chem. Phys.* **120**, 7322 (2004).
- [31] S. Mahapatra, G. A. Worth, H.-D. Meyer, L. S. Cederbaum, and H. Köppel, *J. Phys. Chem. A* **105**, 5567 (2001).
- [32] A. J. C. Varandas, F. B. Brown, C. A. Mead, D. G. Truhlar, and N. C. Blais, *J. Chem. Phys.* **86**, 6258 (1987).
- [33] H.-J. Werner, B. Follmeg, and M. H. Alexander, *J. Chem. Phys.* **89**, 3139 (1988).
- [34] A. Viel and W. Eisfeld, *J. Chem. Phys.* **120**, 4603 (2004).
- [35] W. Eisfeld and A. Viel, *J. Chem. Phys.* **122**, 204317 (2005).
- [36] A. Viel, W. Eisfeld, S. Neumann, W. Domcke, and U. Manthe, *J. Chem. Phys.* **124**, 214306 (2006).
- [37] A. Viel, W. Eisfeld, C. R. Evenhuis, and U. Manthe, *Chem. Phys.* **347**, 331 (2008).
- [38] S. Faraji, H. Köppel, W. Eisfeld, and S. Mahapatra, *Chem. Phys.* **347**, 110 (2008).
- [39] W. Eisfeld, O. Vieuxmaire, and A. Viel, *J. Chem. Phys.* **140**, 224109 (2014).
- [40] H. Köppel, W. Domcke, and L. S. Cederbaum, *Adv. Chem. Phys.* **57**, 59 (1984).
- [41] P. Cattaneo and M. Persico, *Theor. Chem. Acc.* **103**, 390 (2000).
- [42] S. Nangia and D. G. Truhlar, *J. Chem. Phys.* **124**, 124309 (2006).
- [43] Z. H. Li, R. Valero, and D. G. Truhlar, *Theor. Chem. Acc.* **118**, 9 (2007).
- [44] X. Zhu and D. R. Yarkony, *J. Chem. Phys.* **140**, 024112 (2014).
- [45] C. R. Evenhuis and M. A. Collins, *J. Chem. Phys.* **121**, 2515 (2004).

- [46] C. R. Evenhuis, X. Lin, D. H. Zhang, D. Yarkony, and M. A. Collins, *J. Chem. Phys.* **123**, 134110 (2005).
- [47] O. Godsi, C. R. Evenhuis, and M. A. Collins, *J. Chem. Phys.* **125**, 104105 (2006).
- [48] X. Zhu, J. Y. Ma, D. R. Yarkony, and H. Guo, *J. Chem. Phys.* **136**, 234301 (2012).
- [49] N. Wittenbrink, F. Venghaus, D. Williams, and W. Eisfeld, *J. Chem. Phys.* **145**, 184108 (2016).
- [50] T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, *J. Chem. Phys.* **103**, 4129 (1995).
- [51] D. F. R. Brown, M. N. Gibbs, and D. C. Clary, *J. Chem. Phys.* **105**, 7597 (1996).
- [52] K. T. No, B. H. Chang, S. Y. Kim, M. S. Jhon, and H. A. Scheraga, *Chem. Phys. Lett.* **271**, 152 (1997).
- [53] F. V. Prudente, P. H. Acioli, and J. J. S. Neto, *J. Chem. Phys.* **109**, 8801 (1998).
- [54] S. Lorenz, A. Gross, and M. Scheffler, *Chem. Phys. Lett.* **395**, 210 (2004).
- [55] L. M. Raff, M. Malshe, M. Hagan, D. I. Doughan, M. G. Rockley, and R. Komanduri, *J. Chem. Phys.* **122**, 084104 (2005).
- [56] S. Lorenz, M. Scheffler, and A. Gross, *Phys. Rev. B* **73**, 115431 (2006).
- [57] S. Manzhos, X. G. Wang, R. Dawes, and T. Carrington, *J. Phys. Chem. A* **110**, 5295 (2006).
- [58] S. Manzhos and T. Carrington Jr., *J. Chem. Phys.* **125**, 084109 (2006).
- [59] S. Manzhos and T. Carrington Jr., *J. Chem. Phys.* **125**, 194105 (2006).
- [60] J. Behler and M. Parrinello, *Phys. Rev. Lett.* **98**, 146401 (2007).
- [61] S. Manzhos and T. Carrington Jr., *J. Chem. Phys.* **127**, 014103 (2007).
- [62] M. Malshe, R. Narulkar, L. M. Raff, M. Hagan, S. Bukkapatnam, and R. Komanduri, *J. Chem. Phys.* **129**, 044111 (2008).
- [63] S. Manzhos and T. Carrington Jr., *J. Chem. Phys.* **129**, 224104 (2008).
- [64] J. Behler, *J. Chem. Phys.* **134**, 074106 (2011).
- [65] H. T. T. Nguyen and H. M. Le, *J. Phys. Chem. A* **116**, 4629 (2012).
- [66] B. Jiang and H. Guo, *J. Chem. Phys.* **139**, 054112 (2013).
- [67] W. Koch and D. H. Zhang, *J. Chem. Phys.* **141**, 021101 (2014).

Bibliography

- [68] M. Majumder, S. E. Hegger, R. Dawes, S. Manzhos, X.-G. Wang, T. Carrington, Jr., J. Li, and H. Guo, *Mol. Phys.* **113**, 1823 (2015).
- [69] Y. Guan, B. Fu, and D. H. Zhang, *J. Chem. Phys.* **147**, 224307 (2017).
- [70] T. Lenzen and U. Manthe, *J. Chem. Phys.* **147**, 084105 (2017).
- [71] C. Xie, X. Zhu, D. R. Yarkony, and H. Guo, *J. Chem. Phys.* **149**, 144107 (2018).
- [72] T. Lenzen, W. Eisfeld, and U. Manthe, *J. Chem. Phys.* **150**, 244115 (2019).
- [73] Y. Guan, H. Guo, and D. R. Yarkony, *J. Chem. Phys.* **150**, 214101 (2019).
- [74] Y. Wang, C. Xie, H. Guo, and D. R. Yarkony, *J. Phys. Chem. A* **123**, 5231 (2019).
- [75] Y. Guan, D. H. Zhang, H. Guo, and D. R. Yarkony, *Phys. Chem. Chem. Phys.* **21**, 14205 (2019).
- [76] Z. Yin, Y. Guan, B. Fu, and D. H. Zhang, *Phys. Chem. Chem. Phys.* **21**, 20372 (2019).
- [77] Y. Guan and D. R. Yarkony, *J. Phys. Chem. Lett.* **11**, 1848 (2020).
- [78] Y. Guan, H. Guo, and D. R. Yarkony, *J. Chem. Theor. Comp.* **16**, 302 (2020).
- [79] G. Cybenko, *Math. Control Signals Systems* **2**, 303 (1989).
- [80] M. Hagan, H. Demuth, M. Beale, and O. De Jesús, *Neural Network Design (2nd Edition)* (Martin Hagan, 2014), ISBN 9780971732117.
- [81] A. Viel and W. Eisfeld, *Chem. Phys.* **509**, 81 (2018).
- [82] K. Kawaguchi, T. Narahara, R. Fujimori, J. Tang, and T. Ishiwata, *J. Mol. Spectrosc.* **334**, 10 (2017).
- [83] T. Codd, M.-W. Chen, M. Roudjane, J. F. Stanton, and T. A. Miller, *J. Chem. Phys.* **142**, 184305 (2015).
- [84] H. D. Meyer, U. Manthe, and L. S. Cederbaum, *Chem. Phys. Lett.* **165**, 73 (1990).
- [85] U. Manthe, H. D. Meyer, and L. S. Cederbaum, *J. Chem. Phys.* **97**, 3199 (1992).
- [86] U. Manthe, *J. Chem. Phys.* **105**, 6989 (1996).
- [87] A. Weaver, D. W. Arnold, S. E. Bradforth, and D. M. Neumark, *J. Chem. Phys.* **94**, 1740 (1991).
- [88] M. C. Babin, J. A. DeVine, M. DeWitt, J. F. Stanton, and D. M. Neumark, *J. Phys. Chem. Lett.* **11**, 395 (2020).

Appendix

Quantum dynamics and geometric phase in $E \otimes e$ Jahn-Teller systems with general C_{nv} symmetry

Cite as: J. Chem. Phys. **151**, 074302 (2019); <https://doi.org/10.1063/1.5115396>
Submitted: 17 June 2019 . Accepted: 26 July 2019 . Published Online: 19 August 2019

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



View Online



Export Citation



CrossMark

The Journal
of Chemical Physics

Submit Today

The Emerging Investigators Special Collection and Awards
Recognizing the excellent work of early career researchers!



Quantum dynamics and geometric phase in $E \otimes e$ Jahn-Teller systems with general C_{nv} symmetry

Cite as: J. Chem. Phys. 151, 074302 (2019); doi: 10.1063/1.5115396

Submitted: 17 June 2019 • Accepted: 26 July 2019 •

Published Online: 19 August 2019



View Online



Export Citation



CrossMark

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

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: t.weike@uni-bielefeld.de

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

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

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

ABSTRACT

$E \otimes e$ Jahn-Teller (JT) systems are considered the prototype of symmetry-induced conical intersections and of the corresponding geometric phase effect (GPE). For decades, this has been analyzed for the most common case originating from C_{3v} symmetry and these results usually were generalized. In the present work, a thorough analysis of the JT effect, vibronic coupling Hamiltonians, GPE, and the effect on spectroscopic properties is carried out for general C_{nv} symmetric systems (and explicitly for $n = 3-8$). It turns out that the C_{3v} case is much less general than often assumed. The GPE due to the vibronic Hamiltonian depends on the leading coupling term of a diabatic representation of the problem, which is a result of the explicit n , α , and β values of a C_{nv} $E_\alpha \otimes e_\beta$ system. Furthermore, the general existence of n/m ($m \in \mathbb{N}$ depending on n , α , and β) equivalent minima on the lower adiabatic sheet of the potential energy surface (PES) leads to tunneling multiplets of n/m states (state components). These sets can be understood as local vibrations of the atoms around their equilibrium positions within each of the local PES wells symmetrized over all equivalent wells. The local vibrations can be classified as *tangential* or *radial* vibrations, and the quanta in the tangential mode together with the GPE determine the level ordering within each of the vibronic multiplets. Our theoretical predictions derived analytically are tested and supported by numerical model simulations for all possible $E_\alpha \otimes e_\beta$ cases for C_{nv} symmetric systems with $n = 3-8$. The present interpretation allows for a full understanding of the complex JT spectra of real systems, at least for low excitation energies. This also opens a spectroscopic way to show the existence or absence of GPEs.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5115396>

I. INTRODUCTION

The study of vibronic energy levels supported by doubly degenerate electronic states affected by the Jahn-Teller (JT) effect is a long standing subject studied for more than 60 years.^{1,2} In particular, the degenerate E state of a C_{3v} symmetric system has been the subject of many theoretical investigations. Several authors tackled the problem of the $E \otimes e$ case thoroughly.³⁻⁸ They conclude that due to the presence of the conical intersection and the associated geometric phase effect (GPE), the lowest vibronic state of such a system is doubly degenerate. A broad and thorough discussion of the GPE can be found in the excellent review by Mead⁹ and in the

monograph on conical intersections, edited by Domcke, Yarkony, and Köppel.¹⁰ Despite the widespread assumption that such JT systems generally have a double degenerate vibronic ground state, it has been shown for the $E \otimes e$ case of C_{3v} in a couple of studies that this is only valid up to a certain strength of the quadratic couplings.^{11,12} It was shown that with strong quadratic but weak linear couplings, the vibronic ground state is nondegenerate. In most studies, both a linear coupling term and a quadratic vibronic coupling term are considered, resulting in a lower adiabatic potential surface with three identical local wells. For all strengths of the linear and quadratic vibronic coupling terms, the symmetry induced conical intersection at high symmetry configurations is present

and is located in a region usually relevant for the vibronic ground state wave function. It has been shown by Zwanziger and Grant that three additional conical intersections exist on a circle around the symmetry point and at higher energy.⁶ For particularly strong quadratic but weak linear couplings, these additional conical intersections away from the high symmetry point are driven lower in energy such that they become relevant for the vibronic ground state. These additional conical intersections together with the symmetry induced one are the reason for the nondegenerate vibronic ground state. This is coherent with the widespread “rule” that the symmetry of a vibronic level depends on the number of encircled conical intersections (odd or even).

In real systems, like the prototypical NO₃ studied by us thoroughly,^{13–15} the coupling terms are not limited to the linear and quadratic low orders. Additional conical intersections may occur which complicates the picture further. Our previous investigation of the ²E'' state of NO₃ for which the vibronic level ordering pattern is not obvious, motivates the present study.

The aim of the present paper is to investigate the presence of the GPE, its effect on the vibronic quantum dynamics, and the vibronic level structure in C_{*nv*} symmetric Jahn-Teller systems. This is an extension of previous studies only dealing with C_{3*v*}.^{3–8} The cases varying from *n* = 3 to *n* = 8 are considered explicitly in the present work. A general derivation of coupled potential energy surfaces (PESs) of C_{*nv*} symmetric systems is given. It is similar to the procedure developed by us^{13,16,17} and others^{18–21} for the *n* = 3 case. For each C_{*nv*} case, we analyze the conical intersections and their properties (in terms of the GPE) and show that these properties depend on the lowest order coupling terms. To this end, an analytic proof by asymptotic analysis will be given as well as a numerical visualization of the GPE by plotting the vector fields of the adiabatic wave functions. The consequences of the couplings in terms of spectroscopic signature are discussed. The origin of this analyzed effect is that the E ⊗ *e* JT system for C_{*nv*} causes the lower adiabatic PES sheet to have multiple equivalent wells leading to a multiplet structure of the vibronic levels with as many states within each multiplet as the number of wells. The ordering of the levels in these multiplets is indicative of the GPE as will be shown in the present work and as has been recognized in earlier work.^{5,6,9} This systematic study of C_{*nv*} groups with general *n* also accounts for the existence of different *E* representations. All possible combinations of electronic symmetry E_{*α*} with the symmetry of the coupling coordinate *e*_{*β*} are investigated for the E_{*α*} ⊗ *e*_{*β*} JT effect. Our findings and conclusions are supported both by mathematical analysis and numerical quantum dynamics simulations. It turns out that the conclusions drawn from the C_{3*v*} case are less general than widely assumed.

This paper is organized as follows: Sec. II summarizes the relevant symmetry aspects as well as the derivation of the diabatic models. At this point, the diabatic C_{*nv*} symmetric models for arbitrary E_{*α*} ⊗ *e*_{*β*} are derived. Section III focuses specifically on the geometric phase effect induced by the potential energy surfaces and provides a mathematical proof of the GPE properties of general diabatic Hamiltonians and a visualization of the effect. Section IV addresses the effects on the vibronic quantum dynamics, presenting data relevant for spectroscopic

investigations. An analytic analysis and explanation using a Hückel-like next neighbor model is given for the vibronic level structure, and the findings are verified by numerical test calculations. The main aspects and conclusions of the present work are summarized in Sec. V.

II. SYMMETRY AND DIABATIC MODELS

A. Symmetry considerations

The well-known Jahn-Teller effect is a distortion of the molecular geometry due to vibronic coupling within a set of states belonging to a multidimensional irreducible representation (irrep). The best known example is the distortion of a degenerate *E* state of a C_{3*v*} symmetric system along a degenerate set of *e* coordinates, lowering the equilibrium geometry to only C_s due to E ⊗ *e* vibronic coupling. The existence of two-dimensional *E* representations, and thus the existence of degenerate electronic states and vibrational modes, is a general feature of all non-Abelian point groups with a rotational axis of order *n* ≥ 3. The most common molecular point groups of this type are the C_{*nv*} groups which are used here as a general example for our study. The case *n* = 3 has been used extensively in the past to obtain general insight into the effects of the symmetry on the quantum dynamics. However, we will show that this case is much less general than often assumed.

To this end, a closer look at all *E* representations of C_{*nv*} groups is necessary. Quantum mechanical eigenstates and vibrational normal modes must transform like irreps of the molecular symmetry group due to the Wigner theorem.²² Normal modes are just one special case of symmetry-adapted nuclear coordinates suitable to express potential energy surfaces (PESs). Sets of symmetry coordinates belonging to the same two-dimensional irrep *e*_{*β*}, which we label Q_{*βx*} and Q_{*βy*}, can be rotated freely. This can be utilized to transform the symmetry coordinates such that each of the components becomes an eigenfunction of the C_{*n*} rotation operator, the most critical operator of a C_{*nv*} symmetry group. While Q_{*βx*} and Q_{*βy*} are mixed by C_{*n*}, the complex coordinates

$$Q_{\beta\pm} = \frac{1}{\sqrt{2}}(Q_{\beta x} \pm iQ_{\beta y}) = re^{\pm i\beta\phi} \quad (1)$$

are eigenfunctions with respect to C_{*n*} with eigenvalues of $\exp\left(\pm \frac{2i\beta\pi}{n}\right)$. The β in Eq. (1) refers to the kind of the *E* irrep, namely, E_{*β*}. The groups C_{3*v*} and C_{4*v*} only have one *E* irrep, and thus β = 1 is usually omitted, but in general, C_{*nv*} groups have β_{max} irreps E_{*β*} of dimension 2 with

$$\beta_{\max} = \frac{n + n \bmod 2}{2} - 1. \quad (2)$$

For example, C_{5*v*} and C_{6*v*} have two different irreps E₁ and E₂.

The same kind of transformation as for the coordinates Eq. (1) can be applied to the electronic state components of a degenerate electronic state transforming as E_{*α*}, namely,

$$|\psi_{\alpha\pm}\rangle = \frac{1}{\sqrt{2}}(|\psi_{\alpha x}\rangle \pm i|\psi_{\alpha y}\rangle). \quad (3)$$

These are eigenfunctions with respect to \widehat{C}_n with eigenvalues of $\exp(\mp \frac{2i\alpha\pi}{n})$. As has been used in our previous work,¹⁶ both the complex symmetry adapted coordinates and the complex electronic state components are relevant for the derivation of the diabatic models because of their advantageous properties with respect to the rotation operator \widehat{C}_n .

B. General quantum dynamics aspects

The fundamental aspects of the quantum dynamics in any $E_\alpha \otimes e_\beta$ JT system generally can be studied in the two-dimensional subspace of the two e_β nuclear coordinates. It is of advantage to transform the Hamiltonian to polar coordinates which then reads

$$\widehat{H}(r, \phi) = -\frac{1}{2} \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2} \right) + \mathbf{D}(r, \phi), \quad (4)$$

where $\mathbf{D}(r, \phi)$ is the diabatic PES matrix. The coupling between electronic and nuclear motions is described by the off-diagonal elements of the potential matrix \mathbf{D} , which represents the electronic Hamiltonian in a diabatic electronic state basis. \mathbf{D} may be split into an average potential for both state components, $V\mathbf{1}$, and a coupling matrix V_c . The function V and the matrix V_c can be expanded in series, and the general form of the coupling matrix reads

$$V_c(r, \phi) = \sum_{k=0}^{\infty} \sum_{l=-\infty}^{\infty} c_{kl} r^k \begin{pmatrix} 0 & \exp(il\phi) \\ \exp(-il\phi) & 0 \end{pmatrix}. \quad (5)$$

Note that l can be negative. The coefficients c_{kl} depend on the details of the system, especially symmetry can lead to many vanishing coefficients.

It is of advantage to define l -dependent vibronic angular momentum operators

$$\widehat{J}_l = -i \frac{\partial}{\partial \phi} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{l}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (6)$$

which commute with the term of the Fourier expansion in Eq. (5) having the same l value. This operator will not commute with any other expansion term of V_c having a different l value. In the particular case of a V_c limited to a single l expansion term, \widehat{J}_l presents some interesting properties detailed below. In that case, \widehat{J}_l and V_c can be diagonalized simultaneously. The two eigenfunctions of \widehat{J}_l corresponding to the degenerate eigenvalue $\lambda/2$ are given by

$$\frac{1}{\sqrt{2}} \begin{pmatrix} \exp(i\frac{\lambda+l}{2}\phi) \\ \exp(i\frac{\lambda-l}{2}\phi) \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} \exp(i\frac{\lambda+l}{2}\phi) \\ -\exp(i\frac{\lambda-l}{2}\phi) \end{pmatrix}. \quad (7)$$

It can be shown that λ must be an integer of the same parity as l because of the general properties of angular momentum operators and the requirement that the total wave function must be single-valued. The eigenvectors of \widehat{J}_l are also eigenvectors of the Fourier terms of V_c having the same l value with eigenvalues

$$\Lambda = \pm \sum_{k=0}^{\infty} c_{kl} r^k. \quad (8)$$

This equation also shows that any coupling matrix with a coupling element restricted to a single l value will have cylindrical symmetry

($C_{\infty v}$). Equation (7) can be analyzed with respect to rotation of ϕ by 2π and for $\lambda = 0$, this refers to the purely electronic contribution to the wave function and thus to the electronic angular momentum. For odd l , the electronic contribution to the eigenfunctions changes sign and this implies a nontrivial geometric phase of the nuclear part of the wave function to compensate. This is not the case for even l values. These two cases we will refer to as *fermionic* (odd l) and *bosonic* (even l) behavior, respectively.

C. General derivation of diabatic models

We consider a diabatic $E \otimes e$ model for any system of C_{nv} symmetry, where the highest n -fold rotational axis is denoted by C_n . The general symmetry properties of such systems have been presented in Sec. II A. With the previously defined coordinates and states, a general $E_\alpha \otimes e_\beta$ diabatic model can be constructed by expanding the electronic Hamiltonian in the diabatic state basis as

$$\begin{aligned} \widehat{H} &= \sum_{j,k \in \{+,-\}} |\psi_{\alpha j}\rangle \langle \psi_{\alpha j} | \widehat{H} | \psi_{\alpha k}\rangle \langle \psi_{\alpha k} | \\ &= \sum_{j,k \in \{+,-\}} |\psi_{\alpha j}\rangle D_{jk}^\alpha(Q_{\beta+}, Q_{\beta-}) \langle \psi_{\alpha k} |. \end{aligned} \quad (9)$$

The diabatic PES matrix elements can be expanded to arbitrary order as

$$D_{jk}^\alpha(Q_{\beta+}, Q_{\beta-}) = \sum_{\substack{p,q \\ p+q=m}} \frac{1}{m!} c_{jkpq}^\alpha Q_{\beta+}^p Q_{\beta-}^q, \quad (10)$$

and the invariance condition under symmetry transformations allows us to determine the nonvanishing c_{jkpq}^α terms.

In particular, the invariance when considering the \widehat{C}_n rotational transformation

$$\begin{aligned} |\psi_{\alpha\pm}\rangle (Q_{\beta+})^p (Q_{\beta-})^q \langle \psi_{\alpha\pm} | \\ &= |\widehat{C}_n \psi_{\alpha\pm}\rangle (\widehat{C}_n Q_{\beta+})^p (\widehat{C}_n Q_{\beta-})^q \langle \widehat{C}_n \psi_{\alpha\pm} | \\ &= e^{2i\pi \frac{\beta(p-q)}{n}} |\psi_{\alpha\pm}\rangle (Q_{\beta+})^p (Q_{\beta-})^q \langle \psi_{\alpha\pm} |, \end{aligned} \quad (11a)$$

$$\begin{aligned} |\psi_{\alpha+}\rangle (Q_{\beta+})^p (Q_{\beta-})^q \langle \psi_{\alpha-} | \\ &= |\widehat{C}_n \psi_{\alpha+}\rangle (\widehat{C}_n Q_{\beta+})^p (\widehat{C}_n Q_{\beta-})^q \langle \widehat{C}_n \psi_{\alpha-} | \\ &= e^{2i\pi \frac{\beta(p-q)-2\alpha}{n}} |\psi_{\alpha+}\rangle (Q_{\beta+})^p (Q_{\beta-})^q \langle \psi_{\alpha-} | \end{aligned} \quad (11b)$$

implies conditions on both the diagonal and off-diagonal matrix elements. For the diagonal elements of the diabatic matrix, denoted as $V_n(Q_{\beta+}, Q_{\beta-})$, the resulting condition

$$[\beta(p-q)] \bmod n = 0 \quad (12)$$

is independent of α . By contrast, the condition for nonvanishing off-diagonal elements, denoted by $P_{n\alpha}(Q_{\beta+}, Q_{\beta-})$,

$$[\beta(p-q) - 2\alpha] \bmod n = 0 \quad (13)$$

depends on both α and β . The diabatic matrix potential for $E_\alpha \otimes e_\beta$ is thus

$$D_{n\alpha}(Q_{\beta+}, Q_{\beta-}) = V_n(Q_{\beta+}, Q_{\beta-}) \mathbf{1} + \begin{pmatrix} 0 & P_{n\alpha}(Q_{\beta+}, Q_{\beta-}) \\ P_{n\alpha}^*(Q_{\beta+}, Q_{\beta-}) & 0 \end{pmatrix}. \quad (14)$$

One finds that $V_n(Q_{\beta+}, Q_{\beta-})$ is a totally symmetric offset which does not contribute to the coupling.

A more detailed analysis can be performed on the individual terms of the $P_{n\alpha}(Q_{\beta+}, Q_{\beta-})$ functions. In the expansion of $P_{n\alpha}$, one can first factorize out totally symmetric factors $(Q_{\beta+} Q_{\beta-})^j$ which do not change the symmetry properties as evidenced from a closer look at Eq. (13). This yields

$$P_{n\alpha}(Q_{\beta+}, Q_{\beta-}) = \sum_{p,q} c_{pq}^{n\alpha} Q_{\beta+}^p Q_{\beta-}^q = \sum_{p,q} c_{pq}^{n\alpha} (Q_{\beta+} Q_{\beta-})^j Q_{\beta+}^{p-j} Q_{\beta-}^{q-j}. \quad (15)$$

For each term of the double sum, j can always be chosen to equal either p or q , respectively, $p = q$ being excluded by the constants given in Eq. (13). This allows us to rewrite Eq. (15) as

$$P_{n\alpha}(Q_{\beta+}, Q_{\beta-}) = \sum_{q=0}^{\infty} \sum_{p=q+1}^{\infty} c_{pq}^{n\alpha} (Q_{\beta+} Q_{\beta-})^q Q_{\beta+}^{p-q} + \sum_{p=0}^{\infty} \sum_{q=p+1}^{\infty} c_{pq}^{n\alpha} (Q_{\beta+} Q_{\beta-})^p Q_{\beta-}^{q-p} \quad (16a)$$

TABLE I. Lowest monomials of $P_{n\alpha}(Q_{\beta+}, Q_{\beta-})$ for $E_\alpha \otimes e_\beta$ coupling potential of a C_{nv} system. The ‘‘GPE’’ indicates the presence/absence of a nontrivial geometric phase effect in the ‘‘inner region’’ ($\lambda = 0$) (see text).

n	α	β	Monomials				GPE	Radius of ‘‘inner region’’
3	1	1	Q	Q_+^2	Q_-^4	Q_+^5	Yes	$f_{-1}^{31}(0) - f_{+2}^{31}(0) r = 0$
4	1	1	Q_+^2	Q_-^2	Q_+^6	Q_-^6	No	
5	1	1	Q_{1+}^2	Q_{1-}^3	Q_{1+}^7	Q_{1-}^8	No	$f_{+2}^{51}(0) - f_{-3}^{51}(0) r = 0$
	1	2	Q_{2+}	Q_{2-}^4	Q_{2+}^6	Q_{2-}^9	Yes	$f_{+1}^{51}(0) + f_{+1}^{51'}(0) r^2 - f_{-4}^{51}(0) r^3 = 0$
	2	1	Q_1	Q_{1+}^4	Q_{1-}^6	Q_{1+}^9	Yes	$f_{-1}^{52}(0) + f_{-1}^{52'}(0) r^2 - f_{+4}^{52}(0) r^3 = 0$
	2	2	Q_{2+}^2	Q_{2-}^3	Q_{2+}^7	Q_{2-}^8	No	$f_{+2}^{52}(0) - f_{-3}^{52}(0) r = 0$
6	1	1	Q_{1+}^2	Q_{1-}^4	Q_{1+}^8	Q_{1-}^{10}	No	
	1	2	Q_{2+}	Q_{2-}^2	Q_{2+}^4	Q_{2-}^5	Yes	
	2	1	Q_{1-}	Q_{1+}^4	Q_{1-}^8	Q_{1+}^{10}	No	
	2	2	Q_2	Q_{2+}^4	Q_{2-}^4	Q_{2+}^5	Yes	
7	1	1	Q_{1+}^2	Q_{1-}^5	Q_{1+}^9	Q_{1-}^{12}	No	$f_{+2}^{71}(0) + f_{+2}^{71'}(0) r^2 - f_{-5}^{71}(0) r^3 = 0$
	1	2	Q_{2+}	Q_{2-}^6	Q_{2+}^8	Q_{2-}^{13}	Yes	$f_{+1}^{71}(0) + f_{+1}^{71'}(0) r^2 + f_{+1}^{71''}(0) r^4 - f_{-6}^{71}(0) r^5 = 0$
	1	3	Q_{3+}^3	Q_{3-}^4	Q_{3+}^{10}	Q_{3-}^{11}	Yes	$f_{+3}^{71}(0) - f_{-4}^{71}(0) r = 0$
	2	1	Q_{1-}^3	Q_{1+}^4	Q_{1-}^{10}	Q_{1+}^{11}	Yes	$f_{-2}^{72}(0) - f_{+3}^{72}(0) r = 0$
	2	2	Q_{2+}^2	Q_{2-}^5	Q_{2+}^9	Q_{2-}^{12}	No	$f_{-2}^{72}(0) + f_{+2}^{72'}(0) r^2 - f_{-5}^{72}(0) r^3 = 0$
	2	3	Q_3	Q_{3+}^6	Q_{3-}^8	Q_{3+}^{13}	Yes	$f_{-1}^{72}(0) + f_{-1}^{72'}(0) r^2 + f_{-1}^{72''}(0) r^4 - f_{+6}^{72}(0) r^5 = 0$
	3	1	Q_1	Q_{1+}^6	Q_{1-}^8	Q_{1+}^{13}	Yes	$f_{-1}^{73}(0) + f_{-1}^{73'}(0) r^2 + f_{-1}^{73''}(0) r^4 - f_{+6}^{73}(0) r^5 = 0$
	3	2	Q_{2+}^3	Q_{2-}^4	Q_{2+}^{10}	Q_{2-}^{11}	Yes	$f_{+3}^{73}(0) - f_{-4}^{73}(0) r = 0$
	3	3	Q_{3+}^2	Q_{3-}^5	Q_{3+}^9	Q_{3-}^{12}	No	$f_{+2}^{73}(0) + f_{+2}^{73'}(0) r^2 - f_{-5}^{73}(0) r^3 = 0$
8	1	1	Q_{1+}^2	Q_{1-}^6	Q_{1+}^{10}	Q_{1-}^{14}	No	
	1	2	Q_{2+}	Q_{2-}^3	Q_{2+}^5	Q_{2-}^7	Yes	
	1	3	Q_{3-}	Q_{3+}^6	Q_{3-}^{10}	Q_{3+}^{14}	No	
	2	1	Q_{1+}^4	Q_{1-}^4	Q_{1+}^{12}	Q_{1-}^{12}	No	
	2	2	Q_{2+}^2	Q_{2-}^2	Q_{2+}^6	Q_{2-}^6	No	
	2	3	Q_{3+}^4	Q_{3-}^4	Q_{3+}^{12}	Q_{3-}^{12}	No	
	3	1	Q_{1-}	Q_{1+}^5	Q_{1-}^{10}	Q_{1+}^{14}	No	
	3	2	Q_2	Q_{2+}^3	Q_{2-}^5	Q_{2+}^7	Yes	
	3	3	Q_{3+}^2	Q_{3-}^6	Q_{3+}^{10}	Q_{3-}^{14}	No	

$$= \sum_{k=1}^{\infty} Q_{\beta+}^k \underbrace{\sum_{q=0}^{\infty} c_{(q+k)q}^{n\alpha} (Q_{\beta+} Q_{\beta-})^q}_{f_{+k}^{n\alpha}(Q_{\beta+}, Q_{\beta-})} + \sum_{k=1}^{\infty} Q_{\beta-}^k \underbrace{\sum_{p=0}^{\infty} c_{p(p+k)}^{n\alpha} (Q_{\beta+} Q_{\beta-})^p}_{f_{-k}^{n\alpha}(Q_{\beta+}, Q_{\beta-})} \quad (16b)$$

$$= \sum_{k=1}^{\infty} (f_{+k}^{n\alpha}(Q_{\beta+}, Q_{\beta-}) Q_{\beta+}^k + f_{-k}^{n\alpha}(Q_{\beta+}, Q_{\beta-}) Q_{\beta-}^k). \quad (16c)$$

The $Q_{\beta+}^k$ and $Q_{\beta-}^k$ terms of Eq. (16b) are referred to as monomials in the following. The characteristic properties of the potential considered in the present work do only depend on these monomials but not on the totally symmetric prefactors. For all C_{nv} groups with $n = 3, \dots, 8$ and all combinations of E_{α} and e_{β} , the four lowest orders of such monomials are given in Table I. If higher orders are desired, they can be derived easily following the above general rules. Higher orders beyond the typically used lowest two orders are mainly of interest to improve the quality of PES models for explicit systems (like NO_3 studied by us) in regions farther away from the coordinate origin. They do not change the basic properties of the PES but allow us to treat large amplitude motions of the nuclei more accurately. Equation (16c) gives the expansion in the complex representation of the coordinates $Q_{\beta+}$ and $Q_{\beta-}$ in which $\phi_{\beta} = \beta \cdot \phi$. It is convenient to contract the totally symmetric parts in the expression (16c) resulting in

$$P_{n\alpha}(r_{\beta}, \phi_{\beta}) = \sum_k \left[F_{+k}^{n\alpha}(r_{\beta}) e^{ik\phi_{\beta}} + F_{-k}^{n\alpha}(r_{\beta}) e^{-ik\phi_{\beta}} \right]. \quad (17)$$

For practical applications, the functions $F_{\pm k}^{n\alpha}(r_{\beta})$ can be modeled by any totally symmetric and continuous function of r_{β} . The power expansion of the functions $F_{\pm k}^{n\alpha}(r_{\beta})$ will render Eqs. (15) and (17) equivalent. This form is particularly convenient for the analytic investigation of the properties of the diabatic Hamiltonians as will be seen below.

For the numerical models of $C_{nv} E_{\alpha} \times e_{\beta}$ systems studied here, we limit the expansion of $P_{n\alpha}(Q_{\beta+}, Q_{\beta-})$ to the first two monomials, as given in Table I. The use of at least two monomials is crucial in order to obtain PESs with the symmetry induced multiple equivalent potential wells on the lower adiabat as is characteristic for such Jahn-Teller systems. The order of the expansion used for the diagonal average potential $V_n(Q_{\beta+}, Q_{\beta-})$ is chosen such that the resulting adiabatic potential surfaces are bound. Table I also provides the next two higher monomials for the expansion. In addition, the table provides the information about the presence (or absence) of a geometric phase effect (GPE) as deduced from Eq. (7) for $\lambda = 0$ and each n, α, β . The last column in Table I presents formulas to determine the radius at which the additional conical intersections appear and may change the behavior of the system. This is discussed in detail in Sec. III.

The limitation to the first two coupling monomials in the following is entirely due to practical reasons for the analytical and numerical analysis. Adding higher order terms is only needed to improve the PES accuracy for a real system and

will not (or rather is not supposed to) change the fundamental physics. It should also be kept in mind that adding more and more higher order terms in a Taylor expansion might lead to numerical problems and unphysical behavior. For this reason, a new strategy has been developed by us recently, which is based entirely on the lowest coupling orders and artificial neural networks.²³

III. GEOMETRIC PHASE

The knowledge of the general and generic diabatic models for C_{nv} systems now allows us to analyze their properties. One of the most important and widely discussed properties is the geometric phase effect, an effect usually considered to be induced by the existence of a conical intersection. The nontrivial geometric phase in vibronic coupling systems is just a special case of a Berry phase.²⁴ It is often assumed that the “ $E \otimes e$ ” JT coupling generally induces such a conical intersection and thus the geometric phase effect. However, it will be shown in the following that the existence of a geometric phase effect depends on the type of the $E \otimes e$ system, namely, on n, α , and β .

For most of the systems with the above given diabatic models, the coordinate region of interest usually is what we refer to as the “inner region” where the JT coupling induces multiple equivalent potential wells on the lower adiabatic PES. In most cases, the number of equivalent wells is equal to the order n of the principal C_n axis. For $n \geq 6$, there are also cases with only n/m equivalent minima due to the symmetry properties of the e_{β} coordinates (see Sec. IV). For cases dominated by the inner region, it is relevant to look more closely at the PES for short distances from the symmetry point. Therefore, an asymptotic expansion and analysis in the vicinity of the intersection. The totally symmetric part of the potential can be omitted for the following discussion and will not be considered in this derivation.

In the following, we will illustrate the effects of a geometric phase using the above described generic diabatic models in more detail. It is convenient to transform back the electronic component basis to the real space, namely, $\{|\psi_{\alpha x}\rangle, |\psi_{\alpha y}\rangle\}$, in order to represent the effects graphically. To this end, the eigenvectors of the diabatic PES matrix are determined by solving the following equation:

$$\begin{pmatrix} W^{\alpha}(Q_{\beta+}, Q_{\beta-}) - \Lambda & Z^{\alpha}(Q_{\beta+}, Q_{\beta-}) \\ Z^{\alpha}(Q_{\beta+}, Q_{\beta-}) & -W^{\alpha}(Q_{\beta+}, Q_{\beta-}) - \Lambda \end{pmatrix} \mathbf{c}^{\alpha} = \mathbf{0}, \quad (18)$$

where $W^{\alpha}(Q_{\beta+}, Q_{\beta-}) = \text{Re}(P_{n\alpha}(Q_{\beta+}, Q_{\beta-}))$ and $Z^{\alpha}(Q_{\beta+}, Q_{\beta-}) = \text{Im}(P_{n\alpha}(Q_{\beta+}, Q_{\beta-}))$. The n dependence of W and Z has been dropped for better readability. The resulting eigenvectors of this equation can be expressed in the form

$$\mathbf{c}_{\pm}^{\alpha}(Q_{\beta+}, Q_{\beta-}) = \frac{1}{\sqrt{2(1 \mp W^{\alpha}(Q_{\beta+}, Q_{\beta-}))}} \begin{pmatrix} \pm 1 - W^{\alpha}(Q_{\beta+}, Q_{\beta-}) \\ Z^{\alpha}(Q_{\beta+}, Q_{\beta-}) \end{pmatrix} \quad (19)$$

corresponding to the eigenvalues

$$\Lambda = \pm \sqrt{[W^{\alpha}(Q_{\beta+}, Q_{\beta-})]^2 + [Z^{\alpha}(Q_{\beta+}, Q_{\beta-})]^2}. \quad (20)$$

A relevant quantity to plot in order to visualize the role of the orders l in the coupling part $P_{n\alpha}$ is the vector field of the real valued eigenvectors in the plane of the real coordinates $Q_{\beta x}$ and $Q_{\beta y}$. Figure 1 presents the obtained vector fields for a selection of point groups with single principal axis C_n of order $n = 3, 5, 7$ for α, β combinations featuring a leading coupling term of odd order. As already stated, the numerical model systems are limited to the first two leading monomials in the expansion Eq. (17), as listed in Table I.

Figures 1(a)–1(c) present the vector field in the region of low energy around the origin where n equivalent minima on the lower adiabatic PES sheet are located. The characteristic feature of all of these plots is the existence of one or more slits where the vectors change orientation (“sign”) abruptly. These slits are marked clearly by red lines for the convenience of the reader. The case $n = 3$ is the well-known standard case studied for decades. It clearly shows one slit starting from the origin, confirming the established existence of

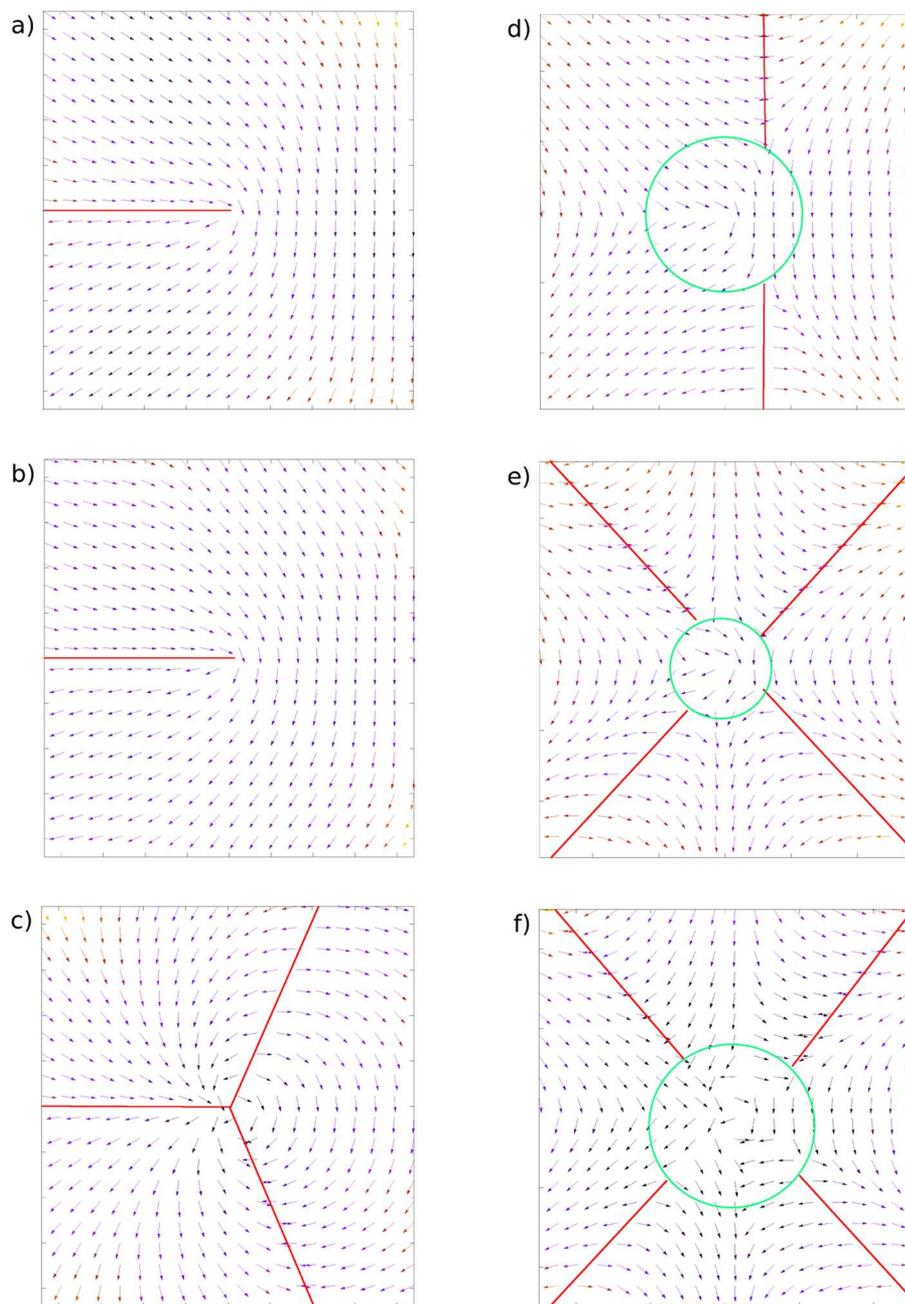


FIG. 1. Vector fields of the eigenvectors corresponding to the lower adiabatic PES plotted in the plane of the $Q_{\beta x}$ and $Q_{\beta y}$ coordinates. (a)–(c) show the inner region for C_{3v} , C_{5v} ($E_1 \otimes e_2$), and C_{7v} ($E_1 \otimes e_3$), respectively. (d)–(f) present the outer region for the same three cases. The green circles indicate the radius separating inner from outer regions.

a geometric phase. The same kind of behavior is found for the case of $n = 5$ for $E_1 \otimes e_2$ ($E_2 \otimes e_1$). Finally, the case $E_1 \otimes e_3$ for $n = 7$ is presented because this is a case in which three slits and a corresponding geometric phase are observed. By contrast, two slits are found for $n = 4$ and C_{5v} ($E_1 \otimes e_1$ or $E_2 \otimes e_2$), respectively, with the consequence that there is no geometric phase in these cases. The corresponding vector fields are presented in Fig. 2. It turns out that the number of slits observed near the origin in the real representation is always equal to the power of the lowest coupling term as is shown below.

To this end, the real valued eigenvectors are analyzed for the limit of small distances from the origin (high symmetry point). In this asymptotic analysis, only the term of the lowest order remains relevant in the expansion of the coupling matrix and therefore the eigenvectors corresponding to the eigenvalues

$$\Lambda = \pm r^k \quad (21)$$

read

$$\mathbf{c}_{\pm}(\phi) = \frac{1}{\sqrt{2(1 \mp \cos(m\phi))}} \begin{pmatrix} \pm 1 - \cos(m\phi) \\ \sin(m\phi) \end{pmatrix}, \quad (22)$$

where

$$m = \beta k. \quad (23)$$

In the following, only the eigenvector \mathbf{c}_- is considered, but the treatment of \mathbf{c}_+ is completely equivalent. We note that the normalization factor becomes singular for $m\phi = 2\pi j + \pi$, $j = 0, 1, \dots, m - 1$. The behavior of the eigenvectors in the vicinity of these singularities is determined by the limit

$$\lim_{\epsilon \rightarrow 0} \mathbf{c}_- \left(\frac{2\pi j + \pi}{m} \pm \epsilon \right) = \lim_{\epsilon \rightarrow 0} \frac{1}{\sqrt{2(1 + \cos(2\pi j + \pi \pm m\epsilon))}} \times \begin{pmatrix} -1 - \cos(2\pi j + \pi \pm m\epsilon) \\ \sin(2\pi j + \pi \pm m\epsilon) \end{pmatrix}. \quad (24)$$

This limit is evaluated by a Taylor expansion up to first order in the denominator and in the numerator, which yields

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \mathbf{c}_- \left(\frac{2\pi j + \pi}{m} \pm \epsilon \right) &= \lim_{\epsilon \rightarrow 0} \frac{1}{\sqrt{2} m^2 \epsilon^2} \begin{pmatrix} 0 \\ \pm m\epsilon \end{pmatrix} \\ &= \lim_{\epsilon \rightarrow 0} \pm \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned} \quad (25)$$

This implies a discontinuity for each angle $\phi = (2\pi j + \pi)/m$, and thus a total of m slits, with m being the order of the coupling term.

While this explains the behavior of the vector fields shown in Figs. 1 and 2 for small distances r from the origin, it fails to describe the $r \gg 1$ outer region. Figures 1(d)–1(f) and 2(b) show the same five systems considered before but now with focus on the region for larger radius $r \gg 1$ and thus for higher energies. The radius at which the additional conical intersections occur is indicated by green circles in Figs. 1 and 2. The $n = 3$ case obviously has two slits starting at this radius whereas the slit starting from the origin vanishes. By contrast, no change is seen for $n = 4$ and the $E_1 \otimes e_1$ ($E_2 \otimes e_2$) case of $n = 5$. The two remaining cases, $n = 5$ $E_1 \otimes e_2$ ($E_2 \otimes e_1$) and $n = 7$ $E_1 \otimes e_3$, result in 4 slits in the outer region. The explanation of this behavior can be gained from a closer look at the two adiabatic PES sheets for these systems. We present two views of the adiabatic PESs plotted in the two-dimensional space of Q_x and Q_y for $n = 3$ in Fig. 3. It becomes apparent from Fig. 3(a) that the PESs fulfill the 3-fold symmetry and invariance properties as required and enforced by construction of the model. In Fig. 3(b), the representation of the bound PESs is restricted to $Q_x > 0$. The well-known conical intersection is seen at the origin and marked by a green circle. An additional conical intersection is clearly visible for larger r and higher energy, also marked by a green circle. Due to the 3-fold symmetry of the PESs, there are three equivalent conical intersections of the latter kind. The formation and disappearance of slits in the vector fields presented in Fig. 1 coincide exactly with these additional conical intersections. Of course, for $n = 3$, the additional conical intersections are known since a long time as was shown in Ref. 6. In the present work, we show that this is a more general feature of C_{nv} systems, at least when modeled in the way described above. Whenever the two used expansion terms are of different orders, this corresponds to the existence of additional conical intersections. The radius r_{CI} where these additional conical intersections occur corresponds to the real positive roots of

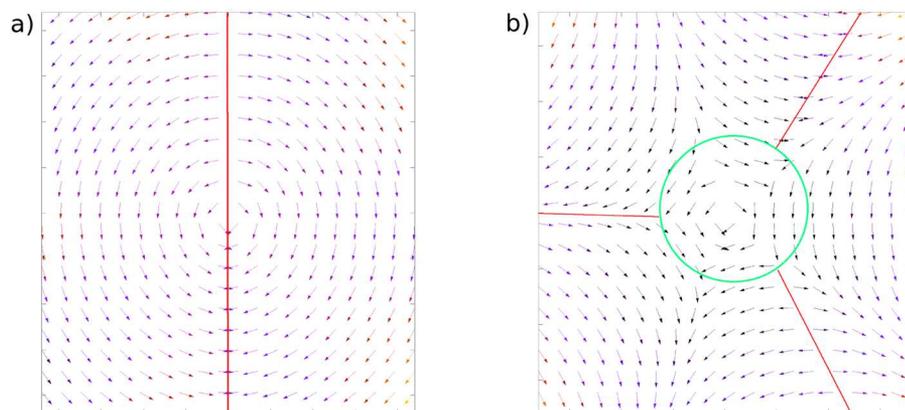


FIG. 2. Vector field of the eigenvectors corresponding to the lower adiabatic PES plotted in the plane of the $Q_{\beta x}$ and $Q_{\beta y}$ coordinates. (a) shows the inner region for C_{5v} ($E_1 \otimes e_1$). (b) presents the outer region for the same case. The green circle indicates the radius separating the inner from outer region.

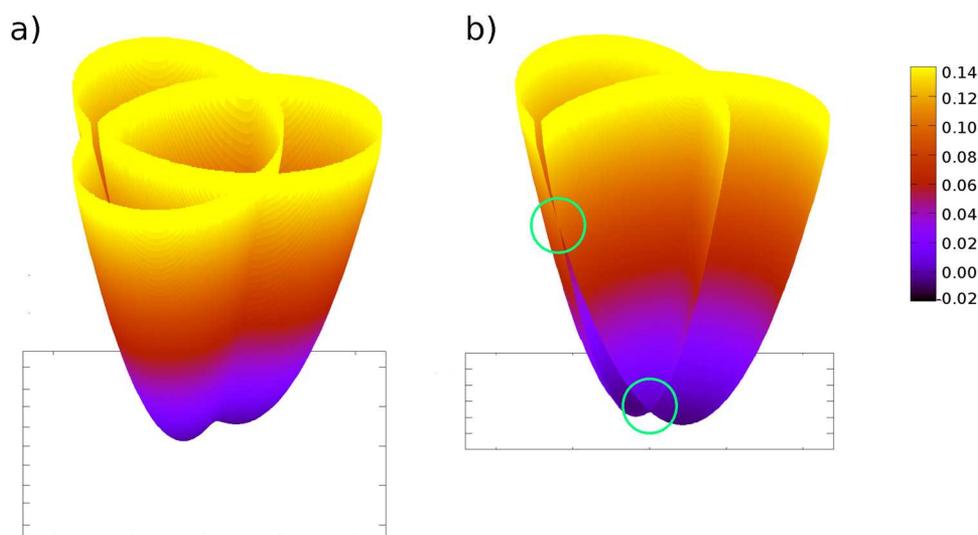


FIG. 3. Adiabatic PESs for $n = 3$ in the two-dimensional space spanned by Q_x and Q_y . (a) illustrates the 3-fold symmetry. (b) presents a view cutting through the origin and clearly showing the conical intersections at the origin and for larger r at high energy (marked by green circles). Color coded energies in arbitrary units.

$P_{n\alpha}(Q_{\beta+}, Q_{\beta-})$. For example, as modeled here for the case $n = 3$, the distance of the additional conical intersection from the symmetry point is simply $r_{CI} = \frac{c_1}{c_2}$ (in agreement with Ref. 6) and for $n = 5$ $E_1 \otimes e_2$ ($E_2 \otimes e_1$), it is $r_{CI} = \sqrt[3]{\frac{c_1}{c_2}}$, where c_1 and c_2 refer to the two expansion coefficients. The radii separating inner and outer regions for all systems treated here can be found approximately by solving for the roots of the functions given in the last column of Table I in which the general totally symmetric expansion functions $f_{k\pm}^\alpha(r_\beta^2)$ [see Eq. (17)] and their derivatives are used.

The existence of additional conical intersections induces two effects. First of all, the additional conical intersections induce local geometric phases, which may play an important role in wave packet propagations. Second, the additional conical intersections will lift the effect of the intersection at the symmetry point (or in general other *inner* intersections). Therefore, the additional conical intersections will change the properties of the vibronic eigenstates. This has been observed before for $n = 3$.^{6,11,12} An odd number of additional conical intersections on a ring around the symmetry point will flip the behavior such that if in the inside of the ring there is no geometric phase, at the outside there will be one and vice versa. An even number of additional conical intersections will not induce such a flip of the properties.

The above discussion gives a clear picture of what happens to the system in the two limiting cases with the wave function either entirely located in the inner or outer region, respectively. However, this clear picture is lost if the wave function extends significantly beyond the dividing line and thus probes the additional conical intersections only partially. In such a case, the impact of the conical intersections and corresponding local geometric phases is not well-defined and the concept of the GPE reaches its limits as a useful concept.

IV. GEOMETRIC PHASE AND SPECTROSCOPY

It has been recognized early on that the existence of a geometric phase has an impact on the vibronic ground state of a Jahn-Teller

system.^{2,3} Model systems with linear and quadratic coupling terms have been shown to have a ground state of E vibronic symmetry in general. An elegant proof for this was given by Ham.⁵ However, it has been shown that the ordering of the first two states switches from E/A (meaning E lower in energy than A) to A/E above a certain value of the quadratic coupling constant.^{11,12} This is due to the fact that for strong quadratic and weak linear coupling, the wave function is not localized in the inner region but extends considerably to the outer region. Thus, it probes four instead of one conical intersection. Furthermore, Ham showed that an alternating pattern should be expected for increasing excitation quanta. We recently found in our investigation of the vibronic eigenstates of the nitrate radical (NO_3) a more complicated and rather intriguing pattern of the vibronic level structure and proposed an explanation for these observations.¹⁴ Here, we generalize these findings beyond C_{3v} and D_{3h} systems.²⁵

As discussed in Sec. III, an n -fold principal axis induces n/m equivalent minima on the lower sheet of the adiabatic PESs for any $E_\alpha \otimes e_\beta$ JT system (m depends on n and β and usually $m = 1$). The local environment around these minima shows a lower symmetry than the symmetry of the full system and if the wells are sufficiently deep the wave function will “localize” (i.e., will have a significantly increased density) in these wells. This is essentially equivalent to the Jahn-Teller theorem stating that the equilibrium geometry is distorted from its highest possible symmetry.²⁶ The molecular vibrations can be conceived as small oscillations of the atoms around their respective equilibrium positions. However, these equilibrium positions refer to one of the equivalent local minima, and the PES in the vicinity of these minima only shows the reduced local symmetry. Therefore, vibrations formally belonging to a degenerate e_β mode in the two dimensional isotropic harmonic oscillator will not degenerate in the local symmetry and thus the corresponding vibronic levels will split. These local vibrations are equivalent for the n/m minima and must be symmetrized according to the global symmetry. This leads to sets of n/m vibronic states for each kind of local vibration, and these

vibronic states transform as irreducible representations of the global symmetry group. If the barriers between the local minima are high enough, the vibronic states form tunneling sets and the corresponding tunneling splittings depend on the well depth and the barrier height between the wells. For more shallow wells or higher excitations, the vibronic level energies may be above the barriers but the general structure of what we call “tunneling sets” or “tunneling multiplets” remains because this is simply a consequence of the global symmetry of the system and of the equivalent potential wells.

This situation can be approximated and analyzed very efficiently with a next neighbor or Hückel model. The n -fold symmetry axis of any C_{nv} point group corresponds to a rotation operator \hat{C}_n transforming one local state residing in one local minimum into the corresponding local state residing in the neighboring PES well. If we further assume that the local states can only interfere with the ones in the neighboring wells, we can write the Schrödinger equation in the next neighbor approximation as

$$\begin{pmatrix} a-E & b & 0 & 0 & \dots & b \\ b & a-E & b & 0 & \dots & 0 \\ 0 & b & a-E & b & \dots & 0 \\ \vdots & & & & \ddots & \vdots \\ b & 0 & 0 & \dots & b & a-E \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (26)$$

Here, a stands for the expectation value of the vibrational Hamiltonian over a purely local eigenstate and b for the integral corresponding to the tunneling interaction with the next neighbor local state. This yields the relation

$$b c_{j-1} + (a-E) c_j + b c_{j+1} = 0, \quad \text{with } 1 < j < n-1. \quad (27)$$

The elements of each eigenvector, c_{j-1} , c_j , and c_{j+1} , are related to each other strictly by symmetry. The resulting n eigenvalues are simply

$$E_k = a + 2b \cos\left(\frac{2\pi k}{n}\right), \quad k = 0, 1, \dots, n-1. \quad (28)$$

For example, for $n = 3$, the eigenvalues would be $a + 2b$ (A), $a - b$ (E), and $b < 0$ for the ground state of systems without the GPE. Thus, the expected level ordering would be (A) below (E). The ordering of the tunneling multiplets only depends on the sign of b , which is a tunneling probability determined from the localized basis functions.

The parameters a and b are determined following the ideas of Wallace developed for the band theory of graphite.²⁷ As an example, the model for C_{3v} is considered, but the results are general and the conclusions for the other point groups are to be obtained by analogy. For the calculation of a and b , the adiabatic potential of the electronic ground state (lower adiabatic PES sheet) is expanded in a Taylor series around a local minimum, using the planar polar coordinates r and ϕ . This yields

$$\begin{aligned} V_-^{ad} &= r^2 - \sqrt{(c_{11}r)^2 + (c_{22}r^2)^2 + 2c_{11}c_{22}r^3 \cos(3\phi)} \\ &= r^2 - (c_{11}r + c_{22}r^2) \sqrt{1 + \frac{2c_{11}c_{22}r}{(c_{11} + c_{22}r)^2} (\cos(3\phi) - 1)}, \end{aligned}$$

which is further approximated by

$$\begin{aligned} V_-^{ad} &\approx r^2 - c_{11}r - c_{22}r^2 - \frac{c_{11}c_{22}r}{c_{11} + c_{22}r} (\cos(3\phi) - 1) \\ &\approx (1 - c_{22}) \left(r - \frac{c_{11}}{2 - 2c_{22}} \right)^2 + \frac{9}{2} \frac{c_{11}c_{22}r^2}{c_{11} + c_{22}r} \phi^2 - \frac{c_{11}^2}{4 - 4c_{22}}. \end{aligned} \quad (29)$$

Defining

$$r_0 = \frac{c_{11}}{2 - 2c_{22}}, \quad (30a)$$

$$\omega_r = \sqrt{2 - 2c_{22}}, \quad (30b)$$

$$I\omega_\phi = \sqrt{9 \frac{c_{11}c_{22}r_0^2}{c_{11} + c_{22}r_0}}, \quad (30c)$$

and

$$V_0 = -\frac{c_{11}^2}{4 - 4c_{22}} \quad (30d)$$

yields the final expansion and definition of the local potential $V^{local}(r, \phi)$ limited to a single well around $\phi = 0$,

$$V_-^{ad} \approx \frac{1}{2} \omega_r (r - r_0)^2 + \frac{1}{2} I\omega_\phi \phi^2 + V_0 \equiv V^{local}(r, \phi). \quad (31)$$

The local Hamiltonian is defined in terms of the local potential as

$$H^{local} = T + V^{local}, \quad (32)$$

where T represents the kinetic energy. This local Hamiltonian corresponds to a two-dimensional harmonic oscillator. The solutions to this system are composed of the well-known one-dimensional harmonic oscillator functions $\Phi_{i, \text{mass}, \text{frequency}}^{HO}(x)$,

$$\Psi_{p_r, p_t}^{local}(r, \phi) = \exp\left(\frac{2i\phi}{3}\right) \Phi_{p_r, 1, \omega_r}^{HO}(r - r_0) \Phi_{p_t, r_0^2, \omega_\phi}^{HO}(\phi). \quad (33)$$

The quantum numbers p_r and p_t correspond to the radial and tangential motion, respectively. In addition, this solution contains a factor $\exp(\frac{2i\phi}{3})$ accounting for the geometric phase directly induced by the fact that C_{3v} is fermionic. When considering a bosonic case, this geometric phase factor is absent. For each combination of p_r and p_t , the three local wave functions $\Psi_{p_r, p_t}^{local}(r, \phi)$, $\Psi_{p_r, p_t}^{local}(r, \phi - \frac{2\pi}{3})$, and $\Psi_{p_r, p_t}^{local}(r, \phi + \frac{2\pi}{3})$ are the solution for each local well, respectively. Each of these triplets of wave functions with the same p_r and p_t forms a basis for a Hückel model like the one given in Eq. (26). The diagonal elements a_{p_r, p_t} are approximated as the local energies,

$$a_{p_r, p_t} = \omega_r \left(p_r + \frac{1}{2} \right) + \omega_\phi \left(p_t + \frac{1}{2} \right) + V_0. \quad (34)$$

The off-diagonal elements of the Hückel matrix are given by the integral

$$b_{p_r, p_t} = \int_0^\infty dr \int_{-\pi}^\pi d\phi \Psi_{p_r, p_t}^{local}(r, \phi + \frac{2\pi}{3}) H \Psi_{p_r, p_t}^{local}(r, \phi). \quad (35)$$

Replacing H by $T + V^{local} + V - V^{local}$ yields

$$\begin{aligned} b_{p_r, p_t} &= a_{p_r, p_t} S_{p_r, p_t} + \int_0^\infty dr \int_{-\pi}^\pi d\phi \Psi_{p_r, p_t}^{local}(r, \phi + \frac{2\pi}{3}) \\ &\quad \times (V - V^{local}) \Psi_{p_r, p_t}^{local}(r, \phi), \end{aligned} \quad (36)$$

where S_{p_r, p_t} is the overlap between two consecutive local wave functions. As usual in the Hückel approximation, this term is neglected such that

$$b_{p_r, p_t} \approx \int_0^\infty dr \int_{-\pi}^\pi d\phi \Psi_{p_r, p_t}^{local} \left(r, \phi + \frac{2\pi}{3} \right) (V - V^{local}) \cdot \Psi_{p_r, p_t}^{local} (r, \phi). \quad (37)$$

Taking into account the geometric phase factors of the local wave functions, the equation simplifies to

$$b_{p_r, p_t} \approx - \int_0^\infty dr \int_{-\pi}^\pi d\phi |\Phi_{p_r, 1, \omega_r}(r - r_0)|^2 \Phi_{p_t, r_0^2, \omega_\phi} \left(\phi + \frac{2\pi}{3} \right) \times \Phi_{p_t, r_0^2, \omega_\phi}(\phi) (V - V^{local}). \quad (38)$$

Because the difference $V - V^{local}$ is always negative, the sign of this integral does only depend on the sign of the product $o = \Phi_{p_t, r_0^2, \omega_\phi} \left(\phi + \frac{2\pi}{3} \right) \Phi_{p_t, r_0^2, \omega_\phi}(\phi)$. The product of these local angular wave functions is positive for even p_t and negative for odd p_t . (This holds strictly at least for the lowest few vibronic levels before the normal mode picture becomes invalid anyway.) The graphical representation given in Fig. 4 clearly shows the link between the sign of b_{p_r, p_t} and the parity of p_t of the local oscillator depending on ϕ . This implies that $b_{p_r, p_t} > 0$ for even p_t and $b_{p_r, p_t} < 0$ for odd p_t . Since the level ordering in each tunneling multiplet is ruled by the sign of b_{p_r, p_t} , we conclude that each tangential excitation inverts the order of the tunneling multiplet with respect to the ordering of the multiplet with one less quanta along ϕ . For C_{3v} , this means that the ground state multiplet is E under A and the multiplet corresponding to the first local tangential excitation is A below E and so on.

The argument for different groups C_{nv} is entirely equivalent keeping in mind that for bosonic cases, the negative sign in Eq. (38) is absent because no geometric phase occurs in this case (for a detailed explanation, see Sec. II B). In a bosonic case, b_{p_r, p_t} is negative for even p_t and positive for odd p_t , while in a fermionic case, b_{p_r, p_t} is positive for even p_t and negative for odd p_t . This explains why fermionic and bosonic cases show opposite behavior regarding the level ordering. The results of the Hückel-like model holds

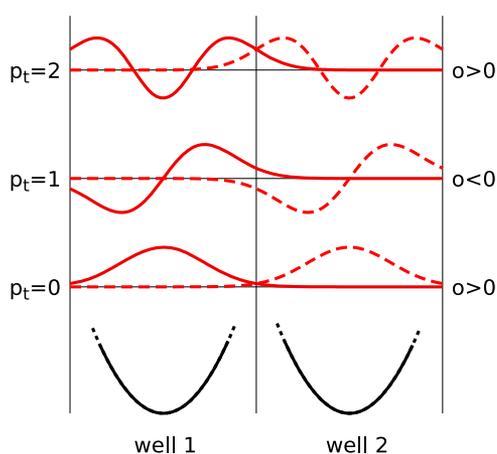


FIG. 4. Local angular wave functions and local potential of two wells in the Hückel model. The product o of the ground state and the second excited state is positive, but the product o of the first excited states is negative.

true even if the energy of the wave function is above the tangential barriers. The tunneling constant b_{p_r, p_t} becomes large, but the general structure is preserved. The simple picture is gradually destroyed if Fermi resonances are present. These occur if the spectrum becomes dense and states of the same symmetry start to be close, which usually happens in all multimode systems beyond the first few quanta of excitation.

In the following, we first will demonstrate this effect by numerical simulations²⁸ for $n = 3$ before discussing higher n . The two leading coupling terms for $n = 3$ are of first and second order (see Table I), and only the combination of the two results in the three equivalent wells. The numerical values used in the following ensure that the densities of the vibronic wave functions outside the inner region are negligible and thus that the inner region characteristics are the only ones relevant for the analysis. The leading first order term induces a geometric phase. The vibrationless ground state of a molecule sitting in one of the wells shows no node of the wave function within the well. However, the global symmetrization over the three wells leads to a vibronic ground state of E symmetry and an excited state of A_2 symmetry together forming a tunneling triplet. It might appear counter-intuitive at first that the vibrationless ground state shows nodes in the wave function. However, these nodes appear between the equivalent wells rather than within the wells. The ordering of E below A as well as the A_2 symmetry of the excited tunneling component is a consequence of the geometric phase effect. The situation is visualized in Fig. 5 where the densities of the vibronic wave functions numerically computed using a prototypical model are plotted.

The densities of the two components of the E state show the expected lack of symmetry with respect to C_3 rotation and cannot be easily interpreted. By contrast, the sum of these two densities shows the rotational symmetry and also demonstrates that the wave function is nodefree within each of the three potential wells. Summing up the E components also allows us to identify the tunneling sets of vibronic states which share the same local vibration as the comparison of Figs. 5(c) and 5(d) demonstrates.

Next, we consider vibrational excitations. It is customary to label the vibrational modes according to the global symmetry. Therefore, it is helpful to map the irreps of the modes in global symmetry onto the corresponding irreps of the local symmetry. The degenerate e modes of global symmetry are of particular interest as these map onto two different irreps in the local symmetry, namely, symmetric (a') or antisymmetric (a''), with respect to the vertical mirror planes σ_v . If the polar coordinate representation of the Hamiltonian equations (4) and (5) is used, the first case would correspond to a radial excitation while the second case would correspond to a tangential excitation. It is obvious that the PES in the vicinity of the equivalent PES minima is rather different in radial and tangential directions, respectively, and this leads to a splitting of the vibronic levels depending on the local excitation direction. Furthermore, the tangential excitation is connected with the tunneling probability (see above), while the radial excitation is not. As a consequence, odd quanta of tangential excitation induce a sign change in the tunneling parameter of the simple next-neighbor model presented above and in our previous study of NO_3 ¹⁴ and thus the order within the tunneling multiplets is inverted. By contrast, any excitation in radial direction as well as even quanta in the tangential mode has no effect on the tunneling parameter thus not affecting the ordering.

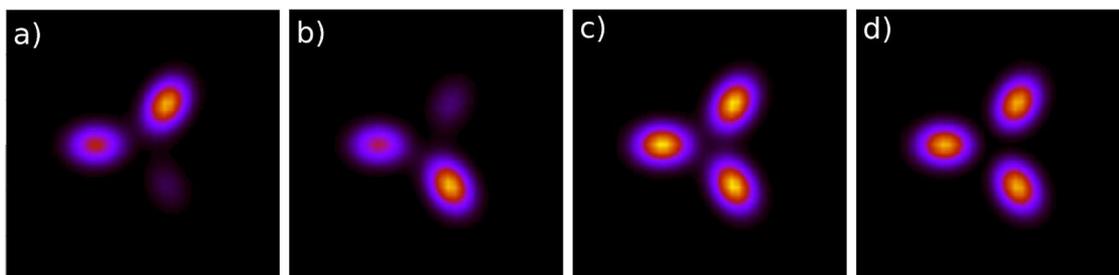


FIG. 5. Density of vibronic state wave functions for $n = 3$ plotted in the Q_x, Q_y plane. (a) First component of the E ground state, (b) second component of the E ground state, (c) sum of first and second components of the E ground state, and (d) A_2 excited tunneling state.

The behavior of the energy levels can be summarized by simple rules. It is customary to interpret vibronic spectra in terms of normal modes transforming as e . In a local isolated deformed oscillator, these normal modes transform like a' and a'' with respect to the local symmetry group C_s . A single radial excitation transforms like a' , and a single tangential excitation transforms like a'' . Symmetrizing these local excitations with respect to the global symmetry group C_{3v} yields the tunneling triplets. This scheme is represented in Fig. 6 for a formally single e excitation. The ordering of the tunneling multiplets depends on the number of tangential quanta of each state. The set ($a' C_s$) corresponding to radial excitation is ordered E/A as in the ground state. By contrast, the set ($a'' C_s$) corresponding to a single tangential excitation is inverted with respect to the ground state pattern and the ordering is A/E . The odd tangential excitation also causes the nondegenerate state to be A_1 rather than A_2 . The construction principle of the vibronic level structure now is clear and can be applied to formally higher excitations in the e modes as well. For example, a double excitation (e)² first leads to a threefold splitting of the levels because the different combinatorial possibilities of the excitations are different for (e_r)², (e_t)², and (e_r)¹(e_t)¹, where e_r and e_t refer to radial and tangential motions, respectively. Each of the three types of local vibrations will result in a tunneling triplet and the level ordering of

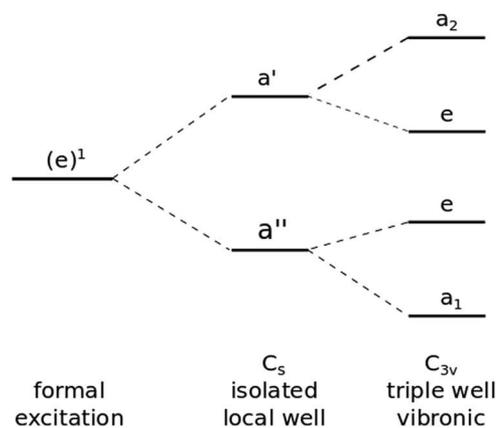


FIG. 6. Vibronic level structure of a formal $(e)^1$ excitation for $n = 3$.

the first two cases (even tangential quanta) will be E/A , while the third case corresponds to odd tangential quanta and thus results in A/E level ordering. This scheme holds for cases for which the lower adiabatic PES sheet shows reasonably pronounced multiwell character.

The behavior of the two different kinds of triplets in C_{3v} (resulting from radial or tangential excitation) with respect to the coupling parameters can be studied systematically, which is illustrated in Fig. 7. Setting linear and quadratic coupling to zero results in an isotropic harmonic oscillator for each of the two electronic components. In this isotropic two dimensional harmonic oscillator, the first few vibronic energy levels are 2-, 4-, 6-, and 8-fold degenerate. Using the irreps of the C_{3v} symmetry group, the ground state is of e symmetry, the first excited state transforms like e , a_1 , and a_2 , the second excited state transforms like e , e , a_1 , a_2 , and so on. The linear coupling term distorts the isotropic harmonic oscillator to the well-known “Mexican hat” shape. As a result, some degeneracies of the isotropic oscillator are lifted. The split energy levels are always twofold degenerate and correspond to the $\pm\lambda/2$ values of the vibronic angular momentum since the cylindrical symmetry is still preserved. As shown in the left half of Fig. 7, the splitting becomes stronger for increasing linear coupling strength. Note, however, that all states can be classified in terms of the C_{3v} irreducible representations and can be grouped into the aforementioned tunneling triplets. For a triplet corresponding to even tangential quanta (black lines in Fig. 7), the e state is always lower in energy than the a_2 state. Examples are the ground state, the states labeled $r^1 t^0$, $r^2 t^0$, and $r^0 t^2$. The state assignments in terms of radial and tangential excitations are given as labels in the right half of Fig. 7. For states corresponding to odd tangential quanta (red lines in Fig. 7), the a_1 part of a triplet is always lower in energy than the corresponding e state. This is true even if the spectrum becomes very dense. When a quadratic term is added, the 2-fold degeneracies of the a_1 and a_2 energies are lifted as shown in the right half of Fig. 7 where the quadratic coupling is increased for a given linear coupling value. In a vibronic coupling model with a linear and a weak quadratic coupling, the splitting of the triplets is quite large as seen near the center of Fig. 7. Therefore, different triplets overlap, but nevertheless, the ordering of each triplet stays invariant (a_2 above e or e above a_1 depending on the number of tangential quanta, i.e., solid below dashed lines for black and dashed below solid for red in Fig. 7). For increasing quadratic coupling, the splitting between the triplet components decreases and its existence becomes more obvious (see near 0.5 to

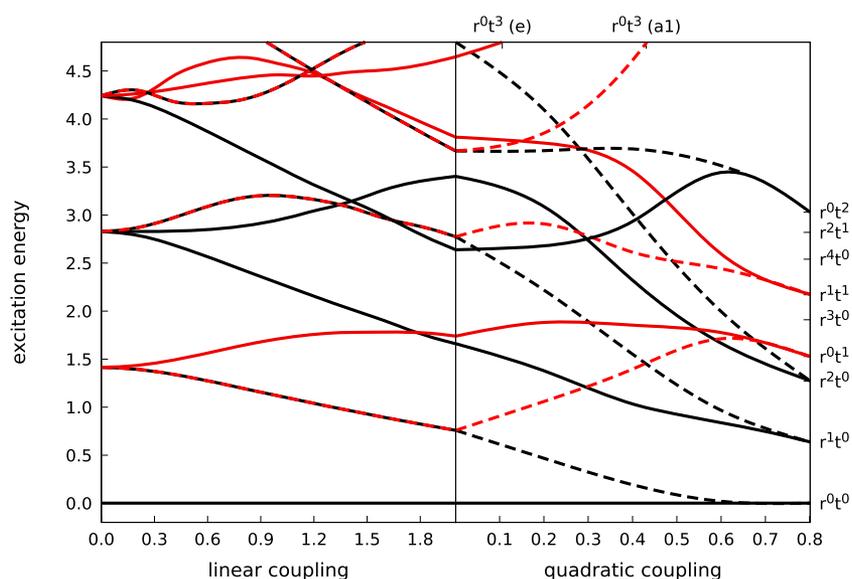


FIG. 7. Vibronic excitation energies of a C_{3v} symmetric JT system. Couplings and energies are given relative to the harmonic constant. For the left half of this figure, the quadratic coupling is set to 0 and the linear coupling varies from 0 to 2. In the right half, the linear coupling is set to 2 and the quadratic coupling is varied from 0 to 0.8. Vibronic states with e symmetry are drawn as solid lines, and vibronic levels of a_1 or a_2 symmetry are marked with dashed lines. Fermionic states are indicated by black lines, and bosonic states are given in red. The states are labeled by their excitation assignment on the right border.

0.6 of quadratic coupling in Fig. 7). The triplets start to be “true” tunneling triplets of the corresponding triple well system. If the quadratic coupling is very strong, the triple well effectively becomes a collection of three isolated oscillators. In this case, tunneling is unlikely and therefore the triplets become essentially degenerate as seen at the right border of Fig. 7. It turns out that for reasonably strong quadratic JT coupling, the Hückel-like picture is more adequate than the picture of a perturbed isotropic harmonic oscillator because the tangential multiwell structure of the potential becomes important.

A confirmation of the predicted patterns is again obtained by numerical model calculations and the analysis of the densities of the vibronic eigenstates. The characteristic densities of the lowest excited vibronic states obtained from a numerical model in which the relevant part of the vibronic wave functions samples mostly the *inner* region are presented in Fig. 8.

The nodal structure of the vibronic wave functions is clearly visible, and radial and tangential excitations can be distinguished easily. The states belonging to a specific tunneling triplet are found by comparing the sum of densities of the E state components with the density of the associated A state. This yields the specific level ordering for each tunneling triplet which is also given in Fig. 8. Figures 8(a), 8(c), 8(e), and 8(h) show states with radial excitations exclusively. The corresponding levels are all ordered E/A_2 like the ground state tunneling triplet because there are no odd tangential quanta. Figures 8(b) and 8(g) show states with tangential excitations exclusively. The states with a single tangential excitation correspond to a level ordering of A_1/E , while the states with a double tangential excitation yield E/A_2 as predicted. Finally, Figs. 8(d) and 8(f) show states with a single tangential node combined with one or two radial nodes, respectively. In both cases, the observed level ordering is A_1/E in agreement with our theoretical prediction. The numerical simulations demonstrate clearly that the theory for the ordering within the tunneling sets outlined above holds for the low lying vibronic states. Note however that the energetic splitting between

different tunneling sets will depend on the specific system. Finally, the above findings hold for low lying states. At higher energies and for real multimode systems, this clear picture will be lost gradually as assignments using excitation quantum numbers will become impossible due to vibronic state interactions (e.g., Fermi resonances). This is a general problem for all multimode vibrational assignments because of the rapidly increasing density of states, which is not a special property of the JT systems studied here. Clear assignments are usually not possible beyond the first few excitation quanta.

Next, the case $n = 4$ is considered for which the two leading coupling terms are of second order, and thus no geometric phase is induced by the PES. Therefore, the vibronic ground state of the system is a nondegenerate A_1 state followed by E and B_1 forming the lowest tunneling quartet. This is indeed observed in our numerical simulations which also include higher excitations. The densities of the first four tunneling sets are displayed in Fig. 9. The first local excitation is found to be a radial one, and thus the same ordering $A_1/E/B_1$ is observed for the vibronic states corresponding to this tunneling set. By contrast, the second single excitation is the tangential one. This results in a tunneling set with a state ordering of $B_2/E/A_2$ which again shows a nondegenerate lowest component. In this case, the inversion of the tunneling multiplet only affects the nondegenerate states exchanging A and B as well as “1” and “2,” respectively. The next set refers to a double excitation in the radial direction for which again the order $A_1/E/B_1$ is expected. Instead, an ordering of $E/A_1/B_1$ is observed in the numerical results. The reason for this apparently is that the four equivalent wells of our numerical example on the lower adiabatic PES sheet are very shallow and a double excitation is already sufficiently above the origin energy at the degeneracy point so that due to the very high density of vibronic states the simple reasoning used above becomes insufficient. State-state interactions (Fermi resonances) can lead to different orderings within the tunneling sets because the density of states strongly increases with energy and thus it is more likely that states of

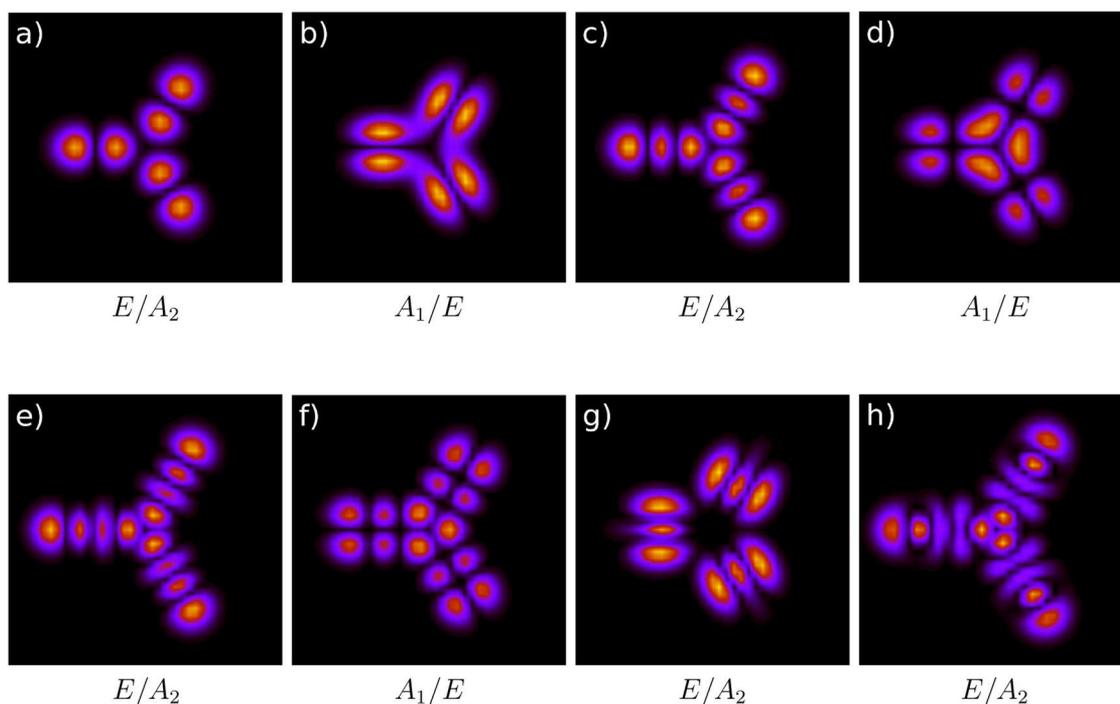


FIG. 8. Density of excited vibronic state wave functions for $n = 3$ plotted in the Q_x, Q_y plane together with the symmetry labels of the energy ordered multiplets. The local excitations are $r^1 t^0$ (a), $r^0 t^1$ (b), $r^2 t^0$ (c), $r^1 t^1$ (d), $r^3 t^0$ (e), $r^2 t^1$ (f), $r^0 t^2$ (g), and $r^4 t^0$ (h).

the same symmetry are in energetic proximity and mix. In fact, the E states show significant mixing with a higher E state in the present example.

The C_{5v} point group is the first case for which two different two-dimensional irreps E_1 and E_2 exist, and thus the situation is considerably more complex than for the previous examples. The leading terms in the diabatic PES models are first order for $E_1 \otimes e_2$ and $E_2 \otimes e_1$ and second order for $E_1 \otimes e_1$ and $E_2 \otimes e_2$. Consequently, the first two combinations will behave fermionic due to the existence of a GPE, while the latter two combinations will have bosonic properties. The effects on the vibronic eigenstates are

displayed in Fig. 10. The first and obvious observation is that the ground state is of A_2 symmetry in the bosonic case [see Fig. 10(a)] and of E symmetry in the fermionic combination [see Fig. 10(e)]. The first two computed excitations in the $E_1 \otimes e_1$ case correspond to radial local excitations, and thus the ordering of the tunneling quintet of $A_2/E_1/E_2$ remains the same. The third local excitation puts a single quantum into the tangential mode and thus leads to an inversion of the tunneling quintet resulting in a level ordering of $E_2/E_1/A_1$. When analyzing the isomorphic case $E_2 \otimes e_2$, the only difference with respect to $E_1 \otimes e_1$ is that the two vibronic E levels E_1 and E_2 are interchanged. As expected, the numerical results for the

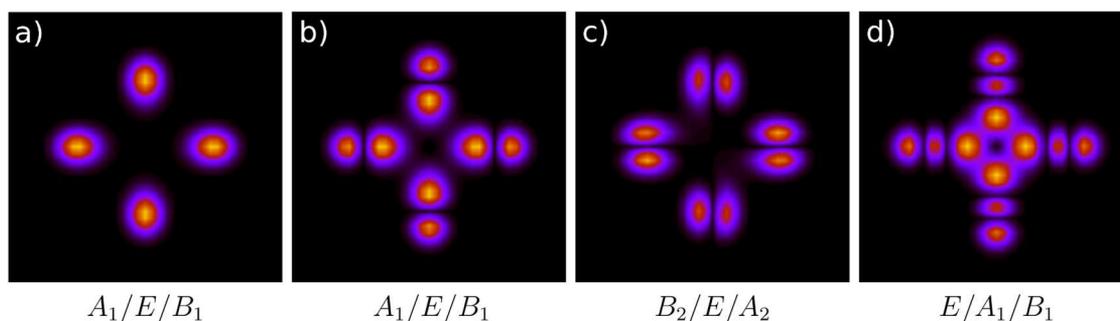


FIG. 9. Density of low lying vibronic state wave functions for $n = 4$ plotted in the Q_x, Q_y plane. The local excitations are $r^0 t^0$ (a), $r^1 t^0$ (b), $r^0 t^1$ (c), and $r^2 t^0$ (d).

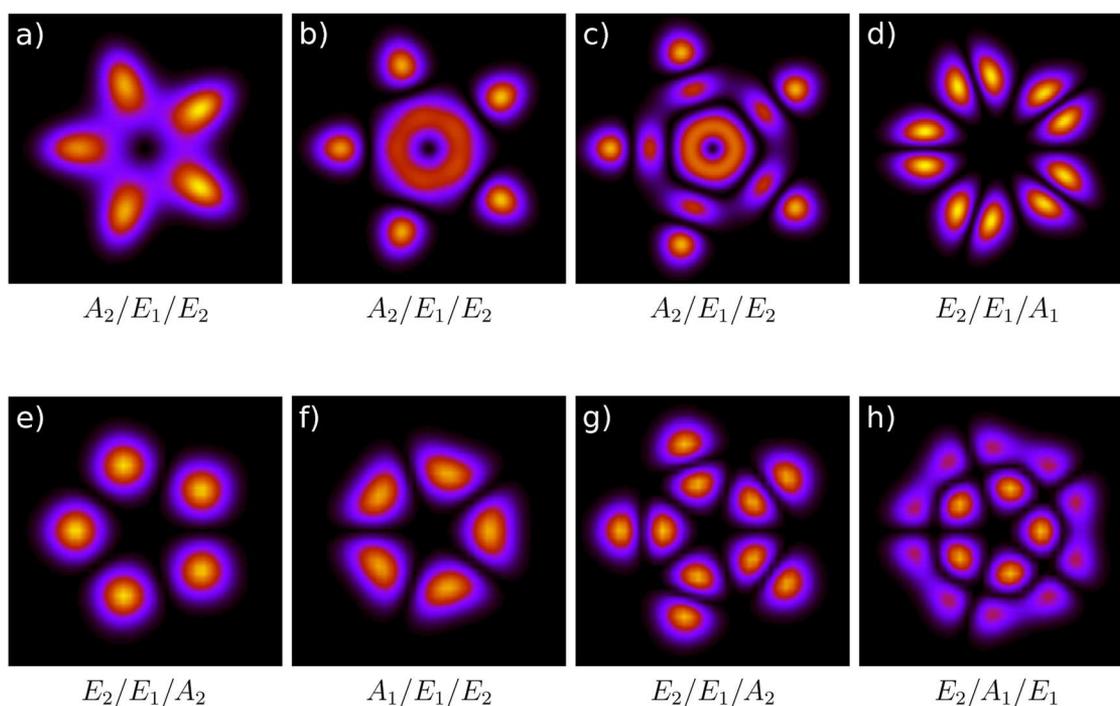


FIG. 10. Density of vibronic state wave functions plotted for $n = 5$ in the Q_x, Q_y plane. (a)–(d) show the $E_1 \otimes e_1$ case; (e)–(h) show the $E_2 \otimes e_1$ case. The local excitations for the $E_1 \otimes e_1$ case are $r^0 t^0$ (a), $r^1 t^0$ (b), $r^2 t^0$ (c), and $r^0 t^1$ (d). The local excitations for the $E_2 \otimes e_1$ case are $r^0 t^0$ (e), $r^0 t^1$ (f), $r^1 t^0$ (g), and $r^1 t^1$ (h).

$E_2 \otimes e_1$ case show a different behavior. The vibronic ground state is E_2 followed by E_1 and A_2 which form the first tunneling set. The first excitation corresponds to a single tangential local excitation and thus inverts the tunneling quintet leading to a level ordering of $A_1/E_1/E_2$. The second excited tunneling set is due to a single radial local excitation and thus has the same level ordering as the ground state, $E_2/E_1/A_2$. The next set of states corresponds to a combination of a single radial and a single tangential excitation in the local wells and thus leads to an inversion again, and would be expected to result in the level ordering $A_1/E_1/E_2$. By contrast, the observed level ordering in our numerical example is $E_2/A_1/E_1$ because the E_2 state is lowered in energy by an interaction with another E_2 level nearby.

The symmetry group C_{6v} provides an interesting new situation because it is the first in the C_{nv} series to contain a non-Abelian subgroup, namely, C_{3v} , which is the most basic fermionic case of a JT system. In addition, C_{6v} has two 2-dimensional representations E_1 and E_2 leading to four different JT cases. $E_1 \otimes e_1$ and $E_2 \otimes e_1$ are bosonic due to a leading coupling term of second order (see Table I). The other two cases $E_1 \otimes e_2$ and $E_2 \otimes e_2$ have a diabatic Hamiltonian isomorphic to that of C_{3v} , and thus the lower adiabatic PES sheet only has three equivalent minima resulting in tunneling triplets for the vibronic states. This is the first case where the n/m equivalent minima correspond to $m > 1$. The e_2 coordinates are symmetric with respect to \widehat{C}_2 rotation, and thus the

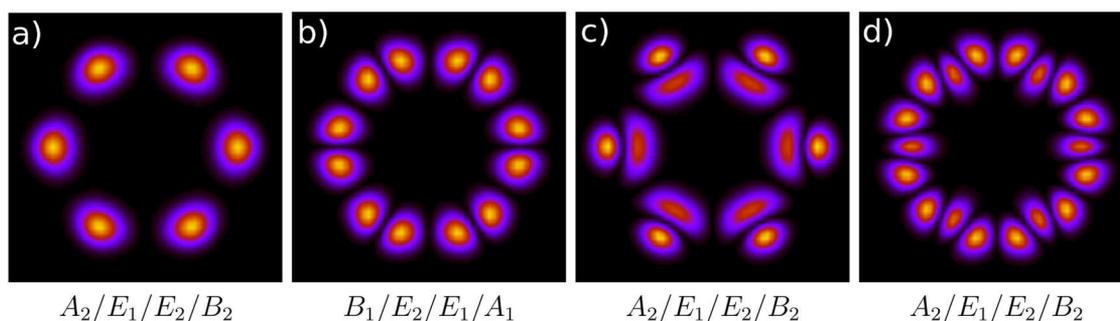


FIG. 11. Density of vibronic state wave functions plotted for $n = 6$ in the Q_x, Q_y plane for the $E_1 \otimes e_1$ case. The local excitations are $r^0 t^0$ (a), $r^0 t^1$ (b), $r^1 t^0$ (c), and $r^0 t^2$ (d).

seemingly missing three potential wells are excluded by the extra symmetry of the coordinates.²⁹ It is also shown easily using the fundamental properties of the C_{6v} irreps that the corresponding vibronic wave functions only can transform as A_1 , A_2 , or E_2 in this case.

The remaining two cases, $E_1 \otimes e_1$ and $E_2 \otimes e_1$, do not show the above idiosyncrasy and are analog to the two bosonic cases for C_{5v} but with the expected six equivalent minima. The corresponding vibronic densities are presented in Fig. 11 together with the level ordering of the tunneling sextets. The vibronic ground state is a nondegenerate A_2 level followed by the two degenerate states E_1

and E_2 and the highest level of the set transforms as B_2 . The densities of these four states do not show any node within the local PES wells. The next tunneling sextet corresponds to a single tangential excitation in the local wells, as shown clearly in Fig. 11(b). This changes the level ordering to $B_1/E_2/E_1/A_1$, similar to the corresponding $E_1 \otimes e_1$ case for C_{5v} [see Fig. 10(d)]. Figure 11(c) shows the next tunneling set corresponding to a single radial excitation in the local wells. Thus, the level ordering is found to be $A_2/E_1/E_2/B_2$. The states corresponding to Fig. 11(d) are due to a double tangential excitation in the local wells and thus again result in the pattern $A_2/E_1/E_2/B_2$.

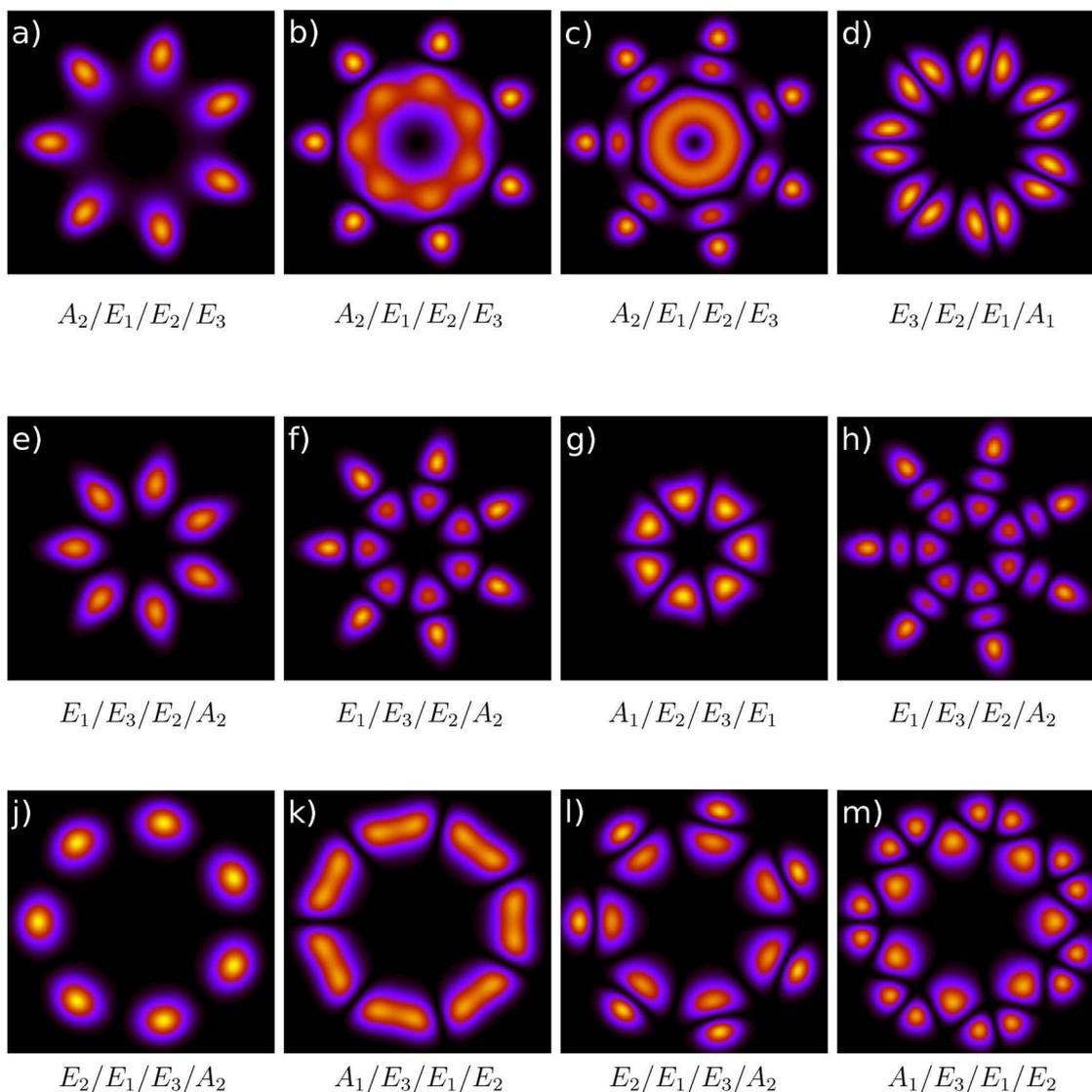


FIG. 12. Density of vibronic state wave functions plotted for $n = 7$ in the Q_x , Q_y plane. (a)–(d) show the $E_1 \otimes e_1$ case, (e)–(h) show the $E_1 \otimes e_2$ case, and (j)–(m) show the $E_1 \otimes e_3$ case. The local excitations for the $E_1 \otimes e_1$ case are $r^0 t^0$ (a), $r^1 t^0$ (b), $r^2 t^0$ (c), and $r^0 t^1$ (d). The local excitations for the $E_1 \otimes e_2$ case are $r^0 t^0$ (e), $r^1 t^0$ (f), $r^0 t^1$ (g), and $r^2 t^0$ (h). The local excitations for the $E_1 \otimes e_3$ case are $r^0 t^0$ (j), $r^0 t^1$ (k), $r^1 t^1$ (l), and $r^1 t^1$ (m).

C_{7v} is the first example for which three different E representations exist, leading to nine distinct $E_\alpha \otimes e_\beta$ combinations. The C_{7v} JT system also shows a new feature in the case of $E_1 \otimes e_3$, namely, a leading term in the diabatic Hamiltonian of third order. The vibronic state densities for the different combinations $E_1 \otimes e_1$, $E_1 \otimes e_2$, and $E_1 \otimes e_3$ are displayed in Fig. 12 for the ground and first few excited states. The $E_1 \otimes e_1$ case with a quadratic leading coupling term does not present a Berry phase. As a consequence, the level ordering for the ground state septet shows a nondegenerate lowest vibronic state (A_2) followed by three doubly degenerate levels ($E_1/E_2/E_3$) [see Fig. 12(a)]. The same ordering is obtained in Figs. 12(b) and 12(c)

which corresponds to a single and a double radial excitation. By contrast, in Fig. 12(d), a single tangential excitation induces an order of $E_3/E_2/E_1/A_1$ for the vibronic levels of the tunneling set. For both $E_1 \otimes e_2$ and $E_1 \otimes e_3$, a geometric phase effect is expected due to the odd power of the leading coupling term (see Table I). As a result, the lowest vibronic levels are doubly degenerate with an E_1 ground state for the case with a linear coupling term $E_1 \otimes e_2$ and a lowest level of E_2 in the case of the cubic leading term for $E_1 \otimes e_3$. In the case of $E_1 \otimes e_2$, the order $E_1/E_3/E_2/A_2$ obtained for the ground state multiplet [Fig. 12(e)] and for the radial excitations [Figs. 12(f) and 12(h)] is modified into $A_1/E_2/E_3/E_1$ when a single tangential excitation is present [Fig. 12(g)]. The third order

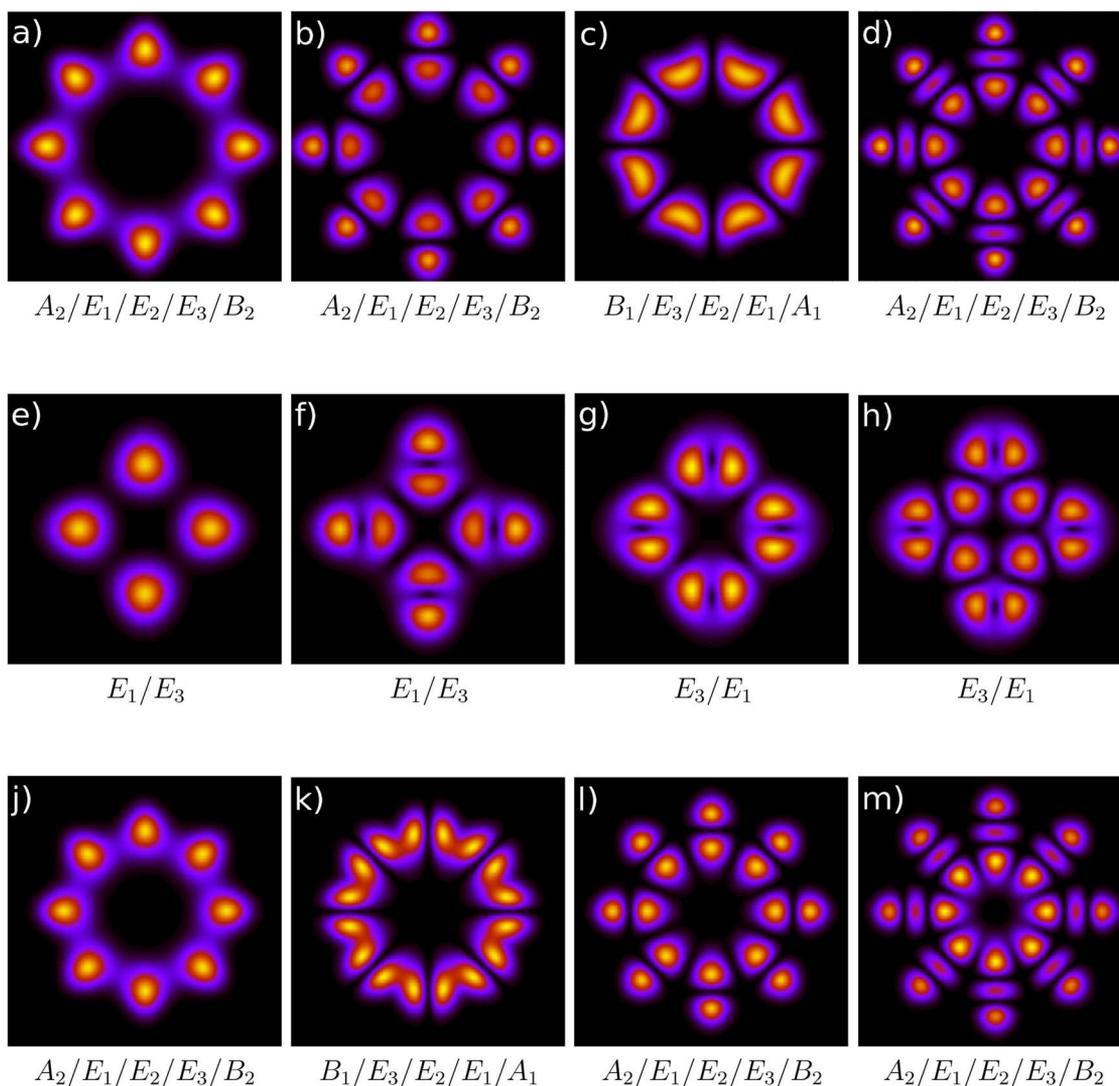


FIG. 13. Density of vibronic state wave functions plotted for $n = 8$ in the Q_x, Q_y plane. (a)–(d) show the $E_1 \otimes e_1$ case, (e)–(h) show the $E_1 \otimes e_2$ case, and (j)–(m) show the $E_2 \otimes e_1$ case. The local excitations for the $E_1 \otimes e_1$ case are $r^0 t^0$ (a), $r^1 t^0$ (b), $r^0 t^1$ (c), and $r^2 t^0$ (d). The local excitations for the $E_1 \otimes e_2$ case are $r^0 t^0$ (e), $r^1 t^0$ (f), $r^0 t^1$ (g), and $r^1 t^1$ (h). The local excitations for the $E_2 \otimes e_1$ case are $r^0 t^0$ (j), $r^0 t^1$ (k), $r^1 t^0$ (l), and $r^2 t^0$ (m).

leading term of $E_1 \otimes e_3$ results in an $E_2/E_1/E_3/A_2$ set of levels for the ground state multiplet and for multiplets corresponding to radial excitations [Figs. 12(j) and 12(l)] and in an $A_1/E_3/E_1/E_2$ level ordering when a single tangential excitation is involved [Figs. 12(k) and 12(m)].

Finally, we present the various cases for C_{8v} for which again three different E representations exist. The nine possible $E_\alpha \otimes e_\beta$ combinations yield four different Hamiltonians of which the $E_2 \otimes e_2$ one is isomorphic to the C_{4v} case. This is due to the fact that C_{4v} is a subgroup of C_{8v} and the situation is similar to the previously seen C_{6v} $E_{1,3} \otimes e_2$ cases. Distinct new cases are found for the isomorphic sets $\{E_{1,3} \otimes e_{1,3}\}$, $\{E_{1,3} \otimes e_2\}$, and $\{E_2 \otimes e_{1,3}\}$. Of these sets, only $\{E_{1,3} \otimes e_2\}$ shows an odd leading coupling term and thus a geometric phase due to the electronic Hamiltonian. The other two sets have even leading coupling terms, and thus the electronic Hamiltonian does not induce a GPE. A symmetry analysis of the vibronic eigenstates is performed using model simulations and the corresponding densities of the first few tunneling sets for the three new, distinct cases are displayed in Fig. 13. The ground state tunneling octet for the bosonic $E_1 \otimes e_1$ case shows a nondegenerate A_2 lowest level, as expected, followed by E_1 , E_2 , E_3 , and B_2 and thus is comparable to the findings for C_{7v} $E_1 \otimes e_1$. The next tunneling set is due to a single radial excitation, clearly visible in Fig. 13(b) and thus has the same level ordering as the ground state set. The first tangential excitation is found for the following tunneling set displayed in Fig. 13(c) and leads to a change of level ordering, which now reads $B_1/E_3/E_2/E_1/A_1$. Finally, a double radial excitation is found and, in agreement with the theoretical prediction, the observed level ordering is back to that of the ground state octet.

In contrast to the $E_1 \otimes e_1$ combination, the Hamiltonian of the $E_1 \otimes e_2$ system has a leading coupling term of odd (first) order and thus the electronic Hamiltonian induces a geometric phase in the ground state multiplet. Furthermore, the e_2 coordinates are invariant under C_2 rotation like in the case of the e_2 coordinates of C_{6v} . This causes the lower adiabatic PES sheet to support only four equivalent minima, and thus the vibronic states show up as tunneling quartets. It is also expected that the ground state should be doubly degenerate, which indeed is found in the numerical example. As seen in Figs. 13(e)–13(h), each tunneling quartet is composed of one E_1 and one E_3 state. The observed ordering is E_1/E_3 for all sets with no or even tangential quanta [Figs. 13(e) and 13(f)] and E_3/E_1 for all cases with odd tangential excitations [Figs. 13(g) and 13(h)].

Last but not least, the $E_2 \otimes e_1$ JT system is investigated for completeness. The leading coupling term of the electronic Hamiltonian is of fourth order, and thus the system is expected to show bosonic character. This is confirmed by the numerical simulation summarized in Figs. 13(j)–13(m). The ground state octet displayed in Fig. 13(j) and the tunneling sets corresponding to radial local excitations exclusively all result in a level ordering of $A_2/E_1/E_2/E_3/B_2$ in analogy to the other bosonic case $E_1 \otimes e_1$ [Figs. 13(a)–13(d)]. The single tangential local excitation visible in Fig. 13(k) leads to a set with a changed level ordering of $B_1/E_3/E_2/E_1/A_1$, also in agreement with the finding for the $E_1 \otimes e_1$ case [Fig. 13(c)].

The above numerical examples clearly confirm the general pattern following from the analytic analysis using the next neighbor approximation.

V. CONCLUSIONS

The quantum dynamics and the geometric phase effect (GPE) are studied in the present work for $E_\alpha \otimes e_\beta$ Jahn-Teller systems of C_{nv} groups with n varying from 3 to 8. The coupled potential energy surfaces (PESs), represented by analytical diabatic 2×2 matrices and derived here for all cases up to $n = 8$, all present at the symmetry coordinate origin a conical intersection or touching point depending on the n , α , and β values. The presence or absence of a GPE around this origin is directly linked to the polynomial order of the leading coupling term. The analytically unraveled characteristics of the coupled surfaces generally show an *inner* and an *outer* region separated by the location of additional conical intersections. These additional intersections do affect the conclusions about the presence or absence of a GPE when considering spatially extended vibronic wave functions for systems for which an odd number of additional conical intersections are present (thus mostly for odd n). By contrast, no change between inner and outer region results for systems with an even number of additional intersections (most of the cases with even n). These effects are studied analytically and are also visualized by plotting the vector fields of the eigenvectors of the diabatic Hamiltonian in the plane of the Cartesian e_β symmetry coordinates. The lines of sign change of the wave functions are clearly visible and are correlated with the order of the leading coupling term of the Hamiltonian. These topological slits are responsible for the presence or absence of the nontrivial topological phase in the adiabatic wave functions.

A second fundamental aspect is the presence of n/m equivalent minima on the lower adiabatic PES sheet for all C_{nv} $E_\alpha \otimes e_\beta$ systems (m depending on n and β). Such a topography leads to the existence of n/m -fold sets of vibronic eigenstates, which are studied in detail in the present work. These sets can be understood in terms of “local vibrations” within each of the local minima and a symmetrization over the equivalent PES wells. The local vibrations can be characterized as tangential or radial motions, respectively, when transformed to polar coordinates. This results in the formation of what we call vibronic tunneling sets, which can be obtained by a simple Hückel type model for weakly interacting local vibrations in a next neighbor approximation. The GPE has a clear impact on the level ordering of the tunneling multiplets, which furthermore depends on the type of local vibration. This effect is due to the sign of the tunneling parameter as shown analytically and changes the level ordering only if the local vibration corresponds to odd excitations of the tangential mode. Since the local environment within each well has a lowered “local symmetry,” the PES is distinctly different in the radial and tangential direction, respectively, leading to an energetic splitting between the corresponding tunneling sets. The GPE, together with the interpretation of local vibrations and symmetrization, is key for understanding the complicated vibronic spectra of real JT systems. The analysis of such spectra could also yield experimental evidence for the impact of the GPE. Any experiment which would provide reliable energy and symmetry data for all levels within a tunneling multiplet would be suitable to confirm the predicted effect of the nontrivial geometric phase on the level ordering.

The theoretical predictions obtained from the above analysis are tested using numerical models restricted to the two lowest

coupling terms for each case and determining the low energy range of the vibronic spectra for all C_{nv} $E_\alpha \otimes e_\beta$ JT systems up to $n = 8$. For all cases studied, the lower adiabatic PES sheet presents n/m equivalent wells as predicted. The signature of the GPE caused by the electronic Hamiltonian is that the vibronic ground state is a degenerate state of E symmetry. This is well known for C_{3v} and general linear vibronic coupling models. As shown here, this is always the case if the leading coupling term is of odd order and the vibronic wave function is dominated by the inner region of the PESs. For $n > 3$, we show that the degeneracy, and more precisely the irreducible representation, of the vibronic ground state does depend on n , α , and β . Furthermore, the numerical results confirm the theoretical interpretation of the tunneling multiplets in terms of symmetrized local vibrations. The different effect of local radial and tangential excitations, respectively, is clearly reflected in a change of the level ordering within each tunneling multiplet. The symmetry analysis of the numerical results provides a full characterization of the computed vibronic states for all cases treated. These results will be of significant help in the understanding of vibronic spectra of real systems.

ACKNOWLEDGMENTS

The authors are grateful for financial support via the PHC/DAAD grant PROCOPE (No. 40442PD) and the FCI (Fonds der Chemischen Industrie). Part of this work was also supported by the Deutsche Forschungsgemeinschaft (DFG).

REFERENCES

- ¹H. C. Longuet-Higgins, U. Öpik, M. H. L. Pryce, and R. A. Sack, *Proc. R. Soc. London, Ser. A* **244**, 1 (1958).
- ²R. Englman, *The Jahn-Teller Effect in Molecules and Crystals* (Wiley-Interscience, New York, 1972).
- ³M. C. M. O'Brien, *Proc. R. Soc. London, Ser. A* **281**, 323 (1964).
- ⁴C. A. Mead, *Chem. Phys.* **49**, 23 (1980).
- ⁵F. S. Ham, *Phys. Rev. Lett.* **58**, 725 (1987).
- ⁶J. W. Zwanziger and E. R. Grant, *J. Chem. Phys.* **87**, 2954 (1987).
- ⁷J. W. Zwanziger, M. Koenig, and A. Pines, *Annu. Rev. Phys. Chem.* **41**, 601 (1990).
- ⁸J. Schön and H. Köppel, *J. Chem. Phys.* **103**, 9292 (1995).
- ⁹C. A. Mead, *Rev. Mod. Phys.* **64**, 51 (1992).
- ¹⁰*Conical Intersections: Electronic Structure, Dynamics and Spectroscopy*, edited by W. Domcke, D. R. Yarkony, and H. Köppel (World Scientific, Singapore, 2004).
- ¹¹H. Koisumi and I. B. Bersuker, *Phys. Rev. Lett.* **83**, 3009 (1999).
- ¹²K. Pae and V. Hizhnyakov, *J. Chem. Phys.* **147**, 084107 (2017).
- ¹³W. Eisfeld, O. Vieuxmaire, and A. Viel, *J. Chem. Phys.* **140**, 224109 (2014).
- ¹⁴W. Eisfeld and A. Viel, *J. Chem. Phys.* **146**, 034303 (2017).
- ¹⁵A. Viel and W. Eisfeld, *Chem. Phys.* **509**, 81 (2018).
- ¹⁶A. Viel and W. Eisfeld, *J. Chem. Phys.* **120**, 4603 (2004).
- ¹⁷W. Eisfeld and A. Viel, *J. Chem. Phys.* **122**, 204317 (2005).
- ¹⁸D. Opalka and W. Domcke, *J. Chem. Phys.* **132**, 154108 (2010).
- ¹⁹S. Bhattacharyya, D. Opalka, L. V. Poluyanov, and W. Domcke, *J. Phys. Chem. A* **118**, 11962 (2014).
- ²⁰C. Roberston and G. A. Worth, *Chem. Phys.* **460**, 125 (2015).
- ²¹T. Zeng and I. Seidu, *Phys. Chem. Chem. Phys.* **19**, 11098 (2017).
- ²²E. Wigner and J. Griffin, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra*, Pure and Applied Physics (Academic Press, 1959).
- ²³D. M. G. Williams and W. Eisfeld, *J. Chem. Phys.* **149**, 204106 (2018).
- ²⁴M. V. Berry, *Proc. R. Soc. London, Ser. A* **392**, 45 (1984).
- ²⁵Note: The general findings presented here for C_{nv} groups are easily extended to general D_{nh} groups. D_{nh} are direct product groups of $C_{nv} \otimes C_s$, and the additional symmetry element (horizontal mirror plane σ_h) does not affect the $E_\alpha \otimes e_\beta$ JT properties.
- ²⁶H. A. Jahn and E. Teller, *Proc. R. Soc. London, Ser. A* **161**, 220 (1937).
- ²⁷P. R. Wallace, *Phys. Rev.* **71**, 622 (1947).
- ²⁸Accurate and fully converged quantum dynamics simulations have been carried out using the discrete variable representation (DVR) approach to set up the 2D Hamiltonian, which then is diagonalized by an accurate short iterative Lanczos algorithm. Two sets of calculations have been performed one with a harmonic oscillator basis (Hermite DVR) and the other with a plane wave basis (Fourier transform DVR). The two different bases were used in order to analyze different aspects of the resulting wave functions more easily.
- ²⁹Note: The e_β coordinate is symmetric with respect to C_β rotation, and $\beta = 2$ is a divisor of $n = 6$ for the C_{6v} $E_\alpha \otimes e_2$ JT system. Thus, only $n/\beta = 3$ equivalent PES wells are obtained.

Neural network diabotization: A new ansatz for accurate high-dimensional coupled potential energy surfaces

David M. G. Williams, and Wolfgang Einfeld

Citation: *J. Chem. Phys.* **149**, 204106 (2018); doi: 10.1063/1.5053664

View online: <https://doi.org/10.1063/1.5053664>

View Table of Contents: <http://aip.scitation.org/toc/jcp/149/20>

Published by the [American Institute of Physics](#)

PHYSICS TODAY

WHITEPAPERS

ADVANCED LIGHT CURE ADHESIVES

Take a closer look at what these environmentally friendly adhesive systems can do

READ NOW

PRESENTED BY
 **MASTERBOND**
ADHESIVES | SEALANTS | COATINGS

Neural network diabatization: A new *ansatz* for accurate high-dimensional coupled potential energy surfaces

David M. G. Williams and Wolfgang Eisfeld^{a)}

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

(Received 24 August 2018; accepted 23 October 2018; published online 28 November 2018)

A new diabatization method based on artificial neural networks (ANNs) is presented, which is capable of reproducing high-quality *ab initio* data with excellent accuracy for use in quantum dynamics studies. The diabatic potential matrix is expanded in terms of a set of basic coupling matrices and the expansion coefficients are made geometry-dependent by the output neurons of the ANN. The ANN is trained with respect to *ab initio* data using a modified Marquardt-Levenberg back-propagation algorithm. Due to its setup, this approach combines the stability and straightforwardness of a standard low-order vibronic coupling model with the accuracy by the ANN, making it particularly advantageous for problems with a complicated electronic structure. This approach combines the stability and straightforwardness of a standard low-order vibronic coupling model with the accuracy by the ANN, making it particularly advantageous for problems with a complicated electronic structure. This novel ANN diabatization approach has been applied to the low-lying electronic states of NO₃ as a prototypical and notoriously difficult Jahn-Teller system in which the accurate description of the very strong non-adiabatic coupling is of paramount importance. Thorough tests show that an ANN with a single hidden layer is sufficient to achieve excellent results and the use of a “deeper” layering shows no clear benefit. The newly developed diabatic ANN potential energy surface (PES) model accurately reproduces a set of more than 90 000 Multi-configuration Reference Singles and Doubles Configuration Interaction (MR-SDCI) energies for the five lowest PES sheets. *Published by AIP Publishing.* <https://doi.org/10.1063/1.5053664>

I. INTRODUCTION

The fundamental understanding of chemical processes and spectroscopy is one of the main goals in chemical physics and theoretical chemistry. A key component in the strive for such understanding is the study of the underlying molecular quantum dynamics. With the ongoing advances of theoretical and experimental treatments of chemical systems alike, fundamental insights regarding chemical dynamics become increasingly available. However, the complexity of experimental data and the intricacy of theoretical treatments vastly increase with the size of the system. Therefore, extending the fundamental insights to larger systems is of great interest and an active field of current research.

One of the core issues in the theoretical treatment of systems beyond triatomics is the development of accurate high-dimensional potential energy surfaces (PESs), which provide the basis for quantum-dynamical investigations. In the case of a single adiabatic electronic state, several strategies for developing higher-dimensional PESs of sufficient accuracy have been established. One class of approaches is based on local interpolation techniques,^{1–5} and another one uses invariant polynomials and least-squares fitting to utilize the permutation symmetry of indistinguishable nuclei.⁶ However, extending these methods to PESs of multiple

excited states proves difficult due to the added complexity of accounting for state-state interactions. Recent attempts to tackle this issue have been formulated by utilizing the modified Shepard interpolation^{7–9} and by using invariant polynomials and complete nuclear permutation-inversion (CNPI) symmetry.^{10,11}

The proper inclusion of excited states requires a number of PESs and their couplings to be represented in an appropriate form. In general, there will be at least some region in nuclear configuration space (NCS) with a high density of electronic states, such that interactions among electronic states become significant, rendering the Born-Oppenheimer (BO) approximation invalid. For such a case, a *quasi-diabatic* representation of the coupled electronic states has been found to be of great advantage.^{12–25} Quasi-diabatic, which will be referred to as “diabatic” from here on for the sake of simplicity, means that the state basis to represent the electronic Hamiltonian is required to preserve the character of the electronic states as much as possible, reducing the nonadiabatic (or derivative) coupling to a degree that it can be neglected safely. Thus, the electronic Hamiltonian will be represented by a non-diagonal PES matrix encoding all state energies and couplings. One advantage is that the matrix elements of the electronic Hamiltonian become simple, well-behaved functions of the nuclear coordinates, reducing the complexity of finding analytic expressions for them. The simplest (and most commonly used) such approach utilizing the diabatic representation is the multi-mode linear vibronic coupling method by

^{a)}wolfgang.eisfeld@uni-bielefeld.de

Köppel, Domcke, and Cederbaum²⁶ as well as similar methods developed by Yarkony and co-workers.^{27–30} These treatments, while suitable for the simulation of certain ultra-fast nonadiabatic processes, are too limited to represent more extended regions of the coupled PESs as required for the treatment of more complicated dynamical processes. Therefore, extensions of vibronic coupling models needed for the treatment of dynamical processes of higher complexity have been developed.^{31–36} Furthermore, several other approaches exist utilizing properties of the adiabatic electronic wave functions for diabaticization.^{18,21,22,24,37–42} Most of these methods do not require a model diabatic matrix, meaning that they do not yield a set of PESs in closed mathematical form. Therefore, a second step would be required to represent the diabatic matrix elements provided by a point-wise diabaticization scheme by some external mathematical model. Representing the diabatic energies and couplings accurately and in closed form as diabatic PESs beyond simple models is still a major problem for systems with more than three atoms.^{43–46} This has been attempted by the aforementioned methods of extending vibronic coupling models in various ways,^{31–36} using modified Shepard interpolation^{7–9} or invariant polynomials and CNPI symmetry¹⁰ as well as by choosing elaborate functional forms for the diabatic matrix elements.^{43–45} However, the number of accurate diabatic surfaces for such systems in the literature remains rather scarce.

Recently an alternative approach for the representation of PESs evolved utilizing artificial neural networks (ANNs),^{47–65} which are trained by various methods usually based on analytic reference data from other models or *ab initio* data. It has been shown mathematically that ANNs are capable of uniformly representing any continuous real function of n dimensions up to arbitrary accuracy.⁶⁶ Therefore, in principle, it should be possible to train an appropriate neural network to represent any PES accurately with respect to the underlying data given. Apart from the high accuracy that can be achieved, ANNs have the advantage that they can be evaluated very efficiently once trained, which plays a key role in quantum dynamics methods such as multi-configuration time-dependent Hartree (MCTDH) where the evaluation of the PES is the most time demanding part.^{67,68} While ANNs are in principle capable of reproducing arbitrary continuous functions, this is in practice of course limited by the acquisition of data, training algorithm, chosen network architecture, and an issue commonly referred to as “over-fitting.” However, ANNs have already been used with impressive results to represent a single PES based on high-level *ab initio*-data.^{63,65} Some first attempts to extend the use of ANNs to diabatic PESs are also known.^{69,70}

In the present work, a novel ANN based diabaticization approach is used to accurately represent the coupled PESs of electronically excited states with strong Jahn-Teller (JT) couplings. For this purpose, an ANN-based diabatic model has been developed and trained to represent the low-lying electronic states of NO_3 . The new method is described here for the first time and the overall quality, the resulting PES model, and the stability of the fitting procedure is analyzed depending on various factors.

II. THEORY

A. Adiabatic and diabatic representation

As pointed out above, we aim for an accurate diabatic representation of the electronic Hamiltonian of a given molecular system. Therefore, the theoretical background of adiabatic and diabatic representations is briefly summarized in the following. The starting point is the total molecular Hamiltonian with electronic degrees of freedom \mathbf{q} and nuclear degrees of freedom \mathbf{Q} reading

$$\hat{H}(\mathbf{q}, \mathbf{Q}) = \hat{T}_{nuc}(\mathbf{Q}) + \underbrace{\hat{T}_{el}(\mathbf{q}) + V_C(\mathbf{q}, \mathbf{Q})}_{\hat{H}_{el}(\mathbf{q}, \mathbf{Q})}. \quad (1)$$

If the nuclear motions are frozen, the nuclear kinetic energy $\hat{T}_{nuc}(\mathbf{Q})$ vanishes and what remains is the electronic Hamiltonian \hat{H}_{el} consisting of the electronic kinetic energy \hat{T}_{el} and the Coulomb potential V_C . The eigenvalues $E_k^a(\mathbf{Q})$ and eigenfunctions $\psi_k^a(\mathbf{q}; \mathbf{Q})$ of \hat{H}_{el} can be evaluated by suitable *ab initio* methods for selected points \mathbf{Q} in the nuclear configuration space. The $E_k^a(\mathbf{Q})$ depend parametrically on the nuclear coordinates \mathbf{Q} which yields the adiabatic potential energy surfaces (PESs) for the electronic states in question. The set of all electronic eigenfunctions ψ_k^a forms the *adiabatic basis* for the complete molecular wave functions. The total wave function thus can be expanded as

$$\Psi_j^a(\mathbf{q}, \mathbf{Q}) = \sum_k \phi_k^a(\mathbf{Q}) \cdot \psi_k^a(\mathbf{q}; \mathbf{Q}). \quad (2)$$

Expanding \hat{H} in terms of $\psi_k^a(\mathbf{q}; \mathbf{Q})$ yields the *adiabatic representation* $\hat{H}^a(\mathbf{Q})$ with all electronic degrees of freedom integrated out

$$\hat{H}_{kj}^a = (\hat{T}_{nuc}(\mathbf{Q}) + E_k^a(\mathbf{Q})) \cdot \delta_{kj} - \hat{\Lambda}_{kj}(\mathbf{Q}). \quad (3)$$

In case that the derivative coupling terms $\hat{\Lambda}(\mathbf{Q})$ are small, they can be ignored which is widely known as Born-Oppenheimer (BO) approximation. However, this condition is often not fulfilled for electronically excited states and $\hat{\Lambda}(\mathbf{Q})$ may even become singular or very large for conical intersections or avoided crossings, respectively.¹²

By contrast, a *diabatic basis* $\{\psi_k^d(\mathbf{q}; \mathbf{Q})\}$ is chosen such that $\hat{\Lambda}$ remains negligibly small by construction so that in the resulting Hamiltonian

$$\hat{H}_{kj}^d = \hat{T}_{nuc} \cdot \delta_{kj} + W_{kj}^d(\mathbf{Q}) \quad (4)$$

the derivative coupling terms can be ignored just like in the BO approximation. The kinetic (derivative) coupling of the adiabatic representation is transformed into a potential coupling in the diabatic representation since the diabatic potential matrix $W^d(\mathbf{Q})$ is not diagonal anymore. Unfortunately, a diabatic basis cannot be defined unambiguously and can only be determined *ab initio* through the computation of adiabatic eigenstates and energies. However, it can be shown that the diabatic matrix elements $W_{kj}^d(\mathbf{Q})$ must be simple and smooth functions of the nuclear coordinates. Therefore, the approximate mathematical representation of $W^d(\mathbf{Q})$ is usually much easier than that of $E^a(\mathbf{Q})$. Finally, the adiabatic basis functions

can be expanded in terms of diabatic basis functions as

$$\psi_l^a(\mathbf{q}; \mathbf{Q}) = \sum_k u_{kl}(\mathbf{Q}) \cdot \psi_k^d(\mathbf{q}; \mathbf{Q}). \quad (5)$$

This relates to a basis transformation $\mathbf{U}(\mathbf{Q})$ which would be exact if both bases would be complete or span the same vector space. In this case, the unitary matrix \mathbf{U} diagonalizes \mathbf{W}^d

$$\mathbf{U}^\dagger \mathbf{W}^d \mathbf{U} = \mathbf{W}^a = \text{diag}(W_j^a) \quad (6)$$

and the eigenvalues exactly reproduce the adiabatic energies $E_k^a(\mathbf{Q})$.

Up to this point no approximations were necessary as all bases were assumed to be complete. However, for any practical application, all bases involved are finite and only a small number of states can be handled. Thus, the adiabatic states are approximately represented in a finite N^d -dimensional diabatic basis by

$$\psi_l^a(\mathbf{q}; \mathbf{Q}) \approx \sum_k u_{kl}^d(\mathbf{Q}) \cdot \psi_k^d(\mathbf{q}; \mathbf{Q}). \quad (7)$$

N^d can be equal or greater than the number of adiabatic states required.

B. Artificial neural networks

In the present work, multilayer perceptron feed-forward neural networks are utilized for the diabaticization of adiabatic molecular energy data. A feed-forward neural network is a function taking a vector $\boldsymbol{\eta}^{(1)}$ as the 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

$$\eta_j^{(k)} = f^{(k)} \left(\underbrace{\beta_j^{(k)} + \sum_l \omega_{jl}^{(k)} \eta_l^{(k-1)}}_{\chi_j^{(k)}} \right). \quad (8)$$

Here, $f^{(k)}$ is a function of one variable $\chi_j^{(k)}$ called the activation function, $\chi_j^{(k)}$ being a weighted sum of the values of the neurons $\eta_l^{(k-1)}$ of the previous layer with an added bias term $\beta_j^{(k)}$. The resulting ANN can be visualized as shown in Fig. 1.

C. Diabatic model (ansatz)

Since the diabatic matrix elements are by nature slowly varying functions of the nuclear coordinates, it is straightforward to expand them as multi-dimensional polynomials as

$$W_{kj}^d(\mathbf{Q}) = \sum_\alpha p_\alpha^{kj} \prod_l Q_l^{n_\alpha l}. \quad (9)$$

The order of the polynomials is given by the sum of the exponents $n = \sum_l n_\alpha l$. Special boundary conditions like asymptotic behaviour or periodicity may be incorporated into the definition of the nuclear coordinates Q_l . The symmetry of the system is conveniently accounted for by using symmetry-adapted

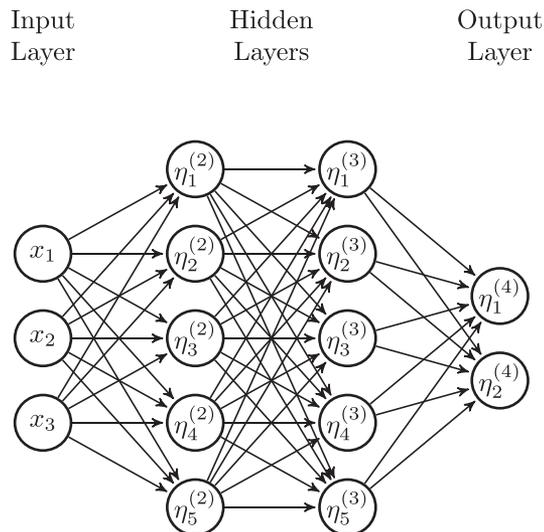


FIG. 1. A neural network with 2 hidden layers. Arrows correspond to weighted contributions. Bias terms and activation functions are not visualized.

coordinates in which case certain expansion parameters p_α^{kj} may vanish or be strictly related to other expansion parameters. In the latter case, it is very convenient to rewrite this expansion in terms of basis matrices as

$$\mathbf{W}^d(\mathbf{Q}) = \sum_L \lambda_L \cdot \mathbf{M}_L(\mathbf{Q}). \quad (10)$$

The symmetry relations are encoded in the basis matrices $\mathbf{M}_L(\mathbf{Q})$ here. This approach is extremely successful and straightforward in the case of *linear vibronic coupling*²⁶ or for other low-order expansions. The order of the expansion terms needs to be increased if higher accuracy of the PES model is required over more extended regions of the nuclear configuration space (NCS). However, the higher-order terms cause several technical problems for the required nonlinear fitting as well as possibly causing artifacts in the shape of the PESs. Hence, the choice of terms included in the model usually requires a trade-off between lower accuracy with higher stability and higher (local) accuracy at the cost of more erratic behavior outside the sampling space. In the present work, we therefore propose a new approach based on an expansion in the basis matrices of lowest orders. The expansion coefficients are first obtained from a standard nonlinear least-squares fit and then are modified by *corrections* provided by ANNs. As will be shown below, this new method overcomes the characteristic oscillation features plaguing typical higher-order polynomial expansions. Furthermore, this approach does not produce the erratic behavior outside of the sampling space, often observed in higher-order polynomial expansions. The new ANN approach, thus, leads to very robust diabatic ANN-PES models.

The fundamental idea of this method is to render the expansion coefficients of Eq. (10) coordinate dependent reading

$$\mathbf{W}^d(\mathbf{Q}) = \sum_L \lambda_L(\mathbf{Q}) \cdot \mathbf{M}_L(\mathbf{Q}), \quad (11)$$

where each coefficient is of the form

$$\lambda_L(\mathbf{Q}) = \begin{cases} \lambda_L^0 & \text{for uncorrected terms} \\ \lambda_L^0 \cdot (1 + \eta_{k_L}^{(f)}(\mathbf{Q})) & \text{otherwise.} \end{cases} \quad (12)$$

The constants λ_L^0 obtained from standard nonlinear fitting procedures are modified by the coordinate dependent outputs of a trained ANN. This way, one can combine the stability and basic qualitative description of low-order expansions while still introducing flexibility to achieve excellent accuracy. Of course, the ANN must be trained properly with respect to *ab initio* reference data and this training also involves a necessary adiabatic-to-diabatic transformation. This requires a modification of standard ANN training algorithms described next.

D. Backpropagation of generalized outputs

All special properties of ANNs aside, they can be viewed as parametrized functions modeled for the specific purpose of closely reproducing an expected output \mathbf{t} for each point in the function's domain. In the most simple case, the function output $\boldsymbol{\eta}^{(f)}$ is modeled to *directly* reproduce the desired output \mathbf{t} , in which case the error can be measured in terms of the difference \mathbf{e} between the two. The problem of finding an optimal set of parameters can then be expressed in terms of minimizing the sum of squares error

$$V = \frac{1}{2} |\mathbf{e}^2|. \quad (13)$$

Applying a (gradient descent) step in parameter space to improve the current parameter set requires derivatives of V with respect to all parameters $\omega_{jl}^{(k)}$ and $\beta_j^{(k)}$. However, thanks to the mathematical structure of feed-forward networks, the derivatives can be easily expressed in terms of derivatives with respect to the weighted sums $\chi_j^{(k)}$ by defining

$$\delta_j^{(k)} := \frac{\partial V}{\partial \chi_j^{(k)}} \quad (14)$$

which yields

$$\frac{\partial V}{\partial \omega_{jl}^{(k)}} = \delta_j^{(k)} \cdot \eta_l^{(k-1)}, \quad (15a)$$

$$\frac{\partial V}{\partial \beta_j^{(k)}} = \delta_j^{(k)}. \quad (15b)$$

These equations generalize naturally for sums over multiple data points by introducing a sum over all points \mathbf{Q}_p which is omitted for brevity. These $\delta_j^{(k)}$ are in return analytically known from the recurrence relation

$$\delta_j^{(k)} = \underbrace{f_j^{(k)}(\chi_j^{(k)})}_{f_j^{(k)}} \cdot \sum_k \omega_{kj}^{(k+1)} \delta_k^{(k+1)}, \quad (16)$$

which terminates at the final layer, yielding

$$\delta_j^{(f)} = -e_j \cdot f_j^{(f)}. \quad (17)$$

Evaluating and applying gradient descent steps by exploiting the recursive nature of the required derivatives seen in Eq. (16) is referred to as *backpropagation*. This scheme is generally applicable as long as the error can be *directly* expressed in

terms of the ANN output. However, in principle, the neural network output $\boldsymbol{\eta}^{(f)}$ can relate to the desired output \mathbf{t} in arbitrarily *indirect* ways. For example, the present case uses the individual output values as coefficients for a diabatic expansion, which in turn produces a matrix that is diagonalized to obtain adiabatic energies comparable to the actual *ab initio* data. This processed form of $\boldsymbol{\eta}^{(f)}$, denoted $\mathcal{D}(\boldsymbol{\eta}^{(f)})$, is now what is actually comparable to \mathbf{t} . Hence we developed a modification allowing for arbitrary differentiable relations $\mathcal{D}(\boldsymbol{\eta}^{(f)})$ between the output layer and physical reference data. While the actual backpropagation remains the same, the final layer now terminates to

$$\begin{aligned} \delta_j^{(f)} &= \frac{\partial V}{\partial \chi_j^{(f)}} = - \sum_l e_l \cdot \frac{\partial \mathcal{D}_l(\boldsymbol{\eta}^{(f)})}{\partial \chi_j^{(f)}} \\ &= - \sum_l e_l \cdot \frac{\partial \mathcal{D}_l(\boldsymbol{\eta}^{(f)})}{\partial \eta_j^{(f)}} \frac{\partial \eta_j^{(f)}}{\partial \chi_j^{(f)}} \\ &= - \sum_l e_l \cdot \frac{\partial \mathcal{D}_l}{\partial \eta_j^{(f)}} \cdot f_j^{(f)}. \end{aligned} \quad (18)$$

In the present case, where $\mathcal{D}(\boldsymbol{\eta}^{(f)})$ refers to the eigenvalues of a diabatic matrix which is parametrized by the output layer $\boldsymbol{\eta}^{(f)}$ as seen in Eqs. (11) and (6), partial derivatives are obtained from numerical differentiation. This is always possible as long as the target \mathbf{t} refers to a continuously differentiable property. Since the diagonalization is carried out numerically, the partial derivatives must be computed numerically as well.

E. (Adapted) Marquardt-Levenberg method

The most basic training algorithm using the backpropagation would be a simple gradient decent approach. However, due to the extreme non-linearity of the fitting problem, it is desirable to choose a more sophisticated algorithm to optimize a given parameter set $\boldsymbol{\Omega}$ (containing all weights and biases) than a gradient descent method can provide. A widely used improvement is to utilize a Marquardt-Levenberg algorithm⁷¹ which approximates the exact Hessian using the Jacobian

$$J_{lk} = \frac{\partial e_l}{\partial \Omega_k} \quad (19)$$

and a damping term λ . The working equation to produce the correction step $\Delta \boldsymbol{\Omega}$ in parameter space then reads

$$\Delta \boldsymbol{\Omega} = (\mathbf{J}^T \mathbf{J} + \lambda \mathbf{1})^{-1} \mathbf{J}^T \mathbf{e}. \quad (20)$$

It can be shown, in analogy to the derivatives of V with respect to individual parameters, that \mathbf{J} can be evaluated analytically. By using backpropagation with a modified termination condition,⁷¹ a single backpropagation yields all J_{lk} for a fixed row index l , meaning one requires as many backpropagations to construct \mathbf{J} as there are physical outputs per data point (e.g., adiabatic energies). In the standard case, each output neuron corresponds to one physical output and the final backpropagation layer $\Delta^{(f,l)}$ reads

$$\Delta_j^{(f,l)} = -\delta_{lj} \cdot f_j^{(f)} \quad (21)$$

for a given l . We again derived a modified version for the case that there is no one to one correspondence of neural network output and physical output but an arbitrary differentiable relation $\mathcal{D}(\boldsymbol{\eta}^{(f)})$ instead. Equation (21) then generalizes to

$$\Delta_j^{(f,l)} = -\frac{\partial \mathcal{D}_l}{\partial \eta_j^{(f)}} \cdot f_j^{(f)}. \quad (22)$$

III. APPLICATIONS

The new method should be widely applicable and can be used for conventional diabaticization-by-*ansatz* problems as well as for the more advanced approaches like the hybrid diabaticization⁷² recently developed by us. A prototypical Jahn-Teller problem is chosen to demonstrate the novel ANN diabaticization. The first three electronic states of NO_3 , \tilde{X}^2A_2' , \tilde{A}^2E'' , and \tilde{B}^2E' , have been studied by us extensively^{35,36,73–75} and thus NO_3 is an ideal benchmark system for the present study. The results of the present work can be compared directly with our latest diabatic PES model.⁷⁵

A. Diabatic model for NO_3

For this first proof-of-principle study of the new ANN diabaticization method, a restricted-dimensional (5D) model of the lowest PESs corresponding to the \tilde{X} , \tilde{A} , and \tilde{B} states of the planar NO_3 radical has been developed. Due to the global D_{3h} symmetry of NO_3 , two pairs of PESs become degenerate at symmetry points and are split for distorted geometries due to strong JT coupling. Thus, the symmetry and couplings have to be accounted for properly. Furthermore, the \tilde{A} state has a rather low dissociation threshold and thus the asymptotic behaviour of the PESs is relevant. To account for the basic asymptotic behavior and a correct description of the complete nuclear transformation symmetry in the underlying low-order model, a set of symmetry-adapted coordinates has been chosen as already described in previous work.^{33,36} First, a set of primitive coordinates is defined comprising the three N–O distances r_i and a set of O–N–O angles α_i . The latter are numbered according to the unique atom not involved in forming the angle. These primitive coordinates are first transformed nonlinearly as

$$m_i = 1 - \exp(-\gamma(r_i - r_0)), \quad (23a)$$

$$\alpha'_i = \frac{\alpha_i - \alpha_0}{r_j r_k}, \quad i \neq j \neq k \quad (23b)$$

to account for the asymptotic behaviour. Here, r_0 and α_0 are the respective distances and angles at the reference point and γ is a chosen Morse-parameter. Then the transformed coordinates are symmetry-adapted to yield a breathing mode a , a degenerate pair of asymmetric stretching modes x_s, y_s , and a degenerate pair of asymmetric bending modes x_b, y_b by

$$a = \sqrt{\frac{1}{3}}(m_1 + m_2 + m_3), \quad (24a)$$

$$x_s = \sqrt{\frac{1}{6}}(2m_1 - m_2 - m_3), \quad (24b)$$

$$y_s = \sqrt{\frac{1}{2}}(m_2 - m_3), \quad (24c)$$

$$x_b = \sqrt{\frac{1}{6}}(2\alpha'_1 - \alpha'_2 - \alpha'_3), \quad (24d)$$

$$y_b = \sqrt{\frac{1}{2}}(\alpha'_2 - \alpha'_3). \quad (24e)$$

With these coordinates, all linear coupling matrices can be constructed with two non-trivial kinds of coupling matrix blocks

$$\boldsymbol{\varepsilon}_{s,b} = \begin{pmatrix} x_{s,b} & y_{s,b} \\ y_{s,b} & -x_{s,b} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\rho}_{s,b} = \begin{pmatrix} x_{s,b} \\ -y_{s,b} \end{pmatrix} \quad (25)$$

accounting for ${}^2E'$ and ${}^2E''$ Jahn-Teller coupling as well as *pseudo*-Jahn-Teller coupling between ${}^2A_2'$ and ${}^2E'$. The diagonal potential $\mathbf{W}^d(\mathbf{Q})$ consists of 12 coordinate-dependent terms up to second order listed in the [Appendix](#). The three constants defining the energy differences between the electronic states at the D_{3h} reference geometry are modified by the ANN as well as the expansion coefficients of the coupling matrices. This yields a total of 9 ANN-corrected basis matrices. In the following, the diabatic electronic basis states are used in order ${}^2A_2'$, ${}^2E''$, and ${}^2E'$. This yields the diabatic matrix

$$\begin{aligned} \mathbf{W}^d(\mathbf{Q}) = & \mathbf{W}_{\text{diag}}^d(\mathbf{Q}) + \begin{pmatrix} \lambda_1 & 0^T & 0^T \\ 0 & \lambda_2 \mathbf{1} & 0 \\ 0 & 0 & \lambda_3 \mathbf{1} \end{pmatrix} \\ & + \begin{pmatrix} 0 & 0^T & 0^T \\ 0 & \lambda_4 \boldsymbol{\varepsilon}_s + \lambda_5 \boldsymbol{\varepsilon}_b & 0 \\ 0 & 0 & \lambda_6 \boldsymbol{\varepsilon}_s + \lambda_7 \boldsymbol{\varepsilon}_b \end{pmatrix} \\ & + \begin{pmatrix} 0 & 0^T & \lambda_8 \boldsymbol{\rho}_s^T + \lambda_9 \boldsymbol{\rho}_b^T \\ 0 & 0 & 0 \\ \lambda_8 \boldsymbol{\rho}_s + \lambda_9 \boldsymbol{\rho}_b & 0 & 0 \end{pmatrix}, \quad (26) \end{aligned}$$

where $\mathbf{W}_{\text{diag}}^d$ contains all uncorrected diagonal terms. The first, second, and third matrix terms encode the ANN-corrected contributions to the diagonal potential terms, the Jahn-Teller coupling blocks, and the *pseudo*-Jahn-Teller coupling blocks, respectively. The corrected coefficients λ_i are of the general form

$$\lambda_j = \lambda_j^0 \cdot (1 + c_j \cdot \eta_j^{(f)}(\mathbf{Q})). \quad (27)$$

Factors of c_j allow further flexibility if particular λ_i^0 (*read*: the reference model terms) reside in a different order of magnitude than the other terms (the coupling terms) and hence require a different treatment.

B. *Ab initio* data and reference fit

The extremely intricate electronic structure of the NO_3 radical requires a rigorous treatment in order to obtain stable results. The *ab initio* data points are taken from previous work⁷⁵ and were computed by Multi-configuration Reference Singles and Doubles Configuration Interaction (MR-SDCI) calculations based on Complete Active Space Self-Consistent Field (CASSCF) reference wave functions using a slightly adapted correlation consistent aug-cc-pVTZ standard basis.⁷⁴ For details of the *ab initio* computations,

see Refs. 36 and 73–75. The data points have been selected by a stochastic approach described in Ref. 76. To this end, normalised random vectors are generated which define linear cuts through the multi-dimensional PESs starting from the ground state equilibrium structure. Then, *ab initio* points are computed along these random directions. By this approach, we can check and ensure that the CASSCF/MRCI calculations converge to the correct result. This could not be achieved by computing single points at random positions due to the very complicated electronic structure of this system (as typical for radicals and excited states in general). The sampled nuclear configuration space is selected such that the resulting PES model is optimally suited for the simulation of spectroscopic properties. The 21 free parameters λ_j^0 of the reference model were fitted against this data set using a Marquardt-Levenberg algorithm incorporated into a genetic algorithm with a resulting root mean square (rms) error of roughly 1700 cm^{-1} .

The large data set of roughly 90 000 *ab initio* reference energies has been partitioned into training sets of 85% of the data and external validation sets of the remaining 15% of the energies. The validation set error is used as a convergence parameter and does not otherwise contribute to the neural network fit, but has been included in the polynomial reference model. This technique is referred to as *early stopping* in the literature.⁷⁷ In the present work, this early stopping technique is implemented in a “relaxed” fashion. Instead of immediately stopping the fit when the reference error does not improve along with the fitting error, the reference error must increase for a fixed number of consecutive cycles (default is 3) before early stopping is executed. The convergence of the fitting error does not suffer from this restriction, as discussed in Sec. IV below. The contributions to the squared error which is minimized during the training of the neural network are weighted with an exponential decay of energy differences of the form

$$\sigma_{ij} = \exp\left(-\alpha(E_j^a(\mathbf{Q}_i) - E_j^a(\mathbf{Q}_0))\right). \quad (28)$$

This scheme allows us to weigh regions in NCS more or less depending on their relevance in the nuclear dynamics and thus helps to improve the accuracy of the fit in the most relevant regions.

C. Description of ANN setup

Apart from the actual fit, several parameters and options regarding the ANN have been taken into consideration to ensure an optimal setup. The final setup discussed in Sec. IV involves ANNs with only one hidden layer of varying sizes. The sigmoid $\tanh(\chi)$ has been chosen as the hidden layer activation function. For the final layer, the identity $\text{Id}(\chi)$ has been selected. Each ANN is trained starting from a set of 100 randomly generated initial guesses for weights and biases which are then optimized according to the working equations from Sec. II E. While modified backpropagation by itself has been tested, it has been found inferior for the current application compared to the modified Marquardt-Levenberg method. Similarly, “deeper” networks with more than one hidden layer have been tested with various size combinations, but no

improvement compared to single layer ANNs was observed. It has been shown by Cybenko that any continuous real function can be represented with arbitrary accuracy by a single layer ANN.⁶⁶ Since this is a consequence of the superposition of sigmoidal functions, this also holds for multi-layer ANNs and has been shown by Cybenko for ANNs with two hidden layers. In principle, it could be more efficient to achieve a certain accuracy with less optimization parameters by using multiple hidden layers of smaller size. Therefore, we tested this possibility by setting up ANNs with one through four hidden layers and layer sizes such that the total number of optimization parameters is roughly the same. Thus, a benefit of “deep layering” ANNs would manifest itself in lower rms errors. From a study involving various ANN topographies listed in greater detail in Sec. IV below, we conclude that increasing the number of hidden layers shows no clear advantage for the present case (see below). Increasing the number of hidden layers can yield slightly better as well as a significantly worse results compared to a single hidden layer. Larger numbers of initial guesses (up to 2000) were found to have no influence on the overall quality of the final model. Various splitting schemes for fitting reference and validation data, respectively, have been studied. The selection of 15% of the original data as validation set has been found to be sufficiently large to reduce over-fitting while not being detrimental to the overall quality of the final model by removing too many data from the training set. Table III provides a quick overview of various choices of data partitioning, which is discussed in Sec. IV. Furthermore, the validation data has been ensured to have no characteristic differences compared to the fitting data, as random reassignment of fitting and reference subsets had no significant effect on the fit quality.

IV. RESULTS AND DISCUSSION

The single layer ANN setup described in Sec. III C yields excellent fit results for a surprisingly low number of hidden layer neurons. The (weighted) rms error of the primitive reference model is fairly large and about the same for both fitting and validation data set (1730 cm^{-1} and 1710 cm^{-1} , respectively). This is to be expected as both data sets have been included equally in the reference fit. The reference model is not flexible enough to represent the surface in a satisfying quantitative way but is sufficient to reproduce the reference data qualitatively. The remaining deviation is corrected by the ANN model yielding PESs of very high accuracy. Various neural network sizes between 20 and 120 hidden layer neurons have been tested as summarized in Table I. The convergence behavior of the best rms error (and the corresponding reference error) is presented in Fig. 2.

Considering the best ANN parametrization found in each set of 100 initial guesses, the fitting rms expresses satisfactory convergence for 60–75 neurons, the best fit from the 75 neuron set yielding a weighted rms error of 38.4 cm^{-1} . While the fitting rms expectedly decreases further for ever increasing network sizes, the difference between fitting and validation error *increases* also. For this reason the 75 neuron network will be considered as the optimal result and primary

TABLE I. Fitting results of best and 10th best out of 100 trained networks for each hidden layer size, depending on the number of hidden layer neurons. Both fitting rms and rms of validation set are shown. All rms errors are weighted and are given in cm^{-1} .

Neurons	Best	Validation	10th best	Validation
20	101	103	107	111
40	68.3	72.9	69.9	74.3
50	56.0	60.5	58.5	64.5
55	51.0	56.4	54.4	59.8
60	48.1	52.0	49.5	54.1
65	44.6	49.8	46.6	51.2
70	41.9	47.7	43.2	49.0
75	38.4	42.6	41.1	45.3
80	36.7	42.4	38.6	44.0
90	34.4	41.1	35.0	41.0
100	31.0	37.3	32.0	38.7
120	26.8	34.5	27.8	34.0

example to alleviate potential concerns of over fitting. Interestingly, the closeness of the validation set and fitting set error (here 38.4 cm^{-1} and 42.6 cm^{-1}) is consistent throughout many parametrizations in each set, as even the 10th percentile of each set of 100 ranked ANNs generally shows the same behavior as the best ANN itself. This is both a testament to the consistency of the model in areas unknown to the fit as it is to the stability of the method because a comparably small number of initial guesses already produces a lot of networks of similar quality. Knowing the basic capabilities of a 75 neuron single layer ANN, it may be interesting to consider comparable networks of “increased depth,” that is, ANNs with multiple hidden layers. To achieve a fair comparison one can increase the number of layers while keeping the number of corresponding formal parameters (weights and biases) approximately constant and keeping all other technical influences of the fit the same. Overall, the errors produced by networks of increasing depth do not show a straightforward improvement compared to the single-layer case, as seen in Table II. While the ANN with three hidden layers does produce slightly better fitting errors,

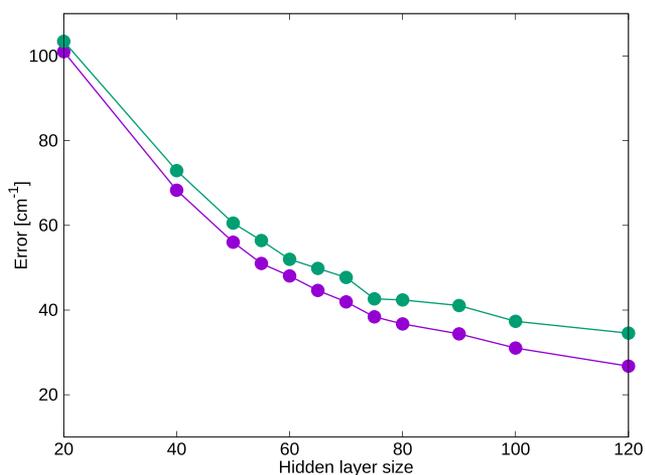


FIG. 2. Weighted root mean square error of the best neural network of a given set as a function of hidden layer size. Purple line represents fitting set error, and green line represents validation set error.

TABLE II. Fitting results of best out of 100 trained networks for each network topography. Both fitting rms and rms of validation set are shown. All rms errors are weighted and are given in cm^{-1} .

Hidden layers	Neurons/Layer	Parameters	Best	Validation
1	75	1134	38.4	42.6
2	27	1170	96.6	104
3	20	1149	26.4	30.2
4	17	1182	75.9	82.5

both two and four hidden layer networks perform significantly worse than the single layer model in this particular setup. Therefore, increasing the number of layers shows no particular advantage compared to the cost of increased complexity.

Another technical factor influencing the fit is that of *early stopping* as described above. In order to investigate potential (negative) impacts on the overall quality of the PES model, the single layer 75 neuron model was fitted again without early stopping and a maximum number of 1000 Marquardt-Levenberg iterations. Both fitting and validation error have been found to show no significant improvement (35.7 cm^{-1} and 40.3 cm^{-1}) at the cost of missing an additional safeguard against over-fitting. Furthermore, the training without early stopping is much less efficient in terms of computer time which might be a big disadvantage for the representation of larger systems and thus larger ANNs.

Apart from the validation set as used in this work, it is not uncommon to withhold an additional part of the data set which is not considered in the fit at all, the so-called *testing set*. Furthermore, choosing the relative sizes of these three data sets may be a point of concern. In particular, increasing the validation (and testing) set(s) may be a way to further decrease the risk of overfitting. The impact of these two factors has been studied for the present case, again using the single layer 75 neuron architecture as a benchmark. Table III shows the results of comparing the present case (85% fit: 15% validation: 0% testing) against a standard literature case (85%: 10%: 5%) as well as the extreme case of choosing a training and validation set of equal size, again with and without the 5% testing set. Due to the extreme similarity of validation and testing rms, no benefit has been observed in withholding data for a testing set. Instead, introducing a testing set at the cost of reducing the validation set has only been found to *increase* the difference between fit and validation for the present case. Similarly, an even split between validation and fitting data has not been found to have any beneficial effect compared to the cost of increased

TABLE III. Fitting results of best out of 100 trained networks (single layer, 75 neurons) for each choice of data split. The full data set is segmented into a set for fitting, a validation set for convergence testing (*early stopping*), and an independent test set. All rms errors are weighted and are given in cm^{-1} .

Split	Best	Validation	Testing
85%: 15%: 0%	38.4	42.6	...
85%: 10%: 5%	39.3	47.0	45.6
50%: 50%: 0%	37.7	53.2	...
50%: 45%: 5%	38.4	50.2	50.2

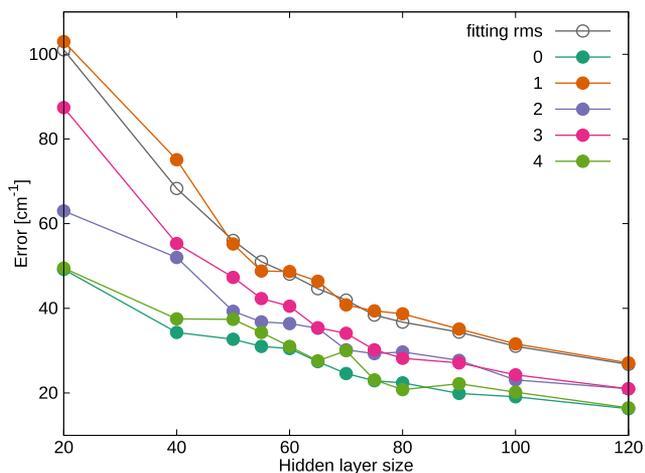


FIG. 3. State-resolved convergence behavior of unweighted root mean square error for *ab initio* data points up to 1 eV above respective reference point energy. States are enumerated from ground state to energetically highest lying state. Gray: weighted fitting set error.

disparity between the two. For this reason, our preferred distribution of the reference data is 85% fit: 15% validation: 0% testing.

Another useful measure for the accuracy of the surface is the unweighted rms error of the adiabatic energies below a certain threshold. Since we are particularly interested in the quality of the surface in the physically relevant region, the cutoff energies have been chosen 1 eV above the reference point energies, which results in about 15 000 *ab initio* values for each electronic state.

Considering again all network sizes as shown in Fig. 3 and Table IV, one finds that the unweighted errors are generally well below the weighted fitting rms for all states except the first (and to a lesser degree third) excited state. This indicates that the physically relevant regions of the surface are reproduced even significantly better than the weighted rms over all *ab initio* data points would indicate. While the convergence of the individual states initially appears more erratic than the total weighted error for small networks, it becomes well-behaved

TABLE IV. Fitting results of best out of 100 trained networks of size 20 to 120, depending on the number of hidden layer neurons. Both fitting rms and rms of validation set are shown. All rms errors are unweighted with a cutoff 1 eV above each state's reference energy. All errors are given in cm^{-1} .

Neurons	State 0	State 1	State 2	State 3	State 4
20	49.2	103	63.0	87.4	49.5
40	34.3	75.1	52.0	55.3	37.5
50	32.7	55.2	39.3	47.3	37.4
55	31.0	48.8	36.8	42.3	34.3
60	30.5	48.7	36.4	40.5	31.0
65	27.4	46.4	35.3	35.4	27.6
70	24.6	40.8	30.2	34.1	30.0
75	22.9	39.4	29.3	30.2	23.1
80	22.4	38.7	29.7	28.2	20.8
90	19.9	35.1	27.7	27.1	22.2
100	19.1	31.6	23.1	24.3	20.2
120	16.3	27.1	21.0	21.0	16.5

for networks approaching acceptable sizes (55 onwards). What remains to be investigated further is the outlying error of the first excited state, which almost appears to dominate the total weighted error for large network sizes. This behavior is easily explained when investigating the topographical complexity of the individual sheets, as the first excited state has a far more complicated multi-minimum structure than the ground state, resulting in more strain for the network. Conversely, the ground state itself is extremely well reproduced due to its significantly simpler shape. With both convergences with respect to network size and initial guess number sufficiently expanded upon, what remains is a deeper analysis of the produced diabatic model.

Given the previously considered 75 neuron network, the cutoff-rms of 30.0 cm^{-1} turns out to be significantly smaller than the weighted rms of 38.4 cm^{-1} . The ground state by itself is reproduced with a corresponding cutoff-rms of only 22.9 cm^{-1} . The performance of the ANN model compared to the polynomial reference model is also demonstrated graphically in Figs. 4 and 5. Figure 4 shows that the ANN model consistently reproduces the *ab initio* data quantitatively over the complete energy range of 6 eV for an arbitrary cut through the 5D surfaces. While the reference model shows significant deviations of over 1 eV, the ANN fit is indistinguishable from the reference energies at the scale of the plot. This also applies to scans from the validation data set, as can be seen in Fig. 5,

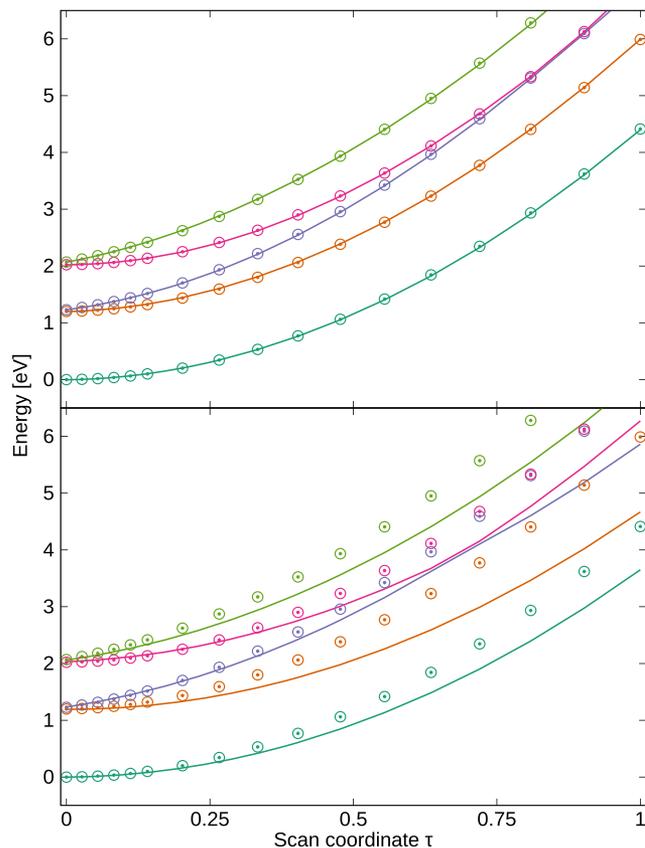


FIG. 4. Comparison of ANN model with underlying primitive reference model along a linear random cut through the 5D surfaces taken out of the training set. Above: energies produced by ANN model. Below: energies produced by polynomial reference model. Circled dots are *ab initio* fitting data. τ is an arbitrary scan parameter.

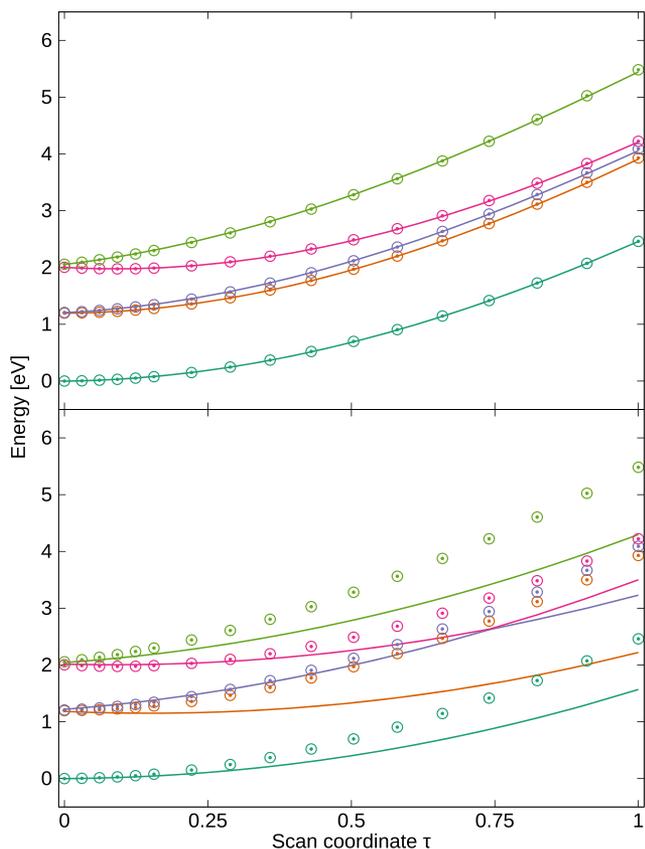


FIG. 5. Comparison of ANN model with underlying primitive reference model using validation set data. Above: energies produced by ANN model. Below: energies produced by polynomial reference model. Circled dots are *ab initio* validation data. Data points belong to a random cut through the 5D surfaces, τ is an arbitrary scan parameter.

indicating consistent behavior beyond the fitting data. Looking at the two sheets of the ${}^2E''$ state as a function of the e' stretch coordinates x_s and y_s , Fig. 6 displays the symmetry-induced three equivalent minima characteristic for the lower adiabatic sheet. While this behavior is to be expected of a quantitatively correct surface, it follows in no way from the underlying reference model. This is due to the fact that *linear* coupling

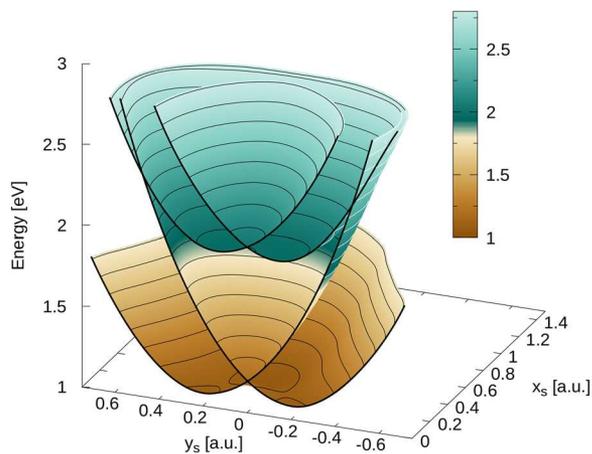


FIG. 6. Cross section of diabatic PESs of the two sheets of the ${}^2E''$ and two sheets of the ${}^2E'$ state of NO_3 depending on the stretching modes x_s and y_s . Contour lines facing outwards are drawn white. The cross section's rim has been marked in black to highlight the present conical intersections.

terms as used in the reference model are only capable (without inclusion of higher order coupling terms) of producing rotationally symmetric “mexican hat”-like potentials. Hence, all characteristic features of the PESs are inherited from the ANN correction. The two adiabatic PES sheets belonging to the \tilde{B}^2E' state are also displayed in Fig. 6. The three equivalent PES wells in the ${}^2E'$ manifold are less pronounced compared to the ${}^2E''$ state and thus not so clearly visible in the figure. The figure also reveals additional intersections between the ${}^2E''$ and ${}^2E'$ PESs which certainly has an impact on the quantum dynamics of the photochemistry of NO_3 after excitation into the \tilde{B} state. These additional conical intersections also become obvious from Fig. 7 in which the adiabatic energies of the five state components are displayed along the 1D bond dissociation coordinate r_{NO} .

The novel ANN approach presented here has one further extremely positive feature beyond its general ability to reproduce the *ab initio* data with very high accuracy. The Taylor expansions of the diabatic matrix used so far tend to produce rapid oscillations and even *unbound areas* in configuration space where no *ab initio* reference data is available. These Taylor expansions require a considerable trade-off between *local accuracy* (higher orders, greater flexibility) and *stability* of the model (lower orders, fewer oscillations). By contrast, the ANN model is found to be completely robust against such oscillations and unbound regions. Considering, for example, a scan along an N–O distance as seen in Fig. 7 which has not been included in the fit, the model shows the ability to produce a complex coupling structure of the higher states without producing oscillations in the repulsive wall. The reason for

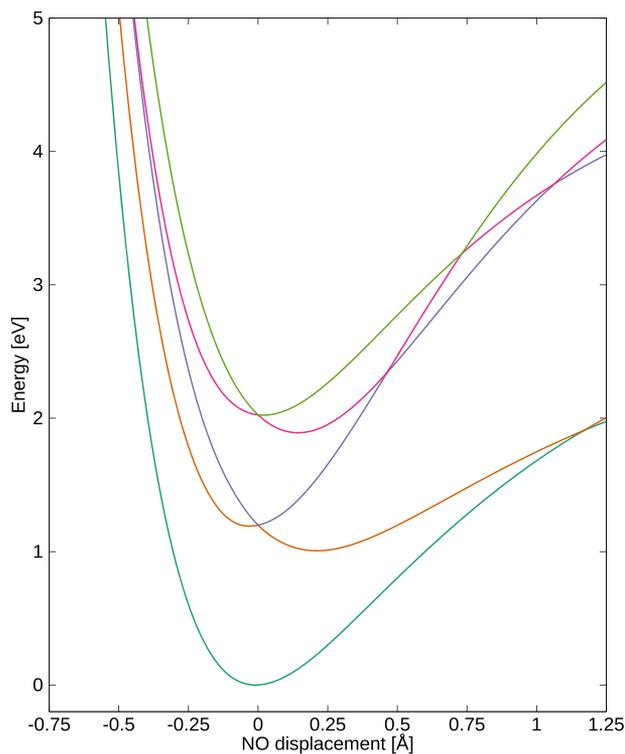


FIG. 7. Scan along NO distance around the reference point with all other coordinates frozen at the reference geometry. The ANN model reproduces conical intersections not included in the reference model and shows appropriate asymptotic behaviour.

this is that the used neural networks, unlike polynomials, are by nature composed of slowly varying sub-expressions (sigmoid functions) which merely increase in number but remain unchanged in character for larger networks while polynomial expansions introduce terms of increasingly erratic behavior which need to be compensated for by all other terms at every reasonable point in coordinate space. Furthermore, it is important to keep in mind that the neural network model does not need to reproduce the coupling terms directly. It only needs to correct a low-order polynomial model which is already intrinsically bound and free of oscillations. Each contribution is scaled in a way such that the neural network corrections reside in the order of magnitude of the expected terms. It turns out that this is an invaluable property of the novel ANN approach when used for quantum dynamics simulations. A demonstration of the strength of the ANN diabaticization method presented here will be given in a forthcoming quantum dynamics study.

V. CONCLUSIONS AND OUTLOOK

A new diabaticization approach is presented to accurately represent the coupled PESs of strongly interacting electronic states. The core idea of this method is to express the diabatic model in terms of a primitive low-order polynomial expansion whose coefficients are then modulated by an artificial neural network (ANN). This approach combines the stability and straightforwardness of a standard low-order vibronic coupling model with the accuracy provided by the ANN as required for reliable quantum dynamics simulations. The new approach is applied to a 5-dimensional 5-state model of planar NO₃. Special attention has been paid to the stability of the fitting algorithm, the overall consistency of the resulting PESs throughout the *ab initio* data, and the quality of the model with respect to the whole data set including points not used for the ANN training as well as the regions which are most relevant for future dynamics calculations, namely, the potential wells. The model has been found to excel in all of the above areas. Furthermore, extensive tests show that a single hidden layer is sufficient for this excellent performance and that “deeper” layering shows no clear benefit. While the primitive and easy to obtain reference model only provides a general bound shape of the potential roughly representing the reference data, the final model is of outstanding quality over an energy range of 6 eV. Despite being intrinsically non-linear, the algorithm produces many high-quality parameter sets even for a small number of initial guesses. Of the roughly 90 000 *ab initio* reference energies obtained from *ab initio* CAS-SCF/MR-SDCI calculations, a subset of 15% has been excluded from the training data, but used as validation set for a convergence test to prevent over-fitting. This validation data set has been found also to be represented with excellent accuracy by the final ANN model despite these data not being used to train the ANN.

The present study demonstrates that the proposed method is capable of producing stable, high-quality PESs based on a straightforward and easy to extend model. What remains is the extension of the model to a full-dimensional description of NO₃, including the umbrella motion as well as corresponding couplings in the diabatic matrix. Furthermore, as a final study

of the resulting PESs, appropriate quantum dynamics calculations are already in progress. As the approach laid out in this paper is generic, applying it to different systems as well as different kinds of coupling (e.g., relativistic coupling) appears to be very promising.

ACKNOWLEDGMENTS

We are grateful to the Deutsche Forschungsgemeinschaft (DFG) for generous financial support. Calculations leading to the results presented here were performed on resources provided by the Paderborn Center for Parallel Computing.

APPENDIX: DIAGONAL MODEL TERMS

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

$$r_{s,b}^2 = x_{s,b}^2 + y_{s,b}^2. \quad (\text{A1})$$

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

$$\mathbf{W}_{\text{diag}}^d(\mathbf{Q}) = \begin{pmatrix} V_1(\mathbf{Q}) & 0^T & 0^T \\ 0 & V_2(\mathbf{Q})\mathbf{1} & 0 \\ 0 & 0 & V_3(\mathbf{Q})\mathbf{1} \end{pmatrix} \quad (\text{A2})$$

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

$$V_i(\mathbf{Q}) = \mu_1^i \cdot a + \mu_2^i \cdot a^2 + \mu_3^i \cdot r_s^2 + \mu_4^i \cdot r_b^2. \quad (\text{A3})$$

- ¹M. A. Collins and D. F. Parsons, *J. Chem. Phys.* **99**, 6756 (1993).
- ²J. Ischtwan and M. A. Collins, *J. Chem. Phys.* **100**, 8080 (1994).
- ³T. S. Ho and H. Rabitz, *J. Chem. Phys.* **104**, 2584 (1996).
- ⁴G. G. Maisuradze, D. L. Thompson, A. F. Wagner, and M. Minkoff, *J. Chem. Phys.* **119**, 10002 (2003).
- ⁵G. G. Maisuradze, A. Kawano, D. L. Thompson, A. F. Wagner, and M. Minkoff, *J. Chem. Phys.* **121**, 10329 (2004).
- ⁶B. J. Braams and J. M. Bowman, *Int. Rev. Phys. Chem.* **28**, 577 (2009).
- ⁷C. R. Evenhuis and M. A. Collins, *J. Chem. Phys.* **121**, 2515 (2004).
- ⁸C. R. Evenhuis, X. Lin, D. H. Zhang, D. Yarkony, and M. A. Collins, *J. Chem. Phys.* **123**, 134110 (2005).
- ⁹O. Godsi, C. R. Evenhuis, and M. A. Collins, *J. Chem. Phys.* **125**, 104105 (2006).
- ¹⁰X. Zhu, J. Y. Ma, D. R. Yarkony, and H. Guo, *J. Chem. Phys.* **136**, 234301 (2012).
- ¹¹C. Xie, J. Ma, X. Zhu, D. H. Zhang, D. R. Yarkony, D. Xie, and H. Guo, *J. Phys. Chem. Lett.* **5**, 1055 (2014).
- ¹²*Conical Intersections: Electronic Structure, Dynamics and Spectroscopy*, edited by W. Domcke, D. R. Yarkony, and H. Köppel (World Scientific, Singapore, 2004).
- ¹³H. C. Longuet-Higgins, *Adv. Spectrosc.* **2**, 429 (1961).
- ¹⁴W. Lichten, *Phys. Rev.* **131**, 229 (1963).
- ¹⁵W. Lichten, *Phys. Rev.* **164**, 131 (1967).
- ¹⁶F. T. Smith, *Phys. Rev.* **179**, 111 (1969).
- ¹⁷M. Baer, *Chem. Phys.* **15**, 49 (1976).
- ¹⁸H. Werner and W. Meyer, *J. Chem. Phys.* **74**, 5802 (1981).
- ¹⁹C. A. Mead and D. G. Truhlar, *J. Chem. Phys.* **77**, 6090 (1982).
- ²⁰C. A. Mead, *J. Chem. Phys.* **78**, 807 (1983).
- ²¹H.-J. Werner, B. Follmeg, and M. H. Alexander, *J. Chem. Phys.* **89**, 3139 (1988).
- ²²T. Pacher, L. S. Cederbaum, and H. Köppel, *J. Chem. Phys.* **89**, 7367 (1988).
- ²³T. Pacher, C. A. Mead, L. S. Cederbaum, and H. Köppel, *J. Chem. Phys.* **91**, 7057 (1989).

- ²⁴T. Pacher, H. Köppel, and L. S. Cederbaum, *J. Chem. Phys.* **95**, 6668 (1991).
- ²⁵T. Pacher, L. S. Cederbaum, and H. Köppel, *Adv. Chem. Phys.* **84**, 293 (1993).
- ²⁶H. Köppel, W. Domcke, and L. S. Cederbaum, *Adv. Chem. Phys.* **57**, 59 (1984).
- ²⁷M. S. Schuurman and D. R. Yarkony, *J. Chem. Phys.* **127**, 094104 (2007).
- ²⁸B. N. Papas, M. S. Schuurman, and D. R. Yarkony, *J. Chem. Phys.* **129**, 124104 (2008).
- ²⁹X. Zhu and D. R. Yarkony, *J. Chem. Phys.* **130**, 234108 (2009).
- ³⁰X. Zhu and D. R. Yarkony, *J. Chem. Phys.* **132**, 104101 (2010).
- ³¹A. 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. Cimraglia, 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).
- ⁶⁵M. Majumder, S. E. Hegger, R. Dawes, S. Manzhos, X.-G. Wang, C. Tucker, Jr., J. Li, and H. Guo, *Mol. Phys.* **113**, 1823 (2015).
- ⁶⁶G. Cybenko, *Math. Control Signals Syst.* **2**, 303 (1989).
- ⁶⁷H. D. Meyer, U. Manthe, and L. S. Cederbaum, *Chem. Phys. Lett.* **165**, 73 (1990).
- ⁶⁸U. Manthe, H. D. Meyer, and L. S. Cederbaum, *J. Chem. Phys.* **97**, 3199 (1992).
- ⁶⁹Y. Guan, B. Fu, and D. H. Zhang, *J. Chem. Phys.* **147**, 224307 (2017).
- ⁷⁰T. Lenzen and U. Manthe, *J. Chem. Phys.* **147**, 084105 (2017).
- ⁷¹M. T. Hagan and M. B. Menhaj, *IEEE Trans. Neural Networks* **5**, 989 (1994).
- ⁷²N. Wittenbrink, F. Venghaus, D. Williams, and W. Eisfeld, *J. Chem. Phys.* **145**, 184108 (2016).
- ⁷³W. Eisfeld and K. Morokuma, *J. Chem. Phys.* **113**, 5587 (2000).
- ⁷⁴W. Eisfeld and K. Morokuma, *J. Chem. Phys.* **114**, 9430 (2001).
- ⁷⁵A. Viel and W. Eisfeld, *Chem. Phys.* **509**, 81 (2018).
- ⁷⁶W. Eisfeld, *J. Chem. Phys.* **134**, 054303 (2011).
- ⁷⁷M. Hagan, H. Demuth, M. Beale, and O. De Jesús, *Neural Network Design*, 2nd ed. (Martin Hagan, 2014), ISBN: 9780971732117, URL: <https://books.google.de/books?id=4EW9oQEACAAJ>.

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



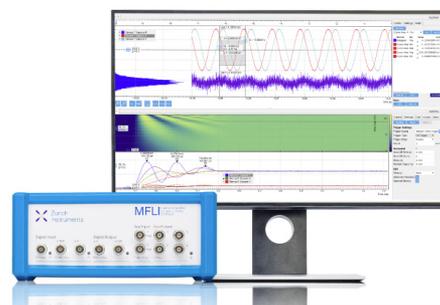
CrossMark

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 $E_i - E_0$	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	μ_i Mean	σ_i 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

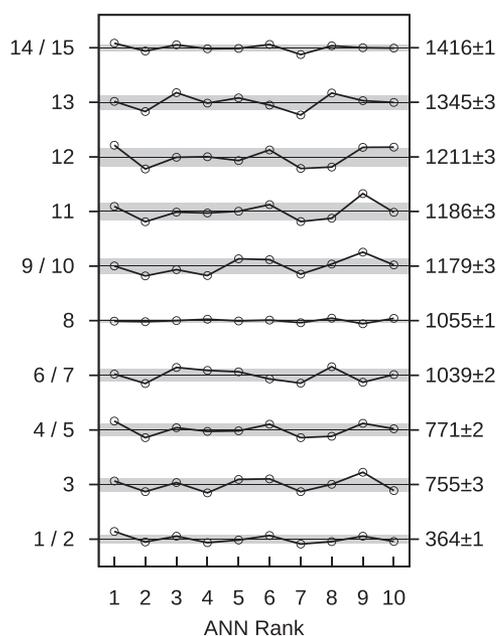


FIG. 1. Visual representation of the first 15 states given in Table II, showing deviations of the 10 best networks from respective average state energies. Level distances, not to scale, correspond to a deviation of 20 cm⁻¹ from average. Gray: Standard deviations σ_i around average energies. Average energies 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 $\tilde{X}^2A'_2$ state of ¹⁴NO₃ (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.

State	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
3	756.3 ± 2.5	5.2	5.8	-5.1	∴	3.0
4/5	774.6 ± 2.2	13.6	10.7	3.1	∴	4.9
6/7	1039.7 ± 2.3	-0.6	0.7	3.9	∴	-4.1
8	1054.9 ± 0.7	0.6	0.2	0.0	∴	-0.4
9/10	1179.3 ± 2.9	9.0	5.9	-6.1	∴	2.8
11	1187.7 ± 3.1	15.3	7.1	-1.0	∴	3.6
12	1215.1 ± 3.4	21.9	13.3	7.6	∴	7.8
13	1344.9 ± 2.6	7.0	5.3	-1.7	∴	-4.3
14/15	1417.7 ± 1.2	4.7	6.2	-0.8	∴	0.8
16/17	1490.0 ± 2.0	-0.1	0.8	1.4	∴	-0.4
18	1498.2 ± 3.8	6.0	4.4	9.6	∴	1.7

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	

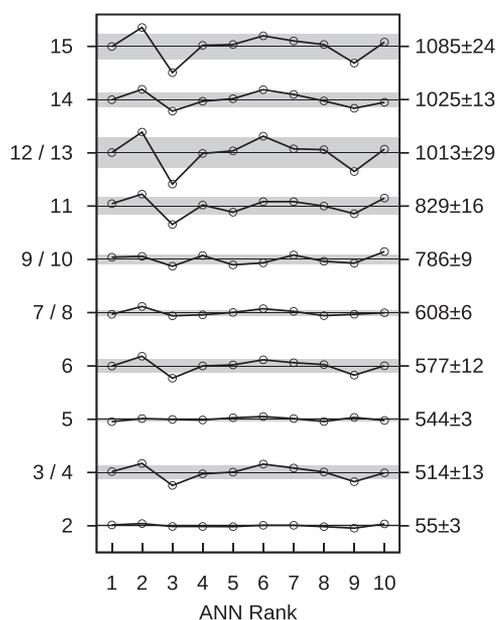


FIG. 2. Visual representation of the first 15 states of Table IV, showing deviations of the 10 best networks from respective average state energies. Level distances, not to scale, correspond to a deviation of 100 cm^{-1} from average. Gray: Standard deviation σ_i around average energies. Excitation energies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

Complete Nuclear Permutation Invariant Artificial Neural Network (CNPI-ANN) Diabatization for the Accurate Treatment of Vibronic Coupling Problems

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

David M. G. Williams* and Wolfgang Eisfeld*

Cite This: <https://dx.doi.org/10.1021/acs.jpca.0c05991>

Read Online

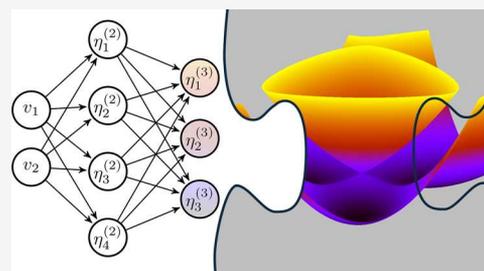
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: A recently developed scheme to produce accurate high-dimensional coupled diabatic potential energy surfaces (PESs) based on artificial neural networks (ANNs) [*J. Chem. Phys.* **2018**, *149*, 204106 and *J. Chem. Phys.* **2019**, *151*, 164118] is modified to account for the proper complete nuclear permutation inversion (CNPI) invariance. This new approach cures the problem intrinsic to the highly flexible ANN representation of diabatic PESs to account for the proper molecular symmetry accurately. It turns out that the use of CNPI invariants as coordinates for the input layer of the ANN leads to a much more compact and thus more efficient representation of the diabatic PES model without any loss of accuracy. In connection with a properly symmetrized vibronic coupling reference model, which is modified by the output neurons of the CNPI-ANN, the resulting adiabatic PESs show perfect symmetry and high accuracy. In the present paper, the new approach will be described and thoroughly tested. The test case is the representation and corresponding vibrational/vibronic nuclear dynamics of the low-lying electronic states of planar NO_3 for which a large number of ab initio data is available. Thus, the present results can be compared directly with the previous studies.



1. INTRODUCTION

The fundamental theoretical understanding of chemical processes by quantum dynamics simulations depends critically on the availability of accurate models for the molecular potential energy, the so-called potential energy surfaces (PESs). The development of single, uncoupled PESs for systems of more than three atoms made impressive progress in the past couple of decades.^{1–6} Particularly the use of artificial neural networks (ANNs) seems to be a very promising approach to achieve high accuracy even for high-dimensional problems.^{7–24} The decisive advantage of ANNs may be that it can be shown mathematically that they are capable of uniformly representing any continuous real function of n dimensions up to arbitrary accuracy.²⁵ Thus, in principle, ANNs should be capable of representing arbitrary PESs up to the limitations of the underlying data. Besides the potential accuracy, they also can be evaluated very efficiently once trained, as their evaluation consists mostly of matrix-vector multiplications. This is an invaluable advantage for the use with quantum dynamics methods such as the multiconfiguration time-dependent Hartree (MCTDH) method,^{26,27} especially when using the correlated discrete variable representation (CDVR)²⁸ scheme for which the evaluation of the PES is the most time-demanding part. Alternatively to the CDVR approach, a sum of products form of the PES model must be used, which is also possible in the ANN framework.^{16,24,29} A

further advantage is that the training of ANNs can be done very efficiently and is much less demanding than other nonlinear optimization approaches. Of course there are also problems of which one of the most prominent is an issue commonly referred to as “over-fitting”. This algorithmic problem, related to the extremely high flexibility and nonlinearity of typical ANNs, limits the accuracy of the function representation and has to be accounted for in the training algorithm. However, the impressive results already obtained by the ANN representation of single molecular PESs based on high-level ab initio data demonstrate the power of this approach.^{14,23,24,29–32}

The situation is way more complicated in the case that multiple and strongly interacting electronic states are involved in the process. For such nonadiabatic processes, not only the PESs corresponding to different states but also the interactions between those states have to be represented by a PES model. One fundamental problem is here that the Born–Oppenheimer

Received: June 30, 2020

Revised: July 31, 2020

Published: August 11, 2020

approximation will break down in regions in nuclear configuration space (NCS), where interactions among electronic states become significant or even singular as is the case for conical intersections.³³ It is well-established by now that the solution in such cases is to use a *quasi-diabatic* representation of the coupled electronic states and corresponding PES model.^{33–49} For the sake of brevity, we will refer to quasi-diabatic representations as “diabatic” representations from here on. The diabatic representation removes singularities in the nonadiabatic couplings and nondifferentiable points on the PESs but is neither unique nor easily available. Many methods have been developed to find the unitary basis transformation between the adiabatic states and energies directly available from ab initio calculations and diabatic states and PES matrices.^{39,42,43,45,47–76} One particularly successful and straightforward approach is the multimode linear vibronic coupling method by Köppel, Domcke, and Cederbaum⁵⁰ and similar approaches.^{70,72,77,78} These models are sufficient to treat many ultrafast nonadiabatic processes but are not accurate enough for many problems that are more complicated and depend critically on more extended regions in the NCS. Thus, the vibronic coupling approach has been extended to higher coupling orders as well.^{84–66,68,69,74} However, the use of higher-order Taylor expansions for the representation of the diabatic PES matrices is quite cumbersome, and therefore alternative representations have been developed as well.^{57,79–86}

The latest development in the pursuit of finding better diabatic PES models is the use of ANNs.^{87–97} Different strategies have been followed of which some are based on the use of complete nuclear permutation inversion (CNPI) invariant polynomials of the input coordinates.^{89,91,92} Our recently developed approach presented in ref 90 and thoroughly analyzed in ref 95 is based on a different strategy and uses all symmetry-adapted valence coordinates of the molecule rather than CNPI invariants. The reason for this is that the approach combines the basic, simple structure of a low-order vibronic coupling *ansatz*, which is tuned by an ANN so that very accurate results are obtained. The high flexibility of the ANN introduces slight symmetry contaminations, which are removed by an a posteriori symmetrization. The new method presented here circumvents the problem of the symmetry contamination by using CNPI invariants as an input layer of the ANN in combination with a standard low-order vibronic coupling reference model.

To stay consistent with our previous studies,^{90,95} the performance of the new approach was tested using planar (SD) NO₃, for which plenty of data and experience are available in the group.^{64,65,74,98} A 5 × 5 diabatic model for the ground and first two excited electronic states of ²A₂′, ²E′′, and ²E′ symmetry is trained, and vibrational/vibronic eigenstates are computed. The latter can be compared to previous theoretical^{69,74,95,99–105} and experimental results.^{106,107}

2. THEORY

2.1. Adiabatic and Diabatic Representation. At the center of any diabaticization approach is the adiabatic-to-diabatic transformation that connects an adiabatic and a diabatic basis of electronic states. There are many established ways to do this, and for a deeper discussion we refer to the literature, for example, ref 33. The necessary information for this basis transformation usually is obtained from adiabatic energies only (“diabatization by *ansatz*”), by using the adiabatic wave function representation alone (“block diagonalization”), or by a combination of both

(“hybrid diabaticization”).⁷⁶ The method presented here is designed for the hybrid diabaticization approach, though in the example calculations, the wave function information is not used. The fundamental requirement utilized in this method is that the adiabatic and diabatic PES matrices are related by the unitary transformation **U** that diagonalizes the diabatic matrix, and the eigenvalues reproduce the adiabatic energies

$$\mathbf{U}^\dagger \mathbf{W}^d \mathbf{U} = \mathbf{W}^a = \text{diag}(W_j^a) \quad (1)$$

The eigenvectors are related to the adiabatic electronic wave functions and can also be used to determine the matrix elements of the diabatic PES matrix **W**^d.

2.2. Artificial Neural Networks. The diabaticization approach presented here utilizes a multilayer perceptron feed-forward neural network. The input layer $\boldsymbol{\eta}^{(1)}$ consists of the molecular coordinates (cf. Section 2.3), which are connected to the hidden layers $\boldsymbol{\eta}^{(k)}$ according to

$$\eta_j^{(k)} = f^{(k)} \left(\underbrace{\beta_j^{(k)} + \sum_l \omega_{jl}^{(k)} \eta_l^{(k-1)}}_{f^{(k)}(\chi_j^{(k)})} \right) \quad (2)$$

The activation functions $f^{(k)}$ used in the present version depend on the layer k . While the output layer ($k = 3$) uses the identity Id, the activation function for the hidden layer reads

$$f^{(2)}(\chi) = \sqrt{3} \cdot \tanh(A \cdot \chi) \quad (3)$$

with $A = \text{artanh}\left(\frac{1}{\sqrt{3}}\right)$. This choice of factors rescales the sigmoid activation function in such a way that its “linear domain” resides in the interval $[-1, 1]$, approximating the identity, while the exponential plateaus reside outside of this interval. The output layer is used to parametrize the diabatic PES matrix (cf. Section 2.4). A simplified version of the neural network utilized (with a reduced number of neurons in each layer) can be found in Figure 1.

The ANN is trained by a modified Marquardt–Levenberg back-propagation method described in ref 90. The modification is necessary, because the output layer cannot be compared directly to the reference data. Only the eigenvalues (and eigenvectors) of the matrix parametrized by the output neurons

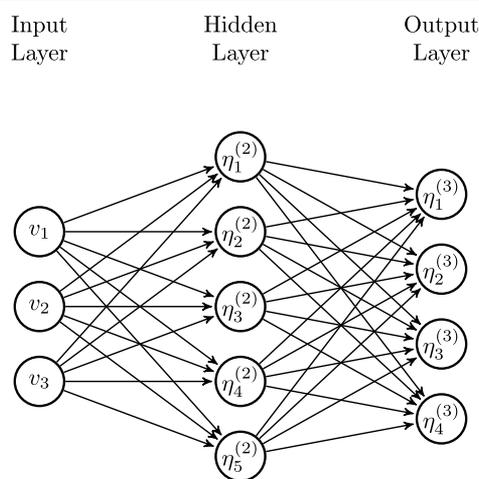


Figure 1. A neural network with one hidden layer. Arrows correspond to weighted contributions. Bias terms and activation functions are not visualized.

can be used to determine the root-mean-square error (RMSE), which is minimized by the method. Therefore, an additional numerical differentiation of the eigenvalues (and eigenvectors) with respect to the output neurons is required.

One notorious problem with the ANN training commonly is called overfitting and is related to the very high flexibility of the ANN. Since the training data are limited, an optimal representation of this reference data might result in a poor representation of data not contained in the reference set. This can be avoided, at least to some degree, by a method called “early stopping”. For this, the complete available data set is partitioned into a reference and a validation set. Only the reference set is used for the actual training to minimize the reference RMSE. In each step also the validation RMSE is computed, and if the latter rises, the training is stopped. The details of how the early stopping is implemented in the present method has been described in ref 90.

While overfitting is a crucial aspect, it is naturally not the only limitation to consider. ANNs are generally rather sensitive toward the data set with which they are trained. A simple, yet important example of this sensitivity lies in the sheer amount of data required, as neural networks tend to extrapolate data rather poorly. However, the use of a robust low-level reference model tuned by an ANN as presented here at least ensures that the extrapolated regions behave reasonably without showing pathologies. Another, more subtle problem related to the input data and the ANN training arises from the range of the input data. Since the present training method relies on derivatives of the activation function with respect to (wrt) its argument, initial convergence will be slowed tremendously if the derivatives are very small. Given the present activation eq 3, this means that χ should, on average, reside in the interval $[-1,1]$. For this purpose, the molecular coordinates $\eta^{(1)}$ are scaled and shifted in such a manner that all $\eta_j^{(1)}$ have a mean of 0 and a standard deviation of 1 across the presented data set. This is achieved by first computing the mean $\bar{\eta}_j$ and standard deviation $\sigma(\eta_j)$ of the coordinates $\eta_j^{(1)}$ over the training data; in other words

$$\bar{\eta}_j = \frac{1}{N_D} \sum_{k=1}^{N_D} \eta_j^{(1)}(\mathbf{Q}_k) \quad (4a)$$

$$\sigma(\eta_j) = \sqrt{\frac{1}{N_D} \sum_{k=1}^{N_D} (\eta_j^{(1)}(\mathbf{Q}_k) - \bar{\eta}_j)^2} \quad (4b)$$

where N_D is the number of ab initio data points, and $\eta_j^{(1)}(\mathbf{Q}_k)$ is the explicit value of $\eta_j^{(1)}$ at the given data point. The normalized input coordinates $\tilde{\eta}_j$ actually used in the training process are then obtained by

$$\tilde{\eta}_j = \frac{\eta_j^{(1)} - \bar{\eta}_j}{\sigma(\eta_j)} \quad (5)$$

which by definition has an average of 0 and a standard deviation of 1 over the given set of ab initio data. Once the training is completed, the optimized parameters corresponding to $\omega_{jl}^{(2)}$ and $\beta_j^{(2)}$ from Equation 2 are modified to obtain an equivalent neural network depending on the original coordinates $\eta_j^{(1)}$, eliminating the need to compute the means and variances when evaluating the PES model during dynamics calculations. The equivalent parameters $\tilde{\omega}_{jl}^{(2)}$ and $\tilde{\beta}_j^{(2)}$ follow from Equation 2 for $k = 2$ by

substituting $\eta_j^{(1)}$ with its normalized version of Equation 5 and read

$$\tilde{\omega}_{jl}^{(2)} = \frac{\omega_{jl}^{(2)}}{\sigma(\eta_l)} \quad (6a)$$

$$\begin{aligned} \tilde{\beta}_j^{(2)} &= \beta_j^{(2)} - \sum_l \frac{\omega_{jl}^{(2)} \bar{\eta}_l}{\sigma(\eta_l)} \\ &= \beta_j^{(2)} - \sum_l \tilde{\omega}_{jl}^{(2)} \bar{\eta}_l \end{aligned} \quad (6b)$$

In the present work this preconditioning is combined with an appropriate choice of random initial weights and biases to ensure better convergence.

2.3. Coordinates and CNPI Invariants. The *ansatz* diabatic PES matrix has to fulfill strict symmetry requirements for the matrix elements, because the represented Hamiltonian must be invariant under all symmetry transformations. However, the off-diagonal elements are not necessarily totally symmetric, which has to be taken into account. In standard vibronic coupling approaches the matrix elements are usually expressed as polynomials in terms of symmetry-adapted coordinates Q according to

$$W_{kj}^d(Q) = \sum_{\alpha} p_{\alpha}^{kj} \prod_l Q_l^{n_{\alpha l}} \quad (7)$$

The expansion coefficients p are strictly related by symmetry, and the coordinates transform according to different irreducible representations of the molecular point group. Alternatively, the diabatic matrix can be constructed from CNPI invariants of the actual CNPI group and the invariants of its permutation subgroups.^{91,108}

The present approach is illustrated for the D_{3h} symmetric NO_3 radical in the following. In this case it is convenient to use symmetry-adapted coordinates for the representation of the vibronic coupling model, because the groups D_{3h} and $S_3 \times C_i$ are isomorphic. The five symmetry-adapted coordinates excluding the umbrella motion read

$$a = \sqrt{\frac{1}{3}} (m_1 + m_2 + m_3) \quad (8a)$$

$$x_s = \sqrt{\frac{1}{6}} (2m_1 - m_2 - m_3) \quad (8b)$$

$$y_s = \sqrt{\frac{1}{2}} (m_2 - m_3) \quad (8c)$$

$$x_b = \sqrt{\frac{1}{6}} (2\alpha'_1 - \alpha'_2 - \alpha'_3) \quad (8d)$$

$$y_b = \sqrt{\frac{1}{2}} (\alpha'_2 - \alpha'_3) \quad (8e)$$

in which the transformed distances m_i and angles α'_i are

$$m_i = 1 - \exp(-\gamma(r_i - r_0)) \quad (9a)$$

$$\alpha'_i = \frac{\alpha_i - \alpha_0}{r_i r_j r_k}, \quad i \neq j \neq k \quad (9b)$$

The coordinate a transforms totally symmetric as a_1' , and the vectors (x, y) transform as e' . The resulting invariants are determined as^{64,65,74,109,110}

$$v^{(0)} = 1 \quad (10a)$$

$$v^{(1)} = 0 \quad (10b)$$

$$v^{(2)} = (x^2 + y^2) \quad (10c)$$

$$v^{(3)} = (2x^3 - 6xy^2) \quad (10d)$$

$$v^{(4)} = (x^4 + 2x^2y^2 + y^4) \quad (10e)$$

⋮

for the first few orders, and the mixed terms of two different e' modes read

$$v_{ee}^{(2,1)} = 2(x_1x_2 + y_1y_2) \quad (11a)$$

$$v_{ee}^{(3,1)} = 2(x_1x_2^2 - x_1y_2^2 - 2x_2y_1y_2) \quad (11b)$$

$$v_{ee}^{(3,2)} = 2(x_1^2x_2 - x_2y_1^2 - 2x_1y_1y_2) \quad (11c)$$

$$v_{ee}^{(4,1)} = 2(x_1^2x_2^2 - x_1^2y_2^2 - x_2^2y_1^2 + y_1^2y_2^2 + 4x_1x_2y_1y_2) \quad (11d)$$

$$v_{ee}^{(4,2)} = 2(x_1^3x_2 + x_1^2y_1y_2 + x_1x_2y_1^2 + y_1^3y_2) \quad (11e)$$

$$v_{ee}^{(4,3)} = 2(x_2^3x_1 + x_1x_2y_2^2 + x_2^2y_1y_2 + y_2^3y_1) \quad (11f)$$

$$v_{ee}^{(4,4)} = x_1^2x_2^2 + x_1^2y_2^2 + x_2^2y_1^2 + y_1^2y_2^2 \quad (11g)$$

⋮

Each of the above invariants may be multiplied by an arbitrary power of the a mode and still remain an invariant. Similarly, if we were to include the umbrella motion θ , we could also multiply each term with arbitrary powers of θ^2 . Since for NO_3 the CNPI group and the D_{3h} point group are isomorphic, the above point group invariants can be used for convenience. In cases in which point group invariants are not sufficient, the CNPI invariants may be constructed following the techniques developed by Braams and Bowman.⁶ Finally, it may be worth noting that, in the present planar case, the CNPI and CNP symmetries are equivalent.

2.4. Diabatic Model (Ansatz). The new approach will be demonstrated using a 5×5 diabatic PES model to represent the electronic Hamiltonian corresponding to the first few electronic states of planar NO_3 , namely, \tilde{X}^2A_2' , \tilde{A}^2E'' , and \tilde{B}^2E' . The diabatic PES model is based on a simple vibronic reference model to simulate the basic physics of the problem, which then is tuned by the ANN to achieve high accuracy. This reference model is expressed in terms of a set of coupling matrices depending on the symmetry-adapted coordinates Q parametrized by expansion coefficients λ and reads

$$\mathbf{W}^d(\mathbf{Q}) = \mathbf{W}_{\text{diag}}^d(\mathbf{Q}) + \begin{pmatrix} \lambda_1 & 0^T & 0^T \\ 0 & \lambda_2 \mathbf{1} & 0 \\ 0 & 0 & \lambda_3 \mathbf{1} \end{pmatrix} + \begin{pmatrix} 0 & 0^T & 0^T \\ 0 & \lambda_4 \mathbf{e}_s + \lambda_5 \mathbf{e}_b & 0 \\ 0 & 0 & \lambda_6 \mathbf{e}_s + \lambda_7 \mathbf{e}_b \end{pmatrix} + \begin{pmatrix} 0 & 0^T & \lambda_8 \boldsymbol{\rho}_s^T + \lambda_9 \boldsymbol{\rho}_b^T \\ 0 & 0 & 0 \\ \lambda_8 \boldsymbol{\rho}_s + \lambda_9 \boldsymbol{\rho}_b & 0 & 0 \end{pmatrix} \quad (12)$$

$\mathbf{1}$ in the above supermatrix is a 2×2 unit matrix, the matrix blocks $\mathbf{e}_{s,b}$ are the well-known linear Jahn–Teller coupling matrices, and the vectors $\boldsymbol{\rho}_{s,b}$ represent the linear pseudo-Jahn–Teller couplings, which are given explicitly as

$$\mathbf{e}_{s,b} = \begin{pmatrix} x_{s,b} & y_{s,b} \\ y_{s,b} & -x_{s,b} \end{pmatrix} \text{ and } \boldsymbol{\rho}_{s,b} = \begin{pmatrix} y_{s,b} \\ x_{s,b} \end{pmatrix} \quad (13)$$

The symmetry-adapted coordinates corresponding to the degenerate asymmetric bending and stretching modes are labeled $x_{s,b}$, $y_{s,b}$, respectively. The diabatic PES matrix eq 12 can be recast in a more compact form as

$$\mathbf{W}^d(\mathbf{Q}) = \sum_L \lambda_L(\mathbf{v}) \cdot \mathbf{M}_L(\mathbf{Q}) \quad (14)$$

utilizing explicit coupling matrices \mathbf{M}_L . The ANN is introduced into this model through the expansion coefficients

$$\lambda_L(\mathbf{v}) = \begin{cases} \lambda_L^0 & \text{for uncorrected terms} \\ \lambda_L^0 \cdot (1 + \eta_L^{(f)}(\mathbf{v})) & \text{otherwise} \end{cases} \quad (15)$$

The constants λ_L^0 , obtained from standard nonlinear fitting procedures, are tuned by the outputs $\eta_L^{(f)}(\mathbf{v})$ of a trained ANN.

Note that, in our previous studies, the ANN was fed with all symmetry-adapted coordinates Q , while in the present approach the invariants v are used. This turns out to make a significant difference. Though the vibronic coupling model with all output neurons $\eta_L^{(f)}$ set to zero fulfills the symmetry transformation invariance exactly, this is not the case anymore when the output neurons depend on the symmetry-adapted coordinates Q . One could say that the ANN simply does not know about the symmetry of the problem except for what it can “learn” from the reference data. The errors are very small ($\sim 1 \text{ cm}^{-1}$) and have been removed by an a posteriori resymmetrization in our previous studies. However, the new approach presented here deletes this issue intrinsically by feeding the ANN with invariants rather than symmetry-adapted coordinates. Thus, on the one hand, regardless of the various equivalent realizations of a molecular geometry related by symmetry transformations, the values of the invariants are the same and thus enforce the proper symmetry of the model. On the other hand, it is not obvious whether or not the invariants contain all the coordinate information represented by the symmetry-adapted coordinates. The lowest-order invariants are a , $v^{(2)}(x_s, y_s)$, $v^{(2)}(x_b, y_b)$, and $v_{ee}^{(2,1)}(x_s, y_s, x_b, y_b)$. This means that there are only four instead of five coordinates available. Furthermore, these invariants do only represent information about the radial displacement of the

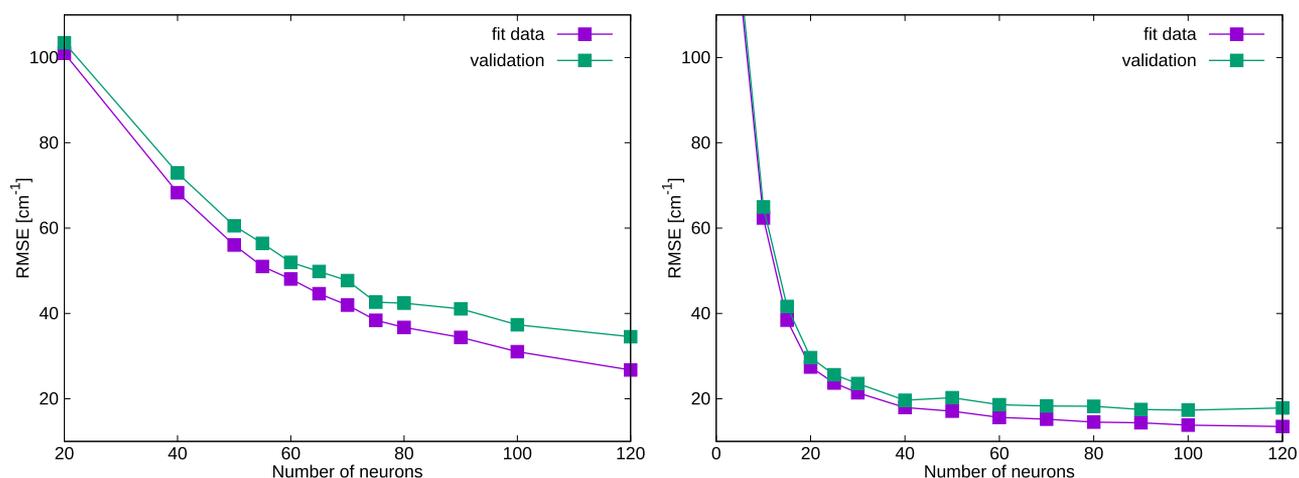


Figure 2. Weighted RMSE of the best neural network of a given set as a function of the hidden-layer size. Purple line represents the fitting set error; green line represents the validation set error. (left) Previous model. (right) New model.

molecular geometry with respect to the conical intersection but none about the tangential displacement with respect to the three equivalent positions around the conical intersection.¹¹⁰ The missing information can be retrieved by utilizing the next higher order of invariants as will be demonstrated by numerical tests. It turns out that adding the four third-order invariants provides sufficient input information for an accurate representation of the full diabatic PES model.

Finally, the cubic terms in the invariants on the input layer of the ANN cause small numerical errors, which only show in the detailed results of the quantum dynamics calculations for the vibrational/vibronic levels of the system. This is avoided by damping the invariants according to

$$v'_k(Q) = c_k \tanh(d_k v_k(Q)) \quad (16)$$

where c_k and d_k are empirical scaling factors chosen according to the physical problem to be treated. A reasonable choice for this scaling appears to be $c_k = d_k^{-1} = 10$, and this assumption is tested numerically in Section 4.3.

This diabatic PES model is designed such that the stability and basic qualitative description of a low-order vibronic coupling expansion is combined with the flexibility of the ANN to achieve excellent accuracy. We found in previous studies that the basic reference model should be restricted to the lowest reasonable orders to avoid numerical problems in the ANN training. This also ensures that the model is free of pathologies. The accuracy is not limited by this reference model but by the ANN tuning the corresponding coupling parameters according to eq 15. What remains is to fit the reference parameters λ_L^0 and to train the ANN with respect to ab initio reference data.

3. COMPUTATIONAL DETAILS

The present method is tested by representing the 5×5 diabatic Hamiltonian for the NO_3 radical. The training of the ANN is based on an extensive set of ab initio reference energies computed by Multiconfiguration Reference Singles and Doubles Configuration Interaction (MR-SDCI) calculations based on Complete Active Space Self-Consistent Field (CASSCF) reference wave functions using a slightly adapted correlation consistent aug-cc-pVTZ standard basis.¹¹¹ All details of the ab initio computations can be found in refs 74, 98, 111, and 112. The sampling approach for the data points is based on the method described in refs 113 and 90.

These reference data are used first to fit the 21 free parameters λ_j^0 of the reference model using a Marquardt–Levenberg algorithm incorporated into a genetic algorithm with a resulting RMSE of $\sim 1700 \text{ cm}^{-1}$. This is the same reference model as used before. Then the reference data set is partitioned into a training set of 85% of the data and an external validation set of the remaining 15% of the energies. The training set is used to train CNPI-ANNs with a single hidden layer of various sizes. The validation set is used in the early stopping mechanism to prevent overfitting as suggested in ref 114. The specific way how early stopping is used here has been described in ref 90. All data are energy weighted as described before in order to get the best accuracy of the PES model in the region of the NCS most relevant for the nuclear dynamics.

The resulting PES models were tested by computing vibrational/vibronic energy levels on the adiabatic ground state as well as the first electronically excited state, which is separable for planar geometries. The second excited state of ${}^2E'$ symmetry cannot be tested easily by the dynamics calculations, because it is coupled to the ${}^2A_2'$ ground state by *pseudo*-JT coupling, and thus vibronic eigenstates would not be accessible easily and therefore are not computed in the present study. The computation of the vibrational/vibronic energy levels was performed using a time-independent Hermite discrete variable representation (Hermite-DVR) method,¹¹⁵ in which the corresponding Hamiltonian was diagonalized by an exact short iterative Lanczos method. Further details can be found in refs 74, 90, 98, and 105.

4. RESULTS AND DISCUSSION

The new scheme is tested and compared to its predecessor laid out in refs 90 and 95. First, both are compared in terms of fitting performance and convergence with respect to the number of parameters needed. In a second step, both models are compared in terms of the dynamics results for both ${}^2A_2'$ ground and ${}^2E''$ excited states, respectively. Finally, the new approach will be assessed by its capability of reproducing the available experimental data,^{106,107} in comparison to previous efforts.

While various ANN parametrizations will be used to provide deeper insights regarding the stability of the produced results (wrt the initial guess), each scheme will be primarily represented by one specific parametrization as a reference case, which turned out to be particularly suitable for the present case. The previous

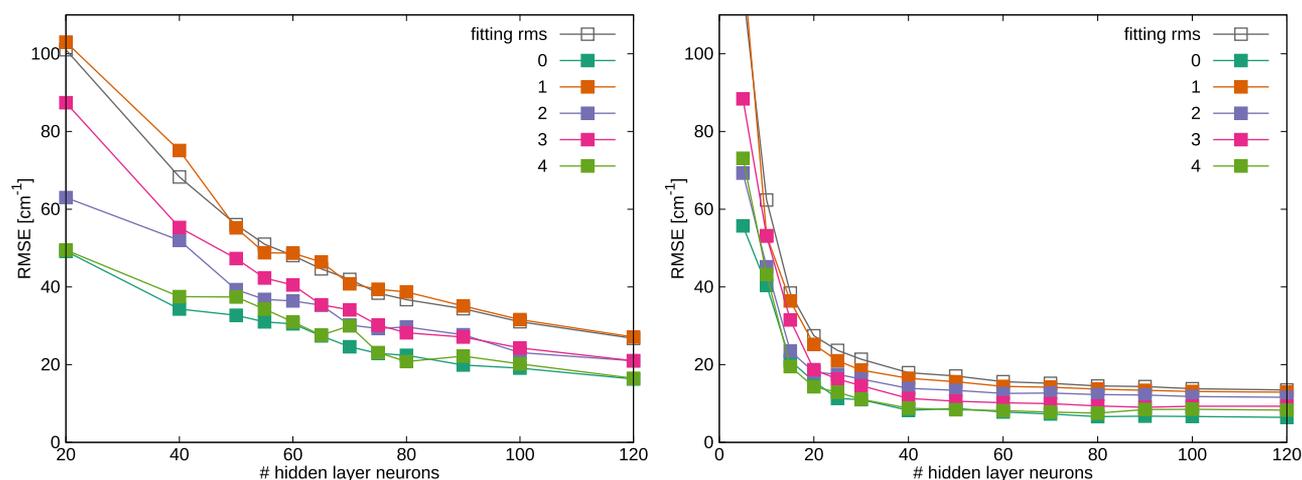


Figure 3. State-resolved convergence behavior of unweighted RMSE for ab initio data points up to 1 eV above respective reference point energy. States are enumerated from the ground state to the energetically highest lying state. Gray: weighted fitting set error. (left) Previous model. (right) Current model.

method will be represented by an ANN model with a single hidden layer of 75 neurons (corresponding to 1134 formal parameters), and the new model has a single hidden layer of 30 neurons (corresponding to 549 formal parameters).

4.1. Fitting Performance. The first and somewhat surprising observation is that the new ANN setup outperforms the previous efforts significantly in terms of compactness and overall fitting results. Various neural network sizes between 5 and 120 hidden-layer neurons have been tested. The convergence behavior of the best obtained RMS error (as well as the corresponding validation error) is presented in Figure 2, together with the analogous graph from ref 90. Not only does the new scheme converge to roughly half of the previous fitting error, convergence itself occurs significantly faster with respect to the ANN size. As in previous work, a sudden increase of the difference between fitting error and validation error is observed at a certain number of hidden-layer neurons. This is a sign for overfitting, and ANN models with a larger hidden layer are considered less reliable. This point is reached for 50 hidden-layer neurons with the new approach compared to 80 neurons for the initial method. In the present study we will focus on the best fitted 30 neuron network, with a fitting RMS of 21.4 cm^{-1} and a validation error of 23.6 cm^{-1} , respectively. This was chosen instead of the 40 neuron network, because both result in almost indistinguishable results, and the smaller ANN is more efficient to test thoroughly. It also should be pointed out that the reference data are obtained from very intricate ab initio calculations rather than an analytic test model, and thus the remaining errors not only reflect possible shortcomings of the present model but also the errors intrinsic to the multireference configuration interaction (MRCI) calculations. We are inclined to estimate that an RMSE on the order of $10\text{--}20 \text{ cm}^{-1}$ might be the lower limit of what is possible to achieve with the present data set.

While the total root-mean-square error provides a measure for how well the given ab initio data are reproduced overall, it provides no information regarding how deviations are distributed. For this purpose, unweighted RMS errors were computed for each adiabatic energy separately below a certain threshold. This threshold was chosen the same way it was in previous work,⁹⁰ 1 eV above the reference point energies, to obtain a measure of how well the model performs in the

physically relevant regions of the individual sheets. A comparison between the convergence behavior of both present and previous models can be found in Figure 3. Apart from inheriting the same smoother, more well-behaved manner of convergence, state-specific errors produced by the new method are significantly closer together, meaning the overall error is far more evenly distributed across the individual electronic states. This is in stark contrast to the previous model, where the error seems largely dominated by the error of the lower ${}^2E''$ sheet. A reasonable explanation for this is that, while its intricate triple-well structure remains the topographically most demanding to reproduce, the inherent symmetry of the modified input coordinates greatly reduces the overall strain on the network. This also explains the significant reduction of required parameters (less than half of the previous reference ANN surface).

4.2. Impact of the Hidden-Layer Size on Dynamics.

While representing ab initio data reliably is a central aspect and strong point of the presented method, it is equally essential for the resulting models to behave consistently in dynamics calculations. To this end, in a first step, fits of neural networks are considered, which differ only in which particular (randomly generated) initial guess was used in the training for the neural network parameters. As mentioned above, these tests are performed with 30 as the number of hidden-layer neurons for this particular study. The resulting surfaces can vary significantly, as the highly nonlinear nature of the model allows for a great deal of different realizations of nearly identical fitting errors. The study itself is performed in analogy to that of ref 95, taking the 10 best-fitted networks (in terms of the RMS error) from a set of 100. These networks only differ marginally in terms of the fitting error, ranging from 21.4 to 22.2 cm^{-1} , meaning they can be considered roughly identical in terms of quality. They are enumerated by their numerical “rank” (starting at 1 for the best network) for the sake of convenience in the following. Figure 4 shows an excerpt of the data resulting from computing vibrational transition energies for the $\tilde{X} \text{ } {}^2A_2'$ state, together with the obtained average transition energies as well as respective standard deviations σ_i . The latter are of particular interest, as they provide a simple measure for how dynamics results “spread” among samples of very similar networks. In other words, these σ_i values are a measure for what kind of

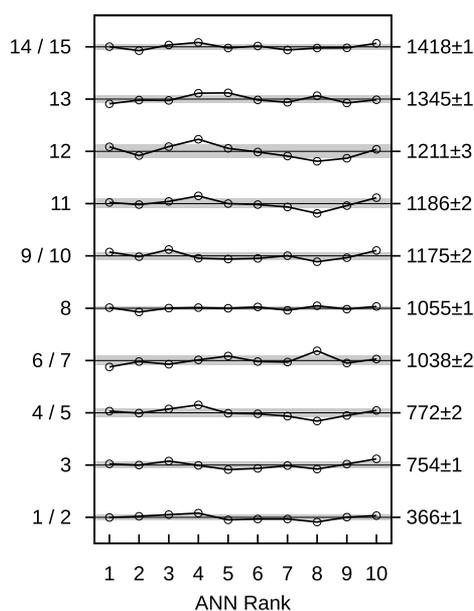


Figure 4. Comparison of the computed transition energies for the $\tilde{X} \ ^2A_2'$ state of $^{14}\text{NO}_3$ (in cm^{-1}) for the 10 best (in terms of the fitting error) ANNs from a set of 100 fitted neural networks, including the reference ANN. Level distances, not to scale, correspond to a deviation 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 inverse centimeters.

deviations are a priori indistinguishable from competing networks of the same architecture as our reference, which is why excitation energies will be denoted with a $\pm\sigma_i$ when appropriate.

Overall, the dynamics results obtained from the new 30 neuron model closely resemble those from the previous 75 hidden-layer neuron model both with respect to the average energies produced as well as with respect to the corresponding standard deviations. Standard deviations are close to 2 cm^{-1} for all 18 states considered in the present work. This more than likely confirms the previous conclusions that the previous model already exhausted the limits of what accuracy can be achieved within the confines of the ab initio data presently available. This indicates that the new approach inherits one of the most favorable aspects of its predecessor, in particular, the consistency of the dynamics results produced and thus the overall tendency of the method to produce robust results of excellent quality.

With the standard deviations σ_i as a measure for significant deviations from the reference network at hand, the findings presented in Table 1 can be contextualized.

The 20 neuron network, with its significantly smaller number of formal parameters, results in vibrational level energies deviating significantly for all states known to be affected by insufficiently flexible ANN models in the past (in particular, 4/5, 11, and 12), suggesting that these states are particularly sensitive regardless of the specifics of the model used. States 9/10, 13, and 14/15 are additionally affected, further suggesting that this network size is insufficient to reproduce the lowest adiabatic sheet accurately. The 25 neuron network, while being more consistent with the reference network overall, still shows significant deviations for states 12 and 13 as well as 6/7, 16/17, and 18, in particular, which were previously indistinguishable from the reference. Most interestingly, the 40 neuron network (corresponding to an increase of formal parameters by $\sim 30\%$)

Table 1. Comparison of Computed Vibrational Transition Energies^a

state	30	20	25	40	60	80
	$E_i - E_0 \pm \sigma_i$	$\delta(E_i - E_0)$				
1/2	365.8 ± 1.0	-0.9	0.2	∴ 0.5	1.4	1.2
3	754.6 ± 1.3	-1.5	-0.5	∴ -0.2	-0.2	0.9
4/5	772.5 ± 1.7	-3.2	-0.8	∴ -0.5	1.2	1.5
6/7	1035.5 ± 1.7	0.4	2.6	∴ 1.9	2.1	2.2
8	1055.6 ± 0.7	-1.7	-1.0	∴ -1.2	-1.2	-1.1
9/10	1176.6 ± 1.5	-2.8	-1.1	∴ -1.9	-1.8	-0.4
11	1186.6 ± 1.9	-2.7	0.4	∴ -0.8	1.8	1.3
12	1212.3 ± 2.5	-6.1	-3.1	∴ -2.3	0.6	1.7
13	1342.7 ± 1.5	-2.3	2.5	∴ 0.5	1.2	1.9
14/15	1417.6 ± 1.0	-2.5	-0.7	∴ -0.7	-0.2	0.0
16/17	1488.1 ± 0.8	0.1	1.8	∴ -1.1	-0.1	0.0
18	1489.8 ± 2.1	0.0	3.6	∴ 0.2	2.8	2.8

^aComparison of the computed vibrational transition energies for the $\tilde{X} \ ^2A_2'$ state of $^{14}\text{NO}_3$ (in cm^{-1}) for different hidden-layer sizes. The vibrational excitation energies of the reference (30) 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 reference network (30). The dotted vertical line indicates the separation between increased and decreased hidden-layer sizes with respect to the reference.

shows little to no significant change when compared to the reference network. This was the reason why the less flexible 30 neuron ANN was chosen as reference model, because it is the more rigorous test for the efficiency of the new model. It turns out that network sizes depending on more than twice the number of formal parameters than the reference network (such as 60 and 80 hidden-layer neurons) show very few significant deviations, with state 18 featuring the largest deviation overall. This is surprising, as overfitting effects were much more apparent in the previous model scheme, where the molecule's symmetry was not directly taken into account by construction. This change in behavior suggests that, at least for the sample of states considered, reproducing the adiabatic ground-state PES accurately is less indicative of the model's overall performance. Apparently, the new CNPI-ANN approach greatly reduces the strain on the network to reproduce the comparably simple, highly symmetric structure of this PES.

Moving on to the significantly more sensitive dynamics of the $^2E''$ first excited state, again first the 10 best networks are considered for the reference hidden-layer size (30 neurons) to estimate the size of a "significant deviation". A visualization of the first 15 states is again provided in Figure 5. Despite the overall improvement of the new model in terms of reproducing ab initio data for this particular electronic state, no change in the overall sensitivity of the resulting transition energies could be observed. However, while the standard deviations σ_i remain rather similar to those produced by the analogous study using the previous model and 75 neuron network for each state considered, even the average energies of the first few states (such as 3/4, 6, and 8) deviate up to 30 cm^{-1} . On the one hand, the latter is to be expected, as the changes in fit performance were the most significant for the coupled surfaces of the $^2E''$ state. The former, on the other hand, can easily be interpreted as the sensitivity of these states being intrinsic to the complexity of the particular system, combined with the limited availability of ab initio data. This interpretation is consistent with our findings in ref 95.

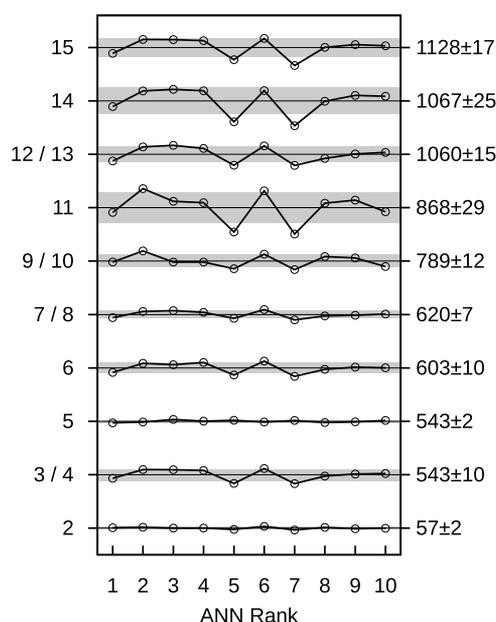


Figure 5. Comparison of the computed transition energies for $^{14}\text{NO}_3$ (in cm^{-1}) on the \tilde{A} state for the 10 best (in terms of fitting error) ANNs in analogy to Figure 4. Level distances, not to scale, correspond to a deviation of 100 cm^{-1} from average. Gray: Standard deviation σ_i around average energies. Excitation energies together with the deviation σ_i are given at the right-hand side.

The obtained standard deviations again provide a context in which the resulting vibronic excitation energies of different network sizes, summarized in Table 2, can be compared. In addition to the energies, this table also provides D_{3h} labels and state assignments based on symmetry information and projected wave function densities. As in the previous work,⁹⁵ state assignments in the degenerate stretch and bend coordinates are given in terms of radial $r_{s,b}^n$ and tangential excitations $t_{s,b}^n$.¹¹⁰ Excitations in the totally symmetric mode are denoted with a^n . While deviations are largely well below 40 cm^{-1} , the 80 neuron

network showed such immense overfitting artifacts that comparison was not deemed feasible, which is why it was omitted. Overall, deviations observed are rather large for different neural network sizes, in particular, for states corresponding to radial stretch excitations r_s . The fact that the radial stretch excitations remain sensitive despite the new model representing the related ab initio data significantly better than before further confirms the previous conclusions that this may be a limitation of the underlying data itself being insufficiently sampled in the dissociative domain of the nuclear configuration space. Apart from that, the overall size of the deviations is consistent with findings from the previous modeling scheme.

On the one hand, considering again the smaller networks first, the 20 hidden-layer neuron network exhibits the most significant deviations for the r_b fundamental and the $t_b^2 a_1''$ state (state 20), with most other states showing deviations close to their respective σ_i . The 25 neuron network, on the other hand, shows expectedly fewer deviations than the 20 neuron network, as the flexibility of the network increases. Both networks remain comparably more similar to the reference ANN overall, at least relative to the deviations found for the same ANN sizes and the corresponding dynamics results for the ground-state calculations.

Intriguingly, this pattern was also observed in the previous model. A plausible explanation for this is that the adiabatic ground state is more sensitive to underfitting, requiring a minimal flexibility to be reproduced though having a simple overall structure. Conversely, the \tilde{A} state appears to be more sensitive to overfitting, having the most complicated structure. This sensitivity toward overfitting can be observed when looking at the 60 neuron network. In this case, deviations are again larger, reaching up to 40 cm^{-1} , with state 20 again deviating the most, relative to the standard deviation. The 40 neuron network, while significantly deviating for a few states (state 20, in particular), is overall very similar to the reference network and can be seen as an overall equivalent choice.

In conclusion, our findings show that the new modeling scheme is overall very similar to its predecessor in terms of how

Table 2. Comparison of Computed Vibronic Transition Energies^a

state	assignment	30 $E_i - E_0 \pm \sigma_i$	20 $\delta(E_i - E_0)$	25 $\delta(E_i - E_0)$	40 $\delta(E_i - E_0)$	60 $\delta(E_i - E_0)$
2	$0 a_1''$	58.1 ± 2.0	2.2	-2.5	:	2.8
3/4	$r_s^1 e''$	536.7 ± 10.5	12.3	15.3	:	13.1
5	$t_b^1 a_2''$	539.5 ± 2.1	2.4	3.5	:	-1.6
6	$r_s^1 a_1''$	595.0 ± 9.9	13.4	17.0	:	14.8
7/8	$t_b^1 e''$	614.3 ± 6.5	11.6	7.3	:	10.0
9/10	$r_b^1 (*) e''$	787.3 ± 11.6	19.5	-0.3	:	3.3
11	$r_b^1 (*) a_1''$	859.4 ± 28.7	36.5	17.2	:	25.7
12	$r_s^1 t_b^1 a_2''$	1047.7 ± 14.7	20.4	23.9	:	17.1
13/14	$r_s^2 e''$	1056.4 ± 24.8	25.0	29.2	:	22.7
15	$r_s^2 a_1''$	1117.3 ± 17.4	26.2	20.7	:	19.3
16/17	$t_b^2 (*) e''$	1121.9 ± 18.5	29.6	30.1	:	28.9
18/19	$r_s^1 t_b^1 (*) e''$	1168.8 ± 5.9	5.3	13.5	:	3.5
20	$t_b^2 (*) a_1''$	1176.2 ± 13.6	25.0	19.9	:	20.2
21/22	$a^1 e''$	1176.3 ± 10.3	18.2	16.3	:	12.9
23	$t_s^1 a_2''$	1271.4 ± 6.7	5.7	8.3	:	-2.6

^aComparison of computed transition energies for $^{14}\text{NO}_3$ (in cm^{-1}) on the \tilde{A} state for different hidden-layer sizes. $\delta(E_i - E_0)$ provides level shifts compared to the reference network (30). 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). The dotted vertical line indicates the separation between increased and decreased hidden-layer sizes with respect to the reference.

it reacts to varying the hidden-layer size and overfitting, while being capable of reproducing ab initio reference data both more accurately and with less than half as many parameters. The primary difference to previous findings (in terms of behavior) is that the model shows a more clear-cut change in behavior as overfitting effects become dominant. However, it is important to note that this does not relevantly affect the stability of the method itself, as these effects occur only for more than twice the number of parameters needed to accurately reproduce the given data, and resides far beyond any suitable model choice. As a consequence, the 30 neuron model remains the focal point of this work, with the 40 neuron model yielding largely equivalent data at the cost of needing significantly more parameters.

4.3. Invariant Input Coordinates and Numerical Stability. So far, only the model's behavior with respect to the hidden-layer size of the neural network has been investigated. However, this is of course not the only free parameter related to the neural network that could be adapted. Most of these parameters, such as the number of hidden layers or particularities regarding the distribution of the data set into the reference and validation set already have been discussed at length in ref 90. What remains is to analyze the impact of the new parameters introduced in Equation 16. In the present case, these parameters were set in such a way that $c_k = d_k^{-1} = 10$. What this means is that $\nu_k'(Q)$ approximates $\nu_k(Q)$ for all values of $|\nu_k|$ less than c_k while approaching $\pm c_k$ as $\nu_k \rightarrow \pm\infty$. In other words, ν_k' serves as an asymptotically damped version of ν_k .

One of the great advantages of neural networks, as we have stressed before, is the homogeneous, bounded nature of the sigmoid functions they are ultimately composed of. This is in stark contrast to the terms of a polynomial expansion, which increase more rapidly in size for increasing polynomial orders, which can cause artifacts caused by rapid oscillations in the undersampled domains of a fit. However, some of this disadvantage is reintroduced by the use of higher-order invariants as input coordinates, which means that small changes in the primitive coordinates translate into large changes in the third-order coordinates. This appears to cause unnecessary numerical noise to the neural network, thus hindering (or even impeding) convergence. The above damping scheme was introduced to remove this problem and make the resulting ANN model more robust numerically. What remains is to address the question of how to choose the resulting parameters c_k . Since the goal of the damping scheme is to avoid artifacts caused by ill-behaved coordinates rapidly increasing in size, the present scheme aims to minimize the risk of numerical issues by keeping ν_k' as limited as possible. This means that the dynamically relevant region of the surface must be represented perfectly and that the damping does not interfere with that region.

Table 3 provides a summary of how the computed \tilde{A} state transitions change for the reference network size (30) for various choices of c_k . Again, labels and deviations wrt the reference network are provided in the same format as that of Table 2, including standard deviations σ_i as a measure for the expected deviations from the reference for equivalent networks (see Section 4.2). No significant impact on the \tilde{X} state could be observed. For this reason the present study will focus solely on the \tilde{A} state. Choosing a significantly smaller value of c_k than 10 causes the resulting network to become incapable of reproducing ab initio reference data, as small c_k begin to dampen ν_k for smaller and smaller displacements from the reference geometry. On the one hand, this loss of quality is so

Table 3. Comparison of Computed Transition Energies^a

state	assignment	$E_i - E_0 \pm \sigma_i$				
		$c_k = 10$	$c_k = 15$	$c_k = 20$	$c_k = 30$	$c_k = 40$
2	$0 a_1''$	58.1 ± 2.0	-0.9	-1.9	1.0	1.5
3/4	$r_s^1 e''$	536.7 ± 10.5	-0.7	8.1	7.6	16.2
5	$t_b^1 a_2''$	539.5 ± 2.1	-1.4	4.5	2.8	4.6
6	$r_s^1 a_1''$	595.0 ± 9.9	2.3	12.5	10.4	17.6
7/8	$t_b^1 e''$	614.3 ± 6.5	-3.0	6.1	6.9	13.6
9/10	$r_b^1 (*) e''$	787.3 ± 11.6	5.8	-6.0	-2.1	-2.3
11	$r_s^1 (*) a_1''$	859.4 ± 28.7	1.4	2.0	7.3	17.7
12	$r_s^1 t_b^1 a_2''$	1047.7 ± 14.7	-3.5	13.9	11.3	23.4
13/14	$r_s^2 e''$	1056.4 ± 24.8	-0.8	22.8	15.1	29.5
15	$r_s^2 a_1''$	1117.3 ± 17.4	1.9	19.2	14.2	25.6
16/17	$t_b^2 (*) e''$	1121.9 ± 18.5	-0.6	26.3	15.8	35.7
18/19	$r_s^1 t_b^1 (*) e''$	1168.8 ± 5.9	-1.8	7.6	2.3	8.8
20	$t_b^2 (*) a_1''$	1176.2 ± 13.6	2.4	14.5	11.0	21.5
21/22	$a^1 e''$	1176.3 ± 10.3	-3.1	19.6	12.8	27.7
23	$t_s^1 a_2''$	1271.4 ± 6.7	1.6	6.1	-2.0	-0.2

^aComparison of the computed transition energies for $^{14}\text{NO}_3$ (in cm^{-1}) on the \tilde{A} state for different damping thresholds $c_k = d_k^{-1}$ used in Equation 16. Columns for c_k values larger than 10 provide level shifts compared to reference network (30 hidden-layer neurons, $c_k = 10$). Labels and standard deviations σ_i taken from Table 2.

significant that obtained dynamics results would fail to provide relevant insight, as the resulting dynamics change too drastically to allow a proper comparison. On the other hand, the behavior of the network changes only marginally when increasing c_k from 10 to 15. This suggests that $c_k = 10$ is indeed large enough to cover the dynamically relevant region, while minimizing the risk of artifacts. Networks with thresholds of 50 or greater (essentially meaning that $\nu_k' \rightarrow \nu_k$) have been found to cause significant artifacts in the surface, rendering further comparisons unfeasible. One way in which these artifacts become noticeable is that dynamics results of formally degenerate level energies begin to show deviations of $\sim 1 \text{ cm}^{-1}$ due to numerical instability. As the damping threshold c_k becomes higher and higher, the computed level energies change increasingly, with deviations from the reference calculation values exceeding σ_i , until a "tipping point" (close to $c_k = 50$) is reached, rendering the obtained dynamics calculations unusable.

In summary, tuning the damping threshold c_k one observes a significant change in the dynamics results obtained, with one limiting case on either side. Each limiting case is accompanied by an abrupt change into pathological behavior. Our interpretation of the data presented in Table 3 is that very large values of the ν_k outside the dynamically relevant region cause the numerical artifacts and thus need to be damped as quickly as possible. This means that the proper values of c_k can be estimated approximately by the inspection of the reference data and should be chosen reasonably small. For this reason, all further results are obtained with a value of $c_k = 10 \forall k$.

4.4. Comparison with Experimental and Previous Theoretical Data. Up to this point it was investigated primarily how the ANN choices like hidden-layer size affects the fit performance as well as quantum dynamics result in comparison with a reference ANN model and how these observations fit into the conclusions drawn from previous efforts. The comparisons aim at estimating errors intrinsic to the model with respect to the differences in excitation energies produced by the ANNs of (in terms of fitting error) indistinguishable quality. These estimates

in return allow to judge different networks in terms of their internal consistency among one another, without the need for an external reference. In this section, the focus will be on the quality of the surfaces based on external consistency with experimental data, considering \tilde{X} and \tilde{A} state vibrational/vibronic transition energies.

4.4.1. 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 4. The experimental frequencies are compared to the

Table 4. Summary of the Comparison^a

assignment	E_{exp}	$E_{\text{theo}}^{\text{ANN,old}}$	$E_{\text{theo}}^{\text{ANN,new}}$	$\delta E_{\text{exp}}^{\text{ANN,old}}$	$\delta E_{\text{exp}}^{\text{ANN,new}}$
$4^1 e'$	365.5	366.8	365.8	-1.3	-0.3
$4^2 a_1'$	752.4	756.3	754.6	-3.9	-2.2
$4^2 e'$	771.8	774.6	772.5	-2.8	-0.7
$1^1 a_1'$	1051.2	1054.9	1055.6	-3.7	-4.4
$3^1 e'$	1055.3	1039.7	1035.5	15.6	19.8
$4^3 e'$	1173.6	1179.3	1176.6	-5.7	-3.
$4^3 a_1'$	1214	1215.1	1212.3	-1.1	1.7
$1^1 4^1 e'$	1413.6	1417.7	1417.6	-4.1	-4.
$3^1 4^1 a_2'$	1491	1344.9	1342.7	146.1	148.3
$3^1 4^1 e'$	1492.4	1490.0	1488.1	2.4	4.3
$3^1 4^1 a_1'$	1499.8	1498.2	1489.8	1.6	10.

^aSummary of the comparison between the experimental and computed transition energies on the \tilde{X}^2A_2' state of $^{14}\text{NO}_3$ (in cm^{-1}). The “assignment” column provides the corresponding D_{3h} labels as well as state assignments including dominant state contributions. $E_{\text{theo}}^{\text{ANN,old}}$ provides excitation energies using the previous (ANN-based) model,^{90,95} and $E_{\text{theo}}^{\text{ANN,new}}$ relates to the ones obtained for the present model. The remaining two columns provide respective differences from the experimental data as listed under E_{exp} , reported from ref 106.

computed transitions for both the novel and previous ANN model. As becomes apparent quickly from the fifth and sixth columns of Table 4, deviations from experiment remain largely similar, with all but one change in deviation being most likely just coincidental and due to the model-intrinsic fluctuations discussed in Section 4.2. The only state to which a noteworthy change can be attributed is the $3^1 4^1 a_1'$ state. However, this deviation by itself does not allow for any meaningful interpretation and still resides well within the limits of what can be expected of the model under present conditions. The new model also reconfirms the long-debated 3^1 fundamental, which can be assigned confidently to a state computed at 1036 cm^{-1} (previously 1040 cm^{-1}), agreeing well with other theoretical treatments^{102–104,116} and recent experimental assignments.^{106,117,118} Similarly, the new model also provides further evidence that the measured energy level of 1492.4 cm^{-1} does not correspond to the 3^1 fundamental but to the combination band $3^1 4^1 e'$. Finally, the large deviation found for the $3^1 4^1 a_2'$ state can only be explained by a misassignment of the experiment as a disagreement of almost 150 cm^{-1} , for a single state appears more than unlikely, especially with excitations even higher in energy not reaching one-tenth of that.

4.4.2. Vibronic Eigenstates in the \tilde{A}^2E'' Electronic State. The available data for the electronically excited $^2E''$ state are restricted to three experimentally measured transitions, as was the case with the previous model. While a greater deal of experimentally observed transitions is indeed available, they all involve umbrella excitations, which cannot be reproduced with the present model being restricted to planar geometries. A comparison of how both old and new ANN models reproduce

available experimental data is provided in Table 5. The two adiabatic PES sheets corresponding to the $^2E''$ state show a

Table 5. Summary of the Comparison^a

assignment	E_{exp}	$E_{\text{theo}}^{\text{ANN,old}}$	$E_{\text{theo}}^{\text{ANN,new}}$	$\delta E_{\text{exp}}^{\text{ANN,old}}$	$\delta E_{\text{exp}}^{\text{ANN,new}}$
$4_0^1 (t_b^1)$	539.5	539.3	539.5	0.2	0.0
$4_0^2 (r_s^1 t_b^1)$	1056.6	1024.8	1047.7	31.8	8.9
$3_0^1 (t_s^1)$	1270.5	1276.1	1271.4	-5.6	-0.9

^aSummary of the comparison between the 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 107, with assignments from Table 2 given in parentheses. $E_{\text{theo}}^{\text{ANN,old}}$ provides excitation energies using the previous (ANN-based) model,^{90,95} and $E_{\text{theo}}^{\text{ANN,new}}$ relates to the ones obtained for the present ANN-based model. The remaining two columns provide the respective differences from the experimental data.

particularly strong Jahn–Teller coupling resulting in a pronounced triple-well structure on the lower adiabatic sheet. As a result, the assignment of the D_{3h} labels based on normal mode excitations are not straightforward. The intricacies of how tunneling and the geometric phase effect affect the resulting spectra as well as the structure of the vibronic eigenstates have been analyzed in detail in refs 105 and 110.

Of the three experimentally known excitations, two have already been reproduced within the intrinsic limitations of the previous model. This continues to be the case for the present method, excellently reproducing experimental data for the $4_0^1 (t_b^1)$ and the $3_0^1 (t_s^1)$ level. In addition, the previously far less well reproduced $r_s^1 t_b^1$ state now is computed with unprecedented agreement with experimental data by the new CNPI-ANN model. While it is impossible to extrapolate too far from such a small data set, it may be concluded that it is a fair assumption that this improvement is indicative of the overall improved quality of the new model, in particular, with respect to the \tilde{A} state.

5. CONCLUSIONS AND OUTLOOK

The recently developed diabaticization method based on a combination of a simple vibronic coupling model and an ANN that tunes the coupling parameters has been improved and thoroughly tested in the present study. The use of symmetry transformation invariants as input coordinates for the ANN intrinsically removes symmetry contaminations. Apparently, the proper symmetry handling by the invariants reduces the stress on the ANN and leads to reduced RMS errors of the represented ab initio data. Furthermore, these improved results are obtained with significantly smaller ANNs. While with the original method an optimal ANN with 75 hidden-layer neurons was determined, the present CNPI-ANN method needs only 30 neurons to yield a PES model with noticeably improved RMSE. Thus, the new method not only is capable of reaching better accuracy but also is considerably more efficient. This can reduce the numerical effort in demanding quantum dynamics calculations considerably and thus can be of tremendous advantage.

The performance of the diabatic PES model obtained with the CNPI-ANN method has been tested by representing an extensive set of ab initio data for the few lowest electronic states $^2A_2'$, $^2E''$, and $^2E'$ of planar NO_3 . The same data set has been used as in our previous studies, and thus the results can be compared directly. In addition to the residual errors in the representation of the ab initio data, vibrational- and vibronic-

level energies have been computed. This allows to assess the sensitivity of the nuclear dynamics results on the various parametrizations of the CNPI-ANN PES model. It turns out that the new CNPI-ANN model behaves very similarly to the previous plain ANN model. In particular, the onset of overfitting can be assessed from the development of the validation error compared to the reference error when increasing the hidden-layer size. For each ANN, many different parametrizations are obtained, which are very similar in terms of RMSE. The corresponding quantum dynamics calculations yield a certain scattering of values for each vibrational/vibronic level that is indicative of the sensitivity of the results on the minute details of the PES model. The observed standard deviations are only on the order of a couple of inverse centimeters for the ${}^2A_2'$ ground-state vibrational levels. The standard deviations computed for the vibronic levels of the excited ${}^2E''$ state, having a much more complicated PES topography, are approximately an order of magnitude larger. Comparison of these dynamics results with the available experimental values shows overall excellent agreement, and most level energies are reproduced more or less within the computed standard deviations. The agreement with experiment is slightly improved over the plain ANN model, which already yielded excellent results. The most noticeable improvements are observed for the ${}^2E''$ state both in fitting RMSE and quantum dynamics results.

The present model limited to 5D planar NO_3 will be extended to full 6D in the near future to be able to investigate the nuclear quantum dynamics and spectroscopy of this intriguing radical in full detail. Of course, the CNPI-ANN diabatization model can be applied also to other vibronic coupling problems. The present results seem quite encouraging that accurate and reliable diabatic PES models can be obtained by the present method in an efficient way.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpca.0c05991>.

Apart from the provided materials, two further tables are available. These provide a full comparison of computed transition energies for the 10 best (in terms of fitting error) ANNs for the reference network on both \tilde{X} and A state. Figures 4 and 5 provide a visual representation of these tables for the first 15 states (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

David M. G. Williams – *Theoretische Chemie, Universität Bielefeld, D-33501 Bielefeld, Germany*; Email: d.williams@uni-bielefeld.de

Wolfgang Eisfeld – *Theoretische Chemie, Universität Bielefeld, D-33501 Bielefeld, Germany*;  orcid.org/0000-0002-3691-730X; Email: wolfgang.eisfeld@uni-bielefeld.de

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jpca.0c05991>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Part of this work was generously supported by the Deutsche Forschungsgemeinschaft.

■ REFERENCES

- (1) Collins, M. A.; Parsons, D. F. Implications of Rotation-Inversion-Permutation Invariance for Analytic Molecular-Potential Energy Surfaces. *J. Chem. Phys.* **1993**, *99*, 6756–6772.
- (2) Ischtwan, J.; Collins, M. A. Molecular-Potential Energy Surfaces by Interpolation. *J. Chem. Phys.* **1994**, *100*, 8080–8088.
- (3) Ho, T. S.; Rabitz, H. A General Method for Constructing Multidimensional Molecular Potential Energy Surfaces from Ab Initio Calculations. *J. Chem. Phys.* **1996**, *104*, 2584–2597.
- (4) Maisuradze, G. G.; Thompson, D. L.; Wagner, A. F.; Minkoff, M. Interpolating Moving Least-Squares Methods for Fitting Potential Energy Surfaces: Detailed Analysis of One-Dimensional Applications. *J. Chem. Phys.* **2003**, *119*, 10002–10014.
- (5) Maisuradze, G. G.; Kawano, A.; Thompson, D. L.; Wagner, A. F.; Minkoff, M. Interpolating Moving Least-Squares Methods for Fitting Potential Energy Surfaces: Analysis of an Application to a Six-Dimensional System. *J. Chem. Phys.* **2004**, *121*, 10329–10338.
- (6) Braams, B. J.; Bowman, J. M. Permutationally Invariant Potential Energy Surfaces in High Dimensionality. *Int. Rev. Phys. Chem.* **2009**, *28*, 577–606.
- (7) Blank, T. B.; Brown, S. D.; Calhoun, A. W.; Doren, D. J. Neural-network models of potential-energy surfaces. *J. Chem. Phys.* **1995**, *103*, 4129–4137.
- (8) Brown, D. F. R.; Gibbs, M. N.; Clary, D. C. Combining ab initio computations, neural networks, and diffusion Monte Carlo: An efficient method to treat weakly bound molecules. *J. Chem. Phys.* **1996**, *105*, 7597–7604.
- (9) Tai No, K.; Ha Chang, B.; Yeon Kim, S.; Shik Jhon, M.; Scheraga, H. A. Description of the potential energy surface of the water dimer with an artificial neural network. *Chem. Phys. Lett.* **1997**, *271*, 152–156.
- (10) Prudente, F. V.; Acioli, P. H.; Neto, J. J. S. The fitting of potential energy surfaces using neural networks: Application to the study of vibrational levels of H_3^+ . *J. Chem. Phys.* **1998**, *109*, 8801–8808.
- (11) Lorenz, S.; Gross, A.; Scheffler, M. Representing high-dimensional potential-energy surfaces for reactions at surfaces by neural networks. *Chem. Phys. Lett.* **2004**, *395*, 210–215.
- (12) Raff, L. M.; Malshe, M.; Hagan, M.; Doughan, D. I.; Rockley, M. G.; Komanduri, R. Ab initio potential-energy surfaces for complex, multichannel systems using modified novelty sampling and feedforward neural networks. *J. Chem. Phys.* **2005**, *122*, No. 084104.
- (13) Lorenz, S.; Scheffler, M.; Gross, A. Descriptions of surface chemical reactions using a neural network representation of the potential-energy surface. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2006**, *73*, 115431.
- (14) Manzhos, S.; Wang, X. G.; Dawes, R.; Carrington, T. A nested molecule-independent neural network approach for high-quality potential fits. *J. Phys. Chem. A* **2006**, *110*, 5295–5304.
- (15) Manzhos, S.; Carrington, T., Jr. A random-sampling high dimensional model representation neural network for building potential energy surfaces. *J. Chem. Phys.* **2006**, *125*, No. 084109.
- (16) Manzhos, S.; Carrington, T., Jr. Using neural networks to represent potential surfaces as sums of products. *J. Chem. Phys.* **2006**, *125*, 194105.
- (17) Behler, J.; Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.
- (18) Manzhos, S.; Carrington, T., Jr. Using redundant coordinates to represent potential energy surfaces with lower-dimensional functions. *J. Chem. Phys.* **2007**, *127*, No. 014103.
- (19) Malshe, M.; Narulkar, R.; Raff, L. M.; Hagan, M.; Bukkapatnam, S.; Komanduri, R. Parametrization of analytic interatomic potential functions using neural networks. *J. Chem. Phys.* **2008**, *129*, No. 044111.
- (20) Manzhos, S.; Carrington, T., Jr. Using neural networks, optimized coordinates, and high-dimensional model representations to obtain a vinyl bromide potential surface. *J. Chem. Phys.* **2008**, *129*, 224104.
- (21) Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *J. Chem. Phys.* **2011**, *134*, No. 074106.

- (22) Nguyen, H. T. T.; Le, H. M. Modified Feed-Forward Neural Network Structures and Combined-Function-Derivative Approximations Incorporating Exchange Symmetry for Potential Energy Surface Fitting. *J. Phys. Chem. A* **2012**, *116*, 4629–4638.
- (23) Jiang, B.; Guo, H. Permutation invariant polynomial neural network approach to fitting potential energy surfaces. *J. Chem. Phys.* **2013**, *139*, No. 054112.
- (24) Koch, W.; Zhang, D. H. Communication: Separable potential energy surfaces from multiplicative artificial neural networks. *J. Chem. Phys.* **2014**, *141*, No. 021101.
- (25) Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control Signals Systems* **1989**, *2*, 303–314.
- (26) Meyer, H. D.; Manthe, U.; Cederbaum, L. S. The multi-configurational time-dependent Hartree approach. *Chem. Phys. Lett.* **1990**, *165*, 73.
- (27) Manthe, U.; Meyer, H. D.; Cederbaum, L. S. Wave-packet dynamics within the multiconfigurational Hartree framework: General aspects and application to NOCl. *J. Chem. Phys.* **1992**, *97*, 3199–3213.
- (28) Manthe, U. A time-dependent discrete variable representation for (multiconfiguration) Hartree methods. *J. Chem. Phys.* **1996**, *105*, 6989–6994.
- (29) Brown, A.; Pradhan, E. Fitting potential energy surfaces to sum-of-products form with neural networks using exponential neurons. *J. Theor. Comput. Chem.* **2017**, *16*, 1730001.
- (30) Majumder, M.; Hegger, S. E.; Dawes, R.; Manzhos, S.; Wang, X.-G.; Carrington, T., Jr; Li, J.; Guo, H. Explicitly correlated MRCl-F12 potential energy surfaces for methane fit with several permutation invariant schemes and full-dimensional vibrational calculations. *Mol. Phys.* **2015**, *113*, 1823–1833.
- (31) Jiang, B.; Li, J.; Guo, H. Potential energy surfaces from high fidelity fitting of ab initio points: the permutation invariant polynomial neural network approach. *Int. Rev. Phys. Chem.* **2016**, *35*, 479–506.
- (32) Pradhan, E.; Brown, A. A ground state potential energy surface for HONO based on a neural network with exponential fitting functions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 22272–22281.
- (33) Domcke, W., Yarkony, D. R., Köppel, H., Eds. *Conical Intersections: Electronic Structure, Dynamics and Spectroscopy*; World Scientific: Singapore, 2004.
- (34) Longuet-Higgins, H. C. Some Recent Developments in the Theory of Molecular Energy Levels. In *Advances in Spectroscopy*; Interscience Publishers, Ltd.: New York, 1961; Vol. 2, p 429.
- (35) Lichten, W. Resonant charge exchange in atomic collisions. *Phys. Rev.* **1963**, *131*, 229.
- (36) Lichten, W. Molecular wave functions and inelastic atomic collisions. *Phys. Rev.* **1967**, *164*, 131.
- (37) Smith, F. T. Diabatic and adiabatic representations for atomic collision problems. *Phys. Rev.* **1969**, *179*, 111.
- (38) Baer, M. Adiabatic and diabatic representations for atom-diatom collisions - Treatment of 3-dimensional case. *Chem. Phys.* **1976**, *15*, 49–57.
- (39) Werner, H.; Meyer, W. MCSCF study of the avoided curve crossing of the two lowest $^1\Sigma^+$ states of LiF. *J. Chem. Phys.* **1981**, *74*, 5802–5807.
- (40) Mead, C. A.; Truhlar, D. G. Conditions for the Definition of a Strictly Diabatic Electronic Basis for Molecular Systems. *J. Chem. Phys.* **1982**, *77*, 6090–6098.
- (41) Mead, C. A. *J. Chem. Phys.* **1983**, *78*, 807.
- (42) Werner, H.-J.; Follmeg, B.; Alexander, M. H. Adiabatic and diabatic potential energy surfaces for collisions of CN with He. *J. Chem. Phys.* **1988**, *89*, 3139.
- (43) Pacher, T.; Cederbaum, L. S.; Köppel, H. Approximately Diabatic States from Block Diagonalization of the Electronic Hamiltonian. *J. Chem. Phys.* **1988**, *89*, 7367–7381.
- (44) Pacher, T.; Mead, C. A.; Cederbaum, L. S.; Köppel, H. Gauge-Theory and Quasidiabatic States in Molecular Physics. *J. Chem. Phys.* **1989**, *91*, 7057–7062.
- (45) Pacher, T.; Köppel, H.; Cederbaum, L. S. Quasidiabatic States from Ab Initio Calculations by Block Diagonalization of the Electronic Hamiltonian - Use of Frozen Orbitals. *J. Chem. Phys.* **1991**, *95*, 6668–6680.
- (46) Pacher, T.; Cederbaum, L. S.; Köppel, H. *Adv. Chem. Phys.* **1993**, *84*, 293.
- (47) Xu, X.; Yang, K. R.; Truhlar, D. G. Diabatic Molecular Orbitals, Potential Energies, and Potential Energy Surface Couplings by the 4-fold Way for Photodissociation of Phenol. *J. Chem. Theory Comput.* **2013**, *9*, 3612–3625.
- (48) Shu, Y.; Kryven, J.; Sampaio de Oliveira-Filho, A. G.; Zhang, L.; Song, G.-L.; Li, S. L.; Meana-Pañeda, R.; Fu, B.; Bowman, J. M.; Truhlar, D. G. Direct diabaticization and analytic representation of coupled potential energy surfaces and couplings for the reactive quenching of the excited $^2\Sigma^+$ state of OH by molecular hydrogen. *J. Chem. Phys.* **2019**, *151*, 104311.
- (49) Parker, K. A.; Truhlar, D. G. Semiglobal diabatic potential energy matrix for the N–H photodissociation of methylamine. *J. Chem. Phys.* **2020**, *152*, 244309.
- (50) Köppel, H.; Domcke, W.; Cederbaum, L. S. Multimode Molecular Dynamics Beyond the Born-Oppenheimer Approximation. *Adv. Chem. Phys.* **1984**, *57*, 59–246.
- (51) Cimiraaglia, R.; Malrieu, J. P.; Persico, M.; Spiegelmann, F. Quasi-Diabatic States and Dynamical Couplings from Ab Initio CI Calculations - A New Proposal. *J. Phys. B: At. Mol. Phys.* **1985**, *18*, 3073–3084.
- (52) Gadea, F. X.; Pelissier, M. *J. Chem. Phys.* **1990**, *93*, 545.
- (53) Domcke, W.; Woywod, C. Diabatic CASSCF Orbitals and Wave-Functions. *Chem. Phys. Lett.* **1993**, *216*, 362–368.
- (54) Yarkony, D. *R. Rev. Mod. Phys.* **1996**, *68*, 985.
- (55) Atchity, G. J.; Ruedenberg, K. Determination of diabatic states through enforcement of configurational uniformity. *Theor. Chem. Acc.* **1997**, *97*, 47–58.
- (56) Thiel, A.; Köppel, H. Proposal and numerical test of a simple diabaticization scheme. *J. Chem. Phys.* **1999**, *110*, 9371–9383.
- (57) Cattaneo, P.; Persico, M. Diabatic and adiabatic potential-energy surfaces for azomethane photochemistry. *Theor. Chem. Acc.* **2000**, *103*, 390–398.
- (58) Köppel, H.; Gronki, J.; Mahapatra, S. Construction scheme for regularized diabatic states. *J. Chem. Phys.* **2001**, *115*, 2377–2388.
- (59) Nakamura, H.; Truhlar, D. G. The direct calculation of diabatic states based on configurational uniformity. *J. Chem. Phys.* **2001**, *115*, 10353–10372.
- (60) Nakamura, H.; Truhlar, D. G. Direct diabaticization of electronic states by the fourfold way. II. Dynamical correlation and rearrangement processes. *J. Chem. Phys.* **2002**, *117*, 5576–5593.
- (61) Abrol, R.; Kuppermann, A. An optimal adiabatic-to-diabatic transformation of the $1^2A'$ and $2^2A'$ states of H_3 . *J. Chem. Phys.* **2002**, *116*, 1035–1062.
- (62) Nakamura, H.; Truhlar, D. G. Extension of the fourfold way for calculation of global diabatic potential energy surfaces of complex, multiarrangement, non-Born-Oppenheimer systems: Application to HNC(O)(S-0,S-1). *J. Chem. Phys.* **2003**, *118*, 6816–6829.
- (63) Worth, G. A.; Cederbaum, L. S. Beyond Born-Oppenheimer: Conical intersections and their impact on molecular dynamics. *Annu. Rev. Phys. Chem.* **2004**, *55*, 127.
- (64) Viel, A.; Eisfeld, W. Effect of Higher-Order Jahn-Teller Coupling on the Nuclear Dynamics. *J. Chem. Phys.* **2004**, *120*, 4603–13.
- (65) Eisfeld, W.; Viel, A. Higher Order $(A + E) \otimes e$ Pseudo Jahn-Teller Coupling. *J. Chem. Phys.* **2005**, *122*, 204317.
- (66) Viel, A.; Eisfeld, W.; Neumann, S.; Domcke, W.; Manthe, U. Photoionization-Induced Dynamics of the Ammonia Cation: Ab Initio Potential-Energy Surfaces and Time-Dependent Wave-Packet Calculations. *J. Chem. Phys.* **2006**, *124*, 214306.
- (67) Viegas, L. P.; Alijah, A.; Varandas, A. J. C. Accurate ab initio based multisheeted double many-body expansion potential energy surface for the three lowest electronic singlet states of H_3^+ . *J. Chem. Phys.* **2007**, *126*, No. 074309.
- (68) Viel, A.; Eisfeld, W.; Evenhuis, C. R.; Manthe, U. Photoionization-Induced Dynamics of the Ammonia Cation Studied by Wave

Packet Calculations Using Curvilinear coordinates. *Chem. Phys.* **2008**, *347*, 331–339.

(69) Faraji, S.; Köppel, H.; Eisfeld, W.; Mahapatra, S. Towards a Higher Order Description of Jahn-Teller Coupling Effects in Molecular Spectroscopy: The \tilde{A}^2E'' State of NO_3 . *Chem. Phys.* **2008**, *347*, 110–119.

(70) Papas, B. N.; Schuurman, M. S.; Yarkony, D. R. Determining quasideiabatic coupled electronic state Hamiltonians using derivative couplings: A normal equations based method. *J. Chem. Phys.* **2008**, *129*, 124104.

(71) Subotnik, J. E.; Yeganeh, S.; Cave, R. J.; Ratner, M. A. Constructing diabatic states from adiabatic states: Extending generalized Mulliken-Hush to multiple charge centers with Boys localization. *J. Chem. Phys.* **2008**, *129*, 244101.

(72) Zhu, X.; Yarkony, D. R. On the determination of optimized, fully quadratic, coupled state quasideiabatic Hamiltonians for determining bound state vibronic spectra. *J. Chem. Phys.* **2009**, *130*, 234108.

(73) Opalka, D.; Domcke, W. High-order expansion of $T(2)\times t(2)$ Jahn-Teller potential-energy surfaces in tetrahedral molecules. *J. Chem. Phys.* **2010**, *132*, 154108.

(74) Eisfeld, W.; Vieuxmaire, O.; Viel, A. Full-dimensional diabatic potential energy surfaces including dissociation: The 2E state of NO_3 . *J. Chem. Phys.* **2014**, *140*, 224109.

(75) Venghaus, F.; Eisfeld, W. Block-diagonalization as tool for the robust diabaticization of high-dimensional potential energy surfaces. *J. Chem. Phys.* **2016**, *144*, 114110.

(76) Wittenbrink, N.; Venghaus, F.; Williams, D.; Eisfeld, W. A new approach for the development of diabatic potential energy surfaces: Hybrid block-diagonalization and diabaticization by ansatz. *J. Chem. Phys.* **2016**, *145*, 184108.

(77) Schuurman, M. S.; Yarkony, D. R. On the vibronic coupling approximation: A generally applicable approach for determining fully quadratic quasideiabatic coupled electronic state Hamiltonians. *J. Chem. Phys.* **2007**, *127*, No. 094104.

(78) Zhu, X.; Yarkony, D. R. Toward eliminating the electronic structure bottleneck in nonadiabatic dynamics on the fly: An algorithm to fit nonlocal, quasideiabatic, coupled electronic state Hamiltonians based on ab initio electronic structure data. *J. Chem. Phys.* **2010**, *132*, 104101.

(79) Evenhuis, C. R.; Collins, M. A. Interpolation of Diabatic Potential Energy Surfaces. *J. Chem. Phys.* **2004**, *121*, 2515–2527.

(80) Evenhuis, C. R.; Lin, X.; Zhang, D. H.; Yarkony, D.; Collins, M. A. Interpolation of Diabatic Potential-Energy Surfaces: Quantum Dynamics on Ab Initio Surfaces. *J. Chem. Phys.* **2005**, *123*, 134110.

(81) Godsi, O.; Evenhuis, C. R.; Collins, M. A. Interpolation of Multidimensional Diabatic Potential Energy Matrices. *J. Chem. Phys.* **2006**, *125*, 104105.

(82) Nangia, S.; Truhlar, D. G. Direct calculation of coupled diabatic potential-energy surfaces for ammonia and mapping of a four-dimensional conical intersection seam. *J. Chem. Phys.* **2006**, *124*, 124309.

(83) Li, Z. H.; Valero, R.; Truhlar, D. G. Improved direct diabaticization and coupled potential energy surfaces for the photodissociation of ammonia. *Theor. Chem. Acc.* **2007**, *118*, 9–24.

(84) Zhu, X.; Ma, J. Y.; Yarkony, D. R.; Guo, H. Computational Determination of the \tilde{A} State Absorption Spectrum of NH_3 and of ND_3 Using a New Quasi-Diabatic Representation of the X and A States and Full Six-Dimensional Quantum Dynamics. *J. Chem. Phys.* **2012**, *136*, 234301.

(85) Xie, C.; Ma, J.; Zhu, X.; Zhang, D. H.; Yarkony, D. R.; Xie, D.; Guo, H. Full-Dimensional Quantum State-to-State Nonadiabatic Dynamics for Photodissociation of Ammonia in its A-Band. *J. Phys. Chem. Lett.* **2014**, *5*, 1055–1060.

(86) Zhu, X.; Yarkony, D. R. Fitting coupled potential energy surfaces for large systems: Method and construction of a 3-state representation for phenol photodissociation in the full 33 internal degrees of freedom using multireference configuration interaction determined data. *J. Chem. Phys.* **2014**, *140*, No. 024112.

(87) Guan, Y.; Fu, B.; Zhang, D. H. Construction of diabatic energy surfaces for LiFH with artificial neural networks. *J. Chem. Phys.* **2017**, *147*, 224307.

(88) Lenzen, T.; Manthe, U. Neural network based coupled diabatic potential energy surfaces for reactive scattering. *J. Chem. Phys.* **2017**, *147*, No. 084105.

(89) Xie, C.; Zhu, X.; Yarkony, D. R.; Guo, H. Permutation invariant polynomial neural network approach to fitting potential energy surfaces. IV. Coupled diabatic potential energy matrices. *J. Chem. Phys.* **2018**, *149*, 144107.

(90) Williams, D. M. G.; Eisfeld, W. Neural network diabaticization: A new ansatz for accurate high-dimensional coupled potential energy surfaces. *J. Chem. Phys.* **2018**, *149*, 204106.

(91) Lenzen, T.; Eisfeld, W.; Manthe, U. Vibronically and spin-orbit coupled diabatic potentials for $X(^2P) + \text{CH}_4 \rightarrow \text{HX} + \text{CH}_3$ reactions: Neural network potentials for $X = \text{Cl}$. *J. Chem. Phys.* **2019**, *150*, 244115.

(92) Guan, Y.; Guo, H.; Yarkony, D. R. Neural network based quasideiabatic Hamiltonians with symmetry adaptation and a correct description of conical intersections. *J. Chem. Phys.* **2019**, *150*, 214101.

(93) Guan, Y.; Zhang, D. H.; Guo, H.; Yarkony, D. R. Representation of coupled adiabatic potential energy surfaces using neural network based quasi-diabatic Hamiltonians: 1,2 ($^2A'$) states of LiFH. *Phys. Chem. Chem. Phys.* **2019**, *21*, 14205–14213.

(94) Yin, Z.; Guan, Y.; Fu, B.; Zhang, D. H. Two-state diabatic potential energy surfaces of ClH_2 based on nonadiabatic couplings with neural networks. *Phys. Chem. Chem. Phys.* **2019**, *21*, 20372–20383.

(95) Williams, D. M. G.; Viel, A.; Eisfeld, W. Diabatic neural network potentials for accurate vibronic quantum dynamics-The test case of planar NO_3 . *J. Chem. Phys.* **2019**, *151*, 164118.

(96) Guan, Y.; Yarkony, D. R. Accurate Neural Network Representation of the Ab Initio Determined Spin-Orbit Interaction in the Diabatic Representation Including the Effects of Conical Intersections. *J. Phys. Chem. Lett.* **2020**, *11*, 1848–1858.

(97) Guan, Y.; Guo, H.; Yarkony, D. R. Extending the Representation of Multistate Coupled Potential Energy Surfaces To Include Properties Operators Using Neural Networks: Application to the 1,2(1A) States of Ammonia. *J. Chem. Theory Comput.* **2020**, *16*, 302–313.

(98) Viel, A.; Eisfeld, W. NO_3 full-dimensional potential energy surfaces and ground state vibrational levels revisited. *Chem. Phys.* **2018**, *509*, 81–90.

(99) Mayer, M.; Cederbaum, L. S.; Köppel, H. Ground state dynamics of NO_3 : Multimode vibronic borrowing including thermal effects. *J. Chem. Phys.* **1994**, *100*, 899.

(100) Okumura, M.; Stanton, J.; Deev, A.; Sommar, J. New insights into the Jahn-Teller effect in NO_3 via the dark \tilde{A}^2E'' state. *Phys. Rev. Lett.* **2006**, *73*, C64–C70.

(101) Mahapatra, S.; Eisfeld, W.; Köppel, H. Effects of multimode Jahn-Teller coupling on the photodetachment spectrum of nitrate anion (NO_3^-). *Chem. Phys. Lett.* **2007**, *441*, 7–15.

(102) Stanton, J. F. On the vibronic level structure in the NO_3 radical: II. Adiabatic calculation of the infrared spectrum. *Mol. Phys.* **2009**, *107*, 1059–1075.

(103) Simmons, C. S.; Ichino, T.; Stanton, J. F. The $v(3)$ Fundamental in NO_3 Has Been Seen Near 1060 cm^{-1} , Albeit Some Time Ago. *J. Phys. Chem. Lett.* **2012**, *3*, 1946–1950.

(104) Homayoon, Z.; Bowman, J. M. Communication: MULTI-MODE calculations of low-lying vibrational states of NO_3 using an adiabatic potential energy surface. *J. Chem. Phys.* **2014**, *141*, 161104.

(105) Eisfeld, W.; Viel, A. Vibronic eigenstates and the geometric phase effect in the $^2E''$ state of NO_3 . *J. Chem. Phys.* **2017**, *146*, No. 034303.

(106) Kawaguchi, K.; Narahara, T.; Fujimori, R.; Tang, J.; Ishiwata, T. Infrared spectroscopy of $2v(4)$ and $v(3)+2v(4)$ bands of the NO_3 radical. *J. Mol. Spectrosc.* **2017**, *334*, 10–21.

(107) Codd, T.; Chen, M.-W.; Roudjane, M.; Stanton, J. F.; Miller, T. A. Jet cooled cavity ringdown spectroscopy of the $\tilde{A}^2E'' \leftarrow X^2A_2'$ transition of the NO_3 radical. *J. Chem. Phys.* **2015**, *142*, 184305.

(108) Lenzen, T.; Manthe, U. Vibronically and spin-orbit coupled diabatic potentials for $X(P) + \text{CH}_4 \rightarrow \text{HX} + \text{CH}_3$ reactions: General

theory and application for $X(P) = F(^2P)$. *J. Chem. Phys.* **2019**, *150*, No. 064102.

(109) Zeng, T.; Seidu, I. Revisiting the $(E + A) \otimes (e + a)$ problems of polyatomic systems with trigonal symmetry: general expansions of their vibronic Hamiltonians. *Phys. Chem. Chem. Phys.* **2017**, *19*, 11098–11110.

(110) Weike, T.; Williams, D. M. G.; Viel, A.; Eisfeld, W. Quantum dynamics and geometric phase in $E \otimes e$ Jahn-Teller systems with general C_{nv} symmetry. *J. Chem. Phys.* **2019**, *151*, No. 074302.

(111) Eisfeld, W.; Morokuma, K. *Ab initio* investigation of the vertical and adiabatic excitation spectrum of NO_3 . *J. Chem. Phys.* **2001**, *114*, 9430–9440.

(112) Eisfeld, W.; Morokuma, K. A detailed study on the symmetry breaking and its effect on the potential surface of NO_3 . *J. Chem. Phys.* **2000**, *113*, 5587.

(113) Eisfeld, W. Highly accurate determination of the electron affinity of SF_6 and analysis of structure and photodetachment spectrum of SF_6^- . *J. Chem. Phys.* **2011**, *134*, No. 054303.

(114) Hagan, M.; Demuth, H.; Beale, M.; De Jesús, O. *Neural Network Design*, 2nd ed.; Martin Hagan, 2014.

(115) Light, J. C.; Carrington, T. Discrete-variable representations and their utilization. *Adv. Chem. Phys.* **2007**, *114*, 263–310.

(116) Stanton, J. F. On the vibronic level structure in the NO_3 radical. I. The ground electronic state. *J. Chem. Phys.* **2007**, *126*, 134309.

(117) Jacox, M. E.; Thompson, W. E. The infrared spectroscopy and photochemistry of NO_3 trapped in solid neon. *J. Chem. Phys.* **2008**, *129*, 204306.

(118) Kawaguchi, K.; Fujimori, R.; Tang, J.; Ishiwata, T. On the vibrational assignment in the ground electronic state of NO_3 . *J. Mol. Spectrosc.* **2015**, *314*, 73–78.

Accurate quantum dynamics simulation of the photodetachment spectrum of the nitrate anion (NO_3^-) based on an artificial neural network diabatic potential model

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

¹⁾ *Univ Rennes, CNRS, IPR (Institut de Physique de Rennes) - UMR 6251, F-35000 Rennes, France*

²⁾ *Theoretische Chemie, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany*

The photodetachment spectrum of the nitrate anion (NO_3^-) is simulated from first principles using wave packet quantum dynamics propagation and a newly developed accurate full-dimensional fully coupled five state diabatic potential model. This model utilizes the recently proposed complete nuclear permutation inversion invariant artificial neural network (CNPI-ANN) diabaticization technique [*J. Phys. Chem. A*, 2020, **124**, 7608]. The quantum dynamics simulations are designed such that temperature effects and the impact of near threshold detachment are taken into account. Thus, the two available experiments at high temperature and at cryogenic temperature using the slow electron velocity-map imaging technique (cryo-SEVI) can be reproduced in very good agreement. These results clearly show the relevance of hot bands as well as vibronic coupling between the $\tilde{X}^2A'_2$ ground state and the \tilde{B}^2E' excited state of the neutral radical. This together with the recent experiment at low temperature gives further support for the proper assignment of the ν_3 fundamental, which has been debated for many years. An assignment of a not yet discussed hot band line is also proposed.

I. INTRODUCTION

One central goal of chemical physics is the detailed and fundamental understanding of the nuclear dynamics of molecular systems. Impressive progress has been achieved in this quest by the combined efforts of experiment and theory over the past decades. Yet, there are still enigmatic systems which, despite being studied intensively, are not understood very well. The nitrate radical (NO_3) is one such case which on the one hand is of high significance due to its importance in atmospheric chemistry¹ and on the other hand shows complicated nuclear dynamics and correspondingly complex spectra. This sparked the interest of many research groups over the years both experimental²⁻⁴⁰ as well as theoretical.⁴¹⁻⁷⁵

One of the reasons for the complicated behaviour of NO_3 seems to be the presence of strong nonadiabatic coupling effects. The two first electronically excited states are of $^2E''$ and $^2E'$ symmetry, respectively, and thus are both subject to an $E \otimes e$ Jahn-Teller (JT) effect. Furthermore, the conical intersection at D_{3h} symmetric nuclear configurations induces a geometric phase due to the singularity of the nonadiabatic coupling. Both cause very complicated spectra corresponding to the excited state manifolds. But even the non-degenerate ground state of $^2A'_2$ symmetry shows spectroscopic features which have been debated for decades. This may be attributed at least in part to significant vibronic couplings to the excited state manifolds. A further complication is that NO_3

is difficult to study both experimentally and theoretically. The bottleneck for detailed theoretical studies of the nuclear dynamics and spectroscopy is the need of a very accurate potential energy surface (PES) model based on advanced electronic structure calculations. Various models have been developed in the past^{40,64-68,70-75} but it has been a stony track to achieve the required quality to reach decisive interpretations in comparison with experimental observations. Our recent developments for producing highly accurate diabatic PES models using artificial neural networks (ANNs) should be a significant step forward to reach that goal.⁷⁶⁻⁷⁸

One experimental breakthrough has been the first photodetachment spectrum of the nitrate anion (NO_3^-) published by the Neumark group in 1991.³⁴ This has been the first direct experimental proof of the existence of the \tilde{A}^2E'' state of the radical, which previously only had been predicted by theory.⁷⁹ This experiment also provided new data for the $\tilde{X}^2A'_2$ state of the radical. However, both partial spectra showed idiosyncrasies which sparked the discussion of their interpretation. A very recent experiment from the same group tried to resolve at least some questions regarding the \tilde{X} state.⁸⁰ By cryogenic cooling some hot bands of the original spectrum could be removed which allows to focus on the remaining unexpected features more clearly. Unfortunately, a corresponding experiment for the \tilde{A} state is not available yet. Thus, one aim of the present study is to simulate the temperature effects on the spectra theoretically and to compare the results with the new experimental data available.

Several attempts to simulate this experiment have been published over the years.^{64,71-73,80} While the \tilde{X} state could be simulated reasonably well even in the first attempt,⁶⁴ the \tilde{A} state took a lot more effort and required

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

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

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

a lot of development to obtain a sufficiently accurate PES model.^{73,81–84} Further development enables us now to use a PES model of unprecedented accuracy to simulate the corresponding spectra.^{76–78,85,86}

II. FULLY COUPLED DIABATIC 6D POTENTIAL ENERGY MODEL

The basic idea behind the model potential used in the present work was detailed in Ref. 78. Its main features and the extension of our approach to the full-dimensional (6D) diabatic PES model for NO₃ are presented here. The underlying *ansatz* for the diabatic matrix is expressed as a sum of diagonal and coupling matrices^{76,82,87} reading

$$\begin{aligned} \mathbf{W}^d(\mathbf{Q}) = & \mathbf{W}_{\text{diag}}^d(\mathbf{Q}) + \begin{pmatrix} \lambda_1 & 0^T & 0^T \\ 0 & \lambda_2 \mathbf{1} & 0 \\ 0 & 0 & \lambda_3 \mathbf{1} \end{pmatrix} \\ & + \theta \cdot \begin{pmatrix} 0 & 0^T & 0^T \\ 0 & 0 & \lambda_{10} \mathbf{1} \\ 0 & \lambda_{10} \mathbf{1} & 0 \end{pmatrix} \\ & + \begin{pmatrix} 0 & 0^T & \lambda_8 \boldsymbol{\rho}_s^T + \lambda_9 \boldsymbol{\rho}_b^T \\ 0 & \lambda_4 \boldsymbol{\varepsilon}_s + \lambda_5 \boldsymbol{\varepsilon}_b & 0 \\ \lambda_8 \boldsymbol{\rho}_s + \lambda_9 \boldsymbol{\rho}_b & 0 & \lambda_6 \boldsymbol{\varepsilon}_s + \lambda_7 \boldsymbol{\varepsilon}_b \end{pmatrix} \\ & + \theta \cdot \begin{pmatrix} 0 & \lambda_{11} \boldsymbol{\rho}_s^T + \lambda_{12} \boldsymbol{\rho}_b^T & 0^T \\ \lambda_{11} \boldsymbol{\rho}_s + \lambda_{12} \boldsymbol{\rho}_b & 0 & \lambda_{13} \boldsymbol{\varepsilon}_s + \lambda_{14} \boldsymbol{\varepsilon}_b \\ 0 & \lambda_{13} \boldsymbol{\varepsilon}_s + \lambda_{14} \boldsymbol{\varepsilon}_b & 0 \end{pmatrix}. \end{aligned} \quad (1)$$

The 2×2 matrices $\boldsymbol{\varepsilon}_{s,b}$ in Eq. (1) are the first-order Jahn-Teller coupling blocks and the vectors $\boldsymbol{\rho}_{s,b}$ are the pseudo-Jahn-Teller coupling blocks given by

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

$x_{s,b}$, $y_{s,b}$ and θ are symmetry-adapted coordinates corresponding to the degenerate asymmetric stretching and bending modes ($x_{s,b}, y_{s,b}$) and the umbrella motion (θ). These coordinates together with the symmetric stretch coordinate a are gathered in the 6D vector of nuclear coordinates \mathbf{Q} . They are defined identically to those used in previous studies^{73,83} and provided in the appendix for the sake of convenience. 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^{73,75,76,83} and reproduced in the appendix. $\mathbf{W}_{\text{diag}}^d(\mathbf{Q})$ is comprised of purely polynomial terms. The remaining four 5×5 matrices of Eq. (1) correspond to the parametrized constant, *linear* Jahn-Teller and *linear* pseudo-Jahn-Teller coupling matrices. Due to symmetry, all inter-state couplings between states of (A', E') and E'' symmetry, respectively, are proportional to the umbrella coordinate θ and hence vanish for planar geometries.

The key for the accuracy of the present model is that the vibronic coupling coefficients, $\lambda_j(\mathbf{Q})$ [$j = 1, \dots, 14$], are not constants like in standard vibronic coupling models. Instead, they depend on all six nuclear coordinates \mathbf{Q} through the use of an ANN with the coordinates as input layer. More specifically, the modified λ_i depend on invariant functions, $\tilde{\mathbf{Q}}$, constructed from \mathbf{Q} . Apart from the inclusion of θ^2 to account for the additional degree of freedom, the invariants ($\tilde{\mathbf{Q}}$) are identical to those used for the 5D ANN-based model described in Ref. 78. These invariants are provided in the Appendix for the convenience of the reader.

The new 6D model is constructed such that for $\theta = 0$ the 5D ANN model from Ref. 78 is reproduced exactly. To this end, the vibronic coupling constants of a reference model λ_j^0 are tuned by the scaled output neurons $\eta_j^{5D}(\tilde{\mathbf{Q}})$. In addition, since the umbrella coordinate is now present, a second ANN, yielding the output neurons $\eta_j^{6D}(\tilde{\mathbf{Q}})$, is introduced to provide the necessary θ -dependence to the λ_i parameters. More specifically, the coefficients λ_{1-3}^0 , which correspond to the adiabatic energies at the reference point, are made coordinate-dependent by

$$\begin{aligned} \lambda_j(\tilde{\mathbf{Q}}) = & \lambda_j^0 \cdot \left(1 + 10^{-5} \cdot \eta_j^{5D}(\tilde{\mathbf{Q}}) \right. \\ & \left. + \theta^2 \cdot 10^{-3} \cdot \eta_j^{6D}(\tilde{\mathbf{Q}}) \right) \quad \forall j \leq 3, \end{aligned} \quad (3)$$

and the coefficients λ_{4-9} , also present in the 5D model, are constructed similarly:

$$\begin{aligned} \lambda_j(\tilde{\mathbf{Q}}) = & \lambda_j^0 \cdot \left(1 + \eta_j^{5D}(\tilde{\mathbf{Q}}) \right. \\ & \left. + \theta^2 \cdot c_j \cdot \eta_j^{6D}(\tilde{\mathbf{Q}}) \right) \quad 4 \leq j \leq 9 \end{aligned} \quad (4)$$

where the η_j^{6D} are the outputs of the additional neural network depending on $\tilde{\mathbf{Q}}$ and which account for the effects of non-planar geometries. The additional scaling factors c_j are introduced to allow for further flexibility if particular reference model terms reside in a different order of magnitude than other terms and hence require a different treatment. The additional parameters λ_{10-14} are defined in analogy to the previously utilized models and correct the new inter-state coupling terms, which vanish for planar geometries and thus were not included in the previous 5D models. These read

$$\lambda_j(\tilde{\mathbf{Q}}) = \lambda_j^0 \cdot \left(1 + c_j \cdot \eta_j^{6D}(\tilde{\mathbf{Q}}) \right) \quad \forall j \geq 10. \quad (5)$$

The constants λ_j^0 are obtained from non-linear least squares fits with respect to accurate *ab initio* reference data and yield a qualitatively correct reference model. In a second step, the two ANNs are trained with respect to the same reference data and provide the \mathbf{Q} -dependent corrections. A global view of the above equations is that the λ_i coefficients are actually modified not by a single but by *two* ANNs, the original θ -independent 5D model from Ref. 78 ($\eta_j^{5D}(\tilde{\mathbf{Q}})$) and the neurons η_j^{6D} from the ANN used in Eq. (5).

The advantage of this approach is that the previously derived 5D ANN model,⁷⁸ which is of excellent quality, is available already and accounts for a large part of the full 6D model. Eqs. 3, 4, and 5 ensure that the coefficients λ_j are constructed in such a way that the model smoothly falls back to the original 5D model for planar geometries. This way, the present, full-dimensional model naturally extends the previous planar model by including out-of-plane motions, while simultaneously correcting the additional contributions of the new vibronic coupling terms. This further serves as a proof of concept that the previously established scheme can be extended without the need of changing the general structure of the *ansatz*.

All neural networks involved in the construction of the model belong to the broad category of feed-forward neural networks. They are functions taking a vector $\tilde{\mathbf{Q}} = \boldsymbol{\eta}^{(1)}$ as input and processing it via intermediate results $\boldsymbol{\eta}^{(k)}$, the so-called hidden layers, to a final (L th) output vector $\boldsymbol{\eta}^{(L)} = \boldsymbol{\eta}^{6D}$ called output layer. The vector elements of each layer are called *neurons*. Each intermediate $\boldsymbol{\eta}^{(k)}$ depends solely on the previous layer $\boldsymbol{\eta}^{(k-1)}$ by

$$\eta_j^{(k)} = f^{(k)}\left(\beta_j^{(k)} + \sum_l \omega_{jl}^{(k)} \eta_l^{(k-1)}\right). \quad (6)$$

$f^{(k)}$ is a nonlinear function of one variable called *activation function*. In the present case, $f^{(k)}$ is chosen to be \tanh for the single unique hidden layer $k = 2$ and as the identity Id for the output ($k = 3$), respectively. $\omega_{jl}^{(k)}$ and $\beta_j^{(k)}$ are called *weights* and *biases*, respectively. They are the fitting parameters of the neural network.

The additional neural network accounting for the θ -dependence of the 6D model contains a single hidden layer with 10 hidden layer neurons, corresponding to a function of approximately 250 formal parameters. While significantly larger (and smaller) networks have been tested, this network size is considered the most suitable. As in our previous work,⁷⁶ this ANN was trained for a set of 100 random initial guesses using a specialized ANN Marquardt-Levenberg method developed in-house. The data set is composed of over 2400 (non-planar) geometries, with a validation data set of about 400 additional geometries. This is a significantly smaller data set than what was used for planar geometries (18,000 geometries), as accurate *ab initio* data for fully asymmetrical (C_1) geometries is significantly more expensive and problematic to attain.

The reference model (including the previous 5D ANN and fitted $\lambda_j^{(0)}$) yields a root mean square (rms) error of 835 cm^{-1} which is significantly higher than the final error after the ANN training of 58.1 cm^{-1} . This fitting error only includes non-planar geometries, as the model is identical to the previous 5D model for planar geometries. The corresponding validation set error is expectedly higher, 70.0 cm^{-1} . While larger networks slightly decrease the resulting rms error by a few inverse cen-

timeters, the validation set error increases significantly, suggesting over fitting. This further motivates our selection of this 10 neuron network as the optimal choice for the present case. The rms error obtained is noticeably higher than what was achieved for the 5D neural network (which had a fitting rms of about 20 cm^{-1}). Most likely, this increase in the rms error is due to the less reliable *ab initio* reference data. Numerous problems with convergence of the electronic structure calculations were observed during the very demanding data acquisition.

A discussion of more technical matters such as potential benefits of using “deeper” neural networks with more hidden layers or the size of the reference data set have been discussed at great length in previous studies^{76–78}, and hence will not be repeated here.

III. COMPUTATIONAL DETAILS

The aim of the present study is to unravel the effects of temperature (hot bands) and vibronic coupling in the low energy regime of the photodetachment spectrum by accurate quantum dynamics simulations. To this end, a time-dependent approach is used for the determination of the detachment spectrum. More specifically, the propagation of a 6-dimensional wavepacket evolving on the 5 coupled PESs of the radical is performed starting from the eigenfunction of each of the lowest vibrational states of NO_3^- . The spectra obtained by the Fourier transform of the resulting autocorrelation functions are subsequently added with the proper Boltzmann weights to simulate the NO_3^- photodetachment spectrum at a given temperature. The MCTDH approach,^{88,89} suitable for the representation of wave functions with large dimensionality is employed for the time propagation of the wave packets. The state average and block diagonalization schemes⁹⁰ are used for the determination of the vibrational eigenfunctions of the anion.

The six internal curvilinear coordinates $\rho^{(\text{cu})}, \vartheta^{(\text{cu})}, \varphi^{(\text{cu})}, \theta^{(\text{cu})}, \phi^{(\text{cu})}, \chi^{(\text{cu})}$, as proposed in ref. 91, are used in the dynamics calculations. The associated *quasi*-exact kinetic operator is approximated by the fourth order Taylor expansion as detailed in ref. 91. Both our previous studies^{74,75} on the first excited state of NO_3 and on the computation of the vibronic levels of the ground state of NO_3 have demonstrated that this choice of coordinates and approximation is efficient and suitable for the study of the NO_3 radical. In the MCTDH approach used, the evaluation of the potential term is carried out using the CDVR scheme⁹² for the anion and using the generalized version as detailed in the appendix of ref. 83 for the evaluation of the diabatic five by five potential matrix of the radical.

The dominant low energy part of the photodetachment experiments^{34,80} is simulated by assuming a vertical transition of one of the NO_3^- vibrational eigenfunctions onto the lowest component corresponding to the ${}^2A_2'$ state of diabatic potential matrix of the neutral system. An ad-

ditional signal is obtained in the low energy domain from the vertical transition to the two E' components of the diabatic model. This second contribution is clearly due to vibronic coupling between the ${}^2A'_2$ ground and ${}^2E'$ excited electronic state of the radical. No significant signal is found in the same energy range when the vertical transition to the two ${}^2E''$ components is treated. A propagation time of 250 fs is found to be sufficient for the comparison of the computed spectrum with the experimental data.

Table I provides the numerical details of the basis set used for the representation of the wave-packets. Careful

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 of hyperspherical type constructed using mass weighted Cartesian coordinates

coord.	n	N	range
$\rho^{(\text{cu})}$	8	64	[635 : 840]
$\vartheta^{(\text{cu})}$	9	96	[0.785 : 1.055]
$\varphi^{(\text{cu})}$	9	64	[0.615 : 0.955]
$\theta^{(\text{cu})}$	6	32	[1.431 : 1.711]
$\phi^{(\text{cu})}$	10	96	[0.820 : 1.295]
$\chi^{(\text{cu})}$	10	64	[2.809 : 3.474]
electronic	5	5	

attention was paid to the definition of the underlying box as well as to the number of single particle functions n and the number of Fourier grid points N to ensure a converged autocorrelation function up to the relevant propagation time. The basis definition is adapted to the excitation to the E' components. A smaller basis would be sufficient for the case of a vertical transition to the lowest diabatic PES only.

The lowest vibrational energy levels of the anion⁷³ as well as their contribution to the final spectrum by means of Boltzmann weights for a 435 Kelvin as in the experimental work³⁴ are presented in Table II. At 10 Kelvin, the only relevant initial vibrational state is the ground state.

When comparing with the experimental data, one must keep in mind the current limits of the theoretical approach. First no overall rotation is explicitly taken into account. The width of the peaks are introduced by an *ad hoc* damping of the autocorrelation function to be Fourier transformed. Second, no *ab initio* information about the photo-detachment cross-section are used. As a substitute, an arbitrary relative weight of the two different electronic state manifolds is used to match the measured spectra of the two experimental setups. And finally, the presented, CPU and memory demanding, computations correspond for the E' manifold to a low intensity part of the full spectrum and are thus more sensitive to round

TABLE II. Ground state energy with respect to the bottom of the potential energy surface and excitation energies for the first six vibrational states of NO_3^- in cm^{-1} . The Boltzmann weights, $B_i = \exp(-E/k_B T)$, at 435 Kelvin are also given in unit of B_1 .

level	description	E_i [cm^{-1}]	B_i/B_1
1	ground state	0 (3038) ^{a)}	1
2	antisym bending	703	0.10
3	antisym bending	703	0.10
4	umbrella	845	0.06
5	sym stretching	1040	0.03
6	antisym stretching	1353	0.01
7	antisym stretching	1353	0.01

a) Zero-point energy in parentheses

off errors. Extensive convergence tests were performed to ensure that the presented data discussed in the following are not artificial.

IV. RESULTS AND DISCUSSION

The present quantum dynamics simulations of the photodetachment spectrum of NO_3^- are designed in order to disentangle unequivocally the effects of temperature and vibronic coupling in the observed spectra. The assignment and interpretation of the spectroscopy of the NO_3^- anion as well as the radical gave rise to several controversies over the past decades. The first photodetachment spectrum published in 1991 was a breakthrough in the understanding of NO_3^- and sparked significant interest.³⁴ The latest slow electron velocity-map imaging (SEVI) experiment⁸⁰ carried out at cryogenic temperature and with various photon energies provides a wealth of new information. The present first principles simulations based on the new and highly accurate diabatic PES model are capable to substantiate the interpretations of those experiments and hopefully end at least some of the debate. An overview of the present simulation results in comparison with these two experiments is given in Fig. 1.

The first issue to be addressed for a complete understanding of the photodetachment spectroscopy of NO_3^- is the influence of temperature and the contribution of hot bands to the spectrum. The absence of one prominent feature (**m**) in the new cryogenically cooled spectrum clearly identifies that peak as hot band. The simulation of the spectrum at a temperature of 435 K, the value estimated from the first photodetachment experiment,³⁴ reproduces this hot band, labeled **m**, and confirms the original assignment as 4_1^1 transition. From Fig. 1 and Fig. 2, the excellent agreement between theory and experiment becomes obvious.

The full photodetachment spectrum is simulated by two partial spectra corresponding to the contributions of

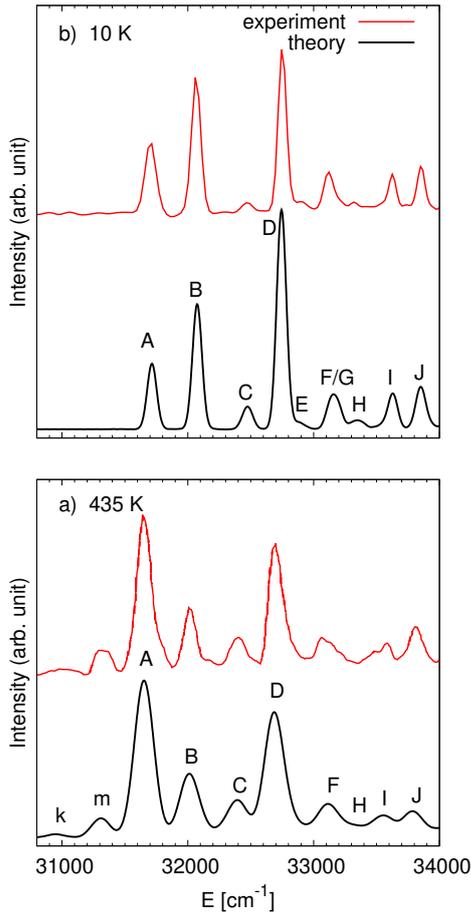


FIG. 1. Comparison of experimental photodetachment spectra with the final quantum dynamics simulation results. Panel a) presents the theoretical and experimental³⁴ spectra at an estimated temperature of 435 K. Panel b) shows results at 10 K in comparison with the new cryo-SEVI results.⁸⁰ A damping time of 200 fs (resp. 100 fs) is used for the 10K (resp. 435K) theoretical spectra. A ratio of 1/150 (resp. 1/30) is used for the \tilde{X}^2A_2' and the \tilde{B}^2E' partial spectra at 10K (resp. 435K)

the \tilde{X}^2A_2' and the \tilde{B}^2E' state, respectively. These two partial spectra are also given in Fig. 2 in order to make their contributions to the experimentally observed spectrum more clear. The reason for this approach is that the detachment cross-sections corresponding to the two different electronic state manifolds are very different and need to be taken into account. The scaling factors due to these different cross-sections and the energy shift due to the lack of size extensivity of the underlying *ab initio* calculations are the only empirical factors used in the present study. Each of these partial spectra can be decomposed further into the contributions originating from different thermally populated initial states in the NO_3^- anion.

The detailed and decomposed partial spectrum corresponding to the \tilde{X} state manifold is presented in panel

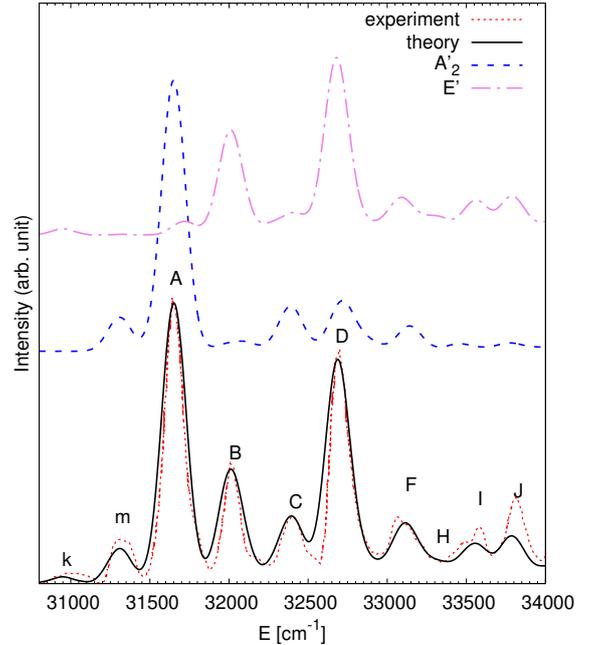


FIG. 2. The theoretically simulated spectrum at $T=435$ K (black) is compared to the experimental hot spectrum from ref. 34 (dotted red). The partial contributions from the \tilde{X}^2A_2' and the \tilde{B}^2E' states are also shown in blue dashed and violet dotted-dashed curves, respectively. The \tilde{B}^2E' contribution is scaled by 30 with respect to the \tilde{X}^2A_2' one.

a) of Figure 3. The photodetachment spectrum obtained when using the vibrational ground state of the anion after a vertical excitation onto the lowest diabatic component is presented in red. The blue spectra are obtained when the initial wavepacket bears one quantum of energy in the anti-symmetric bending of NO_3^- , and the violet trace is obtained when the initial wavepacket is excited in the umbrella mode. The black curve is the resulting spectrum considering a temperature of 435 K thus taking into account the Boltzmann factors of the initial states (see Table II). Because of the lack of size extensivity of the underlying *ab initio* method, the computed absolute energies of the radical and of the anion cannot be compared. For this reason, we present theoretical spectra shifted in energy so that the most intense peak (**A**) obtained matches the experimental data. Only the energy range of the relevant experimental data is shown.

The partial spectra from the wave packet propagations provide information about the entire progression of the corresponding hot band contributions to the total spectrum at 435 K. Some of these peaks superimpose features of the cold spectrum and may lead to a broadening because of slightly shifted transition energies. This is entirely the case for all contributions when a single ν_2 (umbrella) excitation in the anion is considered. Only excitation of ν_4 (asymmetric bend) gives rise to additional features clearly distinguishable from peaks in the

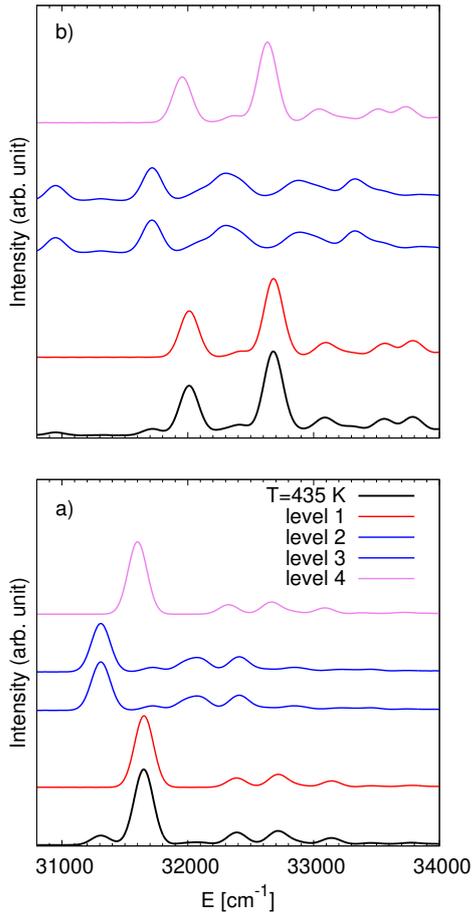


FIG. 3. Partial spectra at 435 K (black) corresponding to excitation to \tilde{X}^2A_2 (a) and to \tilde{B}^2E' (b). The contributions from the different vibrational states of the anion are also displayed: vibrational ground state in red, anti-symmetric bending in blue and umbrella in violet.

cold spectrum. Probably the most interesting of these additional contributions is found at the position of the disputed first peak above the origin 0_0^0 transition labeled **B** in Fig. 1. This feature was attributed to the Franck-Condon (FC) forbidden 4_0^1 transition by Weaver *et al.*³⁴ but assigned to the FC allowed 4_1^3 hot band transition by Yamada and Ross.⁹³ The present results clearly show that there is a corresponding hot band transition at this energy but its intensity is way too small to explain the experimentally observed spectrum. This is also in agreement with the fact that this peak is very prominent in the cryogenically cooled spectrum where hot bands are absent.

Excitations of the same initial NO_3^- vibrational eigenfunctions to the two $^2E'$ components of NO_3 result in spectra provided in panel b) of Figure 3. The vertical axes are not scaled to the same arbitrary unit for the two panels, however. The ground state as well as the ν_2 (umbrella) excitation produce partial spectra with two

main peaks. The first of them is superposed to the 4_1^3 hot band at position **B** (see blue curve in panel a) and by this explaining the higher intensity of this peak in the final spectrum as seen in Fig. 3.

The second prominent feature of this partial spectrum is found at a position roughly 1040 cm^{-1} above the 0_0^0 peak and was assigned originally to the FC allowed 1_0^1 transition by Weaver *et al.*³⁴ (labeled **D** in Fig. 1). However a 1_0^1 transition should not show up in the partial spectrum corresponding to the \tilde{B} state manifold. In fact, we do observe a weak feature slightly to the blue of this peak in the \tilde{X} state partial spectrum but its intensity would be too low to explain the experimentally observed spectrum. When we compare the peak positions of the partial photodetachment spectra with our earlier studies providing accurate vibrational eigenstate energies for the \tilde{X} manifold, we find that FC allowed 1_0^1 transition should correspond to about 1056 cm^{-1} above 0_0^0 while the FC forbidden 3_0^1 transition would correspond to about 1036 cm^{-1} above origin. The latter assignment is in agreement with the interpretation of the new cryo-SEVI spectrum.⁸⁰ Thus, the main intensity of the observed peak **D** apparently originates from the \tilde{B} state manifold and indicates significant vibronic coupling between the two involved electronic states. The present results show that both assignments are correct because the experimentally observed feature is a superposition of both detachment channels.

Some weaker additional features due to the \tilde{B} state manifold are found at higher energies but they are fairly weak compared to the two very prominent peaks discussed above. The temperature effects can be included in the same way as for the \tilde{X} state manifold and a similar behaviour is found. Hot band contributions due to a single quantum in ν_2 (umbrella) in the initial state yields a spectrum very similar to the one obtained for the sole ground state, the only partial spectrum relevant at cryogenic temperature (10 K). By contrast, a single quantum in ν_4 (asymmetric bend) in the initial state results in a very different spectrum. The first prominent peak is slightly shifted to the blue compared to the origin peak in the \tilde{X} state manifold. Next there is a fairly broad feature in the region around 32300 cm^{-1} where there is little intensity originating from the \tilde{X} state manifold. Two similarly broad features are observed above around 32880 and 33330 cm^{-1} .

At lower energy, one very interesting observation is some intensity around $31,000 \text{ cm}^{-1}$ and thus about 700 cm^{-1} below the 0_0^0 origin peak of the experimental spectrum. Apparently, this is a hot band due to a 4_1^1 transition, which would be FC forbidden in the \tilde{X} state manifold but becomes FC allowed in the \tilde{B} state manifold just like the 4_0^1 transition. It appears that this peak labeled **k** in Fig. 1 was not assigned before because the combination of hot band and vibronic coupling is necessary to explain this transition. The hot band character is evident now from the cryo-SEVI experiment while its vibronic coupling origin only becomes clear from the

quantum dynamics simulations.

The various contributions from the \tilde{B} state manifold also show intensity above $33,500\text{ cm}^{-1}$ where there is very little intensity seen from the \tilde{X} state manifold. This results in two visible peaks at positions **I** and **J** which will not disappear at cryogenic temperatures. This is in agreement with experiment as well.

The new cryo-SEVI experiment also offers further data to disentangle the contributions from vibronic coupling between the \tilde{X} and \tilde{B} state by varying the laser frequency near threshold. Thus, Wigner's threshold law can be invoked together with the symmetry rules for allowed detachment channels. It turns out that all FC allowed transitions should result in p-wave electrons having an angular momentum of $l = 1$ while all FC forbidden transitions, which only gain intensity through vibronic coupling to the \tilde{B} state, should produce s-wave electrons with an angular momentum $l = 0$ as was also discussed in Ref. 80. Thus, the energy dependent intensity behavior will scale with $\Delta E^{3/2}$ for \tilde{X} contributions but with $\Delta E^{1/2}$ for \tilde{B} contributions near threshold. Here ΔE is the electron kinetic energy depending on the photon energy used. This effect can be taken into account in the present quantum dynamics simulations and the resulting spectra are presented together with the experimental ones in Fig. 4. In the theoretical simulation, a unique fixed ratio for the contributions of the partial \tilde{X} and \tilde{B} states is used. The variation of the relative peak intensities is only due to the Wigner energy dependence introduced by the $\Delta E^{3/2}$ and $\Delta E^{1/2}$ additional multiplicative factors.

The first observation seen at first glance is that the intensity pattern is grossly different from the high temperature spectrum. The second very important point is that the theoretical simulation taking the near threshold effects into account are in very good agreement with experiment. To show the agreement more clearly, a direct comparison between experimental spectra and the corresponding simulation is presented for each of the six different photon energies used in the experiment in Fig. 5.

This level of agreement is only possible if the assumptions made for the simulation are reasonable of course, giving evidence for both the assumption of the electron wave character as well as the vibronic coupling between the \tilde{X} and \tilde{B} state. Surely, this agreement is not and cannot be perfect. One important effect that is not accounted for is the narrowing of the experimental lines when the photon energy is lowered. In the theoretical simulations, the lines widths are induced by the damping of the autocorrelation functions before Fourier transformation, while in the experiment, overall rotation as well as experimental resolution determine the line widths. Only peak intensities are compared in the present context though the Wigner threshold law affects the detachment cross-section rather than the intensity. So, in the following discussion it should be kept in mind that a peak with a smaller half-width will be more intense for the same cross-section. The simulation shows the same half-widths regardless of the photon energy contrary to the

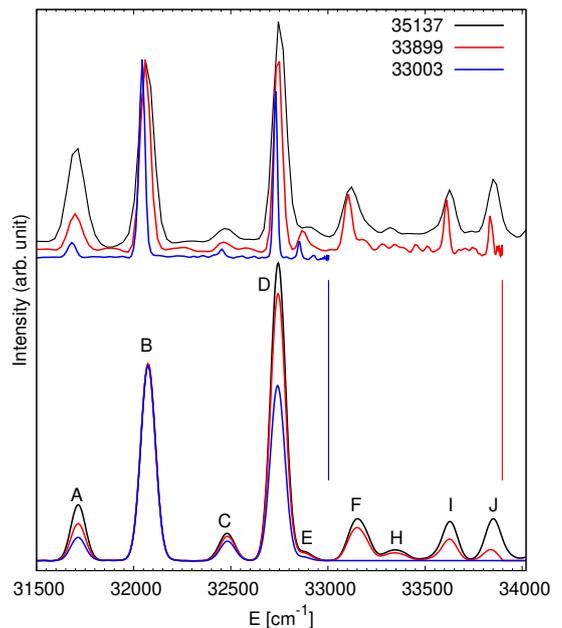


FIG. 4. Experimental (top) and theoretical (bottom) spectra at three photon energies: 35137 cm^{-1} (black), 33899 cm^{-1} (red) and 33003 cm^{-1} (blue). The photon energies are materialized by the vertical red and blue lines. The 35137 cm^{-1} (black) is outside the plotted energy range. The experimental spectra are reproduced from Fig. 3a of ref. 80 with permission from the authors. Both the experimental spectra and the theoretical simulations are scaled such as the maxima of peak **B** are identical for the three photon energies.

experimental features.

In the following, the effect of the near threshold detachment is discussed in detail. The most surprising point is that the origin 0_0^0 peak (**A**) by far is not the most intense feature anymore (see Fig. 4). For all but the two lowest photon energies, peak **D** shows the highest intensity. Peak **B** becomes the dominant feature for the two lowest photon energies, followed by **D**. On the contrary, according to a simple standard FC picture the 0_0^0 transition would be expected as the dominating line of the spectrum because the equilibrium structures of the anion and the radical are very similar. One also would expect this to become even more pronounced when detaching at near threshold but the contrary is the case. Furthermore, both the **A/B** as well as the **A/D** intensity ratio are decreasing significantly rather than increasing when the photon energy is reduced. Note that the intensity of all spectra was normalized with respect to peak **B**. The reason for the above observation clearly is the different electron wave character because **B** and **D** are mainly due to the \tilde{B} state and resulting in s-wave electrons while **A** is entirely due to the \tilde{X} state corresponding to p-wave electrons. The present detailed quantum dynamics simulations show unequivocally that this effect cannot be reproduced without the vibronic coupling effects between

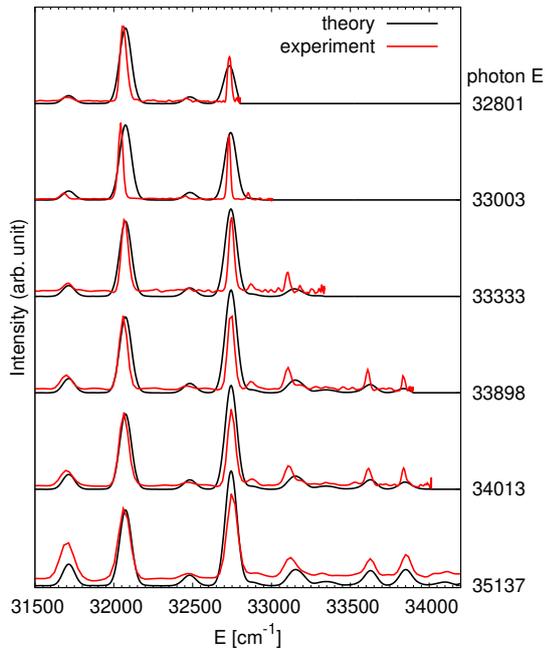


FIG. 5. Experimental (red) and theoretical (black) spectra at various photon energies as given in cm^{-1} along the vertical axis. Experimental⁸⁰ and theoretical spectra are scaled such as the maximum of peak **B** are identical. The fixed contribution of the partial \tilde{X} and \tilde{B} state in the theoretical simulation is identical for all six plots. The variation of the relative peak intensities is only due to the Wigner energy dependence.

these two electronic states. There would be no contribution from the \tilde{B} state manifold in the energy region of the present study without significant vibronic coupling. This conclusion is the same as in previous theoretical studies^{64,80} and is strengthened immensely by the excellent agreement between our detailed simulations of the near threshold effects and the cryo-SEVI experiments.

This agreement is found not only for the **A/B** and **A/D** but also for the **B/D** intensity ratio. In this case, a slight increase is observed when the photon energy is lowered showing that these two peaks belong to the same electronic state manifold and thus l quantum number of the outgoing electrons. It is also clear from the simulations that the corresponding case is $l = 0$ and the intensity of **B** originates dominantly from the \tilde{B} state manifold. This is of particular interest because the assignment of **B** as the FC forbidden 4_0^1 detachment^{34,80} has been disputed by Hirota⁹⁴ based on the interpretation of Yamada and Ross.⁹³ However, the present detailed simulation results clearly support the assignment of feature **B** as 4_0^1 detachment. The near threshold results also clearly support another debated assignment namely that of peak **D**. It is unequivocal that the main contribution to that feature originates from the \tilde{B} state and corresponds to a FC forbidden 3_0^1 detachment. This was already discussed above when analysing the partial spectra from the sim-

ulations. The FC allowed 1_0^1 detachment predicted close by in energy is suppressed in these near threshold spectra and contribute even less to peak **D** than in the hot spectrum. And even in the 435 K spectrum 1_0^1 detachment has only a minor effect compared to FC forbidden 3_0^1 detachment, contrary to the initial assignment.

Finally, there is also an interesting observation for peak **C** in both experiment and simulation. The ratio of **A/C** decreases strongly when the photon energy is lowered. This indicates that **A** and **C** correspond to different l quantum numbers for the outgoing electron wave. In fact, the assignment of peak **C** to 4_0^2 detachment turns out to be correct but the double excitation of a degenerate e' mode leads to a more complex picture for **C** when compared to **A**. The temptation to assume a FC allowed transition corresponding to the \tilde{X} state manifold just like peak **A** does not account for this complexity. Indeed, the present case leads to two different sub levels of a'_1 and e' symmetry. These two sub levels are found to be only 20 cm^{-1} apart in energy and the 4_0^2 (e') sub level corresponds to a FC forbidden transition becoming allowed by coupling to the \tilde{B} state. This latter sub level is less affected by the near threshold effect than the FC allowed a'_1 sub level because the corresponding electron wave belongs to $l = 0$ rather than to $l = 1$. It appears that in the 435 K spectrum this peak is dominated by the 4_0^2 (a'_1) sub level and thus having a higher intensity, while in the cryo-SEVI spectrum this feature is mainly due to the 4_0^2 (e') sub level resulting in a lower but less photon energy dependent intensity.

V. CONCLUSIONS AND OUTLOOK

A new and highly accurate full-dimensional and fully coupled diabatic potential energy (PES) model is presented for the lowest electronic states of NO_3 , namely $\tilde{X}^2A'_2$, \tilde{A}^2E'' , and \tilde{B}^2E' . This model is based on high quality *ab initio* reference data that was diabaticized by the recently developed method of complete nuclear permutation inversion invariant artificial neural network (CNPI-ANN) diabaticization.⁷⁸ A reduced-dimensional model for planar NO_3 was extended by adding a second ANN to account for the out-of-plane motion. This model is extremely efficient and reproduces the *ab initio* reference data very accurately. Quantum wave packet dynamics simulations using this model are performed to study the photodetachment spectroscopy of the NO_3^- anion detaching from the energetic region of the \tilde{X} state of the radical. These calculations are designed such that the two available experiments, an older recording at elevated temperature³⁴ and a very recent SEVI spectrum at cryogenic temperature⁸⁰ are reproduced. To this end, the effect of temperature as well as near threshold effects (Wigner's threshold law) are accounted for. Furthermore, the strong *pseudo*-Jahn-Teller coupling between the \tilde{X} and \tilde{B} state leads to additional contributions in the energetic region of the \tilde{X} state manifold, which con-

tribute considerably to the experimentally observed spectra and are modeled properly in the present simulations. Very good agreement is achieved giving strong evidence to the interpretation of the detailed theoretical data.

Temperature effects are taken into account in a straightforward way to simulate the contribution of hot bands. To this end, partial spectra are computed for several initial vibrational states of the anion and the total spectrum is obtained as a weighted superposition using the Boltzmann factors for a temperature of 435 K. This does not yield a satisfactory spectrum if only the \tilde{X} state manifold is considered. Therefore, a second set of partial spectra is computed accounting for the \tilde{B} state manifold and yielding Franck-Condon (FC) forbidden transitions which become allowed through the strong *pseudo*-Jahn-Teller coupling. Only the weighted superposition of all partial spectra from \tilde{X} state and \tilde{B} state manifold yield a simulated spectrum that agrees with experiment. In fact, the obtained agreement is excellent, giving strong support for the assumptions made in the simulation. The simulations allow to identify the 0_0^0 transition in agreement with experiment and two hot bands to the red of it, namely 4_1^1 and 4_1^0 . The latter transition is FC forbidden and borrows its intensity from the \tilde{B} state apparently. The first prominent peak to the blue is a superposition of the weak FC allowed 4_1^3 hot band and the much stronger FC forbidden 4_0^1 detachment. The present result hopefully will settle the debate about the corresponding assignments. A further strong experimental feature a bit more than 1000 cm^{-1} above the 0_0^0 line also turns out to be a superposition and a weak FC allowed 1_0^1 detachment and a much stronger FC forbidden 3_0^1 transition that borrows intensity from the \tilde{B} state through vibronic coupling. This yields further support from both experiment and theory that the disputed ν_3 fundamental of neutral NO_3 is found around 1040 cm^{-1} rather than at 1492 cm^{-1} .

The evidence for the above assignments is strengthened by the simulation of the new cryo-SEVI spectra in which the near threshold effects (Wigner's threshold law) are taken into account. The two partial spectra for the \tilde{X} and \tilde{B} state correspond to different angular momentum for the detached electrons, $l = 1$ for \tilde{X} and $l = 0$ for \tilde{B} . The different threshold behaviour allows to disentangle which peaks are dominated by which electronic state manifold. The cryogenic temperature of 10 K removes all contributions from hot bands. The experimental cryo-SEVI spectra taken at six different photon energies show very different intensity patterns from the 435 K spectrum, yet they are simulated in very good agreement using the present model. The detailed analysis confirms the above assignments and furthermore identifies another peak as a superposition of weak FC allowed 4_0^2 (a_1') and strong FC forbidden 4_0^2 (e') detachment. All these results clearly show the importance of the vibronic coupling between the \tilde{X} and \tilde{B} state without which it would be impossible to simulate the experimental results that well.

Finally, the very good agreement with experiment obtained for all the spectra indicates the accuracy of the underlying diabatic PES model. The model also includes the \tilde{A} state for which at least a photodetachment spectrum at elevated temperature was recorded and published. It can be expected that the present PES model also will yield detailed simulation results for that spectrum and this shall be studied in the near future. Furthermore, the model allows to investigate the nonadiabatic dynamics after exciting the wave packet to one of the excited states. This may allow to disentangle some further experimental observations, like the \tilde{B} state lifetime, which are not well-understood, yet. We hope that the continued work will lift some more secrets of the enigmatic NO_3 radical in the future.

VI. APPENDIX

A. Coordinates Q

The symmetry-adapted coordinates are constructed from a set of primitive valence coordinates as already described in previous works^{73,83} 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 non-linearly as

$$m_i = 1 - \exp(-\gamma(r_i - r_0)) \quad (7a)$$

$$\alpha'_i = \frac{\alpha_i - \alpha_0}{r_j r_k}, \quad i \neq j \neq k \quad (7b)$$

where r_0 and α_0 are the respective distances and angles at the reference point and γ is a chosen Morse-parameter. 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).

$$a = \sqrt{\frac{1}{3}}(m_1 + m_2 + m_3) \quad (8a)$$

$$x_s = \sqrt{\frac{1}{6}}(2m_1 - m_2 - m_3) \quad (8b)$$

$$y_s = \sqrt{\frac{1}{2}}(m_2 - m_3) \quad (8c)$$

$$x_b = \sqrt{\frac{1}{6}}(2\alpha'_1 - \alpha'_2 - \alpha'_3) \quad (8d)$$

$$y_b = \sqrt{\frac{1}{2}}(\alpha'_2 - \alpha'_3). \quad (8e)$$

The symmetry-adapted out-of-plane umbrella coordinate θ is constructed from the trisector angle ϑ as follows:

$$\theta = \frac{\vartheta - \frac{\pi}{2}}{r_1 r_2 r_3} \quad (9)$$

B. Invariants $\tilde{\mathbf{Q}}$

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

$$r_{s,b}^2 = x_{s,b}^2 + y_{s,b}^2. \quad (10)$$

Furthermore, let the third order invariant terms τ_{jk} with $j, k \in \{s, b\}$ be defined as

$$\tau_{jk} = x_j^2 x_k - x_k y_j^2 - 2x_j y_k y_j. \quad (11)$$

$\tilde{\mathbf{Q}}$ is then composed of the following invariants:

$$\tilde{Q}_1 = a \quad (12a)$$

$$\tilde{Q}_2 = r_s^2 \quad (12b)$$

$$\tilde{Q}_3 = r_b^2 \quad (12c)$$

$$\tilde{Q}_4 = \theta^2 \quad (12d)$$

$$\tilde{Q}_5 = x_s x_b + y_s y_b \quad (12e)$$

$$\tilde{Q}_6 = 10 \cdot \tanh(0.1 \cdot \tau_{ss}) \quad (12f)$$

$$\tilde{Q}_7 = 10 \cdot \tanh(0.1 \cdot \tau_{bb}) \quad (12g)$$

$$\tilde{Q}_8 = 10 \cdot \tanh(0.1 \cdot \tau_{sb}) \quad (12h)$$

$$\tilde{Q}_9 = 10 \cdot \tanh(0.1 \cdot \tau_{bs}) \quad (12i)$$

C. Diagonal model terms

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

$$\mathbf{W}_{\text{diag}}^d(\mathbf{Q}) = \begin{pmatrix} V_1(\mathbf{Q}) & 0^T & 0^T \\ 0 & V_2(\mathbf{Q})\mathbf{1} & 0 \\ 0 & 0 & V_3(\mathbf{Q})\mathbf{1} \end{pmatrix} \quad (13)$$

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.

$$V_i(\mathbf{Q}) = \mu_1^i \cdot a + \mu_2^i \cdot a^2 + \mu_3^i \cdot r_s^2 + \mu_4^i \cdot r_b^2 \quad (14)$$

where r_i^2 are defined in Eq. (10).

ACKNOWLEDGMENTS

The authors are grateful for financial support via the PHC/DAAD grant PROCOPE 40442PD. Part of this work was also generously supported by the Deutsche Forschungsgemeinschaft (DFG). The authors are indebted to Mark C. Babin and Prof. Daniel Neumark for sharing the experimental data as well as for valuable discussions.

- ¹R. P. Wayne, *Chemistry of Atmospheres* (Oxford University Press, Oxford, 2000), 3rd ed.
- ²P. Hautefeuille and J. Chappuis, *C. R. Acad. Sci. Paris* **92**, 80 (1881).
- ³G. Sprenger, *Z. Elektrochem.* **37**, 674 (1931).
- ⁴E. J. Jones and O. R. Wulf, *J. Chem. Phys.* **5**, 873 (1937).
- ⁵G. Schott and N. Davidson, *J. Am. Chem. Soc.* **80**, 1841 (1958).
- ⁶H. S. Johnston and R. A. Graham, *Can. J. Chem.* **52**, 1415 (1974).
- ⁷R. A. Graham and H. S. Johnston, *J. Phys. Chem.* **82**, 254 (1978).
- ⁸D. N. Mitchell, R. P. Wayne, J. P. Allen, R. P. Harrison, and R. J. Twin, *J. Chem. Soc. Faraday Trans. 2* **76**, 785 (1980).
- ⁹D. A. Ramsay, *Proc. Colloq. Spectrosc. Int.* **10**, 583 (1962).
- ¹⁰W. J. Marinelli, D. M. Swanson, and H. S. Johnston, *J. Chem. Phys.* **76**, 2864 (1982).
- ¹¹H. H. Nelson, L. Pasternack, and J. R. McDonald, *J. Phys. Chem.* **87**, 1286 (1983).
- ¹²H. H. Nelson, L. Pasternack, and J. R. McDonald, *J. Chem. Phys.* **79**, 4279 (1983).
- ¹³T. Ishiwata, I. Fujiwara, Y. Naruge, K. Obi, and I. Tanaka, *J. Phys. Chem.* **87**, 1349 (1983).
- ¹⁴T. Ishiwata, I. Tanaka, K. Kawaguchi, and E. Hirota, *J. Chem. Phys.* **82**, 2196 (1985).
- ¹⁵K. Kawaguchi, E. Hirota, T. Ishiwata, and I. Tanaka, *J. Chem. Phys.* **93**, 951 (1990).
- ¹⁶K. Kawaguchi, T. Ishiwata, I. Tanaka, and E. Hirota, *Chem. Phys. Lett.* **180**, 436 (1991).
- ¹⁷T. Ishiwata, I. Tanaka, K. Kawaguchi, and E. Hirota, *J. Mol. Spectrosc.* **153**, 167 (1992).
- ¹⁸E. Hirota, T. Ishiwata, K. Kawaguchi, M. Fujitake, N. Ohashi, and I. Tanaka, *J. Chem. Phys.* **107**, 2829 (1997).
- ¹⁹K. Kawaguchi, T. Ishiwata, E. Hirota, and I. Tanaka, *Chem. Phys.* **231**, 193 (1998).
- ²⁰T. Ishiwata, Y. Nakano, K. Kawaguchi, E. Hirota, and I. Tanaka, *J. Phys. Chem. A* **114**, 980 (2010).
- ²¹K. Kawaguchi, N. Shimizu, R. Fujimori, J. Tang, T. Ishiwata, and I. Tanaka, *J. Mol. Spectrosc.* **268**, 85 (2011).
- ²²R. Fujimori, N. Shimizu, J. Tang, T. Ishiwata, and K. Kawaguchi, *J. Mol. Spectrosc.* **283**, 10 (2013).
- ²³K. Kawaguchi, R. Fujimori, J. Tang, and T. Ishiwata, *J. Phys. Chem. A* **117**, 13732 (2013).
- ²⁴K. Tada, W. Kashiara, M. Baba, T. Ishiwata, E. Hirota, and S. Kasahara, *J. Chem. Phys.* **141**, 184307 (2014).
- ²⁵K. Tada, K. Teramoto, T. Ishiwata, E. Hirota, and S. Kasahara, *J. Chem. Phys.* **142**, 114302 (2015).
- ²⁶E. Hirota, *J. Mol. Spectrosc.* **310**, 99 (2015).
- ²⁷K. Kawaguchi, R. Fujimori, J. Tang, and T. Ishiwata, *J. Mol. Spectrosc.* **314**, 73 (2015).
- ²⁸K. Tada, T. Ishiwata, E. Hirota, and S. Kasahara, *J. Mol. Spectrosc.* **321**, 23 (2016).
- ²⁹K. Kawaguchi, T. Narahara, R. Fujimori, J. Tang, and T. Ishiwata, *J. Mol. Spectrosc.* **334**, 10 (2017).
- ³⁰M. E. Jacox and W. E. Thompson, *J. Chem. Phys.* **129**, 204306 (2008).
- ³¹H. Beckers, H. Willner, and M. E. Jacox, *Chem. Phys. Chem.* **10**, 706 (2009).
- ³²M. E. Jacox and W. E. Thompson, *J. Phys. Chem. A* **114**, 4712 (2010).
- ³³R. R. Friedl and S. P. Sander, *J. Phys. Chem.* **91**, 2721 (1987).
- ³⁴A. Weaver, D. W. Arnold, S. E. Bradforth, and D. M. Neumark, *J. Chem. Phys.* **94**, 1740 (1991).
- ³⁵B. Kim, P. L. Hunter, and H. S. Johnston, *J. Chem. Phys.* **96**, 4057 (1992).
- ³⁶R. T. Carter, K. F. Schmidt, H. Bitto, and J. R. Huber, *Chem. Phys. Lett.* **257**, 297 (1996).
- ³⁷J. Orphal, C. E. Fellows, and P. M. Flaud, *J. Geophys. Res. - Atmosph.* **108**, 4077 (2003).
- ³⁸A. Deev, J. Sommar, and M. Okumura, *J. Chem. Phys.* **122**, 224305 (2005).

- ³⁹K. Takematsu, N. C. Eddingsaas, D. J. Robichaud, and M. Okumura, *Chem. Phys. Lett.* **555**, 57 (2013).
- ⁴⁰T. Codd, M.-W. Chen, M. Roudjane, J. F. Stanton, and T. A. Miller, *J. Chem. Phys.* **142**, 184305 (2015).
- ⁴¹T. E. H. Walker and J. A. Horsley, *Mol. Phys.* **21**, 939 (1971).
- ⁴²A. Lund and K. Thuomas, *Chem. Phys. Lett.* **44**, 569 (1976).
- ⁴³J. F. Olsen and L. Burnelle, *J. Am. Chem. Soc.* **92**, 3659 (1970).
- ⁴⁴N. C. Baird and K. F. Taylor, *Chem. Phys. Lett.* **80**, 83 (1981).
- ⁴⁵P. E. M. Siegbahn, *J. Comput. Chem.* **6**, 182 (1985).
- ⁴⁶B. Kim, H. S. Johnston, D. A. Clabo Jr., and H. F. Schaefer III, *J. Chem. Phys.* **88**, 3204 (1988).
- ⁴⁷R. C. Boehm and L. L. Lohr, *J. Phys. Chem.* **93**, 3430 (1989).
- ⁴⁸R. C. Boehm and L. L. Lohr, *J. Comput. Chem.* **112**, 119 (1991).
- ⁴⁹R. D. Davy and H. F. Schaefer III, *J. Chem. Phys.* **91**, 4410 (1989).
- ⁵⁰B. Kim, B. L. Hammond, W. A. Lester Jr., and H. S. Johnston, *Chem. Phys. Lett.* **168**, 131 (1990).
- ⁵¹V. R. Morris, S. C. Bhatia, and J. H. Hall Jr., *J. Phys. Chem.* **94**, 7414 (1990).
- ⁵²V. R. Morris, S. C. Bhatia, and J. H. Hall, Jr., *J. Phys. Chem.* **95**, 9203 (1991).
- ⁵³J. F. Stanton, J. Gauss, and R. J. Bartlett, *J. Chem. Phys.* **94**, 4084 (1991).
- ⁵⁴J. F. Stanton, J. Gauss, and R. J. Bartlett, *J. Chem. Phys.* **97**, 5554 (1992).
- ⁵⁵U. Kaldor, *Chem. Phys. Lett.* **166**, 599 (1990).
- ⁵⁶U. Kaldor, *Chem. Phys. Lett.* **185**, 131 (1991).
- ⁵⁷A. Stirling, I. Pápai, J. Mink, and D. R. Salahub, *J. Chem. Phys.* **100**, 2910 (1994).
- ⁵⁸L. A. Eriksson, J. Wang, R. J. Boyd, and S. Lunell, *J. Phys. Chem.* **98**, 792 (1994).
- ⁵⁹T. D. Crawford, T. J. Lee, N. C. Handy, and H. F. Schaefer III, *J. Chem. Phys.* **107**, 9980 (1997).
- ⁶⁰C. D. Sherrill, M. S. Lee, and M. Head-Gordon, *Chem. Phys. Lett.* **302**, 425 (1999).
- ⁶¹W. Eisfeld and K. Morokuma, *J. Chem. Phys.* **113**, 5587 (2000).
- ⁶²W. Eisfeld and K. Morokuma, *J. Chem. Phys.* **114**, 9430 (2001).
- ⁶³W. Eisfeld and K. Morokuma, *J. Chem. Phys.* **117**, 4361 (2002).
- ⁶⁴M. Mayer, L. S. Cederbaum, and H. Köppel, *J. Chem. Phys.* **100**, 899 (1994).
- ⁶⁵J. F. Stanton, *J. Chem. Phys.* **126**, 134309 (2007).
- ⁶⁶M. Okumura, J. Stanton, A. Deev, and J. Sommar, *Phys. Scripta* **73**, C64 (2006).
- ⁶⁷J. F. Stanton, *Mol. Phys.* **107**, 1059 (2009).
- ⁶⁸J. F. Stanton and M. Okumura, *Phys. Chem. Chem. Phys.* **11**, 4742 (2009).
- ⁶⁹C. S. Simmons, T. Ichino, and J. F. Stanton, *J. Phys. Chem. Lett.* **3**, 1946 (2012).
- ⁷⁰Z. Homayoon and J. M. Bowman, *J. Chem. Phys.* **141**, 161104 (2014).
- ⁷¹S. Mahapatra, W. Eisfeld, and H. Köppel, *Chem. Phys. Lett.* **441**, 7 (2007).
- ⁷²S. Faraji, H. Köppel, W. Eisfeld, and S. Mahapatra, *Chem. Phys.* **347**, 110 (2008).
- ⁷³W. Eisfeld, O. Vieuxmaire, and A. Viel, *J. Chem. Phys.* **140**, 224109 (2014).
- ⁷⁴W. Eisfeld and A. Viel, *J. Chem. Phys.* **146**, 034303 (2017).
- ⁷⁵A. Viel and W. Eisfeld, *Chem. Phys.* **509**, 81 (2018).
- ⁷⁶D. M. G. Williams and W. Eisfeld, *J. Chem. Phys.* **149**, 204106 (2018).
- ⁷⁷D. M. G. Williams, A. Viel, and W. Eisfeld, *J. Chem. Phys.* **151**, 164118 (2019).
- ⁷⁸D. M. G. Williams and W. Eisfeld, *J. Phys. Chem. A* **124**, 7608 (2020).
- ⁷⁹A. D. Walsh, *J. Chem. Soc.* pp. 2301–2306 (1953).
- ⁸⁰M. C. Babin, J. A. DeVine, M. DeWitt, J. F. Stanton, and D. M. Neumark, *J. Phys. Chem. Lett.* **11**, 395 (2020).
- ⁸¹A. Viel and W. Eisfeld, *J. Chem. Phys.* **120**, 4603 (2004).
- ⁸²W. Eisfeld and A. Viel, *J. Chem. Phys.* **122**, 204317 (2005).
- ⁸³A. Viel, W. Eisfeld, S. Neumann, W. Domcke, and U. Manthe, *J. Chem. Phys.* **124**, 214306 (2006).
- ⁸⁴W. Eisfeld, *J. Chem. Phys.* **134**, 054303 (2011).
- ⁸⁵F. Venghaus and W. Eisfeld, *J. Chem. Phys.* **144**, 114110 (2016).
- ⁸⁶N. Wittenbrink, F. Venghaus, D. Williams, and W. Eisfeld, *J. Chem. Phys.* **145**, 184108 (2016).
- ⁸⁷W. Domcke, D. R. Yarkony, and H. Köppel, eds., *Conical Intersections: Electronic Structure, Dynamics and Spectroscopy* (World Scientific, Singapore, 2004).
- ⁸⁸H. D. Meyer, U. Manthe, and L. S. Cederbaum, *Chem. Phys. Lett.* **165**, 73 (1990).
- ⁸⁹U. Manthe, H. D. Meyer, and L. S. Cederbaum, *J. Chem. Phys.* **97**, 3199 (1992).
- ⁹⁰U. Manthe, *J. Chem. Phys.* **128**, 064108 (2008).
- ⁹¹C. Evenhuis, G. Nyman, and U. Manthe, *J. Chem. Phys.* **127**, 144302 (2007).
- ⁹²U. Manthe, *J. Chem. Phys.* **105**, 6989 (1996).
- ⁹³K. Yamada and S. Ross, S.96th Ann. Meeting Chem. Soc. Jpn., Kyoto, Kyoto, Japan pp. Paper 2E6–28 (2016).
- ⁹⁴E. Hirota, *J. Mol. Spectrosc.* **343**, 81 (2018).

Hiermit versichere ich, die vorgelegte Dissertation selbstständig und ohne unzulässige Hilfe angefertigt zu haben. Die verwendeten Quellen und Hilfstexte sind vollständig angegeben und die Stellen der Arbeit, einschließlich Tabellen und Abbildungen, die anderen Werken im Sinn und Wortlaut entnommen wurden, als Entlehnung kenntlich gemacht. Diese Arbeit wurde nicht bereits zuvor als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht.

Mir ist die geltende Promotionsordnung der Fakultät für Chemie bekannt. Ich versichere weiterhin, weder ähnliche noch andere Abhandlungen als Dissertation an anderen Hochschulen eingereicht zu haben.

Bielefeld, den 17. Dezember 2020.


.....
(David Williams)