Machine Learning: Science and Technology

PERSPECTIVE • OPEN ACCESS

Machine learning and excited-state molecular dynamics

IOP Publishing Mach. Learn.: Sci. Technol. 1 (2020) 043001 https://doi.org/10.1088/2632-2153/ab9c3e

MACHINE LEARNING
Science and Technology

OPEN ACCESS

RECEIVED
20 May 2020

REVISED
4 June 2020

ACCEPTED FOR PUBLICATION
12 June 2020

PUBLISHED
17 September 2020

PERSPECTIVE

Machine learning and excited-state molecular dynamics

Julia Westermayr[1] **and Philipp Marquetand**[1,2,3]

[1] University of Vienna, Faculty of Chemistry, Institute of Theoretical Chemistry, Währinger Str. 17, 1090 Wien, Austria
[2] Vienna Research Platform on Accelerating Photoreaction Discovery, University of Vienna, Währinger Str. 17, 1090 Wien, Austria
[3] University of Vienna, Faculty of Chemistry, Data Science @ Uni Vienna, Währinger Str. 29, 1090 Wien, Austria

E-mail: philipp.marquetand@univie.ac.at

Keywords: machine learning, photodynamics, photochemistry, excited states, quantum chemistry, nonadiabatic couplings, spin-orbit couplings

Abstract

Machine learning is employed at an increasing rate in the research field of quantum chemistry. While the majority of approaches target the investigation of chemical systems in their electronic ground state, the inclusion of light into the processes leads to electronically excited states and gives rise to several new challenges. Here, we survey recent advances for excited-state dynamics based on machine learning. In doing so, we highlight successes, pitfalls, challenges and future avenues for machine learning approaches for light-induced molecular processes.

1. Introduction

Photosynthesis, photovoltaics, the processes that enable our vision or photodamage of biologically relevant molecules, such as DNA or peptides—they all have one thing in common: The underlying processes are governed by a manifold of excited states after the absorption of light [1–18]. They can be studied experimentally via several techniques, such as UV/visible spectroscopy, transient absorption spectroscopy, photoionization spectroscopy or ultrafast electron diffraction [19–27]. However, experimental techniques are to some extent blind to the exact electronic mechanism of photo-induced reactions. In order to get a more comprehensive understanding, theoretical simulations can complement experimental findings and can provide explanations for observed reactions [9]. For instance, simulated UV spectra can be used to unveil the states relevant for photodamage and -stability of molecules [28–36] and the temporal evolution of molecules can be studied via nonadiabatic molecular dynamics (NAMD) simulations [37–46]. The latter gives access to different reaction channels, branching ratios, and excited-state lifetimes and will be the main topic of discussion here.

While experimental techniques require large and costly setups, theoretical simulations require high-performance computing facilities due to expensive electronic structure computations. Especially NAMD simulations are seriously limited by the underlying quantum chemical calculations, making long and experimentally relevant simulation times inaccessible with conventional *ab initio* methods. The larger the molecule becomes, the more electronically excited states are involved in reactions and the more complex their interactions become. This leads to non-linearly increasing costs of quantum chemical calculations and a compromise between accuracy and computational efficiency is indispensable. Relying on such expensive *ab initio* potentials, only a couple of picoseconds can be simulated and the exploration of rare reaction channels is restricted due to bad statistics [17, 43, 47].

Technically, the nuclear part and the electronic part of the calculations can be separated to a large extent. First, the electronic problem is solved leading to potential energies for the nuclei. Afterwards, the nuclei move on these potentials classically or quantum chemically [6, 8, 48–52]. These two subsequent steps can be carried out in every time step (on-the-fly), if classical trajectories are employed. Alternatively, the two steps are separated as much as possible by precomputing the potential energy surfaces (PESs) and then using these precomputed PESs in the subsequent nuclear dynamics. Experimental observables and macroscopic properties can be obtained in follow-up computations or analysis runs. Machine learning (ML) can accelerate the overall simulation process on different levels and at several points. A broad classification of

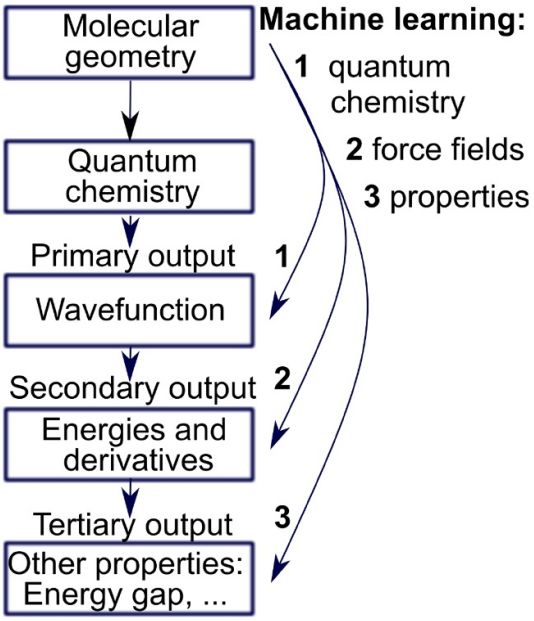

Figure 1. A broad classification of how to use machine learning models to replace different parts of quantum chemistry [53]. Simulations can be enhanced by providing (1) a wavefunction from a machine learning model, (2) a force field by fitting energies and forces or (3) other properties, such as energy gaps or reaction rates by learning the final output of a dynamics simulation directly.

how to use ML models to replace different parts of quantum chemistry to make simulations more efficient is given in figure 1 [53]. The probably most fundamental way is to use ML to solve the Schrödinger equation. This has been done for the ground state by representing the molecular wavefunction on a grid, in a molecular orbital basis or in a Monte-Carlo approach [54–62] and has recently also been applied for the excited states of a one-dimensional model [63]. ML can also be used to reconstruct the wavefunction from near-field spectra [64] or to bypass the Kohn-Sham equation in density functional theory (DFT) [65]. The external potential, functional, electronic density or local density of states can be learned [53, 65–72]. Very recently, Ceriotti and co-workers further introduced a smooth atomic density by defining an abstract chemical environment [73].

By having access to the molecular wavefunction or the electron density, the secondary output, which are energies and forces for the ground state and additionally couplings for the excited states, can be derived efficiently with ML. The coefficients of the ML wavefunctions or the density can further be used as an input for quantum chemical simulations, reducing the number of SCF iterations substantially [59].

Instead of learning the quantum chemistry of systems, the so-called "secondary outputs" [53] can also be mapped directly to a molecular structure, giving rise to so-called ML force fields. By training an ML model on *ab initio* data, the accuracy of quantum chemistry can be combined with the efficiency of conventional force fields for molecular dynamics (MD) simulations in the ground state [74–109]. For the excited states, only a couple of studies are available [106, 110–120, 120–122, 122–127, 127–129]. Nevertheless, the first NAMD simulation with ML dates back to the year 2008, where the scattering of O_2 on Al(III) was studied in a mixed quantum–classical approach considering singlet-triplet transitions [111, 130].

Having access to the excited state energies, "tertiary properties", such as UV spectra [131], band gaps [132–134], HOMO-LUMO gaps or vertical excitation energies [135–138] of molecules can be derived. Again, this tertiary output can also be fit in a direct fashion, which has been done for instance for a light-harvesting system by learning the excitation energy transfer properties [139] or the output of NAMD simulations to find out about the relations of molecular structures and dynamic properties [140]. Moreover, ML has been successfully applied for the inverse design of molecules and materials featuring specific properties, such as defined HOMO-LUMO gaps or catalytic activities. Examples range from the inverse design of photonic materials, to (photo-)catalysts, solar cells or (photo-active) drugs, to name only a few applications [107, 141–150].

Despite the opportunities of ML for the development of groundbreaking new methodologies, current techniques are often limited to certain molecules or specific problems. Methods exist, that extrapolate throughout chemical compound space, see e.g. references [59, 131, 151–160], but usually models fail to go beyond energies and related properties, such as forces, atomization or excitation energies. Further, it is

challenging to predict compounds consisting of atom types strongly different from those inside of the training set. Especially the fitting of the excited-state PESs poses another obstacle, let alone the transferability of excited-state PESs: Not only are ML models restricted to certain molecules or materials [106, 110–119, 119, 120, 120–122, 122–125, 127, 127–129], more often the different energetic states are fit independently from each other with separately trained ML models. As it is clear that the PESs of molecules are not independent of each other, it might also be unsurprising that learning them simultaneously is advantageous for various applications, such as NAMD or spectra predictions. Only a few studies exist that include more energetic states in one ML model and even less treat related properties, such as the vectorial dipole moments or couplings between different PESs in one ML model [85, 118, 131, 154, 161, 162].

However, in our view, the "holy grail" of ML for photochemistry is an ML model that provides all relevant energetic states, forces and properties at once, using derivatives where possible, rather than learning the properties independently. At the very best, this model should be transferable throughout chemical compound space [163] and could be used for molecules of any composition. Given the fact that ML models for the electron wavefunctions of different excited states (or related models within the DFT framework) do not exist for polyatomic systems, this dream has not yet come alive. Hence, we will focus this perspective on ML models that learn the secondary outputs, i.e. excited state PESs, corresponding forces, and nonadiabatic and spin-orbit couplings (NACs and SOCs, respectively) between them. We note that our discussion holds for different spin-multiplicities, although most studies focus on singlet states only. We try to address the recent achievements in the fields of photochemistry using ML and discuss the current challenges and future perspectives to get a step further to a transferable ML model for excited states that treats all properties on the same footing.

We start by discussing the generation of a training set for the treatment of excited state PESs, corresponding forces and couplings and focus on their use in NAMD simulations. Especially, we aim to clarify the differences between excited-state and ground-state properties. We therefore describe the NACs and SOCs that couple different electronic states and highlight their importance for NAMD simulations. Subsequently, state-of-the-art ML models for excited-state PESs are considered along with the challenge of modelling a manifold of energetic states.

2. Generating a training set for excited states

The basis of any successful ML model is a comprehensive and accurate training set that contains the molecular geometries in combination with the corresponding properties that need to be predicted. For the application of ML within NAMD simulations, the training set should contain a molecular geometry and the energies of all energetic states, corresponding forces, and couplings between the states. It is computed with the quantum chemistry method, whose accuracy one wants to obtain. The choice of the quantum chemistry method is a problem on its own and often requires expert knowledge [18, 164]. Simply said, ML can be seen as efficient interpolation between data points with the accuracy of the reference method.

Before we go into detail on how to efficiently create a training set for excited states, we will first discuss the differences to ground state potentials and properties that need to be considered. A major drawback is the fact that a manifold of excited states and thus also the properties between them have to be accounted for. These are NACs between states of same spin multiplicity and SOCs between states of different spin multiplicity as well as transition dipole moments. The fitting of such properties is problematic due to the arbitrary phase of the wavefunction [118, 161]. Therefore, an additional pre-processing might be necessary. Either a diabatization [106, 115–117, 119, 120, 120–122, 122–127, 127], a so-called phase correction [118], or a special learning algorithm [165] renders data learnable. The latter two are described in the following while further details on the former are given in section 3.2.

2.1. Making excited-state data learnable

Compared to energies and forces, NACs, SOCs as well as transition dipole moments result from the wavefunctions of two different electronic states. Due to the non-unique definition of the wavefunction itself, i.e. the fact that multiplication of the electronic wavefunction with a phase factor still gives a valid eigenfunction of the electronic Hamiltonian, leads to an arbitrary phase, which is initiated randomly in a quantum chemical calculation. Consequently, also the sign of the couplings, C_{ij}, can be positive or negative. Here, i and j denote the indices of the involved states. The resulting inconsistencies in the coupling hypersurfaces make it challenging to find a good relation between an ML model, which is per definition a smooth function [166], and those discontinuous raw outputs.

This problem can be illustrated with molecular orbitals of the methylenimmonium cation (figure 2 reproduced from reference [118]). Panel A shows the molecular geometries, which are given as an input to a quantum chemical program. Two molecular orbitals are shown as placeholders for the wavefunctions of two

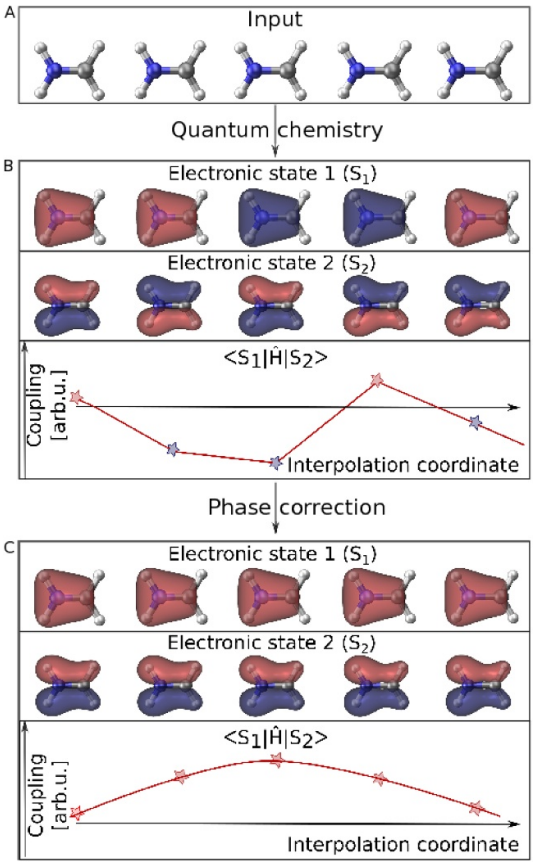

Figure 2. A set of quantum chemical calculations of the methylenimmonium cation, $CH_2NH_2^+$, along the C-N bond. The molecular conformations, which are the input of a quantum chemical calculation, are given in panel A, the orbitals of the computed two electronic states, the S_1 and S_2 state, as well as the corresponding off-diagonal matrix elements between those two states, $\langle S_1 \mid \hat{H} \mid S_2 \rangle$, are given in panel B. As can be seen, the sign of those values arbitrarily switches. Those sign jumps can be removed by applying a phase correction algorithm. Results are given for those elements in panel C. Reproduced from reference [118] under CC-BY, https://creativecommons.org/licenses/by/3.0/.

electronic states, the S_1 and S_2 states. The color of the orbitals can be either blue or red and changes arbitrarily throughout the reaction coordinate. In the same way, also the overall wavefunction (which is difficult to plot) for the respective state changes its phase arbitrarily. As a consequence, also the sign of the couplings, where the product of the two wavefunctions' signs enters, may change randomly (see panel B).

In order to allow for a meaningful ML description of elements resulting from two different electronic states, a data pre-processing is helpful. The former process is termed phase correction [118, 167] and is practicable to remove almost all such inconsistencies in the configurational space of the training set. This phase correction makes the use of conventional training algorithms possible.

To carry out phase correction, a wavefunction overlap computation [168], $\boldsymbol{S} = \langle \Psi_\alpha \mid \Psi_\beta \rangle$, has to be carried out between the wavefunction Ψ_β at every geometry β inside the training set and the wavefunction Ψ_α at a reference geometry α. The phase thus has to be tracked from a pre-defined reference geometry. It often happens that two geometries are dissimilar to each other, so that interpolation between them is necessary, making this process generally more expensive. So-called intruder states give rise to additional problems, since they are so high in energy at the reference geometry that they usually would not be included in the initial calculation. However, they enter the lower energy region at another geometry visited during the NAMD simulations and thus need to be considered in the current calculation step. Hence, they should have been included from the beginning for the phase correction algorithm to work. As a solution to this problem, many electronic states need to be computed from the start. In some cases, where many energetic states lie close to each other and where the photochemistry is complex, phase correction might even be infeasible. The problem of intruder states was also identified by Robertsson *et al* and is well explained for a diabatization procedure in reference [169]. For a more detailed discussion on phase correction, the reader is referred to references [118, 167, 168, 170]. Nevertheless, as given in panel C of figure 2, smooth curves are obtained if phase correction is carried out correctly and these phase-corrected properties can be learned with conventional ML models.

Similarly, a small set of data can be corrected manually and afterwards a cluster growing algorithm can be applied [125, 171]. This algorithm uses Gaussian process regression to continuously add data points to the initially phase-corrected data set. This approach has been employed recently to obtain diabatic transition dipole moments [119]. However, in systems containing many degrees of freedom and many electronic states, a manual correction of the sign of couplings is tedious and the approach has only been applied to small systems, yet [119, 126].

In contrast to the phase-correction procedures described above, an ML-based internal phase-correction during training renders the learning of raw quantum chemical data possible and does not require any pre-processing. However, it requires a modification of the training process itself [165]. In a recent study, we applied such a phase-free training using the deep continuous-filter convolutional-layer neural network (NN) SchNet [156, 172] that we adapted for the treatment of excited states. In contrast to conventional algorithms, where the hyperparameters of the network are optimized to minimize the L_1 or L_2 loss function, here a phase-less loss function is applied. The latter allows the ML model to possess a different phase (or sign) for the learned property than the reference data. Since ML models intrinsically yield smooth curves, the algorithm will then automatically choose a phase for every data point such that smooth coupling curves are produced. This freedom of choice is achieved by calculating the errors between the ML value and all possible phase variations of the reference value and using only the smallest of these errors. The possibilities for phase conventions scales with 2^{N_S-1}, where N_S is the number of considered states. Since the error is computed more often than in conventional ML training, the phase-less loss training becomes more expensive, when including more electronic states. Mathematically, instead of computing the mean squared error, ε_{L_2}, between reference couplings, C_{ij}^{QC}, and predicted couplings, C_{ij}^{ML},

$$\varepsilon_{L_2} = \|C_{ij}^{QC} - C_{ij}^{ML}\|^2, \tag{1}$$

the phase-free error, ε_{ph}, is computed as the minimum of 2^{N_S-1} computed errors:

$$\varepsilon_{ph} = \|C_{ij}^{QC} \cdot p_i^k \cdot p_j^k - C_{ij}^{ML}\|^2 \text{ with } 0 \leq k \leq 2^{N_S-1}. \tag{2}$$

p_i^k and p_j^k are phase factors, giving rise to the sign of state i and j resulting in a phase vector with index k. This adaption of the loss function can remove the influence of any phase during the training process, making the use of raw data possible. A variation of this approach with reduced cost is possible if only one property, i.e. NACs or SOCs, are trained for NAMD simulations. A detailed discussion can be found in reference [165].

It is worth mentioning that, besides the arbitrary phase of the wavefunction, also the Berry phase (or geometric phase) [173] exists. Effects due to the Berry phase can not be accounted for with phase correction. Nevertheless, most often in mixed quantum classical NAMD simulations, the Berry phase can be neglected. As a drawback, some effects, such as interference of nuclear wavefunctions, might not be described correctly with such methods, and thus prevents the application of the corresponding ML properties if those effects are important. In some other dynamics methods and reactions, the Berry phase plays a crucial role and can lead to path-dependent transition probabilities close to conical intersections. This effect is important in quantum dynamics simulations, where problems can be circumvented by using diabatic potentials, which will be described in section 3.2.

2.2. Choosing the right reference method

While many ML potentials for ground-state MD simulations are based on DFT training sets, see e.g. references [77, 79, 86, 174, 175], the training sets for the excited states are mainly obtained with multi-reference methods. Examples are the complete active space self-consistent-field (CASSCF) method [112–114, 120, 122, 127, 140, 165] or multi-reference configuration interaction (MR-CI) schemes [115, 117–119, 126, 161, 165, 176–182]. The advantage of multi-reference methods compared to single reference methods is that photo-dissociation, which is likely to occur in many molecules after their excitation by light, can be treated accurately. In contrast, single-reference methods fail to do so in many cases. However, multi-reference methods are seriously limited by their computational costs [7, 10], calling for an efficient and meaningful training set generation. Therefore, the training set should be as small as possible, but should cover the relevant conformational space of a molecule that is required for accurate NAMD simulations [183]. Accordingly, many recent training sets for MD simulations are built by a so-called "active-learning" [77, 184] or iterative/adaptive sampling scheme [85, 185] that will be described in the following and can be adapted for excited states [118]. From our point of view, it is most favorable to start by computing a small initial training set and to expand it via such an adaptive sampling scheme [77, 85, 118, 184–186].

2.3. Initial training set

If only static calculations are targeted, data bases can be generated efficiently by starting from already existing data sets. As an example, Schwilk *et al* [182] constructed a large data set of 13k carbene structures by randomly choosing 4,000 geometries from the QM9 [187] data set (consisting of 130k organic molecular structures). Hydrogen-atoms were abstracted and singlet and triplet states were optimized. The MR-CI method was subsequently used to compute the energies of the singlet and triplet state and a data set of 13,000 different carbene structures, called QMspin, was obtained, opening avenues to investigate important intermediate geometries critical for organic reaction networks.

As a starting point for all following training-set generation schemes aiming to investigate the temporal evolution of a system, the equilibrium geometry of a molecule can be computed and taken as a reference. The initial training set can then be built up by sampling conformations close to this molecular configuration. In general, every sampling method is possible. Since the normal modes of a molecule are generally important for dynamics, scans along these coordinates can be used to sample different conformations. In cases of small molecules with few degrees of freedom, this process might be a good starting guess for an initial training set [118]. It also makes sense to optimize critical points like excited-state minima, conical intersections and state crossings and to include data along such optimization runs into the training set. The same is advisable for larger systems, but in addition, some other approaches like Wigner sampling [188] or sampling via MD simulations [189, 190] can be considered. To name a few approaches, umbrella sampling [191], trajectory-guided sampling [192], enhanced sampling [193] or metadynamics [194], using a cheap electronic structure method like the semi-empirical tight-binding based quantum chemistry method GFN2-xTB [195], can be employed.

Further, if literature or chemical intuition indicate that certain reactions, like dissociation, take place after photo-excitation, it is also favorable to include those reaction coordinates right from the beginning. The initial training set can easily comprise on the order of 1,000 data points, which might seem like a lot but is reasonable given the large number of data points in commonly used training sets [112–114, 161]. The quality of the initial ML potentials can be assessed by carrying out short scans along different reaction coordinates, such as combinations of normal modes. As soon as the initial training set is large enough, the training set expansion via an adaptive sampling scheme can be started.

2.4. Adaptive sampling for excited-states

ML models fail to predict regions with scarce training data, i.e., their extrapolation capabilities are faint [196]. Since such regions are likely visited during a dynamics simulation, the initial training set then needs to be expanded. A quality control is needed to detect whether unknown conformational regions of the molecule are visited, such that the corresponding structures afterwards can be added to the training set.

This concept was introduced already in 1992 as query by committee [197] and has been used in chemistry in the so-called GROW algorithm of Collins and coworkers [128, 129] as well as in the iterative sampling of Behler [185]. The latter is nowadays well adapted for the ground state [77, 85, 184, 186] and was recently modified for the excited states [118]. The scheme is described in more detail in the following.

To apply the procedure of adaptive sampling, at least two ML models have to be trained independently, e.g. with slightly different hyperparameters or starting weights. An overview of this procedure with two neural networks (NNs) is given as an example in figure 3. At every time step during an MD simulation, the predictions Y_M^p of at least two ML models, M, for a property p (e.g. a potential energy) are compared to each other. To this end, the standard deviation of these predictions with respect to the mean of each property, $\overline{Y}^p$, is computed according to

$$Y_\sigma^p = \sqrt{\frac{1}{M-1}\sum_{m=1}^{M}(Y_m^p - \overline{Y}^p)^2}. \tag{3}$$

This standard deviation is compared to a pre-defined threshold, ε^p, for each trained property. If the standard deviation stays below the threshold, the mean of each property, $\overline{Y}^p$, is forwarded to the MD program to propagate the nuclei. If the threshold is exceeded, the ML prediction is assumed to stem from an undersampled or unknown region of the PESs and is deemed untrustworthy. This conformation has to be included into the training set to guarantee accurate ML PESs. Thus, a quantum chemical reference computation is carried out, the data point is added to the training set, and the ML models are re-trained to execute ML-NAMD simulations on longer time scales. It is sensible to choose a large threshold, ε^p, in the beginning and adaptively make it smaller as the robustness of the ML models increases, giving rise to the name adaptive sampling [85].

Adaptive sampling for excited states differs from adaptive sampling in the ground state in the number of properties that are considered. As illustrated in figure 3, not only the energies must be accurately predicted,

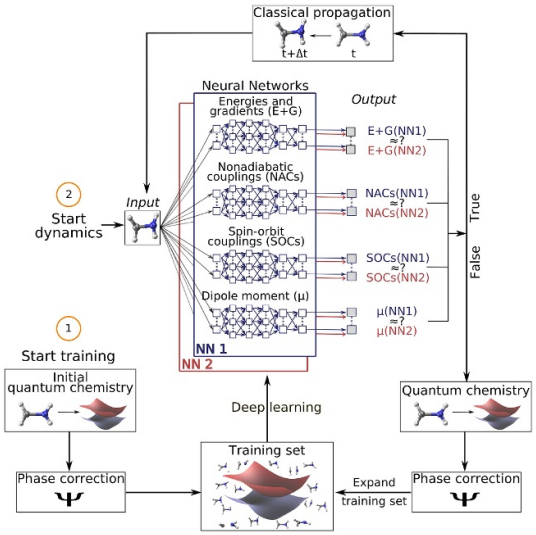

Figure 3. Overview of the adaptive sampling scheme for excited states, reproduced from reference [118] under CC-BY. Adaptive sampling is illustrated exemplarily with two NNs for the methylenimmonium cation, $CH_2NH_2^+$. As a starting point, ML models are trained on the initial training set and ML-based dynamics are executed. At each time step, the predictions of the two deep NNs (NN1 and NN2) are compared to each other for energies (E) and gradients (G), nonadiabatic couplings (NACs), spin-orbit couplings (SOCs), and dipole moments (μ). If the difference between the ML models overcomes a pre-defined, adaptive threshold, the geometry visited at this time step is re-computed with quantum chemistry, added to the training set after phase correction and the ML models are re-trained. Subsequently, a new dynamics cycle is started and this process is repeated until the ML models are deemed to be converged.

but also the couplings and, if necessary, dipole moments. Since more states are considered, an average standard deviation is taken as the mean of the standard deviations of each state in case of energies and gradients,

$$Y_\sigma^p = \frac{1}{N_S}\sum_i^{N_S}\left(\sqrt{\frac{1}{M-1}\sum_{m=1}^{M}(Y_m^{p_i}-\overline{Y}^{p_i})^2}\right), \tag{4}$$

and as the mean of the standard deviations of each pair of states for couplings or dipole moments,

$$Y_\sigma^p = \frac{1}{2N_S^2}\sum_i^{N_S}\sum_j^{N_S}\left(\sqrt{\frac{1}{M-1}\sum_{m=1}^{M}(Y_m^{p_{ij}}-\overline{Y}^{p_{ij}})^2}\right). \tag{5}$$

A separate threshold is set for each of these averaged quantities. If any of the quantities is predicted inaccurately, the data point is recomputed with quantum chemistry, phase corrected and added to the training set. In order to make this process more efficient, not only one MD simulation, but an ensemble of trajectories can be computed. The ML models are only retrained after each of the independent trajectories has reached an untrustworthy region of the PES and after each of the reference calculations has been finished and included in the training set. This makes the parallelization of many trajectories possible [85, 118, 185, 198].

The adaptive sampling scheme should be carried out until the relevant conformational space for photodynamics is sampled sufficiently. However, using more than one ML model for production runs is still favorable. One of us and coworkers observed that the error of predictions decreases with the number of ML models used [85, 198]. We have seen the same trend in a recent study for NAMD simulations. With the adaptive sampling scheme for excited states, we generated a training set of 4,000 data points of the methylenimmonium cation, $CH_2NH_2^+$, to carry out long time-scale NAMD simulations with NNs [118].

2.5. Additional sampling techniques for excited states

Further training sets for NAMD simulations were generated for one-dimensional systems as well as polyatomic molecules. For example, Chen *et al* have computed 90,000 data points via Born-Oppenheimer MD simulations and NAMD simulations, where emphasis was placed on the inclusion of geometries after a transition from one state to another took place [114]. Deep NNs were trained on energies and gradients of this training set to accurately reproduce NAMD simulations.

Hu *et al* [112] used a very similar approach and obtained around 200,000 data points of 6-aminopyrimidine from Born-Oppenheimer MD simulations. They further carried out NAMD simulations

with surface hopping [199, 200], where transitions from one state to another were allowed via so-called hops. The geometries that were visited shortly before a hop took place were used as a starting guess to optimize conical intersections, i.e. critical points of the PESs, where two states become degenerate. Those data points were included to comprehensively sample the regions around a conical intersection. However, the ML models were not accurate enough for NAMD simulations solely based on ML potentials and the authors had to resort to quantum chemistry calculations in critical regions of the PESs.

Dral *et al* [113] generated a training set for a one-dimensional two-state spin-boson model consisting of 10,000 data points with a grid-based method. The training data selection was then based on the structure, rather than on the energy of the molecules. For each data point, a molecular descriptor was computed and the distances of the descriptors were compared. Data points for the training set were chosen to sample the relevant space sufficiently [113, 201]. Compared to random sampling, this method allowed a reduction of training set sizes up to 90 %, which was shown for static calculations of the methyl chloride molecule [201, 202]. A similar structure-based sampling scheme was proposed by Ceriotti *et al* [203].

Additionally, a maximum and minimum value can be computed for each representation of a molecule inside the training set. Every new structure that is obtained throughout an MD run can be compared to those values to get a measure of reliability of ML predictions. If the configuration does not lie within the known region, it can be added to the training set [184, 204]. Very recently, an active learning approach has been proposed to construct PESs without the need of running MD trajectories. The difference between two NN potentials was computed and points were iteratively added at the maxima of this difference surface (or, as phrased in the study, at the minima of the negative difference) [205].

It becomes evident from the diversity of approaches and training set sizes that a general guide on how to compute the training set and how large it should be for NAMD simulations can not be given. It is rather a matter of the efficiency that should be achieved and the computational costs that can be justified. Further, the training set strongly depends on the molecule under investigation. Especially its size, flexibility and the complexity of the light-induced dynamics play an important role. A molecule, whose photodynamics can be described as a two-state problem, such as in a simplified case of ethylene [206, 207], or a molecule, which is rigid, where dynamics mostly lead towards one reaction channel, possibly requires less data points than molecules that exhibit several different reaction channels after photo-excitation.

3. Machine learning nonadiabatic molecular dynamics simulations (ML NAMD)

3.1. Beyond Born-Oppenheimer dynamics

With an accurate training set for excited states at hand, NAMD simulations can be enhanced with ML models in order to enable the dynamics on time scales otherwise unfeasible. The most accurate way to study the dynamics of a molecule would be the full quantum mechanical treatment, which is, however, expensive and limited to a few atoms, even if ML PESs are applied [49, 115–117, 119–125, 127, 127, 208–211]. A mixed quantum classical treatment is thus often preferred, where the motion of the nuclei are treated classically on one of the ML PESs. The mixed quantum classical MD simulation can then be interpreted as a mixed MLMD simulation [165]. The Born-Oppenheimer approximation allows to separate the nuclear from the electronic degrees of freedom. However, this approximation is not valid in the vicinity of avoided state crossings of PESs (or conical intersections, as mentioned before), which play an important role in excited-state dynamics.

In these critical regions of the PESs, ultrafast rearrangement of the motions of the electrons and the nuclei takes place due to strong couplings. As already mentioned, the relevant coupling elements are NACs and SOCs. The NACs (denoted as C^{NAC}) are vectorial properties and can be computed as [48, 212, 213]

$$\begin{aligned} C_{ij}^{\mathrm{NAC}} &\approx \langle \Psi_i \mid \tfrac{\partial}{\partial \mathbf{R}} \Psi_j \rangle \\ &= \tfrac{1}{E_i - E_j} \langle \Psi_i \mid \tfrac{\partial H_{el}}{\partial \mathbf{R}} \mid \Psi_j \rangle \quad \text{for } i \neq j, \end{aligned} \tag{6}$$

neglecting the second order derivatives. Thus, in the vicinity of a conical intersection, the couplings become very large, whereas they are almost vanishing elsewhere. The singularities that arise when two states are degenerate do not only pose an obstacle to quantum chemistry, but consequently also to PESs fitted with ML [112–114, 118]. NACs are nevertheless important properties to determine the direction and probability of internal conversion—a transition from one state to another, where the spin multiplicity does not change [8, 48, 51, 52, 214]. In contrast, the SOCs (denoted as C^{SOC}) are complex-valued properties that determine the rate of intersystem crossing, i.e., the transitions from one state to another, where spin multiplicity does change. In standard quantum chemistry programs, SOCs are given as the off-diagonal elements of the Hamiltonian matrix [8, 52, 215]:

$$C_{ij}^{\mathrm{SOC}} = \langle \Psi_i \mid \hat{H}^{SOC} \mid \Psi_j \rangle. \tag{7}$$

3.2. Fitting diabatic potentials

The numerical difficulties that arise due to discontinuous PESs and singularities of couplings at conical intersections can be circumvented by the use of diabatic potentials instead of adiabatic ones [116, 216–220]. In the diabatic basis, the coupling elements are smooth properties and the arbitrary phase of the wavefunction does not have an impact. This favors the use of diabiatic PESs. Since the output of a quantum chemistry program is generally given in the adiabatic basis, a quasi-diabatization procedure is necessary. Strictly speaking, a diabatization procedure is not possible because e.g. an infinite number of states is needed for an accurate representation. If using a finite number of states, the term quasi-diabatic is employed. For simplicity, we still use the notation of diabatic potentials for quasi-diabatic potentials. Those have been generated with different methods [221] and for small molecules up to date. Examples are the propagation diabatization [222], diabatization by localization [223] or by ansatz [115, 224], diabatization based on couplings or other properties [225–228], configuration uniformity [229], block-diagonalization [230], CI vectors [169] or (partly) on ML [115, 116, 224].

Since several years, (modified) Shepard interpolation is used to fit diabatic potentials [129, 235–238] and also least squares fitting was applied to study the photo dissociation of molecules, such as NH_3 and phenol [239, 240]. In a series, Guo, Yarkony and co-workers developed invariant polynomial NNs [116, 124, 231–234] to address the excited-state dynamics of NH_3 and H_2O by fitting potential energy matrix elements. Absorption spectra as well as branching ratios could be obtained with high accuracy. The same authors further fit the diabatic $1,2^1A$ dipole moment surfaces of NH_3, which can only be fitted accurately if the topography of the PESs is reproduced correctly, validating the previously fitted diabatic potentials [119].

Habershon, Richings and co-workers used Gaussian process regression (in their notation equal to kernel-ridge regression) to fit diabatic potentials to execute on-the-fly dynamics of the butatrien cation with variational Gaussian wavepackets [127]. In another study, they applied an on-the-fly MCTDH scheme (DD-MCTDH) and carried out 4-mode/2-state simulations of pyrazine [120]. By improving the ML approach with a systematic tensor decomposition of kernel ridge regression, the study of 12-mode/2-state dynamics of pyrazine was rendered possible. This achievement remains a huge improvement over current MCTDH simulations in terms of accuracy and efficiency [241].

For the improvement of the diabatization by ansatz procedure, Williams *et al* [115] applied NNs and enabled the fitting of the electronic low lying states of NO_3. The improvement of the diabatization procedure itself is desirable [116, 117], since the generation of meaningful diabatic potentials is often a tedious task and restricts their use tremendously. Up to date, no rigorous diabatization procedure exists that allows the diabatization of adiabatic potentials of polyatomic systems by non-experts in this field [116, 119]. Especially challenging for larger and more complex systems is the number of electronic states within a certain energy range that have to be considered for successful diabatization. An increasing computational effort to provide all relevant electronic states is the result, making diabatization further challenging [169]. Often, more extensive approximations [115, 169, 242], e.g. the linear vibronic coupling model [243] are applied. We refer the reader to references [50, 217, 243–245] for more details on such approaches.

Despite the advantages of diabatic potentials, due to the before-mentioned drawbacks and the fact that the direct output of a quantum chemical calculation is given in the adiabatic basis, on-the-fly NAMD in the adiabatic representation is often the method of choice for large polyatomic systems, which will be discussed in the following.

3.3. Fitting adiabatic potentials

In order to execute NAMD simulations in the adiabatic basis, approximations have to be introduced to account for nonadiabatic transitions between different PESs. A good trade-off between accuracy and efficiency can be achieved with the surface hopping methodology [199, 200], which is often applied in ML-based NAMD simulations [112–114, 118]. In surface hopping, the transitions, or so-called hops, between different states, are computed stochastically and a manifold of trajectories needs to be taken into account to analyze different reaction channels and branching ratios [8, 51, 246] Several algorithms [52, 246–248] are frequently used to compute the hopping probability as well as its direction, with Tully's fewest switching algorithm being among the most popular ones [199, 200]. There, the couplings between adjacent states determine the hopping probability [51]. Other frequently applied algorithms to compute hopping probabilities are the Landau-Zener [249, 250] and the Zhu-Nakamura approximations [247, 248, 251, 252]. Those approximations solely rely on the PESs and omit the computation of wavefunction coefficients and couplings. Other flavors to account for transitions exist, which have, however, not been used in ML based NAMD studies yet. We thus refer the reader to references [48, 51, 52, 170, 199, 246, 251–256] for further information.

3.3.1. NAMD simulations with ML energies and forces

Based on the fewest switches algorithm, one of the first ML NAMD simulation is carried out by Carbogno *et al* [130], where the scattering of O_2 at Al(III) is studied [111, 130, 257]. A set of 3768 carefully selected data points [258] allowed for interpolation of the PESs with NNs [257]. In a first attempt, the authors include a spin-unpolarized singlet PES and a spin-polarized triplet state. Strictly speaking, the output of a quantum chemistry simulation for a singlet and triplet state is spin-diabatic [52, 215] and NAMD simulations ideally should carry out a diagonalization to obtain the spin-adiabatic PESs [52, 215]. The authors took advantage of the adiabatic spin-polarized PES [259] to compute the absolute value of couplings [130]. In this way, the transitions between the states could be approximated using surface hopping omitting the computation of wavefunction dynamics. This study was extended by two-state NAMD simulations of different multiple PESs arising from different spin configurations. Findings suggested a high probability of singlet-to-triplet conversion during scattering experiments with a non-zero probability even at low coupling values [111, 130, 257].

Other studies using the Zhu-Nakamura method [247, 248, 251, 252] to account for nonadiabatic transitions are discussed below. This approximation is based solely on energies and neglects the phase of the wavefunction. As a drawback, PESs are always assumed to couple to each other, when they are close in energy. This holds true for many cases, but one must be aware that strongly and weakly coupled PESs can not be distinguished.

Hu *et al* [112] for example trained separate kernel ridge regression models to fit three singlet states of 6-aminopyrimidine. For learning, they used 65,316 data points comprising the molecular structures and energies of 6-aminopyrimidine with gradients not fitted, but computed afterwards. The data points were obtained from Born-Oppenheimer simulations, which were further clustered into sub-groups, from which the training points were selected randomly. As mentioned before, hopping geometries obtained from reference NAMD simulations were taken to find minimum conical intersections and the latter were also included in the training set. In contrast, Chen *et al* [114] trained two deep NNs on 90,000 data points of two singlet states of CH_2NH. Again, data points were obtained from Born-Oppenheimer MD simulations and NAMD simulations starting from hopping geometries. In both studies, the NAMD simulations of the reference method could be successfully reproduced.

Instead of approximating the hopping probability, the NACs can also be approximated from PESs, gradients and Hessians [208, 260–264]. We made use of this relation and the fact that ML Hessians can be computed efficiently, and carried out NAMD simulations with the surface hopping method for sulfur-dioxide, thioformaldehyde and the methylenimmonium cation [165].

3.3.2. NAMD simulations with ML energies, forces, and couplings

In addition to energies and forces, SOCs need to be fitted with ML models when states of different spin multiplicities become relevant. Furthermore, when approximative schemes for the computation of hopping probabilities fail, the ML models need to learn NACs. One of the first studies, where NACs were fitted, used 1,000 and 10,000 data points to train kernel-ridge regression models to reproduce NAMD simulations of a one-dimensional system. However, especially in critical regions of the PESs, the ML models could not replace quantum chemical calculations and so 13-16% electronic structure calculations were required during an NAMD simulation [113]. The authors highlighted this as a drawback, because efficient simulations should be performed purely with ML and should not rely on intermediate quantum chemical calculations. Moreover, each entry of the NAC vectors was fitted by a separate kernel-ridge regression model, which turned out to be insufficiently accurate.

As indicated before, we also aimed for reproducing NAMD simulations with ML. We employed multi-layer feed-forward NNs trained on 4,000 data points of 3 singlet states of $CH_2NH_2^+$ [118]. Short reference NAMD simulations based on electronic structure calculations could be reproduced. With the ML NAMD, long simulation times on the order of a nanosecond were successfully reached. Significantly different from previous ML NAMD approaches is the smaller size of the training set required to reproduce NAMD simulations. Further, a multi-output ML model was used to fit all NAC vectors between different states of same spin multiplicity at once. We term such models multi-state models. Per definition, kernel-ridge regression, and similar approaches such as linear regression, are single-state models. In order to make multi-state predictions of such models possible, the energetic state has to be encoded explicitly by using for example an additional state kernel. This procedure enables to model several states simultaneously. We studied the use of multi-state descriptors with the QML toolkit [265] for kernel-ridge regression models and showed that a multi-state description is generally superior to a single-state description in terms of accuracy [161].

Lastly, we want to comment on the NACs as vectorial properties. It should be clarified that approaches relating a molecular input directly to NAC values do not provide rotational covariance. This drawback is independent of a single-state treatment, i.e., the use of a separate ML model for each coupling value, or a

multi-state treatment, where all values are represented in one ML model. Very recently, Zhang *et al* applied a symmetry-adapted high-dimensional NN [266] and treated the couplings as derivatives of NN representations. In this case, electronic friction was modelled via ML and applied for MD simulations of molecules at metal surfaces to treat the electron-nuclei coupling in a rotationally covariant manner. For the NAC vectors, we applied a similar strategy (similar also to force-only training for potentials), and implemented them as derivatives of virtual properties (i.e. non-existent in quantum chemistry) built by a deep NN [165].

3.4. Choosing the right descriptor

Many of the aforementioned studies use kernel ridge regression models or NNs in combination with distance-based descriptors [99, 112–114, 118, 267] such as the matrix of inverse distances or the Coulomb matrix [76]. It is worth mentioning that the accuracy of the ML PESs also depends on the type of descriptor. Molecular descriptors that represent atoms in their chemical and structural environment are often superior to those who treat complete molecules [101, 154, 155]. The symmetry functions of Behler [268, 269], their weighted counterparts [175, 270] or the FCHL (Faber-Christensen-Huang-Lilienfeld) representation [101, 155] work very well for NNs and the latter also for kernel-ridge regression and additionally provide permutation invariance.

Further improvement can be provided by message passing NNs [271]. Compared to hand-crafted molecular descriptors, the representation of molecules can be seen as a part of a deep NN and, thus, is generated automatically. For each training set, an accurate descriptor is intrinsically designed, which accounts for the chemical and structural environment of a molecule. Examples for such networks are SchNet [156, 172], the DTNN [133], PhysNet [272], or HIP-NN [273]. For the excited states, the SchNarc [165] approach offers this type of descriptor.

4. Conclusion and outlook

To conclude, ML methods are very powerful and can be used to speed up current MD approaches for the excited states. They have been successfully applied to circumvent existing problems due to the expenses of the underlying electronic structure methods, but there is still a long way ahead of us to make ML applicable for the photodynamics simulations of large and complex systems.

While the fitting of diabatic potentials is generally more favorable, those methods are limited by the challenges that arise in finding meaningful diabatic potentials. Up to date, diabatization procedures are tedious and often not feasible for large and complex systems. ML models have been successfully applied to improve these processes [115–117], but methods to treat large and complex polyatomic systems in the diabatic basis are still lacking.

To investigate the photodynamics of polyatomic molecules, mixed quantum–classical MD simulations in the adiabatic basis thus often remain the method of choice. One advantage is that the direct output of a quantum chemical calculation is given in the adiabatic basis and so the obtained potential energies and forces can be directly fitted with an ML model. By applying approximations for the computation of transition probabilities from one state to another, the photodynamics can be studied efficiently with ML [112, 114, 165]. When approaches aim for additionally fitting the coupling values between different electronic states, inconsistencies in the data need to be considered carefully. Those have to be either removed from the training set or the training process itself has to be adapted in order to achieve successful training. Both approaches have been applied and were used for NAMD simulations [118, 165].

Unfortunately, also mixed quantum–classical MD simulations in the adiabatic basis have been restricted to small, isolated systems or single reactions, when described with ML PESs. At the current stage of research, many challenges remain that need to be tackled when replacing quantum chemical calculations in photodynamics simulations of large and complex polyatomic molecules. The most severe issue is to identify an adequate quantum chemical reference method. Up to date, mainly multi-reference methods have been used to compute training sets for small molecular systems for the excited states. However, the computational effort increases rapidly with the number of atoms and electronic excited states considered making single-reference methods more viable in this regard. The drawback of the latter methods is that their PESs can be even qualitatively wrong in some conformational regions. A tedious exclusion of such regions could in principle allow for a comprehensive training set generation and the accurate fitting of the manifold of energetic states, forces, and couplings between them in the remaining regions. Nonetheless, the resulting ML models might still be inappropriate for dynamics simulations due to the validity in only a restricted conformational space. Furthermore, the PESs from common single-reference methods or approximate methods like Time-Dependent DFT based Tight Binding are likely to be at least quantitatively incorrect. Since even quantitatively small errors in the potential energy can accumulate to completely wrong dynamics

over millions of time steps, the validity of ML models based on such approximate methods for long time scale dynamics is questionable. It is unclear whether such quantitatively wrong potentials can still lead to qualitatively correct trends for the temporal evolution of a system in some cases. Studies comparing such long photodynamics simulations with experimental observables could provide an answer to this question but are not yet available.

The large number of data points that have been required with many recent ML approaches for small polyatomic molecules is particularly concerning in the view of treating larger systems. Here, intelligent algorithms for efficient training data set generation with the smallest possible number of points are required. In addition, energies, forces, and couplings were often trained in separate ML models, leading to unsatisfactory accuracy. The development of an ML model that can treat all properties for photodynamics simulations at once and enables a derivation of relevant properties from ML PESs is clearly desirable. The latter can potentially reduce the amount of required training data to precisely fit quantum chemical properties and further allows to omit their quantum chemical computation during dynamics. The ultimate solution would be an ML multi-reference wave function of a molecular system, but this envisioned dream has not yet been realized.

To enable a full exploitation of the advantages of ML, an ML model should optimally be applicable throughout chemical compound space. While current studies struggle with fitting the excited states of one molecule, the transferability of ML potentials for the excited states is far from being achieved. The description of the photochemistry of a biologically more relevant system, such as a DNA strand or peptide chain, from local contributions of single building blocks, i.e., DNA bases or single amino acids, is one of the biggest benefits an ML model can offer. However, the construction of excited-state PESs from local atomic contributions has not yet been shown to hold and a description of a whole molecular system limits the applicability of the ML model to other systems. At the current stage of research, the small molecules that have been successfully described in their excited states with atom-wise descriptors are not large enough to prove the validity of locally constructed excited state PESs. An estimation of the excited-state locality could also pave the way towards excited-state ML/MM simulations, similar to QM/MM (quantum mechanics/molecular mechanics). Further investigations along the aforementioned avenues are needed to gain insights into the possibilities of ML models to describe the excited states of larger and more complex systems and to fit more than one small, isolated molecule.

Acknowledgments

This work was financially supported by the Austrian Science Fund, W 1232 (MolTag) and the uni:docs program of the University of Vienna (J W). P M thanks the University of Vienna for continuous support, also in the frame of the research platform ViRAPID.

ORCID iDs

Julia Westermayr https://orcid.org/0000-0002-6531-0742
Philipp Marquetand https://orcid.org/0000-0002-8711-1533

References

[1] Cohen B, Crespo-Hernández C E, Hare P M, and Kohler B 2004 *Ultrafast Excited-State Dynamics in DNA and RNA Polymers* (Amsterdam: Elsevier) pp 463–70
[2] Levine B G and Martínez T J 2007 Isomerization through conical intersections *Annu. Rev. Phys. Chem.* **58** 613–34
[3] Turro N J, Ramamurthy V and Scaiano J C 2009 Principles of molecular photochemistry: An introduction *Principles of Molecular Photochemistry: An Introduction* (Mill Valley, CA: University Science Books)
[4] Yarkony D R 2012 Nonadiabatic quantum chemistry - past, present and future *Chem. Rev.* **112** 481–98
[5] Barbatti M, Borin A C and Ullrich S 2014 Photoinduced processes in nucleic acids *Photoinduced Phenomena in Nucleic Acids I, of Topics in Current Chemistry* (Berlin Heidelberg: Springer) vol 355 pp 1–32
[6] Ibele L M, Nicolson A and Curchod B F E 2019 Excited-state dynamics of molecules with classically driven trajectories and gaussians *Mol. Phys.* **118** 8
[7] Nelson T R *et al* 2020 Non-adiabatic excited-state molecular dynamics: Theory and applications for modeling photophysics in extended molecular materials *Chem. Rev.* **120** 2215–87
[8] Mai S and González L 2020 Molecular photochemistry: Recent developments in theory *Angew. Chem. Int. Ed.* in preparation
[9] Matsika S and Krylov A I 2018 Introduction: Theoretical modeling of excited state processes *Chem. Rev.* **118** 6925–6
[10] Lischka H, Nachtigallová D, Aquino A J A, Szalay P G, Plasser F, Machado F B C and Barbatti M 2018 Multireference approaches for excited states of molecules *Chem. Rev.* **118** 7293–361
[11] Ghosh S, Verma P, Cramer C J, Gagliardi L and Truhlar D G 2018 Combining wave function methods with density functional theory for excited states *Chem. Rev.* **118** 7249–92
[12] Norman P and Dreuw A 2018 Simulating x-ray spectroscopies and calculating core-excited states of molecules *Chem. Rev.* **118** 7208–48

[13] Casanova D 2018 Theoretical modeling of singlet fission *Chem. Rev.* **118** 7164–207
[14] Hestand N J and Spano F C 2018 Expanded theory of h- and j-molecular aggregates: The effects of vibronic coupling and intermolecular charge transfer *Chem. Rev.* **118** 7069–163
[15] Penfold T J, Gindensperger E, Daniel C and Marian C M 2018 Spin-vibronic mechanism for intersystem crossing, *Chem. Rev.* **118** 6975–7025
[16] Vacher M *et al* 2018 Chemi- and bioluminescence of cyclic peroxides *Chem. Rev.* **118** 6927–74
[17] Crespo-Otero R and Barbatti M 2018 Recent advances and perspectives on nonadiabatic mixed quantum–classical dynamics *Chem. Rev.* **118** 7026–68
[18] González L and Lindh R 2020 *Quantum Chemistry and Dynamics of Excited States : Methods and Applications* (New York: Wiley)
[19] Harris D C and Bertolucci M D 1989 *Symmetry and Spectroscopy: an Introduction to Vibrational and Electronic Spectroscopy* (New York: Dover Publications)
[20] Ng C-Y 1991 *Vacuum Ultraviolet Photoionization and Photodissociation of Molecules and Clusters* (Singapore: World Scientific)
[21] Zewail A H 1994 *Femtochemistry* (Singapore: World Scientific) pp 3–22
[22] Brixner T, Pfeifer T, Gerber G, Wollenhaupt M and Baumert T 2005 Optimal control of atomic, molecular and electron dynamics with tailored femtosecond laser pulses *Femtosecond Laser Spectroscopy* ed Hannaford P (New York: Springer) pp 225–66
[23] Neves-Petersen M T, Klitgaard Søren, Pascher T, Skovsen E, Polivka T, Yartsev A, Sundström V and Petersen S B 2009 Flash Photolysis of Cutinase: Identification and Decay Kinetics of Transient Intermediates Formed upon UV Excitation of Aromatic Residues *Biophys.* J. **97** 211–26
[24] Iqbal A and Stavros V G 2010 Active participation of $1\pi\sigma^{\star}$ states in the photodissociation of tyrosine and its subunits *J. Phys. Chem. Lett.* **1** 2274–8
[25] Maliş M 2012 Unraveling the mechanisms of nonradiative deactivation in model peptides following photoexcitation of a phenylalanine residue *J. Am. Chem. Soc.* **134** 20340–51 *et al.*
[26] Kowalewski M, Fingerhut B P, Dorfman K E, Bennett K and Mukamel S 2017 Simulating coherent multidimensional spectroscopy of nonadiabatic molecular processes: From the infrared to the x-ray regime *Chem. Rev.* **117** 12165–226
[27] Soorkia S, Jouvet C and Grégoire G 2020 UV photoinduced dynamics of conformer-resolved aromatic peptides *Chem. Rev.* **120** 3296–327
[28] Tajti A, Fogarasi G and Szalay P G 2009 Reinterpretation of the UV spectrum of cytosine: Only two electronic transitions? *ChemPhysChem* **10** 1603–6
[29] Barbatti M, Szymczak J J, Aquino A J A, Nachtigallová D and Lischka H 2011 The decay mechanism of photoexcited guanine–a nonadiabatic dynamics study *J. Chem. Phys.* **134** 014304
[30] Lu Y, Lan Z and Thiel W 2014 Computational modeling of photoexcitation in DNA single and double strands *Photoinduced Phenomena in Nucleic Acids II 356 of Topics in Current Chemistry* (Berlin Heidelberg: Springer) pp 89–122
[31] Ruckenbauer M, Mai S, Marquetand P and González L 2016 Photoelectron spectra of 2-thiouracil, 4-thiouracil and 2,4-dithiouracil *J. Chem. Phys.* **144** 074303
[32] Manathunga M, Yang X, Luk H L, Gozem S, Frutos L M, Valentini A, Ferrè N and Olivucci M 2016 Probing the photodynamics of rhodopsins with reduced retinal chromophores *J. Chem. Theory Comput.* **12** 839–50
[33] Nogueira J J, Plasser F and González L 2017 Electronic delocalization, charge transfer and hypochromism in the UV absorption spectrum of polyadenine unravelled by multiscale computations and quantitative wavefunction analysis *Chem. Sci.* **8** 5682–91
[34] Mai S, Mohamadzade A, Marquetand P, González L and Ullrich S 2018 Simulated and experimental time-resolved photoelectron spectra of the intersystem crossing dynamics in 2-thiouracil *Molecules* **23** 2836
[35] Rauer C, Nogueira J J, Marquetand P and González L 2018 Stepwise photosensitized thymine dimerization mediated by an exciton intermediate *Monatsh. Chem.* **149** 1–9
[36] Zobel J P, Heindl M, Nogueira J J and González L 2018 Vibrational sampling and solvent effects on the electronic structure of the absorption spectrum of 2-nitronaphthalene *J. Chem. Theory Comput.* **14** 3205–17
[37] Barbatti M, Aquino A J A and Lischka H 2006 Ultrafast two-step process in the non-adiabatic relaxation of the CH_2NH_{2+} molecule *Mol. Phys.* **104** 1053–60
[38] Curchod B F E and Tavernelli I 2013 On trajectory-based nonadiabatic dynamics: Bohmian dynamics versus trajectory surface hopping, *J. Chem. Phys.* **138** 184112
[39] Akimov A V, Asahi R, Jinnouchi R and Prezhdo O V 2015 What makes the photocatalytic CO_2 reduction on n-doped Ta_2O_5 efficient: Insights from nonadiabatic molecular dynamics *J. Am. Chem. Soc.* **137** 11517–25
[40] Schapiro I, Roca-Sanjuán D, Lindh R and Olivucci M 2015 A surface hopping algorithm for nonadiabatic minimum energy path calculations *J. Comput. Chem.* **36** 312–20
[41] Rauer C, Nogueira J J, Marquetand P and González L 2016 Cyclobutane thymine photodimerization mechanism revealed by nonadiabatic molecular dynamics *J. Am. Chem. Soc.* **138** 15911–16
[42] Ruckenbauer M, Mai S, Marquetand P and González L 2016 Revealing deactivation pathways hidden in time-resolved photoelectron spectra *Sci. Rep.* **6** 35522
[43] Mai S, Richter M, Marquetand P and González L 2017 The DNA nucleobase thymine in motion – intersystem crossing simulated with surface hopping *Chem. Phys.* **482** 9–15
[44] Mai S and González L 2019 Unconventional two-step spin relaxation dynamics of $[Re(CO)_3(im)(phen)]^+$ in aqueous solution *Chem. Sci.* **10** 10405–11
[45] Horton S L *et al* 2019 Excited state dynamics of CH_2I_2 and CH_2BrI studied with UV pump VUV probe photoelectron spectroscopy *J. Chem. Phys.* **150** 174201
[46] Heim P, Mai S, Thaler B, Cesnik S, Avagliano D, Bella-Velidou D, Ernst W E, González L and Koch M 2020 Revealing ultrafast population transfer between nearly degenerate electronic states *J. Phys. Chem. Lett.* **11** 1443–9
[47] Mai S, Marquetand P and González L 2016 Intersystem crossing pathways in the noncanonical nucleobase 2-thiouracil: A time-dependent picture *J. Phys. Chem. Lett.* **7** 1978–83
[48] Doltsinis N L 2006 *Molecular Dynamics Beyond the Born-Oppenheimer Approximation: Mixed Quantum-Classical Approaches of NIC Series* (Jülich, Germany: John von Neuman Institut for Computing) vol 31
[49] Köppel H, Domcke W and Cederbaum L S 1984 Multimode molecular dynamics beyond the Born-Oppenheimer approximation *Adv. Chem. Phys.* **57** 59–246
[50] Worth G A and Cederbaum L S 2004 Beyond Born-Oppenheimer: Molecular dynamics through a conical intersection *Annu. Rev. Phys. Chem.* **55** 127–58

[51] Richter M, Marquetand P, González-Vázquez J, Sola I and González L 2011 SHARC: Ab initio molecular dynamics with surface hopping in the adiabatic representation including arbitrary couplings *J. Chem. Theory Comput.* **7** 1253–8
[52] Mai S, Marquetand P and González L 2018 Nonadiabatic Dynamics: The SHARC Approach *WIREs Comput. Mol. Sci.* **8** e1370
[53] Chandrasekaran A, Kamal D, Batra R, Kim C, Chen L and Ramprasad R 2019 Solving the electronic structure problem with machine learning *npj Comput. Mater.* **5** 22
[54] Carleo G and Troyer M 2017 Solving the quantum many-body problem with artificial neural networks *Science* **355** 602–6
[55] Saito H 2017 Solving the Bose–Hubbard model with machine learning *J. Phys. Soc. Jpn.* **86** 093001
[56] Nomura Y, Darmawan A S, Yamaji Y and Imada M 2017 Restricted Boltzmann machine learning for solving strongly correlated quantum systems *Phys. Rev.* B **96** 205152
[57] Han J, Zhang L and W E 2018 Solving many-electron Schrödinger equation using deep neural networks
[58] Townsend J and Vogiatzis K D 2019 Data-driven acceleration of the coupled-cluster singles and doubles iterative solver *J. Phys. Chem. Lett.* **10** 4129–35
[59] Schütt K T, Gastegger M, Tkatchenko A, Müller K-R and Maurer R J 2019 Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions *Nat. Commun.* **10** 5024
[60] Pfau D, Spencer J S, Matthews A G de G, and Foulkes W M C 2019 Ab-initio solution of the many-electron Schrödinger equation with deep neural networks arXiv:1909.02487
[61] Hermann J, Schätzle Z and Noé F 2019 Deep neural network solution of the electronic Schrödinger equation arXiv:1909.08423
[62] Gastegger M, McSloy A, Luya M, Schütt K T and Maurer R J 2020 A deep neural network for molecular wave functions in quasi-atomic minimal basis representation *J. Chem. Phys.* **153** 044123
[63] Choo K, Carleo G, Regnault N and Neupert T 2018 Symmetries and many-body excitations with neural-network quantum states *Phys. Rev. Lett.* **121** 167204
[64] Zheng F, Gao X and Eisfeld A 2019 Excitonic wave function reconstruction from near-field spectra using machine learning techniques *Phys. Rev. Lett.* **123** 163202
[65] Brockherde F, Vogt L, Li L, Tuckerman M E, Burke K and Müller K-R 2017 Bypassing the Kohn-Sham equations with machine learning *Nat. Commun.* **8** 872
[66] Hegde G and Bowen R C 2017 Machine-learned approximations to density functional theory Hamiltonians *Sci. Rep.* **7** 42669
[67] Gastegger M, Gonzáez L and Marquetand P 2019 Exploring density functional subspaces with genetic algorithms *Monatsh. Chem.* **150** 173–82
[68] Nelson J, Tiwari R and Sanvito S 2019 Machine learning density functional theory for the Hubbard model *Phys. Rev.* B **99** 075132
[69] Cheng L, Welborn M, Christensen A S and Miller T F 2019 A universal density matrix functional from molecular orbital-based machine learning: Transferability across organic molecules *J. Chem. Phys.* **150** 131103
[70] Lei X and Medford A J 2019 Design and analysis of machine learning exchange-correlation functionals via rotationally invariant convolutional descriptors *Phys. Rev. Materials* **3** 063801
[71] Zhou Y, Wu J, Chen S and Chen G 2019 Toward the exact exchange–correlation potential: A three-dimensional convolutional neural network construct *J. Phys. Chem. Lett.* **10** 7264–9
[72] Kolb B, Lentz L C and Kolpak A M 2017 Discovering charge density functionals and structure-property relationships with prophet: A general framework for coupling machine learning and first-principles methods *Sci. Rep.* **7** 1192
[73] Willatt M J, Musil F and Ceriotti M 2019 Atom-density representations for machine learning *J. Chem. Phys.* **150** 154110
[74] Hobday S, Smith R and Belbruno J 1999 Applications of neural networks to fitting interatomic potential functions *Modell. Simul. Mater. Sci. Eng.* **7** p 397
[75] Bartók A P, Payne M C, Kondor R and Csányi G 2010 Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons *Phys. Rev. Lett.* **104** 136403
[76] Rupp M, Tkatchenko A, Müller K-R and von Lilienfeld O A 2012 Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning *Phys. Rev. Lett.* **108** 058301
[77] Li Z, Kermode J R and De Vita A 2015 Molecular dynamics with on-the-fly machine learning of quantum-mechanical forces *Phys. Rev. Lett.* **114** 096405
[78] von Lilienfeld O A, Ramakrishnan R, Rupp M and Knoll A 2015 Fourier series of atomic radial distribution functions: A molecular fingerprint for machine learning models of quantum chemical properties *Int. J. Quantum Chem.* **115** 1084–93
[79] Gastegger M and Marquetand P 2015 High-dimensional neural network potentials for organic reactions and an improved training algorithm *J. Chem. Theory Comput.* **11** 2187–98
[80] Rupp M, Ramakrishnan R and von Lilienfeld O A 2015 Machine learning for quantum mechanical properties of atoms in molecules *J. Phys. Chem. Lett.* **6** 3309–13
[81] Behler J 2016 Perspective: Machine learning potentials for atomistic simulations *J. Chem. Phys.* **145** 170901
[82] Artrith N and Urban A 2016 An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for TiO_2 *Comput. Mater. Sci.* **114** 135–50
[83] Gastegger M, Kauffmann C, Behler J and Marquetand P 2016 Comparing the accuracy of high-dimensional neural network potentials and the systematic molecular fragmentation method: A benchmark study for all-trans alkanes *J. Chem. Phys.* **144** 194110
[84] Artrith N, Urban A and Ceder G 2017 Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species *Phys. Rev.* B **96** 014112
[85] Gastegger M, Behler J and Marquetand P 2017 Machine learning molecular dynamics for the simulation of infrared spectra *Chem. Sci.* **8** 6924–35
[86] Deringer V L and Csányi G 2017 Machine learning based interatomic potential for amorphous carbon, *Phys. Rev.* B **95** 094203
[87] Botu V, Batra R, Chapman J and Ramprasad R 2017 Machine learning force fields: Construction, validation and outlook *J. Phys. Chem.* C **121** 511–22
[88] Glielmo A, Sollich P and De Vita A 2017 Accurate interatomic force fields via machine learning with covariant kernels *Phys. Rev.* B **95** 214302
[89] Smith J S, Isayev O and Roitberg A E 2017 ANI-1: an extensible neural network potential with dft accuracy at force field computational cost *Chem. Sci.* **8** 3192–203
[90] Fujikake S, Deringer V L, Lee T H, Krynski M, Elliott S R and Csányi G 2018 Gaussian approximation potential modeling of lithium intercalation in carbon nanostructures *J. Chem. Phys.* **148** 241714
[91] Behler J 2017 First principles neural network potentials for reactive simulations of large molecular and condensed systems *Angew. Chem. Int. Edit.* **56** 12828–40

[92] Zong H, Pilania G, Ding X, Ackland G J and Lookman T 2018 Developing an interatomic potential for martensitic phase transformations in zirconium by machine learning *npj Comput. Mater* **4** 48
[93] Wood M A and Thompson A P 2018 Extending the accuracy of the snap interatomic potential form *J. Chem. Phys.* **148** 241721
[94] Chen X, Jørgensen M S, Li J and Hammer B 2018 Atomic energies from a convolutional neural network *J. Chem. Theory Comput.* **14** 3933–42
[95] Bartók A P, Kermode J, Bernstein N and Csányi G 2018 Machine learning a general-purpose interatomic potential for silicon *Phys. Rev.* X **8** 041048
[96] Chmiela S, Sauceda H E, Müller K-R and Tkatchenko A 2018 Towards exact molecular dynamics simulations with machine-learned force fields *Nat. Commun.* **9** 3887
[97] Imbalzano G, Anelli A, Giofré D, Klees S, Behler J and Ceriotti M 2018 Automatic selection of atomic fingerprints and reference configurations for machine-learning potentials *J. Chem. Phys.* **148** 241730
[98] Zhang L, Han J, Wang H, Saidi W A, Car R and E W 2018 End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems *Proc. of the 32nd Int. Conference on Neural Information Processing Systems* NIPS'18, (USA) Curran Associates Inc. pp 4441–51
[99] Zhang L, Han J, Wang H, Car R and E W 2018 Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics *Phys. Rev. Lett.* **120** 143001
[100] Chan H, Narayanan B, Cherukara M J, Sen F G, Sasikumar K, Gray S K, Chan M K Y and Sankaranarayanan S K R S 2019 Machine learning classical interatomic potentials for molecular dynamics from first-principles training data *J. Phys. Chem.* C **123** 6941–57
[101] Faber F A, Christensen A S, Huang B and von Lilienfeld O A 2018 Alchemical and structural distribution based representation for universal quantum machine learning *J. Chem. Phys.* **148** 241717
[102] Wang H and Yang W 2019 Toward building protein force fields by residue-based systematic molecular fragmentation and neural network *J. Chem. Theory Comput.* **15** 1409–17
[103] Gerrits N, Shakouri K, Behler J and Kroes G-J 2019 Accurate probabilities for highly activated reaction of polyatomic molecules on surfaces using a high-dimensional neural network potential: CHD_3 + Cu(111) *J. Phys. Chem. Lett.* **10** 1763–8
[104] Chmiela S, Sauceda H E, Poltavsky I, Müller K-R and Tkatchenko A 2019 GDML: Constructing accurate and data efficient molecular force fields using machine learning *Comput. Phys. Commun.* **240** 38–45
[105] Carleo G, Cirac I, Cranmer K, Daudet L, Schuld M, Tishby N, Vogt-Maranto L and Zdeborová L 2019 Machine learning and the physical sciences *Rev. Mod. Phys.* **91** 045002
[106] Krems R V 2019 Bayesian Machine Learning for Quantum Molecular Dynamics *Phys. Chem. Chem. Phys.* **21** 13392–13410
[107] Deringer V L, Caro M A and Csányi G 2019 Machine learning interatomic potentials as emerging tools for materials science *Adv. Mat.* **31** 1902765
[108] Ward L, Blaiszik B, Foster I, Assary R S, Narayanan B and Curtiss L 2019 Machine learning prediction of accurate atomization energies of organic molecules from low-fidelity quantum chemical calculations *MRS Commun.* **9** 891–9
[109] Noé F, Tkatchenko A, Müller K-R and Clementi C 2020 Machine learning for molecular simulation *Annu. Rev. Phys. Chem.* **71** 361–90
[110] Behler J, Reuter K and Scheffler M 2008 Nonadiabatic effects in the dissociation of oxygen molecules at the Al(111) surface *Phys. Rev.* B **77** 115421
[111] Carbogno C, Behler J, Reuter K and Groß A 2010 Signatures of nonadiabatic O_2 dissociation at Al(111): First-principles fewest-switches study *Phys. Rev.* B **81** 035410
[112] Hu D, Xie Y, Li X, Li L and Lan Z 2018 Inclusion of machine learning kernel ridge regression potential energy surfaces in on-the-fly nonadiabatic molecular dynamics simulation *J. Phys. Chem. Lett.* **9** 2725–32
[113] Dral P O, Barbatti M and Thiel W 2018 Nonadiabatic excited-state dynamics with machine learning *J. Phys. Chem. Lett.* **9** 5660–3
[114] Chen W-K, Liu X-Y, Fang W-H, Dral P O and Cui G 2018 Deep learning for nonadiabatic excited-state dynamics *J. Phys. Chem. Lett.* **9** 6702–8
[115] Williams D M G and Eisfeld W 2018 Neural network diabatization: A new ansatz for accurate high-dimensional coupled potential energy surfaces, *J. Chem. Phys.* **149** 204106
[116] Xie C, Zhu X, Yarkony D R and Guo H 2018 Permutation invariant polynomial neural network approach to fitting potential energy surfaces. IV. coupled diabatic potential energy matrices *J. Chem. Phys.* **149** 144107
[117] Guan Y, Zhang D H, Guo H and Yarkony D R 2019 Representation of coupled adiabatic potential energy surfaces using neural network based quasi-diabatic Hamiltonians: $1,2^2$A' states of LiFH *Phys. Chem. Chem. Phys.* **21** 14205–13
[118] Westermayr J, Gastegger M, Menger M F S J, Mai S, González L and Marquetand P 2019 Machine learning enables long time scale molecular photodynamics simulations *Chem. Sci.* **10** 8100–7
[119] Guan Y, Guo H and Yarkony D R 2020 Extending the representation of multistate coupled potential energy surfaces to include properties operators using neural networks: Application to the $1,2^1$A states of ammonia *J. Chem. Theory Comput.* **16** 302–13
[120] Richings G W and Habershon S 2018 MCTDH on-the-fly: Efficient grid-based quantum dynamics without pre-computed potential energy surfaces, *J. Chem. Phys.* **148** 134116
[121] Alborzpour J P, Tew D P and Habershon S 2016 Efficient and accurate evaluation of potential energy matrix elements for quantum dynamics using gaussian process regression *J. Chem. Phys.* **145** 174112
[122] Richings G W, Robertson C and Habershon S 2019 Improved on-the-fly MCTDH simulations with many-body-potential tensor decomposition and projection diabatization *J. Chem. Theory Comput.* **15** 857–70
[123] Polyak I, Richings G W, Habershon S and Knowles P J 2019 Direct quantum dynamics using variational Gaussian wavepackets and Gaussian process regression *J. Chem. Phys.* **150** 041101
[124] Guan Y, Guo H and Yarkony D R 2019 Neural network based quasi-diabatic Hamiltonians with symmetry adaptation and a correct description of conical intersections *J. Chem. Phys.* **150** 214101
[125] Wang Y, Xie C, Guo H and Yarkony D R 2019 A quasi-diabatic representation of the 1,21a states of methylamine *J. Phys. Chem.* A **123** 5231–41
[126] Guan Y and Yarkony D R 2020 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.* **11** 1848–58
[127] Richings G W and Habershon S 2017 Direct grid-based quantum dynamics on propagated diabatic potential energy surfaces *Chem. Phys. Lett.* **683** 228–33
[128] Netzloff H M, collins M A and Gordon M S 2006 Growing multiconfigurational potential energy surfaces with applications to $X+H_2$ (X=C,N,O) reactions *J. Chem. Phys.* **124** p 154104

[129] Bettens R P A and Collins M A 1999 Learning to interpolate molecular potential energy surfaces with confidence: A Bayesian approach *J. Chem. Phys.* **111** 816–26
[130] Carbogno C, Behler J, Groß A and Reuter K 2008 Fingerprints for Spin-Selection Rules in the Interaction Dynamics of O_2 at Al(111) *Phys. Rev. Lett.* **101** 096104
[131] Ghosh K, Stuke A, Todorović M, Jørgensen P B, Schmidt M N, Vehtari A and Rinke P 2019 Deep learning spectroscopy: Neural networks for molecular excitation spectra *Adv. Sci.* **6** 1801367
[132] Pilania G, Gubernatis J and Lookman T 2017 Multi-fidelity machine learning models for accurate bandgap predictions of solids *Comput. Mat. Sci.* **129** 156–63
[133] Schütt K T, Arbabzadah F, Chmiela S, Müller K R and Tkatchenko A 2017 Quantum-chemical insights from deep tensor neural networks *Nat. Commun.* **8** 13890 EP
[134] Zhuo Y, Mansouri Tehrani A and Brgoch J 2018 Predicting the band gaps of inorganic solids by machine learning *J. Phys. Chem. Lett.* **9** 1668–73
[135] Pereira F, Xiao K, Latino D A R S, Wu C, Zhang Q and Aires-de Sousa J 2017 Machine learning methods to predict density functional theory B3LYP energies of HOMO and LUMO orbitals *J. Chem. Inf. Model.* **57** 11–21
[136] Isayev O, Oses c, Toher c, Gossett E, Curtarolo S and Tropsha A 2017 Universal fragment descriptors for predicting properties of inorganic crystals, *Nat. Commun.* **8** 15679
[137] Pronobis W, Schütt K R, Tkatchenko A and Müller K-R 2018 Capturing intensive and extensive DFT/TDDFT molecular properties with machine learning *Eur. Phys. J.* B **91** 178
[138] Stuke A, Todorović M, Rupp M, Kunkel C, Ghosh K, Himanen L and Rinke P 2019 Chemical diversity in molecular orbital energy predictions with kernel ridge regression *J. Chem. Phys.* **150** 204121
[139] Häse F, Valleau S, Pyzer-Knapp E and Aspuru-Guzik A 2016 Machine learning exciton dynamics *Chem. Sci.* **7** 5139–47
[140] Häse F, Fdez Galván I, Aspuru-Guzik A, Lindh R and Vacher M 2019 How machine learning can assist the interpretation of ab initio molecular dynamics simulations and conceptual understanding of chemistry *Chem. Sci.* **10** 2298–307
[141] O'Boyle N M, Campbell C M and Hutchison G R 2011 Computational design and selection of optimal organic photovoltaic materials *J. Phys. Chem.* C **115** 16200–10
[142] Teunissen J L, De Proft F and De Vleeschouwer F 2017 Tuning the HOMO-LUMO energy gap of small diamondoids using inverse molecular design *J. Chem. Theory Comput.* **13** 1351–65
[143] Liu D, Tan Y, Khoram E and Yu Z 2018 Training deep neural networks for the inverse design of nanophotonic structures *ACS Photonics* **5** 1365–9
[144] Elton D C, Boukouvalas Z, Fuge M D and Chung P W 2019 Deep learning for molecular design - a review of the state of the art *Mol. Syst. Des. Eng.* **4** 828–49
[145] Sanchez-Lengeling B and Aspuru-Guzik A 2018 Inverse molecular design using machine learning: Generative models for matter engineering *Science* **361** 360–5
[146] Goldsmith B R, Esterhuizen J, Liu J-X, Bartel C J and Sutton C 2018 Machine learning for heterogeneous catalyst design and discovery *AIChE J.* **64** 2311–23
[147] Davies D W, Butler K T, Isayev O and Walsh A 2018 Materials discovery by chemical analogy: role of oxidation states in structure prediction *Faraday Discuss.* **211** 553–68
[148] von Lilienfeld O A, Müller K-R and Tkatchenko A 2019 Exploring chemical compound space with quantum-based machine learning arXiv:1911.10084
[149] Freeze J G, Kelly H R and Batista V S 2019 Search for catalysts by inverse design: Artificial intelligence, mountain climbers and alchemists, *Chem. Rev.* **119** 6595–6612
[150] Lee M-H 2020 Robust random forest based non-fullerene organic solar cells efficiency prediction *Org. Electron.* **76** 105465
[151] Montavon G, Rupp M, Gobre V, Vazquez-Mayagoitia A, Hansen K, Tkatchenko A, Müller K-R and von Lilienfeld O A 2013 Machine learning of molecular electronic properties in chemical compound space *New J. Phys.* **15** 095003
[152] Hansen K, Biegler F, Ramakrishnan R, Pronobis W, von Lilienfeld O A, Müller K-R and Tkatchenko A 2015 Machine learning predictions of molecular properties: Accurate many-body potentials and nonlocality in chemical space *J. Phys. Chem. Lett.* **6** 2326–31
[153] Huang B and von Lilienfeld O A 2016 communication: Understanding molecular representations in machine learning: The role of uniqueness and target similarity *J. Chem. Phys.* **145** 161102
[154] Christensen A S and Faber F A and 2019 O A von Lilienfeld, operators in quantum machine learning: response properties in chemical space *J. Chem. Phys.* **150** 064105
[155] Christensen A S, Bratholm L A, Faber F A and O Anatole von Lilienfeld 2020 FCHL revisited: Faster and more accurate quantum machine learning *J. Chem. Phys.* **152** 044107
[156] Schütt K T, Sauceda H E, Kindermans P-J, Tkatchenko A and Müller K-R 2018 SchNet–a deep learning architecture for molecules and materials *J. Chem. Phys.* **148** 241722
[157] Tawfik S A, Isayev O, Spencer M J S and Winkler D A 2019 Predicting thermal properties of crystals using machine learning *Adv. Theory Sim.* 1900208
[158] Christensen A S and von Lilienfeld O A 2019 Operator quantum machine learning: Navigating the chemical space of response properties *CHIMIA* **73** 1028–31
[159] Fias S, Chang K Y S and von Lilienfeld O A 2019 Alchemical normal modes unify chemical space, *J. Phys. Chem. Lett.* **10** 30–9
[160] von Rudorff G F and von Lilienfeld O A 2020 Rapid and accurate molecular deprotonation energies from quantum alchemy *Phys. Chem. Chem. Phys.* **22** 10519–25
[161] Westermayr J, Faber F, Christensen A S, von Lilienfeld A and Marquetand P 2020 Neural networks and kernel ridge regression for excited states dynamics of $CH_2NH_2^+$: From single-state to multi-state representations and multi-property machine learning models *Mach. Learn.: Sci. Technol.* **1** 025009
[162] Zubatyuk R, Smith J S, Leszczynski J and Isayev O 2019 Accurate and transferable multitask prediction of chemical properties with an atoms-in-molecules neural network *Sci. Adv.* **5** eaav6490
[163] von Lilienfeld O A 2018 Quantum machine learning in chemical compound space *Angew. Chem. Int. Edit.* **57** 4164–9
[164] Park J W, Al-Saadon R, MacLeod M K, Shiozaki T and Vlaisavljevich B 2020 Multireference electron correlation methods: journeys along potential energy surfaces *Chem. Rev.* **120** 5878–909
[165] Westermayr J, Gastegger M and Marquetand P 2020 Combining SchNet and SHARC: The SchNarc Machine Learning Approach for Excited-State Dynamics *J. Phys. Chem. Lett.* **11** 3828–34
[166] Goodfellow I, Bengio Y, and Courville A 2016 *Deep Learning* (Cambridge, MA: MIT Press)

[167] Akimov A V 2018 A simple phase correction makes a big difference in nonadiabatic molecular dynamics *J. Phys. Chem. Lett.* **9** 6096–102
[168] Plasser F, Ruckenbauer M, Mai S, Oppel M, Marquetand P and González L 2016 Efficient and flexible computation of many-electron wave function overlaps *J. Chem. Theory Comput.* **12** 1207
[169] Robertson C, González-Vázquez J, corral I, Díaz-Tendero S and Díaz C 2019 Nonadiabatic scattering of N off Au3 clusters: A simple and robust diabatic state manifold generation method for multiconfigurational wavefunctions *J. Comput. Chem.* **40** 794–810
[170] Mai S, Marquetand P and González L 2015 A general method to describe intersystem crossing dynamics in trajectory surface hopping *Int. J. Quantum Chem.* **115** 1215–31
[171] Shu Y *et al* 2019 Direct diabatization 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.* **151** 104311
[172] Schütt K T, Kessel P, Gastegger M, Nicoli K A, Tkatchenko A and Müller K-R 2019 Schnetpack: A deep learning toolbox for atomistic systems *J. Chem. Theory Comput.* **15** 448–55
[173] Ryabinkin I G, Joubert-Doriol L and Izmaylov A F 2017 Geometric phase effects in nonadiabatic dynamics near conical intersections *Acc. Chem. Res.* **50** 1785–93
[174] Behler J, Martoák R, Donadio D and Parrinello M 2008 Metadynamics simulations of the high-pressure phases of silicon employing a high-dimensional neural network potential *Phys. Rev. Lett.* **100** 185501
[175] Gastegger M, Schwiedrzik L, Bittermann M, Berzsenyi F and Marquetand P 2018 WACSF – weighted atom-centered symmetry functions as descriptors in machine learning potentials *J. Chem. Phys.* **148** 241709
[176] Koch W and Zhang D H 2014 Communication: Separable potential energy surfaces from multiplicative artificial neural networks *J. Chem. Phys.* **141** 021101
[177] He D, Yuan J, Li H and Chen M 2016 Global Diabatic Potential Energy Surfaces and Quantum Dynamical Studies for the Li(2p) + $H_2(X^1\Sigma_g^+) \rightarrow LiH(X^1\Sigma^+)$ + H reaction *Sci. Rep.* **6** 25083
[178] Guan Y, Fu B and Zhang D H 2017 Construction of diabatic energy surfaces for LiFH with artificial neural networks *J. Chem. Phys.* **147** 224307
[179] Wang S, Yang Z, Yuan J and Chen M 2018 New Diabatic Potential Energy Surfaces of the NaH_2 System and Dynamics Studies for the Na(3p) + $H_2 \rightarrow$ NaH + H Reaction *Sci. Rep.* **8** 17960
[180] Yuan J, He D, Wang S, Chen M and Han K 2018 Diabatic potential energy surfaces of MgH_2^+ and dynamic studies for the $Mg^+(3p)$ +$H_2 \rightarrow MgH^+$ + H reaction *Phys. Chem. Chem. Phys.* **20** 6638–47
[181] Yin Z, Guan Y, Fu B and Zhang D H 2019 Two-state diabatic potential energy surfaces of ClH_2 based on nonadiabatic couplings with neural networks *Phys. Chem. Chem. Phys.* **21** 20372–83
[182] Schwilk M, Tahchieva D N, and von Lilienfeld O A 2020 Large yet bounded: Spin gap ranges in carbenes, arXiv:2004.10600
[183] Dral P O 2020 Quantum chemistry in the age of machine learning *J. Phys. Chem. Lett.* **11** 2336–47
[184] Botu V and Ramprasad R 2015 Adaptive machine learning framework to accelerate ab initio molecular dynamics *Int. J. Quant. Chem.* **115** 1074–83
[185] Behler J 2015 Constructing high-dimensional neural network potentials: A tutorial review *Int. J. Quantum Chem.* **115** 1032–50
[186] Smith J S, Nebgen B, Lubbers N, Isayev O and Roitberg A E 2018 Less is more: Sampling chemical space with active learning *J. Chem. Phys.* **148** 241733
[187] Raghunathan Ramakrishnan M R, Dral P O and von Lilienfeld O A 2014 Quantum chemistry structures and properties of 134 kilo molecules *Sci. Data* **1** 140022
[188] Wigner E 1932 On the quantum correction for thermodynamic equilibrium *Phys. Rev.* **40** 749–50
[189] Bruccoleri R E and Karplus M 1990 conformational sampling using high-temperature molecular dynamics *Biopolymers* **29** 1847–62
[190] Maximova T, Moffatt R, Ma B, Nussinov R and Shehu A 04 2016 Principles and overview of sampling methods for modeling macromolecular structure and dynamics *PLOS Computat. Biol.* **12** 1–70
[191] Kästner J 2011 Umbrella sampling *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1** 932–42
[192] Tao G 2019 Trajectory-guided sampling for molecular dynamics simulation *Theor. Chem. Acc.* **138** p 34
[193] Yang Y I, Shao Q, Zhang J, Yang L and Gao Y Q 2019 Enhanced sampling in molecular dynamics *J. Chem. Phys.* **151** 070902
[194] Herr J E, Yao K, McIntyre R, Toth D W and Parkhill J 2018 Metadynamics for training neural network model chemistries: A competitive assessment *J. Chem. Phys.* **148** 241710
[195] Grimme S 2019 Exploration of chemical compound, conformer and reaction space with meta-dynamics simulations based on tight-binding quantum chemical calculations *J. Chem. Theory Comput.* **15** 2847–62
[196] Rupp M 2015 Machine learning for quantum mechanics in a nutshell *Int. J. Quantum Chem.* **115** 1058–73
[197] Seung H S, Opper M and Sompolinsky H 1992 Query by committee *Proc. of the Fifth Annual Workshop on Computational Learning Theory* COLT 92 New York, NY, USA Association for Computing Machinery pp 287–294
[198] Gastegger M and Marquetand P 2020 Molecular dynamics with neural-network potentials *Machine Learning Meets Quantum Physics* (Berlin: Springer International) pp 233–52
[199] Tully J C 1990 Molecular dynamics with electronic transitions *J. Chem. Phys.* **93** 1061–71
[200] Tully J C 1991 Nonadiabatic molecular dynamics *Int. J. Quantum Chem.* **40** 299–309
[201] Dral P O, Owens A, Yurchenko S N and Thiel W 2017 Structure-based sampling and self-correcting machine learning for accurate calculations of potential energy surfaces and vibrational levels *J. Chem. Phys.* **146** 244108
[202] Sobol' I M, Asotsky D, Kreinin A and Kucherenko S 2011 construction and comparison of high-dimensional Sobol' generators *Wilmott* **2011** 64–79
[203] Ceriotti M, Tribello G A and Parrinello M 2013 Demonstrating the transferability and the descriptive power of sketch-map *J. Chem. Theory Comput.* **9** 1521–32
[204] Behler J 2011 Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations *Phys. Chem. Chem. Phys.* **13** 17930–55
[205] Lin Q, Zhang Y, Zhao B and Jiang B 2020 Automatically growing global reactive neural network potential energy surfaces: A trajectory-free active learning strategy *J. Chem. Phys.* **152** 154104
[206] Barbatti M, Ruckenbauer M and Lischka H 2005 The photodynamics of ethylene: A surface-hopping study on structural aspects *J. Chem. Phys.* **122** 174307
[207] Hollas D, Sistík L, Hohenstein E G, Martínez T J and Slavícek P 2018 Nonadiabatic ab initio molecular dynamics with the floating occupation molecular orbital-complete active space configuration interaction method *J. Chem. Theory Comput.* **14** 339–50

[208] Köppel H, Gronki J and Mahapatra S 2001 construction scheme for regularized diabatic states *J. Chem. Phys.* **115** 2377–88
[209] Bowman J M, Carrington T and Meyer H 2008 Variational quantum approaches for computing vibrational energies of polyatomic molecules *Mol. Phys.* **106** 2145–82
[210] Meyer H-D, Gatti F and Worth G A 2009 *The Road to* MCTDH ch 2 (New York: Wiley-VCH Verlag GmbH & co. KGaA) pp 9–15
[211] Liu F, Du L, Zhang D and Gao J 2017 Direct learning hidden excited state interaction patterns from ab initio dynamics and its implication as alternative molecular mechanism models *Sci. Rep.* **7** 1–12
[212] Baer M 2002 Introduction to the theory of electronic non-adiabatic coupling terms in molecular systems *Phys. Rep.* **358** 75–142
[213] Lischka H, Dallos M, Szalay P G, Yarkony D R and Shepard R 2004 Analytic evaluation of nonadiabatic coupling terms at the MR-CI level. I. Formalism *J. Chem. Phys.* **120** 7322–9
[214] Ha J-K, Lee I S and Min S K 2018 Surface hopping dynamics beyond nonadiabatic couplings for quantum coherence *J. Phys. Chem. Lett.* **9** 1097–104
[215] Granucci G, Persico M and Spighi G 2012 Surface hopping trajectory simulations with spin-orbit and dynamical couplings *J. Chem. Phys.* **137** 22A501
[216] Tannor D 2006 *Introduction to Quantum Mechanics: A Time-Dependent Perspective* (Sausalito: University Science Books)
[217] Jasper A W, Kendrick B K, Mead C A and Truhlar D G 2004 *Non-Born-Oppenheimer Chemistry: Potential Surfaces, Couplings and Dynamics* (Singapore: World Scientific) pp 329–91
[218] Yarkony D R 2005 On the connectivity of seams of conical intersection: Seam curvature *J. Chem. Phys.* **123** 204101
[219] Zhu X and Yarkony D R 2016 Constructing diabatic representations using adiabatic and approximate diabatic data - coping with diabolical singularities *J. Chem. Phys.* **144** 044104
[220] Wittenbrink N, Venghaus F, Williams D and Eisfeld W 2016 A new approach for the development of diabatic potential energy surfaces: Hybrid block-diagonalization and diabatization by ansatz *J. Chem. Phys.* **145** 184108
[221] Domcke W, Yarkony D R and Köppel H 2004 *Conical Intersections: Electronic Structure, Dynamics and Spectroscopy* eds (Singapore: World Scientific) (https://doi.org/10.1142/5406)
[222] Richings G W and Worth G A 2015 A practical diabatisation scheme for use with the direct-dynamics variational multi-configuration Gaussian method *J. Phys. Chem.* A **119** 12457–70
[223] Accomasso D, Persico M and Granucci G 2019 Diabatization by localization in the framework of configuration interaction based on floating occupation molecular orbitals (FOMO-CI) *ChemPhotoChem* **3** 933–44
[224] Lenzen T and Manthe U 2017 Neural network based coupled diabatic potential energy surfaces for reactive scattering *J. Chem. Phys.* **147** 084105
[225] Subotnik J E, Yeganeh S, Cave R J and Ratner M A 2008 Constructing Diabatic States from Adiabatic States: Extending Generalized Mulliken–Hush to Multiple Charge Centers with Boys localization *J. Chem. Phys.* **129** 244101
[226] Hoyer C E, Parker K, Gagliardi L and Truhlar D G 2016 The DQ and DQØ electronic structure diabatization methods: Validation for general applications *J. Chem. Phys.* **144** 194101
[227] Wittenbrink N, Ndome H and Eisfeld W 2013 Toward spin–orbit coupled diabatic potential energy surfaces for methyl iodide using effective relativistic coupling by asymptotic representation *J. Phys. Chem.* A **117** 7408–20
[228] Varga Z, Parker K A and Truhlar D G 2018 Direct diabatization based on nonadiabatic couplings: the n/d method *Phys. Chem. Chem. Phys.* **20** 26643–59
[229] Nakamura H and Truhlar D G 2002 Direct diabatization of electronic states by the fourfold way. II. dynamical correlation and rearrangement processes *J. Chem. Phys.* **117** 5576–93
[230] Venghaus F and Eisfeld W 2016 Block-diagonalization as a tool for the robust diabatization of high-dimensional potential energy surfaces *J. Chem. Phys.* **144** 114110
[231] Li J, Jiang B and Guo H 2013 Permutation invariant polynomial neural network approach to fitting potential energy surfaces. II. Four-atom systems *J. Chem. Phys.* **139** 204103
[232] Jiang B and Guo H 2013 Permutation invariant polynomial neural network approach to fitting potential energy surfaces *J. Chem. Phys.* **139** 054112
[233] Jiang B and Guo H 2014 Permutation invariant polynomial neural network approach to fitting potential energy surfaces. III. molecule-surface interactions *J. Chem. Phys.* **141** 034109
[234] Jiang B, Li J and Guo H 2016 Potential energy surfaces from high fidelity fitting of ab initio points: the permutation invariant polynomial - neural network approach *Int. Rev. Phys. Chem.* **35** 479–506
[235] Ischtwan J and Collins M A 1994 Molecular potential energy surfaces by interpolation *J. Chem. Phys.* **100** 8080–8
[236] Evenhuis C R and Collins M A 2004 Interpolation of diabatic potential energy surfaces *J. Chem. Phys.* **121** 2515–2527
[237] Evenhuis C and Martínez T J 2011 A scheme to interpolate potential energy surfaces and derivative coupling vectors without performing a global diabatization *J. Chem. Phys.* **135** 224110
[238] Mukherjee S, Bandyopadhyay S, Paul A K and Adhikari S 2013 Construction of diabatic Hamiltonian matrix from ab initio calculated molecular symmetry adapted nonadiabatic coupling terms and nuclear dynamics for the excited states of Na_3 cluster *J. Phys. Chem.* A **117** 3475–95
[239] Zhu X and Yarkony D R 2012 Quasi-diabatic representations of adiabatic potential energy surfaces coupled by conical intersections including bond breaking: A more general construction procedure and an analysis of the diabatic representation *J. Chem. Phys.* **137** 22A511
[240] Zhu X and Yarkony D R 2014 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.* **140** 024112
[241] Richings G W, Robertson C and Habershon S 2019 Improved on-the-fly MCTDH simulations with many-body-potential tensor decomposition and projection diabatization *J. Chem. Theory Comput.* **15** 857–70
[242] Gómez S, Heindl M, Szabadi A and González L 2019 From surface hopping to quantum dynamics and back. finding essential electronic and nuclear degrees of freedom and optimal surface hopping parameters *J. Phys. Chem.* A **123** 8321–32
[243] Köppel H, Domcke W and Cederbaum L S 2004 *in: Conical Intersections* Eds Domcke W, Yarkony D R and Köppel H (New York: World Scientific) eds W Domcke *et al* (Singapore: World Scientific) **15** pp 323–67
[244] Yarkony D R 2004 conical intersections: Their description and consequences *Advanced Series in Physical Chemistry* pp 41–127
[245] Plasser F, Gómez S, Menger M F S J, Mai S and González L 2019 Highly efficient surface hopping dynamics using a linear vibronic coupling model *Phys. Chem. Chem. Phys.* **21** 57–69
[246] Fabiano E, Keal T and Thiel W 2008 Implementation of surface hopping molecular dynamics using semiempirical methods *Chem. Phys.* **349** 334–47

[247] Oloyede P, Mil ikov G and Nakamura H 2006 Generalized trajectory surface hopping method based on the Zhu-Nakamura theory *J. Chem. Phys.* **124** 144110
[248] Ishida T, Nanbu S and Nakamura H 2017 Clarification of nonadiabatic chemical dynamics by the Zhu-Nakamura theory of nonadiabatic transition: from tri-atomic systems to reactions in solutions *Int. Rev. Phys. Chem.* **36** 229–86
[249] Zener C 1932 Non-adiabatic crossing of energy levels *Proc. Roy. Soc. Lond.* A **137** 696–701
[250] Wittig C 2005 The Landau-Zener Formula *J. Phys. Chem.* B **109** 8428–30
[251] Zhu L, Kleiman V, Li X, Lu S P, Trentelman K and Gordon R J 1995 Ultrafast coherent control and destruction of excitons in quantum wells *Phys. Rev. Lett.* **75** 2598–601
[252] Zhu C, Kamisaka H and Nakamura H 2002 New implementation of the trajectory surface hopping method with use of the Zhu–Nakamura theory. II. Application to the charge transfer processes in the 3D DH_2^+ system *J. Chem. Phys.* **116** 3234–47
[253] Granucci G and Persico M 2007 Critical appraisal of the fewest switching algorithm for surface hopping *J. Chem. Phys.* **126** 134114
[254] Malhado J P, Bearpark M J and J T Hynes 2014 Non-adiabatic dynamics close to conical intersections and the surface hopping perspective, *Front. Chem.* **2** 97
[255] Wang L, Akimov A and O V Prezhdo 2016 Recent progress in surface hopping: 2011-2015, *J. Phys. Chem. Lett.* **7** 2100–12
[256] Subotnik J E, Jain A, Landry B, Petit A, Ouyang W and Bellonzi N 2016 Understanding the surface hopping view of electronic transitions and decoherence *Annu. Rev. Phys. Chem.* **67** 387–417
[257] Behler J, Lorenz S and Reuter K 2007 Representing molecule-surface interactions with symmetry-adapted neural networks *J. Chem. Phys.* **127** 014705
[258] Behler J 2004 Dissociation of oxygen molecules on the Al(111) surface *PhD thesis* Technical University Berlin
[259] Behler J, Delley B, Lorenz S, Reuter K and Scheffler M 2005 Dissociation of O_2 at Al(111): The role of spin selection rules *Phys. Rev. Lett.* **94** 036104
[260] Thiel A and Köppel H 1999 Proposal and numerical test of a simple diabatization scheme *J. Chem. Phys.* **110** 9371–83
[261] Köppel H and Schubert B 2006 The concept of regularized diabatic states for a general conical intersection *Mol. Phys.* **104** 1069–79
[262] Maeda S, Ohno K and Morokuma K 2010 Updated branching plane for finding conical intersections without coupling derivative vectors *J. Chem. Theory Comput.* **6** 1538–45
[263] Kammeraad J A and Zimmerman P M 2016 Estimating the derivative coupling vector using gradients *J. Phys. Chem. Lett.* **7** 5074–9
[264] Gonon B, Perveaux A, Gatti F, Lauvergnat D and Lasorne B 2017 On the applicability of a wavefunction-free, energy-based procedure for generating first-order non-adiabatic couplings around conical intersections *J. Chem. Phys.* **147** 114114
[265] Christensen A, Faber F, Huang B, Bratholm L, Tkatchenko A, Müller K and Lilienfeld O 2017 Qml: A python toolkit for quantum machine learning https://github.com/qmlcode/
[266] Zhang Y, Maurer R J and Jiang B 2020 Symmetry-adapted high dimensional neural network representation of electronic friction tensor of adsorbates on metals *J. Phys. Chem.* C **124** 186–95
[267] Wang H, Zhang L, Han J and 2018 DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics *Comput. Phys. Commun.* **228** 178–84 W E
[268] Behler J and Parrinello M 2007 Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces *Phys. Rev. Lett.* **98** 146401
[269] Behler J 2011 Atom-centered symmetry functions for constructing high-dimensional neural network potentials *J. Chem. Phys.* **134** 074106
[270] Herr J E, Koh K, Yao K and Parkhill J 2019 Compressing physics with an autoencoder: Creating an atomic species representation to improve machine learning models in the chemical sciences *J. Chem. Phys.* **151** 084103
[271] Gilmer J, Schoenholz S S, Riley P F, Vinyals O and Dahl G E 2017 Neural message passing for quantum chemistry *Proc. of the 34th Int. Conf. on Machine Learning - Volume 70*, ICML'17 MLR.org pp 1263–72
[272] Unke O T and Meuwly M 2019 PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments and Partial Charges *J. Chem. Theory Comput.* **15** 3678–93
[273] Lubbers N, Smith J S and Barros K 2018 Hierarchical modeling of molecular energies using a deep neural network *J. Chem. Phys.* **148** 241715