

## Machine Learning for Electronically Excited States of Molecules

Julia Westermayr and Philipp Marquetand\*

Cite This: *Chem. Rev.* 2021, 121, 9873–9926

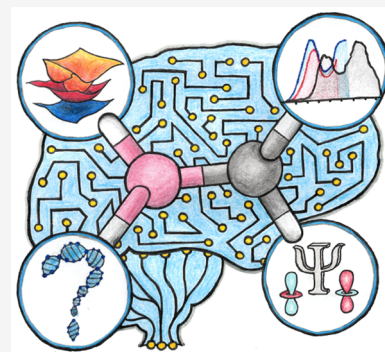
Read Online

ACCESS |

Metrics &amp; More

Article Recommendations

**ABSTRACT:** Electronically excited states of molecules are at the heart of photochemistry, photophysics, as well as photobiology and also play a role in material science. Their theoretical description requires highly accurate quantum chemical calculations, which are computationally expensive. In this review, we focus on not only how machine learning is employed to speed up such excited-state simulations but also how this branch of artificial intelligence can be used to advance this exciting research field in all its aspects. Discussed applications of machine learning for excited states include excited-state dynamics simulations, static calculations of absorption spectra, as well as many others. In order to put these studies into context, we discuss the promises and pitfalls of the involved machine learning techniques. Since the latter are mostly based on quantum chemistry calculations, we also provide a short introduction into excited-state electronic structure methods and approaches for nonadiabatic dynamics simulations and describe tricks and problems when using them in machine learning for excited states of molecules.



## CONTENTS

1. Introduction	9874	5.2.1. Basic Sampling Techniques and Existing Databases	9893
1.1. From Foundations to Applications	9874	5.2.2. Active Learning	9894
1.2. Scope and Philosophy of this Review	9875	5.3. Phase of the Wave Function	9896
2. General Background: From the Ground State to the Excited States	9875	5.3.1. Phase Correction of Adiabatic Data	9897
3. Quantum Chemical Theory and Methods	9877	5.3.2. ML-Based Internal Phase Correction	9898
3.1. Electronic Structure Theory for Excited States	9878	6. Application of ML for Excited States	9899
3.1.1. Wave Function Theory	9878	6.1. Parameters for Quantum Chemistry	9899
3.1.2. Density Functional Theory	9881	6.2. ML of Primary Outputs	9899
3.2. Bases	9882	6.3. ML of Secondary Outputs	9899
3.2.1. Adiabatic (Spin-Diabatic) Basis	9883	6.3.1. ML in the Diabatic Basis	9899
3.2.2. Diabatic Basis	9883	6.3.2. ML in the Adiabatic Basis	9901
3.2.3. Diagonal Basis	9884	6.4. ML of Tertiary Outputs	9904
3.3. Excited-State Dynamics Simulations	9884	6.5. ML-Assisted Analysis	9906
3.3.1. Quantum Nuclear Dynamics	9885	6.6. Open Questions and General Remarks on a Successful ML Method	9906
3.3.2. Mixed Quantum-Classical Molecular Dynamics	9886	7. Conclusion and Future Perspectives	9907
3.4. Dipole Moments and Spectra	9886	Author Information	9909
4. ML Models	9887	Corresponding Author	9909
4.1. ML Models: Type of Regressor	9887	Author	9909
4.1.1. Kernel Methods	9887	Notes	9909
4.1.2. Neural Networks	9888	Biographies	9909
4.2. Descriptors and Features	9889	Acknowledgments	9909
4.2.1. Molecule-wise Descriptors	9890	References	9909
5. Data Sets for Excited States	9891		
5.1. Choosing the Right Reference Method for Excited-State Data	9891		
5.2. Training Set Generation	9892		

Special Issue: Machine Learning at the Atomic Scale

Received: July 17, 2020

Published: November 19, 2020



## 1. INTRODUCTION

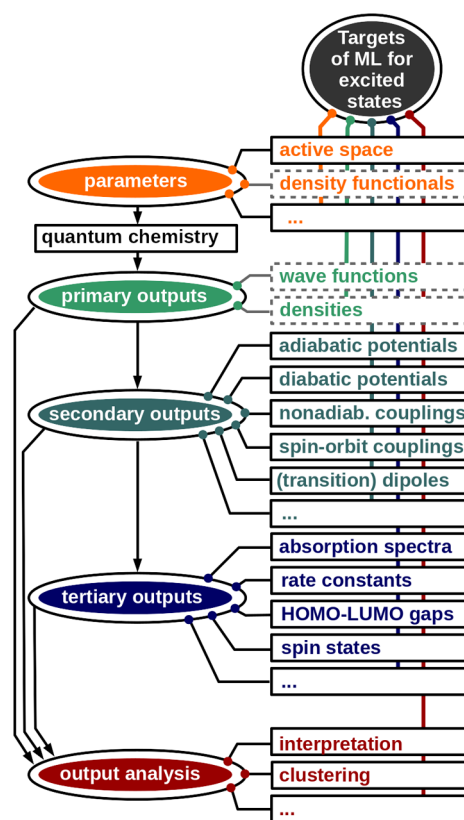
### 1.1. From Foundations to Applications

In recent years, machine learning (ML) has become a pioneering field of research and has an increasing influence on our daily lives. Today, it is a component of almost all applications we use. For example, when we talk to Siri or Alexa, we interact with a voice assistant and make use of natural language processing,<sup>1,2</sup> which was used to perform quantum chemistry calculations recently.<sup>3</sup> ML is applied for example for refugee integration,<sup>4</sup> for playing board games,<sup>5</sup> in medicine,<sup>6</sup> for image recognition<sup>7</sup> or for autonomous driving.<sup>8</sup> A short historical overview over general ML is provided in ref 9.

Recently, ML has also gained increasing interest in the field of quantum chemistry.<sup>10,11</sup> The power of (big) data-driven science is even seen as the “fourth paradigm of science”,<sup>12</sup> which has the potential to accelerate and enable quantum chemical simulations that were considered unfeasible just a few years ago.<sup>13</sup> The reason is, at least in theory, that ML models can learn any input–output relation and offer interpolations thereof at almost no cost while retaining the accuracy of the underlying reference data. With regard to quantum chemical applications, it allows decoupling of the expenses of quantum chemistry calculations from the application, such as dynamics simulations or the computation of different types of spectra. In general, the field of ML in quantum chemistry is progressing faster and faster. In this review, we focus on an emerging part of this field, namely, ML for electronically excited states. In doing so, we concentrate on singlet and triplet states of molecular systems since almost all existing approaches of ML for the excited states focus on singlet states and only a few studies consider triplet states.<sup>14–17</sup> We note that electron detachment or uptake further leads to doublet and quartet states, and even higher spin multiplicities, such as quintets, sextets, etc. are common in transition metal complexes, where an important task is to identify which multiplicity yields the lowest energy and is thus the ground state;<sup>17</sup> see, e.g., refs 18–21.

The theoretical study of the excited states of molecules is crucial to complement experiments and to shed light on many fundamental processes of life and nature.<sup>22–24</sup> For example, photosynthesis,<sup>25,26</sup> human vision,<sup>27,28</sup> photovoltaics,<sup>29–32</sup> or photodamage of biologically relevant molecules are a result of light-induced reactions.<sup>33–35</sup> Experimental techniques such as UV/visible spectroscopy or photoionization spectroscopy<sup>36–43</sup> lack the ability to directly describe the exact electronic mechanisms of photoinduced reactions. The theoretical simulation of the corresponding experiments can go hand-in-hand with experimental results and can provide the missing details of photodamage and -stability of molecules.<sup>42,44–68</sup> However, the computation of the excited states is highly complex and costly, and often necessitates expert knowledge.<sup>69</sup> As ML models have only recently been applied in the field of photochemistry, keeping track of the approaches is still possible, and this field is still in its initial stage.

Because of the multifaceted photochemistry of molecular systems, ML models can target this research field in many different ways, which are summarized in Figure 1. For example, the choice of relevant molecular orbitals for active space selections can be assisted with ML.<sup>71</sup> The fundamentals of



**Figure 1.** Targets of ML for the excited states of molecules (dashed: not yet achieved). All areas of excited-state quantum chemistry (QC) calculations can be enhanced with ML, ranging from input to primary outputs that are used in the computation of secondary outputs, which in turn are employed to calculate tertiary outputs. Analysis can be carried out at all stages. The classification is inspired by the one in ref 70.

quantum chemistry, e.g., to obtain an optimal solution to the Schrödinger equation or density functional theory, can be central ML applications. For the ground state, ML approximations to the molecular wave function<sup>72–80</sup> or the density (functional) of a system exist.<sup>70,80–89</sup> Obtaining a molecular wave function from ML can be seen as the most powerful approach in many perspectives, as any property we wish to know could be derived from it. Unfortunately, such models for the excited states are lacking and have yet been investigated only for a one-dimensional system,<sup>90</sup> leaving much room for improvement.

Most ML studies instead focus on predicting the output of a quantum chemical calculation, the so-called “secondary-output”.<sup>70</sup> Hence, they fit a manifold of energetic states of different spin multiplicities, their derivatives, and properties thereof. With respect to different spin states of molecular systems, only a few studies exist, which predict spins of transition metal complexes<sup>17,91</sup> or singlet and triplet energies of carbenes<sup>14</sup> of different composition or focus on the conformational changes within one molecular system<sup>15,92,93</sup> for the sake of improving molecular dynamics (MD) simulations. The energies of a system in combination with its properties, i.e., the derivatives, the coupling values between them, and the permanent and transition dipole moments,<sup>15,16,92–99</sup> can be used for MD simulations to study the temporal evolution of a system in the ground-state<sup>100–137</sup> and in the excited states.<sup>15,16,92–94,132,138–145,145–149,149–151</sup>

With energies and different properties, tertiary outputs can be computed, such as absorption, ionization or X-ray spectra,<sup>152–155</sup> gaps between highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO), or vertical excitation energies.<sup>156–159</sup>

In addition, quantum chemical outputs can also be analyzed or fitted in a direct way, e.g., reaction kinetics, as results of dynamics simulations can be mapped to a set of molecular geometries and can be predicted with ML models.<sup>160</sup> Excitation energy transfer properties can be learned,<sup>161,162</sup> and structure–property correlations can be explored to design materials with specific properties.<sup>18,31,32,77,133,154,163–170</sup>

## 1.2. Scope and Philosophy of this Review

ML has entered the research field of excited states relatively late, and it might seem that this research field is developing at a slower pace than the exploding field of ML for the electronic ground state.<sup>169,171–174</sup> Important reasons are in our opinion the complexity and high expenses of the underlying reference calculations and the associated complexity of the corresponding ML models, which might make it more suitable to say that ML for the excited states is developing at a similar pace, but toward a much more complex target. Simulation techniques to understand the excited-state processes are not yet viable for many applications at an acceptable cost and accuracy. Therefore, within this review, we also want to highlight the existing problems of quantum chemical approaches that might be solvable with ML and put emphasis on identifying challenges and limitations that hamper the application of ML for the excited states. The young age of this research field leaves much room for improvement and new methods.

This review is structured as follows:

- (1) After a general introduction,
- (2) we will start by discussing the differences between the ground-state potential energy hypersurfaces (PESs) and the excited-state PESs and will also emphasize the difference in their properties in [section 2](#).
- (3) We provide an overview of the theoretical methods that can be used to describe the excited states of molecules. In the forthcoming discussion, we will describe different reference methods with a view to their application in time-dependent simulations, namely, MD simulations.<sup>59,172</sup> It is worth mentioning that, unlike for the ground state, where a lot of different methods can provide reliable reference computations for training, choosing a proper quantum chemistry method for the treatment of excited states is a challenge on its own.<sup>175–177</sup> Many methods require expert knowledge, prohibiting their use further.<sup>178,179</sup> In addition, not any method can provide the necessary properties for any type of application. Subsequently, we aim to review the different flavors of excited-state MD simulations with a focus on nonadiabatic methods that have been enhanced with ML models lately.
- (4) After having provided the basic theoretical background on electronic structure theory and quantum chemical simulation techniques, we go on and summarize the basic ML models applied in studies with a focus on the excited states of molecules. The different types of ML models will help the reader to identify a proper model for a specific purpose. Advantages and disadvantages of certain regressor and descriptors are discussed.

- (5) The theoretical background on electronic structure theory and ML is followed by a discussion on how to generate a comprehensive yet compact training set for the excited states from the quantum chemistry data. We will summarize the existing approaches that are applied to create a full-fledged training set and put emphasis on the bottlenecks of existing methods that can limit also the application of ML. This will provide the reader with the knowledge about starting points for future research questions and clarify where method development is needed. It further provides the basis for the discussion of ML models for the excited states of molecular systems.
- (6) A summary of state-of-the-art ML methods for photochemistry follows. We will differentiate between single-state and multistate ML models and single-property and multiproperty ML models.<sup>95</sup> As mentioned before, ML models can tackle a quantum chemical calculation in many different ways; see [Figure 1](#). The different ML models will be classified in the ways they enhance quantum chemical simulations. Most approaches aim at providing an ML-based force field for the excited states, so we put particular emphasis on this topic. Lastly, the prospects of ML models to revolutionize this field of research and future avenues for ML will be highlighted.

Noteworthy, we focus on the excited states of molecules, as the excited electronic states in the condensed phase are challenging to fit and are thus often not explicitly considered in conventional approaches.<sup>180–185</sup> In solid state physics for example, the electronic states are usually treated as continua. The density of states at the Fermi level,<sup>186</sup> band gaps,<sup>187–189</sup> and electronic friction tensors<sup>125,190,191</sup> have been described with ML models to date, and especially the electronic friction tensor is useful to study the indirect effects of electronic excitations in materials.<sup>192–197</sup> Electron transfer processes as a result of electron–hole-pair excitations can be further investigated along with multiquantum vibrational transitions by discretizing the continuum of electronic states and fitting them (often manually) to reproduce experimental or quantum chemical data in a model Hamiltonian.<sup>183,198–203</sup> Yet, to the best of our knowledge, the excited electronic states in the condensed phase have not been fitted with ML. A recent review on reactive and inelastic scattering processes and the use of ML for quantum dynamics reactions in the gas phase and at a gas-phase interface can be found in [ref 204](#).

Besides the electronic excitations that take place in molecules after light excitation, ML models have successfully entered research fields, which focus on other types of excitations as well. Those are, for example, vibrational or rotational excitations giving rise to Raman spectra or infrared spectra,<sup>41,111,205–209</sup> nuclear magnetic resonance,<sup>210</sup> or magnetism,<sup>211,212</sup> which we will not consider in this review.

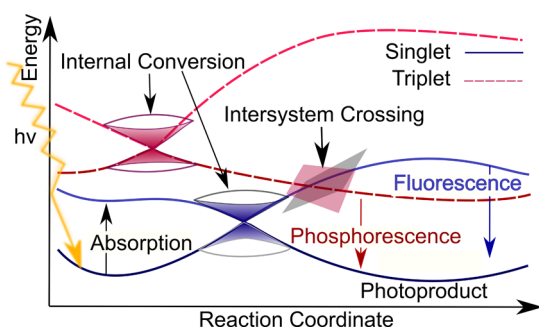
## 2. GENERAL BACKGROUND: FROM THE GROUND STATE TO THE EXCITED STATES

The chemistry we are interested in is not static, but rather depends to a large extent on the changes that matter undergoes. In this regard, it is more intuitive to study the temporal evolution of a system. Much effort has been devoted to develop methods to study the temporal evolution of matter in the ground state potential. As an example, physical functions can be obtained with conventional force fields, such as AMBER,<sup>213</sup> CHARMM,<sup>214</sup> or GROMOS.<sup>215,216</sup> The first ones already date back to the 1940s to 1950s. Such force fields

enable the study of large and complex systems, protein dynamics or binding-free energies on time scales up to a couple of nanoseconds.<sup>180,217–225</sup> However, their applicability is restricted by the limited accuracy and inability to describe bond formation and breaking. Novel approaches, such as reactive force fields exist, but still face the problem of generally low accuracy.<sup>226</sup>

The accuracy of *ab initio* methods can be combined with the efficiency of conventional force fields with ML models. The latter have shown to advance simulations in the ground state considerably and allow for the fitting of almost any input–output relation.<sup>100–134,137,172,227</sup> Accurate and reactive PESs of molecules in the ground state can be obtained with a comprehensive reference data set, which contains the energies, forces, and ground-state properties of a system under investigation. Proper training of an ML model then guarantees that the accuracy of the reference method is retained, while inferences can be made much faster. In this way, they allow for a description of reactions and can overcome the limitations of existing force fields.<sup>135,171,228–232</sup>

Regarding the excited states, processes become much more complex, and the computation of excited state PESs is far more difficult than the computation of the ground state PESs. Figure 2 gives an overview of the excited state processes that will be



**Figure 2.** Excited-state processes that can take place after excitation of a molecule by light. Absorption of light can make the molecule enter a higher electronic singlet state. Internal conversion to another state of same spin-multiplicity and/or intersystem crossing to a triplet state can take place. Radiative emission, i.e., fluorescence and phosphorescence, are possible reactions from an excited singlet and triplet state, respectively.

discussed within this review. As can be seen, several excited states of different spin multiplicity in addition to the ground state have to be accounted for, which feature different local minima, transition states, or saddle points and crossing points. Especially, the latter make a separate treatment of each electronically excited state inaccurate and lead to further challenges that prohibit the straightforward and large-scale use of many existing quantum chemical methods and consequently also existing ML models for the ground state.

As it is visible, processes usually start from a minimum in the electronic ground state (dark-blue line). When light hits a molecule and the incident light coincides with the energy gap between two electronic states, it can be absorbed, and higher electronic states can be reached when dipoles are allowed (here the second excited singlet state in light blue). Internal conversion between states of the same spin multiplicity and/or intersystem crossing between states of different spin multiplicity (here a transition to triplet states indicated as dashed red curves) can prevent the molecule from photodamage.

Nonradiative transitions usually take place on a sub-pico-second time scale. With respect to intersystem crossing, it was long believed that it happens on a longer time scale and is only possible if heavy atoms are part of the molecule.<sup>233,234</sup> However, this belief has been disproved, and today many examples of small molecules or transition metal complexes are known, which show ultrafast intersystem crossing.<sup>178,235–237</sup> In the case of nonradiative transitions, the energy is lost due to molecular vibrations, and the molecule relaxes back to the original starting point in the ground state. However, also photodamage can occur via such nonradiative transitions, where photoproducts can be formed, e.g., by bond breaking and bond formation. When nonadiabatic transitions are not taking place, radiative emission, i.e., fluorescence and phosphorescence, can happen on a much slower time scale, i.e., in the range of nano- to milliseconds.

For small system sizes, such as  $\text{SO}_2$ , highly accurate *ab initio* methods can be applied to describe the excited states, while more crude approximations have to be used for larger systems. The unfavorable scaling of many quantum chemical methods with the size of system under investigation requires this compromise between accuracy and system size. Crude approximations for systems that are larger than several hundreds of atoms become inevitable.<sup>44,178,238</sup>

Additionally, computations of the excited states suffer from being generally less efficient. To name only one central problem: The larger the system becomes, the closer the electronic states lie in energy, and the more excited-state processes can usually take place. The necessary consideration of an increasing number of excited states increases the already substantial computational expenses even more and restricts the use of accurate methods to systems containing only a few dozens of atoms in a reasonable amount of time with current computers. This increasing complexity makes not only the reference computations, but also the application of ML models for the excited states more complicated than for the ground state. At the same time, the application of ML models for the excited states might also be more promising, because higher speed-up can be achieved.

For the excited states, methods similar to force fields, like the linear vibronic coupling (LVC) approach,<sup>239,240</sup> are usually limited to small regions of conformational space and restricted to a single molecule. General force fields that are valid for different molecules in the excited states do not exist. Also the ML analogue, so-called transferable ML models, to fit the excited state PESs of molecules throughout chemical compound space are unavailable to date. Only recently, we have provided a first hint at transferability of the excited states by training an ML model on two isoelectronic molecules.<sup>241</sup> It is clear that an ML model, which is capable of describing the photochemistry of several different molecular systems, e.g., different amino acids or DNA bases of different sizes, is highly desirable. A lot remains to be done in order to achieve this goal, and yet, to the best of our knowledge, no more than a maximum of about 20 atoms and 3 electronic states with a distinct multiplicity have been fitted accurately with ML models (refs 15, 16, 92–94, 96, 132, 138–145, 146–149, 149–151, and 241).

Whether or not the excited states of a molecular system become populated depends on the ability of a molecule to absorb energy in the form of light, or more generally, electromagnetic radiation of a given wavelength. Usually, the so-called resonance condition has to be fulfilled; i.e., the energy

gap between two electronic states has to be equivalent to the photon energy of the incident light. Note however that also multiphoton processes can occur, where several photons have to be absorbed at once to bridge the energy difference between two electronic states.<sup>242–244</sup> Further, the absorption of light not only provides access to one, but most often to a manifold of energetically close-lying states. The number of states that can be excited is related to the range of photon energies that is contained in the electromagnetic radiation. This energy range is inversely proportional to the duration of the electric field, e.g., of a laser pulse, due to the Fourier relation of energy and time.<sup>245</sup> However, the energy range of the photons and the energy difference between the electronic states are not the only factors influencing the absorption of light, which gives rise to questions like: Is the molecule able to absorb light of a considered wavelength? Which of the excited states is populated with the highest probability?

An answer to these questions can be obtained from an analysis of the oscillator strengths for different transitions. In order to make an electronic transition possible, an oscillating dipole must be induced as a result of the interaction of the molecule with light. The oscillator strength,  $f_{ij}^{\text{osc}}$ , between two electronic states,  $i$  and  $j$ , is proportional in atomic units (a.u.) to the respective transition dipole moment,  $\mu_{ij}$ , and the respective energy difference,  $\Delta E_{ij}$ :<sup>246</sup>

$$f_{ij}^{\text{osc}} = \frac{2}{3} \Delta E_{ij} |\mu_{ij}|^2 \quad (1)$$

If the transition dipole moment between two states is zero, no transition is allowed. The reasons can be that a change of the electronic spin would be required, and the transition is thus spin forbidden. Another reason can be the molecular symmetry, leading to symmetry forbidden transitions. The latter are common in molecules that carry an inversion center, and transitions that conserve parity are forbidden.<sup>247</sup> An energetic state is called dark if the transition dipole moment is very small or zero. In contrast, a state is called bright if the transition dipole moment is large. Most often, studies that target the photochemistry of molecules focus on excitation to the lowest brightest singlet state, i.e., the state that absorbs most of the incident energy. The same is true for emission processes. While fluorescence is an allowed transition, phosphorescence is a spin forbidden process, i.e., a triplet-singlet emission in many cases.<sup>248</sup>

After an excitation process, the molecule is considered to move on the excited-state PESs and is expected to undergo further conversions. The excess of energy a molecule carries—as a result of the initial absorption of energy—is most often converted into heat, light, such as fluorescence or phosphorescence, or chemical energy. If the molecule returns to its original state, then the molecule is photostable. Otherwise, either photodamage, such as decomposition, or useful photochemical reactions including bond breaking/formation occur. In all cases, heat or light can be emitted, which can also be harnessed in light-emission applications.<sup>59,249–251</sup> With respect to photostability, ultrafast transitions, in the range of femto- to picoseconds ( $10^{-15}$ – $10^{-12}$  seconds) take place and lead the molecule back to the ground state. This means that the electronic energy is converted into vibrations of the molecule, and the molecule is termed hot. This heat is usually dissipated into the environment, a process that is often neglected in excited-state simulations due to the cost of describing surrounding molecules.

Radiationless transitions from one electronic state to another take place in so-called critical regions of the PESs. As the name already suggests, critical regions are crucial for the dynamics of a molecule, but are also challenging to model accurately. The critical points, where transitions are most likely to occur, are called conical intersections and are illustrated in Figure 2. At these crossing points, PESs computed with quantum chemistry can show discontinuities. These discontinuities can occur also in other excited-state properties and pose an additional challenge for an ML model when fitting excited-state quantities.

In addition to the aforementioned complications of treating a manifold of excited states, also the probability of a radiationless transition between them has to be computed somehow. This probability is usually determined by couplings between two approaching PESs. Between states of the same spin multiplicity, nonadiabatic couplings (NACs) arise, and spin-orbit couplings (SOCs) give rise to the transition probability between states of different spin multiplicities. These couplings are intimately linked to the excited-state PESs and therefore should also be considered with ML. However, only a handful of publications describe couplings with ML,<sup>15,92,94–96,140,145,146,149,252</sup> which highlights the difficulty of providing the necessary reference data as well as the challenges of accurately fitting them. New methods are constantly needed to further enhance this exciting research field.

### 3. QUANTUM CHEMICAL THEORY AND METHODS

In this section, we present some key aspects of quantum theory for excited states, which is the basis of any study focusing on ML for excited states of molecules. We do so because (i) the outcome of the corresponding calculations serve as training data for ML, leaving quantum chemistry and ML thus inseparably connected in many cases and (ii) to clarify the employed nomenclature. We will discuss electronic structure theory (section 3.1), the different bases (section 3.2), i.e., the diabatic, adiabatic, and diagonal bases, the computation of excited-state molecular dynamics simulations (section 3.3) with different flavors of quantum nuclear dynamics and mixed quantum-classical methods, along with the computation of dipole moments and spectra. Experts on these topics may skip directly to section 4, which focuses on ML methods.

In the following, we provide a description of the differences of excited-state computations to calculations for the electronic ground state and the challenges that arise due to the treatment of a manifold of excited states. These challenges also point to issues that are problematic for ML. These explanations will provide the groundwork to evaluate different quantum chemical methods for their use to generate a training set for ML and to use it for different types of applications, such as excited-state MD simulations. Naturally, we can only provide a general idea of this field and refer the interested reader to pertinent textbooks and reviews, such as refs 178, 253–264.

In order to follow a consistent notation within this review, we try to explain all basic concepts with notations that are frequently used in the literature. Currently, a zoo of different notations for the same property can be found. For example, the NACs, or derivative couplings, are sometimes referred to as so-called interstate couplings, i.e., couplings between two states multiplied with the corresponding energy gap between those two states,<sup>144</sup> while in other works interstate couplings refer to off-diagonal elements of the Hamiltonian in another basis,

where the potential energies are no eigenvalues of the electronic Schrödinger equation. We want to avoid a confusion of different notations and thus provide a consistent definition below. For the excited states, a number of different electronic states are required. Throughout this review, we adopt the following labeling convention for different electronic states: The lower case Latin letters,  $i, j$ , etc. will be used to denote different electronic states. The abbreviations  $N_S$ ,  $N_M$ , and  $N_A$  will indicate the number of states, molecules, and atoms, respectively.

The foundation for the following sections is a separation of electronic and nuclear degrees of freedom, which is based on the work of Born and Oppenheimer.<sup>265</sup> However, the famous Born–Oppenheimer approximation is later (partly) lifted, and the coupling between electrons and nuclei is taken into account in nonadiabatic dynamics simulations.

### 3.1. Electronic Structure Theory for Excited States

The main goal when carrying out an electronic structure calculation is usually to compute the potential energy and other physicochemical properties of a compound. We distinguish between two overarching theories to achieve this goal: wave function theory (WFT) and density functional theory (DFT) as outlined, e.g., by Kohn in his Nobel lecture.<sup>266</sup>

The basis of WFT, as for any electronic structure calculation, is the electronic Schrödinger equation<sup>267,268</sup> with the electronic Hamilton operator,  $\hat{H}_{el}$ , and the  $N$ -electron wave function  $\Psi_i(\mathbf{R}, \mathbf{r})$  of electronic state  $i$ , which is dependent on the electronic coordinates  $\mathbf{r}$  and parametrically dependent on the nuclear coordinates,  $\mathbf{R}$ :

$$\hat{H}_{el}(\mathbf{R}, \mathbf{r})|\Psi_i(\mathbf{R}, \mathbf{r})\rangle = E_i|\Psi_i(\mathbf{R}, \mathbf{r})\rangle \quad (2)$$

From the wave function, i.e., the eigenvector of this eigenvalue equation, any property of the system under investigation can be derived. How to solve the electronic Schrödinger equation exactly to obtain the potential energy of an electronic state  $i$ ,  $E_i$ , is known in theory. However, from a practical point of view, the computation is infeasible for molecules that are more complex than for example  $H_2$ ,  $He_2^+$ , and similar systems.<sup>269</sup> In order to make the computation of larger and more complex systems viable, approximated wave functions are introduced.

In contrast to WFT, DFT reformulates the energy of a system in terms of the ground state electron density rather than the  $N$ -electron wave function, and the energy is expressed as a functional thereof. The advantage of DFT is a rather high accuracy at a rather low computational cost. If DFT is applied properly, it is considered as one of the most efficient ways to obtain reliable and reasonably accurate results of molecules up to hundreds of atoms. In solid state physics, DFT is even the workhorse of most studies aiming to describe ground state properties.<sup>270</sup> However, the problem is that the equations to be solved are unknown. The missing piece is the exact exchange–correlation functional of a system. To date, researchers have come up with many different approximations to this functional that can be used to treat specific problems, but a universal functional capable of describing different problems equally accurately has not yet been found. Moreover, there is no systematic way to improve a density functional. The results obtained with DFT therefore critically depend on the choice of the functional.<sup>269,271</sup>

In the following sections, we will describe both theories and focus on the excited states of molecules. We will start to cover *ab initio* methods, which means that they are derived from first-principles without parametrization. We mention the basic underlying concepts here because we believe them to be essential in order to generate training data and carry out ML for excited states. Furthermore, these methods present starting points for an ML approximation to the excited-state wave function or density, which is still lacking, to the best of our knowledge. Nevertheless, such an ML model would be extremely powerful and could provide a solution to many existing quantum chemical problems.

**3.1.1. Wave Function Theory.** The basis of all discussed *ab initio* methods is the Hartree–Fock method. The  $N$ -electron wave function is represented by a single Slater determinant,  $\phi_0$ , which makes  $N$  coupled one-electron problems out of the  $N$ -body problem. This Slater determinant is the antisymmetric product of one-electron wave functions, the spin orbitals, which can be atomic, molecular, or crystal orbitals, depending on the system. In the case of molecular (or also crystal) orbitals, they are usually expanded as a linear combination of atomic orbitals, where the expansion coefficients are optimized during the calculation. In order to do so efficiently, the atomic orbitals are themselves expanded with the help of a basis set. The  $N$ -electron wave function is therefore obtained as a double expansion. Two approximations are applied, which is the use of a finite basis set to represent the atomic orbitals and in turn also the molecular orbitals on the one hand and the use of a single Slater determinant on the other hand. This usually gives a poor description of a system under investigation, due to a lack of electronic correlation.

Electronic correlation describes how much the motion of an electron is influenced by all other electrons. Since the Hartree–Fock method can be seen as a mean-field theory, where an electron “feels” only the average of the other electrons, correlation is quantified by the correlation energy, which is the difference between the Hartree–Fock energy and the exact energy of a system.

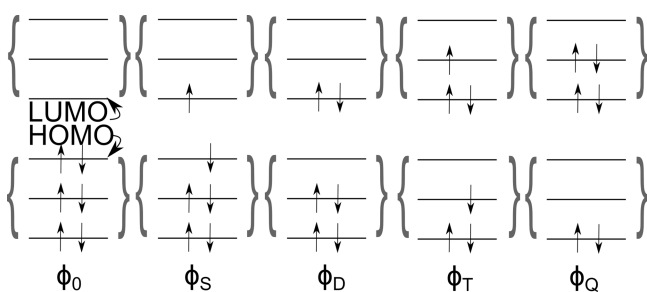
Unsurprisingly, all further discussed quantum chemical methods aim at improving the Hartree–Fock method. They can be seen as different flavors of the same solution to the problem: They all include more determinants in one way or another. Accordingly, the wave function is expanded as a linear combination of determinants, where a determinant consists of molecular orbitals, which are expanded in atomic orbitals. This ansatz contains two types of coefficients that can be optimized, the ones for the determinants and the ones yielding the molecular orbitals. If the latter are kept the same for different determinants, we speak of a single-reference wave function. If both types of coefficients are adapted, we speak of a multireference wave function. Similarly, the electron correlation is also divided into two parts, termed dynamic correlation and static correlation. Single-reference methods improve on the dynamic correlation, while a multireference wave function allows for static correlation. However, the separation is not so strict, as can be seen by the following fact: Both the aforementioned single-reference variant and the multireference variant become equivalent when including an infinite number of terms and deliver the exact solution to the Schrödinger equation if also an infinite basis set is used.

**Configuration Interaction.** In the case of single-reference methods, the orbitals obtained from the reference calculation (usually Hartree–Fock) are kept fixed. Since usually more

orbitals than the number of electrons in the system are calculated, the possibility of constructing different Slater determinants from these orbitals exist, which can be used for expanding the actual wave function:<sup>272,273</sup>

$$|\Psi_i\rangle = \sum_I c_{iI} |\phi_I\rangle \quad (3)$$

Each Slater determinant is weighted by a coefficient,  $c_{iI}$ . These coefficients can be obtained variationally by minimizing the total energy under the constraint of fixed orbitals, ending up in the configuration interaction (CI) methods.  $\phi_0$  is the reference, Hartree–Fock, wave function. In principle, the exact solution can be obtained by considering all possible Slater determinants in combination with a complete basis set. The use of all possible configurations is called full-CI and represents the case when all electrons are arranged in all possible ways. Practicable methods require truncation; e.g., CIS (CI singles) or CISD (CIS and doubles) are frequently used, where only single excitations or additionally double excitations are accounted for, respectively. Figure 3 gives a schematic



**Figure 3.** Different arrangements of electrons in molecular orbitals giving rise to the configuration interaction (CI) method. Inclusion of excited configurations in addition to the ground-state, reference determinant,  $\phi_0$ , allows one to go beyond the Hartree–Fock method. Electrons are excited into higher electronic orbitals, and Slater determinants are indicated using the letters S, D, T, and Q, which refer to single, double, triplet, and quadruple excitations.

overview of the improvements of CI that one can apply. A huge advantage of these methods is that how to obtain the exact solution is known and that they are systematically improvable. However, truncated CI does not scale correctly with the system size and is therefore not size-extensive and also not size-consistent (i.e., the energy of two fragments A and B at large distance computed together,  $E(A+B)$ , is not equal to the sum of the energies of the fragments from separate calculations,  $\neq E(A) + E(B)$ ).<sup>274</sup>

The CI scheme can be employed to improve the ground-state wave function by mixing the Hartree–Fock determinant and determinants of different electron configurations. In the same way, also wave functions of excited states can be computed. Then, the coefficients,  $c_{iI}$ , are optimized for higher eigenvalues of the electronic Hamiltonian instead of the first one. Beginners in the field then often get confused by terms such as single excitation in comparison to the first excited state. A single excitation determinant (see Figure 3) can be part of the wave function for the first excited state but can also be a part of the ground-state wave function.

**Electron Propagator Methods.** Another class of methods that we shortly want to mention here are electron propagator methods that are based on one electron Green's function and are a variant of perturbation theory schemes. One popular

method that is based on Green's function one electron propagator approach is the algebraic diagrammatic construction scheme to second-order perturbation theory (ADC(2)).<sup>275</sup> ADC(2) is a single-reference method and can be used to efficiently compute excited states of molecules. It offers a good compromise between computational efficiency and accuracy, while being systematically improvable (higher order variants such as ADC(2)-x or ADC(3) exist). The time evolution of a systems polarizability is obtained by applying the polarization propagation, which contains information on a system's excited states.<sup>272,276–279</sup> The ground-state energy of ADC(2) is based on Møller–Plesset perturbation theory of second order,<sup>280,281</sup> MP2, where the latter can formally be shown to include double excitations for the improvement of Hartree–Fock; see ref 272. The dependence of ADC(2) on MP2 gives rise to instabilities in regions, where excited states come close to the ground state, or homolytic dissociation takes place. The excited states of bound molecules are described with reasonable accuracy. Compared to multireference CI methods (see below), the black box behavior of ADC(2) is a clear advantage.<sup>275</sup>

**Coupled Cluster.** The current gold standard of ab initio methods for the ground state is the family of coupled cluster (CC) methods. CC is often referred to as the size-extensive and size-consistent version of CI. The different electronic configurations accounting for single or double excitations (such as in CIS and CISD for example) are obtained by applying an excitation operator,  $\hat{T}$ :<sup>282</sup>

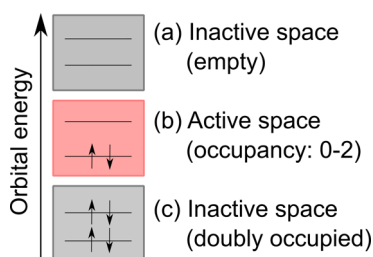
$$|\Psi_{CC}\rangle = e^{\hat{T}} |\phi_0\rangle = \left( \hat{T} = 1 + \hat{T} + \frac{1}{2!} \hat{T}^2 + \frac{1}{3!} \hat{T}^3 + \dots \right) |\phi_0\rangle \quad (4)$$

Similarly to CI, this operator can be truncated. If  $\hat{T} = \hat{T}_1 + \hat{T}_2$ , single and double excitations are accounted for. When using the same number of determinants, CC usually converges faster than CI.

Excited states can be computed in a single-reference approach by equation-of-motion-CC (EOM-CC), where the excited-state wave function is written as an excitation operator times the ground-state wave function. For further details, see, e.g. refs 283 or 284.

**CASSCF.** The problem of missing static correlation in the Hartree–Fock approach is tackled by a multireference ansatz for the wave function. Not only coefficients, but also orbitals are optimized.<sup>271</sup> This treatment is important for many excited-state problems, but also some transition metal complexes in their ground state, transition states, or homolytic bond-breaking with the dissociation of the  $N_2$  molecule being a notoriously difficult example.<sup>285,286</sup> An accurate ML training set for many chemical problems in the excited states often calls for such methods.

The multiconfigurational self-consistent field (MCSCF) method can be seen as the multireference counterpart to the Hartree–Fock method.<sup>287</sup> One of the most popular variants of MCSCF methods is the Complete Active Space SCF (CASSCF),<sup>288,289</sup> where important atomic orbitals and electrons are selected giving rise to an active space. An example is shown in Figure 4. According to this scheme, the orbitals are split into an inactive, doubly occupied part, an active part, and an inactive, empty part. Within the active space, a full CI (FCI) computation is carried out. The active space has to be chosen manually by selecting a number of active electrons and active orbitals. CASSCF is no black box



**Figure 4.** Electrons and orbitals of an arbitrary system to exemplify the active space needed for many multireference methods. (a) The highest, not considered, molecular orbitals are inactive and always empty. (c) The lowest, not considered, molecular orbitals are always doubly occupied. (b) The active space is shown with two active electrons in two active orbitals.

method, and a meaningful active space selection is the full responsibility of the user. As an advantage, CASSCF can describe static correlation well, which is necessary in systems with nearly degenerated configurations with respect to the reference Slater determinant. For completeness, state-averaging (i.e., SA-CASSCF) is most often applied, where states belonging to the same symmetry are averaged. Another variant of MCSCF methods is restricted active space SCF (RASSCF), which is very similar to CASSCF, but within RASSCF the active space is restricted and no FCI computation is carried out.<sup>272</sup>

**DMRG.** As an alternative to deal with large active spaces, the density matrix renormalization group (DMRG) can be used.<sup>290</sup> A DMRG-SCF calculation is similar to a CASSCF calculation, but instead of a FCI solution of the active space, an approximated solution with DMRG is obtained to avoid the exponential scaling of the computational costs with the number of active orbitals.<sup>291–296</sup> Very recently, transcorrelated DMRG (tcDMRG) was introduced for strongly correlated systems.<sup>297</sup>

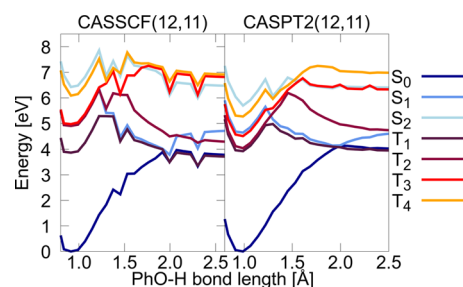
**MR-CI.** Even higher accuracy can be obtained with multireference CI methods,<sup>253,298,299</sup> such as MR-CISD, that additionally add single and double excitations out of the active space and are therefore based on CASSCF wave functions. With this approach, electronic correlation, i.e., static and dynamic correlation, can be treated.

**CASPT2.** Alternatively, complete-active-space perturbation theory of second order, CASPT2,<sup>300–302</sup> can correct electronic correlation effects via treating multireference problems with perturbation theory. This variant of multireference perturbation theory methods uses the CASSCF wave function as the zeroth order wave function. CASPT2 can be applied to each state separately (single-state (SS)-CASPT2) or correlated states can be mixed at second order resulting in a multistate perturbation treatment (MS-CASPT2).<sup>300–302</sup> Other perturbation approaches for multireference problems exist, like the *n*-electron valence state perturbation theory (NEVPT2).<sup>303–305</sup>

**MRCC.** In addition to multireference methods based on CI, multireference variants of CC approaches exist. A relatively efficient implementation is for example the Mk-MRCC approach of Mukherjee and co-workers<sup>306</sup> or the Brillouin-Wigner approach,<sup>307</sup> which is however not size extensive. Noticeably, the development of multireference CC approaches is a rather young research field compared to other excited-state methods, and the computation of properties and forces is not well explored. Many studies therefore focus on the simulation of energies of low-lying states with MRCC methods. Additionally, such methods suffer from algebraic complexity

and numerical instabilities. Interested readers who seek for a more extensive summary of existing MRCC methods are referred to refs 253, 308, and 309.

**Challenges.** The probably biggest drawback of the aforementioned multireference methods is that their protocols are very demanding. Finding a proper active space is a tedious task that often requires expert knowledge. Too small active spaces can lead to inaccurate energies, and problems with so-called intruder states are common. Those are electronic states that are high in energy at a reference molecular geometry, but become very low in energy at another molecular geometry, that is visited along a reaction coordinate. The active space then changes along this path. This behavior can result in inconsistent potential energies. In the case of CASPT2, the configurations of intruder states can lead to large contributions in the second-order energy, making the assumption of small perturbations invalid. Especially for describing molecular systems with many energetically close-lying states and for the generation of a training set for ML, such inconsistencies are problematic. Figure 5 shows an example of potential energy



**Figure 5.** Potential energy curves of the three lowest singlet ( $S_0$ – $S_2$ ) and the four lowest triplet state ( $T_1$ – $T_4$ ) of the amino acid tyrosine along the O–H bond length of the hydroxy group located at the phenyl ring (Ph–OH) computed with CASSCF(12,11)/ano-rcc-pVDZ and CASPT2(12,11)/ano-rcc-pVDZ.<sup>311</sup>

curves of three singlet states and four triplet states of tyrosine computed with (a) CASSCF(12,11) and (b) CASPT2(12,11), where 12 refers to the number of active electrons and 11 to the number of active orbitals. We used OpenMolcas<sup>310</sup> to compute an unrelaxed scan along the reaction coordinate, which is a stretching of the O–H bond located at the phenyl-ring of tyrosine.

Intruder states are no exception. Actually, they are quite common in small- to medium-sized organic molecules. A large enough reference space can mitigate this problem, but makes computations almost infeasible. The computational costs increase exponentially with the number of active orbitals. In many cases, the improved accuracy due to a larger active space cannot justify the considerably higher expenses. At its best and with massively parallel simulations, an active space of about 20 electrons in 20 orbitals can be treated,<sup>312</sup> which is impracticable for many applications, such as dynamics simulations. For medium-sized molecules, the active space that would be required for a given simulation might even be way too large to be feasible for calculations in a static picture. With respect to ML, a model that can detect such regions along reaction coordinates would be very helpful. Indeed, the dipole moment or nonadiabatic couplings are commonly used to identify a change in a character of a state. Monitoring of these properties along reaction coordinates could potentially help to identify such regions. To the best of our knowledge,



such a model does not yet exist. What is known so far is that ML can provide a smooth interpolation of such cusps in energy potentials as long as the density of electronic states is not too high.<sup>94,171</sup>

Worth mentioning at this point are also Rydberg states, which often need to be considered in small- to medium-sized molecules. Rydberg states can be strongly interlaced with valence excited states. In such cases, the active space needs to be large enough to treat both the valence and Rydberg molecular orbitals. Additionally, the one electron basis set should be flexible enough to describe both types of orbitals. This increases the computational costs additionally. More details on the inclusion of Rydberg states in simulations can be found in refs 313–316. Especially in such cases, an ML representation of the electronic wave function to reduce the computational time would be highly beneficial.

A promising tool to eliminate the complex choice of active orbitals is autoCAS.<sup>317–319</sup> It provides a measure of the entanglement of molecular orbitals that is based on DMRG. In principle, there is no prerequisite for the active space selection. If possible, the reference space should be the full valence in order to identify the relevant orbitals and electrons. If the full valence space is too large for a DMRG-CI calculation, one or several smaller chemically sensible reference active spaces should be selected to be able to analyze the importance of all orbitals.<sup>317–319</sup> As an alternative, ML can be used to determine an active space.<sup>71</sup>

**3.1.2. Density Functional Theory.** A complementary view on how to obtain the energy of a system is provided by DFT. DFT dates back to 1964, when it was formulated by Hohenberg and Kohn<sup>320</sup> entirely in terms of the electron density,  $\eta(\vec{r})$ . A one-to-one correspondence between this density and an external potential,  $v(\vec{r})$  exists and the potential acts on the electron density. The energy can be formulated in terms of a universal functional,  $F[\eta(\vec{r})]$ , of the electron density, which is independent of the external potential. In this way, the energy of a system's ground state can be computed with the following equation:

$$E[\eta(\vec{r})] = \int v(\vec{r})\eta(\vec{r}) d\vec{r} + F[\eta(\vec{r})] \quad (5)$$

The most widely used implementations of DFT rely on the Kohn–Sham approach.<sup>321</sup> In fact, Kohn–Sham DFT is so successful that it is often simply referred to as DFT. In this approach, an auxiliary wave function in the form of a Slater determinant is employed. Since a single Slater determinant is the exact solution for a system of noninteracting electrons, this DFT approach can be seen as describing a system of noninteracting electrons that are forced to behave as if they were interacting. The latter effect can be achieved only by an unknown modification of the Hamiltonian or rather of the aforementioned functional. In other words, a Slater determinant as wave function ansatz is exact, but the Hamiltonian can only be approximated, in contrast to Hartree–Fock, where the true electronic Hamiltonian is used, but the Slater determinant is only an approximate wave function.

The functional  $F[\eta(\vec{r})]$  can be separated into Coulombic interactions and a non-Coulombic part. The latter can further be divided into two terms: the kinetic energy of the noninteracting electrons and the exchange–correlation part, which describes the interaction of electrons and thus also corrects the kinetic energy by the difference of the real kinetic energy and the kinetic energy of the fictitious system of

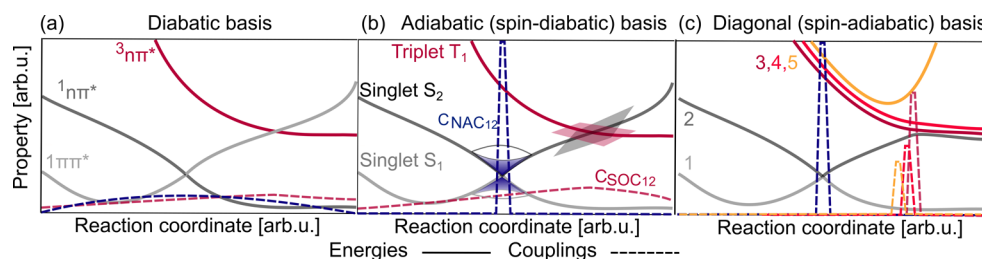
noninteracting electrons. The exchange–correlation functional is the part of DFT that is unknown, and finding the exchange–correlation functional remains the Holy grail of DFT.

In principle, if the exact functional was known, the exact ground-state energy of a system could be computed. Unfortunately, it is not known, and the success of a DFT calculation critically depends on the approximation that is used to the unknown exchange–correlation functional.

**Excited States.** As explained above, the electron density is computed from a single reference Kohn–Sham wave function, i.e., the one of noninteracting electrons with the density of the real system. This single-reference wave function makes DFT a single-reference method. In fact, most failures of DFT are a consequence of an improper description of static correlation.<sup>271</sup> In order to describe excited states, the time-dependent (TD) version of DFT, namely, TDDFT, can be used. The foundation of this theory was laid in the 1980s with the Runge–Gross theorems,<sup>322</sup> which can be regarded as analogies to the Hohenberg–Kohn theorems. They are based on the assumption that a one-to-one correspondence exists also between a time-dependent potential and a time-dependent electron density in this potential. A system can therefore be completely described by its time-dependent density. Also in the time-dependent case, the variational principle for the density is proposed.

The most widely used approach of TDDFT is linear response TDDFT (LR-TDDFT). Again, often TDDFT is used as a synonym for LR-TDDFT due to its extensive use. Within this theory and the KS approximation, no time dependent density is necessary to compute excitation energies and excited state properties. Linear response theory can be directly applied to the ground state density.<sup>323,324</sup> Casida's formulation of this theory is the most popular one and gives rise to random-phase approximation pseudoeigenvalue equations, which are also known as the Casida equations. Within the adiabatic approximation, they are implemented efficiently in many existing electronic structure programs. The Tamm–Dancoff approximation<sup>325,326</sup> further simplifies the equations to an eigenvalue problem, resulting in the counterpart to CIS.<sup>327</sup> Especially in cases when the time evolution of a system is studied, the Tamm–Dancoff approximation is beneficial since it leads to more stable computations close to critical regions of the PES.<sup>269,328</sup>

**Advantages and Disadvantages.** The advantage of LR-TDDFT is its computational efficiency. The reasonable accuracy if a proper functional is chosen makes this approach often the method of choice to study the photochemistry of medium-sized to large and complex systems, which are not feasible to treat with costly multireference WFT based methods.<sup>253,329,330</sup> Shortcomings of LR-TDDFT are the incorrect dimensionality of conical intersections, which are, however, one of the most important regions during non-adiabatic MD simulations.<sup>331–336</sup> The incorrect dimensionality of conical intersections with standard TDDFT implementations leads to a qualitatively incorrect description of such critical regions. The missing couplings can be corrected for example with the CI-corrected Tamm–Dancoff approximation<sup>337</sup> or the hole–hole Tamm–Dancoff approximation,<sup>338</sup> which can recover the missing couplings and provide correct dimensionality at conical intersections. Alternatively, an incorporation of the spin-restricted ensemble-referenced Kohn–Sham method into the tight-binding TDDFT ap-



**Figure 6.** (a) Example of three potential energy curves ordered by their character along with respective potential couplings between different states shown by dashed lines. (b) Two singlets ( $E_i$  and  $E_j$ ) and one triplet state ( $E_k$ ) including coupling values (with vectorial properties,  $C_{NAC_{ij}}$  shown by their norm) in the adiabatic basis, in which the triplet state crosses singlet states. (c) The diagonal, or spin-adiabatic, basis, in which all states are ordered by their energy and are spin-mixed. Kinetic couplings are shown by their norm. Note that the ground state is not shown. Potential energy curves are represented using solid lines and couplings using dashed lines.

proach<sup>339</sup> can be used to describe conical intersections with reasonable accuracy.

In addition, one should be aware that by definition, double excitations cannot be accounted for with LR-TDDFT. The computation of double excitations can be achieved by using a frequency dependent exchange kernel, which is known as dressed TDDFT.<sup>340,341</sup> Alternatively, spin-flip TDDFT<sup>342–344</sup> can be used, where a triplet state is taken as a reference state, and single excitations are treated with a flip in the electron's spin. However, spin-contamination is quite common within these methods. In general, the description of double excitations from a multireference state would be more favorable, although spin-flip TDDFT is often considered to be a multireference method. In order to compute specific orbital occupations and consequently excitations and charge-transfer states, an alternative approximation exists, which is known as the  $\Delta$ -SCF approach. In this theory, the electrons are forced into specific KS orbitals. The SCF is applied to converge the energy with respect to this configuration.<sup>345–347</sup> Other multireference variants of TDDFT exist too. However, their description is beyond the scope of this review, and we refer the reader to recent reviews covering this topic in much more detail.<sup>253,348</sup> The accuracy of (TD)DFT simulations for the ground state<sup>349</sup> or excitation energies and absorption spectra of organic molecules could be improved by ML corrections, which were obtained from the genetic algorithm and NN approach (GANN),<sup>350</sup> support vector machines,<sup>351</sup> or AdaBoost ensemble correction models<sup>352</sup> for example.

Last but not least, we briefly want to discuss the most critical part of a DFT calculation, which is the proper choice of the exchange-correlation functional. In the case of excited states, the treatment of valence excitations, Rydberg states and long-range charge transfer excitations on the same footing are highly problematic. While hybrid (meta-) generalized gradient approximation (GGA) or range-separated hybrid functionals<sup>353</sup> are for example well suited for vertical excitations and the latter also for Rydberg states, global hybrid meta GGA or range-separated hybrid GGA functionals are better to describe charge transfer.<sup>269,354</sup> Most often, functionals are accurate for one specific problem, but they fail to describe others. Although much effort has been devoted to develop functionals, finding a universal functional for DFT is still far from being achieved.<sup>180,253,269,328</sup> ML could be particularly powerful to advance DFT and TDDFT in this regard. For example, models exist to predict the energy of a system based on self-consistent charge densities or external potentials. Kohn–Sham DFT can be circumvented, and density functionals can be constructed

from ML models. Besides the density-potential, energy-density maps, and whole functionals can be learned.<sup>82,355–357</sup>

In summary, it should be stressed that, in general, there is not only one single solution to a particular problem, but that many possible ways can be considered, which lead to an equivalent description of a particular problem. Considering the excited states of molecules, it should be mentioned that it is of utmost importance to think carefully about the photochemical processes that may occur in order to find the most appropriate method for most of the assumed reactions. It often happens that within the same molecular system, one method can describe a certain photochemical reaction quite well, while another reaction can be described better with another method. However, the mixing of methods is not practicable for standard applications. Recently, studies on ML models have emerged that combine the different strengths of several methods, e.g.  $\Delta$ -learning techniques<sup>358,359</sup> or transfer learning.<sup>360</sup> These methods could be well-suited solutions for many future applications to overcome the current limitations of existing quantum mechanical methods for the excited states. Even more than for ground state properties, the quality of the excited states depends critically on the ability of a method to describe the different possible reaction—as a consequence of the larger accessible configuration space of a molecular system. Even for medium-sized systems, it should be clear that a suitable method may already be computationally impracticable, and a balance between accuracy and computational effort has to be found.

### 3.2. Bases

The potentials computed with the aforementioned methods for different nuclear geometries can be represented in different bases, which are connected by unitary transformations. An example of five states in different bases is given in Figure 6. Note that often a system in a certain basis is also referred to as being in a certain picture or representation; here we will not use the term representation in order to not confuse the reader with molecular representations used in ML. As it is visible in the figure, we focus on three types of bases: (a) the diabatic basis, (b) the adiabatic (spin-diabatic) basis, i.e., the direct output of standard electronic structure programs, (c) the diagonalized version of (a) and (b), i.e., the spin-adiabatic basis. Throughout the literature, different names are given to these bases, which are summarized in Table 1. They stem from a partition of the total wave function into a sum of electronic and nuclear contributions, which can be written for all bases as

**Table 1. Commonly Used Names of Bases for the Excited-State Potential Energy Surfaces<sup>a</sup>**

a	b	c
diabatic	adiabatic	diagonal
crude diabatic	spin-diabatic	spin-adiabatic
spectroscopic	MCH	field-adiabatic
quasi-diabatic	field-free	field-dressed

<sup>a</sup>Based on refs 239, 256, and 362–365. The labels a, b, and c are consistent with Figure 6.

$$\Psi(\mathbf{r}, \mathbf{R}, t) = \sum_i \psi_i^{\text{basis}}(\mathbf{r}, \mathbf{R}) \chi_i^{\text{basis}}(\mathbf{R}, t) \quad (6)$$

In principle, the total wave function can be expanded in infinitely many different bases. The electronic part  $\psi_i^{\text{basis}}(\mathbf{r}, \mathbf{R})$  corresponds to the eigenfunctions of the electronic Hamiltonian only for one of the bases (namely, the one from column B of Table 1). Associated with these functions are the corresponding potentials, depicted for a model system in Figure 6. Note that a different approach is taken in the exact factorization method,<sup>361</sup> where the total wave function is expanded only in a single product, i.e., without the sum in eq 6, giving rise to only one (time-dependent) potential.

**3.2.1. Adiabatic (Spin-Diabatic) Basis.** The direct output of an electronic structure calculation usually provides the eigenenergies and eigenfunctions of the electronic Hamiltonian. In many cases, only one spin multiplicity is calculated. If this procedure is repeated along a nuclear coordinate, potential curves result that are termed adiabatic. Adiabatic means “not going through” (from greek a = not, dia = through, batos = passable), and, indeed, the potentials never cross when considering one multiplicity. This situation is schematically illustrated in Figure 6b for singlet  $S_i$  and singlet  $S_j$ .

Within one multiplicity,  $3N_A$ -dimensional adiabatic PESs are obtained that are strictly ordered by energy. Hence, the states are usually denominated with the first letter of the multiplicity and a number as subscript, e.g.,  $S_0$ ,  $S_1$ , etc. For states of the same multiplicity, critical points and seams exist. These regions of the PESs are referred to as conical intersection (seams), in which the corresponding states become degenerate. Such features make adiabatic PESs nonsmooth functions of the atomic coordinates, which make them difficult to predict with the intrinsically smooth regressors of ML. At a conical intersection, the approaching potential energy curves form a cone, and the NACs, denoted as  $C_{\text{NAC}_{ij}}$ , between them show singularities as a result of the inverse proportionality to the vanishing energy gap:<sup>298,366</sup>

$$C_{\text{NAC}_{ij}} \approx \left\langle \Psi_i \left| \frac{\partial}{\partial \mathbf{R}} \Psi_j \right. \right\rangle = \frac{1}{E_i - E_j} \left\langle \Psi_i \left| \frac{\partial H_{el}}{\partial \mathbf{R}} \right| \Psi_j \right\rangle \text{ for } i \neq j \quad (7)$$

Second-order derivatives are neglected here, as is done in many quantum chemistry programs that compute NAC vectors. The blue dashed curve in Figure 6a illustrates the norm of the NAC vector,  $C_{\text{NAC}_{ij}}$ , that couples the states  $S_i$  and  $S_j$ . At the avoided crossing points of the states, the NAC norm shows a sharp spike, but is almost vanishing elsewhere. If more than one multiplicity is considered, the term adiabatic is not adequate anymore, because potentials of different multiplicity might cross through each other. This situation is then called diabatic with respect to the spin multiplicities, or spin-diabatic

in short. For example, singlets are adiabatic among each other, triplets are adiabatic among each other, but singlets are diabatic with respect to triplets. However, also the diabatic basis (see Figure 6a and also below) qualifies as spin-diabatic. Because of this nomenclature issue, which even gets experts confused sometimes, we refer to this basis as “Molecular Coulomb Hamiltonian” (MCH) because it is obtained from the eigenfunctions and eigenvalues of the nonrelativistic electronic Hamiltonian, where only Coulomb interactions are considered.

As an example, a crossing of a singlet state and a triplet state is shown in Figure 6b. As it is visible, the triplet components, which are defined by different magnetic quantum numbers, are degenerate. The states are coupled by SOCs (denoted as  $C_{\text{SOC}_{ij}}$ ), which are usually obtained as smooth potential couplings with standard quantum chemistry programs:<sup>256,365,367</sup>

$$C_{\text{SOC}_{jk}} = \langle \Psi_j | \hat{H}^{\text{SO}} | \Psi_k \rangle \quad (8)$$

These couplings are single real-valued or complex-valued properties.<sup>368,369</sup> Whether they are complex or not depends on the electronic structure program employed, but they can be converted into each other.<sup>256,257,368,370</sup> It is important to know in which way the SOCs are presented by an electronic structure program in order to find the best possible solution for ML approximations.<sup>15</sup>

$\hat{H}^{\text{SO}}$  in eq 8 is the spin-orbit Hamilton operator, which describes the relativistic effect due to interactions of the electron-spin with the orbital angular momentum, allowing states of different spin-multiplicities to couple.<sup>370–372</sup> Note that also SOCs between different states of the same multiplicity exist except for singlets. No exact expression on how to include relativistic effects into the many-body equations has been found yet. Among the most popular approximations used is the Breit equation,<sup>373</sup> applying an adapted Hamiltonian instead of the electronic Hamiltonian, which comprises, among other terms, a relativistic part. This additional part of the Hamiltonian accounts for spin-orbit effects and is proportional to the atomic charge,<sup>257,368,370,372,373</sup> leading to the belief that SOCs would only be relevant in systems with heavy atoms.<sup>233,234</sup> Today, it is known that spin-orbit effects also play a crucial role in many other molecular systems and are important for intersystem crossing between states of different spin multiplicities.<sup>178,235–237</sup>

The states in the MCH basis can also be coupled via external electric-magnetic fields, e.g., by sunlight or a laser. The corresponding couplings stem from the transition dipole moments multiplied with the electric field. Since the effect of the field is not included in the potentials but as off-diagonal potential couplings, the MCH basis is also called field-free.<sup>362–364,374</sup> However, also the diabatic basis qualifies as field-free.

**3.2.2. Diabatic Basis.** In the diabatic basis, the electronic wave function is not parametrically dependent on the nuclear coordinates. Diabatic potentials usually need to be determined from adiabatic potentials and are not unique; i.e., they rely on the method and the reference point, which is chosen in the adiabatic basis to fit diabatic potentials.<sup>239,256</sup> The advantage of diabatic potentials is that singularities in NACs are removed together with nondifferentiable points of the PESs. Note that such a strictly diabatic basis for polyatomic systems does not exist in practice, and only approximated, so-called, quasi-diabatic, PESs can be fit, where NACs are almost removed, but

not completely. In the literature, quasi-diabatic PESs are most often referred to as diabatic ones, so we will also use this notation here.

An example of a system in the diabatic basis is given in Figure 6a, and commonly used notations can be found in Table 1 in the first column. In regions where an avoided crossing is present in the adiabatic basis, the coupled diabatic potential energy curves cross. Since the electronic wave function of a state is ideally independent of the nuclear coordinates, its character is conserved. Consequently, the states are labeled according to their character and multiplicity, e.g., as  $^1\pi\pi^*$  or according to symmetry labels. Similar to the character, also spectroscopically important quantities like the dipole moment are mostly conserved or vary smoothly along the nuclear coordinates. Therefore, spectroscopic experiments can easily be interpreted when using the diabatic basis, which is thus sometimes also called spectroscopic basis. Note that sometimes labels like  $S_1$ , etc. are used also when referring to the diabatic basis, especially in experimental papers when an identification of the wave function's character has not been carried out and only one geometry is considered. However, at a different geometry, the energetic order of the states might have changed such that a state previously labeled as  $S_2$  might now be lower in energy than a state previously labeled as  $S_1$ . Furthermore, this labeling scheme in the diabatic basis can lead to confusion with the labels from the MCH basis, and we suggest to reserve it only for the MCH basis.

Because of the mostly conserved characters and the crossing of states, diabatic potentials are smooth functions of the nuclear coordinates, in contrast to adiabatic potentials. A diabatic PES is thus highly favorable for several numerical applications including ML.

The MCH and diabatic bases can be interconverted by a unitary transformation

$$\Psi^{\text{MCH}}(\mathbf{r}, \mathbf{R}) = \mathbf{U}(\mathbf{R})\Psi^{\text{diab}}(\mathbf{r}, \mathbf{R}) \quad (9)$$

with a unitary matrix,  $\mathbf{U}$ , that is determined up to an arbitrary sign (as a result of the arbitrary sign of the wave function, which will be discussed in detail in section 5.3). In the case of two states,  $\mathbf{U}$ , is a rotation matrix:

$$\mathbf{U} = \begin{pmatrix} \cos \theta(\mathbf{R}) & -\sin \theta(\mathbf{R}) \\ \sin \theta(\mathbf{R}) & \cos \theta(\mathbf{R}) \end{pmatrix} \quad (10)$$

and is dependent on the rotation angle,  $\theta$ . Accordingly, the peaky NACs, which are obtained as derivative couplings (also called kinetic couplings) in the MCH basis, are converted to smooth potential couplings in the diabatic basis. The smooth SOC from the MCH basis become even smoother (ideally constant) in the diabatic basis.

While one can straightforwardly apply diagonalization to convert diabatic PESs to adiabatic PESs (and similarly adiabatic PESs to diagonal PESs), a dilemma arises when one wants to take the inverse way to obtain diabatic PESs from adiabatic ones (and similarly adiabatic PESs from diagonal ones). In fact, finding diabatic PESs is highly complex and most often requires expert knowledge. To date, only small molecules could be represented with accurate diabatic potentials, and developing a method to automatically generate diabatic PESs remains an active field of research. Existing methods to obtain diabatic potentials require human input and are mostly applicable to small systems and certain reaction coordinates. Early pioneering works can be found in refs 239

and 375. Today, a lot more variants exist. Examples are diabaticization using explicit derivative couplings,<sup>376,377</sup> the propagation diabaticization procedure,<sup>378</sup> diabaticization by localization,<sup>379</sup> Procrustes diabaticization,<sup>252</sup> or diabaticization by ansatz.<sup>142,380,381</sup> Further, methods can be based on couplings or other properties,<sup>382–385</sup> configuration uniformity,<sup>386</sup> block-diagonalization,<sup>387,388</sup> a hybrid diagonalization combining block-diagonalization and diabaticization by ansatz,<sup>389,390</sup> CI vectors,<sup>391</sup> whereas some procedures are carried out at least partly using ML.<sup>142,143,380,390,392</sup>

**3.2.3. Diagonal Basis.** As the name indicates, the diagonal basis can be obtained by a diagonalization from the MCH or diabatic bases. In this case, a strictly adiabatic picture is obtained, where states never cross.<sup>256</sup> Accordingly, the concept of multiplicity for a single state is lost because the state might be of singlet character in one region and of triplet character in another region. Therefore, the basis is also called spin-mixed or spin-adiabatic.<sup>257,369,393</sup> The states are strictly ordered by energy and can be labeled simply with numbers (see Figure 6c). The resulting wave functions are eigenfunctions of the relativistic electronic Hamiltonian.<sup>236,256,257</sup> These eigenfunctions as well as the eigenenergies can be also obtained directly with, e.g., relativistic two-component or four-component calculations,<sup>394</sup> instead of via diagonalization.

In this basis, the effect of the SOC is incorporated into the PESs to a large extent. What remains are localized kinetic couplings, which are similar in nature to the NACs in the MCH basis. An example is given in Figure 6c. The parts of the potentials that correspond to the different triplet components in the MCH basis are split energetically in the diagonal basis. In the case of small SOC, the diagonal potentials look similar to the MCH potentials. However, if the SOC are strong, potentials that are degenerate in the MCH basis can be easily shifted apart by 1 eV in the diagonal basis. Such splittings are then also experimentally observable, and the diagonal basis yields a more intuitive interpretation of these experiments.<sup>43,395,396</sup>

As mentioned above, the states in the MCH basis can also be coupled via electromagnetic fields. A diagonalization of the potential matrix then yields so-called field-dressed states or light-induced potentials, which can also be termed field-adiabatic.<sup>362,374,397–399</sup> Since the fields are usually time-dependent, the most important axis along which the potentials in this field-dressed basis need to be plotted is time.<sup>374</sup>

In principle, all these bases are equivalent but only if an infinite number of terms is considered in eq 6. In practice, potentials represented in different bases have different advantages for dynamics simulations, especially in combination with different approximations made in the different dynamics methods as outlined below.

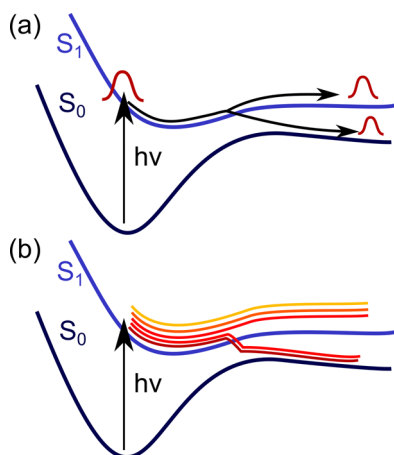
### 3.3. Excited-State Dynamics Simulations

In order to investigate the temporal evolution of an isolated molecular system in the excited states, the time-dependent Schrödinger equation has to be solved:<sup>254</sup>

$$i\hbar \frac{\partial \Psi(\mathbf{r}, \mathbf{R}, t)}{\partial t} = \hat{H}(\mathbf{r}, \mathbf{R})\Psi(\mathbf{r}, \mathbf{R}, t) \quad (11)$$

From a technical point of view, a sequence of time steps is computed, where in every step the electronic problem is solved to yield potentials, which determine the forces acting on the nuclei such that the nuclear equations of motion can be solved for the current time step.

Ideally, the nuclei are treated quantum mechanically. In this case, the PESs are usually computed in advance and either interpolated or stored on a grid for later use. The hope is that ML can improve the interpolation of potentials drastically. Such global PESs are needed because a wave function is employed for the nuclei, which extends over a range of nuclear coordinates at the same time (see Figure 7a). An overview over corresponding dynamics methods is given in section 3.3.1.



**Figure 7.** Excited-state dynamics can be treated with (a) quantum approaches, where wave functions are used for the nuclei, or (b) classical approaches, based on trajectories.

The nuclear dynamics can also be approximated classically while quantum potentials are used; i.e., mixed quantum classical dynamics (MQCD) simulations are carried out. Such methods are discussed in section 3.3.2. Since the classical nuclear trajectories are defined only at one nuclear geometry at a time (see Figure 7b), on-the-fly calculations of the potential energies are possible. An on-the-fly scheme is computationally advantageous if the number of visited geometries during the dynamics is smaller than the number of points needed to represent the conformational space on a grid or via interpolation.<sup>236,256,258,331,332,367,400</sup> No fitting of PESs is necessary in an on-the-fly approach, but fitted PESs can still be used as an alternative. Since ML approaches provide such interpolated potentials, the amount of training points generated with quantum chemistry must be less than the number of points needed in an on-the-fly approach in order to be advantageous. This demand is satisfied, e.g., for long time scales or if many trajectories are necessary.

In the following, we will shortly discuss the different types of describing nuclear motion and the opportunities of ML models to enhance the respective dynamics simulations.

**3.3.1. Quantum Nuclear Dynamics.** The computational cost of an exact nuclear dynamics simulation scales exponentially with the nuclear degrees of freedom. Hence, simulations are limited to small systems, typically containing less than five atoms.<sup>59,368,401,402</sup> Still, the calculation of the PESs of the molecule can be a rather expensive part of the whole scheme, and the use of ML algorithms is advisable even for such small systems.

To treat larger systems, approximations have to be invoked.<sup>402</sup> A prominent approach that can be converged to the exact solution is the multiconfigurational time-dependent Hartree (MCTDH) approach.<sup>48,403–405</sup> Its high efficiency stems from the use of time-dependent basis functions to

represent the nuclear wave functions. Nonetheless, the computations are computationally costly, and the nuclear degrees of freedom are often reduced to only a few important key coordinates,<sup>239,406</sup> where classical simulations can help identify the latter.<sup>407</sup> Whether quantum dynamics of such reduced-dimensionality models are better than using classical dynamics of a full-dimensional system is still under debate and probably depends on the system. In the case of quantum dynamics, the potentials need to be presented to the algorithm in the diabatic basis, mostly due to numerical stability (e.g., smooth couplings are easier to integrate than singular ones). For more than 20 years, the (modified) Shepard interpolation has been used to fit diabatic potentials.<sup>151,408–411</sup> Notably, the grow algorithm<sup>151</sup> can be used to efficiently generate the database of points upon which the interpolation is based. However, it is clearly desirable to treat larger systems, and ML models like neural networks (NNs) promise higher performance or more flexibility in such cases.<sup>143,146,147,149,378,392,412–414</sup>

More recently, on-the-fly methods addressing quantum dynamics have been developed.<sup>145,415–417</sup> They mostly rely on a combination of Gaussians to represent the nuclear wave function.<sup>258</sup> For example, the variational multiconfiguration Gaussian method (dd-vMCG)<sup>418</sup> offers a variational and thus accurate solution for the equations of motion. Also full multiple spawning<sup>45,401,419</sup> can be regarded as fully quantum mechanically by describing the wave function with a number of time-dependent Gaussian functions that follow classical trajectories with quantum mechanically determined time-dependent coefficients. In its more affordable ab initio multiple spawning variant, more approximations are introduced such that the results sometimes draw near the classical solutions.<sup>420,421</sup> Further related methods exist, like the ab initio multiple cloning method,<sup>422</sup> or the thawed Gaussian approximation.<sup>423</sup>

Another class of dynamics methods are semiclassical approaches, which allow the inclusion of quantum effects in the classical dynamics of nuclei, such as quantum mechanical tunnelling or coherence.<sup>424</sup> Note that these methods, where the nuclear dynamics is treated semiclassically, should not be confused with the MQCD approaches (see below) that are also often termed semiclassical (because the nuclei are treated classically and the electrons quantum-mechanically). The semiclassical dynamics methods range from the initial value representation,<sup>425,426</sup> adapted with the Zhu–Nakamura approach leading to the Zhu–Nakamura–Herman–Kluk initial value representation,<sup>427</sup> to path integral approaches.<sup>428</sup>

The path integral formalism is especially interesting when the quantum and classical degrees of freedom should be coupled in a dynamically consistent manner. By using so-called ring-polymers, i.e., replica of the original classical system, a deviation of the nuclear dynamics from the classical path can be obtained, and the time evolution of a system including nuclear quantum effects can be investigated. However, ring-polymer dynamics suffer from high computational efforts as a consequence of the large number of replica required. Accelerated formalisms exist, which are for example implemented in the Python wrapper i-PI,<sup>429,430</sup> which allow interfacing path-integral methods with programs that provide PESs, but are mostly dedicated to the electronic ground state. To date, only a few implementations of semiclassical methods in atomistic simulation software are available. Compared to classical mechanics, the computational costs increase by a factor of about 10–100.<sup>424,431,432</sup>

### 3.3.2. Mixed Quantum-Classical Molecular Dynamics.

While semiclassical methods are promising to simulate the dynamics of molecular systems containing up to tens of atoms highly accurately, the study of larger systems is still dominated by computationally cheaper MQCD methods, where the nuclear motion is treated fully classically.<sup>59,431–433</sup> In contrast to quantum dynamics, the motion of the nuclei can be computed very fast using classical mechanics, and the computation of the PESs, on which the nuclei are assumed to move, remains the time limiting step. In this sense, ML models have a huge potential to enhance MQCD simulations by providing the electronic PESs and enabling the investigation of reactions that are not feasible with conventional approaches.<sup>15,175,434,435</sup> In fact, most studies to date that describe photochemistry with ML aim to replace the quantum chemical calculation of the PESs in MQCD approaches.

The most popular MQCD method is trajectory surface hopping,<sup>436–438</sup> schematically represented in Figure 7b. A manifold of independent trajectories is required to obtain statistically relevant results and to mimic the extended nuclear wave functions. For a single trajectory, the nuclei move classically on one of the quantum potentials, and hence only one state is considered to be active, but transitions between different states are allowed.<sup>439</sup>

Different approaches exist to determine the probability of such a transition, also called hop or jump in surface hopping methods. To this aim, different quantities are needed that are commonly provided in the MCH basis, as it is the direct outcome of a quantum chemical simulation. One of the first implementations to compute the hopping probability is based on the Landau–Zener formalism.<sup>440,441</sup> On the basis of the Landau–Zener formula, the potential energy differences are used to determine the hopping probability. No information about couplings is required, which implies that the approach must fail for states that do not couple but lie close in energy. Very similar to this approach is the Zhu–Nakamura theory.<sup>442–445</sup> Also here, the computation of couplings is omitted, and only information about PESs is used. Among the mostly used hopping algorithm is Tully's fewest switches algorithm,<sup>436</sup> which is valid for many cases and based on the NACs between different PESs. An extension to other couplings is provided, e.g., in the surface hopping including arbitrary couplings (SHARC) method.<sup>236</sup> When couplings are considered, an internal transformation from the MCH basis to the diagonal basis is most advantageous because the localized couplings of the diagonal picture precisely indicate where the few switches of the fewest switches approach should take place. In cases where the PESs are fit in advance, either with ML models or other types of analytical functions, the use of a diabatic basis is favorable (because of the Berry phase, see below), but should be transformed to the diagonal picture for the calculation of hopping probabilities. Other flavors to account for transitions exist. However, they have not been applied in simulations with ML algorithms yet. Interested readers are therefore referred to refs 47, 236, 256, 257, 332, 436, 443, and 446–450 for further information.

The bottleneck of approaches that require NACs is that the computation of the couplings remains one of the most expensive parts of a quantum chemical calculation. The computational effort to compute a NAC vector is comparable to that of a force calculation. However, more NACs are present than there are forces, i.e.,  $N_S \times (N_S - 1)/2$  NACs need to be computed, whereas  $N_S$  forces are needed (respectively with

entries for the Cartesian coordinates of each nucleus). Note that in the case of fitted PESs with ML, all of these vectors have to be computed for each data point. Conventional approaches with an ab initio on-the-fly evaluation of the PESs can make use of the fact that only one active state needs to be considered at a certain time step. Many MD programs therefore only require a computation of the forces of the active state and the respective couplings arising from this state.

Note that despite the benefits of MQCD simulations, they obey microreversibility only approximately,<sup>451</sup> and effects due to coherences or tunneling necessitate additional considerations as a consequence of the classical treatment of nuclear motion.<sup>452</sup>

A more approximate approach is the Ehrenfest dynamics method, also referred to as mean-field trajectory method. It is often used for large systems and also frequently in material science.<sup>183,194</sup> The Ehrenfest method is based on the approximation that nuclei move classically on an average potential, rather than switching from one specific state to another.<sup>332,453,454</sup> Because of the treatment of each electronic state separately, surface hopping methods allow the accurate bifurcation into different reaction channels, while such effects are neglected in a mean-field treatment of PESs.

The main limitation of MQCD approaches is the expensive evaluation of ab initio potentials, which allows dynamics simulations only for up to a couple of picoseconds. In addition, rare reaction channels are hardly explored as a result of usually bad statistics.<sup>257,455,456</sup> In this sense, MQCD simulations offer a perfect place for ML to enter this field of research and advance it significantly. The fast evaluation of the ML PESs can help to explore different reaction channels and to obtain accurate reaction kinetics. Observables and macroscopic properties can be computed directly or with postprocessing as well as analysis runs and offer another fulcrum for ML. The computed observables should then be directly compared to experiments.

### 3.4. Dipole Moments and Spectra

An important property for comparing experiment and theory is the dipole moment. The permanent dipole moment of the ground state is a frequent target of studies with ML.<sup>111,172,457–467</sup> The permanent dipole moment,  $\mu_i$  (or  $\mu_{ii}$ ), of a state  $i$  can be obtained via the dipole moment operator (see eq 13 below) or as the sum of partial charges,  $q_{a,i}$  of atom  $a$  in state  $i$ , and the vector that describes the distance of the position of atom  $a$  to the center of mass of the molecule,  $r_a^{\text{CM}}$ :

$$\mu_i = \sum_a^{N_A} q_{a,i} r_a^{\text{CM}} \quad (12)$$

It can be used for the computation of infrared spectra with MD simulations. The spectrum is then obtained as the Fourier transform of the time autocorrelation function of the time derivative of the dipole moment.<sup>468</sup>

In contrast to the ground state, excited-state simulations often make use of the transition dipole moments, which are computed from the dipole moment operator within many quantum chemistry programs:

$$\mu_{ij} = \langle \Psi_i | \hat{\mu} | \Psi_j \rangle \quad (13)$$

The ground state dipole moment can differ strongly from those in the excited states, due to a frequency shift and altered electron distribution upon light-excitation.<sup>469</sup>

Transition and permanent dipole moments are the target of some very recent studies, see refs 16, 94, 95, 97, and 241. Worth mentioning is the charge model of ref 111 originally proposed for the ground state and adapted by us for an arbitrary number of excited states,<sup>241</sup> where point charges are never learned directly, but instead are inferred as latent variables by an NN dipole model making use of  $r_{\alpha}^{\text{CM}}$ .<sup>172</sup> In this way, the rotational covariance can be preserved.

Noticeably, the computation of absolute values of permanent and transition dipole moments is very challenging even when highly accurate quantum chemistry methods are employed and experimental values are hardly reproduced.<sup>95,470</sup> However, also experimental studies provide absolute values only in a few cases. Most computational studies therefore do not aim to reproduce the absolute values of transition dipole moments but rather use relative values to obtain reasonably accurate absorption spectra, which can be compared to experiments.<sup>55,254,256,471–473</sup> Since many molecules absorb in the UV, the terms UV spectra and absorption spectra are often used interchangeably. However, absorption can take place in many regions of the electromagnetic spectrum, including, e.g., X-rays, where rather core electrons than valence electrons are excited.<sup>474</sup>

As already mentioned shortly, absorption spectra can be obtained from a calculation of excited-state energies and oscillator strengths, which are proportional to the squared transition dipole moments. Noticeably, the transition dipole moment is only defined up to an arbitrary sign as a result of the arbitrary phase of the wave function (see section 5.3).<sup>16,94</sup> To circumvent this ill-definition, oscillator strengths or the lengths of dipole vectors can be fitted with ML. However, this workaround can be problematic if explicit field–dipole interactions should be considered with ML models, as the relative orientation of the field vector and the dipole vector can be important in this case. In theory, when computing UV/visible absorption spectra, the absorption related to a type of transition of one molecular geometry results in a delta function. However, in experiments lines are usually broadened due to the experimental technique and instrument used and because the molecules vibrate and collide with each other. Therefore, the computation of spectra requires the evaluation of transition properties of not only one conformation, but of many thousands. Wigner sampling<sup>475</sup> or sampling via MD can help to achieve better agreement with experiment in the case of absorption spectra, but the line shapes due to different vibrational lifetimes can usually not be described. Therefore, the lines are broadened using Gaussian, Lorentzian, or Voigt functions, which can approximate effects due to Doppler broadening, broadening due to the uncertainty of the lifetimes of excited states, or a combination of these effects, respectively.<sup>246,476</sup>

## 4. ML MODELS

Besides the reference method to compute the training set, which defines the highest possible accuracy an ML model can attain, the type of regressor and the descriptor to represent a molecule to the ML model also play important roles.<sup>477</sup> Improper choices of regressors and descriptors can result in inaccurate ML models.

### 4.1. ML Models: Type of Regressor

Given the vast number of ML algorithms applied in the field of computational chemistry, one might ask which one to use or

adapt for photochemistry. As recent studies applying ML for quantum chemistry have shown, many possible choices of ML approaches exist, and there is no single solution. Nevertheless, a trend can be observed: Many studies that use ML in the research field of quantum chemistry employ labeled data sets, i.e., supervised learning techniques. Within supervised learning, one can distinguish between regression and classification. Classification aims at finding patterns and at grouping data into certain clusters.<sup>478</sup> Those types of ML models are often used, e.g., in spam filters, in medicine to diagnose diseases,<sup>479,480</sup> or in food research, e.g., to guarantee a certain wine quality or origin.<sup>481</sup> Examples of applied classification models in the field of computational chemistry are for example support vector machines, random forests, or decision trees used, e.g., to classify enzymes<sup>482</sup> or for the selection of an active space.<sup>71,483</sup>

More often than classification models, regression models are applied to assist the search for a solution of a quantum chemical problem. Regression is used to fit functions that can relate a molecular input,  $X$ , to a quantum chemical output,  $Y$ . The simplest relation that can be assumed is linear. However, most quantum chemical problems cannot be accurately described with a linear function as given in eq 14. Linear regression is not seen as a universal approximator, while this quality has been proven, e.g., for NNs.<sup>484,485</sup> This is why we do not consider linear regression as an ML method in this review. Note that seemingly accurate fits can be achieved with linear regression when using more inputs (descriptors, features) than data points, but such an agreement rather resembles spurious correlation than real learning. Nonetheless, linear models are the foundation of many ML approaches and can serve as a baseline model to evaluate the minimum accuracy that an ML model should obtain.<sup>94</sup> In the linear relation,

$$Y = b + wX \quad (14)$$

the regression coefficients, also known as weights,  $w$ , and biases,  $b$ , are tailored for a given problem under investigation. Here, ordinary least-squares minimization can be applied to find these coefficients. The process of finding the optimal relation between  $X$  and  $Y$  is termed training. The coefficients are optimized by minimizing a so-called loss function,  $L$ , which monitors the error between the original property,  $Y^{\text{QC}}$ , and the predicted property by the ML model,  $Y^{\text{ML}}$ , with respect to the training instances. Most often, the  $L_1$  loss or the  $L_2$  loss is used as an indicator for the training convergence. The  $L_1$  monitors the MAE and the  $L_2$  loss the mean squared error (MSE) of predictions:

$$L_2 = \frac{1}{N_M} \sum_{\beta}^{N_M} (Y_{\beta}^{\text{ML}} - Y_{\beta}^{\text{QC}})^2 \quad (15)$$

The Greek letter  $\beta$  runs over all molecules,  $N_M$ , inside the training set. In principle, any error estimate can be used to train an ML model and find suitable regression coefficients.

An example specifically developed for excited-state problems is the aforementioned phase-less loss (see section 5.3.2).<sup>15</sup> Such adapted loss functions and also conventional ones are employed in different types of ML models. In the following, we focus on the two most widely used models for the description of the excited states: Kernel methods and NNs.

**4.1.1. Kernel Methods.** Kernel methods<sup>486</sup> are based on a similarity measure between data points. Examples are KRR or GPR, which go beyond linear regression by applying the kernel trick and ridge regression. Ridge regression is used to find the

weights, which differs from linear regression by a regularization term,  $\lambda$ :

$$w = (K + \lambda I)^{-1} Y^{QC} \quad (16)$$

$Y^{QC}$  refers to the training data and  $K$  to the kernel matrix.

The kernel trick makes it possible to apply ridge regression to nonlinearly separable data by mapping them into a higher-dimensional feature space, in which the data points are linearly separable. Therefore, a kernel function,  $k$ , e.g., a Gaussian or Laplacian, is placed on each compound to measure the distance to all of the other compounds in the training set. The kernel function defines the nonlinearity of the model. A property of a query compound,  $\alpha$ , can be obtained as the weighted sum of regression coefficients and kernel instances:

$$Y^{ML}(X_\alpha) = \sum_{\beta}^{N_M} w_{\beta} K(X_\alpha, X_{\beta}) \quad (17)$$

The size of the kernel matrix is dependent on the number of training points, and hence the depth of the model is inherently linked to the size of the training set, which is why they are called “nonparametric”.<sup>478,487</sup>

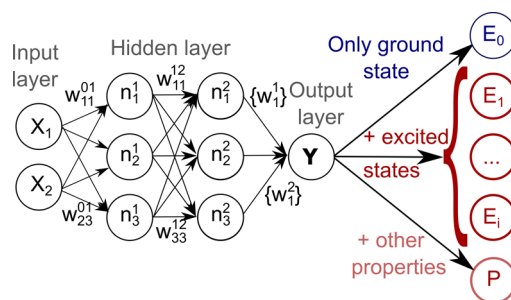
An advantage of kernel methods is that they mainly contain two hyperparameters, i.e., internal model parameters, which need to be optimized for proper training. Most important are the width of the nonlinear kernel function,  $\sigma$ , and the regularization. The latter is used to prevent the model from overfitting—the case when the model fits training data including noise almost exactly and fails to accurately predict data points not included in the training set but stemming from an interpolative regime. As quantum chemical data are most often noise-free, the regularization term is usually small.

Especially for excited-state data, there are, however, systematic errors that can be seen as noise: Inconsistencies in potential energy curves along certain reaction coordinates are quite common and NACs are singular at crossing points of two PESs.<sup>171</sup> In these cases, care should be taken in order to avoid overfitting. A powerful way to mitigate the problem of overfitting, e.g., for kernel methods, is, for instance,  $k$ -fold cross-validation.<sup>478,488</sup> Although kernel methods are generally said to be resistant to overfitting,<sup>95,489</sup> we shortly want to discuss  $k$ -fold cross validation here. By applying  $k$ -fold cross-validation, the whole data set is split into a *training* and *test* set. The *test* set is held back until the optimal hyperparameters are found and the remaining *training* set is split into  $k$  parts, which is most often 5 or 10. The kernel method is trained on  $k-1$  parts of the *training* set, while the hyperparameters are optimized in order to minimize the error the model makes on the last part of the *training* set, i.e., the *validation* data in this case. The procedure is carried out  $k$ -times, and each time the *validation* data consist of another part of the *training* set. This procedure allows the optimization of hyperparameters without being biased by the error on the test set. In case data are sparse and expensive,  $k$ -fold cross-validation is a powerful technique to find optimal hyperparameters of an ML model without the need to compute massive amounts of additional data for validation and testing. The accuracy of the final ML model is then assessed by computing the error on the *test* set. This procedure is similarly valid for NNs, whereas overfitting is most often mitigated by applying an early stopping mechanism, which will be discussed in the context of NNs in the next paragraph.

As the optimization of hyperparameters is often a tedious task, kernel methods with their few hyperparameters are easier to use than, e.g., NNs with many hyperparameters. Nonetheless, kernel methods can provide almost exact solutions of problems under investigation.<sup>127</sup> A drawback is, however, that the inversion of the kernel matrix can become expensive and even be rendered infeasible on current computers due to increasing memory requirements with increasing training set size.<sup>95</sup>

Further, kernel methods are usually defined to only map an input to a single output. Therefore, they can treat only one electronic state at a time in standard implementations and, thus, can be referred to as single-state models. A single-state treatment requires a separate ML model for each electronic state or for each property resulting of a pair of states, whereas a multistate ML model describes all electronic states and properties resulting from different pairs of states at once.<sup>95,175</sup> Hence, in their standard implementation, the treatment of several excited states necessitates the use of several kernel models, which is commonly done in the research field of quantum chemistry.<sup>139,140,149,490,491</sup> The description of forces is possible for the ground state or a single excited state and is implemented, e.g., in the QML toolkit using KRR and the Faber–Christensen–Huang–Lilienfeld (FCHL) representation,<sup>466</sup> in the symmetric gradient domain ML (sGDML)<sup>122,122</sup> method, with smooth overlaps of atomic positions (SOAP)<sup>492</sup> for GPR,<sup>121</sup> or GAP<sup>101</sup> originally developed for materials, but recently also applied for molecules, e.g., fluid methane.<sup>493</sup>

**4.1.2. Neural Networks.** Another prominent approach in ML is the use of NNs as highly flexible parametric functions, which can fit huge amounts of data and can map a molecular input to many quantum chemical outputs.<sup>95</sup> The simplest form of NNs is the multilayer feed-forward NNs, which are schematically represented in Figure 8.



**Figure 8.** Schematic representation of a multilayer feed-forward NN with inputs,  $X$ , nodes,  $n$ , and outputs,  $Y$ . In the usual implementation for the fitting of PESs, the NN maps a molecular geometry to the ground state, which could be similarly done for any other single state. In the case a manifold of excited states is described, one molecular input can also be mapped to a vector of different excited states, and additionally, other properties can be included. The forces are treated as derivatives of the NN potentials with respect to Cartesian coordinates.

As it is visible in Figure 8, the width of the model is dependent on the number of nodes,  $n_r^t$ , which are connected to each other using weights,  $w_{rs}^{tu}$ . The indices refer to a connection between node  $r$  and node  $s$  from layer  $t$  and layer  $u$ , respectively. The number of nodes and hidden layers can be chosen independently of the training set size.



Because of the highly flexible functional form of NNs, highly complex relationships can be fit, but an analytical solution to find the weights is not available (in contrast to KRR). A numerical solution can be obtained with stochastic gradient algorithms, which are frequently applied to obtain a stepwise update of the weights:

$$w_{k+1} = w_k - l_r \nabla L_2(w) \quad (18)$$

The gradient of the loss function as given in eq 15 with respect to the weights is multiplied with a so-called learning rate,  $l_r$ . This hyperparameter is deemed one of the most important hyperparameters used for training.<sup>10,494</sup> In order to obtain an optimal solution, the learning rate needs to be chosen properly. Algorithms such as AdaGrad<sup>495</sup> or Adam<sup>496</sup> can automatically adapt the learning rate during training. Further, the second-order derivatives can be included into algorithms, which is for instance done in the global extended Kalman filter,<sup>497</sup> in its parallel variant,<sup>498</sup> or the element-decoupled variant.<sup>105</sup> The loss function can be adapted so that more than only one property can be trained at once. This is often done to include the forces in the training process.

In general, NNs possess various hyperparameters like the learning rate, regularizers, number of nodes, etc. As a consequence, an extensive hyperparameter search complicates the use of NNs and makes them more complex to apply than kernel methods. One common hyperparameter optimization procedure is random grid search.<sup>10,494</sup> Therefore, the hyperparameters of the model are randomly shuffled, and ML models are trained using these parameters. After training, the errors on a validation set are compared to each other, and the hyperparameter space, which has to be explored, can be narrowed. This procedure is beneficial if it is repeated several times while narrowing the space of hyperparameters every time. As hyperparameter optimization is an optimization problem, algorithms, such as Bayesian optimization<sup>499–502</sup> have been designed for this task and are frequently applied for deep ML models and kernel methods. Because of the tedious procedure of manually tuning hyperparameters or the expert knowledge, which is required in most cases to find optimal hyperparameters, ML models have been designed which automatically learn optimal hyperparameters and only require little human intervention, see, e.g., refs 15,77, 461, 465, 503–508.

In addition, NNs are prone to overfitting. Therefore, during training, it can be beneficial to split the *training* data into two parts, typically in a ratio of 9:1, with the first part being directly used for training (i.e., adjusting the weights) and the second only for validation. In every training epoch, the error of the ML model on the training and validation data is compared. As soon as the error on the validation data increases, the training is stopped. This process is known as early stopping and is, besides drop-out or the comparison with less complex ML architectures, a powerful tool to prevent an NN from overfitting.<sup>10,478</sup>

Besides simple multilayer feed-forward NNs, high-dimensional variants exist. These networks comprise several atomic NNs, which represent atoms in their chemical and structural environment and are thus also called atomistic NNs. Each local atomic contribution,  $E_a$ , can be summed up to provide the energy of the whole system,  $E$ , which is well-known to work for the ground-state PESs:

$$E = \sum_{a=1}^{N_a} E_a \quad (19)$$

and was originally implemented by Behler to construct high-dimensional NN potentials.<sup>509</sup> Embedded-atom NNs<sup>508</sup> are similar to high-dimensional NNs in their way of constructing the energy of a system. They differ in the underlying descriptors to the ones of Behler. Atomic contributions to the energy are dependent on the embedded density of atoms and are summed up according to eq 19. These embedded density-like descriptors are approximated from atomic orbitals.

Independent of a simple or an atomistic architecture, the model can be used to fit a single output or a vector of many outputs at the same time. For ground state problems, a single-state model is usually used, which maps an input to a single output, e.g., the PES of the ground state. Oftentimes, this single-state fashion is adapted to fit different excited states with different NN models.<sup>16,141,204,510</sup> However, it has been shown that including more excited-states in one model can be advantageous,<sup>95</sup> as the excited-states are inherently linked to each other and so are the excited-state properties.<sup>178</sup> Treating many excited states can be referred to as multistate model, and the inclusion of more properties can result in a multiproperty model.<sup>77,95,97,175,191</sup> The different properties can be weighted with respect to their magnitudes or importance for a given chemical problem under investigation, such that the best possible accuracy can be obtained.<sup>15</sup>

Convolutional NNs represent another class of networks and are most often applied in image or speech recognition,<sup>511–513</sup> but can also be adapted to process a molecular input and identify an optimal molecular descriptor. This type of network can be combined in an end-to-end fashion with an architecture, which fits this generated molecular representation to a query output.<sup>461,465,503,507,514</sup>

An important ingredient of all these ML models is the descriptor, which is mapped to the output. In most studies, the descriptor is one of many different possibilities to represent a molecule, which will be discussed in the next section.

#### 4.2. Descriptors and Features

Electronic structure methods can process and uniquely identify molecules using, e.g., Cartesian coordinates. In contrast, such types of inputs are not optimal for ML models as the same molecular geometry, but translated or rotated, could only be mapped to the same output with great effort and unnecessary computational cost. Hence, a molecular descriptor should fulfill the following requirements: It should be translationally, rotationally, and permutationally invariant as well as differentiable.<sup>104</sup> It should also be unique with respect to the relative spatial arrangement of atoms, universally applicable for any kind of system, and computationally efficient.<sup>477</sup> However, a descriptor can be more than that; it can already include a part of the mapping, e.g., from a molecular structure to an energy. It can thus ease the task of the regressor and help to attain the best possible accuracy for a given training set.

The ways to represent a molecule to an ML model can be classified roughly into two categories: molecule-wise descriptors, which represent the molecule as a whole to the ML model, and atom-wise descriptors, which represent atoms in their chemical and structural environment and build up a property using local contributions.<sup>104,515</sup> Both ways in describing a molecular system have their merits and pitfalls

and will be discussed along with their applications in recent studies for the excited states in the following.

**4.2.1. Molecule-wise Descriptors.** The distance matrix is one of the simplest descriptors that preserves rotational and translational invariance. Most often it is used in its inverse form with distances between atoms  $a$  and  $b$ ,

$$D_{ab} = \frac{1}{\|r_a - r_b\|} \quad (20)$$

giving rise to the symmetric inverse distance matrix,  $\mathbf{D}$ . Because of the ill-definition of diagonal elements, which are not differentiable, the diagonal elements are excluded, and only the upper or lower triangular matrix is used to represent a molecule to an ML model.<sup>491</sup> Since the Hamiltonian contains distances rather in the denominator, it makes sense to also use the matrix of inverse distances.<sup>94</sup> The matrix of inverse distances is very similar to the Coulomb Matrix,  $\mathbf{C}$ :<sup>102</sup>

$$C_{ab} = \begin{cases} 0.5Z_a^{2.4} & \text{if } a = b \\ \frac{Z_a Z_b}{\|r_a - r_b\|} & \end{cases} \quad (21)$$

but the Coulomb matrix additionally considers fixed point charges,  $Z$ , as employed in classic force fields. These types of descriptors are frequently used in ML studies for the excited states. For example, MLMD simulations in the excited states could be advanced using these simple descriptors<sup>94,95,139,140</sup> and were also accurate enough to fit NNs and KRR models for excited-state properties.<sup>94,95,162,358,490,510</sup> Distance based descriptors are further implemented in several program packages that have been used for photodynamics simulations with KRR. For example, MLAtom<sup>516</sup> contains the Coulomb Matrix and a representation that includes all nuclear pairs in form of normalized inverted internuclear distances.<sup>517</sup> The QML toolkit<sup>518</sup> includes the Coulomb matrix in addition to other representations, such as bag of bonds.<sup>519</sup> Another variant is polynomials formed from inverse distances.<sup>94</sup>

These molecule-wise descriptors have the advantage of being easy to use and implement. Especially for small molecular systems and with regard to the training of an ML model, they are cheap. However, they might miss some important information based on angular distributions. Currently, it is also investigated whether representations based on two-body or three-body terms are accurate enough to uniquely identify a molecule.<sup>520</sup>

A problematic issue of the aforementioned types of distance-based molecular descriptors is that they are not permutationally invariant.<sup>104,175,514,515</sup> This problem can be mitigated by data augmentation, i.e., randomly permutation of atoms by mixing of matrix rows, which results in more data points for the same molecular input. The additional amount of data increases rapidly with the system size and could lead to long training times.<sup>514,515</sup> Alternatively, another metric than the commonly used  $L_1$  or  $L_2$  norms can be employed, the so-called Wasserstein metric, which was tested with the Coulomb matrix.<sup>521</sup>

Permutation invariant polynomials (PIPs), introduced by Bowman and co-workers,<sup>522–524</sup> are frequently applied in a PIP-NN approach by Yarkony, Guo, and co-workers to investigate ground state<sup>412–414</sup> and photochemical problems.<sup>143,144,147,392</sup> The advantage of these polynomials is that they are invariant to permutation of atoms and inversion.<sup>147</sup>

They comprise single-valued functions,  $p_{ab}$ , such as logarithmic or Morse-like functions, which incorporate internuclear distances,  $r_{ab}$ . The PIP vector,  $\mathbf{G}$ , is obtained applying a symmetrization operator,  $\hat{S}$ , accounting for possible permutation operations:

$$\mathbf{G} = \hat{S} \prod_{a < b}^{N_A} p_{ab} \quad (22)$$

with an example of  $p_{ab}$ :

$$p_{ab} = e^{-cr_{ab}} \quad (23)$$

Evidently, additional hyperparameters such as  $c$  have to be optimized, and the choice of PIPs is generally not unique.<sup>413,525</sup> It is worth mentioning that the internuclear distances are redundant for molecules with more than four atoms, but this redundancy does not affect the description of the PES.

Some studies suggest that molecule-wise descriptors might be superior to atom-wise descriptors as the bonds are possibly better represented by internuclear distances.<sup>95,158,175</sup> A negative aspect of molecule-wise descriptors is, however, that they can only treat one molecular system, because the input size is fixed. The input dimension could, in principle, be defined according to the largest system included in the training set, but this would lead to unnecessarily large input vectors for smaller systems, which would then contain many zero values.<sup>509,514</sup> The training of more ML models, each for one specific system size, is one possible solution,<sup>162</sup> but obviously necessitates the training and evaluation of more than one ML model.

**Atom-wise Descriptors.** In contrast, atom-wise representations allow for a fitting of molecules of arbitrary size and composition. Such descriptors are state-of-the-art for ground-state problems. The main principle for the design of atom-wise descriptors is to fit a reference property as the sum of atomic contributions, as given in eq 19 for the molecular energy of a system. The molecule is thus split into atoms, which are represented in their chemical and structural local environment. Usually, these types of descriptors rely on a cutoff function, which defines the sphere around an atom, which is deemed to be important and is therefore considered when modeling the atomic local environment. Commonly used examples are the SOAP,<sup>492</sup> atom-centered symmetry functions (ACSF),<sup>509</sup> weighted ACSFs,<sup>526,527</sup> embedded-atom density-like descriptors,<sup>508</sup> moment tensor potentials,<sup>528</sup> spectral neighbor analysis potentials (SNAPs),<sup>529</sup> or the FCHL representation.<sup>127,530</sup> Interatomic distance between atoms is considered very important for the design of the descriptor and is most often included in representations in the form of radial distribution functions, so-called two-body terms. They are often used together with angular distribution functions, i.e., three-body terms. It is further beneficial to include one-body terms, i.e., the element types of atoms and hence the stoichiometry.<sup>127,461,526,527</sup> Most often, higher order terms than three-body terms are not included due to increasing costs and little improvements in accuracy.<sup>514</sup> Lately, some models have emerged, which take the embedded electron density<sup>508</sup> or interactions of pairs of atoms<sup>77,80</sup> into account. To allow a scalable and accurate representation of larger molecules, geometric moments could further serve the construction of descriptors. They have been formulated using pairwise distance vectors.<sup>531</sup>

Much effort is further devoted to reduce the costs of the descriptors. For instance, weighted ACSFs,<sup>526,527</sup> embedded-atom density-like,<sup>508</sup> moment tensor,<sup>528</sup> and descriptors used for SNAPs<sup>529</sup> reduce the amount of many-body terms. Especially tantalizing is the reduction of the sum over many-body terms into a product of two-body terms in the last three mentioned descriptors. Very recently, Zhang and Jiang propose so-called piecewise switching function based descriptors<sup>532</sup> for embedded-atom NNs, which scale linearly with the number of neighboring atoms.

The description of PESs from atomic contributions is beneficial in order to treat systems of arbitrary sizes and to use systematic molecular fragmentation methods.<sup>109</sup> Admittedly, the validity of this approach is not so clear for the excited-states, and consequently, such representations are less frequently used in ML studies targeting the excited states. To date, only small molecules have been fitted with atom-wise representations, which are too small to prove the validity of excited-state PESs, which are constructed from local atomic contributions. To the best of our knowledge, the largest molecule fitted with atom-wise descriptors contained 12 atoms and was *N*-methylacetamide.<sup>97</sup> Other molecules were  $\text{CH}_2\text{NH}_2^+$ ,<sup>15,95</sup>  $\text{CH}_2\text{NH}$ ,<sup>141</sup>  $\text{SO}_2$ ,<sup>15</sup> or  $\text{CSH}_2$ .<sup>15</sup> Further studies are needed to demonstrate whether an atom-wise construction of excited-state properties and PESs is possible or not. Nevertheless, this approach is most powerful for studies that aim to describe large and complex systems, which could potentially be described from smaller building blocks. For instance, the construction of a DNA double strand or a peptide could be, at least in principle, constructed from ML models that are trained on their smaller subsystems, i.e., DNA bases and amino acids, respectively. Unfortunately, we are far away from having achieved a description of large molecular systems for the excited states, let alone the construction of accurate PESs of medium-sized molecular systems, such as DNA bases or amino acids.

**Other Types of Descriptors.** Besides the benefits that high-dimensional ML models offer for the fitting of PESs of molecules, descriptors are not restricted to the aforementioned examples. In general, any type of descriptor might be suitable for a given problem. Applied descriptors range from topological and binary features generated from SMILES strings<sup>533</sup> to normal modes, which are often used as a coordinate system and descriptors to fit diabatic PESs (refs 16, 99, 136, 143, 145, 145–147, 149, 392, 534). Other types of molecular features besides structure-based ones, e.g., electro-negativity, bond-order, oxidation states, ..., <sup>17,71</sup> are also used.

**Automatically Generated Descriptors.** The selection of an optimal descriptor and the optimization of the related parameters for this descriptor are not trivial tasks and require expert knowledge in many cases.<sup>514</sup> A way to circumvent an extensive parameter search is offered by the aforementioned message passing NNs,<sup>503</sup> which include the descriptor parameters in the network architecture. In this way, they automatically fit the optimal parameters of a descriptor for a given problem, i.e., training set under investigation. Such tailored descriptors can guarantee highly accurate solutions if the NN model is trained properly. PhysNet,<sup>504</sup> HIP-NN,<sup>505</sup> DeepMD,<sup>506</sup> or Deep Tensor NN (DTNN),<sup>507</sup> which forms the basis of the deep learning model SchNet,<sup>461,465</sup> which in turn is used within the SchNarc approach for excited states,<sup>15</sup> are examples of such NNs.

## 5. DATA SETS FOR EXCITED STATES

The basis of any successful ML model is a comprehensive and accurate training set that can describe the required conformational space of a molecule comprehensively and accurately with as little noise as possible.<sup>535</sup> While electronic structure theory for ground state problems is almost free of noise, the same cannot be said so easily for problems in the excited states. “Bad points with abrupt changes”<sup>16</sup> within ab initio calculations for the excited states are frequently observed, which can occur even far away from any critical point of the PESs and are difficult to detect.<sup>15,16,94</sup> The amount of noise in the reference data depends not only on the chosen method (and in the case of multireference methods on the selected active space), but also on the number of electronic states considered and the photochemistry of the molecule under investigation.

### 5.1. Choosing the Right Reference Method for Excited-State Data

Many existing training sets for ML in quantum chemistry are based on DFT.<sup>103,105,112,526,536–538</sup> The ease of use and low computational costs of DFT-based methods make them suitable to treat large systems with acceptable accuracy. In fact, DFT is the workhorse of many studies solving ground-state problems. In contrast, TDDFT has not yet managed to equal DFT for the treatment of excited-state problems. Consequently, training sets for the excited states are less frequently computed with TDDFT<sup>93,97,98,162,358,539</sup> and rely most often on multireference methods. Examples of applied methods are CASSCF<sup>15,139–141,145,146,149,160</sup> or MR-CI schemes,<sup>14–16,92,94–96,142,144,540–545</sup> where the latter method is more expensive than the former and therefore limited to describe small systems.

In general, the computation of excited-state PESs is much more expensive than the computation of the ground state potential of the same molecule. Not only highly accurate ab initio methods have to be applied for many systems, but also forces and couplings are required for the considered states. A high density of electronic states present in a molecular system can thus increase the costs of a calculation considerably. In this regard, an active, efficient, and meaningful training set generation is indispensable, especially when photodynamics simulations are the target of a study.

Keeping in mind that the quality of the reference data confines the quality of an ML model, several key questions can be identified when designing a study based on ML potentials. We believe the following questions to be important for the selection of a suitable reference method:

- (1) What is the goal of an ML model, and what properties must it predict in order to benefit from the advantages that ML can offer? Are only energy gaps of different electronic states to the electronic ground state necessary, or are gaps between other states and couplings between them also relevant? Especially, the description of couplings requires further consideration, as they cannot be calculated with all quantum chemistry methods and additionally face the problem of random sign jumps along different reaction coordinates.<sup>92,94,546</sup>
- (2) How many excited states are relevant, and which method is computationally affordable to treat the amount of states required? A comparison with experiment and the computation of vertical excitation spectra

with reference methods can help to obtain an answer to this question.

- (3) How large is the system under investigation, and how complex are the excited state processes that are considered to be important? This question is important in order to identify whether single reference methods such as LR-TDDFT or ADC(2) make sense for certain reactions that might occur. While large and flexible molecules with a lot of energetically close-lying states can give rise to a multifaceted photochemistry including dissociation, homolytic bond-breaking, and bond-formation, the dynamics of rigid molecules might only be dominated by one main reaction channel and lose the additional energy in the form of molecular vibrations. The complexity of the excited-state processes can help to estimate the number of necessary data points to describe the relevant configurational space of the molecule.

In case multireference methods are necessary to describe many different excited-state processes of a molecule, the training set generation can become unfeasible. For example, 356 data points were computed for the 15-atom cyclopentoxo molecule with MR-CISD(5,3)/cc-pVD(T)Z.<sup>96</sup> Respective calculations comprised 19,302,445 configuration state functions, and one reaction coordinate could be fitted in the diabatic basis. We also ran into a similar problem when fitting the excited states of the amino acid tyrosine containing 24 atoms, which also requires a multireference treatment. The size of the active space and the number of states needed for an accurate description made multireference methods such as CASSCF or CASPT2 computationally too expensive; see Figure 5. In these cases, the computation of an ample training set is far too expensive with multireference methods, and the quantum chemistry calculations remain the bottleneck even when using ML.

In addition to the aforementioned intricacies to build up a meaningful, yet accurate training set for the excited states, the process is further complicated by the arbitrary phase of the wave function. As a consequence, excited-state properties resulting from two different electronic states, such as transition dipole moments or couplings between different electronic states,<sup>15,16,92,94,95,546</sup> are not uniquely defined and cannot simply be fitted with conventional ML models. Either an additional data preprocessing, termed phase correction, or an adaption of the learning algorithm has to be incorporated to render data learnable with ML models. Details on how to correct these data in advance or during training will be discussed in section 5.3. The subsequent discussion will be dedicated to the training set generation and common training sets applied to date.

## 5.2. Training Set Generation

The requirements and desirable specifications for a training set can vary strongly, dependent on the type of application: When the focus of a study is the investigation of the huge chemical space and the search for certain patterns thereof or the design of new molecules with targeted properties, usually the training set should be as large as possible to cover as many molecules as possible. In the best case, the data points are computed with high accuracy, and this reference method is accurate for the excited states of many different types of systems. In terms of accuracy and general applicability, ab initio methods are more suitable, as they do not require the selection of a density functional, which might be accurate for some cases, but fail for

others. However, the costs and complexity of highly accurate multireference ab initio methods limit their applicability, so that TDDFT remains the method of choice when making predictions throughout the chemical compound space.<sup>152,358,547</sup>

The problems of TDDFT have been discussed very recently by Thawani et al.<sup>548</sup> The authors developed a data set for relevant photoswitches, which are useful, e.g., for medical applications or renewable energy technologies. To this aim, photochemical properties of azobenzenes and associated derivatives were manually extracted from experimental papers. The  $\pi\pi^*$  and  $n\pi^*$  transitions turned out to be key to accurately describe a molecule with photoswitching activity. Different ML models were subsequently trained using different types of descriptors to fingerprint these compounds. Comparable accuracy to TDDFT could be achieved as well as superior performance to human chemists in predicting these transitions. This work highlights very well how data sets can be generated from experiments and provides a practical, useful tool for chemists not versed in TDDFT.

Besides this data set based on experimental data, the most widely applied approach to generate a training set for screening purposes or for the exploration of chemical compound space is to start from an existing (ground-state) database that already covers a large chemical space of certain types of molecules. In this way, not much effort has to be devoted into the exploration of chemical space and structure optimizations to get the most stable conformations of different molecules.

For the purpose of ML-based excited-state dynamics simulations, things look quite different. Note that for photodynamics simulations, only molecule-specific ML models exist until now, which can potentially develop into a universal excited-state force field, but much remains to be done to achieve this goal. Indeed, the generalization of the excited state PESs and corresponding couplings is expected to be a highly complex task, especially due to the problematic generalization of excited states.<sup>94</sup> A comparison of the isoelectronic molecules  $\text{CH}_2\text{NH}_2^+$  and  $\text{C}_2\text{H}_4$  can serve as an example. Their conical intersection between the first excited singlet state and the ground state is accompanied by a rotation along the dihedral angle, which could lead to very similar photoinitiated processes. However, higher-lying excited states are ordered completely different in both molecules and excitation leads to completely different photodynamics.<sup>28,94,549–558</sup> Particularly promising in order to achieve the goal of an excited-state ML force field is the construction of excited-state potentials from atom-wise contributions, i.e., an ML model which learns the surrounding of an atom rather than the molecule as its whole. Promising models are multilayer energy-based fragment methods similar to ref 536 in combination with high-dimensional NNs<sup>509</sup> and density-like descriptors,<sup>508</sup> or automatically learned descriptors based on geometric information.<sup>461,528,529</sup> Especially, ML for wave functions like the SchNOrb model<sup>77,80</sup> could be helpful in this regard and are further interesting for dynamics simulations to compute wave function overlaps from ML. Further, it is important to encode the charge of the molecule in order to treat molecules of same composition, but different electronic charges. However, as it stands, existing ML models for photodynamics simulations are developed to investigate the photoinitiated processes of one specific molecule, which is why we focus on this goal in the following discussion.

Overall, we arrive at the following wish list for the training set, which has been identified also for MD in the ground state:<sup>103,117,515,559</sup> (1) The training set should be as small as possible to keep the number of reference calculations at a minimum. (2) At the same time, the relevant conformational space of the molecule that is required for the reaction under investigation should be sampled comprehensively.<sup>94,117,359,515,535</sup>

Keeping this in mind, an efficient procedure to obtain relevant molecular structures has to be applied. A large number of schemes to achieve this goal have been proposed, which are mainly based on two different strategies: One approach is to simulate MD in the ground and excited states with the reference method and putting much effort into covering critical regions of the PESs comprehensively.<sup>139–141</sup> Structure-based sampling or subsequent clustering is beneficial in this case.<sup>139,140,517,560,561</sup> The other strategy is to use an active learning approach, which decreases the number of necessary reference calculations considerably, but is usually more time-consuming.<sup>515</sup> Noticeable, within ML for quantum chemistry, active learning often refers to an approach, where an initial training set is used to fit an ML model, and this previously learned information is applied to expand the training set.<sup>435</sup> The latter approach is often carried out with the help of MD simulations. Simulation of many trajectories on-the-fly and estimation of the reliability of the ML-fitted PESs at each time step are powerful to identify under-sampled or unknown regions of the PESs. Retraining of the ML models as a data point is added to the training set is required, which makes such a procedure generally expensive. Recently, this active learning procedure has also been adapted in a trajectory-free way,<sup>435,562</sup> which can reduce the costs for the training set generation considerably. The different strategies to generate an ample training set for the excited states will be discussed in the following.

**5.2.1. Basic Sampling Techniques and Existing Databases.** To find patterns within certain groups of molecules, to explore chemical space, and to develop new methods that can fit for example different properties of molecules, such as the valence density used in DFT,<sup>82</sup> or large molecules from small building blocks,<sup>109</sup> a good starting point is often considered to be an already existing database. Prominent examples are the QM databases, namely, QM7, QM7b, QM8, and QM9,<sup>457</sup> which have been used in a large number of publications to date and provide a benchmark for many ML studies.<sup>14,152,461,466,467,526,530,563–565</sup> Especially the QM9<sup>457</sup> data set containing more than 133k small organic molecular structures and corresponding DFT energies, enthalpies, harmonic frequencies, and dipole moments (to name only a few properties) is very popular among the scientific community and has also been used in challenges on kaggle, where researchers and laypersons all over the world can compete against each other to find the most suitable solution to a given task. Prices up to several thousand dollars are quite common.<sup>566</sup> In a similar spirit, the QM9 IPAM ML 2016 challenge requires predicting the energies of QM9 from only 100 training points within chemical accuracy (error of ~0.05 eV).<sup>567</sup>

All aforementioned databases originate from GDB databases<sup>568–570</sup> and are often a subset thereof. The chemical universe GDB databases have been designed using molecular graphs to sample a comprehensive space of molecular

structures for the search of new lead compounds in drug design.<sup>570</sup>

One of the first databases available for the scientific community to treat the excited states of molecules is most probably the QM7b<sup>571</sup> data set, which contains the excitation energies computed with TDDFT for a total amount of >14k molecules with atoms C, N, O, H, S, and Cl. This data set is based on the molecular geometries of the QM7<sup>102,570</sup> data set plus an additional amount of 7211 molecules containing a chlorine atom. The excitation energies of the first singlet state and other properties were recomputed for each optimized molecular geometry. Very similar, the QM8<sup>358</sup> database was developed, based on the GDB-17 database.<sup>572</sup> This data set can be used for the computation of vertical excitation spectra. It hence includes not only the vertical excitation energies of the first excited singlet state, but also the corresponding oscillator strengths. Oscillator strengths are also reported in an autogenerated data set for optoelectronic materials with DFT.<sup>547</sup> Note that the oscillator strength is computed from the squared transition dipole moment, and hence an arbitrary phase factor cancels out and the data does not have to be preprocessed. In addition to the TDDFT energies, CCSD energies are reported, having enabled the development of the so-called  $\Delta$ -learning approach—a powerful way to obtain the accuracy of highly accurate ab initio methods with only a small amount of respective reference calculations. Two ML models are trained in this approach, one on a less accurate method and another one on the difference between the less accurate and higher sophisticated method.<sup>573</sup> This scheme can also be applied multiple times to achieve increasing accuracy with little additional computational effort<sup>359</sup> and has been adapted for spectroscopy in the condensed phase as well.<sup>153</sup>

The QM9 data set has further been the basis of a very recently constructed data set for singlet and triplet states of >13k carbene structures, termed QMspin.<sup>14</sup> A total of 4000 geometries from the QM9 data set were randomly selected, hydrogen atoms were subtracted, and singlet and triplet states were optimized using CASSCF(2,2)/cc-pVDZ-F12 and open-shell restricted KS-DFT with the B3LYP<sup>574,575</sup> functional, respectively. The MR-CI method was subsequently used to compute the electronic energies of singlet and triplet states. This data set has been used to investigate structural and electronic relationships in carbenes, which are important intermediates in many organic reaction networks.<sup>14</sup>

The OE62<sup>576</sup> database, a benchmark data set applicable for spectroscopy, is another descent of several existing data sets, such as the QM8 and QM9 data sets. It consists of >61k organic molecules able to form crystals including up to 174 non-hydrogen atoms. Reported are the orbital energies of molecules computed with DFT/PBE.<sup>577</sup>

Another database, which also contains excited state data, is the PubChemQC database.<sup>578</sup> It contains over three million molecules, whose structures are reported along with the energies at DFT/B3LYP/6-31G\* level of theory. In addition, the excitation energies of at least three million structures are reported for the 10 energetically lowest-lying singlet states at TDDFT/B3LYP/6-31G\* level of theory.

A simple strategy was carried out by Kolb et al.,<sup>579</sup> who used an existing analytical PES to create an ML potential: They randomly sampled data points, trained an ML model, and added more points in regions with deviations from the original PES. Other strategies have been carried out mainly for the fitting of ground state potentials and for materials which are

however also relevant to consider for the excited states. One novel, suitable strategy is for example “de novo exploration” of PESs using a similarity measure provided by ML models.<sup>580</sup> At least for material discovery, this method can be used to omit any additional active learning procedure to converge PESs. Similar strategies are ab initio random structure searching (AIRSS),<sup>581</sup> particle swarm optimization with the CALYPSO method,<sup>582</sup> or USPEX, an evolutionary algorithm.<sup>583</sup> The latter methods are used in the condensed phase and for inorganic systems mainly, but can be adapted for the search of molecular structure, and the desire to compute as few reference data points as possible is also relevant in this case.

A different approach to build a training set is to employ molecule-generating ML models,<sup>165,584,585</sup> such as the recently developed GschNet.<sup>586</sup> Alternatively, MD simulations with the reference method can provide a good starting point for training.<sup>122,507,563</sup> For example, Ye et al.<sup>510</sup> sampled 70k conformations for *N*-methylacetamide via MD simulations with the OPLS force field<sup>587</sup> within GROMACS<sup>588</sup> for subsequent UV spectra calculations. We have applied a similar scheme to generate a training set of SO<sub>2</sub> based on an LVC model.<sup>240</sup> Surface hopping MD simulations with the SHARC method<sup>236,256,589</sup> were carried out with the reference method LVC(MR-CISD) ending up with >200k data points of different conformations of SO<sub>2</sub>.<sup>15</sup> Because of the crude sampling and low cost of the reference method, no emphasis was put on clustering the training set into a smaller, still comprehensive set.

A total of 90k data points were required in an ML-based surface hopping study of CH<sub>2</sub>NH with the Zhu-Nakamura method. Reference data for the ground and first excited singlet state, S<sub>0</sub> and S<sub>1</sub>, were generated with CASSCF(2,2)/6-31G via ground-state and surface hopping MD simulations. The latter method was applied to sample the regions around conical intersections between the S<sub>0</sub> and S<sub>1</sub> state.<sup>141</sup>

Similarly, Hu et al.<sup>139</sup> sampled 200k data points of 6-aminopyrimidine using ground-state and surface hopping MD with CASSCF(10,8)/6-31G\*. State-averaging over three singlet states was applied. In addition, structures that led to hops between different states were used as starting points to find minimum energy conical intersections, and clustering was carried out to reduce the amount of data for training.

One way to select data points more efficiently is a structure-based sampling scheme, as proposed for instance by Ceriotti et al. with sketch map,<sup>560,590,591</sup> an algorithm for dimensionality reduction of atomistic MD simulations or enhanced sampling simulations. Likewise, Dral et al.<sup>140</sup> applied a grid-based sampling method to construct PESs of a model spin-boson Hamiltonian to execute surface hopping MD with KRR. The energetically low-lying regions of the PESs were first sampled via an inexpensive method, and subsequently the distances between the molecular structures were computed. In this way, 10 000 data points were obtained.<sup>140,517</sup> ML models trained on only 1000 data points were accurate enough to reproduce reference dynamics. This approach was compared with random sampling for the methyl chloride molecule and was shown to reduce the amount of training data needed up to 90% for static calculations.<sup>517,561</sup>

Another technique to explore PESs, which is frequently applied for the electronic ground state, is geometry optimization. In the past few years, effort has been devoted toward acceleration of the optimization with ML; see e.g. refs.<sup>592–597</sup> Most often, single-state ML models, which are

mainly based on GPR, are employed to optimize minima, transition states, minimum energy paths, and many more. Very recently, Raggi et al.<sup>597</sup> used internal coordinates and variance-restriction instead of commonly applied Cartesian coordinates in combination with a step restriction. The use of internal coordinates removes transitional and rotational variance and allows for different length scales. The variance measure can be directly obtained from GPR models to restrict the step in the geometry optimization process. The latter enables an exploration of large geometry displacements in some cases, i.e., dependent on the acceptable variance.

**5.2.2. Active Learning.** As shown in the previous section, training sets with the respective equilibrium structure of a large number of molecules are very powerful for investigating the huge chemical space or for the design of new molecules. However, the usefulness of such training sets for photodynamics is rather questionable. The reason for this deficiency is that, especially in MD simulations in the excited states, the excess of energy carried by a molecule very quickly leads to conformations that are far beyond the equilibrium structure and most likely far away from originally sampled structures. The formation and breaking of bonds is quite common in photodynamics simulations and is usually only accessible from an excited, dissociative state. The use of photodynamics simulations with the reference method could solve this problem, but is not feasible if specific reactions occur on a rather slow time scale or if many different processes take place.<sup>59,171,175,178,257,367</sup> As previous studies have shown, inefficient sampling techniques lead to a huge amount of data, which still does not guarantee that the training set is comprehensive enough for excited-state MLMD simulations. In fact, ML models fail dramatically in undersampled and extrapolative regions of the PESs. A smarter sampling technique is advantageous in these cases in order to efficiently identify such undersampled regions and build trustworthy ML models.

Active learning, where ML “asks” for its training data, is one solution to create a data set more efficiently. An example from chemistry is the adaptation of an initially generated training set due to an uncertainty measure for ML models trained on this initial training set. This concept has already been introduced in 1992 as query by committee<sup>598</sup> and has been adapted for quantum chemistry quite fast due to the required fitting and interpolation of PESs for grid-based quantum dynamics simulations. Pioneering works by Collins and co-workers<sup>150,151,408,599</sup> applied modified Shepard interpolation to fit PESs and iteratively adapt them in out-of-confidence regions using the GROW algorithm.<sup>599,600</sup> The first version similar to query by committee for the fitting of NNs was proposed by Artrith and Behler.<sup>601</sup> Since then, several sampling techniques have been developed that are based on MD and an extension of databases using interpolation moving least-squares,<sup>602,603</sup> permutation invariant polynomial fitting,<sup>522,524</sup> and different ML models for the ground state<sup>101,103,111,508,515,559,604–616</sup> and also excited states.<sup>15,94,140</sup>

As active learning starts from already trained ML models, an initial training set has to be provided. Some strategies to provide this initial reference data set will be discussed, following strategies applied to adapt this initial training set. Note that all previously discussed methods can be similarly applied to generate an initial training set. Although we cannot give a general guide on how large a training set should be, in our cases, it was beneficial to cover approximately one-fourth

of the training set with data obtained from initial sampling and the rest with data obtained from active learning techniques. About 300–400 data points per degree of freedom of a molecule turned out to be sufficient at least for small molecules.

**Initial Training Set.** In general, an initial training set can be obtained in many different ways. As photoinitiated MD simulations usually start from vertical excitation of the ground state equilibrium geometry, this structure is commonly used as the starting point and reference geometry for the training set generation. In principle, any technique can be applied to then add conformations to obtain a preliminary training set. A good starting guess is to use normal modes of a molecule, as they are generally important for dynamics. In two recent works, we carried out scans along different normal modes and combinations thereof to sample conformations of small molecules.<sup>15,94</sup> Normal modes were also sampled for generating ANI-1 NN PESs.<sup>115</sup> For the excited states, it is generally favorable to include critical regions of the molecule in the initial training set by carrying out optimization of these geometries and including the calculations into the training set.<sup>94,139</sup>

When small molecules are targeted, this initial training set can already be comprehensive to start the training of ML models and adapt the training set based on an uncertainty measure provided by the ML models.<sup>94</sup> In case more flexible and larger molecules are studied that give rise to a complex photochemistry and a high density of states including different spin multiplicities, a small initial training set might not be sufficient, and a larger conformational space of the molecule needs to be sampled. This can be done for example via Wigner sampling<sup>475</sup> and also with MD simulations in the ground state.<sup>617,618</sup> Suitable methods are for example umbrella sampling,<sup>619</sup> trajectory-guided sampling,<sup>620</sup> enhanced sampling,<sup>621</sup> or metadynamics<sup>622</sup> in combination with a cheap electronic structure method such as the semiempirical tight-binding based quantum chemistry method GFN2-xTB<sup>623</sup> or existing ground-state force fields. A large amount of different geometries can be created very fast and inexpensively, which then can be clustered to exclude similar conformations of the molecule to keep the number of reference simulations at a minimum. The selected data points for the training set can then be computed with the chosen reference method, whose accuracy is targeted with ML. Additionally, if certain reaction coordinates have been shown to be important in experiments or previous studies, then it is favorable to include data from scans along these reaction coordinates.<sup>96,175</sup>

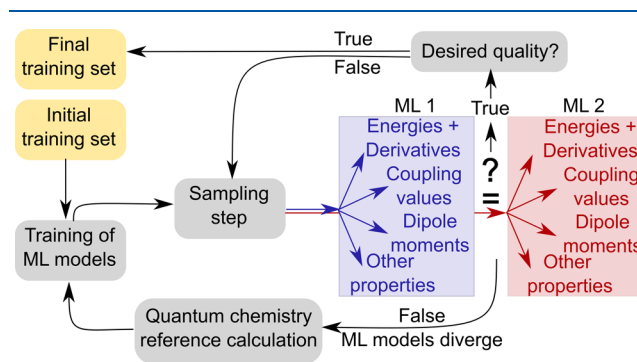
As soon as meaningful ML models can be obtained from the initial training set, active learning techniques can be applied to enlarge the set. What number of data points turns out to be sufficient for the initial training set is dependent on a lot of different factors, such as the size and flexibility of the molecule under investigation, the number of excited electronic states described, and the ML model and descriptor applied.<sup>94,95</sup> In order to give a ballpark figure, we note that we used approximately 1000 data points as initial training set for small molecules in recent studies using deep multilayer feed-forward NNs.<sup>15,94</sup>

**Strategies for Actively Expanding the Training Set.** The next step in active learning is to expand the initial training set by adding points from out-of-confidence regions. The detection of these undersampled regions can be done in

many different ways, whereby most approaches rely on MD simulations.

Among the most popular strategies is the iterative sampling scheme of Behler,<sup>515</sup> originally developed for fitting ground-state PESs. Today, it is widely used, see for example refs,<sup>103,559,624</sup> and has been modified as a so-called adaptive sampling approach.<sup>111</sup> The latter has been adapted by us for the generation of a training set for the excited state PESs of molecules including couplings.<sup>94</sup> The basis of almost any iterative or adaptive sampling scheme is a similarity measure to judge whether a molecular geometry can be predicted reliably with ML models or not. While kernel methods intrinsically provide a measure of similarity for each molecular geometry, NNs do not. Therefore, adaptive sampling with NNs requires at least two ML models. In the case of KRR or GPR, two ML models can be used as well, but are not necessarily needed. Indeed, the statistical uncertainty estimate of the predictions remains a huge advantage of GPR models.<sup>525,535,625</sup> As a remark, from a materials' perspective, Gaussian approximation potentials (GAPs) can be used as a similarly useful tool to provide such an uncertainty measure.<sup>121,626</sup>

The adaptive sampling scheme for the excited states is illustrated in Figure 9 and exemplified with two ML models.



**Figure 9.** Adaptive sampling scheme illustrated using two ML models (blue and red blocks). The active learning procedure starts from an initial, preliminary training set (yellow), which is used to train ML models. A sampling step, e.g., a time step of an MD simulation, is executed: The ML models take the molecular geometry of the sampling step as an input and predict the energies of the considered excited states, their derivatives, and additional required photochemical properties. In case the predictions of the ML models are deemed to be different, quantum chemical reference calculations are carried out, ML models are retrained, and the serial steps are carried out again. This procedure is executed until the desired quality of the ML PESs is attained in order to sufficiently describe the chemical problem under investigation.

The whole process starts with an initial training set, which is used to train the two (or more) preliminary ML models. These models differ in their initial weights or model parameters. The resulting dissimilar ML architectures guarantee that the ML models do not predict the exact same number for a given molecular input. The hypothesis underlying this scheme is that inferences of different ML models trained on the same training set will be similar to each other as long as an interpolative regime is given. The inferences of the ML models are inaccurate and should differ from each other to a much larger extent if a molecular input lies in an unknown or undersampled region of the PESs.

In order to find such regions, sampling steps are carried out, e.g., by running (excited-state) MD simulations based on the

mean of the inferences made by the different ML models for energies,  $\bar{E}^{\text{ML}}$ , forces,  $\bar{F}^{\text{ML}}$ , and if required also couplings,  $\bar{C}^{\text{ML}}$ . In each sampling step, the variances for each predicted property are computed. In the present example, energies and forces are treated together as  $\sigma_{E+F}^{\text{ML}}$  (but can also be used separately), separately from variance of the couplings  $\sigma_C^{\text{ML}}$ . If a variance exceeds a predefined threshold, the ML models diverge, and the predictions are deemed untrustworthy.  $N_{\text{ML}}$  refers to the number of different ML models,  $\zeta$ , used for adaptive sampling:

$$\sigma_{E+F}^{\text{ML}} = \frac{1}{N_S} \sum_i \left( \sqrt{\frac{1}{N_{\text{ML}} - 1} \sum_{\zeta=1}^{N_{\text{ML}}} (E_{i\zeta}^{\text{ML}} - \bar{E}_i^{\text{ML}})^2} + \sqrt{\frac{1}{N_{\text{ML}} - 1} \sum_{\zeta=1}^{N_{\text{ML}}} \left( \frac{1}{3N_A} \sum_a (F_{i\zeta,a}^{\text{ML}} - \bar{F}_a^{\text{ML}})^2 \right)} \right) \quad (24)$$

$$\sigma_C^{\text{ML}} = \frac{1}{2N_S^2} \sum_i \sum_j \sqrt{\frac{1}{N_{\text{ML}} - 1} \sum_{\zeta=1}^{N_{\text{ML}}} (C_{ij\zeta}^{\text{ML}} - \bar{C}_{ij}^{\text{ML}})^2} \quad (25)$$

Note that the variance is averaged over all states for energies and forces and over all pairs of states for couplings, that are described with the ML models. As a variant, each state could also be treated separately. However, as the different electronic states are not independent of each other, a mean-treatment is assumed to be advantageous.<sup>95</sup>

Each data point that is predicted with a variance larger than the predefined threshold for a given property, is recomputed with the reference quantum chemistry method and added to the training set. In this way, under-sampled or generally unknown regions of the PESs are identified. Whenever the variance of each property is within the range that is thought to be reliable, the mean of the inferences is forwarded to the MD program to propagate the nuclei and continue MLMD simulations. The name adaptive sampling is based on the recommendation to choose a rather large threshold in the beginning of the adaptive sampling procedure and to adapt this threshold to smaller values as the ML models become more accurate and robust.<sup>111</sup> A first estimate for the initial value of a threshold can be obtained from the mean absolute error (MAE) of the corresponding ML model on the initial training set.

In principle, adaptive sampling can be carried out for every property, that should be represented with ML potentials, and is not restricted to energies, forces, and couplings. Similarly, it does not need to be executed with excited-state dynamics, but could also be done with ground-state MD or any sampling method that is considered to be suitable.

As a negative side effect, this procedure is generally more time-consuming than many other sampling techniques because ML models have to be trained each time a new data point is added to the training set. To apply adaptive sampling in a more efficient way, it is advantageous to execute not only one ML trajectory, but many hundred trajectories in parallel, as it is usually done in MD simulations. The ML models should then only be retrained, when all ML-based trajectories have reached an undersampled conformational region.<sup>94,111,515</sup> Despite the higher complexity of adaptive sampling compared to random sampling, it can reduce the number of required data points for

MLMD simulations substantially. In this regard, also the computational costs for the training set generation can be kept at a minimum.

Adaptive sampling was carried out successfully to generate a training set of 4000 data points of  $\text{CH}_2\text{NH}_2^+$  containing three singlet states and couplings. ML-based surface hopping MD simulation could be carried out on long time scales using the average of two deep NNs. The concept of iterative sampling also proved beneficial for the long MD simulation to guarantee accurate ML potentials throughout the production run. Here, the threshold was not adapted anymore, and the MD was continued from the current geometry after a training cycle was completed.<sup>94</sup> In addition, the average of more NNs turned out to be more accurate than the prediction of only one NN, which was also shown in ref 111.

Another quality control besides the property-based one proposed by Behler can be obtained by comparing the molecular structures at each time step as done by Dral et al.<sup>140,517</sup> and Ceriotti et al.<sup>560</sup> A combination of a structure-based and property-based detection of sparsely sampled regions of the PESs has been done by Zhang et al. and Guo et al.<sup>392,607,627–629</sup> Very recently, an alternative approach has been applied with NNs by Lin et al.<sup>435</sup> that does not require MD simulations. It is based on the finding that the negative of the squared difference surface obtained from NNs approaches zero in regions, where no data points are available.<sup>603</sup> Therefore, new points can be computed at the minima of the negative squared difference surfaces of at least two NNs (or, equivalently, at local maxima of the squared difference surface). This method is supposed to be very efficient in cases, where different conformations are separated by large energy barriers or strongly stabilized local minima are common. MD simulations would take a long time to overcome the potential barriers and reach the region of unknown molecular structures.<sup>435</sup>

The idea behind this technique is similar to previous works with GPR. A measure of confidence can be provided with GPR models, which enables the search of regions with large variance in the ML predictions. In these regions, data points can be added to build up a training set.<sup>562,630–632</sup> Similarly, Bayesian Optimisation Structure Search (BOSS) has been proposed for constructing energy landscapes of organic and inorganic interfaces.<sup>633</sup> A combination of different approaches has also been applied by Häse et al.,<sup>162</sup> who fitted TDDEFT excited-state energies of a light-harvesting system. Given a large enough, error-free, and comprehensive data set, ML has the potential to determine known and unknown (un)physical laws within the data.<sup>634</sup>

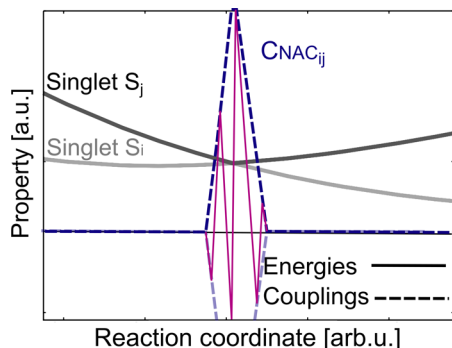
### 5.3. Phase of the Wave Function

In contrast to ground state properties, excited-state properties such as transition dipole moments, NACs, or SOCs arise from two different electronic states. As a consequence of the arbitrary phase of the wave function of each electronic state, properties resulting from two different states carry an arbitrary sign, which makes them generally double-valued. In the case of vectorial properties, such as dipole moments or coupling vectors, the whole vector can be multiplied by +1 or -1 and is still a valid solution. Similarly, single valued properties, such as SOCs obtained from electronic structure programs, can be multiplied by +1 or -1 and are equally correct. This additional complexity prohibits that conventional ML algorithms learn such raw data of quantum chemistry and hampers the training



process to find a proper relation between a molecular geometry and the excited-state property.<sup>94,546</sup>

A one-dimensional example of this problem is illustrated for the NAC (exemplified using one single value along the reaction coordinate) that couples an excited singlet state,  $S_j$ , and a second excited singlet state,  $S_i$ , in Figure 10. A positively signed



**Figure 10.** NAC value between singlet state  $S_i$  and  $S_j$  in the MCH basis. A consistent sign along the reaction path of couplings is shown by blue dashed lines. The direct output of a quantum chemical calculation is shown by a magenta line.

function of atomic coordinates is shown by dashed blue lines with a cusp at the point at which the two singlet states are degenerate. Such a smooth function (besides the sharp spike at the conical intersection) is highly desirable when fitting with ML models is aimed for. It is worth mentioning that a consistent negative sign (light-blue dashed line) along this reaction coordinate is equally correct and that it is desirable to seek for one global sign. However, the direct output of a quantum chemistry program along this reaction coordinate looks more similar to the dashed magenta line in-between the blue curves. As one can imagine, no proper training can be guaranteed with these inconsistent data. Note that existing MD programs for the excited states usually track such phase jumps within electronic wave functions in order to account for nonadiabatic transitions correctly.<sup>256</sup>

The idea of phase tracking can also be applied in ML in order to thwart the problems due to the arbitrariness within coupling or dipole elements. Some algorithms have been developed to remove the arbitrary sign jumps and provide smooth functions of atomic coordinates.<sup>15,16,94,635</sup> Noticeably, the properties obtained after a transformation to the diabatic basis are already smoothly varying functions of atomic coordinates.<sup>369</sup> However, the challenges arising due to the arbitrary phase of the wave function still persist, because the inconsistencies within adiabatic properties have to be removed in order to make the diabaticization process feasible.<sup>16,92</sup>

It is worth mentioning at this point that also another kind of phase exists that cannot be eliminated in the aforementioned way. It is called the Berry phase or geometric phase. After a loop was performed in space around a conical intersection and returning to the original point, a change in the phase of the wave function of  $\pi$  can be observed; i.e., the same point is only reached after two loops around the conical intersection. Neglecting this effect can lead to false transition probabilities, depending on the dynamics method and the system. While in most cases in MQCD the Berry phase can be safely neglected, this is not possible in quantum dynamics simulations. A diabatic basis is advantageous in this case because the Berry phase is absent in this picture. However, the Berry phase still

has to be kept in mind, when fitting diabatic potentials from adiabatic ones.<sup>636–641</sup>

**5.3.1. Phase Correction of Adiabatic Data.** First ML studies on dynamics in the adiabatic basis omitted a preprocessing and were unable to reproduce reference results based on ML alone,<sup>140</sup> or avoided the phase problem by using the Zhu–Nakamura method.<sup>139,141</sup> Evidently, potentials and forces can be learned with conventional ML approaches, but adaptations or a preprocessing of data is necessary to learn coupling elements or transition dipole moments. Independent of the purpose, the fitting of adiabatic quantities<sup>94,546</sup> or the diabaticization of adiabatic data with property-dependent diabaticization schemes,<sup>16</sup> adiabatic data have to be corrected to remove the arbitrary sign jumps that are due to the arbitrary phase of the wave function. Several ways for these corrections exist, which have been shown to work well for different excited-state problems.

One possibility is to preprocess data according to the wave function overlap—between the wave functions from a geometry of interest and a reference geometry—for each electronic state. This process is termed phase correction<sup>256,546</sup> and has been applied by us in order to generate a training set for three singlet states of  $\text{CH}_2\text{NH}_2^+$ <sup>94</sup> and two singlet and two triplet states of  $\text{CSH}_2$ . SOCs,<sup>15</sup> NACs,<sup>15,94,95</sup> and transition dipole moments<sup>94,95</sup> could be fitted in the adiabatic basis with deep NNs and kernel ridge regression (KRR).<sup>15,94,95</sup> Very recently, Zhang et al.<sup>97</sup> applied this procedure to describe transition dipole moments of *N*-methylacetamide.

The wave function overlap matrix,  $\mathbf{S}$ , with size  $N_S \times N_S$ , is computed between two molecular geometries  $\alpha$  and  $\beta$ :<sup>642</sup>

$$\mathbf{S} = \langle \Psi_\alpha | \Psi_\beta \rangle \quad (26)$$

In many cases along a given reaction path, the off-diagonal elements of the overlap matrix are very close to zero, and the diagonal elements are very close to +1 or −1, indicating whether the phase of a state has changed along this path or not. Whenever a new state enters along the reaction path or adiabatic states switch their character, which is common after passing through a conical intersection for example, the off-diagonal elements provide the relevant phase information instead of the diagonal elements. Taking all these effects into account, a phase vector,  $\mathbf{p}$ , can be derived for each given molecular geometry. A property resulting from electronic state  $i$  and  $j$  has to be multiplied by the corresponding phase factors of these states.<sup>94</sup>

An advantage of this algorithm is that it does not require any manual fitting of data. However, this procedure has to be carried out for every data point included in the training set with respect to one predefined reference wave function. This reference wave function can be for example the wave function of the ground-state equilibrium structure of the molecule and needs to be identified to guarantee an almost globally consistent sign of elements. During a photoinitiated simulation, it is common that geometries quickly start to differ from the reference geometry. The wave function overlap then tends to zero and cannot provide information about the correct sign of a certain electronic state. In this case, the phase must be propagated from the reference geometry on with  $n$  interpolation steps. The phase vector applicable for the correction of the data point to be included in the training set is then obtained by multiplication with all previously obtained phase vectors,  $p_0$  to  $p_{n-1}$ :

$$p = \prod_{\alpha=0}^{n-1} p_{\alpha} \quad (27)$$

Intruder states prohibit a proper tracking because their wave function is absent at the earlier geometries. Hence, a phase correction may be rendered infeasible for systems with a high density of states.

In order to obtain the correct phase, more states can be included in the simulations, which however increases the computational cost. A solution is to take many electronic states into account only close to the reference geometry. The amount of states can then be reduced along a given reaction coordinate, and relevant states can be disentangled from irrelevant ones. Further, it makes sense to save the already phase-corrected wave functions of several geometries in addition to the reference geometry. Whenever a new data point should be included into the training set, the distance to each saved data point can be computed in order to find the closest available structure and reduce the amount of interpolation steps.<sup>94,175</sup>

This problem has also been recognized by Robertson et al.<sup>391</sup> for a diabatic process, where a sufficiently large vector space of the CAS wave function is required for proper diabaticization. The overlaps of electronic states can be maximized by rotation of CI vectors of CAS wave function states. A similar version to use the information on CI vectors for diabaticization was applied by Williams et al.,<sup>142</sup> who used NNs to assist the diabaticization process of adiabatic NO<sub>3</sub> potentials.

Another way to correct the sign of data points was carried out by Guan et al.,<sup>16</sup> who fitted diabatic 1,2<sup>1</sup>A PESs and dipole moment surfaces of NH<sub>3</sub> from MR-CISD/aug-cc-pVTZ data with NNs. The diabatic PESs were taken from a previous study and obtained with the Zhu–Yarkony diabaticization procedure.<sup>377,643,644</sup> By diagonalization, the rotation matrix defined in eq 10 could be obtained, which connects the diabatic and the adiabatic basis (see eq 9). The adiabatic dipole moments,  $\mu^{\text{MCH}}$ , could then be transformed into the diabatic basis using the unitary matrix,  $\mathbf{U}$ :

$$\mu^{\text{diab}} = \mathbf{U}\mu^{\text{MCH}}\mathbf{U}^{\dagger} \quad (28)$$

As the unitary matrix  $\mathbf{U}$  is only defined up to an arbitrary sign, the signs of the diabatic dipole moments have to be corrected in order to provide a consistent diabatic dipole moment surface. This correction has been done with a so-called cluster growing algorithm.<sup>635</sup>

The cluster growing algorithm requires an initial set of phase corrected data points. In this work, 347 data points were adjusted manually for this purpose. Subsequently, a Gaussian process regression (GPR) model<sup>645</sup> was fitted to these data points. The signs of the rest of the data points to be corrected were then adjusted with the GPR model. Several iterations were carried out, where each iteration aims for the inclusion of close-lying points to the cluster, leading to the name “cluster growing” algorithm.<sup>148</sup>

The singularities in regions close to conical intersections can make this algorithm fail. Therefore, data points in such regions have been removed by setting a threshold. Data points with energy gaps lower than this threshold were excluded from the cluster. The regions around conical intersections could not be fitted as comprehensively as other regions of the PESs. As another drawback, the authors note that the initial manual

fitting of the signs is a tedious task, especially when larger systems and more dimensions are described.

Two of the authors also fitted diabatic PESs of two singlet states and one triplet state as well as the SOCs between singlets and triplets of formaldehyde, CH<sub>2</sub>O, with NNs.<sup>92</sup> The electronic structure reference method was MR-CISD/cc-pVTZ. The diabatic potentials were obtained using an adapted version of the Boys localization.<sup>382</sup> The energy differences between two states were incorporated in the equations in order to remove earlier identified diabolic singularities.<sup>148</sup> The range of  $\pi$ , which the rotation angle for the diabaticization covers, guarantees a proper treatment of the Berry phase. The diabaticization procedure further requires consistent transition dipole moments, which were adjusted manually for this purpose. The diabatic SOCs were then obtained as a linear combination of the adiabatic SOCs by applying the same rotation matrix as for the energies. One separate NN function was used to fit each coupling value and electronic state separately.

It becomes clear that only a small number of works on this topic exist. At the moment, many problems remain unsolved for generating a training set that properly accounts for both types of phases, the arbitrary phase and the Berry phase, and is applicable for large systems with many states. An automatic phase correction procedure without the need of manual input would be very advantageous, especially when larger and more flexible systems are treated. Further developments are needed.

**5.3.2. ML-Based Internal Phase Correction.** One step toward a routine application of ML for photochemical studies and an easier training set generation with quantum chemistry is an ML-based internal phase correction, which has been implemented by us into the SchNarc approach for photochemical simulations.<sup>15</sup> In contrast to the phase correction algorithm to correct the training data, this procedure renders the learning of inconsistent quantum chemical data possible. A modification of the training process, termed phase-free training, is required for this purpose.<sup>15</sup> We implemented this training algorithm in a combination of the deep continuous-filter convolutional-layer NN SchNet,<sup>461,465</sup> adapted for excited states, and the MD program SHARC.<sup>236,256,589</sup>

Similar to standard training algorithms, parameters of an ML model are optimized in order to minimize a cost function. Most frequently, the L<sub>1</sub> or L<sub>2</sub> loss functions are applied, which take the mean absolute error or mean squared error between predicted and reference data into account. The phase-free training algorithm uses a phase-less loss function, which includes all trained properties at once and additionally removes the influence of the random phase switches. In this way, the computational costs for the training set generation can be reduced.

Compared to the previously reported ML models for photochemistry, where each state was fitted independently,<sup>16,92,141</sup> SchNarc is capable of describing all PESs at once, including the elements resulting from different pairs of states. This results in an overall loss function with several terms, where each term is weighted with a different trade-off value,  $t$ , that can be defined manually:

$$L_{\text{ph}} = t_{\text{E}} \left\| E^{\text{QC}} - E^{\text{ML}} \right\|^2 + t_{\text{F}} \left\| F^{\text{QC}} - F^{\text{ML}} \right\|^2 + t_{\text{SOC}} \cdot L_{\text{SOC}} + t_{\text{NAC}} \cdot L_{\text{NAC}} \quad (29)$$

If only energies ( $E$ ) and forces ( $F$ ) are fitted, then the loss function is equal to a linear combination of  $L_2$  loss functions for energies and forces.<sup>15,111</sup> The parts of the SOCs and NACs are

$$L_{\text{SOC}} = \min(|\varepsilon_{\text{SOC}}^\kappa|) \text{ with } 0 \leq \kappa \leq 2^{N_s-1} \quad (30)$$

and

$$L_{\text{NAC}} = \min(|\varepsilon_{\text{NAC}}^\kappa|) \text{ with } 0 \leq \kappa \leq 2^{N_s-1} \quad (31)$$

respectively. The error for SOCs and NACs that enters the loss function is the minimum error that can be achieved when trying out all possible combinations of phases for each pair of states, i.e.,  $2^{N_s-1}$  possible solutions. The algorithm takes into account that the signs of SOCs and NACs coupling different pairs of states depend on each other.

The error function containing all possible solutions for SOCs,  $\varepsilon_{\text{SOC}}^\kappa$ , and NACs,  $\varepsilon_{\text{NAC}}^\kappa$ , can be obtained as follows:

$$\varepsilon_{\text{SOC}}^\kappa = \frac{1}{N_s^2} \sum_{i=1}^{N_s} \sum_{i \neq j}^{N_s} \left\| C_{\text{SOC}_{ij}}^{\text{QC}} - C_{\text{SOC}_{ij}}^{\text{ML}} \cdot p_i^\kappa \cdot p_j^\kappa \right\|^2$$

with  $0 \leq \kappa \leq 2^{N_s-1}$  (32)

$$\varepsilon_{\text{NAC}}^\kappa = \frac{1}{N_s^2} \sum_{i=1}^{N_s} \sum_{i \neq j}^{N_s} \frac{1}{N_A} \sum_{a=1}^{N_A} \left\| C_{\text{NAC}_{ij,a}}^{\text{QC}} - C_{\text{NAC}_{ij,a}}^{\text{ML}} \cdot p_i^\kappa \cdot p_j^\kappa \right\|^2$$

with  $0 \leq \kappa \leq 2^{N_s-1}$  (33)

This phase-less loss procedure does not require any preprocessing of training data. Quantum chemistry calculations can be directly fitted with this adaption of the loss function. The power of this approach is that, once a given phase vector for a data point has been found, it can be directly applied to correct the arbitrary signs of other properties, such as transition dipole moments. If other properties are targeted, the loss function applied for NACs can be similarly used for other vectorial properties, and the loss function applied for SOCs can be used for any other single- or complex-valued element of arbitrary sign.<sup>15</sup> However, as a consequence of the higher complexity of the loss function, the training process is generally more expensive. The computational effort required for training can be reduced if only one type of coupling is treated within MD simulations. In these cases, a simpler adaption of the phase-free loss is also applicable.<sup>15</sup>

## 6. APPLICATION OF ML FOR EXCITED STATES

In this section, we review ML studies of excited states and their properties. We aim to show how they have been employed to improve static and dynamics calculations and focus on the used type of regressor, descriptor, training set, and property. We will classify the approaches according to Figure 1.

### 6.1. Parameters for Quantum Chemistry

Traditionally, the user decides whether a multireference method is needed or a single-reference method is sufficient to describe a chemical problem. Recently, Kulik and co-workers<sup>176</sup> have presented an NN model based on a semisupervised virtual adversarial training approach and a diagnostic inputs training set<sup>177</sup> to learn the multireference character of molecular systems. The authors developed a decision engine to detect strongly correlated systems, which can be applied in a high-throughput screening fashion. Their work can potentially pave the way toward automatic selection

of a proper reference method—a tool that is urgently needed, especially in the research field of ML for excited states.

Besides this seminal work, ML can help to select an active space for multireference methods. Jeong et al.<sup>71</sup> developed an ML protocol for classification based on XGBoost<sup>483</sup> to allow for a “black box” use of many multireference methods by automatically selecting the relevant active space for molecular systems. The tedious selection of active orbitals and active electrons can thus be avoided. The accuracy of this approach was demonstrated for diatomic molecules in the dissociation limit, and the molecules were represented via the molecular orbital bond order and the average electronegativity of the system.

### 6.2. ML of Primary Outputs

To the best of our knowledge, no ML models for providing primary outputs of quantum chemistry exist for excited states (see Figure 1). Targeting the primary output of a quantum chemistry simulation, i.e., the  $N$ -electron wave function, or providing ML density (functionals) is far from trivial even for ground-state problems.<sup>72–80,82,90,355,646–649</sup> However, such an approach for excited states could solve many problems and allow for wave function analysis, providing additional insights like the excited state characters.<sup>650</sup> Therefore, we expect such models to appear in the near future.

### 6.3. ML of Secondary Outputs

In the following, we summarize the contributions of ML models that fit the secondary output of quantum chemical calculations, i.e., PESs, SOCs, NACs, and transition as well as permanent dipole moments in the adiabatic and diabatic basis (Figure 1). The prediction of the manifold quantities (see Figure 2) can be done in two ways, i.e., in a single-state fashion and in a multistate fashion.<sup>95</sup> The applicability of such ML models to the simulation of photodynamics will be discussed.

**6.3.1. ML in the Diabatic Basis.** Diabatic PESs have been fitted with ML and related methods for more than 25 years.<sup>151,408</sup> An advantage of diabatic PESs is their smoothness, which is perfectly matched by ML models built upon smooth functions. However, the tedious procedure to generate diabatic PESs remains. Some effort is therefore devoted to develop ML-assisted diabaticization procedures and eliminate this limiting step.

*Diabatization.* Williams et al.<sup>142</sup> incorporated NNs into diabaticization by ansatz and fit diabatic  $\text{NO}_3$  PESs. The ground state vibrational energy levels were computed, and subsequently, the authors used the diabatic potentials for quantum dynamics simulations in five dimensions.<sup>381</sup> The diabaticization procedure was further modified to properly account for complete nuclear permutation inversion (CNPI) invariance.<sup>390</sup> To this aim, the molecular input was replaced by CNPI invariant coordinates. Recently, Shen and Yarkony<sup>96</sup> fit two diabatic potentials of the cyclopentoxy radical,  $\text{C}_5\text{H}_9\text{O}$ , and one state of cyclopentoxide,  $\text{C}_5\text{H}_9\text{O}^-$ , with 356 data points sampled from scans along different reaction coordinates. The diabaticization was assisted with NNs. Because of the high dimensionality of the system, the authors resort to application of regularization in the fitting algorithm and an adapted loss function to obtain an accurate representation of two-state diabatic PESs with NNs. This novel strategy is envisioned for the computation of the photoelectron spectrum of cyclopentoxide.<sup>96</sup> Fitting 39 degrees of freedom in the diabatic basis is a huge improvement in this research field. The authors further note that a comprehensive sampling of the full relevant

PESs in such high dimensional space is problematic. PIP-NNs were used to introduce a new diabaticization procedure recently.<sup>651</sup> The ground and first excited state of ammonia served as a test molecule. Four separate NNs were trained to fit the parts of the diabatic Hamiltonian, while the loss function was formulated in the adiabatic basis and adiabatic energies, forces, and adapted derivative couplings obtained from the diabatic fitted Hamiltonian after diagonalization were mapped to the adiabatic reference data.

Recently, Shu et al.<sup>652</sup> proposed a new, semiautomatic diabaticization approach based on two training sets—one is formed from adiabatic energies, and one contains a selected number of diabatic potential energy matrix elements, which are assumed to be known at some molecular conformations, e.g., dissociated geometries or the equilibrium structure. The authors further demonstrate that results from diabaticization in lower dimensions can be used for higher-dimensional diabaticization. The diabaticization scheme was tested using a two-state analytical model and two states of thiophenol with adiabatic data obtained from Extended Multi-Configuration Quasi-Degenerate Perturbation Theory (XMC-QDPT).<sup>653</sup>

Because of the aforementioned problems, a description of medium-sized to large molecules with diabatic potentials is often done with more crude approximations.<sup>142,407</sup> An example is the LVC model,<sup>239</sup> with its one-shot variant,<sup>240</sup> or the exciton model.<sup>182,654</sup> For more details on this topic, the reader is referred to refs 64, 239, 331, and 655–657. The Frenkel exciton Hamiltonian can be used to describe light-harvesting systems.<sup>182,654</sup> Such a Hamiltonian was constructed for the investigation of the excited state energies of bacteriochlorophylls of the Fenna–Matthews–Olson complex. Multilayer feed-forward NNs with the Coulomb matrix as a molecular descriptor could accelerate the construction of such Hamiltonians for the prediction of excited-state energies.<sup>162</sup> The effective Hamiltonian of the whole complex was subsequently used to predict excitation energy transfer times and efficiencies. Therefore, Häse et al. used exciton Hamiltonians as an input to NNs, which were trained to reproduce excitation transfer times and efficiencies of pigments in the complex. The excitation energy transfer properties for the training set were computed via hierarchical equation of motion technique,<sup>658</sup> which is costly and thus limited due to the large number of pigments that need to be computed. By using ML to learn the relation between a Frenkel Hamiltonian and excitation energy transfer properties, large-scale screening studies are enabled and an efficient design and search of novel excitonic devices becomes possible.<sup>64</sup> The hyperparameters of the ML model were optimized by applying a Bayesian optimization algorithm. Overall, the accuracy of the model was in excellent agreement to reference data, and out-of-sample Hamiltonians could be computed reasonably well for geometries close to those inside of the training set.<sup>161</sup> Recently, Krämer et al. have used DFT Tight Binding data to train a KRR model to simulate the exciton transfer properties of anthracene. Although the semiempirical reference method is computationally efficient compared to *ab initio* methods, an acceleration was achieved that is concluded to be even more pronounced in larger systems. A perspective on ML for the prediction of phenomena related to light-harvesting systems is provided in ref 659.

**Fitting Diabatic Potentials and Properties.** Given diabatic PESs, ML models can be used to fit them. KRR models are often employed for this task, due to their ease of use and ability to provide accurate predictions, as mentioned above. Recent

studies by Habershon and co-workers focus on interpolation of diabatic PESs and their use for grid-based quantum dynamics methods, i.e., variational Gaussian wavepackets and MCTDH. The butatriene cation has been investigated in two dimensions comprising two electronic states.<sup>149</sup> The description of this molecule has been recently advanced with a new diabaticization scheme, namely, Procrustes diabaticization. The method was evaluated with two-state direct-dynamics MCTDH (DD-MCTDH) simulations of LiF and applied to four electronic states of butatriene.<sup>252</sup> Some of the authors also carried out DD-MCTDH 4-mode/2-state<sup>145</sup> and subsequently 12-mode/2-state dynamics of pyrazine.<sup>146</sup> The investigation of the higher-dimensional space of pyrazine could be achieved by systematic tensor decomposition of KRR and advances conventional MCTDH simulations considerably with respect to accuracy and computational efficiency. Further, the method was applied to investigate the ultrafast photodynamics of mycosporine-like amino acids, which are suitable as ingredients in sunscreens due to their photochemical properties and photostability.<sup>660</sup> However, the reduced 6-dimensional and 14-dimensional DD-MCTDH simulations with KRR interpolated PESs were unable to reproduce the expected ultrafast photodynamics, which had been observed in previously performed surface hopping calculations and is typical for sunscreen ingredients. The authors note that the inclusion of more adiabatic states for the diabaticization procedure and the consideration of additional relevant modes can lead to more accurate results. All of the reference simulations were carried out at the CASSCF level of theory with KRR fitted diabatic PESs.

In addition to KRR models, NNs were also used to describe diabatic PESs. Seminal works include PIP-based NNs by Guo, Yarkony, and co-workers. Absorption spectra and the dynamics of excited states of NH<sub>3</sub> and H<sub>2</sub>O could be studied by fitting potential energy matrix elements.<sup>143,147,392,412–414,525</sup> Subsequently, some of the authors fit the dipole moments corresponding to the diabatic 1,2<sup>1</sup>A surface of NH<sub>3</sub>.<sup>16</sup> SOCs of formaldehyde were learned with NNs in the diabatic picture.<sup>92</sup> A total of 341 data points were used for training of SOCs. A singlet and a triplet state in the adiabatic basis were transformed to diabatic states using Boys localization.<sup>382</sup> Since this diabaticization is based on transition dipole moments, the respective properties of the excited states had to be phase corrected. The authors proved the accuracy of their fitted PESs and emphasized the usability of the ML models to describe full-dimensional quantum dynamics.<sup>16,92,525</sup> Very recently, they investigated the OH + H<sub>2</sub> reaction, i.e., the nonadiabatic quenching of the hydroxyl radical colliding with molecular hydrogen. Four diabatic potentials including forces and couplings were fitted using a least-squares fitting procedure. A total of 1345 data points of 1,2,3<sup>2</sup>A adiabatic PESs were computed with MR-CISD.<sup>525</sup>

The aforementioned ML models are single-state models. Each energetic state and each coupling or dipole moment value resulting from different pairs of states are fitted with a separate ML model. While this yields justifiable accuracy for energies and diabatic coupling values,<sup>95</sup> dipole moments are vectorial properties and need to preserve rotational covariance.<sup>97,241</sup>

As the aforementioned studies show, ML models are generally powerful to advance quantum dynamics simulations for the excited states and can also assist the construction of effective Hamiltonians. However, currently, diabatic PESs cannot simply be fit for systems with arbitrary size and

arbitrary complexity. The diabaticization remains a methodological bottleneck, where additional developments are needed.

The investigation of medium-sized to larger molecular systems, especially the investigation of their temporal evolution, is more often carried out in the adiabatic basis using on-the-fly simulations. An increasing number of recent studies focus on fitting such adiabatic PESs. The inconsistencies in adiabatic properties make such quantities generally more challenging to fit, which is why this field of research gained a lot of attention relatively late, i.e., only in the last 3 years—after early examples dating back already to 1999.<sup>661</sup>

### 6.3.2. ML in the Adiabatic Basis. Surface Hopping MD.

The first NN models for MQCD calculations probably date back to the year 2008.<sup>98</sup> Nonadiabatic MD simulations were carried out with NN-interpolated PESs to investigate O<sub>2</sub> scattered from Al(111). Symmetry functions were used as descriptors.<sup>662</sup> A spin-unpolarized singlet and a spin-polarized triplet state at DFT level of theory were fitted with 3768 data points.<sup>662,663</sup> This two-state spin-diabatic problem allowed for evaluation of coupling values and singlet–triplet transitions with the fewest switches surface hopping approach.<sup>436,437</sup> In a later study, another adiabatic spin-polarized PES was included, and coupling values were computed between singlets and triplets<sup>664</sup> and evaluated from constructed Hamiltonian matrices.<sup>93</sup> MD simulations were executed using a manifold of ML-fitted PESs according to different spin-configurations.<sup>93,98</sup> The studies showed that singlet–triplet transitions are highly probable during the scattering event of O<sub>2</sub> on Al(111). As it has been shown later by embedded correlated wave function computations,<sup>665</sup> the activation barrier of O<sub>2</sub> on Al(111) is rather due to charge transfer than spin-flip as described above. The description of the activation barrier has been improved later<sup>666</sup> using six-dimensional PESs parametrized using the London–Eyring–Polanyi–Sato function.<sup>667</sup>

After the two studies by Carbogno et al., the interest in advancing ML-based MQCD simulations for the excited states in the adiabatic basis increased mainly in the last three years. One of the first works during this time was conducted by Hu et al.,<sup>139</sup> who investigated the nonadiabatic dynamics of 6-aminopyrimidine with KRR and the Coulomb matrix. Because of the many degrees of freedom of the molecule and including three singlet states, a large amount of training data was required (>65k data points). Coupling values were not fitted, but, instead, the Zhu–Nakamura approach was used to compute hopping probabilities.

Later, Dral et al.<sup>140</sup> applied KRR models to accurately fit a two-state spin-Boson Hamiltonian and reproduce reference dynamics using 1000 and 10 000 data points. NAC vectors were fit in a single-state fashion. During dynamics simulations, conformations close to critical regions were computed with the reference method instead of the ML model in order to allow for accurate transitions.

In another study, Chen et al.<sup>141</sup> used two separate deep NNs to fit the energies and forces of two adiabatic singlet states of CH<sub>2</sub>NH. About 90k data points were used to generate these single-state models. Using the Zhu–Nakamura approach to account for hopping probabilities, the reference dynamics could be reproduced, and quantum chemical calculations were replaced completely during the dynamics.

Cui and co-workers<sup>668</sup> further developed a multilayer energy-based fragmentation method to study the excited-

state dynamics and photochemistry of larger systems. This scheme composes a molecular system into a photochemically active (inner) region and a photochemically inert (outer) region. In the original scheme, the active region and the interactions with the outer region are described with the multireference method CASSCF, whereas the outer region is treated with DFT. This decomposition of the total energy of a system allows one to treat larger systems, which cannot be described fully with CASSCF. Compared to quantum mechanics/molecular mechanics (QM/MM) schemes, the energy-based fragmentation method can be rigorously derived from a many-body energy expansion. Different truncation levels and methods can be used for each intralayer and interlayer interactions, so that its accuracy and efficiency can be controlled according to the need. The authors simulated two-state photodynamics of CH<sub>3</sub>N=NCH<sub>3</sub> (inner region) including five water molecules (outer region) without the use of ML. The Zhu–Nakamura approximation to model hopping probabilities in nonadiabatic MD simulations was applied.<sup>668</sup> In order to make the simulations more efficient, the authors replaced the DFT calculations with deep multilayer feed-forward NNs using a distance-based descriptor,<sup>125,506</sup> hence, they describe the ground state energies and forces of the photochemically inert region with ML and describe the S<sub>1</sub> and S<sub>0</sub> state of the inner region with CASSCF. The hybrid ML multilayer energy-based fragmentation method can reproduce the photodynamics of the system.<sup>536</sup> Subsequently, the deep NNs were replaced with embedded-atom NNs,<sup>508</sup> and accurate second derivatives could be computed efficiently.<sup>537</sup>

Recently, we sought to fit NACs and transition and permanent dipole moments in addition to energies and forces of three singlet states of the methylenimmonium cation, CH<sub>2</sub>NH<sub>2</sub><sup>+</sup>, using deep NNs and the matrix of inverse distances as a molecular descriptor.<sup>94</sup> We were able to perform ML-enhanced excited-state MD simulations with hopping probabilities based on ML-fitted NACs. NNs replaced the reference method MR-CISD completely during the dynamics, which is one of the key factors behind a successful MLMD study as it allows one to completely decouple the costs of the expensive reference method from the dynamics simulations. The accuracy of the ML-approximated PESs and couplings was further assessed by comparing the populations in the different excited states of the reference dynamics, of the ML models, and of the same quantum chemical reference method, namely, MR-CISD, but with a slightly different basis set. This comparison helped to estimate the meaning of “good agreement” for population dynamics. Root mean squared deviations of the nuclear geometries tracked over a short time scale also helped in assessing the success of the MLMD model in reproducing dynamics before using it for, e.g., simulation of longer time scales and better statistics. Such long time scale photodynamics simulations for 1 ns were achieved using the mean of two NN models in approximately two months, whereas the reference method would have taken an estimated 19 years to compute the dynamics for 1 ns on the same computer. This study demonstrated the possibility of MLMD simulations to go beyond time scales of conventional methods. As another benefit of the ML models, it was shown that a large ensemble of trajectories could be calculated, still at a lower cost than a few trajectories with the reference method.<sup>94</sup>

Recently, Li et al.<sup>669</sup> built on these recently developed techniques and developed the MD program PyRAI<sup>2</sup>MD (Python Rapid Artificial Intelligence *Ab Initio* MD). NNs

were trained on the  $S_1$  and  $S_0$  states of  $CF_3-CH=CH-CF_3$  (hexafluoro-2-butene) at the CASSCF(2,2)/cc-pVDZ level of theory in order to enable 10 ns photodynamics simulations. Generation of an initial training set via Wigner sampling, optimizations of critical regions, and short time-scale trajectories in addition to adaptive sampling resulted in a training set of 6232 data points. The descriptor was generated from inverse distances and (dihedral) angles, and a phase-less loss function<sup>15</sup> was applied to render NACs learnable.

The performance of KRR in comparison to NNs was assessed by us together with von Lilienfeld and co-workers.<sup>95</sup> The operator formalism<sup>670</sup> and the FCHL representation<sup>127,530</sup> were used to fit the three singlet states of  $CH_2NH_2^+$  using the previously generated training set of 4000 data points. A single-state treatment and a multistate treatment for predicting energies were compared. To this aim, a multistate KRR approach as developed with an additional kernel that encodes the quantum energy levels. The accuracy of KRR models could be improved using this extended approach.<sup>95</sup> The KRR models were further compared to deep NN models regarding their ability to predict dipole moments and NACs. While NNs yielded a slightly higher accuracy at the largest available training set size, KRR models exhibited a steeper learning curve, and hence more efficient learning. The multireference quantum chemical potential energy curves could be faithfully reproduced with KRR models and NN models for the three singlet energies of  $CH_2NH_2^+$  at first sight. Interestingly, the small differences lead to correctly predicted dynamics with the NNs, while the KRR model was unable to reproduce the reference dynamics. Hence, small differences between the reference method and ML models, especially in critical regions of the PESs, can lead to completely wrong photodynamics simulations.<sup>95</sup> The different performance of NNs and KRR models was proposed to be a result of the parametric dependence of the depth of NNs and the nonparametric dependence of the depth of KRR models.

Further, it was shown that inconsistencies along QC PESs, which are common especially close to conical intersections, were not reproduced by ML models.<sup>94,171</sup> The problematic fitting of nonsmooth functions representing NACs, i.e., their singularities at conical intersections, was circumvented by employing so-called smooth couplings. To arrive at the latter, the NACs were multiplied with the respective energy gaps. Hence, accurately trained energies were also required in this approach. For prediction, the smooth fitted couplings were subsequently divided by the inferred energy gaps of the ML model. In this way, the training process became more robust.<sup>95</sup> For some quantum chemical reference methods and some chemical problems, it might be beneficial to remove data points from conformations very close to a conical intersection. Although these regions need to be represented well, such a procedure can reduce the amount of data points with cusps in energy potentials and thus problematic data points in the training set.<sup>15,148</sup>

In order to omit the extensive hyperparameter search of the descriptor and regressor, we further developed the SchNarc approach for photodynamics,<sup>15</sup> which is based on SchNet.<sup>461,465</sup> SchNarc allows for (1) a description of SOCs, (2) an NAC approximation based on ML-fitted PESs, their first and second derivatives with respect to Cartesian coordinates, and (3) a phase-free training algorithm to enable a training of raw quantum chemical data. The SchNarc approach is based on the message passing NN SchNet,<sup>461,465</sup>

which was adapted by us for the treatment of a manifold of excited electronic states. Additionally, this model can describe dipole moments using the charge model of ref 111, also adapted for excited-states. All excited-state properties can be described in one ML model in a multistate fashion. The performance of SchNarc was evaluated with surface hopping dynamics: Three singlet and three triplet states of  $SO_2$  were computed with ML models for 700 fs, and the underlying PESs were based on an “one-shot” LVC(MR-CISD) model.<sup>240</sup>  $CSH_2$  was investigated using two singlets and two triplet states for 3 ps at the CASSCF level of theory representing slow population transfer, and the performance of SchNarc to reproduce ultrafast transitions during dynamics was assessed using  $CH_2NH_2^+$  with the aforementioned training set. The hopping probabilities were computed according to ML-fitted SOCs and NACs—the latter being fitted in a rotationally covariant way as derivatives of virtual ML properties and approximated from ML PESs. In all cases, excellent agreement with the reference method could be achieved. Noticeably, all the aforementioned photodynamics studies with ML models<sup>15,94,95,139–141</sup> make use of Tully’s fewest switches surface hopping approach with hopping probabilities based on coupling values or approximated schemes.<sup>436,437</sup>

It is further worth mentioning that ML models, which provide energies and derivatives, can be used to optimize reaction coordinates to find, e.g., local minima or minimum energy conical intersections.<sup>94,171,241</sup> In general, a successful MLMD study should make it possible to investigate reactions at longer time or larger length scales than complementary studies with quantum chemical reference methods or should enable large scale screenings, which would not be feasible with the reference method. With respect to different ML models, kernel methods are usually faster during training, but take longer for predictions. In contrast, deep NNs require longer training times and more complex hyperparameter optimization, but once trained, they can predict properties and derivatives thereof extremely fast.<sup>175,671</sup> In the following paragraphs, we compare the timings of some electronic structure methods with predictions made by the deep NN model SchNarc.

*Exemplary Timings of Single Point Calculations.* The speed-up of simulations is one of the main arguments employed for promoting ML in quantum chemistry. In order to get an idea about the computational time used in different calculations, we provide an example here. In order to get a better feeling of the speed-up one can achieve for excited-state energies and derivatives of medium-sized molecules, the topic will be discussed here. In the next paragraph, timings for dynamics simulations of smaller molecules, which have been achieved up to date, will be provided. Different electronic structure methods and the SchNarc model are used to compute energies and derivatives of 5 singlet and 8 triplet states of tyrosine (24 atoms) on a 2x Intel Xeon E5-2650 v3 CPU and GeForce GTX 1080 Ti GPU. The training of the SchNarc model on the same GPU took about 7 days. One advantage of ML in this case is that it provides access to properties, which are not available with any electronic structure theory, such as NACs, approximated from first and second derivatives of ML PESs.<sup>15</sup> As it becomes visible from the examples shown in Table 2, the most time-consuming parts are the derivatives with second derivatives being far more expensive than first derivatives. These simulations can be accelerated a lot by ML models. Further, a NAC computation is comparable to a gradient computation. However, there are

**Table 2. Comparison of the Timings to Compute Different Excited-State Properties of the Molecule Tyrosine with Different Electronic Structure Methods and SchNarc<sup>15a</sup>**

method	processor	time [s]
Energies (13 States)		
ADC(2)	CPU	2,160
TDDFT	CPU	724
CASSCF	CPU	5,719
CASPT2	CPU	7,972
SchNarc	CPU	1.5
SchNarc	GPU	0.03
Energies + Gradients (13 states)		
ADC(2)	CPU	9,280
TDDFT	CPU	5,938
CASPT2	CPU	129,389
SchNarc	CPU	2
SchNarc	GPU	0.1
Hessian (1 State, Frequency)		
ADC(2)	CPU	11,760
SchNarc	CPU	97
SchNarc	GPU	15
Approximated NACs (All States) (Not Implemented for QC Methods)		
SchNarc	CPU	1,260
SchNarc	GPU	186

<sup>a</sup>2x Intel Xeon E5-2650 v3 CPUs and GeForce GTX 1080 Ti GPUs are used for computations. In order to keep the notation short ADC(2) refers to ADC(2)/def2-SVP, CASSCF to CASSCF(10,9)/ano-rcc-vDZP, CASPT2 to MS-CASPT2(10,9)/ano-rcc-vDZP, and TDDFT to TDDFT/B3LYP/def2-SVP. The programs Turbomole<sup>672</sup> (for ADC(2)), openMolcas<sup>310</sup> (for CASSCF and CASPT2), and ORCA<sup>673</sup> (for TDDFT) were used.

not only  $N_c$  NAC calculations to carry out, but  $N_c \times (N_c - 1)$  calculations.

**Exemplary Timings for MLMD, LVC Dynamics, and MQCD.** The timings of surface hopping MD with analytical PESs (from LVC), quantum chemical PESs, and ML-fitted PESs based on fitted and approximated NACs from Hessians can be found for three exemplary molecules in Table 3.

Obviously, crude excited-state force fields like the LVC model are faster than ML models, e.g., for SO<sub>2</sub>. We note that even such force field implementations can probably still be streamlined for speed but will always be more expensive than

**Table 3. Comparison of the Timings to Compute 100 fs with the Surface Hopping Including Arbitrary Couplings (SHARC)<sup>236,256,589</sup> method<sup>a</sup>**

	100 fs dynamics [s/CPU]		
	MLMD1	MLMD2	reference
SO <sub>2</sub>	10	12	2–3
CH <sub>2</sub> NH <sub>2</sub> <sup>+</sup>	24	250	74,224
CSH <sub>2</sub>	14	16	104

<sup>a</sup>For SO<sub>2</sub> and CH<sub>2</sub>NH<sub>2</sub><sup>+</sup>, three singlet states are described and for CSH<sub>2</sub> two singlet and two triplet states. The molecule SO<sub>2</sub> is approximated using a highly efficient LVC model,<sup>240</sup> while the underlying reference method to describe the excited states of CH<sub>2</sub>NH<sub>2</sub><sup>+</sup> is MR-CISD/aug-cc-pVDZ and of CSH<sub>2</sub> is CASSCF-(6,5)/def2-SVP. SchNarc is used for the MLMD simulations. Once energies, forces, and NACs are trained and predicted (MLMD1), and once NACs are approximated from first- and second-order derivatives of ML PESs (MLMD2). 2x Intel Xeon E5-2650 v3 CPUs are used.<sup>15</sup>

ground-state MD simulations, where it would take approximately 0.005 s to simulate 100 fs for the gas-phase methylenimmonium cation, CH<sub>2</sub>NH<sub>2</sub><sup>+</sup>, using a state-of-the-art program like Amber.<sup>213</sup>

However, dynamics based on highly accurate quantum chemical calculations can be accelerated significantly with ML-fitted PESs, e.g., SchNarc models for CH<sub>2</sub>NH<sub>2</sub><sup>+</sup> based on MR-CISD/aug-cc-pVDZ.<sup>15</sup> The speedup is higher if NACs are learned directly (MLMD1) compared to when they are approximated from Hessians (MLMD2). A lot of Hessian evaluations are required in this example because ultrafast transitions occur in CH<sub>2</sub>NH<sub>2</sub><sup>+</sup>. The second-order derivatives reduce the efficiency by a factor of about 10. Nevertheless, Hessian calculations of ML-PESs can be accelerated by a factor of about 5–10 using a GPU (dependent on the molecule and GPU used).

Table 3 further shows that a cheaper underlying reference method, such as CASSCF(6,5)/def2-SVP used for CSH<sub>2</sub>, does not allow for such a significant speedup. In this example, however, the difference between simulations with learned NACs and approximated NACs is small because the dynamics of CSH<sub>2</sub> is characterized by slow population transfer. Hence, less Hessian evaluations are required to estimate the hopping probabilities.

The time required to train a SchNarc model on a GeForce GTX 1080 Ti GPU is approximately 11 h for energies and forces of 3 singlet states with 3000 data points of CH<sub>2</sub>NH<sub>2</sub><sup>+</sup>, about 13 h for energies, forces, and SOC of 2 singlet and 2 triplet states using 4000 data points of CSH<sub>2</sub> and about 4 h for energies and forces of 3 singlet states of SO<sub>2</sub> using 5000 data points.

**Dipole Moments and Atomic Charges.** In addition to the investigation of the temporal evolution of some systems in the excited states, permanent and transition dipole moments have been computed with ML models. As mentioned before, in our earlier approaches, we fitted permanent and transition dipole moments as single values with NNs and KRR—strictly speaking, we were neglecting the rotational covariance of the vectors (since rotations were negligible in these simulations).<sup>94,95</sup> The NN and KRR models for dipole moments have been evaluated and compared to quantum chemical reference dipole moments using learning curves and MAEs. Their potential to compute UV spectra was emphasized.

The use of dipole moments to actually simulate UV spectra was demonstrated by Jiang, Mukamel, and co-workers using *N*-methylacetamide, a model system to investigate peptide bonds.<sup>97,510</sup> They evaluated the ability of ML to describe transition dipole moments at TDDFT level of theory. In a first attempt,<sup>510</sup> the authors predicted dipole vectors as independent values. Fourteen internal coordinates in combination with multilayer feed-forward NNs were used to predict transition energies of *N*-methylacetamide. XYZ representations served as an input for fitting ground state dipole moments. The Coulomb matrix was employed to fit transition dipole moments for the  $n\pi^*$  and  $\pi\pi^*$  transitions, but did not lead to sufficiently accurate results. Higher accuracy was obtained by replacing the atomic charges in the Coulomb matrix (eq 21) with charges from natural population analysis. The choice of descriptors was justified by screening different types of descriptors for prediction of different properties. In a later work, some of the authors used embedded-atom NNs to predict transition dipole moments from atomic contributions in a rotationally covariant way. The dipole moment vector

between two states  $i$  and  $j$  was obtained as a linear combination of three contributions:

$$\mu_{ij} = \mu_T^i + \mu_T^j + \mu_T^3 \quad (34)$$

$\mu_T^i$  and  $\mu_T^j$  were modeled using the charge model of ref 111. A third contribution,  $\mu_T^3$ , was obtained as the cross product of  $\mu_T^i$  and  $\mu_T^j$ :

$$\mu_T^3 = \sum_a^{N_a} q_a^3 (\mu_T^i \times \mu_T^j) \quad (35)$$

$\mu_T^i$ ,  $\mu_T^j$ , and  $q_a^3$  were outputs of the same embedded-atom NN.

Recently, we extended the SchNarc model to describe permanent and transition dipole moments of an arbitrary number of electronic states and pairs of states as vectorial properties in a single ML model in addition to the excited-state energies, forces, and couplings.<sup>241</sup> Also this model is based on the charge model of ref 111 and allows one to predict latent partial charges for excited states. This charge model relying on atom-wise descriptors thus preserves the correct direction of the permanent and transition dipole moments.

#### 6.4. ML of Tertiary Outputs

The secondary outputs, such as dipole moments or excited state energies, can be used to calculate oscillator strengths (eq 1) and energy gaps (Figure 1d). These properties can serve for the modeling of UV absorption spectra. UV spectra were computed in the previously described studies of *N*-methylacetamid with the ML fitted transition dipole moments. Jiang, Mukamel, and co-workers<sup>510</sup> applied the transition dipole moment and additionally fitted  $n\pi^*$  and  $\pi\pi^*$  excitation energies to compute UV spectra of this molecule with NNs. Subsequently, some of the authors<sup>97</sup> used these excitation energies and the transition dipole moments to model a Frenkel exciton Hamiltonian for proteins using amino acid residues and peptide bonds. This effective Hamiltonian could further be used to approximate UV spectra of proteins. The interaction between amino acid residues and peptides was neglected, so only the isolated peptide excitation energies, i.e., those of *N*-methylacetamid, and the respective transition dipole moments were needed to construct the Hamiltonian. The authors made use of the dipole–dipole approximation<sup>674</sup> and applied embedded-atom NNs. High transferability and predictive power were obtained.

In addition, the transition dipole moments of SchNarc<sup>241</sup> have been used to predict UV/visible spectra of the methylenimmonium cation and ethylene. Both molecules were trained simultaneously. Although they differ in their photochemistry, i.e., the first excited singlet state is bright in the case of ethylene, whereas it is dark in the case of the methylenimmonium cation, with an opposite behavior observed for the second excited singlet state, ML models trained on both molecules were slightly superior to ML models trained on single molecules. In contrast to the previous models, where each property and electronic state or transition energy is fitted separately in one model, the SchNarc model can treat excited-state energies, forces, permanent dipole moments, and transition dipole moments in one model simultaneously, while being able to fit any predefined number of electronic states of different spin multiplicities. According to eq 12, dipole moments are obtained as the sum of atomic contributions, which are obtained from latent partial charges inferred by the ML model multiplied with the vector of an atom with respect

to the center of mass. In this way, direct access to the atomic charges is provided, which should, in principle, also be possible with the embedded-atom NN used in ref 97, but has not been evaluated. As the charge distribution in a molecule is highly dependent on the underlying partitioning scheme, a comparison of the different schemes is not straightforward. The Hirshfeld charges<sup>675</sup> are often considered more accurate than, e.g., Mulliken charges.<sup>676</sup> Hirshfeld charges were compared to the latent partial charges of the ML model and found to agree well.<sup>172,241</sup> The ML charges further were used to compute electrostatic potentials and charge redistribution after light excitation.<sup>241</sup>

Because of the excellent agreement of the ML predictions and the reference data, the transferability of SchNarc for the excited state energies and dipole moments was evaluated. Therefore, electrostatic potentials and UV/visible spectra of aminomethylene ( $\text{CHNH}_2$ ) and methylenimine ( $\text{CH}_2\text{NH}$ ) were computed with SchNarc, and qualitatively correct results could be obtained. This result left us to conclude that at least similar structured molecules can be predicted with ML models even though they are not included in the training set. The high costs and complexity of the underlying multireference quantum chemical method hampered the exploration of more molecules and the transferability of SchNarc toward a larger chemical space.<sup>241</sup>

Ramakrishnan et al.<sup>358</sup> predicted excitation energies of the lowest-lying two excited singlet states,  $S_1$  and  $S_2$ , as well as corresponding oscillator strengths obtained from TDDFT calculations with KRR. The QM8<sup>572</sup> database was used consisting of 20k organic molecules. With the  $\Delta$ -learning approach, CC2 accuracy could be obtained. Very recently, Xue et al.<sup>490</sup> assessed the performance of KRR models with the normalized inverse distances as a molecular descriptor to predict absorption spectra of benzene and a derivative of acridine containing 38 atoms. Therefore, the authors learned the excited-state energy gaps of several states and the corresponding oscillator strengths in a single-state fashion. Applying a nuclear ensemble approach, the absorption cross sections could be computed at TDDFT accuracy using a fraction of ensemble points.

Pronobis et al.<sup>158</sup> compared two-body, three-body, and automatically designed descriptors to learn TDDFT HOMO–LUMO gaps as well as first and second vertical excitation energies. More than 20k molecules of the QM9 database<sup>457,572</sup> were selected for this purpose, and learning curves were used to evaluate the learning behavior of different ML models. While atom-wise descriptors worked well for HOMO–LUMO gaps, the authors concluded that the accuracy of predicted transition energies is not sufficiently accurate and suggested that advanced nonlocal descriptors might be necessary to achieve higher accuracy. They further proposed the idea of encoding information about the electronic state in the ML model.<sup>158</sup> Indeed, our recent study, in which we compared the performance of KRR and NN models with atom-wise and molecule-wise descriptors demonstrated that encoding of the energy level is advantageous.<sup>95</sup>

Recently, Kang et al.<sup>533</sup> used 500 000 molecules of the PubChemQC<sup>578</sup> database to train a random forest model on the excitation energy and the oscillator strength corresponding to the electronic state with the highest oscillator strength. Ten singlet states, as available in the PubChemQC database, were evaluated for that purpose. The authors used simplified molecular-input line-entry system (SMILES) strings and



converted them into descriptors. The descriptors comprised several topological<sup>677</sup> and binary<sup>678</sup> fingerprints, which were calculated with the help of the RDKit library.<sup>679</sup> The authors compared the prediction accuracy to the aforementioned models and stated that their model outperformed previous ML models in the task of predicting accurate oscillator strengths and excitation energies for the most probable transition in organic molecules. Analysis of important features led the authors identify that nitrogen-containing heterocycles are important for high oscillator strengths in molecules. The authors concluded that their study could serve the design of new fluorophores with high oscillator strengths.<sup>533</sup>

Ghosh et al.<sup>152</sup> used multilayer feed-forward NNs, convolutional NNs and DTNNs to fit the 16 highest occupied orbital energies from DFT, where the respective eigenvalues are broadened by Gaussians with a full width at half-maximum of 0.5 eV. The resulting spectra are probably comparable to ionization potentials in line with Koopmans' theorem. Geometries from the QM7b<sup>570,571</sup> and QM9<sup>457,572</sup> database were used for training, and predictions were tested using 10k additional diastereomers, which were also used by Ramakrishnan et al.<sup>358</sup> to evaluate the  $\Delta$ -learning approach. The convolutional NNs with the Coulomb matrix and DTNNs with an automatically generated representation outperformed the simpler NNs. Overall, good agreement to reference DFT calculations could be achieved.<sup>152</sup>

Markland and co-workers<sup>539</sup> trained NNs with atom-centered Chebyshev polynomial descriptors<sup>110</sup> on the TDDFT/CAM-B3LYP/6-31+G\*  $S_0$ – $S_1$  energy gap of the deprotonated trans-thiophenyl-*p*-coumarate (chromophore of yellow protein) in water and Nile red chromophore in water and benzene. Farthest point sampling<sup>123</sup> was used to select about 2000 data points from a larger set of 36 000 data points and was compared to random sampling. The authors assessed the performance of three different ML approaches to compute absorption spectra, spectral densities, and two-dimensional electronic spectra. One model (hidden solvation) completely ignored any environmental effects and only described the chromophore, another model (indirect solvation) incorporated environmental effects within a 5 Å cutoff of the atomistic descriptor for the chromophore, and a third model (direct solvation) treated the whole system, i.e., the chromophore and the atoms of the solvent, explicitly. As expected, the hidden solvation model turned out to be insufficiently accurate for systems with strong solvent–chromophore interactions, but was comparable to the hidden solvation model when describing Nile red chromophore in benzene. The indirect solvation and direct solvation models were comparable to each other, but with respect to the computational efficiency, the indirect solvation model was beneficial. This model could reproduce reference linear absorption spectra, spectral densities, and could capture spectral diffusion of two-dimensional electronic spectra of all treated chromophores.<sup>539</sup>

Penfold and co-workers<sup>155</sup> applied deep multilayer feed-forward NNs to prove the ability of ML to predict X-ray absorption spectra (XAS), which provide a wealth of information on the geometry and electronic structure of chemical systems, especially in the near-edge structure region. Note that X-ray free-electron laser spectroscopy can further be used to generate ultrashort X-ray pulses to investigate photodynamics simulations in real-time. The training set for the prediction of Fe K-edge X-ray near-edge structure spectra contained 9040 data points. The inputs for NNs were

generated using local radial distributions around the Fe absorption site of arbitrary systems taken from the Materials Project Database.<sup>680</sup> Qualitatively accurate peak positions and intensities could be obtained computationally efficient, and the structural refinement of nitrosylmyoglobin and  $[\text{Fe}(\text{bpy})_3]^{2+}$  was assessed with NNs. The authors noted that future development is needed to accurately capture structures far from equilibrium as well as irregularities in the bulk. The spectral shapes and other properties of X-ray laser pulsed from free-electron laser facilities could be predicted by Sánchez-González et al.<sup>681</sup> with NNs and support vector regression.

Another study was executed by Aarva et al.,<sup>682</sup> who focused on XAS and X-ray photoelectron spectra of functionalized amorphous carbonaceous materials. By clustering of DFT data with unsupervised ML techniques average fingerprint spectra of distinct functionalized surfaces could be obtained. The authors use GPR. Similarly to the aforementioned state encoding,<sup>95</sup> the authors encoded the electronic structure, i.e., the  $\Delta$ -Kohn–Sham values (core–electron binding energies), in a Gaussian kernel. This kernel was then linearly combined with a structure-based kernel based on the SOAP<sup>30</sup> descriptor. The spectra computed from the different clusters were used to fit experimental spectra allowing for an approximation to the composition of experimental samples on a semiquantitative level. The so-called fingerprint spectra, which enabled the differentiation of the spectral signatures, were assessed in a previous study using different models for amorphous carbon,<sup>683</sup> among them an ML fitted PES using GPR.<sup>112,684</sup>

Kulik and co-workers<sup>17</sup> used deep NNs to predict the spin-state ordering in transition metal complexes to determine the spin of the lowest lying energetic state in open-shell systems. The determination of spin states is important to evaluate catalytic and material properties of metal complexes. Descriptors based on a selection of empirical features were used to capture the bonding in inorganic molecular systems. The performance of descriptors including different features was assessed for a set of octahedral complexes with first-row transition metals. The most important features were identified to be the atom, which connects the ligand to the metal, its environment and its electronegativity, the metal identity and its oxidation state, as well as the formal charge and denticity of the ligand.<sup>685</sup> The ML models were tested on spin-crossover complexes and could assign the correct spin in most cases. Additionally, ML models were applied for the discovery of inorganic complexes<sup>686–689</sup> Similarly, Behler and co-workers<sup>91</sup> applied high-dimensional NNs to predict spin states of transition metal oxides, which are for instance important for lithium ion batteries. In addition, the atomic oxidation state could be predicted.

The inverse design of molecules with specific properties was further targeted by Schütt et al.,<sup>77</sup> who developed SchNOrb, a deep NN model based on SchNet. The automatically generated descriptor was extended with a description of atom pairs in their chemical and structural environment. An analytic representation of the electronic structure of a molecular system was obtained in a local atomic orbital representation. The analytic derivatives of the electronic structure allowed for optimization of electronic properties. This was demonstrated by minimizing and maximizing the HOMO–LUMO gap of malonaldehyde.<sup>563</sup> Besides, the ML method was used to predict the lowest 20 molecular orbitals of ethanol at the DFT level of theory, to investigate proton

transfer in malonaldehyde using ground-state dynamics and to analyze bond order and partial charges of uracil.

Bayesian NN models were applied by Häse et al.<sup>160</sup> to relate molecular geometries to the outcome of nonadiabatic MD simulations obtained with CASSCF. Normal modes with and without velocities of initial conditions served as an input for NN models. Velocities in addition to normal modes as descriptors improved the accuracy of ML models slightly, pointing out that normal modes contain already enough information for the sake of their study. The dissociation times of 1,2-dioxetane obtained from nonadiabatic MD simulations was the targeted output. The NNs could faithfully reproduce dissociation times and further provided a measure of uncertainty. The authors noted that their method could be particularly interesting for analysis of MLMD simulations.

Lately, regression and ML models have emerged to assist the prediction of the quantum yield, which is a property targeted in studies for the design of photoactive materials, such as organic light emitting diodes (OLEDs), molecules useful in phototherapy, solar cells, or biomedical labeling. The yield of fluorescence and phosphorescence can be targeted. It can be determined from the ratio of the rate of nonradiative emission and the rate of radiative emission. As radiative emission usually takes place on time scales in the range of nano- to milliseconds, theoretical methods to determine their rates, i.e., dynamics simulations, are limited. Approximations can be made using static calculations, such as it has been done by Kohn et al.,<sup>690</sup> who developed a semiempirical method to compute the fluorescence quantum yield of chromophores of molecules using TDDFT. The coefficients, which could not be determined from theory, were fit to experimental data. Qiu et al.<sup>691</sup> studied aggregation induced emission of triphenylamine compounds using a support vector machine (SVM). The charge of three carbon atoms adjacent to a central nitrogen atom served as an input to the classifier. Inactive and active materials could be identified and in combination with DFT leading to aggregation induced emission were investigated. Different types of ML models, such as KRR and NNs, were further compared to predict emission and absorption wavelengths and luminescence quantum yields of organic dyes that show fluorescence after excitation. Different fingerprints were compared generated from an external software and data was obtained from experiments.<sup>692</sup> The tested model emphasized the possibility to combine experimental data with ML algorithms to enable large scale screening and the design of novel materials and compounds.

### 6.5. ML-Assisted Analysis

The aforementioned studies have shown that ML enables the simulation of MD simulations and spectra predictions at low computational costs. The computational efficiency allows for enhanced statistics, i.e., in the case of MD simulations a huge number of trajectories and the simulations on long time scales.<sup>15,94</sup> Therefore, subsequent analyses of production runs can become a time limiting step of studies. This problem was identified in the aforementioned study on the dissociation times of 1,2-dioxetane by Häse et al.<sup>160</sup> Therefore, the authors further used their method to interpret the outcomes of nonadiabatic MD simulations. 1,2-Dioxetane is the target of their study as it is the smallest molecule known to show chemiluminescence after nonadiabatic transitions from the ground state to an excited state. The chemiluminescent properties of this compound were related to its decomposition

rate into two formaldehyde molecules, which was also identified to be relevant in an earlier work of some of the authors.<sup>693</sup> By analysis of the ML models that fit the dissociation times, correlations could be observed between the normal modes and the dissociation times. For example, the modes corresponding to the symmetric C–O bond stretchings, and simultaneous planarization of the two formaldehyde moieties were found to be relevant for the accurate prediction of dissociation times. It was further emphasized by the authors that although the findings of NNs were expected and obey physical laws, ML models were helpful to extract relevant information of large amount of data and could potentially serve as an inspiration to humans.

Recently, some of the authors used classification algorithms to further analyze the different types of geometries identified in the *ab initio* MD simulations of the decomposition reaction of dioxetane,<sup>694</sup> which can lead to successful dissociation or frustrated dissociation. Both in this study and in previous work,<sup>695</sup> it was found that the planarization of the two formaldehyde moieties is key for dissociation of dioxetane.<sup>694</sup>

Time-resolved experimental photoluminescence spectra could be analyzed with the LumiML software developed by Đorđević et al.,<sup>696</sup> who applied linear regression models to learn from computer-generated photoluminescence data. The software was employed to predict decay rate distributions<sup>697</sup> of perovskite nanocrystals from data generated with femtosecond broadband fluorescence upconversion spectroscopy.<sup>698</sup> The authors highlighted the applicability of their method to enhance studies on the optimization and design of optical devices and further noted that their approach can also be used to analyze transient absorption spectra. Aspuru-Guzik and co-workers<sup>154</sup> applied Bayesian NNs to find correlations of nanoaggregates with electronic coupling in semiconducting materials using absorption spectra. In general, the analysis of experimental spectra and the inverse design of compounds is most frequently applied in the research field of material science. Their description goes beyond the scope of this review, and the reader is referred to refs 164–168, and 170.

### 6.6. Open Questions and General Remarks on a Successful ML Method

The most important open questions in this field are in our opinion the following: what is generally necessary to go to larger length scales, and which reference method can be used in order to describe the excited-state energies and properties accurately for large systems? While multireference methods suffer from high costs and varying active spaces for different molecules, single-reference methods cannot describe reactions and the formation and breaking of bonds accurately. While it could be shown that long time scale photodynamics simulations in the range of nanoseconds are possible at high accuracy with ML, it is not clear how to go to time scales of seconds or minutes, which might be even more relevant. What is necessary to combine both long time scales and large length scales? How many data points are needed in order to describe the excited states of large molecules with ML on long time scales? How can we construct transition properties, such as couplings, from atomic contributions in a universal way that is valid not only for one molecule? As all of the above-mentioned questions are not answered yet, and it is not clear how to develop a universal ML force field for the excited states, we can only conclude by trying to summarize the key factors for a meaningful ML study, which focuses on one molecular system

in the case of dynamics simulations and on many molecules in case of static simulations:

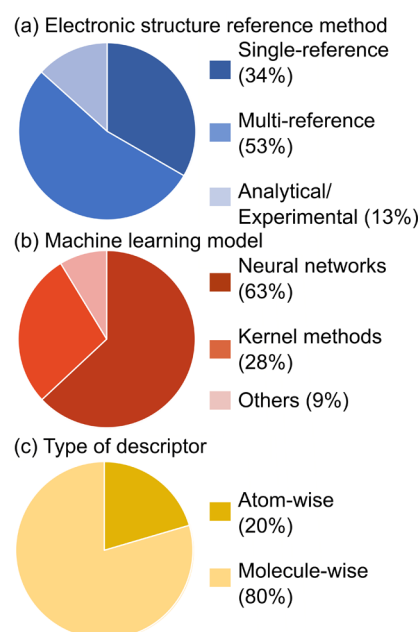
- (1) As a first step, the relevant processes that are assumed to happen after light excitation in a molecule should be evaluated, and the reference method should be decided based on these findings. In general, it might be easier—for the practitioner as well as for the ML model learning the excited states of a molecule—to use a single-reference method with a black box character, which is additionally less expensive compared to multireference methods. However, multireference methods cannot be circumvented in many cases.
- (2) As soon as the reference method has been identified, an initial training set should be computed, which samples the region around the equilibrium configuration comprehensively in the case of dynamics simulations. In case different molecules are treated, an existing database could be used, and molecular conformations could be extracted and relevant properties recomputed. Whenever transition properties are needed, it is important to either apply a phase correction in advance or adapt the loss function of the ML model.
- (3) The choice of the ML model is dependent on the type of study: If different molecules are treated, it is beneficial to use atom-wise descriptors. Whenever only one molecule is used, it cannot be said in advance what type of descriptor is better suited. The same accounts for kernel methods and neural networks—both have their merits and pitfalls. Hyperparameters should be tuned for the given problem under investigation.
- (4) As soon as the model is trained, the accuracy should be checked, usually by computing the error on a separate test set. It should be further assessed whether the model is overfitting or not. If so, a less complex model might be more suitable.
- (5) The success of an ML study can be evaluated whenever a speed-up can be obtained by applying ML instead of the reference method, e.g., longer time scales can be reached or more molecules can be scanned to design new molecules with targeted functions. A single prediction made by an ML model should be much less expensive than the reference method, but should maintain its accuracy.

## 7. CONCLUSION AND FUTURE PERSPECTIVES

In the past few years, machine learning (ML) has started to slowly enter the research field of photochemistry, especially the photochemistry of molecular systems. Although this field of research is rather young compared to ML for the electronic ground-state, some groundbreaking works have already shown the potential of ML models to significantly accelerate and improve existing simulation techniques. So far, most studies provide a proof of concept using small molecular systems or model systems. Different applications are targeted and will also be aimed at in the future, ranging from dynamics with excited-state ML potentials via absorption spectra to the interpretation of data, see Figure 1.

Analyzing the different studies reviewed here, some trends in the choice of reference methods, ML models, and descriptors can be observed. These trends are illustrated in Figure 11.

The pie chart in Figure 11a shows the used reference methods for the computation of a training set to describe the



**Figure 11.** Pie diagrams summarizing the reference methods used for the training set generation, the chosen ML models, and the type of descriptors for the description of the excited states with ML. Analysis is based on 45 studies of ML for the excited states.

excited states or excited-state properties of molecules. As can be seen, about half of the training sets are computed with multireference methods (refs 14–16, 71, 92, 94–96, 139, 141–147, 149, 160, 252, 392, 412, 413, and 660). The employed single-reference approaches are exclusively based on DFT (refs 17,77, 93, 97, 98, 152, 158, 358, 510, 533, 539, 547, and 683). Analytical methods or experimental data are also applied.<sup>140,154,161,162,696</sup>

When restricting the analysis to studies targeting dynamics, the fraction that employs multireference methods even increases. About 72% of all dynamics studies use multireference methods to compute the training data for ML models. A total of 14% of the studies use single-reference methods, and an equally large portion apply model Hamiltonians or analytical potentials. This shows that most chemical problems for the investigation of the excited states of molecules require multireference accuracy.

Recent studies of ML-based photodynamics simulations have shown that many thousands of data points are necessary to describe a few excited-state potentials of small molecular systems. To the best of our knowledge, the dynamics in the excited states with ML for molecules with more than 12 atoms in full dimensions has not yet been investigated.<sup>139,145,146</sup> Especially, the huge number of data points is concerning in this case, as larger molecules with more energetic states and a complex photochemistry could require many more data points. A meaningful training set generation, which can be achieved with active learning, adaptive sampling, and structure-based sampling techniques, is thus essential for dynamics simulations.<sup>94,111,515,559</sup> Clustering of molecular geometries obtained from dynamics simulations with a cheap method further is beneficial for selecting important reference geometries.<sup>139,140,517,560,561</sup> Still, the high costs and the complexity of multireference methods to compute an ample training set for ML also hamper the application of ML models to fit the excited states of larger polyatomic systems, whose accurate

photochemical description is often additionally complicated by a high density of electronic states.

Single reference methods, such as time-dependent DFT, are advantageous with respect to the computational costs of the training set, but suffer from qualitatively incorrect PESs in some conformational regions of molecules, such as dissociative regions. In principle, these conformational regions could be excluded from the training set, and the remaining conformational space could be interpolated using ML, but the training set would then remain incomplete and so would the dynamics. Schemes like the  $\Delta$ -learning approach<sup>358</sup> or transfer learning<sup>360</sup> could be helpful in this regard. These approaches might be useful to let ML models learn from single-reference data and adjust their accuracy according to multireference methods. The direct use of approximated methods, such as time-dependent DFT-based tight binding, is most likely not suitable for photodynamics on long time scales, because such approaches might easily be quantitatively incorrect. Of particular concern is then the accumulation of quantitatively tiny errors in the underlying potentials toward wrong dynamics trends. At the current stage of research, it is not clear whether such approximate potentials can provide qualitatively correct trends for reaction dynamics.<sup>175</sup>

In addition to the aforementioned problems, the training set generation is complicated by the arbitrariness of the signs of coupling values and properties resulting from two different electronic states.<sup>15,16,92,94,95,97</sup> This arbitrariness has to be removed in order to make data learnable with conventional methods. Such a correction scheme is termed phase correction and has been applied to correct coupling values and dipole moments.<sup>16,92,94,97,546</sup> An alternative phase correction training algorithm has been shown to be beneficial with respect to the costs of the training set generation and has enabled the learning of raw quantum chemical data.<sup>15</sup>

Figure 11b shows which ML models are applied in the discussed studies. About two-thirds rely on NNs, whereby simple multilayer feed-forward NNs are most often employed. Several research fields were advanced with NN-fitted functions: photodynamics simulations (refs 15, 93–95, 98, 140, 141, 143, 144, 147, 392, 412, and 413), spectra predictions and analysis,<sup>97,152,155,241,490,510,539</sup> excited-state properties,<sup>15–17,92,95,97,510</sup> diabaticization procedures,<sup>96,142</sup> interpretation of reaction outcomes,<sup>160,696</sup> and the prediction of HOMO–LUMO gaps or gaps between energetic states.<sup>77,152,158</sup> In these studies, between one and seven hidden layers with varying numbers of nodes were used. KRR methods were mainly applied to interpolate diabatic potentials<sup>145,146,149,252,660</sup> and in studies focusing on more than one molecular systems.<sup>358</sup> In general, only a few studies focused on extrapolation throughout chemical compound space in the excited states. Yet, only the energies, HOMO–LUMO gaps, or spectra based on fitted oscillator strengths could be predicted using a single ML model for different molecules.<sup>17,155,158,358</sup> Decision trees were used to select an active space for diatomic molecules<sup>71</sup> and semisupervised classification was applied to assess whether a molecule require multireference or single-reference treatment.<sup>176</sup>

One drawback of recently developed ML models is that they are molecule-specific and thus not universal. In part, this issue is related to the used molecular descriptors. As can be seen in Figure 11c, most studies apply descriptors that capture molecules as a whole. The few studies, which describe PESs and properties of molecular systems from atomic contribu-

tions, either treat small molecular systems<sup>15,95,97</sup> or predict properties related to the ground-state equilibrium structure of a molecular system or to electronic ground state calculations, e.g., the HOMO–LUMO gaps.<sup>77,152,158</sup> Because of the limited transferability of existing ML models to predict the excited state PESs and properties of different molecular systems, an extrapolation throughout chemical compound space is hindered in many cases. Nevertheless, in order to preserve rotational covariance especially in transition properties, such as transition dipole moments or NAC values, atom-wise descriptors have proven to be more successful.<sup>97,241</sup>

In order to fully exploit the advantages that ML models offer and to achieve the aforementioned goal of a transferable ML model for the excited states, a highly versatile descriptor is required, which can describe atoms in their chemical and structural environment and enable an ML model to treat molecules of arbitrary size and composition. It would be highly desirable if an ML model could then describe the photochemistry of large systems, which are too expensive to compute with precise multireference methods, using only small building blocks, i.e., small enough ones to describe their electronic structure accurately. For example, the excited states of proteins or DNA strands could potentially be predicted from contributions of amino acids or DNA bases, respectively, which is most often done using effective model Hamiltonians to date.<sup>55</sup> A local description of the excited-state PESs and their properties derived from the ML-fitted PESs could further provide a way toward excited-state ML/MM simulations alike QM/MM techniques.<sup>175,536,668</sup> Unfortunately, it is not yet known whether the excited-state PESs and properties can be constructed from atomic contributions or not.<sup>175</sup>

In studies comparing different ML models, it was even suggested that nonlocal descriptors might be needed or that the electronic state has to be encoded explicitly in the molecular representation to enable a transferable description of the excited states with ML.<sup>95,158</sup>

To conclude, the reviewed studies focus on almost all aspects of excited-state quantum chemistry and improve them successfully: ML models can help to choose a proper active space for multireference methods, and they predict secondary and tertiary outputs of quantum chemical calculations and help in the interpretation of theoretical studies. ML models push the boundaries of computed time scales<sup>94</sup> and are used to investigate and analyze the huge amount of data we produce every day in experiments or with high-performance computers.<sup>160,696</sup>

It should be emphasized once more that the recent studies show that the goal of ML is not to replace existing methods completely, but to provide a way to improve them. In fact, ML models for the excited states at their current stage are far from replacing existing quantum chemical methods, and they are also far from being routine. Without human intervention, ML cannot solve existing problems, and much remains to be done to describe systems beyond single, isolated molecules.

To the best of our knowledge, what is still missing is the proof that ML can provide an approximation to the multireference wave function of a molecular system. Such an achievement would be a great advancement in the research field of photochemistry, as any property we wish to know could possibly be derived from the ML wave function. An ML representation of the electronic structure would further be beneficial to allow for an inverse design of molecules with specific properties, which has been shown to be feasible for the

ground state of a molecular system.<sup>77</sup> The optimization of photochemical properties with respect to molecular geometries would be useful for many exciting research fields, e.g., photocatalysis,<sup>166</sup> photosensitive drug design,<sup>699</sup> or photo-voltaics.<sup>29,30</sup>

The multifaceted photochemistry offers a perfect playground for ML models. It may be important to highlight that, despite the negative image ML has suffered in some research communities, it cannot be denied that it opens up many new ways and possibilities to improve simulations and make studies feasible that were considered unattainable only a few years, if not only months ago.<sup>535</sup> The computational efficiency and high flexibility of deep learning models can lead this research field toward simulations of long time and large length scales. The possibilities ML models offer are far from being exhausted. The enormous chemical space, estimated to consist of more than  $10^{60}$  molecules,<sup>700</sup> and the desire to develop methods, which could develop into a universal approximator, make ML models perfectly suited to advance this research field. The possibility of deep ML models to process a huge amount of data can even assist the interpretation and analysis<sup>160,696</sup> of many photochemical studies and can help to explore unknown physical relations and be a source of potential human inspiration.

## AUTHOR INFORMATION

### Corresponding Author

**Philipp Marquetand** – Institute of Theoretical Chemistry, Faculty of Chemistry, Vienna Research Platform on Accelerating Photoreaction Discovery, and Data Science @ Uni Vienna, University of Vienna, 1090 Vienna, Austria; [orcid.org/0000-0002-8711-1533](https://orcid.org/0000-0002-8711-1533); Email: [philipp.marquetand@univie.ac.at](mailto:philipp.marquetand@univie.ac.at)

### Author

**Julia Westermayr** – Institute of Theoretical Chemistry, Faculty of Chemistry, University of Vienna, 1090 Vienna, Austria; [orcid.org/0000-0002-6531-0742](https://orcid.org/0000-0002-6531-0742)

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.chemrev.0c00749>

### Notes

The authors declare no competing financial interest.

### Biographies

Julia Westermayr works as a postdoctoral research fellow at the University of Warwick (United Kingdom) in the research group of Assoc.-Prof. Dr. Reinhard Maurer. She studied Chemistry at the University of Vienna (Austria) and received her Ph.D. degree in 2020 in theoretical chemistry under the supervision of Priv.-Doz. Dr. Philipp Marquetand and Univ.-Prof. Dr. Dr. h.c. Leticia González, for which she was awarded the uni:docs fellowship of the University of Vienna. Her research interests include the photochemistry of molecules and materials and the development of machine learning models to accelerate nonadiabatic molecular dynamics simulations.

Philipp Marquetand studied chemistry at the University of Heidelberg and the University of Würzburg, where he completed his Ph.D. under the supervision of Prof. Dr. Volker Engel in 2007. He did a postdoctoral stay at the École Normale Supérieure in Paris with Prof. James T. Hynes until 2010. Afterwards, he pursued his habilitation mentored by Univ.-Prof. Dr. Dr. h.c. Leticia González at the University of Jena and the University of Vienna. Since 2017, he has been a senior scientist and university teacher (Privatdozent) at the

University of Vienna. His research interests comprise machine learning with a focus on electronically excited states, computational photochemistry, and excited-state ab initio molecular dynamics. He is a developer of the SHARC program package for nonadiabatic dynamics. Furthermore, he is interested in interactions of matter with strong fields including multiphoton ionization and quantum dynamics.

## 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. We thank P. A. Sánchez-Murcia for help in setting up the quick Amber simulation for MD timings.

## REFERENCES

- (1) Kępuska, V.; Bohouta, G. Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). *IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, 2018; pp 99–103.
- (2) Hoy, M. B. Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Med. Ref. Serv. Q.* **2018**, *37*, 81–88.
- (3) Raucci, U.; Valentini, A.; Pieri, E.; Weir, H.; Seritan, S.; Martinez, T. J. ChemVox: Voice-Controlled Quantum Chemistry. *ChemRxiv* **2020**, DOI: [10.26434/chemrxiv.13054154.v1](https://doi.org/10.26434/chemrxiv.13054154.v1).
- (4) Silver, D.; et al. Mastering the Game of Go with Deep Neural Networks and Tree Search. *Nature* **2016**, *529*, 484–489.
- (5) Bansak, K.; Ferwerda, J.; Hainmueller, J.; Dillon, A.; Hangartner, D.; Lawrence, D.; Weinstein, J. Improving Refugee Integration through Data-Driven Algorithmic Assignment. *Science* **2018**, *359*, 325–329.
- (6) Leung, M. K. K.; Delong, A.; Alipanahi, B.; Frey, B. J. Machine Learning in Genomic Medicine: A Review of Computational Problems and Data Sets. *Proc. IEEE* **2016**, *104*, 176–197.
- (7) Shen, D.; Wu, G.; Suk, H.-I. Deep Learning in Medical Image Analysis. *Annu. Rev. Biomed. Eng.* **2017**, *19*, 221–248.
- (8) Chen, C.; Seff, A.; Kornhauser, A.; Xiao, J. DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving. *The IEEE International Conference on Computer Vision (ICCV)*, 2015.
- (9) Yang, X.; Wang, Y.; Byrne, R.; Schneider, G.; Yang, S. Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. *Chem. Rev.* **2019**, *119*, 10520–10594.
- (10) Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press, 2016.
- (11) Gómez-Bombarelli, R.; Aspuru-Guzik, A. In *Handbook of Materials Modeling: Methods: Theory and Modeling*; Andreoni, W., Yip, S., Eds.; Springer International Publishing: Cham, 2018; pp 1–24.
- (12) Agrawal, A.; Choudhary, A. Perspective: Materials Informatics and Big Data: Realization of the “Fourth Paradigm” of Science in Materials Science. *APL Mater.* **2016**, *4*, 053208.
- (13) Aspuru-Guzik, A.; Lindh, R.; Reiher, M. The Matter Simulation (R)evolution. *ACS Cent. Sci.* **2018**, *4*, 144–152.
- (14) Schwilk, M.; Tahchieva, D. N.; von Lilienfeld, O. A. Large yet Bounded: Spin Gap Ranges in Carbenes. *arXiv* **2020**, 2004.10600.
- (15) Westermayr, J.; Gastegger, M.; Marquetand, P. Combining SchNet and SHARC: The SchNarc Machine Learning Approach for Excited-State Dynamics. *J. Phys. Chem. Lett.* **2020**, *11*, 3828–3834.
- (16) Guan, Y.; Guo, H.; Yarkony, D. R. Extending the Representation of Multistate Coupled Potential Energy Surfaces to Include Properties Operators using Neural Networks: Application to the  $1,2^1A$  States of Ammonia. *J. Chem. Theory Comput.* **2020**, *16*, 302–313.
- (17) Taylor, M. G.; Yang, T.; Lin, S.; Nandy, A.; Janet, J. P.; Duan, C.; Kulik, H. J. Seeing Is Believing: Experimental Spin States from

Machine Learning Model Structure Predictions. *J. Phys. Chem. A* **2020**, *124*, 3286–3299.

(18) Kulik, H. J. Making Machine Learning a Useful Tool in the Accelerated Discovery of Transition Metal Complexes. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2020**, *10*, No. e1439.

(19) Power, P. P. Stable Two-Coordinate, Open-Shell (d1–d9) Transition Metal Complexes. *Chem. Rev.* **2012**, *112*, 3482–3507.

(20) Bousseksou, A.; Molnár, G.; Matouzenko, G. Switching of Molecular Spin States in Inorganic Complexes by Temperature, Pressure, Magnetic Field and Light: Towards Molecular Devices. *Eur. J. Inorg. Chem.* **2004**, *2004*, 4353–4369.

(21) Li, H.; Feng, H.; Sun, W.; Fan, Q.; King, R. B.; Schaefer, H. F. First-Row Transition Metals in Binuclear Cyclopentadienylmetal Derivatives of Tetramethyleneethane:  $\eta^3$ ,  $\eta^3$  versus  $\eta^4$ ,  $\eta^4$  Ligand–Metal Bonding Related to Spin State and Metal–Metal Bonds. *Organometallics* **2014**, *33*, 3489–3499.

(22) Barbatti, M.; Borin, A. C.; Ullrich, S. Photoinduced Phenomena in Nucleic Acids I; *Topics in Current Chemistry*; Springer Berlin Heidelberg, 2014; Vol. 355; pp 1–32.

(23) Turro, N. J.; Ramamurthy, V.; Scaiano, J. C. *Principles of Molecular Photochemistry: An Introduction*, 2009.

(24) Cohen, B.; Crespo-Hernández, C. E.; Hare, P. M.; Kohler, B. *Ultrafast Excited-State Dynamics in DNA and RNA Polymers*; Elsevier: Amsterdam, 2004; Chapter Ultrafast Excited-State Dynamics in DNA and RNA Polymers, pp 463–470.

(25) Cerullo, G.; Polli, D.; Lanzani, G.; De Silvestri, S.; Hashimoto, H.; Cogdell, R. J. Photosynthetic Light Harvesting by Carotenoids: Detection of an Intermediate Excited State. *Science* **2002**, *298*, 2395–2398.

(26) Cheng, Y.-C.; Fleming, G. R. Dynamics of Light Harvesting in Photosynthesis. *Annu. Rev. Phys. Chem.* **2009**, *60*, 241–262.

(27) Herbst, J.; Heyne, K.; Diller, R. Femtosecond Infrared Spectroscopy of Bacteriorhodopsin Chromophore Isomerization. *Science* **2002**, *297*, 822–825.

(28) Tapavicza, E.; Tavernelli, I.; Rothlisberger, U. Trajectory Surface Hopping within Linear Response Time-Dependent Density-Functional Theory. *Phys. Rev. Lett.* **2007**, *98*, 023001.

(29) Mathew, S.; Yella, A.; Gao, P.; Humphry-Baker, R.; Curchod, B. F. E.; Ashari-Astani, N.; Tavernelli, I.; Rothlisberger, U.; Nazeeruddin, M. K.; Grätzel, M. Dye-Sensitized Solar Cells with 13% Efficiency Achieved Through the Molecular Engineering of Porphyrin Sensitizers. *Nat. Chem.* **2014**, *6*, 242–247.

(30) Bartók, A. P.; De, S.; Poelking, C.; Bernstein, N.; Kermode, J. R.; Csányi, G.; Ceriotti, M. Machine Learning Unifies the Modeling of Materials and Molecules. *Sci. Adv.* **2017**, *3*, e1701816.

(31) O'Boyle, N. M.; Campbell, C. M.; Hutchison, G. R. Computational Design and Selection of Optimal Organic Photovoltaic Materials. *J. Phys. Chem. C* **2011**, *115*, 16200–16210.

(32) Lee, M.-H. Robust Random Forest Based Non-Fullerene Organic Solar Cells Efficiency Prediction. *Org. Electron.* **2020**, *76*, 105465.

(33) Schultz, T.; Samoylova, E.; Radloff, W.; Hertel, I. V.; Sobolewski, A. L.; Domcke, W. Efficient Deactivation of a Model Base Pair via Excited-State Hydrogen Transfer. *Science* **2004**, *306*, 1765–1768.

(34) Schreier, W. J.; Schrader, T. E.; Koller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. Thymine Dimerization in DNA Is an Ultrafast Photoreaction. *Science* **2007**, *315*, 625–629.

(35) Rauer, C.; Nogueira, J. J.; Marquetand, P.; González, L. Cyclobutane Thymine Photodimerization Mechanism Revealed by Nonadiabatic Molecular Dynamics. *J. Am. Chem. Soc.* **2016**, *138*, 15911–15916.

(36) Harris, D. C.; Bertolucci, M. D. *Symmetry and Spectroscopy: an Introduction to Vibrational and Electronic Spectroscopy*; Dover Publications: New York, 1989.

(37) Ng, C.-Y. *Vacuum Ultraviolet Photoionization and Photodissociation of Molecules and Clusters*; World Scientific, 1991.

(38) Zewail, A. H. *Femtochemistry: Ultrafast Dynamics of the Chemical Bond*; World Scientific, 1994; pp 3–22.

(39) Brixner, T.; Pfeifer, T.; Gerber, G.; Wollenhaupt, M.; Baumert, T. In *Femtosecond Laser Spectroscopy*; Hannaford, P., Ed.; Springer-Verlag: New York, 2005; pp 225–266.

(40) Iqbal, A.; Stavros, V. G. Active Participation of  $1\pi\sigma^*$  States in the Photodissociation of Tyrosine and its Subunits. *J. Phys. Chem. Lett.* **2010**, *1*, 2274–2278.

(41) Kowalewski, M.; Fingerhut, B. P.; Dorfman, K. E.; Bennett, K.; Mukamel, S. Simulating Coherent Multidimensional Spectroscopy of Nonadiabatic Molecular Processes: From the Infrared to the X-Ray Regime. *Chem. Rev.* **2017**, *117*, 12165–12226.

(42) Soorkia, S.; Juvet, C.; Grégoire, G. UV Photoinduced Dynamics of Conformer-Resolved Aromatic Peptides. *Chem. Rev.* **2020**, *120*, 3296–3327.

(43) Liu, Y.; et al. Spectroscopic and Structural Probing of Excited-State Molecular Dynamics with Time-Resolved Photoelectron Spectroscopy and Ultrafast Electron Diffraction. *Phys. Rev. X* **2020**, *10*, 021016.

(44) Matsika, S.; Krylov, A. I. Introduction: Theoretical Modeling of Excited State Processes. *Chem. Rev.* **2018**, *118*, 6925–6926.

(45) Martínez, T. J. Insights for Light-Driven Molecular Devices from *Ab Initio* Multiple Spawning Excited-State Dynamics of Organic and Biological Chromophores. *Acc. Chem. Res.* **2006**, *39*, 119–126.

(46) Barbatti, M.; Sellner, B.; Aquino, A. J. A.; Lischka, H. In *Radiation Induced Molecular Phenomena in Nucleic Acids*; Shukla, M., Leszczynski, J., Eds.; Challenges and Advances in Computational Chemistry and Physics; Springer Netherlands, 2008; Vol. 5, pp 209–235.

(47) Subotnik, J. E.; Jain, A.; Landry, B.; Petit, A.; Ouyang, W.; Bellonzi, N. Understanding the Surface Hopping View of Electronic Transitions and Decoherence. *Annu. Rev. Phys. Chem.* **2016**, *67*, 387–417.

(48) Curchod, B. F. E.; Martínez, T. J. *Ab Initio* Nonadiabatic Quantum Molecular Dynamics. *Chem. Rev.* **2018**, *118*, 3305–3336.

(49) Ashfold, M. N. R.; Bain, M.; Hansen, C. S.; Ingle, R. A.; Karsili, T. N. V.; Marchetti, B.; Murdock, D. Exploring the Dynamics of the Photoinduced Ring-Opening of Heterocyclic Molecules. *J. Phys. Chem. Lett.* **2017**, *8*, 3440–3451.

(50) Tajti, A.; Fogarasi, G.; Szalay, P. G. Reinterpretation of the UV Spectrum of Cytosine: Only Two Electronic Transitions? *ChemPhysChem* **2009**, *10*, 1603–1606.

(51) Barbatti, M.; Szymczak, J. J.; Aquino, A. J. A.; Nachtigallová, D.; Lischka, H. The Decay Mechanism of Photoexcited Guanine – A Nonadiabatic Dynamics Study. *J. Chem. Phys.* **2011**, *134*, 014304.

(52) Lu, Y.; Lan, Z.; Thiel, W. Photoinduced Phenomena in Nucleic Acids II; *Topics in Current Chemistry*; Springer Berlin Heidelberg, 2014; Vol. 356, pp 89–122.

(53) Ruckebauer, M.; Mai, S.; Marquetand, P.; González, L. Photoelectron Spectra of 2-Thiouracil, 4-Thiouracil, and 2,4-Dithiouracil. *J. Chem. Phys.* **2016**, *144*, 074303.

(54) Manathunga, M.; Yang, X.; Luk, H. L.; Gozem, S.; Frutos, L. M.; Valentini, A.; Ferrè, N.; Olivucci, M. Probing the Photodynamics of Rhodopsins with Reduced Retinal Chromophores. *J. Chem. Theory Comput.* **2016**, *12*, 839–850.

(55) Nogueira, J. J.; Plasser, F.; González, L. Electronic Delocalization, Charge Transfer and Hypochromism in the UV Absorption Spectrum of Polyadenine Unravelling by Multiscale Computations and Quantitative Wavefunction Analysis. *Chem. Sci.* **2017**, *8*, 5682–5691.

(56) Mai, S.; Mohamadzade, A.; Marquetand, P.; González, L.; Ullrich, S. Simulated and Experimental Time-Resolved Photoelectron Spectra of the Intersystem Crossing Dynamics in 2-Thiouracil. *Molecules* **2018**, *23*, 2836.

(57) Rauer, C.; Nogueira, J. J.; Marquetand, P.; González, L. Stepwise photosensitized thymine dimerization mediated by an exciton intermediate. *Monatsh. Chem.* **2018**, *149*, 1–9.

(58) Zobel, J. P.; Heindl, M.; Nogueira, J. J.; González, L. Vibrational Sampling and Solvent Effects on the Electronic Structure

of the Absorption Spectrum of 2-Nitronaphthalene. *J. Chem. Theory Comput.* **2018**, *14*, 3205–3217.

(59) Nelson, T. R.; White, A. J.; Bjorgaard, J. A.; Sifain, A. E.; Zhang, Y.; Nebgen, B.; Fernandez-Alberti, S.; Mozyrsky, D.; Roitberg, A. E.; Tretiak, S. Non-adiabatic Excited-State Molecular Dynamics: Theory and Applications for Modeling Photophysics in Extended Molecular Materials. *Chem. Rev.* **2020**, *120*, 2215–2287.

(60) Vacher, M.; Fdez. Galván, I.; Ding, B.-W.; Schramm, S.; Berraud-Pache, R.; Naumov, P.; Ferré, N.; Liu, Y.-J.; Navizet, I.; Roca-Sanjuán, D.; Baader, W. J.; Lindh, R. Chemi- and Bioluminescence of Cyclic Peroxides. *Chem. Rev.* **2018**, *118*, 6927–6974.

(61) Pathak, S.; et al. Tracking the Ultraviolet Photochemistry of Thiophenone During and Beyond the Initial Ultrafast Ring Opening. *Nat. Chem.* **2020**, *12*, 795–800.

(62) Maria Teresa Neves-Petersen, S. P.; Gajula, G. P. *UV Light Effects on Proteins: From Photochemistry to Nanomedicine, Molecular Photochemistry - Various Aspects*; IntechOpen, 2012; Chapter 7.

(63) Cadet, J.; Grand, A.; Douki, T. Photoinduced Phenomena in Nucleic Acids II; *Topics in Current Chemistry*; Springer Berlin Heidelberg, 2014; Vol. 356; pp 249–275.

(64) Segatta, F.; Cupellini, L.; Garavelli, M.; Mennucci, B. Quantum Chemical Modeling of the Photoinduced Activity of Multichromophoric Biosystems. *Chem. Rev.* **2019**, *119*, 9361–9380.

(65) Landry, B. R.; Subotnik, J. E. Quantifying the Lifetime of Triplet Energy Transfer Processes in Organic Chromophores: A Case Study of 4-(2-Naphthylmethyl)benzaldehyde. *J. Chem. Theory Comput.* **2014**, *10*, 4253–4263.

(66) Schultz, T.; Quenneville, J.; Levine, B.; Toniolo, A.; Martínez, T. J.; Lochbrunner, S.; Schmitt, M.; Shaffer, J. P.; Zgierski, M. Z.; Stolow, A. Mechanism and Dynamics of Azobenzene Photoisomerization. *J. Am. Chem. Soc.* **2003**, *125*, 8098–8099.

(67) Toniolo, A.; Olsen, S.; Manohar, L.; Martínez, T. J. Conical Intersection Dynamics in Solution: The Chromophore of Green Fluorescent Protein. *Faraday Discuss.* **2004**, *127*, 149–163.

(68) Domcke, W.; Yarkony, D.; Köppel, H. *Conical Intersections: Theory, Computation and Experiment*; *Advanced Series in Physical Chemistry*; World Scientific Publishing Company, 2011.

(69) Serrano-Andrés, L.; Merchán, M. Quantum chemistry of the excited state: 2005 overview. *J. Mol. Struct.: THEOCHEM* **2005**, *729*, 99–108.

(70) Chandrasekaran, A.; Kamal, D.; Batra, R.; Kim, C.; Chen, L.; Ramprasad, R. Solving the Electronic Structure Problem with Machine Learning. *npj Comput. Mater.* **2019**, *5*, 22.

(71) Jeong, W.; Stoneburner, S. J.; King, D.; Li, R.; Walker, A.; Lindh, R.; Gagliardi, L. Automation of Active Space Selection for Multireference Methods via Machine Learning on Chemical Bond Dissociation. *J. Chem. Theory Comput.* **2020**, *16*, 2389–2399.

(72) Carleo, G.; Troyer, M. Solving the Quantum Many-Body Problem with Artificial Neural Networks. *Science* **2017**, *355*, 602–606.

(73) Saito, H. Solving the Bose–Hubbard Model with Machine Learning. *J. Phys. Soc. Jpn.* **2017**, *86*, 093001.

(74) Nomura, Y.; Darmawan, A. S.; Yamaji, Y.; Imada, M. Restricted Boltzmann Machine Learning for solving strongly correlated quantum systems. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *96*, 205152.

(75) Han, J.; Zhang, L.; E, W. Solving Many-Electron Schrödinger Equation using Deep Neural Networks. *J. Comput. Phys.* **2019**, *399*, 108929.

(76) Townsend, J.; Vogiatzis, K. D. Data-Driven Acceleration of the coupled-Cluster Singles and Doubles Iterative Solver. *J. Phys. Chem. Lett.* **2019**, *10*, 4129–4135.

(77) Schütt, K. T.; Gastegger, M.; Tkatchenko, A.; Müller, K.-R.; Maurer, R. J. Unifying Machine Learning and quantum chemistry with a deep neural network for molecular wavefunctions. *Nat. Commun.* **2019**, *10*, 5024.

(78) Pfau, D.; Spencer, J. S.; Matthews, A. G. D. G.; Foulkes, W. M. C. Ab initio solution of the many-electron Schrödinger equation with deep neural networks. *Phys. Rev. Research* **2020**, *2*, 033429.

(79) Hermann, J.; Schätzle, Z.; Noé, F. Deep-Neural-Network Solution of the Electronic Schrödinger Equation. *Nat. Chem.* **2020**, *12*, 891–897.

(80) Gastegger, M.; McSloy, A.; Luya, M.; Schütt, K. T.; Maurer, R. J. A Deep Neural Network for Molecular Wave Functions in Quasi-Atomic Minimal Basis Representation. *J. Chem. Phys.* **2020**, *153*, 044123.

(81) Hegde, G.; Bowen, R. C. Machine-Learned Approximations to Density Functional Theory Hamiltonians. *Sci. Rep.* **2017**, *7*, 42669.

(82) Brockherde, F.; Vogt, L.; Li, L.; Tuckerman, M. E.; Burke, K.; Müller, K.-R. Bypassing the Kohn-Sham Equations with Machine Learning. *Nat. Commun.* **2017**, *8*, 872.

(83) Gastegger, M.; González, L.; Marquetand, P. Exploring Density Functional Subspaces with Genetic Algorithms. *Monatsh. Chem.* **2019**, *150*, 173–182.

(84) Nelson, J.; Tiwari, R.; Sanvito, S. Machine Learning Density Functional Theory for the Hubbard Model. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2019**, *99*, 075132.

(85) Cheng, L.; Welborn, M.; Christensen, A. S.; Miller, T. F. A Universal Density Matrix Functional from Molecular Orbital-Based Machine Learning: Transferability Across Organic Molecules. *J. Chem. Phys.* **2019**, *150*, 131103.

(86) Lei, X.; Medford, A. J. Design and Analysis of Machine Learning Exchange–Correlation Functionals via Rotationally Invariant Convolutional Descriptors. *Phys. Rev. Materials* **2019**, *3*, 063801.

(87) Zhou, Y.; Wu, J.; Chen, S.; Chen, G. Toward the Exact Exchange–Correlation Potential: A Three-Dimensional Convolutional Neural Network Construct. *J. Phys. Chem. Lett.* **2019**, *10*, 7264–7269.

(88) Kolb, B.; Lentz, L. C.; Kolpak, A. M. Discovering Charge Density Functionals and Structure-Property Relationships with PROPhet: A General Framework for Coupling Machine Learning and First-Principles Methods. *Sci. Rep.* **2017**, *7*, 1192.

(89) Willatt, M. J.; Musil, F.; Ceriotti, M. Atom-Density Representations for Machine Learning. *J. Chem. Phys.* **2019**, *150*, 154110.

(90) Choo, K.; Carleo, G.; Regnault, N.; Neupert, T. Symmetries and Many-Body Excitations with Neural-Network Quantum States. *Phys. Rev. Lett.* **2018**, *121*, 167204.

(91) Eckhoff, M.; Lausch, K. N.; Blöchl, P. E.; Behler, J. Predicting Oxidation and Spin States by High-Dimensional Neural Networks: Applications to Lithium Manganese Oxide Spinels. *arXiv* **2020**, 2007.00335.

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

(93) Carbogno, C.; Behler, J.; Reuter, K.; Groß, A. Signatures of Nonadiabatic O<sub>2</sub> Dissociation at Al(111): First-Principles Fewest-Switches Study. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2010**, *81*, 035410.

(94) Westermayr, J.; Gastegger, M.; Menger, M. F. S. J.; Mai, S.; González, L.; Marquetand, P. Machine Learning Enables Long Time Scale Molecular Photodynamics Simulations. *Chem. Sci.* **2019**, *10*, 8100–8107.

(95) Westermayr, J.; Faber, F. A.; Christensen, A. S.; von Lilienfeld, O. A.; Marquetand, P. Neural Networks and Kernel Ridge Regression for Excited States Dynamics of CH<sub>2</sub>NH<sub>2</sub><sup>+</sup>: From Single-State to Multi-State Representations and Multi-Property Machine Learning Models. *Mach. Learn.: Sci. Technol.* **2020**, *1*, 025009.

(96) Shen, Y.; Yarkony, D. R. Construction of Quasi-diabatic Hamiltonians That Accurately Represent *Ab Initio* Determined Adiabatic Electronic States Coupled by Conical Intersections for Systems on the Order of 15 Atoms. Application to Cyclopentoxide Photoelectron Detachment in the Full 39 Degrees of Freedom. *J. Phys. Chem. A* **2020**, *124*, 4539–4548.

(97) Zhang, Y.; Ye, S.; Zhang, J.; Hu, C.; Jiang, J.; Jiang, B. Efficient and Accurate Simulations of Vibrational and Electronic Spectra with

Symmetry-Preserving Neural Network Models for Tensorial Properties. *J. Phys. Chem. B* **2020**, *124*, 7284–7290.

(98) Carbogno, C.; Behler, J.; Groß, A.; Reuter, K. Fingerprints for Spin-Selection Rules in the Interaction Dynamics of O<sub>2</sub> at Al(111). *Phys. Rev. Lett.* **2008**, *101*, 096104.

(99) Polyak, I.; Richings, G. W.; Habershon, S.; Knowles, P. J. Direct Quantum Dynamics using Variational Gaussian Wavepackets and Gaussian Process Regression. *J. Chem. Phys.* **2019**, *150*, 041101.

(100) Hobday, S.; Smith, R.; Belbruno, J. Applications of Neural Networks to Fitting Interatomic Potential Functions. *Modell. Simul. Mater. Sci. Eng.* **1999**, *7*, 397.

(101) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403.

(102) Rupp, M.; Tkatchenko, A.; Müller, K.-R.; von Lilienfeld, O. A. Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning. *Phys. Rev. Lett.* **2012**, *108*, 058301.

(103) Li, Z.; Kermode, J. R.; De Vita, A. Molecular Dynamics with On-the-Fly Machine Learning of Quantum-Mechanical Forces. *Phys. Rev. Lett.* **2015**, *114*, 096405.

(104) von Lilienfeld, O. A.; Ramakrishnan, R.; Rupp, M.; Knoll, A. Fourier Series of Atomic Radial Distribution Functions: A Molecular Fingerprint for Machine Learning Models of Quantum Chemical Properties. *Int. J. Quantum Chem.* **2015**, *115*, 1084–1093.

(105) Gastegger, M.; Marquetand, P. High-Dimensional Neural Network Potentials for Organic Reactions and an Improved Training Algorithm. *J. Chem. Theory Comput.* **2015**, *11*, 2187–2198.

(106) Rupp, M.; Ramakrishnan, R.; von Lilienfeld, O. A. Machine Learning for Quantum Mechanical Properties of Atoms in Molecules. *J. Phys. Chem. Lett.* **2015**, *6*, 3309–3313.

(107) Behler, J. Perspective: Machine Learning Potentials for Atomistic Simulations. *J. Chem. Phys.* **2016**, *145*, 170901.

(108) Artrith, N.; Urban, A. An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for TiO<sub>2</sub>. *Comput. Mater. Sci.* **2016**, *114*, 135–150.

(109) Gastegger, M.; Kauffmann, C.; Behler, J.; Marquetand, P. 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.* **2016**, *144*, 194110.

(110) Artrith, N.; Urban, A.; Ceder, G. Efficient and Accurate Machine-Learning Interpolation of Atomic Energies in Compositions with Many Species. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *96*, 014112.

(111) Gastegger, M.; Behler, J.; Marquetand, P. Machine Learning Molecular Dynamics for the Simulation of Infrared Spectra. *Chem. Sci.* **2017**, *8*, 6924–6935.

(112) Deringer, V. L.; Csányi, G. Machine Learning Based Interatomic Potential for Amorphous Carbon. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *95*, 094203.

(113) Botu, V.; Batra, R.; Chapman, J.; Ramprasad, R. Machine Learning Force Fields: Construction, Validation, and Outlook. *J. Phys. Chem. C* **2017**, *121*, 511–522.

(114) Glielmo, A.; Sollich, P.; De Vita, A. Accurate Interatomic Force Fields via Machine Learning with Covariant Kernels. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *95*, 214302.

(115) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: An Extensible Neural Network Potential with DFT Accuracy at Force Field. *Computational Cost. Chem. Sci.* **2017**, *8*, 3192–3203.

(116) Fujikake, S.; Deringer, V. L.; Lee, T. H.; Krynski, M.; Elliott, S. R.; Csányi, G. Gaussian Approximation Potential Modeling of Lithium Intercalation in Carbon Nanostructures. *J. Chem. Phys.* **2018**, *148*, 241714.

(117) Behler, J. First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems. *Angew. Chem., Int. Ed.* **2017**, *56*, 12828–12840.

(118) Zong, H.; Paliana, G.; Ding, X.; Ackland, G. J.; Lookman, T. Developing an Interatomic Potential for Martensitic Phase Trans-

formations in Zirconium by Machine Learning. *npj comput Mater.* **2018**, *4*.

(119) Wood, M. A.; Thompson, A. P. Extending the Accuracy of the SNAP Interatomic Potential Form. *J. Chem. Phys.* **2018**, *148*, 241721.

(120) Chen, X.; Jørgensen, M. S.; Li, J.; Hammer, B. Atomic Energies from a Convolutional Neural Network. *J. Chem. Theory Comput.* **2018**, *14*, 3933–3942.

(121) Bartók, A. P.; Kermode, J.; Bernstein, N.; Csányi, G. Machine Learning a General-Purpose Interatomic Potential for Silicon. *Phys. Rev. X* **2018**, *8*, 041048.

(122) Chmiela, S.; Sauceda, H. E.; Müller, K.-R.; Tkatchenko, A. Towards Exact Molecular Dynamics Simulations with Machine-Learned Force Fields. *Nat. Commun.* **2018**, *9*, 3887.

(123) Imbalzano, G.; Anelli, A.; Giofré, D.; Klees, S.; Behler, J.; Ceriotti, M. Automatic Selection of Atomic Fingerprints and Reference Configurations for Machine-Learning Potentials. *J. Chem. Phys.* **2018**, *148*, 241730.

(124) Zhang, L.; Han, J.; Wang, H.; Saidi, W. A.; Car, R.; Weinan, E. End-to-end Symmetry Preserving Inter-atomic Potential Energy Model for Finite and Extended Systems. In *Proceedings of the 32Nd International conference on Neural Information Processing Systems, USA, 2018*; pp 4441–4451.

(125) Zhang, L.; Han, J.; Wang, H.; Car, R.; E, W. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Phys. Rev. Lett.* **2018**, *120*, 143001.

(126) Chan, H.; Narayanan, B.; Cherukara, M. J.; Sen, F. G.; Sasikumar, K.; Gray, S. K.; Chan, M. K. Y.; Sankaranarayanan, S. K. R. S. Machine Learning Classical Interatomic Potentials for Molecular Dynamics from First-Principles Training Data. *J. Phys. Chem. C* **2019**, *123*, 6941–6957.

(127) Faber, F. A.; Christensen, A. S.; Huang, B.; von Lilienfeld, O. A. Alchemical and Structural Distribution Based Representation for Universal Quantum Machine Learning. *J. Chem. Phys.* **2018**, *148*, 241717.

(128) Wang, H.; Yang, W. Toward Building Protein Force Fields by Residue-Based Systematic Molecular Fragmentation and Neural Network. *J. Chem. Theory Comput.* **2019**, *15*, 1409–1417.

(129) Gerrits, N.; Shakouri, K.; Behler, J.; Kroes, G.-J. Accurate Probabilities for Highly Activated Reaction of Polyatomic Molecules on Surfaces Using a High-Dimensional Neural Network Potential: CHD<sub>3</sub> + Cu(111). *J. Phys. Chem. Lett.* **2019**, *10*, 1763–1768.

(130) Chmiela, S.; Sauceda, H. E.; Poltavsky, I.; Müller, K.-R.; Tkatchenko, A. sGDML: Constructing Accurate and Data Efficient Molecular Force Fields using Machine Learning. *Comput. Phys. Commun.* **2019**, *240*, 38–45.

(131) Carleo, G.; Cirac, I.; Cranmer, K.; Daudet, L.; Schuld, M.; Tishby, N.; Vogt-Maranto, L.; Zdeborová, L. Machine Learning and the Physical Sciences. *Rev. Mod. Phys.* **2019**, *91*, 045002.

(132) Krems, R. V. Bayesian Machine Learning for Quantum Molecular Dynamics. *Phys. Chem. Chem. Phys.* **2019**, *21*, 13392–13410.

(133) Deringer, V. L.; Caro, M. A.; Csányi, G. Machine Learning Interatomic Potentials as Emerging Tools for Materials Science. *Adv. Mater.* **2019**, *31*, 1902765.

(134) Ward, L.; Blaiszik, B.; Foster, I.; Assary, R. S.; Narayanan, B.; Curtiss, L. Machine Learning Prediction of Accurate Atomization Energies of Organic Molecules from Low-Fidelity Quantum Chemical Calculations. *MRS Commun.* **2019**, *9*, 891–899.

(135) Noé, F.; Tkatchenko, A.; Müller, K.-R.; Clementi, C. Machine Learning for Molecular Simulation. *Annu. Rev. Phys. Chem.* **2020**, *71*, 361–390.

(136) Alborzpour, J. P.; Tew, D. P.; Habershon, S. Efficient and Accurate Evaluation of Potential Energy Matrix Elements for Quantum Dynamics using Gaussian Process Regression. *J. Chem. Phys.* **2016**, *145*, 174112.

(137) Cheng, Z.; Zhao, D.; Ma, J.; Li, W.; Li, S. An On-the-Fly Approach to Construct Generalized Energy-Based Fragmentation Machine Learning Force Fields of Complex Systems. *J. Phys. Chem. A* **2020**, *124*, 5007–5014.



- (138) Behler, J.; Reuter, K.; Scheffler, M. Nonadiabatic Effects in the Dissociation of Oxygen Molecules at the Al(111) Surface. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2008**, *77*, 115421.
- (139) Hu, D.; Xie, Y.; Li, X.; Li, L.; Lan, Z. Inclusion of Machine Learning Kernel Ridge Regression Potential Energy Surfaces in On-the-Fly Nonadiabatic Molecular Dynamics Simulation. *J. Phys. Chem. Lett.* **2018**, *9*, 2725–2732.
- (140) Dral, P. O.; Barbatti, M.; Thiel, W. Nonadiabatic Excited-State Dynamics with Machine Learning. *J. Phys. Chem. Lett.* **2018**, *9*, 5660–5663.
- (141) Chen, W.-K.; Liu, X.-Y.; Fang, W.-H.; Dral, P. O.; Cui, G. Deep Learning for Nonadiabatic Excited-State Dynamics. *J. Phys. Chem. Lett.* **2018**, *9*, 6702–6708.
- (142) Williams, D. M. G.; Eisfeld, W. Neural Network Diabatization: A New Ansatz for Accurate High-Dimensional Coupled Potential Energy Surfaces. *J. Chem. Phys.* **2018**, *149*, 204106.
- (143) Xie, C.; Zhu, X.; Yarkony, D. R.; Guo, H. Permutation Invariant Polynomial Neural Network Approach to Fitting Potential Energy Surfaces. IV. Coupled Diabatic Potential Energy Matrices. *J. Chem. Phys.* **2018**, *149*, 144107.
- (144) Guan, Y.; Zhang, D. H.; Guo, H.; Yarkony, D. R. Representation of Coupled Adiabatic Potential Energy Surfaces using Neural Network Based Quasi-Diabatic Hamiltonians: 1,2 <sup>2</sup>A' States of LiFH. *Phys. Chem. Chem. Phys.* **2019**, *21*, 14205.
- (145) Richings, G. W.; Habershon, S. MCTDH on-the-Fly: Efficient Grid-Based Quantum Dynamics without Pre-Computed Potential Energy Surfaces. *J. Chem. Phys.* **2018**, *148*, 134116.
- (146) Richings, G. W.; Robertson, C.; Habershon, S. Improved on-the-Fly MCTDH Simulations with Many-Body-Potential Tensor Decomposition and Projection Diabatization. *J. Chem. Theory Comput.* **2019**, *15*, 857–870.
- (147) Guan, Y.; Guo, H.; Yarkony, D. R. Neural Network Based Quasi-Diabatic Hamiltonians with Symmetry Adaptation and a Correct Description of Conical Intersections. *J. Chem. Phys.* **2019**, *150*, 214101.
- (148) Wang, Y.; Xie, C.; Guo, H.; Yarkony, D. R. A Quasi-Diabatic Representation of the 1,2<sup>1</sup>A States of Methylamine. *J. Phys. Chem. A* **2019**, *123*, 5231–5241.
- (149) Richings, G. W.; Habershon, S. Direct Grid-Based Quantum Dynamics on Propagated Diabatic Potential Energy Surfaces. *Chem. Phys. Lett.* **2017**, *683*, 228–233.
- (150) Netzloff, H. M.; Collins, M. A.; Gordon, M. S. Growing Multiconfigurational Potential Energy Surfaces with Applications to X + H<sub>2</sub> (X = C, N, O) Reactions. *J. Chem. Phys.* **2006**, *124*, 154104.
- (151) Bettens, R. P. A.; Collins, M. A. Learning to Interpolate Molecular Potential Energy Surfaces with Confidence: A Bayesian Approach. *J. Chem. Phys.* **1999**, *111*, 816–826.
- (152) Ghosh, K.; Stuke, A.; Todorović, M.; Jørgensen, P. B.; Schmidt, M. N.; Vehtari, A.; Rinke, P. Deep Learning Spectroscopy: Neural Networks for Molecular Excitation Spectra. *Adv. Sci.* **2019**, *6*, 1801367.
- (153) Kananenka, A. A.; Yao, K.; Corcelli, S. A.; Skinner, J. L. Machine Learning for Vibrational Spectroscopic Maps. *J. Chem. Theory Comput.* **2019**, *15*, 6850–6858.
- (154) Roch, L. M.; Saikin, S. K.; Häse, F.; Friederich, P.; Goldsmith, R. H.; León, S.; Aspuru-Guzik, A. From Absorption Spectra to Charge Transfer in Nanoaggregates of Oligomers with Machine Learning. *ACS Nano* **2020**, *14*, 6589.
- (155) Rankine, C. D.; Madkhali, M. M. M.; Penfold, T. J. A Deep Neural Network for the Rapid Prediction of X-Ray Absorption Spectra. *J. Phys. Chem. A* **2020**, *124*, 4263–4270.
- (156) Pereira, F.; Xiao, K.; Latino, D. A. R. S.; Wu, C.; Zhang, Q.; Aires-de Sousa, J. Machine Learning Methods to Predict Density Functional Theory B3LYP Energies of HOMO and LUMO Orbitals. *J. Chem. Inf. Model.* **2017**, *57*, 11–21.
- (157) Isayev, O.; Oses, c.; Toher, c.; Gossett, E.; Curtarolo, S.; Tropsha, A. Universal Fragment Descriptors for Predicting Properties of Inorganic Crystals. *Nat. Commun.* **2017**, *8*, 15679.
- (158) Pronobis, W.; Schütt, K. R.; Tkatchenko, A.; Müller, K.-R. Capturing Intensive and Extensive DFT/TDDFT Molecular Properties with Machine Learning. *Eur. Phys. J. B* **2018**, *91*, 178.
- (159) Stuke, A.; Todorović, M.; Rupp, M.; Kunkel, C.; Ghosh, K.; Himanen, L.; Rinke, P. Chemical Diversity in Molecular Orbital Energy Predictions with Kernel Ridge Regression. *J. Chem. Phys.* **2019**, *150*, 204121.
- (160) Häse, F.; Fdez. Galván, I.; Aspuru-Guzik, A.; Lindh, R.; Vacher, M. How Machine Learning can Assist the Interpretation of *Ab Initio* Molecular Dynamics Simulations and Conceptual Understanding of Chemistry. *Chem. Sci.* **2019**, *10*, 2298–2307.
- (161) Häse, F.; Kreisbeck, C.; Aspuru-Guzik, A. Machine Learning for Quantum Dynamics: Deep Learning of Excitation Energy Transfer Properties. *Chem. Sci.* **2017**, *8*, 8419–8426.
- (162) Häse, F.; Valleau, S.; Pyzer-Knapp, E.; Aspuru-Guzik, A. Machine Learning Exciton Dynamics. *Chem. Sci.* **2016**, *7*, 5139–5147.
- (163) Teunissen, J. L.; De Proft, F.; De Vleschouwer, F. Tuning the HOMO–LUMO Energy Gap of Small Diamondoids Using Inverse Molecular Design. *J. Chem. Theory Comput.* **2017**, *13*, 1351–1365.
- (164) Liu, D.; Tan, Y.; Khoram, E.; Yu, Z. Training Deep Neural Networks for the Inverse Design of Nanophotonic Structures. *ACS Photonics* **2018**, *5*, 1365–1369.
- (165) Elton, D. C.; Boukouvalas, Z.; Fuge, M. D.; Chung, P. W. Deep Learning for Molecular Design – A Review of the State of the Art. *Mol. Syst. Des. Eng.* **2019**, *4*, 828–849.
- (166) Sanchez-Lengeling, B.; Aspuru-Guzik, A. Inverse Molecular Design using Machine Learning: Generative Models for Matter Engineering. *Science* **2018**, *361*, 360–365.
- (167) Goldsmith, B. R.; Esterhuizen, J.; Liu, J.-X.; Bartel, C. J.; Sutton, C. Machine Learning for Heterogeneous Catalyst Design and Discovery. *AIChE J.* **2018**, *64*, 2311–2323.
- (168) Davies, D. W.; Butler, K. T.; Isayev, O.; Walsh, A. Materials Discovery by Chemical Analogy: Role of Oxidation States in Structure Prediction. *Faraday Discuss.* **2018**, *211*, 553–568.
- (169) Anatole von Lilienfeld, O.; Müller, K.-R.; Tkatchenko, A. Exploring Chemical compound Space with Quantum-Based Machine Learning. *Nat. Rev. Chem.* **2020**, *4*, 347–358.
- (170) Freeze, J. G.; Kelly, H. R.; Batista, V. S. Search for Catalysts by Inverse Design: Artificial Intelligence, Mountain Climbers, and Alchemists. *Chem. Rev.* **2019**, *119*, 6595–6612.
- (171) Cartwright, H. M., Ed. *Machine Learning in Chemistry*; Theoretical and Computational Chemistry Series; The Royal Society of Chemistry, 2020.
- (172) Gastegger, M.; Marquetand, P. In *Machine Learning Meets Quantum Physics*; Schütt, K. T., Chmiela, S., von Lilienfeld, O. A., Tkatchenko, A., Tsuda, K., Müller, K.-R., Eds.; Springer International Publishing: Cham, 2020; pp 233–252.
- (173) Schütt, K. T., Chmiela, S., von Lilienfeld, O. A., Tkatchenko, A., Tsuda, K., Müller, K.-R., Eds. *Machine Learning Meets Quantum Physics*; Springer International Publishing, 2020.
- (174) Janet, J. P.; Kulik, H. J. *Machine Learning in Chemistry*; American Chemical Society: Washington, DC, USA, 2020.
- (175) Westermayr, J.; Marquetand, P. Machine Learning and Excited-State Molecular Dynamics. *Mach. Learn.: Sci. Technol.* **2020**, *1*, 043001.
- (176) Duan, C.; Liu, F.; Nandy, A.; Kulik, H. J. Semi-supervised Machine Learning Enables the Robust Detection of Multireference Character at Low Cost. *J. Phys. Chem. Lett.* **2020**, *11*, 6640–6648.
- (177) Duan, C.; Liu, F.; Nandy, A.; Kulik, H. J. Data-Driven Approaches Can Overcome the Cost–Accuracy Trade-Off in Multireference Diagnostics. *J. Chem. Theory Comput.* **2020**, *16*, 4373–4387.
- (178) González, L.; Lindh, R. *Quantum Chemistry and Dynamics of Excited States: Methods and Applications*; John Wiley and Sons Ltd, 2020.
- (179) Park, J. W.; Al-Saadon, R.; MacLeod, M. K.; Shiozaki, T.; Vlaisavljevich, B. Multireference Electron Correlation Methods: Journeys along Potential Energy Surfaces. *Chem. Rev.* **2020**, *120*, 5878.

- (180) Akimov, A. V.; Prezhdo, O. V. Large-Scale Computations in Chemistry: A Bird's Eye View of a Vibrant Field. *Chem. Rev.* **2015**, *115*, 5797–5890.
- (181) Frutos, L. M.; Andruniów, T.; Santoro, F.; Ferré, N.; Olivucci, M. Tracking the Excited-State Time Evolution of the Visual Pigment with Multiconfigurational Quantum Chemistry. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 7764–7769.
- (182) Menger, M. F. S. J.; Plasser, F.; Mennucci, B.; González, L. Surface Hopping within an Exciton Picture. An Electrostatic Embedding Scheme. *J. Chem. Theory Comput.* **2018**, *14*, 6139–6148.
- (183) Dou, W.; Subotnik, J. E. Nonadiabatic Molecular Dynamics at Metal Surfaces. *J. Phys. Chem. A* **2020**, *124*, 757–771.
- (184) Dou, W.; Nitzan, A.; Subotnik, J. E. Frictional Effects Near a Metal Surface. *J. Chem. Phys.* **2015**, *143*, 054103.
- (185) Tavernelli, I. Electronic Density Response of Liquid Water using Time-Dependent Density Functional Theory. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2006**, *73*, 094204.
- (186) Schütt, K. T.; Glawe, H.; Brockherde, F.; Sanna, A.; Müller, K. R.; Gross, E. K. U. How to Represent Crystal Structures for Machine Learning: Towards Fast Prediction of Electronic Properties. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, *89*, 205118.
- (187) Lee, J.; Seko, A.; Shitara, K.; Nakayama, K.; Tanaka, I. Prediction Model of Band Gap for Inorganic Compounds by Combination of Density Functional Theory Calculations and Machine Learning Techniques. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2016**, *93*, 115104.
- (188) Zhuo, Y.; Mansouri Tehrani, A.; Brgoch, J. Predicting the Band Gaps of Inorganic Solids by Machine Learning. *J. Phys. Chem. Lett.* **2018**, *9*, 1668–1673.
- (189) Pilania, G.; Gubernatis, J.; Lookman, T. Multi-Fidelity Machine Learning Models for Accurate Bandgap Predictions of Solids. *Comput. Mater. Sci.* **2017**, *129*, 156–163.
- (190) Spiering, P.; Shakouri, K.; Behler, J.; Kroes, G.-J.; Meyer, J. Orbital-Dependent Electronic Friction Significantly Affects the Description of Reactive Scattering of N<sub>2</sub> from Ru(0001). *J. Phys. Chem. Lett.* **2019**, *10*, 2957–2962.
- (191) Zhang, Y.; Maurer, R. J.; Jiang, B. Symmetry-Adapted High Dimensional Neural Network Representation of Electronic Friction Tensor of Adsorbates on Metals. *J. Phys. Chem. C* **2020**, *124*, 186–195.
- (192) Zhang, Y.; Maurer, R. J.; Guo, H.; Jiang, B. Hot-Electron Effects during Reactive Scattering of H<sub>2</sub> from Ag(111): The Interplay between Mode-Specific Electronic Friction and the Potential Energy Landscape. *Chem. Sci.* **2019**, *10*, 1089–1097.
- (193) Head-Gordon, M.; Tully, J. C. Molecular Dynamics with Electronic Frictions. *J. Chem. Phys.* **1995**, *103*, 10137–10145.
- (194) Douglas-Gallardo, O. A.; Berdakin, M.; Frauenheim, T.; Sánchez, C. G. Plasmon-Induced Hot-Carrier Generation Differences in Gold and Silver Nanoclusters. *Nanoscale* **2019**, *11*, 8604–8615.
- (195) Yin, R.; Zhang, Y.; Jiang, B. Strong Vibrational Relaxation of NO Scattered from Au(111): Importance of the Adiabatic Potential Energy Surface. *J. Phys. Chem. Lett.* **2019**, *10*, 5969–5974.
- (196) Rittmeyer, S. P.; Bukas, V. J.; Reuter, K. Energy dissipation at metal surfaces. *Adv. Phys.-X* **2018**, *3*, 1381574.
- (197) Therrien, A. J.; Kale, M. J.; Yuan, L.; Zhang, C.; Halas, N. J.; Christopher, P. Impact of Chemical Interface Damping on Surface Plasmon Dephasing. *Faraday Discuss.* **2019**, *214*, 59–72.
- (198) Wodtke, A. M.; Tully, J. C.; Auerbach, D. J. Electronically Non-Adiabatic Interactions of Molecules at Metal Surfaces: Can we Trust the Born–Oppenheimer Approximation for Surface Chemistry? *Int. Rev. Phys. Chem.* **2004**, *23*, 513–539.
- (199) Park, G. B.; Krüger, B. C.; Borodin, D.; Kitsopoulos, T. N.; Wodtke, A. M. Fundamental Mechanisms for Molecular Energy Conversion and Chemical Reactions at Surfaces. *Rep. Prog. Phys.* **2019**, *82*, 096401.
- (200) Jiang, B.; Guo, H. Dynamics in Reactions on Metal Surfaces: A Theoretical Perspective. *J. Chem. Phys.* **2019**, *150*, 180901.
- (201) Shenvi, N.; Roy, S.; Tully, J. C. Nonadiabatic Dynamics at Metal Surfaces: Independent-Electron Surface Hopping. *J. Chem. Phys.* **2009**, *130*, 174107.
- (202) Shenvi, N.; Roy, S.; Tully, J. C. Dynamical Steering and Electronic Excitation in NO Scattering from a Gold Surface. *Science* **2009**, *326*, 829–832.
- (203) Dou, W.; Schinabeck, C.; Thoss, M.; Subotnik, J. E. A broadened classical master equation approach for treating electron-nuclear coupling in non-equilibrium transport. *J. Chem. Phys.* **2018**, *148*, 102317.
- (204) Jiang, B.; Li, J.; Guo, H. High-Fidelity Potential Energy Surfaces for Gas Phase and Gas-Surface Scattering Processes from Machine Learning. *J. Phys. Chem. Lett.* **2020**, *11*, 5120–5131.
- (205) Buhrke, D.; Hildebrandt, P. Probing Structure and Reaction Dynamics of Proteins Using Time-Resolved Resonance Raman Spectroscopy. *Chem. Rev.* **2020**, *120*, 3577–3630.
- (206) Raimbault, N.; Grisafi, A.; Ceriotti, M.; Rossi, M. Using Gaussian Process Regression to Simulate the Vibrational Raman Spectra of Molecular Crystals. *New J. Phys.* **2019**, *21*, 105001.
- (207) Hu, W.; Ye, S.; Zhang, Y.; Li, T.; Zhang, G.; Luo, Y.; Mukamel, S.; Jiang, J. Machine Learning Protocol for Surface-Enhanced Raman Spectroscopy. *J. Phys. Chem. Lett.* **2019**, *10*, 6026–6031.
- (208) Lussier, F.; Thibault, V.; Charron, B.; Wallace, G. Q.; Masson, J.-F. Deep Learning and Artificial Intelligence Methods for Raman and Surface-Enhanced Raman Scattering. *TrAC, Trends Anal. Chem.* **2020**, *124*, 115796.
- (209) Fu, W.; Hopkins, W. S. Applying Machine Learning to Vibrational Spectroscopy. *J. Phys. Chem. A* **2018**, *122*, 167–171.
- (210) Aires-de Sousa, J.; Hemmer, M. C.; Gasteiger, J. Prediction of <sup>1</sup>H NMR Chemical Shifts Using Neural Networks. *Anal. Chem.* **2002**, *74*, 80–90.
- (211) Taguchi, A. T.; Evans, E. D.; Dikanov, S. A.; Griffin, R. G. Convolutional Neural Network Analysis of Two-Dimensional Hyperfine Sublevel Correlation Electron Paramagnetic Resonance Spectra. *J. Phys. Chem. Lett.* **2019**, *10*, 1115–1119.
- (212) Cobas, C. NMR Signal Processing, Prediction, and Structure Verification with Machine Learning Techniques. *Magn. Reson. Chem.* **2020**, *58*, 512–519.
- (213) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An overview of the Amber biomolecular simulation package. *WIREs Computational Molecular Science* **2013**, *3*, 198–210.
- (214) Brooks, B. r.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- (215) Eichenberger, A. P.; Allison, J. R.; Dolenc, J.; Geerke, D. P.; Horta, B. A. C.; Meier, K.; Oostenbrink, C.; Schmid, N.; Steiner, D.; Wang, D.; van Gunsteren, W. F. GROMOS++ Software for the Analysis of Biomolecular Simulation Trajectories. *J. Chem. Theory Comput.* **2011**, *7*, 3379–3390.
- (216) Reif, M. M.; Hünenberger, P. H.; Oostenbrink, C. New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field. *J. Chem. Theory Comput.* **2012**, *8*, 3705–3723.
- (217) Perthold, J. W.; Petrov, D.; Oostenbrink, C. Towards Automated Free Energy Calculation with Accelerated Enveloping Distribution Sampling (A-EDS). *J. Chem. Inf. Model.* **2020**, DOI: 10.1021/acs.jcim.0c00456.
- (218) Öhlknecht, C.; Lier, B.; Petrov, D.; Fuchs, J.; Oostenbrink, C. Correcting Electrostatic Artifacts due to Net-Charge Changes in the Calculation of Ligand Binding Free Energies. *J. Comput. Chem.* **2020**, *41*, 986–999.
- (219) Michlits, H.; Lier, B.; Pfanzagl, V.; Djinović-Carugo, K.; Furtmüller, P. G.; Oostenbrink, C.; Obinger, C.; Hofbauer, S. Actinobacterial Coproheme Decarboxylases Use Histidine as a Distal Base to Promote Compound I Formation. *ACS Catal.* **2020**, *10*, 5405–5418.
- (220) Brunk, E.; Rothlisberger, U. Mixed Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulations of Biological Systems in Ground and Electronically Excited States. *Chem. Rev.* **2015**, *115*, 6217–6263.

- (221) Bedrov, D.; Piquemal, J.-P.; Borodin, O.; MacKerell, A. D.; Roux, B.; Schröder, C. Molecular Dynamics Simulations of Ionic Liquids and Electrolytes Using Polarizable Force Fields. *Chem. Rev.* **2019**, *119*, 7940–7995.
- (222) Sosso, G. C.; Chen, J.; Cox, S. J.; Fitzner, M.; Pedevilla, P.; Zen, A.; Michaelides, A. Crystal Nucleation in Liquids: Open Questions and Future Challenges in Molecular Dynamics Simulations. *Chem. Rev.* **2016**, *116*, 7078–7116.
- (223) Venable, R. M.; Krämer, A.; Pastor, R. W. Molecular Dynamics Simulations of Membrane Permeability. *Chem. Rev.* **2019**, *119*, 5954–5997.
- (224) Marrink, S. J.; Corradi, V.; Souza, P. C.; Ingólfsson, H. I.; Tieleman, D. P.; Sansom, M. S. Computational Modeling of Realistic Cell Membranes. *Chem. Rev.* **2019**, *119*, 6184–6226.
- (225) *Biomolecular Simulations: Methods and Protocols*; Monticelli, L., Salonen, E., Eds.; Methods in Molecular Biology (Methods and Protocols), Vol. 924; Humana Press: Totowa, NJ, 2013.
- (226) Senftle, T. P., et al. The ReaxFF Reactive Force-Field: Development, Applications and Future Directions. *npj Comput. Mater.* **2016**, *2*.
- (227) Saucedo, H. E.; Chmiela, S.; Poltavsky, I.; Müller, K.-R.; Tkatchenko, A. In *Machine Learning Meets Quantum Physics*; Schütt, K. T., Chmiela, S., von Lilienfeld, O. A., Tkatchenko, A., Tsuda, K., Müller, K.-R., Eds.; Springer International Publishing: Cham, 2020; pp 277–307.
- (228) Noé, F. In *Machine Learning Meets Quantum Physics*; Schütt, K. T., Chmiela, S., von Lilienfeld, O. A., Tkatchenko, A., Tsuda, K., Müller, K.-R., Eds.; Springer International Publishing: Cham, 2020; pp 331–372.
- (229) Glielmo, A.; Zeni, C.; Fekete, Á.; De Vita, A. In *Machine Learning Meets Quantum Physics*; Schütt, K. T., Chmiela, S., von Lilienfeld, O. A., Tkatchenko, A., Tsuda, K., Müller, K.-R., Eds.; Springer International Publishing: Cham, 2020; pp 67–98.
- (230) Abbott, A. S.; Turney, J. M.; Zhang, B.; Smith, D. G. A.; Altarawy, D.; Schaefer, H. F. PES-Learn: An Open-Source Software Package for the Automated Generation of Machine Learning Models of Molecular Potential Energy Surfaces. *J. Chem. Theory Comput.* **2019**, *15*, 4386–4398.
- (231) Hellström, M.; Behler, J. In *Machine Learning Meets Quantum Physics*; Schütt, K. T., Chmiela, S., von Lilienfeld, O. A., Tkatchenko, A., Tsuda, K., Müller, K.-R., Eds.; Springer International Publishing: Cham, 2020; pp 253–275.
- (232) Vargas-Hernández, R. A.; Krems, R. V. In *Machine Learning Meets Quantum Physics*; Schütt, K. T., Chmiela, S., von Lilienfeld, O. A., Tkatchenko, A., Tsuda, K., Müller, K.-R., Eds.; Springer International Publishing: Cham, 2020; pp 171–194.
- (233) Pyykko, P. Relativistic Effects in Structural Chemistry. *Chem. Rev.* **1988**, *88*, 563–594.
- (234) Neese, F.; Petrenko, T.; Ganyushin, D.; Olbrich, G. Advanced Aspects of *Ab Initio* Theoretical Optical Spectroscopy of Transition Metal complexes: Multiplets, spin-orbit coupling and resonance Raman intensities. *Coord. Chem. Rev.* **2007**, *251*, 288–327.
- (235) Neese, F. Efficient and Accurate Approximations to the Molecular Spin-Orbit Coupling Operator and their Use in Molecular G-Tensor Calculations. *J. Chem. Phys.* **2005**, *122*, 034107.
- (236) Richter, M.; Marquetand, P.; González-Vázquez, J.; Sola, I.; González, L. SHARC: *Ab Initio* Molecular Dynamics with Surface Hopping in the Adiabatic Representation Including Arbitrary couplings. *J. Chem. Theory Comput.* **2011**, *7*, 1253–1258.
- (237) Mai, S.; Marquetand, P.; González, L. A General Method to Describe Intersystem Crossing Dynamics in Trajectory Surface Hopping. *Int. J. Quantum Chem.* **2015**, *115*, 1215–1231.
- (238) Yarkony, D. R. Nonadiabatic Quantum Chemistry - Past, Present, and Future. *Chem. Rev.* **2012**, *112*, 481–498.
- (239) Köppel, H.; Domcke, W.; Cederbaum, L. S. In *Conical Intersections*; Domcke, W., Yarkony, D. R., Köppel, H., Eds.; World Scientific: New York, 2004.
- (240) Plasser, F.; Gómez, S.; Menger, M. F. S. J.; Mai, S.; González, L. Highly Efficient Surface Hopping Dynamics using a Linear Vibronic Coupling Model. *Phys. Chem. Chem. Phys.* **2019**, *21*, 57–69.
- (241) Westermayr, J.; Marquetand, P. Deep Learning for UV Absorption Spectra with SchNarc: First Steps Toward Transferability in Chemical Compound Space. *J. Chem. Phys.* **2020**, *153*, 154112.
- (242) He, G. S.; Tan, L.-S.; Zheng, Q.; Prasad, P. N. Multiphoton Absorbing Materials: Molecular Designs, Characterizations, and Applications. *Chem. Rev.* **2008**, *108*, 1245–1330.
- (243) Marquetand, P.; Weinacht, T.; Rozgonyi, T.; González-Vázquez, J.; Geizler, D.; González, L. In *Advances in Multiphoton Processes and Spectroscopy*; Fujimura, Y., Ed.; World Scientific, Singapore, 2014; Vol. 21; pp 1–54.
- (244) Tagliamonti, V.; Sándor, P.; Zhao, A.; Rozgonyi, T.; Marquetand, P.; Weinacht, T. Nonadiabatic Dynamics and Multiphoton Resonances in Strong-Field Molecular Ionization with Few-Cycle Laser Pulses. *Phys. Rev. A: At., Mol., Opt. Phys.* **2016**, *93*, 051401.
- (245) M. Wollenhaupt, A. A.; Baumert, T. In *Springer Handbook of Lasers and Optics*; Träger, F., Ed.; Springer Science and Business Media, LLC: New York, 2007; Chapter 12, pp 937–983.
- (246) Hilborn, R. C. Einstein Coefficients, Cross Sections, f Values, Dipole Moments, and All That. *Am. J. Phys.* **1982**, *50*, 982–986.
- (247) Andrews, D. L. *Molecular Photophysics and Spectroscopy*; 2053-2571; Morgan & Claypool Publishers, 2014; pp 9–1 to 9–4.
- (248) Baryshnikov, G.; Minaev, B.; Ågren, H. Theory and Calculation of the Phosphorescence Phenomenon. *Chem. Rev.* **2017**, *117*, 6500–6537.
- (249) Silva, G. L.; Ediz, V.; Yaron, D.; Armitage, B. A. Experimental and Computational Investigation of Unsymmetrical Cyanine Dyes: Understanding Torsionally Fluorogenic Dyes. *J. Am. Chem. Soc.* **2007**, *129*, 5710–5718.
- (250) Hartschuh, A.; Pedrosa, H. N.; Novotny, L.; Krauss, T. D. Simultaneous Fluorescence and Raman Scattering from Single Carbon Nanotubes. *Science* **2003**, *301*, 1354–1356.
- (251) Terenziani, F.; Katan, C.; Badaeva, E.; Tretiak, S.; Blanchard-Desce, M. Enhanced Two-Photon Absorption of Organic Chromophores: Theoretical and Experimental Assessments. *Adv. Mater.* **2008**, *20*, 4641–4678.
- (252) Richings, G. W.; Habershon, S. A New Diabatization Scheme for Direct Quantum Dynamics: Procrustes Diabatization. *J. Chem. Phys.* **2020**, *152*, 154108.
- (253) Lischka, H.; Nachtigallová, D.; Aquino, A. J. A.; Szalay, P. G.; Plasser, F.; Machado, F. B. C.; Barbatti, M. Multireference Approaches for Excited States of Molecules. *Chem. Rev.* **2018**, *118*, 7293–7361.
- (254) Tannor, D. *Introduction to Quantum Mechanics: A Time-Dependent Perspective*; University Science Books: Sausalito, 2006.
- (255) Weinacht, T.; Pearson, B. *Time-Resolved Spectroscopy: An Experimental Perspective*; CRC Press: New York, 2019.
- (256) Mai, S.; Marquetand, P.; González, L. Nonadiabatic Dynamics: The SHARC Approach. *WIREs Comput. Mol. Sci.* **2018**, *8*, No. e1370.
- (257) Crespo-Otero, R.; Barbatti, M. Recent Advances and Perspectives on Nonadiabatic Mixed Quantum–Classical Dynamics. *Chem. Rev.* **2018**, *118*, 7026–7068.
- (258) Ibele, L. M.; Nicolson, A.; Curchod, B. F. E. Excited-State Dynamics of molecules with classically driven trajectories and Gaussians. *Mol. Phys.* **2020**, *118*, No. e1665199.
- (259) Yonehara, T.; Hanasaki, K.; Takatsuka, K. Fundamental Approaches to Nonadiabaticity: Toward a Chemical Theory beyond the Born–Oppenheimer Paradigm. *Chem. Rev.* **2012**, *112*, 499–542.
- (260) Casida, M.; Huix-Rotllant, M. Progress in Time-Dependent Density-Functional Theory. *Annu. Rev. Phys. Chem.* **2012**, *63*, 287–323.
- (261) Maitra, N. T. Perspective: Fundamental Aspects of Time-Dependent Density Functional Theory. *J. Chem. Phys.* **2016**, *144*, 220901.

- (262) Szalay, P. G.; Müller, T.; Gidofalvi, G.; Lischka, H.; Shepard, R. Multiconfiguration Self-Consistent Field and Multireference Configuration Interaction Methods and Applications. *Chem. Rev.* **2012**, *112*, 108–181.
- (263) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons, Ltd, 2014.
- (264) Roos, B. O.; Lindh, R.; Malmqvist, P. Å.; Veryazov, V.; Widmark, P. *Multiconfigurational Quantum Chemistry*; John Wiley & Sons, Ltd, 2016.
- (265) Born, M.; Oppenheimer, R. Zur Quantentheorie der Molekeln. *Ann. Phys.* **1927**, *389*, 457–484.
- (266) Kohn, W. Nobel Lecture: Electronic Structure of Matter – Wave Functions and Density Functionals. *Rev. Mod. Phys.* **1999**, *71*, 1253–1266.
- (267) Schrödinger, E. An Undulatory Theory of the Mechanics of Atoms and Molecules. *Phys. Rev.* **1926**, *28*, 1049–1070.
- (268) Erwin-Schrödinger – Nobel Lecture. <https://www.nobelprize.org/prizes/phy-sics/1933/schrodinger/lecture/>.
- (269) Yu, H. S.; Li, S. L.; Truhlar, D. G. Perspective: Kohn-Sham Density Functional Theory Descending a Staircase. *J. Chem. Phys.* **2016**, *145*, 130901.
- (270) Maurer, R. J.; Freysoldt, C.; Reilly, A. M.; Brandenburg, J. G.; Hofmann, O. T.; Björkman, T.; Lebègue, S.; Tkatchenko, A. Advances in Density-Functional Calculations for Materials Modeling. *Annu. Rev. Mater. Res.* **2019**, *49*, 1–30.
- (271) Benavides-Riveros, C. L.; Lathiotakis, N. N.; Marques, M. A. L. Towards a Formal Definition of Static and Dynamic Electronic Correlations. *Phys. Chem. Chem. Phys.* **2017**, *19*, 12655–12664.
- (272) Szabo, A.; Ostlund, N. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Books on Chemistry, Dover Publications, 2012.
- (273) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons, Ltd, 2014; Chapter 10, pp 433–522.
- (274) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons, Ltd, 2014; Chapter 11, pp 523–597.
- (275) Dreuw, A.; Wormit, M. The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *WIREs Comput. Mol. Sci.* **2015**, *5*, 82–95.
- (276) von Niessen, W.; Schirmer, J.; Cederbaum, L. Computational Methods for the One-Particle Green's Function. *Comput. Phys. Rep.* **1984**, *1*, 57–125.
- (277) Linderberg, J.; Öhrn, Y. *Propagators in Quantum Chemistry*; John Wiley & Sons, Ltd, 2005; Chapter 2, pp 3–6.
- (278) Melin, J.; Ayers, P.; Ortiz, J. The Electron-Propagator Approach to Conceptual Density-Functional Theory. *Proc. - Indian Acad. Sci., Chem. Sci.* **2005**, *117*, 387–400.
- (279) Corzo, H. H.; Ortiz, J. V. In *Löwdin Vol.*; Sabin, J. R., Brändas, E. J., Eds.; Advances in Quantum Chemistry; Academic Press, 2017; Vol. 74; pp 267 – 298.
- (280) Möller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.
- (281) Bartlett, R. J. Many-Body Perturbation Theory and Coupled Cluster Theory for Electron Correlation in Molecules. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359–401.
- (282) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons, Ltd, 2014; Chapter 13, pp 648–723.
- (283) Izsák, R. Single-Reference Coupled Cluster Methods for Computing Excitation Energies in Large Molecules: The Efficiency and Accuracy of Approximations. *WIREs Comput. Mol. Sci.* **2020**, *10*, No. e1445.
- (284) Krylov, A. I. Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker's Guide to Fock Space. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433–462.
- (285) Parrill, A.; Lipkowitz, K. *Reviews in Computational Chemistry*, Vol. 31; Reviews in Computational Chemistry; Wiley, 2018.
- (286) Pacifici, L. L. A., Verdichio, M. In *Computational Science and Its Applications – ICCSA 2013*; Murgante, B., et al., Eds.; Springer: Berlin, 2013; Vol. 7971.
- (287) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons, Ltd, 2014; Chapter 12, pp 598–647.
- (288) Roos, B. O.; Taylor, P. R.; Sigbahn, P. E.M. A Complete Active Space SCF Method (CASSCF) using a Density Matrix Formulated Super-CI Approach. *Chem. Phys.* **1980**, *48*, 157–173.
- (289) Roos, B. O.; Siegbahn, P. E. M. A Direct CI Method with a Multiconfigurational Reference State. *Int. J. Quantum Chem.* **1980**, *17*, 485–500.
- (290) Baiardi, A.; Reiher, M. The Density Matrix Renormalization Group in Chemistry and Molecular Physics: Recent Developments and New Challenges. *J. Chem. Phys.* **2020**, *152*, 040903.
- (291) Chan, G. K.-L.; Van Voorhis, T. Density-Matrix Renormalization-Group Algorithms with Nonorthogonal Orbitals and Non-Hermitian Operators, and Applications to Polyenes. *J. Chem. Phys.* **2005**, *122*, 204101.
- (292) Zgid, D.; Nooijen, M. The Density Matrix Renormalization Group Self-Consistent Field Method: Orbital Optimization with the Density Matrix Renormalization Group Method in the Active Space. *J. Chem. Phys.* **2008**, *128*, 144116.
- (293) Keller, S.; Dolfi, M.; Troyer, M.; Reiher, M. An Efficient Matrix Product Operator Representation of the Quantum Chemical Hamiltonian. *J. Chem. Phys.* **2015**, *143*, 244118.
- (294) Knecht, S.; Keller, S.; Autschbach, J.; Reiher, M. A Nonorthogonal State-Interaction Approach for Matrix Product State Wave Functions. *J. Chem. Theory Comput.* **2016**, *12*, 5881–5894.
- (295) Freitag, L.; Knecht, S.; Angeli, C.; Reiher, M. Multireference Perturbation Theory with Cholesky Decomposition for the Density Matrix Renormalization Group. *J. Chem. Theory Comput.* **2017**, *13*, 451–459.
- (296) Freitag, L.; Ma, Y.; Baiardi, A.; Knecht, S.; Reiher, M. Approximate Analytical Gradients and Nonadiabatic Couplings for the State-Average Density Matrix Renormalization Group Self-Consistent-Field Method. *J. Chem. Theory Comput.* **2019**, *15*, 6724–6737.
- (297) Baiardo, A.; Reiher, M. Transcorrelated Density Matrix Renormalization Group, 2020.
- (298) Lischka, H.; Dallos, M.; Szalay, P. G.; Yarkony, D. R.; Shepard, R. Analytic Evaluation of Nonadiabatic coupling Terms at the MR-CI Level. I. Formalism. *J. Chem. Phys.* **2004**, *120*, 7322–7329.
- (299) Lischka, H.; et al. The Generality of the GUGA MRCI Approach in COLUMBUS for Treating Complex Quantum Chemistry. *J. Chem. Phys.* **2020**, *152*, 134110.
- (300) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-Order Perturbation Theory with a CASSCF Reference Function. *J. Phys. Chem.* **1990**, *94*, 5483–5488.
- (301) Andersson, K.; Malmqvist, P.; Roos, B. O. Second-Order Perturbation Theory with a Complete Active Space Self-Consistent Field Reference Function. *J. Chem. Phys.* **1992**, *96*, 1218–1226.
- (302) Finley, J.; Malmqvist, P.-A.; Roos, B. O.; Serrano-Andrés, L. The Multi-State {CASPT2} Method. *Chem. Phys. Lett.* **1998**, *288*, 299–306.
- (303) Angeli, C.; Cimraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of N-Electron Valence States for Multireference Perturbation Theory. *J. Chem. Phys.* **2001**, *114*, 10252–10264.
- (304) Roemelt, M.; Guo, S.; Chan, G. K.-L. A Projected Approximation to Strongly Contracted N-Electron Valence Perturbation Theory for DMRG Wavefunctions. *J. Chem. Phys.* **2016**, *144*, 204113.
- (305) Guo, Y.; Sivalingam, K.; Valeev, E. F.; Neese, F. Explicitly Correlated N-Electron Valence State Perturbation Theory (NEVPT2-F12). *J. Chem. Phys.* **2017**, *147*, 064110.
- (306) Maitra, R.; Sinha, D.; Mukherjee, D. Unitary Group Adapted State-Specific Multi-Reference Coupled Cluster Theory: Formulation and Pilot Numerical Applications. *J. Chem. Phys.* **2012**, *137*, 024105.

- (307) Mășik, J.; Hubač, I. In *Multireference Brillouin-Wigner Coupled-Cluster Theory. Single-Root Approach*; Sabin, J. R., Zerner, M. C., Brändas, E., Wilson, S., Maruani, J., Smeyers, Y., Grout, P., McWeeny, R., Eds.; *Advances in Quantum Chemistry*; Academic Press, 1998; Vol. 31; pp 75–104.
- (308) Musiał, M.; Perera, A.; Bartlett, R. J. Multireference Coupled-Cluster Theory: The Easy Way. *J. Chem. Phys.* **2011**, *134*, 114108.
- (309) Evangelista, F. A. Perspective: Multireference Coupled Cluster Theories of Dynamical Electron Correlation. *J. Chem. Phys.* **2018**, *149*, 030901.
- (310) Fdez. Galván, I.; et al. OpenMolcas: From Source Code to Insight. *J. Chem. Theory Comput.* **2019**, *15*, 5925–5964.
- (311) Roos, B.; Lindh, R.; Malmqvist, P.-Å.; Veryazov, V.; Widmark, P.-O. Main Group Atoms and Dimers Studied with a new Relativistic ANO Basis Set. *J. Phys. Chem. A* **2004**, *108*, 2851–2858.
- (312) Vogiatzis, K. D.; Ma, D.; Olsen, J.; Gagliardi, L.; de Jong, W. A. Pushing Configuration-Interaction to the limit: Towards Massively Parallel MCSCF Calculations. *J. Chem. Phys.* **2017**, *147*, 184111.
- (313) Kato, H.; Baba, M. Dynamics of Excited Molecules: Predissociation. *Chem. Rev.* **1995**, *95*, 2311–2349.
- (314) Merer, A. J.; Mulliken, R. S. Ultraviolet Spectra and Excited States of Ethylene and its Alkyl Derivatives. *Chem. Rev.* **1969**, *69*, 639–656.
- (315) Ashfold, M. N. R.; Langford, S. R. In *The Role of Rydberg States in Spectroscopy and Photochemistry: Low and High Rydberg States*; Sándorfy, C., Ed.; Springer Netherlands: Dordrecht, 1999; pp 23–56.
- (316) Merkt, F. Molecules in High Rydberg States. *Annu. Rev. Phys. Chem.* **1997**, *48*, 675–709.
- (317) Stein, C. J.; Reiher, M. Automated Selection of Active Orbital Spaces. *J. Chem. Theory Comput.* **2016**, *12*, 1760–1771.
- (318) Stein, C. J.; Reiher, M. Measuring Multi-Configurational Character by Orbital Entanglement. *Mol. Phys.* **2017**, *115*, 2110–2119.
- (319) Stein, C. J.; Reiher, M. Automated Identification of Relevant Frontier Orbitals for Chemical Compounds and Processes. *Chimia* **2017**, *71*, 170–176.
- (320) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871.
- (321) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.
- (322) Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52*, 997–1000.
- (323) Zangwill, A.; Soven, P. Density-Functional Approach to Local-Field Effects in Finite Systems: Photoabsorption in the Rare Gases. *Phys. Rev. A: At., Mol., Opt. Phys.* **1980**, *21*, 1561–1572.
- (324) Chong, D. P. *Recent Advances in Density Functional Methods*; World Scientific, 1995.
- (325) Tamm, I. Relativistic Interaction of Elementary Particles. *J. Phys. (Moscow)* **1945**, *9*, 449.
- (326) Dancoff, S. M. Non-Adiabatic Meson Theory of Nuclear Forces. *Phys. Rev.* **1950**, *78*, 382–385.
- (327) Hirata, S.; Head-Gordon, M. Time-Dependent Density Functional Theory within the Tamm–Dancoff Approximation. *Chem. Phys. Lett.* **1999**, *314*, 291–299.
- (328) Casida, M. E. Time-Dependent Density-Functional Theory for Molecules and Molecular Solids. *J. Mol. Struct.: THEOCHEM* **2009**, *914*, 3–18.
- (329) Cordova, F.; Doriol, L. J.; Ipatov, A.; Casida, M. E.; Filippi, C.; Vela, A. Troubleshooting Time-Dependent Density-Functional Theory for Photochemical Applications: Oxirane. *J. Chem. Phys.* **2007**, *127*, 164111.
- (330) Goerigk, L.; Casanova-Paéz, M. The Trip to the Density Functional Theory Zoo Continues: Making a Case for Time-Dependent Double Hybrids for Excited-State Problems. *Aust. J. Chem.* **2020**, DOI: 10.1071/CH20093.
- (331) Worth, G. A.; Cederbaum, L. S. Beyond Born-Oppenheimer: Molecular Dynamics Through a conical Intersection. *Annu. Rev. Phys. Chem.* **2004**, *55*, 127–158.
- (332) Doltsinis, N. L. *Molecular Dynamics Beyond the Born-Oppenheimer Approximation: Mixed Quantum-Classical Approaches*; NIC Series; John von Neuman Institut for Computing, 2006; Vol. 31; pp 389–409.
- (333) Levine, B. G.; Ko, C.; Quenneville, J.; Martínez, T. J. Conical Intersections and Double Excitations in Time-Dependent Density Functional Theory. *Mol. Phys.* **2006**, *104*, 1039–1051.
- (334) Levine, B. G.; Martínez, T. J. Isomerization Through Conical Intersections. *Annu. Rev. Phys. Chem.* **2007**, *58*, 613–634.
- (335) Tapavicza, E.; Tavernelli, I.; Rothlisberger, U.; Filippi, C.; Casida, M. E. Mixed Time-Dependent Density-Functional Theory/Classical Trajectory Surface Hopping Study of Oxirane Photochemistry. *J. Chem. Phys.* **2008**, *129*, 124108.
- (336) Jacquemin, D.; Adamo, C. In *Density-Functional Methods for Excited States*; Ferré, N., Filatov, M., Huix-Rotllant, M., Eds.; Springer International Publishing: Cham, 2016; pp 347–375.
- (337) Li, S. L.; Marenich, A. V.; Xu, X.; Truhlar, D. G. Configuration Interaction-Corrected Tamm–Dancoff Approximation: A Time-Dependent Density Functional Method with the Correct Dimensionality of Conical Intersections. *J. Phys. Chem. Lett.* **2014**, *5*, 322–328.
- (338) Bannwarth, C.; Yu, J. K.; Hohenstein, E. G.; Martínez, T. J. Hole-Hole Tamm-Dancoff-Approximated Density Functional Theory: A Highly Efficient Electronic Structure Method Incorporating Dynamic and Static Correlation. *J. Chem. Phys.* **2020**, *153*, 024110.
- (339) Lee, I. S.; Filatov, M.; Min, S. K. Formulation and Implementation of the Spin-Restricted Ensemble-Referenced Kohn–Sham Method in the Context of the Density Functional Tight Binding Approach. *J. Chem. Theory Comput.* **2019**, *15*, 3021–3032.
- (340) Maitra, N. T.; Zhang, F.; Cave, R. J.; Burke, K. Double Excitations within Time-Dependent Density Functional Theory Linear Response. *J. Chem. Phys.* **2004**, *120*, 5932–5937.
- (341) Elliott, P.; Goldson, S.; Canahui, C.; Maitra, N. T. Perspectives on Double-Excitations in TDDFT. *Chem. Phys.* **2011**, *391*, 110–119.
- (342) Katriel, J.; Zahariev, F.; Burke, K. Symmetry and Degeneracy in Density Functional Theory. *Int. J. Quantum Chem.* **2001**, *85*, 432–435.
- (343) Shao, Y.; Head-Gordon, M.; Krylov, A. I. The Spin–Flip Approach within Time-Dependent Density Functional Theory: Theory and Applications to Diradicals. *J. Chem. Phys.* **2003**, *118*, 4807–4818.
- (344) Lee, S.; Shostak, S.; Filatov, M.; Choi, C. H. Conical Intersections in Organic Molecules: Benchmarking Mixed-Reference Spin–Flip Time-Dependent DFT (MRSF-TD-DFT) vs Spin–Flip TD-DFT. *J. Phys. Chem. A* **2019**, *123*, 6455–6462.
- (345) Gavnholt, J.; Olsen, T.; Engelund, M.; Schiøtz, J.  $\Delta$  Self-Consistent Field Method to obtain Potential Energy Surfaces of Excited Molecules on Surfaces. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2008**, *78*, 075441.
- (346) Maurer, R. J.; Reuter, K. Assessing Computationally Efficient Isomerization Dynamics:  $\Delta$ -SCF Density-Functional Theory Study of Azobenzene Molecular Switching. *J. Chem. Phys.* **2011**, *135*, 224303.
- (347) Maurer, R. J.; Reuter, K. Excited-State Potential-Energy Surfaces of Metal-Adsorbed Organic Molecules from Linear Expansion  $\Delta$ -Self-Consistent Field Density-Functional Theory ( $\Delta$ SCF-DFT). *J. Chem. Phys.* **2013**, *139*, 014708.
- (348) Ghosh, S.; Verma, P.; Cramer, C. J.; Gagliardi, L.; Truhlar, D. G. Combining Wave Function Methods with Density Functional Theory for Excited States. *Chem. Rev.* **2018**, *118*, 7249–7292.
- (349) Wang, X.; Wong, L.; Hu, L.; Chan, C.; Su, Z.; Chen, G. Improving the Accuracy of Density-Functional Theory Calculation: The Statistical Correction Approach. *J. Phys. Chem. A* **2004**, *108*, 8514–8525.
- (350) Li, H.; Shi, L.; Zhang, M.; Su, Z.; Wang, X.; Hu, L.; Chen, G. Improving the Accuracy of Density-Functional Theory Calculation:



- (393) Mai, S.; Richter, M.; Marquetand, P.; González, L. *Excitation of Nucleobases from a Computational Perspective II: Dynamics*, 2014.
- (394) Liu, W. *Essentials of Relativistic Quantum Chemistry*. *J. Chem. Phys.* **2020**, *152*, 180901.
- (395) Horton, S. L.; Liu, Y.; Forbes, R.; Makhija, V.; Lausten, R.; Stolow, A.; Hockett, P.; Marquetand, P.; Rozgonyi, T.; Weinacht, T. Excited state dynamics of CH<sub>2</sub>I<sub>2</sub> and CH<sub>2</sub>BrI studied with UV pump VUV probe photoelectron spectroscopy. *J. Chem. Phys.* **2019**, *150*, 174201.
- (396) Horton, S. L.; Liu, Y.; Chakraborty, P.; Marquetand, P.; Rozgonyi, T.; Matsika, S.; Weinacht, T. Strong-Field- Versus Weak-Field-Ionization Pump-Probe Spectroscopy. *Phys. Rev. A: At, Mol, Opt. Phys.* **2018**, *98*, 053416.
- (397) Sussman, B. J.; Townsend, D.; Ivanov, M. Y.; Stolow, A. Dynamic Stark Control of Photochemical Processes. *Science* **2006**, *314*, 278–281.
- (398) Marquetand, P.; Richter, M.; González-Vázquez, J.; Sola, I.; González, L. Nonadiabatic *Ab Initio* Molecular Dynamics Including Spin-Orbit Coupling and Laser Fields. *Faraday Discuss.* **2011**, *153*, 261–273.
- (399) Bajo, J. J.; González-Vázquez, J.; Sola, I.; Santamaria, J.; Richter, M.; Marquetand, P.; González, L. Mixed Quantum-Classical Dynamics in the Adiabatic Representation to Simulate Molecules Driven by Strong Laser Pulses. *J. Phys. Chem. A* **2012**, *116*, 2800–2807.
- (400) Köuppel, H.; Domcke, W.; Cederbaum, L. S. Multimode Molecular Dynamics Beyond the Born-Oppenheimer Approximation. *Adv. Chem. Phys.* **2007**, *57*, 59–246.
- (401) Ben-Nun, M.; Martínez, T. J. *Advances in Chemical Physics*; John Wiley & Sons, Ltd, 2002; pp 439–512.
- (402) Guo, H.; Yarkony, D. R. Accurate Nonadiabatic Dynamics. *Phys. Chem. Chem. Phys.* **2016**, *18*, 26335–26352.
- (403) Beck, M.; Jäckle, A.; Worth, G.; Meyer, H.-D. The Multiconfiguration Time-Dependent Hartree (MCTDH) Method: A Highly Efficient Algorithm for Propagating Wavepackets. *Phys. Rep.* **2000**, *324*, 1–105.
- (404) Yeager, D. L.; Jørgensen, P. A Multiconfigurational Time-Dependent Hartree-Fock Approach. *Chem. Phys. Lett.* **1979**, *65*, 77–80.
- (405) Manthe, U. Wavepacket Dynamics and the Multi-Configurational Time-Dependent Hartree Approach. *J. Phys.: Condens. Matter* **2017**, *29*, 253001.
- (406) Eng, J.; Gourlaouen, C.; Gindensperger, E.; Daniel, C. Spin-Vibronic Quantum Dynamics for Ultrafast Excited-State Processes. *Acc. Chem. Res.* **2015**, *48*, 809–817.
- (407) Gómez, S.; Heindl, M.; Szabadi, A.; González, L. 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* **2019**, *123*, 8321–8332.
- (408) Ischtwan, J.; Collins, M. A. Molecular Potential Energy Surfaces by Interpolation. *J. Chem. Phys.* **1994**, *100*, 8080–8088.
- (409) Evenhuis, C. R.; Collins, M. A. Interpolation of Diabatic Potential Energy Surfaces. *J. Chem. Phys.* **2004**, *121*, 2515–2527.
- (410) Evenhuis, C.; Martínez, T. J. A Scheme to Interpolate Potential Energy Surfaces and Derivative Coupling Vectors without Performing a Global Diabatization. *J. Chem. Phys.* **2011**, *135*, 224110.
- (411) Mukherjee, S.; Bandyopadhyay, S.; Paul, A. K.; Adhikari, S. Construction of Diabatic Hamiltonian Matrix from *Ab Initio* Calculated Molecular Symmetry Adapted Nonadiabatic Coupling Terms and Nuclear Dynamics for the Excited States of Na<sub>3</sub> Cluster. *J. Phys. Chem. A* **2013**, *117*, 3475–3495.
- (412) Li, J.; Jiang, B.; Guo, H. Permutation Invariant Polynomial Neural Network Approach to Fitting Potential Energy Surfaces. II. Four-Atom Systems. *J. Chem. Phys.* **2013**, *139*, 204103.
- (413) Jiang, B.; Guo, H. Permutation Invariant Polynomial Neural Network Approach to Fitting Potential Energy Surfaces. *J. Chem. Phys.* **2013**, *139*, 054112.
- (414) Jiang, B.; Guo, H. Permutation Invariant Polynomial Neural Network Approach to Fitting Potential Energy Surfaces. III. Molecule-Surface Interactions. *J. Chem. Phys.* **2014**, *141*, 034109.
- (415) Worth, G.; Robb, M.; Lasorne, B. Solving the Time-Dependent Schrödinger Equation for Nuclear Motion in One Step: Direct Dynamics of Non-Adiabatic Systems. *Mol. Phys.* **2008**, *106*, 2077–2091.
- (416) Persico, M.; Granucci, G. An Overview of Nonadiabatic Dynamics Simulations Methods, with Focus on the Direct Approach Versus the Fitting of Potential Energy Surfaces. *Theor. Chem. Acc.* **2014**, *133*, 1526.
- (417) Komarova, K. G.; Remacle, F.; Levine, R. On the Fly Quantum Dynamics of Electronic and Nuclear Wave Packets. *Chem. Phys. Lett.* **2018**, *699*, 155–161.
- (418) Lasorne, B.; Robb, M. A.; Worth, G. A. Direct Quantum Dynamics using Variational Multi-Configuration Gaussian Wavepackets. Implementation Details and Test Case. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3210–3227.
- (419) Ben-Nun, M.; Martínez, T. J. Photodynamics of Ethylene: *Ab Initio* Studies of Conical Intersections. *Chem. Phys.* **2000**, *259*, 237–248.
- (420) Curchod, B. F. E.; Rauer, C.; Marquetand, P.; González, L.; Martínez, T. J. Communication: GAIMS—Generalized *Ab Initio* Multiple Spawning for Both Internal Conversion and Intersystem Crossing Processes. *J. Chem. Phys.* **2016**, *144*, 101102.
- (421) Mignolet, B.; Curchod, B. F. E. A Walk Through the Approximations of *Ab Initio* Multiple Spawning. *J. Chem. Phys.* **2018**, *148*, 134110.
- (422) Freixas, V. M.; Fernandez-Alberti, S.; Makhov, D. V.; Tretiak, S.; Shalashilin, D. An *Ab Initio* Multiple Cloning Approach for the Simulation of Photoinduced Dynamics in Conjugated Molecules. *Phys. Chem. Chem. Phys.* **2018**, *20*, 17762–17772.
- (423) Begušić, T.; Patoz, A.; Šulc, M.; Vanček, J. On-the-Fly *Ab Initio* Three Thawed Gaussians Approximation: A Semiclassical Approach to Herzberg-Teller Spectra. *Chem. Phys.* **2018**, *515*, 152–163.
- (424) Markland, T.; Ceriotti, M. Nuclear Quantum Effects Enter the Mainstream. *Nat. Rev. Chem.* **2018**, *2*, 0109.
- (425) Miller, W. H. Classical S Matrix: Numerical Application to Inelastic Collisions. *J. Chem. Phys.* **1970**, *53*, 3578–3587.
- (426) Ceotto, M.; Atahan, S.; Shim, S.; Tantardini, G. F.; Aspuru-Guzik, A. First-Principles Semiclassical Initial Value Representation Molecular Dynamics. *Phys. Chem. Chem. Phys.* **2009**, *11*, 3861–3867.
- (427) Nakamura, H.; Nanbu, S.; Teranishi, Y.; Ohta, A. Development of Semiclassical Molecular Dynamics Simulation Method. *Phys. Chem. Chem. Phys.* **2016**, *18*, 11972–11985.
- (428) Gao, X.; Saller, M. A. C.; Liu, Y.; Kelly, A.; Richardson, J. O.; Geva, E. Benchmarking Quasiclassical Mapping Hamiltonian Methods for Simulating Electronically Nonadiabatic Molecular Dynamics. *J. Chem. Theory Comput.* **2020**, *16*, 2883–2895.
- (429) Ceriotti, M.; More, J.; Manolopoulos, D. E. i-PI A Python Interface for *Ab Initio* Path Integral Molecular Dynamics Simulations. *Comput. Phys. Commun.* **2014**, *185*, 1019–1026.
- (430) Kapil, V.; et al. i-PI 2.0: A Universal Force Engine for Advanced Molecular Simulations. *Comput. Phys. Commun.* **2019**, *236*, 214–223.
- (431) Thoss, M.; Miller, W. H.; Stock, G. Semiclassical Description of Nonadiabatic Quantum Dynamics: Application to the S<sub>1</sub>–S<sub>2</sub> Conical Intersection in Pyrazine. *J. Chem. Phys.* **2000**, *112*, 10282–10292.
- (432) Lee, M. K.; Huo, P.; Coker, D. F. Semiclassical Path Integral Dynamics: Photosynthetic Energy Transfer with Realistic Environment Interactions. *Annu. Rev. Phys. Chem.* **2016**, *67*, 639–668.
- (433) Stock, G.; Thoss, M. Semiclassical Description of Non-adiabatic Quantum Dynamics. *Phys. Rev. Lett.* **1997**, *78*, 578–581.
- (434) Weinreich, J.; Römer, A.; Paleico, M. L.; Behler, J. Properties of  $\alpha$ -Brass Nanoparticles. 1. Neural Network Potential Energy Surface. *J. Phys. Chem. C* **2020**, *124*, 12682.

- (435) Lin, Q.; Zhang, Y.; Zhao, B.; Jiang, B. Automatically Growing Global Reactive Neural Network Potential Energy Surfaces: A Trajectory-Free Active Learning Strategy. *J. Chem. Phys.* **2020**, *152*, 154104.
- (436) Tully, J. C. Molecular Dynamics with Electronic Transitions. *J. Chem. Phys.* **1990**, *93*, 1061–1071.
- (437) Tully, J. C. Nonadiabatic Molecular Dynamics. *Int. J. Quantum Chem.* **1991**, *40*, 299–309.
- (438) Tully, C. J. Mixed Quantum–Classical Dynamics. *Faraday Discuss.* **1998**, *110*, 407–419.
- (439) S. Mai, L. G., Marquetand, P. In *Quantum Chemistry and Dynamics of Excited States: Methods and Applications*; González, L., Lindh, R., Eds.; Wiley, 2020, in press.
- (440) Zener, C. Non-Adiabatic Crossing of Energy Levels. *Proc. R. Soc. London A* **1932**, *137*, 696–702.
- (441) Wittig, C. The Landau-Zener Formula. *J. Phys. Chem. B* **2005**, *109*, 8428–8430.
- (442) Zhu, C.; Kamisaka, H.; Nakamura, H. Significant Improvement of the Trajectory Surface Hopping Method by the Zhu–Nakamura Theory. *J. Chem. Phys.* **2001**, *115*, 11036–11039.
- (443) Zhu, C.; Kamisaka, H.; Nakamura, H. 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<sub>2</sub><sup>+</sup> System. *J. Chem. Phys.* **2002**, *116*, 3234–3247.
- (444) Oloyede, P.; Mil'nikov, G.; Nakamura, H. Generalized Trajectory Surface Hopping Method Based on the Zhu–Nakamura Theory. *J. Chem. Phys.* **2006**, *124*, 144110.
- (445) Ishida, T.; Nanbu, S.; Nakamura, H. 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.* **2017**, *36*, 229–286.
- (446) Heberle, A. P.; Baumberg, J. J.; Kohler, K. Ultrafast coherent control and Destruction of Excitons in Quantum Wells. *Phys. Rev. Lett.* **1995**, *75*, 2598–2601.
- (447) Granucci, G.; Persico, M. Critical Appraisal of the Fewest Switching Algorithm for Surface Hopping. *J. Chem. Phys.* **2007**, *126*, 134114.
- (448) Fabiano, E.; Keal, T.; Thiel, W. Implementation of Surface Hopping Molecular Dynamics using Semiempirical Methods. *Chem. Phys.* **2008**, *349*, 334–347.
- (449) Malhado, J. P.; Bearpark, M. J.; Hynes, J. T. Non-Adiabatic Dynamics Close to Conical Intersections and the Surface Hopping Perspective. *Front. Chem.* **2014**, *2*, 97.
- (450) Wang, L.; Akimov, A.; Prezhdo, O. V. Recent Progress in Surface Hopping: 2011–2015. *J. Phys. Chem. Lett.* **2016**, *7*, 2100–2112.
- (451) Subotnik, J. E.; Rhee, Y. M. On Surface Hopping and Time-Reversal. *J. Phys. Chem. A* **2015**, *119*, 990–995.
- (452) Hammes-Schiffer, S.; Tully, J. C. Proton Transfer in Solution: Molecular Dynamics with Quantum Transitions. *J. Chem. Phys.* **1994**, *101*, 4657–4667.
- (453) Sawada, S.-I.; Nitzan, A.; Metiu, H. Mean-Trajectory Approximation for Charge- and Energy-Transfer Processes at Surfaces. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1985**, *32*, 851–867.
- (454) Li, X.; Tully, J. C.; Schlegel, H. B.; Frisch, M. J. *Ab Initio* Ehrenfest Dynamics. *J. Chem. Phys.* **2005**, *123*, 084106.
- (455) Mai, S.; Marquetand, P.; González, L. Intersystem Crossing Pathways in the Noncanonical Nucleobase 2-Thiouracil: A Time-Dependent Picture. *J. Phys. Chem. Lett.* **2016**, *7*, 1978–1983.
- (456) Mai, S.; Richter, M.; Marquetand, P.; González, L. The DNA Nucleobase Thymine in Motion – Intersystem Crossing Simulated with Surface Hopping. *Chem. Phys.* **2017**, *482*, 9–15.
- (457) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Quantum Chemistry Structures and Properties of 134 Kilo Molecules. *Sci. Data* **2014**, *1*, 140022.
- (458) Artrith, N.; Morawietz, T.; Behler, J. High-Dimensional Neural-Network Potentials for Multicomponent Systems: Applications to Zinc Oxide. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2011**, *83*, 153101.
- (459) Huang, B.; von Lilienfeld, O. A. Communication: Understanding Molecular Representations in Machine Learning: The Role of Uniqueness and Target Similarity. *J. Chem. Phys.* **2016**, *145*, 161102.
- (460) Yao, K.; Herr, J. E.; Toth, D. W.; Mckintyre, R.; Parkhill, J. The TensorMol-0.1 Model Chemistry: A Neural Network Augmented with Long-Range Physics. *Chem. Sci.* **2018**, *9*, 2261–2269.
- (461) Schütt, K. T.; Saucedo, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. SchNet – A Deep Learning Architecture for Molecules and Materials. *J. Chem. Phys.* **2018**, *148*, 241722.
- (462) Nebgen, B.; Lubbers, N.; Smith, J. S.; Sifain, A. E.; Lohkov, A.; Isayev, O.; Roitberg, A. E.; Barros, K.; Tretiak, S. Transferable Dynamic Molecular Charge Assignment Using Deep Neural Networks. *J. Chem. Theory Comput.* **2018**, *14*, 4687–4698.
- (463) Sifain, A. E.; Lubbers, N.; Nebgen, B. T.; Smith, J. S.; Lohkov, A. Y.; Isayev, O.; Roitberg, A. E.; Barros, K.; Tretiak, S. Discovering a Transferable Charge Assignment Model Using Machine Learning. *J. Phys. Chem. Lett.* **2018**, *9*, 4495–4501.
- (464) Schütt, K. T.; Gastegger, M.; Tkatchenko, A.; Müller, K.-R. *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*; Springer International Publishing, 2019; pp 311–330.
- (465) Schütt, K. T.; Kessel, P.; Gastegger, M.; Nicoli, K. A.; Tkatchenko, A.; Müller, K.-R. SchNetPack: A Deep Learning Toolbox For Atomistic Systems. *J. Chem. Theory Comput.* **2019**, *15*, 448–455.
- (466) Christensen, A. S.; Faber, F. A.; von Lilienfeld, O. A. Operators in Quantum Machine Learning: Response Properties in Chemical Space. *J. Chem. Phys.* **2019**, *150*, 064105.
- (467) Veit, M.; Wilkins, D. M.; Yang, Y.; DiStasio, R. A.; Ceriotti, M. Predicting Molecular Dipole Moments by Combining Atomic Partial Charges and Atomic Dipoles. *J. Chem. Phys.* **2020**, *153*, 024113.
- (468) Thomas, M.; Brehm, M.; Fligg, R.; Vöhringer, P.; Kirchner, B. Computing Vibrational Spectra from *Ab Initio* Molecular Dynamics. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6608–6622.
- (469) Wilke, J.; Wilke, M.; Meerts, W. L.; Schmitt, M. Determination of Ground and Excited State Dipole Moments via Electronic Stark Spectroscopy: 5-Methoxyindole. *J. Chem. Phys.* **2016**, *144*, 044201.
- (470) Tennyson, J. Perspective: Accurate Ro-Vibrational Calculations on Small Molecules. *J. Chem. Phys.* **2016**, *145*, 120901.
- (471) Marquetand, P.; Nogueira, J.; Mai, S.; Plasser, F.; González, L. Challenges in Simulating Light-Induced Processes in DNA. *Molecules* **2017**, *22*, 49.
- (472) Nogueira, J. J.; González, L. Computational Photophysics in the Presence of an Environment. *Annu. Rev. Phys. Chem.* **2018**, *69*, 473–497.
- (473) Barbatti, M.; Granucci, G.; Persico, M.; Ruckebauer, M.; Vazdar, M.; Eckert-Maksić, M.; Lischka, H. The on-the-Fly Surface-Hopping Program System Newton-X: Application to *Ab Initio* Simulation of the Nonadiabatic Photodynamics of Benchmark Systems. *J. Photochem. Photobiol., A* **2007**, *190*, 228–240.
- (474) Norman, P.; Dreuw, A. Simulating X-Ray Spectroscopies and Calculating Core-Excited States of Molecules. *Chem. Rev.* **2018**, *118*, 7208–7248.
- (475) Wigner, E. On The Quantum Correction for Thermodynamic Equilibrium. *Phys. Rev.* **1932**, *40*, 749–750.
- (476) Thorne, A. P. *Spectrophysics*; Springer: Dordrecht, 1983.
- (477) Haghghatlatari, M.; Li, J.; Heidar-Zadeh, F.; Liu, Y.; Guan, X.; Head-Gordon, T. Learning to Make Chemical Predictions: The Interplay of Feature Representation, Data, and Machine Learning Methods. *Chem.* **2020**, *6*, 1527.
- (478) Bishop, C. M. *Pattern Recognition and Machine Learning*, 1st ed.; Springer: New York, 2006.
- (479) Halama, N. Machine Learning for Tissue Diagnostics in Oncology: Brave New World. *Br. J. Cancer* **2019**, *121*, 431–433.
- (480) Bychkov, D.; Linder, N.; Turkki, R.; Nordling, S.; Kovanen, P. E.; Verrill, C.; Walliander, M.; Lundin, M.; Haglund, C.; Lundin, J.



Deep Learning Based Tissue Analysis Predicts Outcome in Colorectal Cancer. *Sci. Rep.* **2018**, *8*, 3395.

(481) Gómez-Meiré, S.; Campos, C.; Falqué, E.; Díaz, F.; Fdez-Riverola, F. Assuring the Authenticity of Northwest Spain White Wine Varieties using Machine Learning Techniques. *Food Res. Int.* **2014**, *60*, 230–240.

(482) Watanabe, N.; Murata, M.; Ogawa, T.; Vavricka, C. J.; Kondo, A.; Ogino, C.; Araki, M. Exploration and Evaluation of Machine Learning-Based Models for Predicting Enzymatic Reactions. *J. Chem. Inf. Model.* **2020**, *60*, 1833–1843.

(483) Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; New York, NY, USA, 2016; pp 785–794.

(484) Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst* **1989**, *2*, 303–314.

(485) Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Networks* **1991**, *4*, 251–257.

(486) Hofmann, T.; Schölkopf, B.; Smola, A. J. Kernel Methods in Machine Learning. *Ann. Statist.* **2008**, *36*, 1171–1220.

(487) Raschka, S.; Mirjalili, V. *Python Machine Learning*, 3rd ed.; Packt Publishing, 2019.

(488) Hansen, K.; Montavon, G.; Biegler, F.; Fazli, S.; Rupp, M.; Scheffler, M.; von Lilienfeld, O. A.; Tkatchenko, A.; Müller, K.-R. Assessment and Validation of Machine Learning Methods for Predicting Molecular Atomization Energies. *J. Chem. Theory Comput.* **2013**, *9*, 3404–3419.

(489) Rupp, M. Machine learning for quantum mechanics in a nutshell. *Int. J. Quantum Chem.* **2015**, *115*, 1058–1073.

(490) Xue, B.-X.; Barbatti, M.; Dral, P. O. Machine Learning for Absorption Cross Sections. *J. Phys. Chem. A* **2020**, *124*, 7199–7210.

(491) Ramakrishnan, R.; von Lilienfeld, O. A. *Reviews in Computational Chemistry*; John Wiley & Sons, Ltd, 2017; Chapter 5, pp 225–256.

(492) Bartók, A. P.; Kondor, R.; Csányi, G. On Representing Chemical Environments. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2013**, *87*, 184115.

(493) Veit, M.; Jain, S. K.; Bonakala, S.; Rudra, I.; Hohl, D.; Csányi, G. Equation of State of Fluid Methane from First Principles with Machine Learning Potentials. *J. Chem. Theory Comput.* **2019**, *15*, 2574–2586.

(494) Glorot, X.; Bengio, Y. Understanding the Difficulty of Training Deep Feedforward Neural Networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*; Chia Laguna Resort: Sardinia, Italy, 2010; pp 249–256.

(495) Duchi, J.; Hazan, E.; Singer, Y. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *J. Mach. Learn. Res.* **2011**, *12*, 2121–2159.

(496) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, abs/1412.6980, 1412.6980.

(497) Puskorius, G. V.; Feldkamp, L. A. Decoupled extended Kalman filter training of feedforward layered networks. *IJCNN-91-Seattle International Joint Conference on Neural Networks*, 1991; pp 771–777.

(498) Singraber, A.; Morawietz, T.; Behler, J.; Dellago, C. Parallel Multistream Training of High-Dimensional Neural Network Potentials. *J. Chem. Theory Comput.* **2019**, *15*, 3075–3092.

(499) Srinivas, N.; Krause, A.; Kakade, S.; Seeger, M. Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design. *Proceedings of the 27th International Conference on Machine Learning*; Madison, WI, USA, 2010; pp 1015–1022.

(500) Wu, J.; Chen, X.-Y.; Zhang, H.; Xiong, L.-D.; Lei, H.; Deng, S.-H. Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization. *J. Electron. Sci. Technol.* **2019**, *17*, 26–40.

(501) Perrone, V.; Shen, H.; Seeger, M.; Archambeau, C.; Jenatton, R. *Learning Search Spaces for Bayesian Optimization: Another View of Hyperparameter Transfer Learning*, 2019.

(502) Stuke, A.; Rinke, P.; Todorović, M. *Efficient Hyperparameter Tuning for Kernel Ridge Regression with Bayesian Optimization*, 2020.

(503) Gilmer, J.; Schoenholz, S. S.; Riley, P. F.; Vinyals, O.; Dahl, G. E. Neural Message Passing for Quantum Chemistry. *Proceedings of the 34th International Conference on Machine Learning*, 2017; pp 1263–1272, Volume 70.

(504) Unke, O. T.; Meuwly, M. PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges. *J. Chem. Theory Comput.* **2019**, *15*, 3678–3693.

(505) Lubbers, N.; Smith, J. S.; Barros, K. Hierarchical Modeling of Molecular Energies using a Deep Neural Network. *J. Chem. Phys.* **2018**, *148*, 241715.

(506) Wang, H.; Zhang, L.; Han, J.; E, W. DeePMD-kit: A Deep Learning Package for Many-Body Potential Energy Representation and Molecular Dynamics. *Comput. Phys. Commun.* **2018**, *228*, 178–184.

(507) Schütt, K. T.; Arbabzadah, F.; Chmiela, S.; Müller, K. R.; Tkatchenko, A. Quantum-Chemical Insights from Deep Tensor Neural Networks. *Nat. Commun.* **2017**, *8*, 13890.

(508) Zhang, Y.; Hu, C.; Jiang, B. Embedded Atom Neural Network Potentials: Efficient and Accurate Machine Learning with a Physically Inspired Representation. *J. Phys. Chem. Lett.* **2019**, *10*, 4962–4967.

(509) Behler, J. Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials. *J. Chem. Phys.* **2011**, *134*, 074106.

(510) Ye, S.; Hu, W.; Li, X.; Zhang, J.; Zhong, K.; Zhang, G.; Luo, Y.; Mukamel, S.; Jiang, J. A Neural Network Protocol for Electronic Excitations of N-Methylacetamide. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 11612–11617.

(511) LeCun, Y.; Bengio, Y. *The Handbook of Brain Theory and Neural Networks*; The MIT Press: Cambridge, MA, USA, 1995; pp 255–257.

(512) Krizhevsky, A.; Sutskever, I.; Hinton, G. E. *ImageNet Classification with Deep Convolutional Neural Networks*, 2012, 1097–1105.

(513) Sainath, T. N.; Kingsbury, B.; Saon, G.; Soltau, H.; Mohamed, A.-r.; Dahl, G.; Ramabhadran, B. Deep Convolutional Neural Networks for Large-scale Speech Tasks. *Neural Networks* **2015**, *64*, 39–48.

(514) Schütt, K. *Learning Representations of Atomistic Systems with Deep Neural Networks*. Doctoral Thesis, Technische Universität Berlin, Berlin, 2018.

(515) Behler, J. Constructing High-Dimensional Neural Network Potentials: A Tutorial Review. *Int. J. Quantum Chem.* **2015**, *115*, 1032–1050.

(516) Dral, P. O. MLatom: A Program Package for Quantum Chemical Research Assisted by Machine Learning. *J. Comput. Chem.* **2019**, *40*, 2339–2347.

(517) Dral, P. O.; Owens, A.; Yurchenko, S. N.; Thiel, W. Structure-Based Sampling and Self-Correcting Machine Learning for Accurate Calculations of Potential Energy Surfaces and Vibrational Levels. *J. Chem. Phys.* **2017**, *146*, 244108.

(518) Christensen, A.; Faber, F.; Huang, B.; Bratholm, L.; Tkatchenko, A.; Müller, K.; Lilienfeld, O. *QML: A Python Toolkit for Quantum Machine Learning*, <https://github.com/qmlcode/qml>, 2017.

(519) Hansen, K.; Biegler, F.; Ramakrishnan, R.; Pronobis, W.; von Lilienfeld, O. A.; Müller, K.-R.; Tkatchenko, A. Machine Learning Predictions of Molecular Properties: Accurate Many-Body Potentials and Nonlocality in Chemical Space. *J. Phys. Chem. Lett.* **2015**, *6*, 2326–2331.

(520) Pozdnyakov, S. N.; Willatt, M. J.; Bartók, A. P.; Ortner, C.; Csányi, G.; Ceriotti, M. Incompleteness of Atomic Structure Representations. *Phys. Rev. Lett.* **2020**, *125*, 166001.

(521) Çaylak, O.; von Lilienfeld, O. A.; Baumeier, B. Wasserstein Metric for Improved Quantum Machine Learning with Adjacency Matrix Representations. *Mach. Learn.: Sci. Technol.* **2020**, *1*, 03LT01.

- (522) Braams, B. J.; Bowman, J. M. Permutationally Invariant Potential Energy Surfaces in High Dimensionality. *Int. Rev. Phys. Chem.* **2009**, *28*, 577–606.
- (523) Bowman, J. M.; Czakó, G.; Fu, B. High-Dimensional *Ab Initio* Potential Energy Surfaces for Reaction Dynamics Calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 8094–8111.
- (524) Qu, C.; Yu, Q.; Bowman, J. M. Permutationally Invariant Potential Energy Surfaces. *Annu. Rev. Phys. Chem.* **2018**, *69*, 151–175.
- (525) Malbon, C. L.; Zhao, B.; Guo, H.; Yarkony, D. R. On the Nonadiabatic Collisional Quenching of OH(A) by H<sub>2</sub>: A Four Coupled Quasi-Diabatic State Description. *Phys. Chem. Chem. Phys.* **2020**, *22*, 13516–13527.
- (526) Gastegger, M.; Schwiedrzik, L.; Bittermann, M.; Berzsenyi, F.; Marquetand, P. wACSF – Weighted Atom-Centered Symmetry Functions as Descriptors in Machine Learning Potentials. *J. Chem. Phys.* **2018**, *148*, 241709.
- (527) Herr, J. E.; Koh, K.; Yao, K.; Parkhill, J. Compressing Physics with an Autoencoder: Creating an Atomic Species Representation to Improve Machine Learning Models in the Chemical Sciences. *J. Chem. Phys.* **2019**, *151*, 084103.
- (528) Shapeev, A. V. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Multiscale Model. Simul.* **2016**, *14*, 1153–1173.
- (529) Thompson, A.; Swiler, L.; Trott, C.; Foiles, S.; Tucker, G. Spectral Neighbor Analysis Method for Automated Generation of Quantum-Accurate Interatomic Potentials. *J. Comput. Phys.* **2015**, *285*, 316–330.
- (530) Christensen, A. S.; Bratholm, L. A.; Faber, F. A.; Anatole von Lilienfeld, O. FCHL Revisited: Faster and More Accurate Quantum Machine Learning. *J. Chem. Phys.* **2020**, *152*, 044107.
- (531) Zaverkin, V.; Kästner, J. Gaussian Moments as Physically Inspired Molecular Descriptors for Accurate and Scalable Machine Learning Potentials. *J. Chem. Theory Comput.* **2020**, *16*, 5410–5421.
- (532) Zhang, Y.; Hu, C.; Jiang, B. *Bridging the Efficiency Gap Between Machine Learned Potentials with ab initio Accuracy and Classical Force Fields*, 2020; <https://arxiv.org/pdf/2006.16482>.
- (533) Kang, B.; Seok, C.; Lee, J. Prediction of Molecular Electronic Transitions using Random Forests. *ChemRxiv* **2020**, DOI: 10.26434/chemrxiv.12482840.v1.
- (534) Richings, G. W.; Habershon, S. Direct Quantum Dynamics Using Grid-Based Wave Function Propagation and Machine-Learned Potential Energy Surfaces. *J. Chem. Theory Comput.* **2017**, *13*, 4012–4024.
- (535) Dral, P. O. Quantum Chemistry in the Age of Machine Learning. *J. Phys. Chem. Lett.* **2020**, *11*, 2336–2347.
- (536) Chen, W.-K.; Fang, W.-H.; Cui, G. Integrating Machine Learning with the Multilayer Energy-Based Fragment Method for Excited States of Large Systems. *J. Phys. Chem. Lett.* **2019**, *10*, 7836–7841.
- (537) Chen, W.-K.; Zhang, Y.; Jiang, B.; Fang, W.-H.; Cui, G. Efficient Construction of Excited-State Hessian Matrices with Machine Learning Accelerated Multilayer Energy-Based Fragment Method. *J. Phys. Chem. A* **2020**, *124*, 5684.
- (538) Behler, J.; Martonak, R.; Donadio, D.; Parrinello, M. Metadynamics Simulations of the High-Pressure Phases of Silicon Employing a High-Dimensional Neural Network Potential. *Phys. Rev. Lett.* **2008**, *100*, 185501.
- (539) Chen, M. S.; Zuehlsdorff, T. J.; Morawietz, T.; Isborn, C. M.; Markland, T. E. Exploiting Machine Learning to Efficiently Predict Multidimensional Optical Spectra in Complex Environments. *J. Phys. Chem. Lett.* **2020**, *11*, 7559–7568.
- (540) Koch, W.; Zhang, D. H. Communication: Separable Potential Energy Surfaces from Multiplicative Artificial Neural Networks. *J. Chem. Phys.* **2014**, *141*, 021101.
- (541) He, D.; Yuan, J.; Li, H.; Chen, M. Global Diabatic Potential Energy Surfaces and Quantum Dynamical Studies for the Li(2p) + H<sub>2</sub>(X<sup>1</sup>Σ<sub>g</sub><sup>+</sup>) → LiH(XI<sup>2+</sup>) + H Reaction. *Sci. Rep.* **2016**, *6*, 25083.
- (542) Guan, Y.; Fu, B.; Zhang, D. H. Construction of Diabatic Energy Surfaces for LiFH with Artificial Neural Networks. *J. Chem. Phys.* **2017**, *147*, 224307.
- (543) Wang, S.; Yang, Z.; Yuan, J.; Chen, M. New Diabatic Potential Energy Surfaces of the NaH<sub>2</sub> System and Dynamics Studies for the Na(3p) + H<sub>2</sub> → NaH + H Reaction. *Sci. Rep.* **2018**, *8*, 17960.
- (544) Yuan, J.; He, D.; Wang, S.; Chen, M.; Han, K. Diabatic Potential Energy Surfaces of MgH<sub>2</sub><sup>+</sup> and Dynamic Studies for the Mg<sup>+</sup>(3p) + H<sub>2</sub> → MgH<sup>+</sup> + H Reaction. *Phys. Chem. Chem. Phys.* **2018**, *20*, 6638–6647.
- (545) Yin, Z.; Guan, Y.; Fu, B.; Zhang, D. H. Two-State Diabatic Potential Energy Surfaces of CH<sub>2</sub> Based on Nonadiabatic Couplings with Neural Networks. *Phys. Chem. Chem. Phys.* **2019**, *21*, 20372–20383.
- (546) Akimov, A. V. A Simple Phase correction Makes a Big Difference in Nonadiabatic Molecular Dynamics. *J. Phys. Chem. Lett.* **2018**, *9*, 6096–6102.
- (547) Beard, E. J.; Sivaraman, G.; Vázquez-Mayagoitia, A.; Vishwanath, V.; Cole, J. M. Comparative Dataset of Experimental and Computational Attributes of UV/Vis Absorption Spectra. *Sci. Data* **2019**, *6*. DOI: 10.1038/s41597-019-0306-0
- (548) Thawani, A. R.; Griffiths, R.-R.; Jamasb, A.; Bourached, A.; Jones, P.; McCorkindale, W.; Aldrick, A.; Lee, A. The Photoswitch Dataset: A Molecular Machine Learning Benchmark for the Advancement of Synthetic Chemistry. *ChemRxiv* **2020**, DOI: 10.26434/chemrxiv.12609899.v1.
- (549) Barbatti, M.; Ruckebauer, M.; Lischka, H. The Photo-dynamics of Ethylene: A Surface-Hopping Study on Structural Aspects. *J. Chem. Phys.* **2005**, *122*, 174307.
- (550) Tavernelli, I.; Tapavicza, E.; Rothlisberger, U. Nonadiabatic Coupling Vectors within Linear Response Time-Dependent Density Functional Theory. *J. Chem. Phys.* **2009**, *130*, 124107.
- (551) Tavernelli, I.; Tapavicza, E.; Rothlisberger, U. Non-Adiabatic Dynamics using Time-Dependent Density Functional Theory: Assessing the Coupling Strengths. *J. Mol. Struct.: THEOCHEM* **2009**, *914*, 22–29.
- (552) Barbatti, M.; Aquino, A. J. A.; Lischka, H. Ultrafast Two-Step Process in the Non-Adiabatic Relaxation of the CH<sub>2</sub>NH<sub>2</sub> Molecule. *Mol. Phys.* **2006**, *104*, 1053–1060.
- (553) Tao, H.; Allison, T. K.; Wright, T. W.; Stooke, A. M.; Khurmi, C.; van Tilborg, J.; Liu, Y.; Falcone, R. W.; Belkacem, A.; Martínez, T. J. Ultrafast internal conversion in ethylene. I. The excited state lifetime. *J. Chem. Phys.* **2011**, *134*, 244306.
- (554) Allison, T. K.; Tao, H.; Glover, W. J.; Wright, T. W.; Stooke, A. M.; Khurmi, C.; van Tilborg, J.; Liu, Y.; Falcone, R. W.; Martínez, T. J.; Belkacem, A. Ultrafast internal conversion in ethylene. II. Mechanisms and pathways for quenching and hydrogen elimination. *J. Chem. Phys.* **2012**, *136*, 124317.
- (555) Mori, T.; Glover, W. J.; Schuurman, M. S.; Martínez, T. J. Role of Rydberg States in the Photochemical Dynamics of Ethylene. *J. Phys. Chem. A* **2012**, *116*, 2808–2818.
- (556) Sellner, B.; Barbatti, M.; Müller, T.; Domcke, W.; Lischka, H. Ultrafast Non-Adiabatic Dynamics of Ethylene including Rydberg States. *Mol. Phys.* **2013**, *111*, 2439–2450.
- (557) Barbatti, M.; Lan, Z.; Crespo-Otero, R.; Szymczak, J. J.; Lischka, H.; Thiel, W. Critical Appraisal of Excited State Nonadiabatic Dynamics Simulations of 9H-Adenine. *J. Chem. Phys.* **2012**, *137*, 22A503.
- (558) Hollas, D.; Šišťík, L.; Hohenstein, E. G.; Martínez, T. J.; Slaviček, P. Nonadiabatic *Ab Initio* Molecular Dynamics with the Floating Occupation Molecular Orbital-Complete Active Space Configuration Interaction Method. *J. Chem. Theory Comput.* **2018**, *14*, 339–350.
- (559) Botu, V.; Ramprasad, R. Adaptive Machine Learning Framework to Accelerate *Ab Initio* Molecular Dynamics. *Int. J. Quantum Chem.* **2015**, *115*, 1074–1083.
- (560) Ceriotti, M.; Tribello, G. A.; Parrinello, M. Demonstrating the Transferability and the Descriptive Power of Sketch-Map. *J. Chem. Theory Comput.* **2013**, *9*, 1521–1532.

- (561) Sobol', I. M.; Asotsky, D.; Kreinin, A.; Kucherenko, S. Construction and Comparison of High-Dimensional Sobol' Generators. *Wilmott* **2011**, *2011*, 64–79.
- (562) Uteva, E.; Graham, R. S.; Wilkinson, R. D.; Wheatley, R. J. Active Learning in Gaussian Process Interpolation of Potential Energy Surfaces. *J. Chem. Phys.* **2018**, *149*, 174114.
- (563) Chmiela, S.; Tkatchenko, A.; Sauceda, H. E.; Poltavsky, I.; Schütt, K. T.; Müller, K.-R. Machine Learning of Accurate Energy-Conserving Molecular Force Fields. *Sci. Adv.* **2017**, *3*, e1603015.
- (564) Kim, H.; Park, J.; Choi, S. Energy Refinement and Analysis of Structures in the QM9 Database via a Highly Accurate Quantum Chemical Method. *Sci. Data* **2019**, *6*, 109.
- (565) Glavatskikh, M.; Leguy, J.; Hunault, G.; Cauchy, T.; Da Mota, B. Dataset's Chemical Diversity Limits the Generalizability of Machine Learning Predictions. *J. Cheminf.* **2019**, *11*, 69.
- (566) <https://www.kaggle.com/c/champs-scalar-coupling/> (accessed 2020-05-01).
- (567) von Lilienfeld, O. A. QM9 challenge. <https://twitter.com/ProfvLilienfeld/status/1073179005854121984>, 2018.
- (568) Fink, T.; Bruggesser, H.; Reymond, J.-L. Virtual Exploration of the Small-Molecule Chemical Universe Below 160 Da. *Angew. Chem., Int. Ed.* **2005**, *44*, 1504–1508.
- (569) Fink, T.; Reymond, J.-L. Virtual Exploration of the Chemical Universe up to 11 Atoms of C, N, O, F: Assembly of 26.4 Million Structures (110.9 Million Stereoisomers) and Analysis for New Ring Systems, Stereochemistry, Physicochemical Properties, Compound Classes, and Drug Discovery. *J. Chem. Inf. Model.* **2007**, *47*, 342–353.
- (570) Blum, L. C.; Reymond, J.-L. 970 Million Druglike Small Molecules for Virtual Screening in the Chemical Universe Database GDB-13. *J. Am. Chem. Soc.* **2009**, *131*, 8732.
- (571) Montavon, G.; Rupp, M.; Gobre, V.; Vazquez-Mayagoitia, A.; Hansen, K.; Tkatchenko, A.; Müller, K.-R.; von Lilienfeld, O. A. Machine Learning of Molecular Electronic Properties in Chemical Compound Space. *New J. Phys.* **2013**, *15*, 095003.
- (572) Ruddigkeit, L.; van Deursen, R.; Blum, L. C.; Reymond, J.-L. Enumeration of 166 Billion Organic Small Molecules in the Chemical Universe Database GDB-17. *J. Chem. Inf. Model.* **2012**, *52*, 2864–2875.
- (573) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Big Data Meets Quantum Chemistry Approximations: The  $\Delta$ -Machine Learning Approach. *J. Chem. Theory Comput.* **2015**, *11*, 2087–2096.
- (574) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.
- (575) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098–3100.
- (576) Stuke, A.; Kunkel, C.; Golze, D.; Todorović, M.; Margraf, J. T.; Reuter, K.; Rinke, P.; Oberhofer, H. Atomic Structures and Orbital Energies of 61,489 Crystal-Forming Organic Molecules. *Sci. Data* **2020**, *7*, 58.
- (577) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (578) Nakata, M.; Shimazaki, T. PubChemQC Project: A Large-Scale First-Principles Electronic Structure Database for Data-Driven Chemistry. *J. Chem. Inf. Model.* **2017**, *57*, 1300–1308.
- (579) Kolb, B.; Zhao, B.; Li, J.; Jiang, B.; Guo, H. Permutation Invariant Potential Energy Surfaces for Polyatomic Reactions using Atomistic Neural Networks. *J. Chem. Phys.* **2016**, *144*, 224103.
- (580) Bernstein, N.; Csányi, G.; Deringer, V. L. De Novo Exploration and Self-Guided Learning of Potential-Energy Surfaces. *npj Comput. Mater.* **2019**, *5*.
- (581) Deringer, V. L.; Pickard, C. J.; Csányi, G. Data-Driven Learning of Total and Local Energies in Elemental Boron. *Phys. Rev. Lett.* **2018**, *120*, 156001.
- (582) Tong, Q.; Xue, L.; Lv, J.; Wang, Y.; Ma, Y. Accelerating CALYPSO Structure Prediction by Data-Driven Learning of a Potential Energy Surface. *Faraday Discuss.* **2018**, *211*, 31–43.
- (583) Podryabinkin, E. V.; Tikhonov, E. V.; Shapeev, A. V.; Oganov, A. R. Accelerating Crystal Structure Prediction by Machine-Learning Interatomic Potentials with Active Learning. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2019**, *99*, 064114.
- (584) Yao, Z.; Sanchez-Lengeling, B.; Bobbitt, N. S.; Bucior, B. J.; Kumar, S. G. H.; Collins, S. P.; Burns, T.; Woo, T. K.; Farha, O.; Snurr, R. Q.; Aspuru-Guzik, A. Inverse Design of Nanoporous Crystalline Reticular Materials with Deep Generative Models. *ChemRxiv* **2020**, DOI: 10.26434/chemrxiv.12186681.v1.
- (585) Krenn, M.; Hase, F.; Nigam, A.; Friederich, P.; Aspuru-Guzik, A. Self-Referencing Embedded Strings (SELFIES): A 100% robust molecular string representation. *Mach. Learn.: Sci. Technol.* **2020**, *1*, 045024.
- (586) Gebauer, N.; Gastegger, M.; Schütt, K. In *Advances in Neural Information Processing Systems 32*; Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., Garnett, R., Eds.; Curran Associates, Inc., 2019; pp 7566–7578.
- (587) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (588) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
- (589) Mai, S.; Richter, M.; Ruckebauer, M.; Oettel, M.; Marquetand, P.; González, L. SHARC2.0: Surface Hopping Including Arbitrary Couplings – Program Package for Non-Adiabatic Dynamics, sharc-md.org, 2018.
- (590) Ceriotti, M.; Tribello, G. A.; Parrinello, M. Simplifying the Representation of Complex Free-Energy Landscapes using Sketch-Map. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13023–13028.
- (591) Tribello, G. A.; Ceriotti, M.; Parrinello, M. Using Sketch-Map Coordinates to Analyze and Bias Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 5196–5201.
- (592) Garijo del Río, E.; Mortensen, J. J.; Jacobsen, K. W. Local Bayesian Optimizer for Atomic Structures. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2019**, *100*, 104103.
- (593) Denzel, A.; Kästner, J. Gaussian Process Regression for Geometry Optimization. *J. Chem. Phys.* **2018**, *148*, 094114.
- (594) Koistinen, O.-P.; Åsgeirsson, V.; Vehtari, A.; Jónsson, H. Nudged Elastic Band Calculations Accelerated with Gaussian Process Regression Based on Inverse Interatomic Distances. *J. Chem. Theory Comput.* **2019**, *15*, 6738–6751.
- (595) Koistinen, O.-P.; Åsgeirsson, V.; Vehtari, A.; Jónsson, H. Minimum Mode Saddle Point Searches Using Gaussian Process Regression with Inverse-Distance Covariance Function. *J. Chem. Theory Comput.* **2020**, *16*, 499–509.
- (596) Meyer, R.; Hauser, A. W. Geometry Optimization using Gaussian Process Regression in Internal Coordinate Systems. *J. Chem. Phys.* **2020**, *152*, 084112.
- (597) Raggi, G.; Galván, I. F.; Ritterhoff, C. L.; Vacher, M.; Lindh, R. Restricted-Variance Molecular Geometry Optimization Based on Gradient-Enhanced Kriging. *J. Chem. Theory Comput.* **2020**, *16*, 3989–4001.
- (598) Seung, H. S.; Opper, M.; Sompolinsky, H. Query by Committee. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*; New York, NY, USA, 1992; pp 287–294.
- (599) Collins, M. Molecular Potential-Energy Surfaces for Chemical Reaction Dynamics. *Theor. Chem. Acc.* **2002**, *108*, 313–324.
- (600) Godsi, O.; Collins, M. A.; Peskin, U. Quantum Grow—A Quantum Dynamics Sampling Approach for Growing Potential Energy Surfaces and Nonadiabatic Couplings. *J. Chem. Phys.* **2010**, *132*, 124106.
- (601) Artrith, N.; Behler, J. High-Dimensional Neural Network Potentials for Metal Surfaces: A Prototype Study for Copper. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2012**, *85*, 045439.

- (602) Dawes, R.; Thompson, D. L.; Guo, Y.; Wagner, A. F.; Minkoff, M. Interpolating Moving Least-Squares Methods for Fitting Potential Energy Surfaces: Computing High-Density Potential Energy Surface Data from Low-Density *Ab Initio* Data Points. *J. Chem. Phys.* **2007**, *126*, 184108.
- (603) Dawes, R.; Thompson, D. L.; Wagner, A. F.; Minkoff, M. Interpolating Moving Least-Squares Methods for Fitting Potential Energy Surfaces: A Strategy for Efficient Automatic Data Point Placement in High Dimensions. *J. Chem. Phys.* **2008**, *128*, 084107.
- (604) Lorenz, S.; Groß, A.; Scheffler, M. Representing High-Dimensional Potential-Energy Surfaces for Reactions at Surfaces by Neural Networks. *Chem. Phys. Lett.* **2004**, *395*, 210–215.
- (605) Raff, L. M.; Malshe, M.; Hagan, M.; Doughan, D. I.; Rockley, M. G.; Komanduri, R. *Ab Initio* Potential-Energy Surfaces for Complex, Multichannel Systems using Modified Novelty Sampling and Feedforward Neural Networks. *J. Chem. Phys.* **2005**, *122*, 084104.
- (606) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.
- (607) Chen, J.; Xu, X.; Xu, X.; Zhang, D. H. A Global Potential Energy Surface for the  $\text{H}_2 + \text{OH} \leftrightarrow \text{H}_2\text{O} + \text{H}$  Reaction using Neural Networks. *J. Chem. Phys.* **2013**, *138*, 154301.
- (608) Jiang, B.; Guo, H. Dynamics of Water Dissociative Chemisorption on Ni(111): Effects of Impact Sites and Incident Angles. *Phys. Rev. Lett.* **2015**, *114*, 166101.
- (609) Shen, X.; Chen, J.; Zhang, Z.; Shao, K.; Zhang, D. H. Methane Dissociation on Ni(111): A Fifteen-Dimensional Potential Energy Surface using Neural Network Method. *J. Chem. Phys.* **2015**, *143*, 144701.
- (610) Shao, K.; Chen, J.; Zhao, Z.; Zhang, D. H. Communication: Fitting Potential Energy Surfaces with Fundamental Invariant Neural Network. *J. Chem. Phys.* **2016**, *145*, 071101.
- (611) Cui, J.; Krems, R. V. Efficient Non-Parametric Fitting of Potential Energy Surfaces for Polyatomic Molecules with Gaussian Processes. *J. Phys. B: At., Mol. Opt. Phys.* **2016**, *49*, 224001.
- (612) Kolb, B.; Marshall, P.; Zhao, B.; Jiang, B.; Guo, H. Representing Global Reactive Potential Energy Surfaces Using Gaussian Processes. *J. Phys. Chem. A* **2017**, *121*, 2552–2557.
- (613) Kolb, B.; Luo, X.; Zhou, X.; Jiang, B.; Guo, H. High-Dimensional Atomistic Neural Network Potentials for Molecule–Surface Interactions: HCl Scattering from Au(111). *J. Phys. Chem. Lett.* **2017**, *8*, 666–672.
- (614) Huang, S.-D.; Shang, C.; Zhang, X.-J.; Liu, Z.-P. Material Discovery by Combining Stochastic Surface Walking Global Optimization with a Neural Network. *Chem. Sci.* **2017**, *8*, 6327–6337.
- (615) Zhou, X.; Nattino, F.; Zhang, Y.; Chen, J.; Kroes, G.-J.; Guo, H.; Jiang, B. Dissociative Chemisorption of Methane on Ni(111) using a Chemically Accurate Fifteen Dimensional Potential Energy Surface. *Phys. Chem. Chem. Phys.* **2017**, *19*, 30540–30550.
- (616) Podryabinkin, E. V.; Shapeev, A. V. Active Learning of Linearly Parametrized Interatomic Potentials. *Comput. Mater. Sci.* **2017**, *140*, 171–180.
- (617) Bruccoleri, R. E.; Karplus, M. Conformational Sampling using High-Temperature Molecular Dynamics. *Biopolymers* **1990**, *29*, 1847–1862.
- (618) Maximova, T.; Moffatt, R.; Ma, B.; Nussinov, R.; Shehu, A. Principles and Overview of Sampling Methods for Modeling Macromolecular Structure and Dynamics. *PLOS Comput. Biol.* **2016**, *12*, 1–70.
- (619) Kästner, J. Umbrella Sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 932–942.
- (620) Tao, G. Trajectory-Guided Sampling for Molecular Dynamics Simulation. *Theor. Chem. Acc.* **2019**, *138*, 34.
- (621) Yang, Y. I.; Shao, Q.; Zhang, J.; Yang, L.; Gao, Y. Q. Enhanced Sampling in Molecular Dynamics. *J. Chem. Phys.* **2019**, *151*, 070902.
- (622) Herr, J. E.; Yao, K.; McIntyre, R.; Toth, D. W.; Parkhill, J. Metadynamics for Training Neural Network Model Chemistries: A Competitive Assessment. *J. Chem. Phys.* **2018**, *148*, 241710.
- (623) Grimme, S. Exploration of Chemical compound, conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations. *J. Chem. Theory Comput.* **2019**, *15*, 2847–2862.
- (624) Smith, J. S.; Nebgen, B.; Lubbers, N.; Isayev, O.; Roitberg, A. E. Less is More: Sampling Chemical Space with Active Learning. *J. Chem. Phys.* **2018**, *148*, 241733.
- (625) Dral, P. O. Quantum chemistry assisted by machine learning; *Adv. Quantum Chem.*; Academic Press, 2020.
- (626) Bernstein, N.; Bhattarai, B.; Csányi, G.; Drabold, D. A.; Elliott, S. R.; Deringer, V. L. Quantifying Chemical Structure and Machine-Learned Atomic Energies in Amorphous and Liquid Silicon. *Angew. Chem., Int. Ed.* **2019**, *58*, 7057–7061.
- (627) Xu, X.; Chen, J.; Zhang, D. H. Global Potential Energy Surface for the  $\text{H} + \text{CH}_4 \leftrightarrow \text{H}_2 + \text{CH}_3$  Reaction using Neural Networks. *Chin. J. Chem. Phys.* **2014**, *27*, 373–379.
- (628) Li, J.; Guo, H. Communication: An Accurate Full 15 Dimensional Permutationally Invariant Potential Energy Surface for the  $\text{OH} + \text{CH}_4 \rightarrow \text{H}_2\text{O} + \text{CH}_3$  Reaction. *J. Chem. Phys.* **2015**, *143*, 221103.
- (629) Jiang, B.; Guo, H. Six-Dimensional Quantum Dynamics for Dissociative Chemisorption of  $\text{H}_2$  and  $\text{D}_2$  on Ag(111) on a Permutation Invariant Potential Energy Surface. *Phys. Chem. Chem. Phys.* **2014**, *16*, 24704–24715.
- (630) Toyoura, K.; Hirano, D.; Seko, A.; Shiga, M.; Kuwabara, A.; Karasuyama, M.; Shitara, K.; Takeuchi, I. Machine-Learning-Based Selective Sampling Procedure for Identifying the Low-Energy Region in a Potential Energy Surface: A Case Study on Proton Conduction in Oxides. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2016**, *93*, 054112.
- (631) Guan, Y.; Yang, S.; Zhang, D. H. Construction of Reactive Potential Energy Surfaces with Gaussian Process Regression: Active Data Selection. *Mol. Phys.* **2018**, *116*, 823–834.
- (632) Vargas-Hernández, R. A.; Guan, Y.; Zhang, D. H.; Krems, R. V. Bayesian Optimization for the Inverse Scattering Problem in Quantum Reaction Dynamics. *New J. Phys.* **2019**, *21*, 022001.
- (633) Todorović, M.; Gutmann, M. U.; Corander, J.; Rinke, P. Bayesian Inference of Atomistic Structure in Functional Materials. *npj Comput. Mater.* **2019**, *5*, 35.
- (634) Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine Learning for Molecular and Materials Science. *Nature* **2018**, *559*, 547–555.
- (635) Shu, Y.; Kryven, J.; Sampaio de Oliveira-Filho, A. G.; Zhang, L.; Song, G.-L.; Li, S. L.; Meana-Pañeda, R.; Fu, B.; Bowman, J. M.; Truhlar, D. G. Direct 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.* **2019**, *151*, 104311.
- (636) Yarkony, D. R. On the Consequences of Nonremovable Derivative Couplings. I. The Geometric Phase and Quasidiabatic States: A Numerical Study. *J. Chem. Phys.* **1996**, *105*, 10456–10461.
- (637) Yarkony, D. R. On the Role of Conical Intersections in Photodissociation. V. Conical Intersections and the Geometric Phase in the Photodissociation of Methyl Mercaptan. *J. Chem. Phys.* **1996**, *104*, 7866–7881.
- (638) Ryabinkin, I. G.; Joubert-Doriol, L.; Izmaylov, A. F. When Do We Need to Account for the Geometric Phase in Excited State Dynamics? *J. Chem. Phys.* **2014**, *140*, 214116.
- (639) Gherib, R.; Ryabinkin, I. G.; Izmaylov, A. F. Why Do Mixed Quantum-Classical Methods Describe Short-Time Dynamics through Conical Intersections So Well? Analysis of Geometric Phase Effects. *J. Chem. Theory Comput.* **2015**, *11*, 1375–1382.
- (640) Ryabinkin, I. G.; Joubert-Doriol, L.; Izmaylov, A. F. Geometric Phase Effects in Nonadiabatic Dynamics Near Conical Intersections. *Acc. Chem. Res.* **2017**, *50*, 1785–1793.
- (641) Xie, C.; Malbon, C. L.; Guo, H.; Yarkony, D. R. Up to a Sign. The Insidious Effects of Energetically Inaccessible Conical Intersections on Unimolecular Reactions. *Acc. Chem. Res.* **2019**, *52*, 501–509.

- (642) Plasser, F.; Ruckebauer, M.; Mai, S.; Oettel, M.; Marquetand, P.; González, L. Efficient and Flexible Computation of Many-Electron Wave Function Overlaps. *J. Chem. Theory Comput.* **2016**, *12*, 1207.
- (643) Zhu, X.; Yarkony, D. R. Quasi-Diabatic Representations of Adiabatic Potential Energy Surfaces Coupled by Conical Intersections including Bond Breaking: A More General Construction Procedure and an Analysis of the Diabatic Representation. *J. Chem. Phys.* **2012**, *137*, 22A511.
- (644) Zhu, X.; Yarkony, D. R. On the Representation of Coupled Adiabatic Potential Energy Surfaces using Quasi-Diabatic Hamiltonians: A Distributed Origins Expansion Approach. *J. Chem. Phys.* **2012**, *136*, 174110.
- (645) Rasmussen, C. E. In *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia*, February 2–14, 2003, Tübingen, Germany, August 4–16, 2003, Revised Lectures; Bousquet, O., von Luxburg, U., Rätsch, G., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2004; pp 63–71.
- (646) Zheng, F.; Gao, X.; Eisfeld, A. Excitonic Wave Function Reconstruction from Near-Field Spectra Using Machine Learning Techniques. *Phys. Rev. Lett.* **2019**, *123*, 163202.
- (647) Fabrizio, A.; Grisafi, A.; Meyer, B.; Ceriotti, M.; Corminboeuf, C. Electron Density Learning of Non-Covalent Systems. *Chem. Sci.* **2019**, *10*, 9424–9432.
- (648) Grisafi, A.; Fabrizio, A.; Meyer, B.; Wilkins, D. M.; Corminboeuf, C.; Ceriotti, M. Transferable Machine-Learning Model of the Electron Density. *ACS Cent. Sci.* **2019**, *5*, 57–64.
- (649) Fabrizio, A.; Briling, K.; Grisafi, A.; Corminboeuf, C. Learning (from) the Electron Density: Transferability, Conformational and Chemical Diversity. *Chimia* **2020**, *74*, 232–236.
- (650) Mai, S.; Plasser, F.; Dorn, J.; Fumanal, M.; Daniel, C.; González, L. Quantitative Wave Function Analysis for Excited States of Transition Metal Complexes. *Coord. Chem. Rev.* **2018**, *361*, 74–97.
- (651) Hong, Y.; Yin, Z.; Guan, Y.; Zhang, Z.; Fu, B.; Zhang, D. H. Exclusive Neural Network Representation of the Quasi-Diabatic Hamiltonians Including Conical Intersections. *J. Phys. Chem. Lett.* **2020**, *11*, 7552–7558.
- (652) Shu, Y.; Truhlar, D. G. Diabatization by Machine Intelligence. *J. Chem. Theory Comput.* **2020**, *16*, 6456–6464.
- (653) Granovsky, A. A. Extended Multi-Configuration Quasi-Degenerate Perturbation Theory: The New Approach to Multi-State Multi-Reference Perturbation Theory. *J. Chem. Phys.* **2011**, *134*, 214113.
- (654) Mennucci, B.; Cappelli, C.; Guido, C. A.; Cammi, R.; Tomasi, J. Structures and Properties of Electronically Excited Chromophores in Solution from the Polarizable Continuum Model Coupled to the Time-Dependent Density Functional Theory. *J. Phys. Chem. A* **2009**, *113*, 3009–3020.
- (655) Jasper, A. W.; Kendrick, B. K.; Mead, C. A.; Truhlar, D. G. *Modern Trends in Chemical Reaction Dynamics*; World Scientific, 2004; pp 329–391.
- (656) Yarkony, D. R. In conical intersections; Domcke, W., Yarkony, D. R., Köppel, H., Eds.; *Advanced Series in Physical Chemistry*; World Scientific, 2004; Vol. 15.
- (657) Cupellini, L.; Bondanza, M.; Nottoli, M.; Mennucci, B. Successes & Challenges in the Atomistic Modeling of Light-Harvesting and its Photoregulation. *Biochim. Biophys. Acta, Bioenerg.* **2020**, *1861*, 148049.
- (658) Kreisbeck, C.; Kramer, T.; Rodríguez, M.; Hein, B. High-Performance Solution of Hierarchical Equations of Motion for Studying Energy Transfer in Light-Harvesting Complexes. *J. Chem. Theory Comput.* **2011**, *7*, 2166–2174.
- (659) Haese, F.; Roch, L. M.; Friederich, P.; Aspuru-Guzik, A. Designing and Understanding Light-Harvesting Devices with Machine Learning. *Nat. Commun.* **2020**, *11*, 4587.
- (660) Richings, G. W.; Robertson, C.; Habershon, S. Can We Use on-the-Fly Quantum Simulations to Connect Molecular Structure and Sunscreen Action? *Faraday Discuss.* **2019**, *216*, 476–493.
- (661) Thompson, K.; Martínez, T. J. Ab initio/interpolated quantum dynamics on coupled electronic states with full configuration interaction wave functions. *J. Chem. Phys.* **1999**, *110*, 1376–1382.
- (662) Behler, J.; Lorenz, S.; Reuter, K. Representing Molecule-Surface Interactions with Symmetry-Adapted Neural Networks. *J. Chem. Phys.* **2007**, *127*, 014705.
- (663) Behler, J. *Dissociation of Oxygen Molecules on the Al(111) Surface*. Ph.D. Thesis, Technical University Berlin, 2004.
- (664) la Cour Jansen, T.; Rettrup, S.; Sarma, C.; Snijders, J.; Palmieri, P. On the Evaluation of Spin-Orbit Coupling Matrix Elements in a Spin-Adapted Basis. *Int. J. Quantum Chem.* **1999**, *73*, 23–27.
- (665) Libisch, F.; Huang, C.; Carter, E. A. Embedded Correlated Wavefunction Schemes: Theory and Applications. *Acc. Chem. Res.* **2014**, *47*, 2768–2775.
- (666) Yin, R.; Zhang, Y.; Libisch, F.; Carter, E. A.; Guo, H.; Jiang, B. Dissociative Chemisorption of O<sub>2</sub> on Al(111): Dynamics on a Correlated Wave-Function-Based Potential Energy Surface. *J. Phys. Chem. Lett.* **2018**, *9*, 3271–3277.
- (667) Martin-Gondre, L.; Crespos, C.; Larregaray, P.; Rayez, J.; van Ootegem, B.; Conte, D. Is the LEPS Potential Accurate Enough to Investigate the Dissociation of Diatomic Molecules on Surfaces? *Chem. Phys. Lett.* **2009**, *471*, 136–142.
- (668) Chen, W.-K.; Fang, W.-H.; Cui, G. A Multi-Layer Energy-Based Fragment Method for Excited States and Nonadiabatic Dynamics. *Phys. Chem. Chem. Phys.* **2019**, *21*, 22695–22699.
- (669) Li, J.; Reiser, P.; Eberhard, A.; Friederich, P.; Lopez, S. Nanosecond Photodynamics Simulations of a Cis-Trans Isomerization Are Enabled by Machine Learning. *ChemRxiv* **2020**, DOI: 10.26434/chemrxiv.13047863.v1.
- (670) Christensen, A. S.; von Lilienfeld, O. A. Operator Quantum Machine Learning: Navigating the Chemical Space of Response Properties. *Chimia* **2019**, *73*, 1028–1031.
- (671) Gastegger, M. *Artificial Intelligence in Theoretical Chemistry*. Ph.D. Thesis, University of Vienna, 2017.
- (672) University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, 2010, TURBOMOLE GmbH; available from <http://www.turbomole.com>.
- (673) Neese, F. The ORCA Program System. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (674) Kasha, M.; Rawls, H. R.; El-Bayoumi, M. A. The Exciton Model in Molecular Spectroscopy. *Pure Appl. Chem.* **1965**, *11*, 371–392.
- (675) Hirshfeld, F. Bonded-atom fragments for describing molecular charge densities. *Theoret. Chim. Acta* **1977**, *44*, 129–138.
- (676) Mulliken, R. S. Electronic Population Analysis on LCAO-MO Molecular Wave Functions. I. *J. Chem. Phys.* **1955**, *23*, 1833–1840.
- (677) Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742–754.
- (678) Durant, J. L.; Leland, B. A.; Henry, D. R.; Nourse, J. G. Reoptimization of MDL Keys for Use in Drug Discovery. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1273–1280.
- (679) Landrum, G. *RDKit: Open-Source Cheminformatics Software*, 2016.
- (680) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Mater.* **2013**, *1*, 011002.
- (681) Sanchez-Gonzalez, A. Accurate Prediction of X-Ray Pulse Properties from a Free-Electron Laser Using Machine Learning. *Nat. Commun.* **2017**, *8*, 15461.
- (682) Aarva, A.; Deringer, V. L.; Sainio, S.; Laurila, T.; Caro, M. A. Understanding X-Ray Spectroscopy of Carbonaceous Materials by Combining Experiments, Density Functional Theory, and Machine Learning. Part II: Quantitative Fitting of Spectra. *Chem. Mater.* **2019**, *31*, 9256–9267.

(683) Aarva, A.; Deringer, V. L.; Sainio, S.; Laurila, T.; Caro, M. A. Understanding X-ray Spectroscopy of Carbonaceous Materials by Combining Experiments, Density Functional Theory, and Machine Learning. Part I: Fingerprint Spectra. *Chem. Mater.* **2019**, *31*, 9243–9255.

(684) Deringer, V. L.; Caro, M. A.; Jana, R.; Aarva, A.; Elliott, S. R.; Laurila, T.; Csányi, G.; Pastewka, L. Computational Surface Chemistry of Tetrahedral Amorphous Carbon by Combining Machine Learning and Density Functional Theory. *Chem. Mater.* **2018**, *30*, 7438–7445.

(685) Janet, J. P.; Kulik, H. J. Predicting Electronic Structure Properties of Transition Metal Complexes with Neural Networks. *Chem. Sci.* **2017**, *8*, 5137–5152.

(686) Janet, J. P.; Duan, C.; Yang, T.; Nandy, A.; Kulik, H. J. A Quantitative Uncertainty Metric Controls Error in Neural Network-Driven Chemical Discovery. *Chem. Sci.* **2019**, *10*, 7913–7922.

(687) Janet, J. P.; Gani, T. Z. H.; Steeves, A. H.; Ioannidis, E. I.; Kulik, H. J. Leveraging Cheminformatics Strategies for Inorganic Discovery: Application to Redox Potential Design. *Ind. Eng. Chem. Res.* **2017**, *56*, 4898–4910.

(688) Janet, J. P.; Chan, L.; Kulik, H. J. Accelerating Chemical Discovery with Machine Learning: Simulated Evolution of Spin Crossover Complexes with an Artificial Neural Network. *J. Phys. Chem. Lett.* **2018**, *9*, 1064–1071.

(689) Janet, J. P.; Ramesh, S.; Duan, C.; Kulik, H. J. Accurate Multiobjective Design in a Space of Millions of Transition Metal Complexes with Neural-Network-Driven Efficient Global Optimization. *ACS Cent. Sci.* **2020**, *6*, 513–524.

(690) Kohn, A. W.; Lin, Z.; Van Voorhis, T. Toward Prediction of Nonradiative Decay Pathways in Organic Compounds I: The Case of Naphthalene Quantum Yields. *J. Phys. Chem. C* **2019**, *123*, 15394–15402.

(691) Qiu, J.; et al. Prediction and Understanding of AIE Effect by Quantum Mechanics-Aided Machine-Learning Algorithm. *Chem. Commun.* **2018**, *54*, 7955–7958.

(692) Ju, C.-W.; Bai, H.; Li, B.; Liu, R. Machine Learning Enables Highly Accurate Predictions of Photophysical Properties of Organic Fluorescent Materials: Emission Wavelengths and Quantum Yields. *ChemRxiv* **2020**, DOI: 10.26434/chemrxiv.12111060.v2.

(693) Vacher, M.; Farahani, P.; Valentini, A.; Frutos, L. M.; Karlsson, H. O.; Fdez. Galván, I.; Lindh, R. How Do Methyl Groups Enhance the Triplet Chemiexcitation Yield of Dioxetane? *J. Phys. Chem. Lett.* **2017**, *8*, 3790–3794.

(694) Häse, F.; Galván, I. F.; Aspuru-Guzik, A.; Lindh, R.; Vacher, M. Machine Learning for Analysing *Ab Initio* Molecular Dynamics Simulations. *J. Phys.: Conf. Ser.* **2020**, *1412*, 042003.

(695) Vacher, M.; Brakestad, A.; Karlsson, H. O.; Fdez. Galván, I.; Lindh, R. Dynamical Insights into the Decomposition of 1,2-Dioxetane. *J. Chem. Theory Comput.* **2017**, *13*, 2448–2457.

(696) Đorđević, N.; Beckwith, J. S.; Yarema, M.; Yarema, O.; Rosspeintner, A.; Yazdani, N.; Leuthold, J.; Vauthey, E.; Wood, V. Machine Learning for Analysis of Time-Resolved Luminescence Data. *ACS Photonics* **2018**, *5*, 4888–4895.

(697) Abramavicius, D.; Jiang, J.; Bulheller, B. M.; Hirst, J. D.; Mukamel, S. Simulation Study of Chiral Two-Dimensional Ultraviolet Spectroscopy of the Protein Backbone. *J. Am. Chem. Soc.* **2010**, *132*, 7769–7775.

(698) Zhang, X.-X.; Würth, C.; Zhao, L.; Resch-Genger, U.; Ernstring, N. P.; Sajadi, M. Femtosecond Broadband Fluorescence Upconversion Spectroscopy: Improved Setup and Photometric Correction. *Rev. Sci. Instrum.* **2011**, *82*, 063108.

(699) Ahmad, I.; Ahmed, S.; Anwar, Z.; Sheraz, M. A.; Sikorski, M. Photostability and Photostabilization of Drugs and Drug Products. *Int. J. Photoenergy* **2016**, *2016*, 1–19.

(700) Dobson, C. M. Chemical Space and Biology. *Nature* **2004**, *432*, 824–828.

## NOTE ADDED AFTER ASAP PUBLICATION

This paper was published ASAP on November 19, 2020, with an incorrect version of Figure 10. The corrected version was posted on November 23, 2020.