On application of deep learning to simplified quantum-classical dynamics in electronically excited states

IOP Publishing Mach. Learn.: Sci. Technol. 2 (2021) 035039 https://doi.org/10.1088/2632-2153/abfe3f

MACHINE LEARNING
Science and Technology

OPEN ACCESS

RECEIVED
15 December 2020

REVISED
20 April 2021

ACCEPTED FOR PUBLICATION
5 May 2021

PUBLISHED
14 July 2021

PAPER

On application of deep learning to simplified quantum-classical dynamics in electronically excited states

Evgeny Posenitskiy[1,*], **Fernand Spiegelman**[2] **and Didier Lemoine**[1]

[1] Laboratoire Collisions Agrégats et Réactivité (LCAR), UMR5589 Université de Toulouse (UPS) and CNRS, 118 Route de Narbonne, F-31062 Toulouse, France
[2] Laboratoire de Chimie et Physique Quantiques (LCPQ), UMR5626 Université de Toulouse (UPS) and CNRS, 118 Route de Narbonne, F-31062 Toulouse, France
[*] Author to whom any correspondence should be addressed.

E-mail: posenitskiy@irsamc.ups-tlse.fr

Keywords: non-adiabatic dynamics, excited states, TD-DFTB, machine learning

Supplementary material for this article is available online

Abstract

Deep learning (DL) is applied to simulate non-adiabatic molecular dynamics of phenanthrene, using the time-dependent density functional based tight binding (TD-DFTB) approach for excited states combined with mixed quantum–classical propagation. Reference calculations rely on Tully's fewest-switches surface hopping (FSSH) algorithm coupled to TD-DFTB, which provides electronic relaxation dynamics in fair agreement with various available experimental results. Aiming at describing the coupled electron-nuclei dynamics in large molecular systems, we then examine the combination of DL for excited-state potential energy surfaces (PESs) with a simplified trajectory surface hopping propagation based on the Belyaev–Lebedev (BL) scheme. We start to assess the accuracy of the TD-DFTB approach upon comparison of the optical spectrum with experimental and higher-level theoretical results. Using the recently developed SchNetPack (Schütt *et al* 2019 *J. Chem. Theory Comput.* **15** 448–55) for DL applications, we train several models and evaluate their performance in predicting excited-state energies and forces. Then, the main focus is given to the analysis of the electronic population of low-lying excited states computed with the aforementioned methods. We determine the relaxation timescales and compare them with experimental data. Our results show that DL demonstrates its ability to describe the excited-state PESs. When coupled to the simplified BL scheme considered in this study, it provides reliable description of the electronic relaxation in phenanthrene as compared with either the experimental data or the higher-level FSSH/TD-DFTB theoretical results. Furthermore, the DL performance allows high-throughput analysis at a negligible cost.

1. Introduction

Modeling the properties and evolution of large molecular systems is a challenging task in chemical and biological sciences, which usually requires solving the stationary or the time-dependent Schrödinger equation. Recently, machine learning (ML) has appeared as a promising tool that can be used to fully or partially avoid electronic structure calculations in atomistic simulations [1–6]. ML methodologies may differ depending on the objective and can be categorized in three main branches: (i) supervised learning, (ii) unsupervised learning and (iii) reinforcement learning. The major focus in reinforcement learning is given to finding a balance between exploration (of uncharted areas) and exploitation (of current knowledge) [7]. The main objective of the unsupervised learning is to look for previously undetected patterns in a dataset. The supervised learning aims to find optimal parameters for a function that maps an input to an output based on a given set of input–output pairs (also called the training set). Ideally, the supervised ML model should infer some knowledge from the training set in order to perform reasonably well on new data entries.

For example, the global structure optimization can be enhanced by each of the aforementioned techniques [8–12]. Artificial neural networks (NNs) have been actively studied and used in various fields of science, thus forming a deep learning (DL) domain of ML [13]. In particular, it has been proven that multilayer feed-forward NNs are 'universal approximators' [14, 15], i.e. they can approximate an unknown multidimensional function with an arbitrary accuracy based on a set of known function values. The more hidden layers and nodes are included, the more flexible functional form of the network is obtained. The supervised ML models can be used to construct complex interatomic potentials [2, 16] that can in turn be used to perform extensive molecular dynamics simulations [17–19]. Ideally, this would correspond to an *ab initio* accuracy at the computational cost of a force field. Some supervised ML models include forces in the training process, which drastically improves the quality of a machine-learned potential energy surface (PES) [20–23]. Alternatively, NNs have also been used to correct density functional theory (DFT) [24] and density functional based tight-binding (DFTB) [25] results based on a Δ-ML [26] approach. Several software packages are available for applications of supervised ML to atomistic simulations, e.g. MLatom [27], DeePMD-kit [28] and SchNetPack [29].

Addressing electronically excited states and non-adiabatic dynamics in excited states obviously proves both methodologically and computationally much more demanding than merely ground electronic state properties and dynamics, which involves among other things complex shapes of PES and the occurrence of conical intersections. Several theoretical approaches have been devised to incorporate non-adiabatic effects in the dynamics. The most accurate type is fully *ab initio* and deals with quantum dynamics of nuclei, such as the multi-configuration time-dependent Hartree (MCTDH) method [30, 31] propagating time-dependent wavepackets. Alternatively, mixed quantum–classical schemes have been developed based on a classical description of the nuclear motion, e.g. the mean-field propagation [32] or the trajectory surface hopping (TSH) approach [33–35]. Within the generalized TSH picture, the nuclear wavepacket evolution is simulated by an ensemble of independent classical trajectories and non-adiabatic effects are taken into account via a probability to switch from the current electronic state to another one. Recently, a DL architecture SchNet [23] has been interfaced with SHARC [36] code for non-adiabatic molecular dynamics [37]. Other studies have also reported excited-state dynamics using supervised ML schemes [38–42]. Some of them require an *a priori* knowledge about the location of conical intersections in order to include more training points in this region [38, 39], while others attempt to include the non-adiabatic couplings in the ML model [40, 41], which is another challenging and computationally demanding task. Alternatively, direct quantum dynamics can be performed within the Bayesian framework [43] or using the MCTDH propagation on a machine-learned PES [44]. Recent progress regarding ML applications to electronically excited states has been reviewed by Westermayr *et al* [45].

The main motivation of the present work is to investigate the ability of DL schemes at describing non-adiabatic dynamics in relatively large molecular systems with no *a priori* knowledge about the topology of the excited-state PESs, based on on-the-fly blind propagation involving all nuclear degrees of freedom and as many excited states as needed. Electronic structure calculations can become computationally extremely demanding for medium and large polyatomic molecules even within the commonly used linear response time-dependent density functional theory (TD-DFT) approach [46]. This has led us to choose the approximate DFTB [47–50] scheme that can be applied to large systems, due to the use of parameterized (precomputed) integrals and of a minimal valence basis set. The linear response time-dependent density functional based tight binding (TD-DFTB) approach has been developed in order to access excited states [51]. It has been coupled to nuclear dynamics via Tully's fewest-switches surface hopping (FSSH) scheme and has been applied to medium-sized systems (namely a few tens of atoms) by several groups [52–56]. Recently, we showed that this combined FSSH/TD-DFTB approach satisfactorily describes the relaxation of highly excited electronic states in polyacenes [55, 57]. Yet, achieving statistically meaningful sampling of the initial conditions and/or propagating long-lasting trajectories beyond a few hundreds of femtoseconds proves to be numerically involved. Moreover, for large systems beyond a few hundreds or a few thousands atoms, the TD-DFTB scheme may no longer prove computationally feasible. In addition, the computational cost of non-adiabatic coupling calculations scales as powers of the the number of involved states and the system size, respectively. Thus, for large complexes, especially with high density of states [58], the generation of training data for non-adiabatic couplings might be overly demanding. Also, ML potentials are not at present well transferable from small to large systems and there may be a need to avoid the explicit use of non-adiabatic couplings in the dynamics, so that consideration of simpler hopping algorithms is of particular interest. Historically, the Landau–Zener (LZ) approximation [59, 60] has been proposed to evaluate the hopping probability at diabatic state crossings. The LZ formulation has been adapted by Belayev and Lebedev [61] in order to deal with adiabatic states. Alternatively, Zhu and Nakamura [62, 63] have proposed an improved LZ

formula for the switching probability. In recent years, multidimensional extensions of the Zhu-Nakamura (ZN) theory have appeared in the literature [64, 65].

The present study is dedicated to the application of DL through the use of the SchNet architecture, to simplify quantum–classical dynamics in excited electronic states. The goals of the present study are to (i) obtain accurate machine-learned energies and forces for excited states based on TD-DFTB reference data; (ii) perform molecular dynamics and incorporate non-adiabatic effects via a simplified TSH approach without any *a priori* knowledge about conical intersections or non-adiabatic couplings; (iii) benchmark the simplified TSH scheme against the reference FSSH algorithm on the one hand and available experimental data on the other hand. The paper is organized as follows. In the next section, we briefly outline the basics of the TD-DFTB approach as well as some details regarding the SchNet architecture. Next, we present and discuss the results required for the benchmark study on phenanthrene, first assessing the accuracy of TD-DFTB and SchNet methods and then analyzing the electronic relaxation computed with different hopping schemes. Complementary on-the-fly analysis, i.e. along some selected trajectories, has been performed in order to shed light on the origin of the observed deviations between the simplified TSH and the FSSH approaches. Finally, the conclusions and perspectives are given.

2. Methods

In this section, we present a brief outline of the TD-DFTB and DL methods to calculate excited state properties and several TSH approaches to non-adiabatic dynamics, which are considered in this study.

2.1. Electronic structure calculations

The self-consistent charge DFTB was developed by Elstner *et al* [50] as an extension of the original DFTB framework for the ground state [47, 48]. It is based on the second order expansion of the Kohn–Sham DFT total energy around a reference electronic density, so the final expression for the DFTB total energy of a system with M atoms reads

$$E_{\mathrm{SCC}} = \sum_{i=1}^{N_{\mathrm{occ}}} n_i \sum_{A=1}^{M} \sum_{\mu \in A} \sum_{B=1}^{M} \sum_{\nu \in B} c_{\mu i} H^0_{\mu\nu} c_{\nu i} + \frac{1}{2} \sum_{A=1}^{M} \sum_{B=1}^{M} \Delta q_A \gamma_{AB} \Delta q_B + E_{\mathrm{rep}}, \tag{1}$$

where n_i is the occupation number of ith molecular orbital (MO), μ and ν are the Kohn–Sham atomic orbital (AO) indices, Δq_A is the Mulliken charge of atom A, E_{rep} is the atomic repulsive contribution, $c_{\mu i}$ are the Kohn–Sham MO coefficients and γ_{AB} describes the Coulomb interaction between spherically symmetric charge distributions centered on atoms A and B with a short-range exchange-correlation contribution. The total energy E_{SCC} is further minimized following the self-consistent procedure as proposed in the original work [50]. The remarkable computational efficiency of the DFTB approach comes from the fact that AO overlap and Hamiltonian matrix elements as well as repulsive potentials can be calculated only once and tabulated for a set of interatomic distances between different pairs of elements. In practice, they are actually precomputed based on some reference electronic structure method (most often DFT) and stored in an external file.

Linear response TD-DFTB was developed by Niehaus *et al* [51] as a DFTB analogue of the conventional linear response TD-DFT [46]. Excitation energies are given as eigenvalues Ω_I of the following matrix equation:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \Omega_I \begin{pmatrix} \mathbf{I} & 0 \\ 0 & -\mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}, \tag{2}$$

where **I** is the identity matrix, **A** and **B** are matrices with the elements given by

$$A_{ia,jb} = (\varepsilon_a - \varepsilon_i)\delta_{ij}\delta_{ab} + 2K^{\Sigma}_{ia,jb}; \tag{3}$$

$$B_{ia,jb} = 2K^{\Sigma}_{ia,jb}; \tag{4}$$

and indices i, j and a, b denoting the occupied and virtual MOs with energies ε_i and ε_a, respectively; $\Sigma = S(T)$ if singlet (triplet) excited states have to be computed. The coupling matrix elements $K^{S}_{ia,jb}$ for singlet TD-DFTB transitions can be calculated using the generalized Mulliken approximation as follows [51]

$$K^{S}_{ia,jb} = \sum_{A=1}^{M}\sum_{B=1}^{M} q^{ia}_{A}\gamma_{AB}q^{jb}_{B}, \tag{5}$$

where q^{ia}_{A} are Mulliken atomic transition charges. It is worth mentioning that the TD-DFTB absorption spectrum neither involves doubly excited states, nor Rydberg states.

In order to perform molecular dynamics simulations coupled to TD-DFTB for electronic structure calculations, excited state energy gradients have to be developed. Their derivation for linear response TD-DFTB relies on the so-called Z-vector method, which was initially applied by Furche and Ahlrichs [66, 67] to compute analytical forces within the TD-DFT approach. It was further adapted by Heringer *et al* [68, 69] to compute the gradients of the TD-DFTB excitation energy Ω_I. Alternatively, one can follow derivations of TD-DFTB gradients with the long-range correction from [53].

Application of DL to molecular dynamics is essentially based on learning complex multidimensional PES, which has been reported for ground [23] and excited [37] states. We rely on a recently developed DL architecture called SchNet [23], which is implemented in the open-source Python package SchNetPack [29]. There are several reasons why this particular implementation has been chosen. First of all, SchNet is a continuous-filter convolutional NN, which automatically generates filters that map one hidden layer to another one based on pairwise interactions between a given atom and the surrounding ones [70]. It has been shown to yield accurate molecular representations and energies, especially when forces are included in the training [23]. Furthermore, it satisfies all the required symmetries and the resulting energies are rotationally and translationally invariant, while forces are rotationally equivariant, thus providing energy-conserving force models [29]. A shifted softplus is used as an activation function:

$$\mathrm{ssp}(x) = \ln\left(\frac{\mathrm{e}^{x}+1}{2}\right), \tag{6}$$

due to an infinite order of continuity, which allows to obtain smooth PES, forces and derivative properties. SchNet is trained by minimizing the cost function J, which is computed as follows

$$J\left([E,\mathbf{F}_1,\ldots,\mathbf{F}_M],[\widetilde{E},\widetilde{\mathbf{F}}_1,\ldots,\widetilde{\mathbf{F}}_M]\right) = \rho||E-\widetilde{E}||^2 + \frac{1}{M}\sum_{A=1}^{M}\left|\left|\mathbf{F}_A - \left(-\frac{\partial\widetilde{E}}{\partial\mathbf{R}_A}\right)\right|\right|^2, \tag{7}$$

if both energies E and atomic forces $\mathbf{F}$ are included in the training. In the equation above, ρ is the trade-off between the energy and force loss [71], quantities with tilde denote the model (SchNet) predictions and plain symbols denote reference data. The second term in equation (7) can be neglected if only energies are provided. The initial dataset is split into training, validation and test sets. The training set is used to optimize model parameters in order to minimize the aforementioned cost function. The validation set is used for an early stopping in order to prevent overfitting, which is a common issue in ML applications when model predictions are very accurate for the training data but significantly less accurate for new inputs [45, 72]. The test set is used to evaluate the final model accuracy. More details about the SchNet architecture and practical aspects can be found in [70].

2.2. Non-adiabatic molecular dynamics

We now briefly go through the details of the quantum–classical TSH method for incorporating the non-adiabatic effects. In the TSH approach, the nuclear wavepacket motion is simulated by an ensemble of independent classical trajectories. Each trajectory evolves on a single electronic state at a given time with a probability to switch (hop) from the current state to another one.

In the Tully's FSSH approach, the switching between excited states is controlled by the electron dynamics. Substituting the electronic wavefunction expanded in a basis of adiabatic electronic states into the time-dependent electronic Schrödinger equation, one derives the following equation for the propagation of the complex expansion coefficients $C_J(t)$ [33]:

$$i\hbar\frac{dC_J(t)}{dt} = C_J(t)E_J(t) - i\hbar\sum_{K\neq J} C_K(t)D_{JK}(t), \tag{8}$$

where E_J is the adiabatic energy of state J, D_{JK} is the non-adiabatic coupling between states J and K, which is calculated using a finite difference method as follows [73]

$$D_{JK}(t+\Delta t/2) \approx \frac{1}{2\Delta t}[\langle\psi_J(t)|\psi_K(t+\Delta t)\rangle - \langle\psi_J(t+\Delta t)|\psi_K(t)\rangle]. \tag{9}$$

In the equation above and hereafter, Δt denotes the nuclear time step that is used to propagate a trajectory. It is important to apply the decoherence correction on C_J since the propagation of equation (8) in FSSH is overcoherent, which means that electronic coherences $C_I C_J^*$ do not vanish after passing through the region of strong coupling between states I and J. Decoherence corrections have been shown to be crucial in a number of applications [74, 75], and we rely on the commonly used energy-based correction [74] of the C_J coefficients.

The FSSH probability to switch from the active state I to another state K during the electronic time step $\Delta\tau$ (used to propagate equation (8)) is estimated from the following equation [73]

$$P_{IK}^{\mathrm{FS}}(\tau) = \max\left[0; -2\Delta\tau \frac{\mathrm{Re}\,(C_I^*(\tau)C_K(\tau))}{|C_I|^2} D_{IK}(\tau)\right], \tag{10}$$

where $|C_I|^2$ is the electronic population of a given excited state I. A uniform random number $0 < \xi < 1$ is generated at each nuclear time step to determine whether the hop from the active state I to another state K is allowed from the quantum point of view [73]. The hop is accepted if the following condition is fulfilled

$$\sum_{J=1}^{K-1} P_{IJ}^{\mathrm{FS}} < \xi \leqslant \sum_{J=1}^{K} P_{IJ}^{\mathrm{FS}} \qquad K \neq 1; \tag{11}$$

$$\xi \leqslant P_{I1}^{\mathrm{FS}} \qquad \text{otherwise.} \tag{12}$$

More details regarding the implementation of the Tully's FSSH scheme can be found in [52–54] for the TD-DFTB approach in general and in [55] for the particular implementation in the deMon-Nano [76] code that has been used in this study.

Even though the FSSH is a commonly used approach to non-adiabatic dynamics, it has a number of limitations. For example, the wavefunctions overlap in equation (9) can be computed differently within density-based methods like TD-DFT(B) [55, 77], decoherence corrections are somewhat ad hoc [74] and there are numerical instabilities in non-adiabatic couplings that can require additional diabatization of equation (8) [53, 78]. However, it is possible to avoid the aforementioned issues of the FSSH method if there is no electronic equation to propagate and no non-adiabatic couplings to compute. This is one of the reasons why results of simplified methods based on the LZ approximation still appear in the literature [38, 39, 58, 61, 64, 65, 79]. These schemes are essentially two-state models with one of the following switching probabilities: (i) based on the Belyaev–Lebedev (BL) approach [61] or (ii) based on the ZN theory [62, 63]. The BL hopping probability can be computed as follows

$$P_{IJ}^{\mathrm{BL}} = \exp\left(-\frac{\pi}{2\hbar}\sqrt{\frac{Z_{IJ}^3}{\ddot{Z}_{IJ}}}\right), \tag{13}$$

where $Z_{IJ} = |E_I - E_J|$ is the adiabatic energy gap between states I and J and $\ddot{Z}_{IJ}$ is its second time derivative at the crossing point t_c, i.e. at the local minimum of Z_{IJ}. It is worth mentioning that the equation above is applicable only at the diabatic curve-crossing point.

One of the key differences between the simplified TSH and the FSSH is that the hopping probability P_{IJ}^{BL} is evaluated only at the local minimum of the adiabatic energy gap Z_{IJ} and only between neighboring states, which is a prerequisite of the underlying LZ approximation. A switch from an active state I to state J is performed if P_{IJ}^{BL} is larger than a random number $0 < \xi < 1$. Notably, the aforementioned simplified TSH scheme naturally accounts for decoherence effects [58]. On the other hand, the electronic population $|C_I|^2$ of excited state S_I cannot be computed directly since coefficients C_I are generally missing in the LZ approach. However, the population can be approximated based on the fractional occupation that is a ratio N_I/N_{traj}, where N_I is the number of trajectories currently (at a given moment of time) running on a PES of S_I and N_{traj} is the total number of trajectories.

To conserve the total energy after hopping, the nuclear velocities are rescaled uniformly by a factor β following the energy conservation law

$$E_I + E_{\mathrm{kin}} = E_J + \beta^2 E_{\mathrm{kin}}. \tag{14}$$

Thus, the switch of two states can be still rejected if the energy gap $E_J - E_I > E_{\text{kin}}$. Such hops are called 'frustrated' or classically forbidden hops. Alternatively, one may perform a more accurate velocity rescaling based on energy gradients in order to conserve the total angular momentum [64, 80].

2.3. Computational details

Dynamics in low-lying excited states of phenanthrene ($C_{14}H_{10}$) has been chosen as a test case due to both its relatively large molecular size and the diversity of available experimental results, furthermore giving rise to some ambiguity [81] in their interpretation. Also, phenanthrene is among the most abundant species observed in meteorites [82].

The first part of the discussion in the next section is devoted to the comparison of TD-DFTB and TD-DFT absorption spectra. The TD-DFT absorption spectrum was computed with Gaussian 09 package [83] using BLYP functional and 6-31G(d,p) basis set.

All DFTB calculations presented in this study have been performed with the deMon-Nano [76] code using the mio-1-1 [50] set of parameters. The ability to reproduce energies and geometries corresponding to low-lying excited states of organic molecules has been critically assessed for different sets of DFTB parameters in [84, 85].

Three TD-DFTB datasets have been generated independently for the S_2, S_3 and S_4 adiabatic excited states of phenanthrene. These particular states have been chosen for the reasons given in the next section. Each generated dataset contains 10 000 points that have been sampled from a single *NVT* trajectory equilibrated at $T = 500$ K in the corresponding excited state during 50 ps using a chain of five Nosé–Hoover thermostats and $\Delta t = 0.5$ fs. This temperature allows to sample a wider region of the PES, which is desirable due to the absence of any *a priori* knowledge about conical intersections between the states of interest. Excited state forces have been included in all generated datasets in order to benefit from a relatively low computational cost of reference TD-DFTB calculations and to enhance the training performance.

Next, each dataset is provided to SchNetPack with additional parameters that specify the network topology, i.e. the number of interaction blocks and the number of features, which are specified in section 3.2. In fact, SchNet consists of a representation network (containing interaction blocks) and a prediction network [23]. We have only changed the topology of the former while the latter was kept fixed to the default pyramidal configuration, namely three dense layers with 24 input nodes (equal to the number of atoms) and one output node. Additionally, one may activate the graphical processing units (GPUs) and/or include forces in the training. All SchNet models have been trained with a mini-batch stochastic gradient descent, using ADAM optimizer [86] with mini-batches of 100 samples, 1000 training epochs and the trade-off $\rho = 1$ in order to put more emphasis on the energy loss [23]. The trade-off parameter has not been fine-tuned since the default value was found to produce relatively accurate models. The learning rate is initially set to 10^{-4} and can decay exponentially with a ratio 0.8 down to 10^{-6} with a patience of 25 steps. The maximum number of epochs is set to 5000. The total number of model parameters varies between 238 465 and 886 529 depending on the SchNet topology for phenanthrene, thereby achieving a computational regime taking full advantage of the use of GPUs.

It is worth mentioning that by default SchNet training and evaluation is performed in single precision for floating point numbers. While this causes no issues during the training step, application of the resulting models to TSH simulations with small time steps might be affected. In particular, the BL approach requires smooth PES since hops are only possible at the local minima. We have performed TSH simulations with SchNet models evaluated in single or double precision. Convergence issues and numerical artefacts associated with the use of single-precision SchNet models are discussed in the supplementary information (is available online at stacks.iop.org/MLST/2/035039/mmedia).

Initial conditions for the TSH simulations were sampled from the thermal distribution of the ground state. A single trajectory was equilibrated at $T = 300$ K during 50 ps using a chain of five Nosé–Hoover thermostats and $\Delta t = 0.5$ fs. Snapshots were taken every 50 fs to be further used as initial conditions.

Coupled FSSH/TD-DFTB simulations have been performed with the deMon-Nano code and serve as reference calculations for the simplified TSH scheme. Each classical trajectory has been propagated using the Velocity Verlet algorithm with $\Delta t = 0.25$ fs during 1 ps and the conical intersection threshold [55] set to 10 meV for excited singlet states. Equation (8) is integrated using a 4th order Runge-Kutta algorithm with an electronic time step $\Delta\tau = 0.048$ a.s. In order to be consistent, the same set of initial conditions has been used for Tully's FSSH and for simplified TSH scheme.

The SchNetPack is natively interfaced with the atomic simulation environment (ASE) [87]—an open-source set of Python tools for atomistic simulations. This is why the coupled TSH/SchNet simulations have been performed using the molecular dynamics driver of the ASE with the BL switching probability

computed on the fly. The ASE driver calls the pre-trained SchNet model to compute adiabatic energies and forces at each nuclear time step, which are further used to run the nuclear dynamics (using Velocity Verlet algorithm) and to detect the local minimum of the energy gap. Due to the fact that the simplified TSH scheme requires data from three consecutive steps (to detect the local minimum of Z_{IJ} and to compute $\ddot{Z}_{IJ}$ with finite differences), one has to actually come back to step t_c and make the $[t_c] \rightarrow [t_c + \Delta t]$ step again on a new PES if the hop was accepted at t_c. The second derivative of the energy gap $\ddot{Z}_{IJ}$ at the crossing point t_c is computed using finite differences as follows

$$\ddot{Z}_{IJ}|_{t=t_c} \approx \frac{Z_{IJ}(t_c + \Delta t) - 2Z_{IJ}(t_c) + Z_{IJ}(t_c - \Delta t)}{\Delta t^2}. \tag{15}$$

We have benchmarked our implementation of the BL scheme for a one-dimensional model problem, namely Tully's simple avoided crossing model from [33]. The computed transition probabilities (see figure S2 in the supplementary information) are in good agreement with the work of Hanasaki *et al* [65] and the observed deviations are associated with the use of different particle mass.

Finally, experimental results for phenanthrene have been extracted from [82, 88–91]. We consider more specifically the full width at half maximum (FWHM) assessed from early studies performed in supersonic jets [82, 88, 89] and the relaxation time from pump-probe experiments of Blanchet *et al* [90]. The approximate decay time τ_{app} can be derived from the FWHM as follows [81]

$$\tau_{\text{app}} = \frac{1}{2\pi c \Gamma}, \tag{16}$$

where Γ is the FWHM in cm^{-1} and c is the velocity of light.

3. Results

This section is dedicated to the results of several TSH models applied to the electronic relaxation in phenanthrene. It includes: (i) the validation of the TD-DFTB excited states upon comparison of the computed absorption spectrum *versus* TD-DFT, CASPT2 and experimental results; (ii) the application of the SchNet model to learn energies and forces based on the TD-DFTB reference datasets; (iii) the analysis of the results of the FSSH/TD-DFTB simulations and their comparison with experimental findings; and finally (iv) the comparison of the simplified TSH simulations based on the BL scheme, with the FSSH/TD-DFTB results, in order to assess the relevance and performance of non-adiabatic dynamics based upon machine-learned PESs with no *a priori* knowledge about couplings and conical intersections.

3.1. Absorption spectra of phenanthrene

First, the positions of several low-lying singlet states of neutral phenanthrene together with the corresponding FWHMs and decay times (where available) are deduced from the various experimental UV spectra and summarized in table 1.

We have computed the absorption spectrum of neutral phenanthrene at both the TD-DFTB and the TD-DFT levels of theory. These results together with selected experimental values are compiled in figure 1. The agreement between TD-DFTB and TD-DFT spectra is reasonably good and both methods predict positions of two experimentally observed bright states around 284 and 266 nm with remarkable accuracy. It is important to note that the band at 284 nm (4.37 eV) corresponds to the $S_0 \rightarrow S_3$ transition within both TD-DFT and TD-DFTB, whereas the available experimental studies attribute it to the $S_0 \rightarrow S_2$ excitation. However, this might be due to the relatively low oscillator strength ($\approx 10^{-3}$) of S_1 compared to S_2 (0.06–0.07). For more details, see table S1 in supplementary information. Moreover, the TD-DFTB ratio κ of oscillator strengths for S_3/S_2 and S_3/S_1 amounts to 1.5 and 79.4, respectively, the latter value being in very good agreement with $\kappa = 84$ reported by Amirav *et al* [88]. As for TD-DFT, κ for S_3/S_2 and S_3/S_1 is 1.4 and 113.7, respectively. The TD-DFTB results can also be compared with the *ab initio* CASPT2 calculations performed for neutral phenanthrene by González-Luque *et al* [93]. The first bright state in CASPT2 is S_3 with an excitation energy equal to 4.37 eV, which means that the TD-DFTB ordering is actually reasonable, while dark S_1 and S_2 states can be found at 3.42 and 4.26 eV, respectively. More recently, Nazari *et al* [94] have performed a detailed analysis of the PESs and of the ultrafast dynamics of monomeric phenanthrene and of some of its derivatives. Whereas both the TD-DFT and CASPT2 absorption spectra are in reasonably good agreement with previous studies, the ordering of the states in [94] can be somewhat misleading since the higher-lying excited states (S_3 and above) have been discarded.

Table 1. Selected results from the experiments with neutral phenanthrene in different environments.

Environment	Attributed transition	Position, nm	FWHM, cm^{-1}	Decay time, ps
Supersonic jet[a]	$S_0 \rightarrow S_1$	340.9	5.8	0.92
	$S_0 \rightarrow S_2$	282.6	11.1	0.48
Supersonic jet[b]	$S_0 \rightarrow S_1$	340.9	—	—
	$S_0 \rightarrow S_2$	282.6	12.6	0.42
Isopentane glass at 4 K[c]	$S_0 \rightarrow S_2$	—	67 ± 7	0.08
Supersonic jet[d]	$S_0 \rightarrow S_2$	—	12.0	0.44
Ne matrix at 4 K[e]	$S_0 \rightarrow S_1$	341.1	—	—
	$S_0 \rightarrow S_2$	284.3	—	—
	$S_0 \rightarrow S_3$	273.4	—	—
	$S_0 \rightarrow S_4$	262.4	—	—
	$S_0 \rightarrow S_5$	243.0	—	—
	$S_0 \rightarrow S_6$	229.0	—	—
Gas phase[f]	$S_0 \rightarrow S_1$	341.6	—	—
	$S_0 \rightarrow S_2$	283.7	—	0.52

[a] Amirav *et al* [88].
[b] Ohta and Baba [89].
[c] Dick and Nickel [91].
[d] Brechignac and Hermine [82].
[e] Salama *et al* [92].
[f] Blanchet *et al* [90].

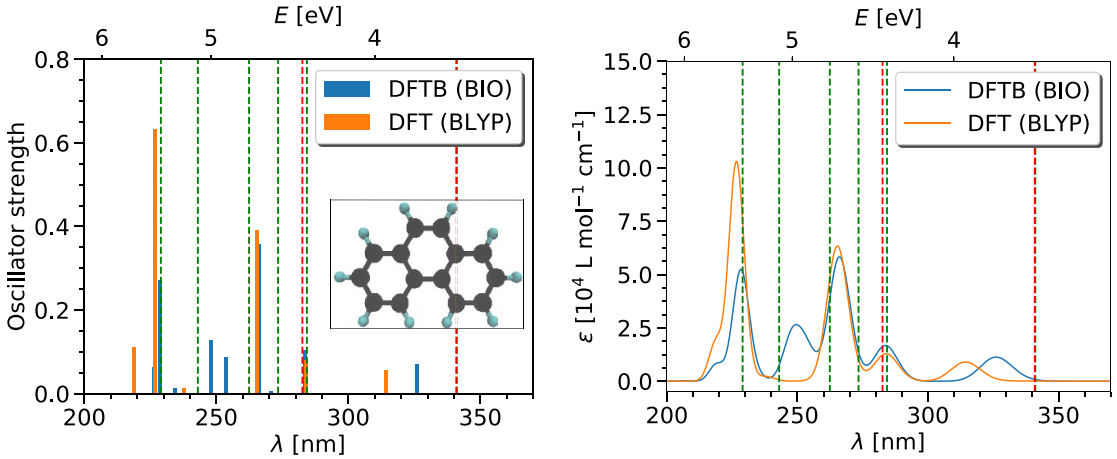

Figure 1. Absorption spectra (sticks in the left panel, convoluted in the right panel) of neutral phenanthrene computed with TD-DFT (orange) and TD-DFTB (blue) at the equilibrium geometry. All spectra have been truncated to 6.0 eV. Green and red vertical dashed lines indicate positions of some excited states from the experiments in Ne matrix [92] and in the supersonic jet [88, 89], respectively (see table 1). Inset in the left panel is the balls and sticks representation of phenanthrene ($C_{14}H_{10}$) where black and cyan balls correspond to carbon and hydrogen atoms, respectively. Adapted from Ref. [95].

3.2. SchNet training

The performance of several SchNet models is summarized in table 2. Notably, all of them achieve chemical accuracy (errors smaller than 1 kcal mol^{-1} or 0.043 eV) on the TD-DFTB energies of phenanthrene. This is not surprising, taking into account the outstanding performance of SchNet for prediction of both ground state [23] and excited state [37] properties. It is worth mentioning that the real accuracy of the machine-learned quantities may vary depending on the underlying electronic structure method that has been used for training. For example, SchNet errors have exceeded 1 kcal mol^{-1} for wavefunction-based CASSCF data [37]. However and more generally, errors are larger for the SchNet model #6, trained without forces, compared to model #3, which is consistent with previous studies [23, 96]. Furthermore, including forces in the SchNet training allows to optimize the model much faster due to a significant amount of information provided by $3M$ energy gradients. The error does not vary significantly between models with three and six interactions, which has already been pointed out by Schütt *et al* [23] for the QM9 [97–99] dataset. Yet, the difference in accuracy of SchNets with 1, 2 and 3 interactions can be more pronounced [1]. Notably, using 3000 training and 1000 validation points, model #5 with 256 features achieves similar accuracy as model #4

Table 2. Parameters of the considered SchNet models and the corresponding training times and errors. All models have used forces (except #6) and one NVIDIA® V100 GPU in the training. The MAEs (RMSEs) for energies and forces are given in eV and eV $Å^{-1}$, respectively. The MAE and RMSE have been computed for a test set (2000 structures for all models except #5 and #7 where the remaining 6000 and 3000 structures, respectively, have been used).

Model parameters							Dataset sizes	
#	State	Features	Interactions	Training set	Validation set	Runtime, h	Property	MAE (RMSE)
1	S_2	128	6	6000	2000	14	Energy	0.014 (0.018)
							Force	0.075 (0.126)
2	S_2	128	3	6000	2000	15	Energy	0.018 (0.023)
							Force	0.084 (0.138)
3	S_3	128	6	6000	2000	13	Energy	0.016 (0.022)
							Force	0.068 (0.113)
4	S_3	256	3	6000	2000	15	Energy	0.014 (0.018)
							Force	0.062 (0.106)
5	S_3	256	3	3000	1000	5	Energy	0.017 (0.022)
							Force	0.075 (0.124)
6	S_3	128	6	6000	2000	40	Energy	0.029 (0.038)
							Force	—
7	S_4	128	6	6000	1000	13	Energy	0.023 (0.030)
							Force	0.137 (0.239)

that has been trained with twice as much data. Overall, 128 features and three interactions seem to be a reasonable choice for the SchNet architecture. All the aforementioned conclusions indicate an importance of the hyperparameter tuning for practical applications of NNs. In fact, thoroughly crafted NNs can achieve remarkable accuracy even on relatively small training sets [70]. For all SchNet-based calculations presented below (sections 3.4 and 3.5), we have used models #1, #3 and #7 for S_2, S_3 and S_4 excited states, respectively.

3.3. FSSH/TD-DFTB simulations

In this subsection we assess the validity of the reference FSSH/TD-DFTB approach in the study of the radiationless electronic relaxation from the S_3 state of phenanthrene, by comparing the computed decay times with the experimental results from table 1. An ensemble of 250 trajectories has been launched and analyzed, which can be viewed in figure 2. Notably, the shapes of the average population curves change slightly as the number of trajectories is increased from 100 to 250. Also, the computed decay times are 745 and 860 fs for 100 and 250 trajectories, respectively. Hence, convergence might not be fully achieved with respect to ensemble size but should at least be good enough. The S_3 population curve can be roughly divided into two parts. The first one is located in the time window [0;400] fs and corresponds to a relatively rapid decay of the initial population from S_3 toward higher- and lower-lying states. It can be approximated by an exponential decay rate of 550–600 fs. The population transfer after 400 fs appears to be slower than the initial one. These results are in qualitative agreement with the complex multichannel relaxation, which is expected for phenanthrene according to [94]. As pointed out by Nazari *et al* [94], an internal conversion from the first bright singlet state consists of fast (100 fs) and slow (600 fs) contributions. The same group has also performed FSSH/TD-DFT simulations in the two lowest excited singlet states of a phenanthrene derivative. Their results are in qualitatively good agreement with experimental observations, despite neglecting the transitions toward higher-lying excited states. However, the major fraction (80%) of the population is rapidly (within less than 200 fs) transferred to the lowest excited state, in contrast with our FSSH/TD-DFTB simulations. Furthermore, the TD-DFT energy gaps (computed with ωB97XD and BHandHLYP functionals) [94] between the two lowest excited states are smaller than the ones in the CASPT2 [93, 94] or TD-DFTB (this work) spectra. This is consistent with faster and more complete (compared to the aforementioned FSSH/TD-DFTB results) relaxation reported by Nazari *et al* in [94].

Overall, the FSSH/TD-DFTB approach applied to the electronic relaxation of neutral phenanthrene is in reasonably good agreement with the available experimental data. It is worth mentioning that comparison of dynamical simulations with experiments must be carried out with caution. Indeed, experimental access to excited-state dynamics is most often not direct, depending on the experimental probes and setups, involving either transient spectroscopy or time-resolved photo-electron spectroscopy. In the latter case, the probe signal stems from the ionization and realistic comparison with experiment may require additional ingredients in the simulation such as ionization probabilities from excited states [100, 101].

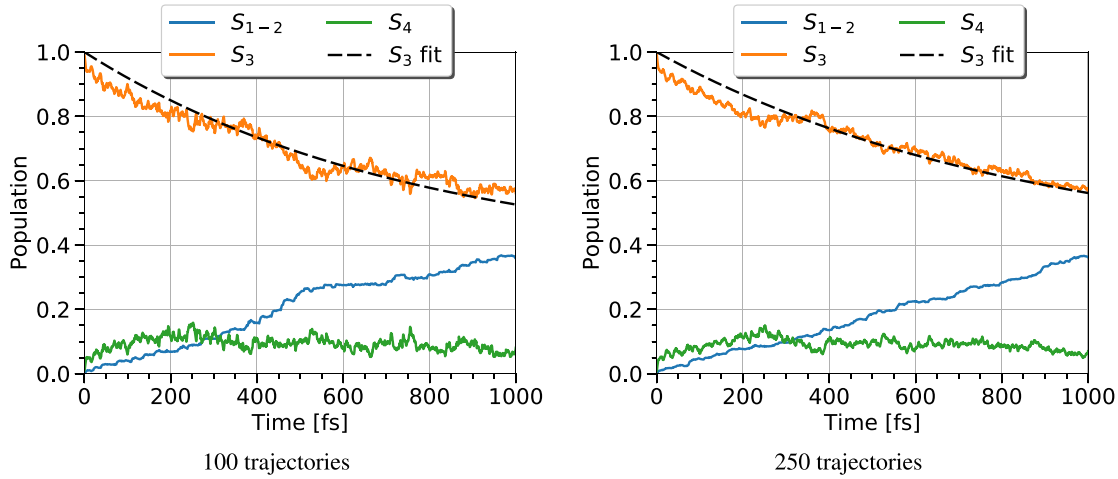

Figure 2. Populations of the low-lying singlet excited states in phenanthrene averaged over an ensemble of 100 (left panel) or 250 (right panel) FSSH/TD-DFTB trajectories. The initial state is S_3. Reproduced from Ref. [95].

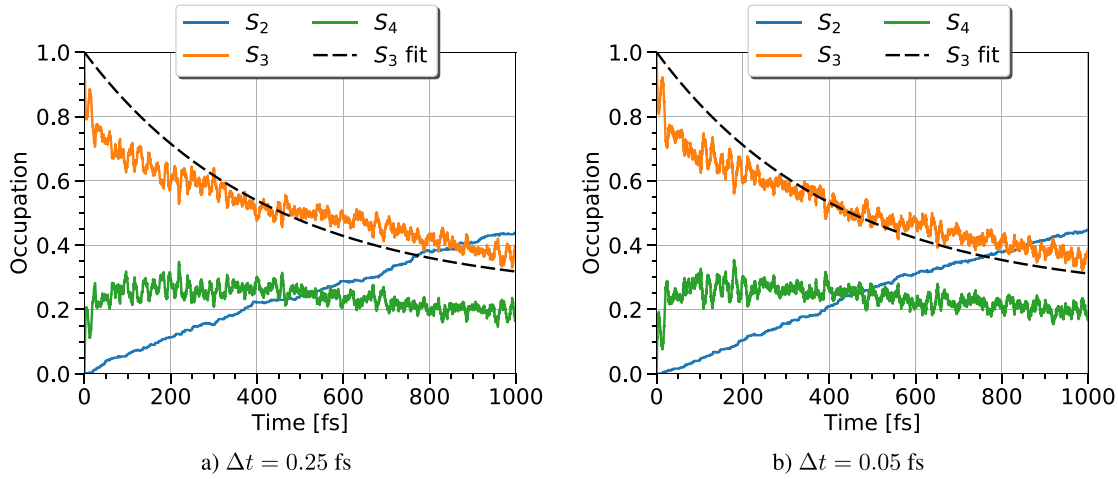

Figure 3. Occupations of the low-lying singlet excited states in phenanthrene averaged over an ensemble of 500 trajectories. The initial state is S_3. Each trajectory has been propagated with an indicated time step Δt on machine-learned PESs coupled to the BL hopping probability.

3.4. TSH/SchNet simulations

We start with the analysis of the TSH simulations with the BL hopping probability coupled to the SchNet-learned quantities (energies and forces) for two propagation steps, such as depicted in figure 3. Clearly, there are no significant differences in the calculated electronic relaxation. The computed decay time is about 416 and 421 fs with $\Delta t = 0.25$ and $\Delta t = 0.05$ fs, respectively, which means that the TSH dynamics is converged with time step $\Delta t = 0.25$ fs. The population transfer simulated with the BL/SchNet approach is slower than the one observed in the reference FSSH/TD-DFTB calculation, which is related to the fact that the number of BL-induced surface hops is significantly larger compared to that of FSSH [79]. It is worth mentioning another pitfall of the simplified TSH schemes that has been pointed out by Smith and Akimov [58], namely that BL-based approaches are well suited to study the dynamics in manifolds of states with energy gaps below 0.1–0.2 eV, but may fail for states separated by gaps $\geqslant$0.5 eV. Nonetheless in the present case, the BL/Schnet decay time seems to be in even closer agreement with the four experimental values in the range [0.42–0.52] ps (see table 1), with respect to the reference FSSH calculation. Yet, a direct comparison with our theoretical results may be only semi-quantitative, depending on the experimental probe mechanism (see discussion in the last paragraph of section 3.3).

Additional TSH calculations have been performed with the TD-DFTB energies and forces instead of the machine-learned ones. The goal was to assess whether the aforementioned results are a direct consequence of using ML-based quantities or is due to the particular TSH implementation presented in this work. Figure 4(b) shows the occupation dynamics for an ensemble of 100 trajectories propagated using TD-DFTB

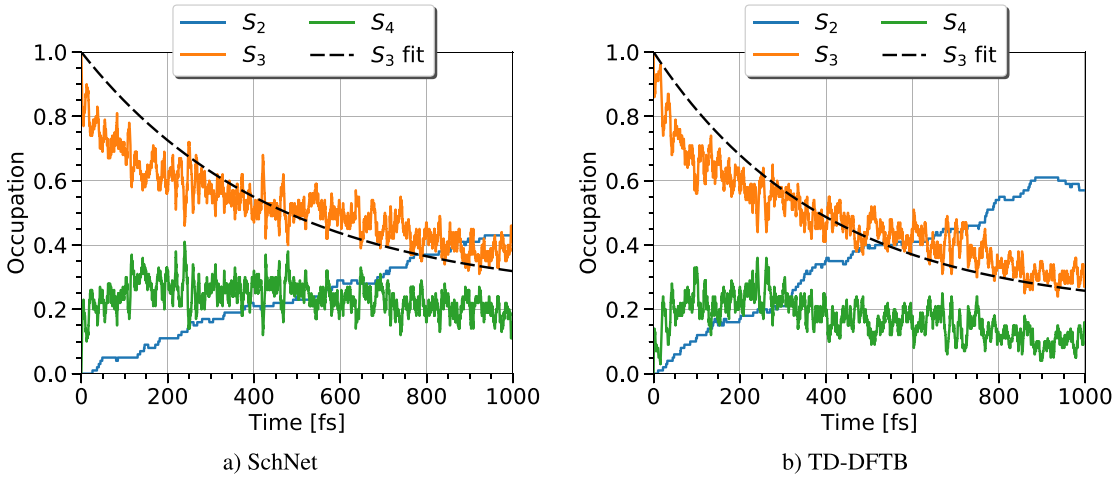

Figure 4. Occupations of the low-lying singlet excited states in phenanthrene averaged over an ensemble of 100 trajectories. Each trajectory has been propagated with the time step $\Delta t = 0.25$ fs on PES provided by (a) SchNet or (b) TD-DFTB coupled to the BL hopping probability.

energies and forces coupled to the BL hopping probability. The results are qualitatively similar to those obtained with the machine-learned quantities (see figure 4(a)), yet with more rapid population transfer from S_3 to S_2.

3.5. Discussion

To gain further insight into the performance of the TSH algorithms, we evaluated the dependence of the number of successful surface switches on the instantaneous adiabatic energy gap at the hopping event, which is compiled in figure 5. The number of hops peaks at small energy gaps of about 0.02 eV and about 0.01 eV for BL and FSSH methods, respectively. Notably, the BL-induced switches are more frequent than the FSSH ones, despite the constraint that switches are only possible at the local minimum of the adiabatic energy gap. This is also consistent with the previous study on excited-state dynamics of a phenol molecule [79].

We have performed additional analysis for one selected trajectory (see figure 6) in order to demonstrate the numerical issues associated with the use of single-precision SchNet models. This trajectory has been propagated in S_3 with $\Delta t = 0.05$ fs, coupled to TD-DFTB or SchNet (evaluated in single or double precision) for electronic structure calculations starting from the same set of initial conditions and with surface switches disabled. The adiabatic energy gap Z_{23} between the S_3 and S_2 states and its second time derivative $\ddot{Z}_{23}$ at each local minimum has been computed on the fly. As expected, the global energy gap evolution is not affected by the change of the SchNet precision. On the other hand, the distribution of $\ddot{Z}_{23}$ values is very different for trajectory runs based on SchNet. In particular, the number of SchNet-visited local minima (each point in the bottom row corresponds to the finite difference calculation at the detected local minimum) is significantly larger for the single precision case compared to either TD-DFTB or double precision propagation. As can be seen from the inset in figure 6, SchNet predicts several local minima when using single precision and zero minimum using double precision within the short considered period of time between 1.5 and 2.5 fs. Thus, we conclude that the TSH simulations should be performed with double-precision SchNet (see supplementary information for a more detailed discussion).

Finally, the computational costs are compared for the approaches considered in this study. A single FSSH-trajectory propagation with $\Delta t = 0.25$ fs (4000 steps) takes approximately 33 h when four states are included and 13 h with three states (mean values averaged over 200 trajectories). Notably, the ML-based propagation of three states takes only about 25–30 min with $\Delta t = 0.25$ fs (4000 steps) and about 65–70 min with $\Delta t = 0.1$ fs (10 000 steps), which is already remarkable gain of the CPU time. However, one should also take into account the time consumed for the training step of DL (SchNet) models, which varied between 4 and 15 h depending on both the number of training points and SchNet hyperparameters (see table 2). On the other hand, the training has to be done only once and the resulting SchNet model can be used in principal for an arbitrary number of trajectories. Thus, the TSH/SchNet simulations can be at least one order-of-magnitude faster than the FSSH/TD-DFTB ones while avoiding exponential scaling with respect to the number of excited states included in the propagation.

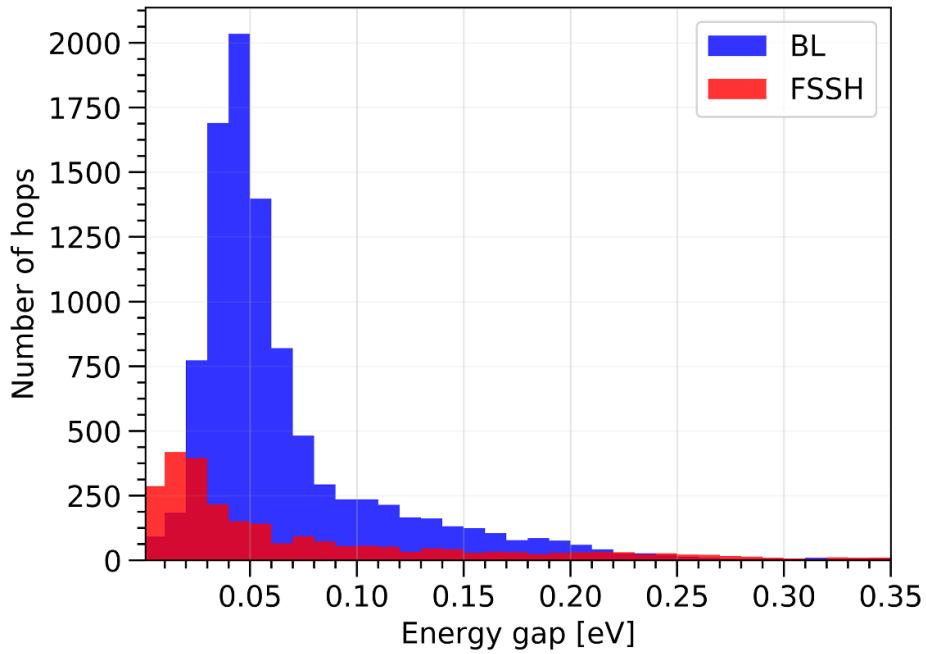

Figure 5. Distribution of the number of successful surface hops computed with the BL (blue) and Tully's FSSH (red) approaches *versus* the corresponding adiabatic energy gap. Data collected from 200 trajectories.

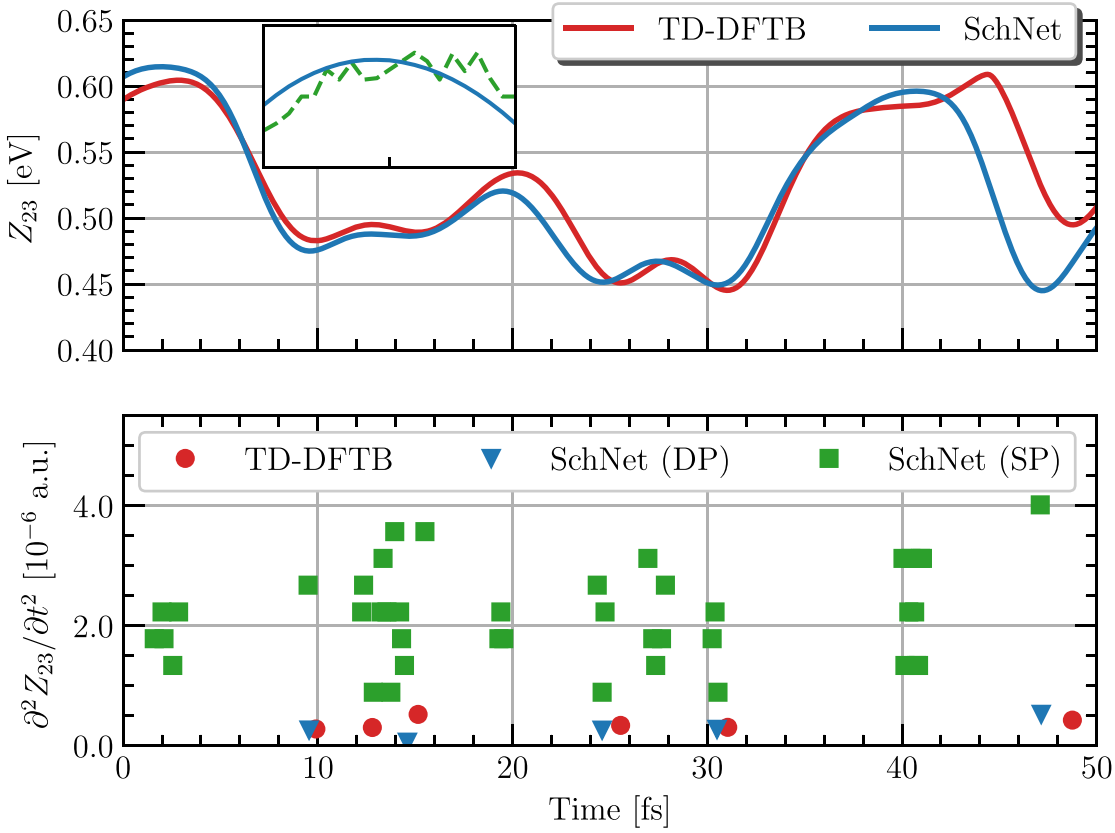

Figure 6. Adiabatic energy gap Z_{23} between the S_3 and S_2 states (top row) and its second time derivative at the local minima (bottom row) along the selected trajectory propagated with $\Delta t = 0.05$ fs computed with TD-DFTB (red), single-precision (SP) or double-precision (DP) SchNet models (green or blue, respectively). The inset corresponds to adiabatic energy gaps in the time window [1.5:2.5] fs computed with DP (solid blue line) or SP (green dashed line) SchNet models, see main text for discussion.

4. Conclusion

A detailed theoretical study dedicated to the application of DL to the non-adiabatic molecular dynamics of neutral phenanthrene, based on TD-DFTB calculations and simplified TSH switching probability has been presented. The results of the BL approach coupled to machine-learned PESs have been compared with those from Tully's FSSH/TD-DFTB approach and with available experimental data.

First of all, the accuracy of the TD-DFTB method has been assessed by calculating the absorption spectrum of neutral phenanthrene and upon comparison with higher-level electronic structure and available experimental data. The agreement for the low-lying excited singlet states is reasonably good. In a second step, the electronic relaxation from the bright S_3 state via the cascade of radiationless transitions has been investigated. The detailed analysis reveals good agreement between the simulated decay and the experimental results. Notably, the computed electronic relaxation in phenanthrene is significantly slower than what has been reported in our previous studies on polyacenes [55, 57], which is well correlated with the large energy gap between the initially excited S_3 state and the one below in energy in the TD-DFTB spectrum. This study brings some insight in the photophysics of phenanthrene and has been considered as a reference for the evaluation of simplified approaches to non-adiabatic molecular dynamics. The present results are of interest for astrochemistry and laboratory experiments concerned with atto- or femto-second laser spectroscopy of carbonaceous molecules. Consequently and due to the diversity of available experimental and theoretical results, phenanthrene can be viewed as a benchmark system to either study photophysical processes or test advanced theoretical tools.

The performance of several DL (SchNet) models trained on the TD-DFTB datasets has been evaluated. Overall, the SchNet architecture has been found able to accurately reproduce the complex multidimensional PESs of phenanthrene, especially when forces are included in the training. Notably, all of them achieve chemical accuracy (errors smaller than 1 kcal mol^{-1} or 0.043 eV) on the TD-DFTB energies of phenanthrene. The mean absolute errors could be lowered to less than 0.02 eV for the energy and 0.08 eV $Å^{-1}$ for the forces. One should precise the limits of the present conclusions. The real accuracy of the supervised ML methods is determined by the training data. Thus, the aforementioned SchNet models naturally inherit all advantages and disadvantages of the underlying TD-DFTB approach. The main drawback of some supervised ML algorithms is that there is no confidence in the predictions for inputs that lie beyond the space of training points. Nevertheless, this should not be an issue in this study since there is no fragmentation or isomerization involved.

Finally, a major focus has been given to the results computed with a simplified TSH scheme on SchNet-learned PESs. The convergence of the BL approach has been investigated. This TSH scheme coupled to SchNet for electronic structure calculations provides reliable results when SchNet is evaluated in double precision whereas the TSH dynamics propagated with small time steps fails to converge when the default single precision is used. The computed decay time for the S_3 state of phenanthrene is about 0.42 ps that is somewhat shorter than the decay in the reference FSSH calculations (0.55–0.60 ps). Nevertheless, both calculated decay times are in reasonably good agreement with the four experimental values in the range (0.42–0.52) ps (see table 1). Indeed, a direct comparison with our theoretical results may be only semi-quantitative, depending on the experimental probe mechanism. As far as we know, this study is the first to investigate the electronic relaxation in excited states based on machine-learned quantities derived for a system with more than 20 atoms.

The extreme acceleration provided by DL will prove invaluable for investigating larger systems with numerous excited states and/or to allow for longer lasting trajectories or a much denser sampling of the initial conditions. There is no doubt that simple hopping schemes can be of valuable interest in parallel with the alternative route consisting in the application of supervised ML models to fit the non-adiabatic couplings thus enabling the use of the Tully's FSSH approach [37, 40].

Data availability statement

Codes developed in this work are available on github.com/q-posev/ml-tsh. All TD-DFTB datasets, pre-trained SchNet models (#1, #3, #7 from table 2) and initial conditions used in this work are openly available in Ref. [102].

The data that support the findings of this study are openly available at the following URL/DOI: 10.5281/zenodo.4266393.

Acknowledgments

We acknowledge the European Union (EU) and Horizon 2020 funding awarded under the Marie Skłodowska-Curie action to the EUROPAH consortium, Grant No. 722346. This work was performed using HPC resources from CALMIP (Grant No. 2020-P18019). E P acknowledges kind hospitality during his stay at the Aarhus University. The authors are grateful to the referees for several valuable comments that contributed to significantly improve the manuscript.

Conflicts of interest

There are no conflicts to declare.

ORCID iDs

Evgeny Posenitskiy https://orcid.org/0000-0002-1623-0594
Fernand Spiegelman https://orcid.org/0000-0001-9412-2866

References

[1] Schütt K T, Arbabzadah F, Chmiela S, Müller K R and Tkatchenko A 2017 Quantum-chemical insights from deep tensor neural networks *Nat. Commun.* **8** 13890
[2] Deringer V L, Caro M A and Gábor C 2019 Machine learning interatomic potentials as emerging tools for materials science *Adv. Mater.* **31** 1902765
[3] Westermayr J and Marquetand P 2020 Machine learning and excited-state molecular dynamics *Mach. Learn.: Sci. Technol.* **1** 043001
[4] Dral P O 2020 Quantum chemistry in the age of machine learning *J. Phys. Chem. Lett.* **11** 2336–47
[5] Unke O T and Meuwly M 2019 Physnet: A neural network for predicting energies, forces, dipole moments and partial charges *J. Chem. Theory Comput.* **15** 3678–93
[6] Meuwly M Transformative applications of machine learning for chemical reactions (arXiv:2101.03530)
[7] Kaelbling L P, Littman M L, and Moore A W 1996 Reinforcement learning: a survey *CoRR* cs.AI/9605103
[8] Sørensen K H, Jørgensen M S, Bruix A and Hammer B 2018 Accelerating atomic structure search with cluster regularization *J. Chem. Phys.* **148** 241734
[9] Jørgensen M S, Groves M N and Hammer B 2017 Combining evolutionary algorithms with clustering toward rational global structure optimization at the atomic scale *J. Chem. Theory Comput.* **13** 1486–93
[10] Jacobsen T L, Jørgensen M S and Hammer B 2018 On-the-fly machine learning of atomic potential in density functional theory structure optimization *Phys. Rev. Lett.* **120** 026102
[11] Bisbo M K and Hammer B 2020 Efficient global structure optimization with a machine-learned surrogate model *Phys. Rev. Lett.* **124** 086102
[12] Jørgensen M S, Mortensen H L, Meldgaard S A, Kolsbjerg E L, Jacobsen T L, Sorensen K H and Hammer B 2019 Atomistic structure learning *J. Chem. Phys.* **151** 054111
[13] Gasteiger J and Zupan J 1993 Neural networks in chemistry *Angewandte Chemie Int. Edn English* **32** 503–27
[14] Cybenko G 1989 Approximation by superpositions of a sigmoidal function *Math. Control Signals Syst.* **2** 303–14
[15] Hornik K, Stinchcombe M and White H 1989 Multilayer feedforward networks are universal approximators *Neural Netw.* **2** 359–66
[16] Behler J and Parrinello M 2007 Generalized neural-network representation of high-dimensional potential-energy surfaces *Phys. Rev. Lett.* **98** 146401
[17] Gastegger M, Behler J and Marquetand P 2017 Machine learning molecular dynamics for the simulation of infrared spectra *Chem. Sci.* **8** 6924–35
[18] Shakouri K, Behler J, Meyer J and Kroes G-J 2017 Accurate neural network description of surface phonons in reactive gas–surface dynamics: N2 + ru(0001) *J. Phys. Chem. Lett.* **8** 2131–6
[19] Zhang Y, Zhou X and Jiang B 2019 Bridging the gap between direct dynamics and globally accurate reactive potential energy surfaces using neural networks *J. Phys. Chem. Lett.* **10** 1185–91
[20] Chiriki S, Jindal S and Bulusu S S 2017 Neural network potentials for dynamics and thermodynamics of gold nanoparticles *J. Chem. Phys.* **146** 084314
[21] Zhenwei Li, Kermode J R and Alessandro D V 2015 Molecular dynamics with on-the-fly machine learning of quantum-mechanical forces *Phys. Rev. Lett.* **114** 096405
[22] Zhang L, Han J, Wang H, Car R and Weinan E 2018 Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics *Phys. Rev. Lett.* **120** 143001
[23] Schütt K T, Sauceda H E, Kindermans P-J, Tkatchenko A and Müller K-R 2018 SchNet—a deep learning architecture for molecules and materials *J. Chem. Phys.* **148** 241722
[24] Ramakrishnan R, Hartmann M, Tapavicza E and von Lilienfeld O A 2015 Electronic spectra from TDDFT and machine learning in chemical space *J. Chem. Phys.* **143** 084111
[25] Zhu J, Vuong V Q, Sumpter B G and Irle S 2019 Artificial neural network correction for density-functional tight-binding molecular dynamics simulations *MRS Commun.* **9** 867–73
[26] Ramakrishnan R, Dral P O, Rupp M and von Lilienfeld O A 2015 Big data meets quantum chemistry approximations: the δ-machine learning approach *J. Chem. Theory Comput.* **11** 2087–96
[27] Dral P O 2019 MLatom: a program package for quantum chemical research assisted by machine learning *J. Comput. Chem.* **40** 2339–47

[28] Wang H, Zhang L, Han J and Weinan E 2018 Deepmd-kit: a deep learning package for many-body potential energy representation and molecular dynamics *Comput. Phys. Commun.* **228** 178–84
[29] Schütt K T, Kessel P, Gastegger M, Nicoli K A, Tkatchenko A and Müller K-R 2019 SchNetPack: a deep learning toolbox for atomistic systems *J. Chem. Theory Comput.* **15** 448–55
[30] Meyer H-D, Manthe U and Cederbaum L S 1990 The multi-configurational time-dependent Hartree approach *Chem. Phys. Lett.* **165** 73–8
[31] Beck M H, Jäckle A, Worth G A and Meyer H-D 2000 The multiconfiguration time-dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets *Phys. Rep.* **324** 1–105
[32] Ehrenfest P 1927 Bemerkung über die angenäherte gültigkeit der klassischen mechanik innerhalb der quantenmechanik *Z. Phys.* **45** 455–7
[33] Tully J C 1990 Molecular dynamics with electronic transitions *J. Chem. Phys.* **93** 1061–71
[34] Fabiano E, Keal T W and Thiel W 2008 Implementation of surface hopping molecular dynamics using semiempirical methods *Chem. Phys.* **349** 334–47
[35] Crespo-Otero R and Barbatti M 2018 Recent advances and perspectives on nonadiabatic mixed quantum–classical dynamics *Chem. Rev.* **118** 7026–68
[36] Mai S, Marquetand P and Leticia G 2018 Nonadiabatic dynamics: the SHARC approach *WIREs Comput. Mol. Sci.* **8** e1370
[37] Westermayr J, Gastegger M and Marquetand P 2020 Combining SchNet and SHARC: the SchNarc machine learning approach for excited-state dynamics *J. Phys. Chem. Lett.* **11** 3828–34
[38] Chen W-K, Liu X-Y, Fang W-H, Dral P O and Cui G 2018 Deep learning for nonadiabatic excited-state dynamics *J. Phys. Chem. Lett.* **9** 6702–8
[39] Hu D, Xie Y, Li X, Li L and Lan Z 2018 Inclusion of machine learning kernel ridge regression potential energy surfaces in on-the-fly nonadiabatic molecular dynamics simulation *J. Phys. Chem. Lett.* **9** 2725–32
[40] Dral P O, Barbatti M and Thiel W 2018 Nonadiabatic excited-state dynamics with machine learning *J. Phys. Chem. Lett.* **9** 5660–3
[41] Westermayr J, Michael Gastegger M F Menger S J, Mai S, González L and Marquetand P 2019 Machine learning enables long time scale molecular photodynamics simulations *Chem. Sci.* **10** 8100–7
[42] Westermayr J, Faber F A, Christensen A S, von Lilienfeld O A and Marquetand P 2020 Neural networks and kernel ridge regression for excited states dynamics of $CH_2NH_2^+$: from single-state to multi-state representations and multi-property machine learning models *Mach. Learn.: Sci. Technol.* **1** 025009
[43] Krems R V 2019 Bayesian machine learning for quantum molecular dynamics *Phys. Chem. Chem. Phys.* **21** 13392–410
[44] Richings G W and Habershon S 2017 Direct quantum dynamics using grid-based wave function propagation and machine-learned potential energy surfaces *J. Chem. Theory Comput.* **13** 4012–24
[45] Westermayr J and Marquetand P 2020 Machine learning for electronically excited states of molecules *Chem. Rev.* (https://doi.org/10.1021/acs.chemrev.0c00749)
[46] Casida M E 1995 *Time-Dependent Density Functional Response Theory for Molecules* pp 155–92
[47] Porezag D, Frauenheim T, Köhler T, Seifert G and Kaschner R 1995 Construction of tight-binding-like potentials on the basis of density-functional theory: application to carbon *Phys. Rev.* B **51** 12947–57
[48] Seifert G, Porezag D and Frauenheim T 1996 Calculations of molecules, clusters and solids with a simplified LCAO-DFT-LDA scheme *Int. J. Quantum Chem.* **58** 185–92
[49] Elstner M and Seifert G 2014 Density functional tight binding *Phil. Trans. R. Soc.* A **372** 20120483
[50] Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S and Seifert G 1998 Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties *Phys. Rev.* B **58** 7260–8
[51] Niehaus T A, Suhai S, Della Sala F, Lugli P, Elstner M, Seifert G and Frauenheim T 2001 Tight-binding approach to time-dependent density-functional response theory *Phys. Rev.* B **63** 085108
[52] Pal S, Trivedi D J, Akimov A V, Aradi B, Frauenheim T and Prezhdo O V 2016 Nonadiabatic molecular dynamics for thousand atom systems: a tight-binding approach toward PYXAID *J. Chem. Theory Comput.* **12** 1436–48
[53] Humeniuk A and Roland M 2017 DFTBaby: a software package for non-adiabatic molecular dynamics simulations based on long-range corrected tight-binding TD-DFT(B) *Comput. Phys. Commun.* **221** 174–202
[54] Stojanović L, Aziz S G, Hilal R H, Plasser F, Niehaus T A and Barbatti M 2017 Nonadiabatic dynamics of cycloparaphenylenes with TD-DFTB surface hopping *J. Chem. Theory Comput.* **13** 5846–60
[55] Posenitskiy E, Rapacioli M, Lepetit B, Lemoine D and Spiegelman F 2019 Non-adiabatic molecular dynamics investigation of the size dependence of the electronic relaxation in polyacenes *Phys. Chem. Chem. Phys.* **21** 12139–49
[56] Uratani H and Nakai H 2020 Non-adiabatic molecular dynamics with divide-and-conquer type large-scale excited-state calculations *J. Chem. Phys.* **152** 224109
[57] Posenitskiy E, Rapacioli M, Lemoine D and Spiegelman F 2020 Theoretical investigation of the electronic relaxation in highly excited chrysene and tetracene: the effect of armchair vs zigzag edge *J. Chem. Phys.* **152** 074306
[58] Smith B and Akimov A V 2020 Hot electron cooling in silicon nanoclusters via Landau–Zener nonadiabatic molecular dynamics: size dependence and role of surface termination *J. Chem. Theory Comput.* **11** 1456–65
[59] Zener C and Fowler R H 1932 Non-adiabatic crossing of energy levels *Proc. R. Soc.* A **137** 696–702
[60] Ter Haar D 1965 9—a theory of energy transfer. ii *Collected Papers of L.D. Landau*, ed (Oxford: Pergamon) pp 63–6
[61] Belyaev A K and Lebedev O V 2011 Nonadiabatic nuclear dynamics of atomic collisions based on branching classical trajectories *Phys. Rev.* A **84** 014701
[62] Zhu C and Nakamura H 1992 The two-state linear curve crossing problems revisited. ii. Analytical approximations for the stokes constant and scattering matrix: the Landau–Zener case *J. Chem. Phys.* **97** 8497–514
[63] Zhu C and Nakamura H 1993 The two-state linear curve crossing problems revisited. iii. Analytical approximations for stokes constant and scattering matrix: nonadiabatic tunneling case *J. Chem. Phys.* **98** 6208–22
[64] Yu L, Xu C, Lei Y, Zhu C and Wen Z 2014 Trajectory-based nonadiabatic molecular dynamics without calculating nonadiabatic coupling in the avoided crossing case: trans–cis photoisomerization in azobenzene *Phys. Chem. Chem. Phys.* **16** 25883–95
[65] Hanasaki K, Kanno M, Niehaus T A and Kono H 2018 An efficient approximate algorithm for nonadiabatic molecular dynamics *J. Chem. Phys.* **149** 244117
[66] Furche F and Ahlrichs R 2002 Adiabatic time-dependent density functional methods for excited state properties *J. Chem. Phys.* **117** 7433–47

[67] Furche F and Ahlrichs R 2004 Erratum: 'adiabatic time-dependent density functional methods for excited state properties' [j. chem. phys. 117, 7433 (2002)]
J. Chem. Phys. **121** 12772–3
[68] Heringer D, Niehaus T A, Wanko M and Frauenheim T 2007 Analytical excited state forces for the time-dependent density-functional tight-binding method *J. Comput. Chem.* **28** 2589–601
[69] Heringer D, Niehaus T A, Wanko M and Frauenheim T 2011 Erratum: 'analytical excited state forces for the time-dependent density-functional tight-binding method' [j. comp. chem. 28, 2589] *J. Comput. Chem.* **33** 593
[70] Schütt K, Kindermans P-J, Felix H E S, Chmiela S, Tkatchenko A and Müller K-R *et al* 2017 SchNet: a continuous-filter convolutional neural network for modeling quantum interactions *Advances in Neural Information Processing Systems 30*, ed I Guyon (Red Hook, NY: Curran Associates, Inc.) pp 991–1001
[71] Pukrittayakamee A, Malshe M, Hagan M, Raff L M, Narulkar R, Bukkapatnum S and Komanduri R 2009 Simultaneous fitting of a potential-energy surface and its corresponding force fields using feedforward neural networks *J. Chem. Phys.* **130** 134101
[72] Behler J 2015 Constructing high-dimensional neural network potentials: a tutorial review *Int. J. Quantum Chem.* **115** 1032–50
[73] Hammes-Schiffer S and Tully J C 1994 Proton transfer in solution: molecular dynamics with quantum transitions *J. Chem. Phys.* **101** 4657–67
[74] Granucci G and Persico M 2007 Critical appraisal of the fewest switches algorithm for surface hopping *J. Chem. Phys.* **126** 134114
[75] Subotnik J E, Ouyang W and Landry B R 2013 Can we derive Tully's surface-hopping algorithm from the semiclassical quantum Liouville equation? Almost, but only with decoherence *J. Chem. Phys.* **139** 214107
[76] Heine T *et al* 2009 Demonnano (available at: http://demon-nano.ups-tlse.fr/)
[77] Werner U, Mitrić R, Suzuki T and Bonačić-Koutecký V 2008 Nonadiabatic dynamics within the time dependent density functional theory: ultrafast photodynamics in pyrazine *Chem. Phys.* **349** 319–24
[78] Granucci G, Persico M and Toniolo A 2001 Direct semiclassical simulation of photochemical processes with semiempirical wave functions *J. Chem. Phys.* **114** 10608–15
[79] Xie W and Domcke W 2017 Accuracy of trajectory surface-hopping methods: test for a two-dimensional model of the photodissociation of phenol *J. Chem. Phys.* **147** 184114
[80] Bonhommeau D, Viel A and Halberstadt N 2005 Dissociative ionization of neon clusters nen, n=3 to 14: a realistic multisurface dynamical study *J. Chem. Phys.* **123** 054316
[81] Ermolaev V L 2001 Ultrafast nonradiative transitions between higher excited states in organic molecules *Russ. Chem. Rev.* **70** 471–90
[82] Brechignac P and Hermine P 1994 Photophysical studies of jet-cooled PAHs: emission spectra, lifetimes and van der Waals clusters astrophysical implications *AIP Conf. Proc* vol 312 pp 613–28
[83] Frisch M J *et al* 2009 *Gaussian 09 (Revision D.01)* (Wallingford, CT: Gaussian Inc.)
[84] Wanko M, Garavelli M, Bernardi F, Niehaus T A, Frauenheim T and Elstner M 2004 A global investigation of excited state surfaces within time-dependent density-functional response theory *J. Chem. Phys.* **120** 1674–92
[85] Fihey A and Jacquemin D 2019 Performances of density functional tight-binding methods for describing ground and excited state geometries of organic molecules *J. Chem. Theory Comput.* **15** 6267–76
[86] Kingma D P and Ba J 2017 Adam: a method for stochastic optimization (arXiv:1412.6980)
[87] Larsen A H *et al* 2017 The atomic simulation environment—a Python library for working with atoms *J. Phys.: Condens. Matter.* **29** 273002
[88] Amirav A, Sonnenschein M and Jortner J 1984 Statistical-limit line broadening in the s2 state of phenanthrene in supersonic jets *J. Phys. Chem.* **88** 5593–6
[89] Ohta N and Baba H 1986 Electronic spectra and intramolecular dynamics of phenanthrene in a supersonic free jet *Mol. Phys.* **59** 921–39
[90] Blanchet V, Lochbrunner S, Schmitt M, Shaffer J P, Larsen J J, Zgierski M Z, Seideman T and Stolow A 2000 Towards disentangling coupled electronic–vibrational dynamics in ultrafast non-adiabatic processes *Faraday Discuss.* **115** 33–48
[91] Dick B and Nickel B 1986 Determination of femtosecond lifetimes of higher excited singlet states by means of transient photophysical hole-burning. The s2 state of phenanthrene *Chem. Phys.* **110** 131–43
[92] Salama F, Joblin C and Allamandola L J 1994 Electronic absorption spectroscopy of matrix-isolated polycyclic aromatic hydrocarbon cations. ii. The phenanthrene cation (C14H10+) and its 1-methyl derivative *J. Chem. Phys.* **101** 10252–62
[93] González-Luque R, Serrano-Andrés L, Merchán M and Fülscher M P 2003 Theoretical characterization of the absorption spectra of phenanthrene and its radical cation *Theor. Chem. Acc.* **110** 224–32
[94] Nazari M *et al* 2019 Ultrafast dynamics in polycyclic aromatic hydrocarbons: the key case of conical intersections at higher excited states and their role in the photophysics of phenanthrene monomer *Phys. Chem. Chem. Phys.* **21** 16981–8
[95] Posenitskiy E 2020 Non-adiabatic molecular dynamics of PAH-related complexes Theses, Université Paul Sabatier - Toulouse III
[96] Chmiela S, Alexandre Tkatchenko H E, Sauceda I P, Schütt K T and Müller K-R 2017 Machine learning of accurate energy-conserving molecular force fields *Sci. Adv.* **3** e1603015
[97] Reymond J-L 2015 The chemical space project *Acc. Chem. Res.* **48** 722–30
[98] Blum L C and Reymond J-L 2009 970 million druglike small molecules for virtual screening in the chemical universe database GDB-13 *J. Am. Chem. Soc.* **131** 8732–3
[99] Ramakrishnan R, Dral P, Rupp M and Anatole von L 2014 Quantum chemistry structures and properties of 134 kilo molecules *Sci. Data* **1** 140022
[100] Tao H *et al* 2011 Ultrafast internal conversion in ethylene. i. The excited state lifetime *J. Chem. Phys.* **134** 244306
[101] Ruckenbauer M, Mai S, Marquetand P and Leticia G 2016 Revealing deactivation pathways hidden in time-resolved photoelectron spectra *Sci. Rep.* **6** 35522
[102] Posenitskiy E 2020 Phenanthrene: TD-DFTB datasets, pre-trained SchNet models and initial coniditions for TSH (https://doi.org/10.5281/zenodo.4266393)