HAL
open science

Non-adiabatic molecular dynamics of PAH-related complexes

Evgeny Posenitskiy

To cite this version:

Evgeny Posenitskiy. Non-adiabatic molecular dynamics of PAH-related complexes. Quantum Physics [quant-ph]. Université Toulouse 3 - Paul Sabatier, 2020. English. NNT : . tel-02983135v1

HAL Id: tel-02983135

https://theses.hal.science/tel-02983135v1

Submitted on 29 Oct 2020 (v1), last revised 5 Jan 2021 (v2)

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

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

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Evgeny POSENITSKIY

Le 29 septembre 2020

Dynamique moléculaire non-adiabatique des complexes de type PAH

Ecole doctorale : **SDM - SCIENCES DE LA MATIERE - Toulouse**

Spécialité : **Physique**

Unité de recherche :
LCAR-IRSAMC - Laboratoire Collisions Agrégats Réactivité

Thèse dirigée par
Didier LEMOINE et Fernand SPIEGELMAN

Jury
Mme Valérie VALLET, Rapporteure
M. Rodolphe VUILLEUMIER, Rapporteur
M. Thomas NIEHAUS, Examinateur
Mme Christine JOBLIN, Examinatrice
M. Didier LEMOINE, Directeur de thèse
M. Fernand SPIEGELMAN, Co-directeur de thèse
M. Mathias RAPACIOLI, Invité

Non-adiabatic molecular dynamics of PAH-related complexes

Evgeny Posenitskiy

Toulouse, France
Submitted: **August 2020**
Defended: **September 2020**

Abstract

Polycyclic Aromatic Hydrocarbons (PAHs) have been proposed as main carriers of diffuse interstellar bands that are observed in the interstellar medium. This has motivated an extensive study of their photophysical and photochemical response to UV irradiation. Underlying competing mechanisms drive the evolution of gas in the interstellar medium. The main objective of this thesis is to describe and to get theoretical insight in the energy relaxation mechanisms in large PAH molecules *via* extensive non-adiabatic molecular dynamics simulations coupled to the linear response Time-Dependent Density Functional based Tight Binding (TD-DFTB) approach of the excited states.

Prerequisite substantial development was made in the DFTB deMon-Nano package (http://demon-nano.ups-tlse.fr), firstly with the implementation of analytical gradients of potential energy surfaces (PESs) and of non-adiabatic couplings within the TD-DFTB scheme. Next, the Tully's fewest-switches trajectory surface hopping (FSSH) algorithm has been adapted and coupled to the TD-DFTB scheme in order to take into account non-adiabatic transitions. After detailed methodological considerations and comparison with higher-level electronic structure methods, the first full-scale application is dedicated to non-adiabatic molecular dynamics of linearly cata-condensed PAHs. Electronic relaxation from the brightest excited state has been simulated for neutral polyacenes with 2 to 7 aromatic cycles. The results display a striking alternation in decay times of the brightest singlet state computed for polyacenes with up to 6 aromatic cycles, which is correlated with a qualitatively similar alternation of energy gaps between the brightest state and the state lying just below in energy. Next, the influence of geometry on relaxation has been investigated through the comparison of two isomers: armchair-edge chrysene *versus* zigzag-edge tetracene. After assessing the performance of DFTB parameter sets, the main focus is given to the analysis of the electronic relaxation from the brightest excited state, which is located around 270 nm for both isomers. The results show that the electronic population of the brightest excited state in chrysene decays an order-of-magnitude faster than that in tetracene. This is correlated with a significant difference in energy gaps between the brightest state and the state lying just below in energy, which is consistent with the previous conclusions for polyacenes.

A last major development concerns the use of Machine Learning (ML) algorithms that have been proposed as a way to avoid most of the computationally-demanding electronic structure calculations. It aims to assess the performance of neural networks algorithms applied to excited-state dynamics. Electronic relaxation in neutral phenanthrene has been chosen as a

test case due to the diversity of available experimental results. Several neural networks have been trained with different parameters and their respective accuracy and efficiency analyzed. In addition, approximate trajectory surface hopping schemes have been interfaced to ML-based PESs and gradients, resulting in non-adiabatic dynamics simulations at a negligible cost. Various simplified hopping approaches have been compared with FSSH. Overall, ML is found to be a highly promising tool for nanosecond-long molecular dynamics in excited states.

This PhD research opens new avenues to investigate theoretical photophysics of large molecular complexes. Last but not least, the theoretical tools developed and implemented in deMon-Nano in a modular way can be further combined with other advanced (such as Configuration Interaction) DFTB techniques better adapted to charge-transfer states.

Résumé

Les hydrocarbures aromatiques polycycliques (PAH) ont été proposés comme porteurs principaux de bandes interstellaires diffuses observées dans le milieu interstellaire, motivant des études approfondies de leur réponse photophysique et photochimique au rayonnement UV. Les mécanismes sous-jacents en compétition déterminent l'évolution du gaz dans le milieu interstellaire. L'objectif principal de cette thèse est de décrire et de comprendre les mécanismes de relaxation dans des PAHs de grande taille, par des simulations de dynamique moléculaire non-adiabatique, couplées à l'approche de la réponse linéaire "Time-Dependent Density Functional based Tight Binding" (TD-DFTB) des états excités.

Des développements substantiels, prérequis ont été effectués dans le code DFTB deMon-Nano (http://demon-nano.ups-tlse.fr), d'abord avec le calcul des gradients analytiques des surfaces d'énergie potentielle (PES) et des couplages non-adiabatiques des états TD-DFTB. Puis, l'algorithme de trajectoire à sauts de surface minimaux (FSSH) de Tully a été adapté à l'approche TD-DFTB afin de prendre en compte les effets non-adiabatiques. Après comparaison avec des méthodes de structure électronique de référence, la première application est dédiée à la dynamique non-adiabatique de PAHs cata-condensés linéairement. La relaxation électronique de l'état excité le plus brillant a été simulée pour des polyacènes neutres constitués de 2 à 7 cycles aromatiques. Les résultats montrent une alternance marquée dans les temps de dépopulation de l'état initial pour les polyacènes contenant jusqu'à 6 cycles aromatiques, ce qui est corrélé avec une alternance des écarts d'énergie entre l'état initial et l'état situé juste dessous. Puis, l'influence de la géométrie sur la relaxation a été étudiée en comparant deux isomères, le chrysène de type "armchair-edge" et le tétracène de type "zigzag-edge". Après évaluation des paramétrages DFTB, la relaxation électronique à partir de l'état excité le plus brillant, situé autour de 270 nm pour les deux isomères, à été analysée. Les résultats montrent que la population électronique excitée du chrysène décroît un ordre de grandeur plus rapidement que celle du tétracène. Ceci est aussi corrélé à une différence significative des écarts d'énergie entre l'état initial et l'état situé juste dessous.

Un dernier développement majeur concerne l'utilisation d'algorithmes "Machine Learning" (ML) proposés comme un moyen d'éviter la plupart des calculs de structure électronique, très coûteux en temps calcul. Les performances d'algorithmes de réseaux de neurones appliqués à la dynamique des états excités ont été évaluées. Le cas de la relaxation électronique dans le phénanthrène neutre a été choisi comme test en raison de divers résultats expérimentaux disponibles. L'apprentissage de plusieurs réseaux de neurones a été effectué et leurs précision

et efficacité analysés. De plus, des approximations de trajectoires à sauts de surface ont été interfacées à l'approche ML, résultant en un coût négligeable des simulations de dynamique non-adiabatique. L'efficacité des diverses approches simplifiées a été comparée à FSSH. Dans l'ensemble, ML se révèle un outil très prometteur pour la dynamique dans les états excités à l'échelle de la nanoseconde.

Ce travail de thèse ouvre de nouvelles voies pour étudier la photophysique théorique de complexes moléculaires de grande taille. Enfin, les outils développés et implémentés dans deMon-Nano, de manière modulaire, peuvent être combinés avec d'autres approches DFTB sophistiquées (tel que "Configuration Interaction") plus adaptées aux états à transfert de charge.

Contents

Page

Abstract **iii**

Résumé **v**

1 Introduction **1**

1.1 A few words about photophysics . . . 2

1.2 The PAH hypothesis . . . 5

1.2.1 Laboratory spectroscopy of PAHs . . . 7

1.3 Theoretical considerations . . . 7

1.3.1 Non-adiabatic effects (or a few more words about photophysics) . . . 9

1.3.2 Bridging dynamical simulations and spectroscopic observations . . . 11

1.4 Objectives and outline . . . 12

2 Methods **13**

2.1 Density Functional Theory (DFT) . . . 14

2.1.1 The Kohn-Sham approach . . . 15

2.1.2 Density Functional based Tight-Binding (DFTB) . . . 17

2.2 Time-Dependent DFT(B) . . . 21

2.2.1 Linear Response TD-DFT(B) . . . 22

2.2.2 Analytical gradients for TD-DFT(B) . . . 25

2.3 Non-adiabatic molecular dynamics . . . 27

2.3.1 Fewest-Switches Trajectory Surface Hopping (FSSH) . . . 28

2.3.2 FSSH coupled to TD-DFTB . . . 32

3 Non-adiabatic molecular dynamics of large cata-condensed PAHs **35**

3.1 Introduction . . . 35

3.2 Analysis of the absorption spectra . . . 39

3.2.1 Polyacenes: from naphthalene to heptacene . . . 41

3.2.2 Tetracene *versus* chrysene . . . 42

3.3 Size dependence of the ultrafast relaxation . . . 44

3.3.1 Polyacenes: odd number of aromatic cycles . . . 45

3.3.2 Polyacenes: even number of aromatic cycles . . . 46

3.3.3 Discussion . . . 47

3.4 Shape dependence of the ultrafast relaxation 52
3.4.1 FSSH dynamics of highly excited tetracene *versus* chrysene. 52
3.4.2 Discussion 53
3.5 Conclusion 56

4 Testing Machine Learning potentials for simplified non-adiabatic dynamics 59
4.1 Introduction 59
4.2 Neural Networks. 62
4.3 Related work. 67
4.4 Computational details 68
4.5 Results 72
4.5.1 Absorption spectra of phenanthrene 73
4.5.2 Training SchNet 75
4.5.3 FSSH/TD-DFTB simulations 77
4.5.4 TSH/SchNet simulations 80
4.6 Conclusion 85

5 Conclusion 87

A Example of the deMon-Nano input 91

B Supplementary spectroscopic data 93

C Supplementary population analysis 97

D Interface with the Atomic Simulation Environment (ASE) 99

Bibliography 103

List of Publications 123

List of Figures 124

List of Tables 133

Résumé étendu en français 135

Acknowledgements 157

1

Introduction

One of the main objectives in astrophysics is to understand and describe the evolution of the universe and galaxies within. Modern models assume that this evolution is mostly governed by balance between the formation and destruction of stars within each galaxy.[1]

Stars are formed from collapsing clouds of condensed gas and dust, the so-called molecular clouds, and evolve differently depending on their mass. For instance, the existence of a massive star ends with an explosion called supernova. This is followed by release of the heavy elements that had been formed inside the star. Evolution of stars that are not that massive is governed by the stellar winds that drag away molecules and dust from the circumstellar environments into the Interstellar Medium (ISM). The matter travels in the ISM and is gathered into denser clouds, namely the molecular clouds. Once a critical mass is accumulated, the molecular cloud may collapse and form stars. This couples the life cycle of the stars to the evolution of gas and dust, thus forming a closed loop that is schematically represented in Figure 1.1.

The ISM of our galaxy consists of two parts: gas (about 90% hydrogen, 10% helium and a small fraction of carbon, oxygen, nitrogen and some other elements) and dust nanoparticles (mostly silicates and carbon). In dense molecular clouds, the gas temperature varies from ∼1000 K in the external layers illuminated by UV photons to ∼10 K in the inner regions. This change in temperature is mostly governed by the penetration depth of the stellar UV radiation. These photodissication regions[2] (also called photon-dominated regions or PDRs) associated with star forming areas gained a lot of attention in recent years. In the cold inner regions, the H_2 and CO survive and can be used to track the evolution of molecular gas. On the contrary, UV radiation may dissociate these molecules and ionize released fragments (for example, forming C^+) in the warmer regions. All the aforementioned regions represent a unique playground for astrochemistry and astrophysics research that aims to understand the evolution of the ISM but also to study molecular processes.

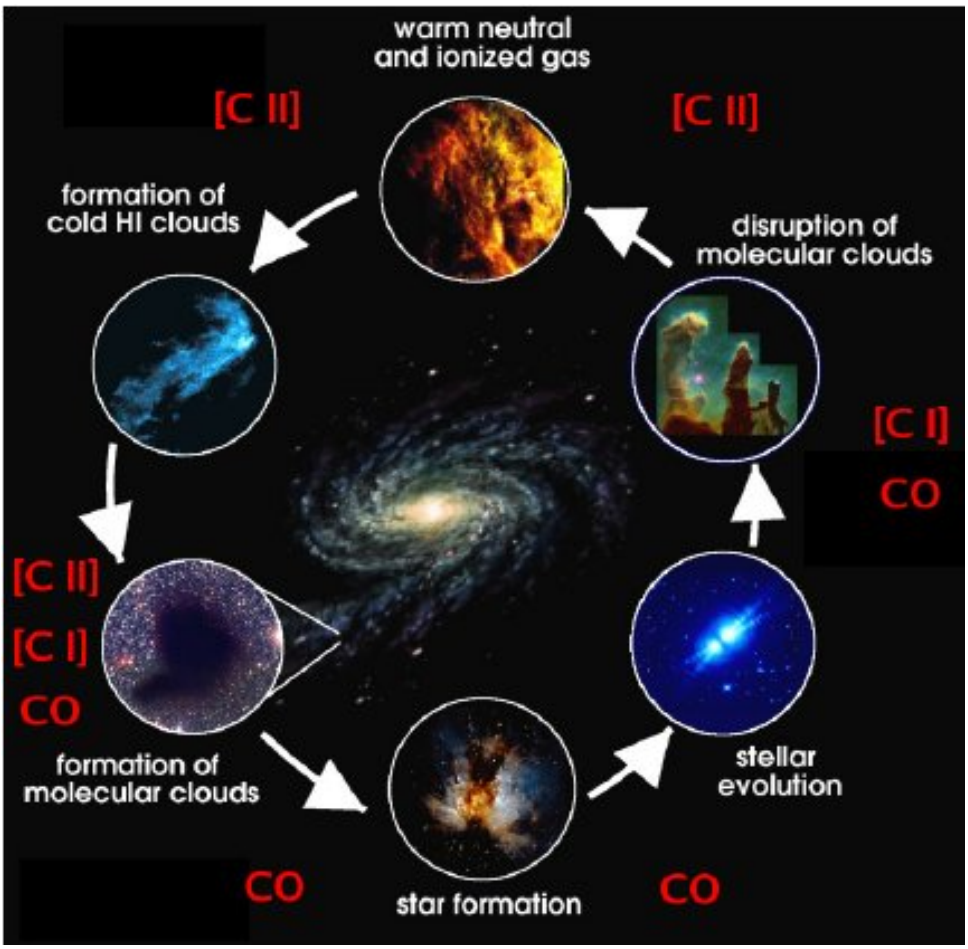


Figure 1.1: Sketch of the various stages in the life cycle of stars and matter in the galaxy. Reproduced from Kulesa *et al.* (2012). Opportunities for Terahertz Facilities on the High Plateau. Proceedings of the International Astronomical Union, 8(S288), 256-263 with permission from Cambridge University Press.

1.1 A few words about photophysics

One way to trace the evolution of remote objects (e.g. interstellar regions) is *via* spectroscopical analysis. Owing to the technological advancements in various fields, many telescopes have been built and numerous observations have been performed in recent years. In particular, Herschel and Spitzer missions provided spectral and spatial information about the infrared (IR) emission for many astrophysical environments in several galaxies.[3]

Some basic photophysical concepts have to be introduced in order to better understand the evolution of molecular complexes following absorption of photons. Figure 1.2 shows several types of relaxation upon absorption of an ultraviolet (UV) light by an isolated molecule. The case of dissociation/fragmentation has not been considered here even though it is responsible for a significant amount of the interstellar chemistry. In fact, there are many chemical reactions that play a crucial role in the evolution of the ISM (see Table 1.1). It is also worth mentioning photoisomerization as one of the most important photochemical processes.[5]

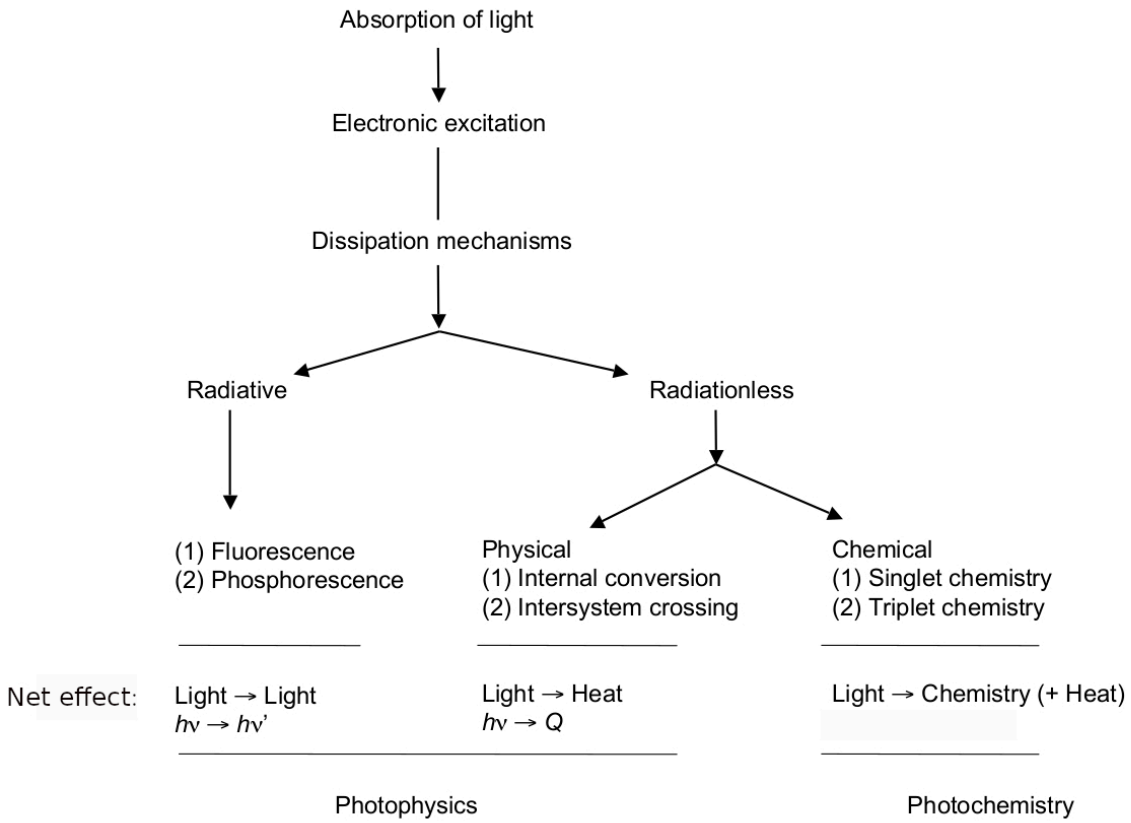


Figure 1.2: Possible decay pathways for a photoinduced process according to Ref. [4].

Table 1.1: Important chemical reactions in the ISM.

Bond formation	
Radiative association	$A + B \rightarrow AB + h\nu$
Grain surface formation	$A + B : g \rightarrow AB + g$
Associative detachment	$A^- + B \rightarrow AB + e^-$
Bond breaking	
Photodissociation	$AB + h\nu \rightarrow A + B$
Dissociative recombination	$AB^+ + e^- \rightarrow A + B$
Collisional dissociation	$AB + M \rightarrow X + Y + M$
Rearrangement reactions	
Ion-molecule exchange	$A^+ + BC \rightarrow AB^+ + C$
Charge transfer	$A^+ + B \rightarrow A + B^+$
Neutral collisions	$A + BC \rightarrow AB + C$

Consider now a molecular system in its ground (equilibrium) state. Once a photon is absorbed, the system gains an excess of energy that promotes it to an excited state. The subsequent relaxation can be roughly divided in two branches depending whether there is a photon emission or not:

- The system may reach a certain excited state after a cascade of radiationless transitions and emit a photon from this state. This emitted photon will carry an excess of energy, thus bringing the system back to the ground state. Fluorescence (phosphorescence) corresponds to an emission from an excited singlet (triplet) state. This is a radiative process;

- Alternatively, a cascade of transitions may bring the system all the way back to the ground state. Thus, an excess of energy is translated into the heat. Since no emission is involved, this process is non-radiative.

The Jablonski diagram (see Figure 1.3) is commonly used to represent the complexity of photophysical processes. It takes into account the vibrational relaxation (radiative transitions between vibrational modes), which is responsible for the IR emission. The number of vibrational degrees of freedom is $3N$–6 for a molecule with N atoms. It is worth mentioning that the vibrational relaxation in an isolated system is driven by the coupling between the initally excited and the remaining vibrational modes. Thus, the IR spectroscopy is a powerful tool that can be used to distinguish the chemical families of complex molecular systems. While the IR emission can be observed in the ground state, the fluorescence (phosphorescence) is related to excited states only. The timescales of the latter processes are significantly (by several orders of magnitude) lower than in case of an internal conversion or vibrational relaxation. If the emitted photon carries energy in the range of 380 to 750 nm, then such light falls in the visible part of the spectrum (often called UV-vis or visible UV).

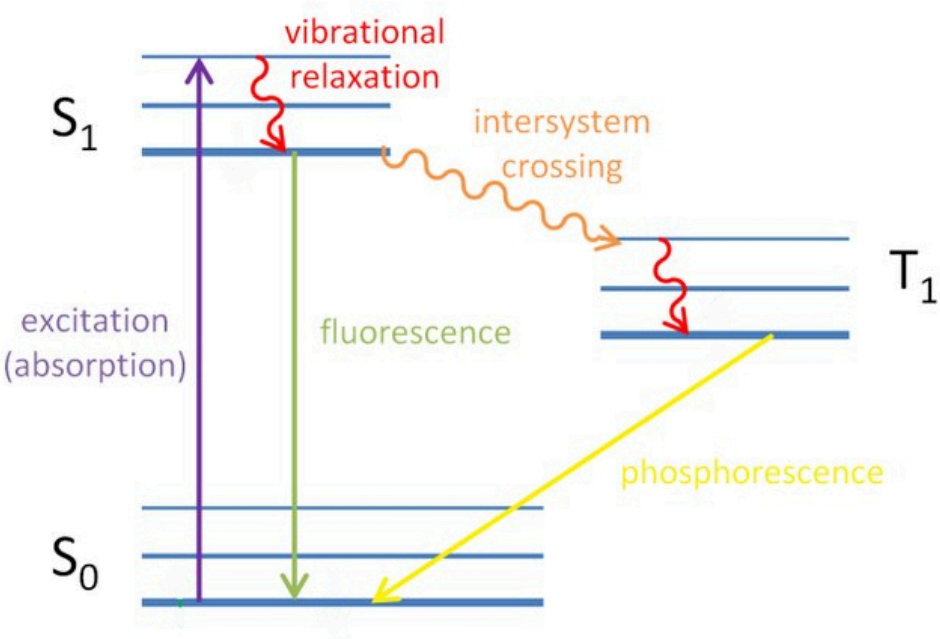


Figure 1.3: Simplified Jablonski diagram representing different relaxation pathways following absorption of a UV photon. Intersystem crossing (ISC) is a radiationless transition between singlet and triplet states.

1.2 The PAH hypothesis

The analysis of the dust emission in astrophysical objects has indicated the presence of some characteristic features in the mid-infrared (around 3.3, 6.2, 7.7, 8.6 and 11.3 μm). These bands are now known as Aromatic Infrared Bands (AIBs). More bands are observed with a higher resolution in several astronomical objects (see Figure 1.4).

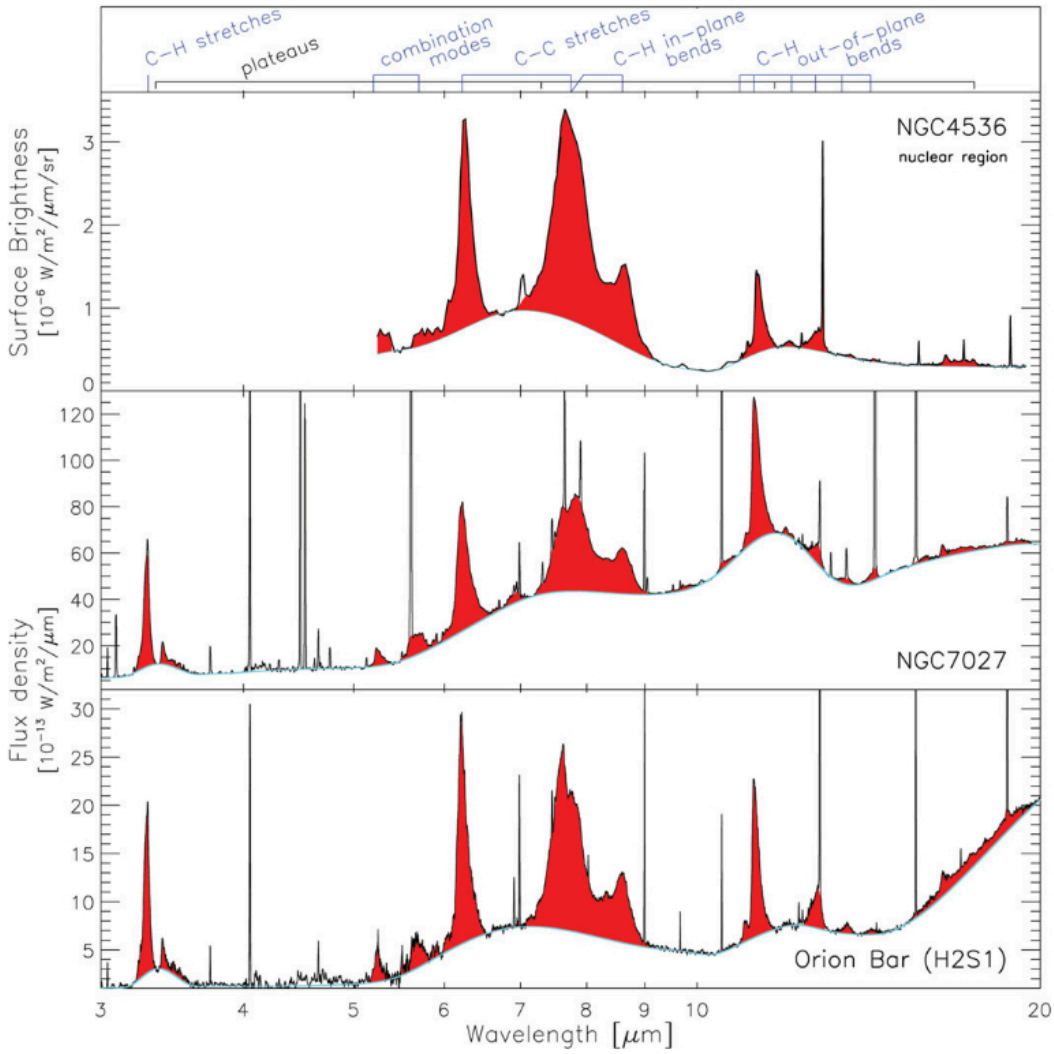


Figure 1.4: The ISO-SWS spectra of the nucleus of the galaxy NGC 4536, planetary nebula NGC 7027 and the photo-dissociation region at the Orion bar. Aromatic modes of the major PAH vibrations are indicated on the top axis. Reproduced from Els Peeters (2011). The PAH Hypothesis after 25 Years. Proceedings of the International Astronomical Union, 7(S280), 149-161 with permission from Cambridge University Press.

AIBs were initially attributed to very small grains.[6] Later on, Léger and Puget[7] and Allamandola *et al.*[8] proposed that Polycyclic Aromatic Hydrocarbons (PAHs, see Figure 1.5) are the carriers of these bands. PAHs are organic molecules that are composed of two or more carbonaceous hexagons (aromatic carbon rings) with peripheral hydrogens. These rings result

from the sp^2 bonds between carbon atoms and grant the aromaticity to the corresponding molecule. Some PAHs can be called peri-condensed if there are carbon atoms that are shared between three rings and cata-condensed if carbon atoms are located in one or maximum two successive rings. The PAH hypothesis is based on the assumption that PAH molecules reach high temperatures following the absorption of a UV photon and then radiatively cool down *via* an IR emission. They can survive the UV radiation due to their remarkable stability. Notably, some AIBs have been attributed to aliphatic (methylated) groups by Joblin *et al.*[9] More details about the PAH hypothesis can be found in Refs. [10, 11] and references therein.

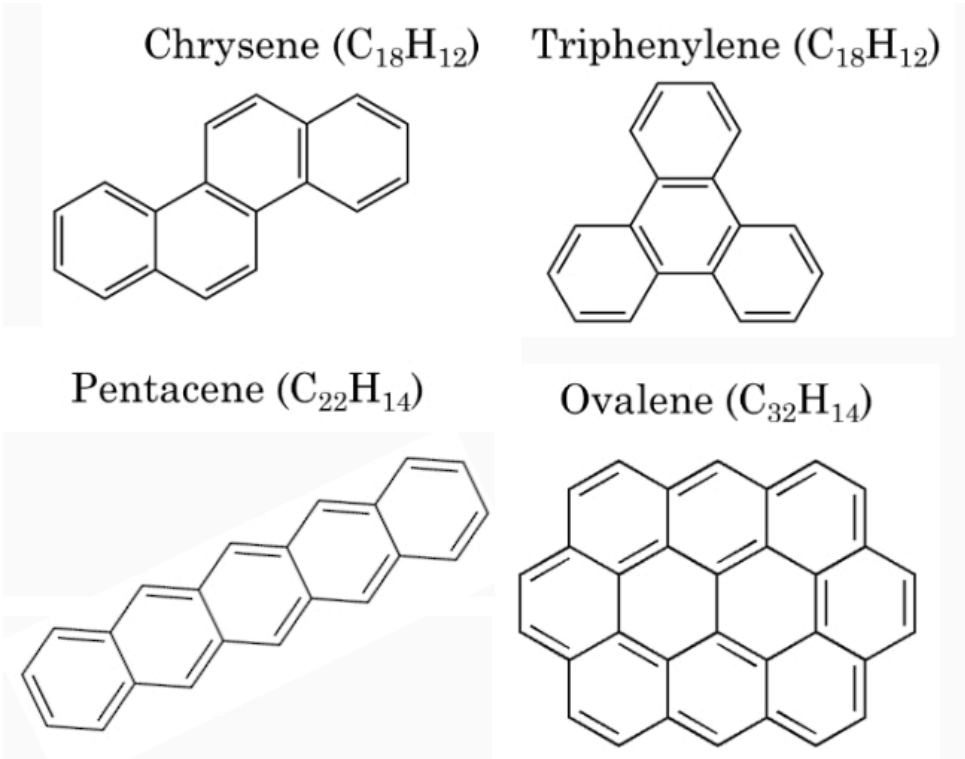


Figure 1.5: Several PAH molecules. Left and right columns correspond to cata- and peri-condensed morphologies, respectively.

The aforementioned PAH model suffers from the lack of identification of individual species. However, PAHs also have absorption bands in the visible and UV range. In fact, Diffuse Interstellar Bands (DIBs), that fall essentially in the visible range,[12] might be due to ionized PAH compounds[13–17] and these bands may hold important clue for the identification of individual interstellar PAHs. This motivated many experimental and theoretical investigations of static[13,18–21] and dynamical[22–26] properties of charged PAHs. However, less information about the photophysics of large neutral PAH molecules is available, despite the fact that they cannot be excluded from contributing to the DIBs.[14]

Up to date, not a single PAH has been unambiguously assigned as a DIB carrier. Furthermore, Peeters *et al.*[27] have noticed that the shape of AIBs does not vary significantly in different environments. This leads to an open discussion on whether they originate from a limited

number of large and compact PAHs (the so-called "grandPAH" hypothesis)[28] or from a random mixture of many different compounds[29]. On the other hand, the role of derivative species like nitrogenated PAHs (nitro-PAHs), PAH clusters and PAH complexes with metals is yet to be precised.

It is worth mentioning two laboratory studies that have confirmed C_{60}^{+}[30], C_{60} and C_{70}[31] as the carriers of DIBs. Recent work by Palotás *et al.*[32] on protonated buckminsterfullerene $C_{60}H^{+}$ has indicated its possible contribution to the IR emission in the regions with high abundances of C_{60}. Another recent theoretical study has addressed the dual photophysical behaviour of C_{60}^{+} upon near-IR and UV excitations.[33]

1.2.1 Laboratory spectroscopy of PAHs

It is hard to imagine further development and assesment of the PAH hypothesis without the support from laboratory research. This has motivated many spectroscopical studies on PAH compounds both at the experimental and theoretical levels (see Ref. [11] and references therein). Numerous infrared and ultraviolet spectra have been measured or calculated for neutral and charged species of different sizes. These results have been compiled in spectroscopic atlases like Refs. [34, 35] and free online databases developed by Malloci *et al.*[36] and Bauschlicher *et al.*[37] (hosted by OAC-Cagliari[1] and NASA AMES[2], respectively).

PAHs may also provide a catalytic surface for H_2 formation in the ISM.[38] The group of Prof. Liv Hornekær from Aarhus University has studied the hydrogenation patterns of neutral PAHs with the help of scanning tunneling microscopy, mass spectroscopy and theoretical calculations. The superhydrogenated coronene ($C_{24}H_{36}$)[39] and pentacene ($C_{22}H_{36}$)[40] have been demonstrated to act as possible catalysts for molecular hydrogen formation under interstellar conditions. Work on hydrogenated buckminsterfullerene C_{60} is currently in progress.

The EUROPAH Marie Skłodowska-Curie project[3] aims to understand the role that PAHs play in the physics and chemistry of the ISM. The present PhD thesis is part of this collaborative effort that brings together 16 PhD students from 9 universities.

1.3 Theoretical considerations

Quantum Chemistry and Materials Science are rapidly developing fields. Many useful molecular or solid state properties can be derived from a solution of the time-independent Schrödinger equation. However, solving this equation for many-body systems is a challenging and computationally demanding problem. One breakthrough in the field was the development of the

[1] http://astrochemistry.oa-cagliari.inaf.it/database/pahs.html
[2] http://www.astrochem.org/pahdb
[3] http://www.europah.eu

Hartree-Fock (HF) approach in the late 1920s. The main idea behind it is that the exact N-body wavefunction of the system can be approximated by a single Slater determinant. By invoking the variational method, one can derive a set of N coupled equations for N spin orbitals. A self-consistent solution of these equations yields the HF wavefunction and energy of the system. The original HF approach as well as the so-called post-HF wavefunction methods are often limited to small systems. Later on, Hohenberg and Kohn[41] proved that the ground state electronic energy is determined completely by the electron density. This theorem lies in the foundation of what is called Density Functional Theory (DFT) nowadays. Even though a different density in principle yields to a different ground state energy, the functional that maps one to another remains unknown. For more details, see Chapter 2.

Understanding the photophysics of PAHs is a crucial and challenging problem for modern Quantum Chemistry. Naphthalene (the smallest PAH) has 10 carbon and 8 hydrogen atoms or 48 valence electrons. Taking into account that calculation of some integrals in Quantum Chemistry has the $O(n^4)$ complexity, where n is the number of electrons (actually up to $O(n^8)$ for wavefunction-based methods[42]), performing theoretical simulations on PAHs can be cumbersome if not unfeasible. Nevertheless, studies have been done both at the post-HF[43] and DFT[44,45] levels of theory.

Moreover, the MAD group from Université Toulouse III - Paul Sabatier and CNRS has addressed a broad range of theoretical challenges on PAH-related compounds. I would like to highlight some of them, in particular (i) anharmonic infrared spectroscopy of cationic [SiPAH] and [FePAH][46,47], (ii) statistical dissociation/fragmentation[48–53] and (iii) isomerization pathways and barriers[54]. More recently, Dubosq *et al.*[55,56] have studied four structural families of C_{60} carbon clusters that might contribute to the AIBs[55] and the DIBs in the UV[56] range of spectra. On the other hand, structural[57], thermal[58] and ionization[59] properties of pyrene clusters have been investigated in great details.

The majority of the aforementioned studies has been performed with a Density Functional based Tight-Binding (DFTB) method that is a simplified and computationally efficient (compared to the conventional DFT) approach for electronic structure calculations. The DFTB, combined with molecular dynamics, allows for an exhaustive exploration of Potential Energy Surfaces (PESs) for relatively large compounds. Despite its success to describe the static and dynamical properties of some PAHs in the ground state, information about excited-state dynamics is generally missing. This is not an issue when one considers relatively slow (~ns or ~μs) processes. However, ultrafast (~as or ~fs) phenomena have gained a lot of attention in recent years.[22,23,25,26] Accurate description of the underlying quantum effects requires access to molecular dynamics beyond the Born-Oppenheimer approximation. Implementation and application of such non-adiabatic molecular dynamics coupled to the DFTB formalism is the main topic of the present PhD thesis.

1.3.1 Non-adiabatic effects (or a few more words about photophysics)

Modelling the evolution of an extended molecular system following the absorption of a UV photon is a challenging task, which requires a fine description of the energy spreading over many electronic and nuclear degrees of freedom. Understanding the involved processes is however of primary interest to characterize the photostability or photochemistry as well as the competition between radiative and non-radiative relaxation channels.[5]

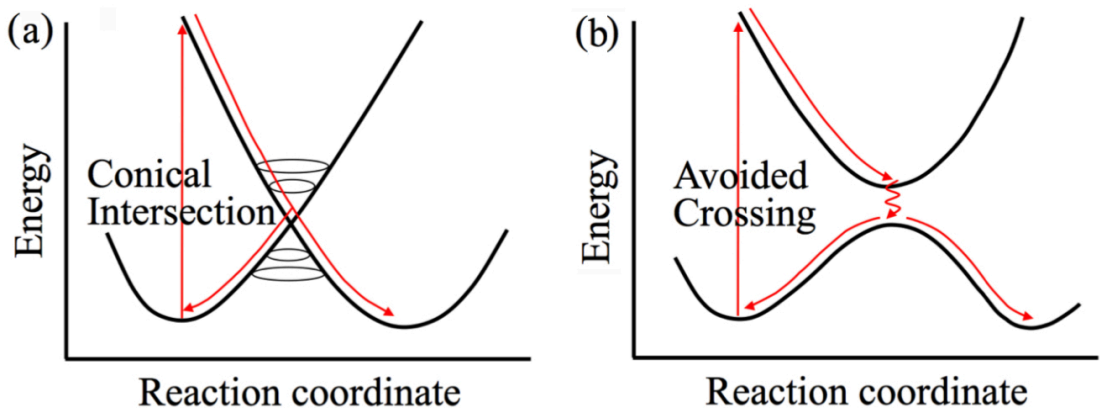


Figure 1.6: Schematic potential energy profiles for relaxation pathways through (a) a conical intersection and (b) an avoided crossing. Adapted from the HDR thesis of Dr. Martial Boggio-Pasqua.[60]

Standard molecular dynamics usually relies on the Born-Oppenheimer approximation, which retains a single adiabatic electronic state in the expansion of a wavefunction of the system. This approach works fine as long as nuclei are considered as slowly moving particles (due to their higher mass compared to electrons). In the Born-Oppenheimer approximation, the nuclear dynamics obeys classical Newton's equations of motion, while electrons are propagated using the time-dependent Schrödinger equation. However, in the regions where two or more electronic PESs are getting close (e.g. in the vicinity of a conical intersection, see Figure 1.6), the nuclear and electronic degrees of freedom evolve with similar timescales and cannot be treated separately. This breakdown of the Born-Oppenheimer approximation requires to incorporate the non-adiabatic effects in order to render reliable quantum dynamics. Many approaches have been developed in recent years both at the *ab initio*[61] and mixed quantum-classical[62–64] levels of theory. One consists in propagating the time-dependent wavepacket as done in the Multi Configuration Time-Dependent Hartree (MCTDH) scheme.[65–67] This approach, which makes use of the *ab initio* PES fitted on a grid, usually allows to take into account only few electronic states and vibrational degrees of freedom due to the computational complexity of the propagation. Alternative schemes have been developed, relying on a classical description of nuclei, the two most popular ones being the mean-field propagation[68] and the Trajectory Surface Hopping (TSH). Schematic relation between different approaches to non-adiabatic dynamics is depicted in Figure 1.7. Starting from the exact time-dependent Schrödinger

equation, the full molecular problem can be solved *via* (i) a Born-Huang expansion (as done in the MCTDH), (ii) an exact factorization of the molecular wavefunction or (iii) propagation of the density *via* Liouville equations. In the Born-Huang branch, the MCTDH combined with frozen Gaussian wavepackets approach renders the vMCG approximation, which, under certain conditions, converges to the Multiple Spawning.[69] In the exact factorization branch, a trajectory approximation of the nuclear wavepacket leads to a coupled-trajectory mixed quantum-classical method[70], in which trajectories are coupled by quantum forces. If these quantum forces are neglected, the method reduces to the mean-field Ehrenfest approach.[71] From the quantum-classical Liouville equations, assuming unique trajectories, large nuclear velocities and modifying the electronic density matrix, one can derive the fewest-switches trajectory surface hopping scheme.[72] For more details, see Ref. [73] and references therein.

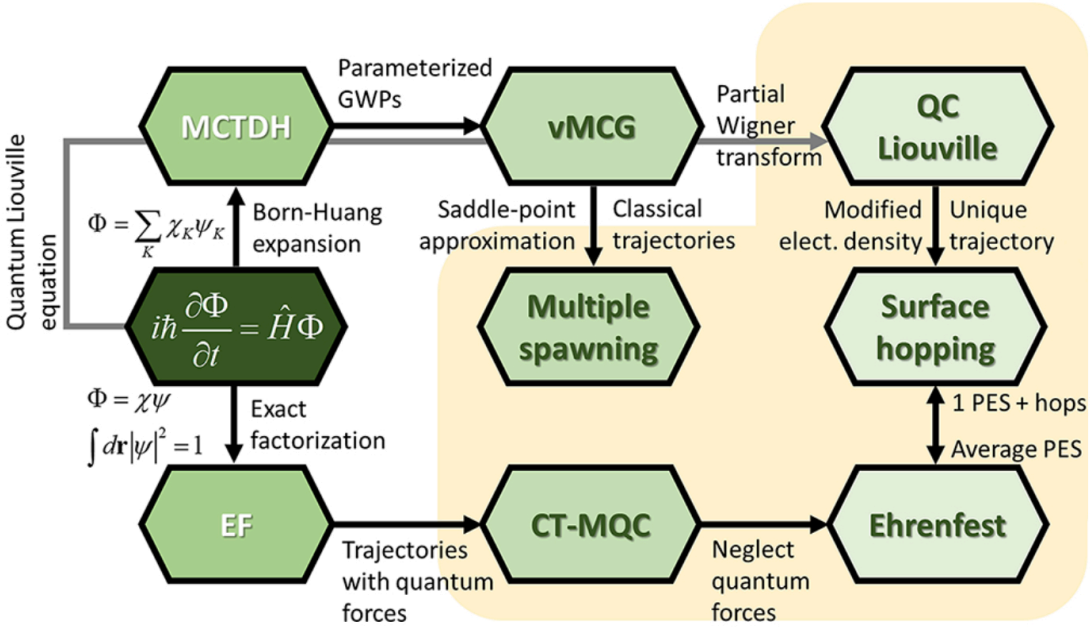


Figure 1.7: Schematic relation between different approaches to non-adiabatic molecular dynamics. EF stands for Exact Factorization, GWP for Gaussian WavePacket and CT-MQC for Coupled-Trajectory Mixed Quantum-Classical approach. Adapted with permission from Rachel Crespo-Otero and Mario Barbatti (2018). Recent Advances and Perspectives on Nonadiabatic Mixed Quantum-Classical Dynamics. Chemical Reviews 118(15), 7026-7068). Copyright (2020) American Chemical Society.

The TSH approach is based on the idea that nuclei evolve on a given PES with a probability to switch states in the regions of strong non-adiabatic coupling. Historically, a Landau-Zener approach[74,75] has been developed for diabatic states' crossing. Later on, Belyaev and Lebedev[76] adapted the Landau-Zener hopping probability for adiabatic states, which is better suited for practical applications due to the fact that Quantum Chemistry codes usually provide adiabatic energies and gradients. Alternatively, Zhu and Nakamura[77,78] have developed an improved Landau-Zener formula for the switching probability. This approach, also called

the Zhu-Nakamura theory, has gained a lot of attention in recent years partially due to its algorithmic simplicity and an absence of non-adiabatic coupling calculations. Yu *et al.*[79] have developed a multidimensional extension of the Zhu-Nakamura theory and applied it to study the cis-trans isomerization of azobenzene. Their scheme has been further improved by Hanasaki *et al.*[80] and validated based on several analytical and *ab initio* models. Both studies indicate reasonable agreement with other theoretical approaches and available experimental data, especially taking into account an absence of non-adiabatic coupling terms. More recently, the Zhu-Nakamura theory has been combined with Machine Learning techniques,[81,82] thus opening a promising direction for development and application of a long-scale quantum dynamics. Yet, one of the most commonly used TSH approaches is the Tully's fewest-switches scheme.[83,84] It incorporates the electronic population and non-adiabatic coupling terms in the hopping probability, thus rendering a more accurate description of the electron dynamics. More details can be found in Section 2.3 of this thesis and Refs. [62, 64, 73]. It is worth mentioning that the calculation of non-adiabatic couplings consumes a significant amount of computational time even within the Time-Dependent Density Functional Theory (TD-DFT)[85] approach. The DFTB method has been extended by Niehaus *et al.*[86] to access electronic excited states leading to the TD-DFTB approach. It has been later on coupled to the nuclear dynamics *via* trajectory surface hopping schemes to study the non-adiabatic relaxation of large molecular systems.[87–90]

1.3.2 Bridging dynamical simulations and spectroscopic observations

Several studies at Imperial College London have been dedicated to the photophysics of cationic naphthalene[18], pyrene[19] and perylene[20]. The presence of easily accessible sloped conical intersections suggests that naphthalene and pyrene radical cations are highly photostable, with an ultrafast non-radiative decay back to the initial ground state geometry. These studies are based on accurate but computationally demanding electronic structure methods like CASSCF and CASPT2. Notably, no accessible conical intersection with the ground state has been found in cationic perylene, which is consistent with experimental[91] observations.

More recently, owing to the rapid development of High-Performance Computing centers, the research community has gained an access to large computational resources that allowed for an exhaustive exploration of PESs. While this is certainly possible with conventional static-like approaches, dynamical simulations represent a good alternative. Thus, excited-state dynamics comes into play to make its contribution in the field of spectroscopy. In particular, it is worth mentioning several theoretical works by the group of Prof. Susanta Mahapatra from University of Hyderabad. In 2010, Reddy *et al.*[22] have performed the MCTDH propagation of the six low-lying electronic states of cationic naphthalene and anthracene. They have also analyzed the vibronic structures and determined non-adiabatic relaxation pathways. The computed time-dependent spectra are in excellent agreement with experimental[92,93] observations. Later on, the same group has performed MCTDH calculations on a series of cationic polyacenes up

to hexacene, demonstrating the importance of the non-adiabatic effects in the spectroscopical analysis of these compounds.[23] It is worth mentioning that computed decay times can be attributed to the corresponding absorption band widths.[94] In 2015, Marciniak *et al.*[25] have reported the results of their femto-astrochemistry experiment. Ultrafast relaxation times (few tens of fs) following XUV pump and IR probe pulses (the so-called pump-probe scheme) have been observed for cationic naphthalene, anthracene, tetracene and pyrene. These results have been supported by the MCTDH calculations and the computed decay times are in good agreement with the experimental values. This remarkable synergy of experiment and theory has been further used to reveal a somewhat counter-intuitive increase of relaxation timescales with increasing state energy and to observe the XUV-induced vibronic coherence in naphthalene.[26] Alternative pump-probe experimental setups have been used by Noble *et al.*[95] to investigate the ultrafast relaxation of a UV-excited pyrene and by Blanchet *et al.*[96–98] to study the coupled electronic-vibrational dynamics in phenantrene[96] and azulene[97,98].

1.4 Objectives and outline

The main objective of this thesis is the implementation of the non-adiabatic molecular dynamics coupled to the DFTB approach for electronic structure calculations and its application to study the photophysics of large PAH molecules.

First, excited state gradients have to be developed in order to propagate a classical trajectory on a given excited PES. Afterwards, a mixed quantum-classical approach incorporating the non-adiabatic effects has to be implemented. Chapter 2 gives an overview of some computational methods that are used in Quantum Chemistry. In particular, one can find methodological and implementation details related to the Tully's fewest-switches TSH scheme (FSSH, Section 2.3) coupled to the TD-DFTB approach (Section 2.2) for the electronic structure calculations in the deMon-Nano[99] code.

Second, an ultrafast radiationless decay can be simulated in order to shed more light on the photostability of PAH-related complexes. Theoretical investigation of the non-adiabatic molecular dynamics of neutral cata-condensed PAHs is summarized in Chapter 3 of this thesis. Computational efficiency of the DFTB approach allows for an exhaustive exploration over many different parameters such as size (Section 3.3) or shape (Section 3.4) of the corresponding molecules. Detailed discussion about observed effects and underlying mechanisms can be found in Subsections 3.3.3 and 3.4.2. These studies are followed by a more methodological work summarized in Chapter 4. It is focused on the analysis of the state-of-the-art Machine Learning models (Subsection 4.5.2) coupled to several simplified TSH schemes (Subsection 4.5.4) for non-adiabatic dynamics.

Finally, conclusions and perspectives are outlined in Chapter 5.

2
Methods

Quantum Chemistry and Materials Science are rapidly developing fields. Prediction of molecular or solid state stationary properties usually requires finding a solution of the time-independent Schrödinger equation

$$\hat{H}\Psi = E\Psi, \tag{2.1}$$

where $\hat{H}$ is the nonrelativistic Hamiltonian operator for a combined system of electrons and nuclei described by position vectors $\mathbf{r}_i$ and $\mathbf{R}_A$, respectively. The Hamiltonian operator in atomic units for N electrons and M nuclei can be written as follows

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \frac{1}{2}\sum_{A=1}^{M}\frac{\nabla_A^2}{M_A} - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}, \tag{2.2}$$

where $r_{iA} = |\mathbf{r}_{iA}| = |\mathbf{r}_i - \mathbf{R}_A|$, $r_{ij} = |\mathbf{r}_{ij}| = |\mathbf{r}_i - \mathbf{r}_j|$, $R_{AB} = |\mathbf{R}_{AB}| = |\mathbf{R}_A - \mathbf{R}_B|$, M_A is the ratio of the mass of the nucleus A to the mass of an electron, Z_A is an atomic number of nucleus A and index of the Laplacian operator ∇^2 refers to the diferrentiation with respect to the corresponding electronic or nuclear coordinates. In Eqn. 2.2, second derivative terms are the kinetic energy of electrons and nuclei, respectively, $1/r_{iA}$ term represents the Coulomb attraction between electrons and nuclei and two last terms describe the repulsion between electrons and nuclei, respectively.

Nuclei are heavier than electrons and, in general, their motion is much slower. Thus, one can further break the system in two parts and describe electrons on a quantum level as moving in the field of fixed nuclei. These assumptions form the Born-Oppenheimer approximation – one of the key concepts in Quantum Chemistry. Following this approximation, the nuclear kinetic energy term can be neglected in Eqn. 2.1 and the nuclear repulsion is considered constant, thus only contributing to the eigenvalues but not affecting the eigenvectors of Eqn. 2.1. The remaining terms form an electronic Hamiltonian

$$\hat{H}_{\text{el}} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}. \tag{2.3}$$

Thus, the initial eigenvalue problem (2.1) boils down to the following equation

$$\hat{H}_{\text{el}}\Psi_{\text{el}} = E_{\text{el}}\Psi_{\text{el}}, \tag{2.4}$$

where $\Psi_{\text{el}} = \Psi_{\text{el}}(\mathbf{r}, \mathbf{R})$ is the electronic wavefunction, which depends explicitly on electronic coordinates and parametrically on the nuclear ones, as does the electronic energy E_{el}. The total energy of the system with fixed nuclei can be computed as follows

$$E_{\text{tot}} = E_{\text{el}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} = E_{\text{el}} + E_{\text{nn}}. \tag{2.5}$$

Analytical solutions exist for some model systems with a few electrons. However, finding a solution for polyatomic molecules is more involved and requires additional approximations beyond the conventional Born-Oppenheimer one. Depending on the system as well as the desired property, methods may differ significantly. Nevertheless, they can be divided in two groups: (i) wavefunction-based and (ii) density-based approaches. I will briefly go through the former ones and will focus more on the latter in Section 2.1.

Historically, the Hartree-Fock approximation was central to the Quantum Chemistry. Later on, schemes based on the Coupled-Cluster (CC) theory and the Configuration Interaction (CI) method demonstrated high accuracy on small molecules and became a common reference for computational methods. Recently, multiconfigurational Complete Active Space methods (e.g. CASSCF and CASPT2) have been used to compute low-lying excited states[100,101] and even to simulate the excited-state dynamics[102,103] of medium-sized molecular systems. However, it seems that the choice of an active space requires certain experience and intuition. Not to mention that the computational complexity of the above-mentioned schemes grows exponentially with the number of electrons, thus making them unfeasible for some practical applications. More detailed review of the wavefunction-based methods can be found in Refs. [104, 105].

2.1 Density Functional Theory (DFT)

One of the major breakthroughs in Quantum Chemistry was the proof by Hohenberg and Kohn[41] that the ground state electronic energy is determined completely by the electron density. This theorem lies in the foundation of what is called Density Functional Theory (DFT) nowadays. Even though a different density in principle yields to a different ground state energy, the functional connecting these two quantities remains unknown. From now on, unless stated

otherwise, I consider only closed-shell systems with the same number of α and β electrons, so the spin index can be neglected.

2.1.1 The Kohn-Sham approach

The rapid development of DFT-based methods would not be feasible without the approach proposed by Kohn and Sham[106] in 1965. It is based on the idea that the electronic kinetic energy should be calculated from an auxiliary set of orbitals, which are used to represent the electron density. Within the Kohn-Sham approach, the energy as a functional of the electron density $\rho = \rho(\mathbf{r})$ can be written as follows

$$E_{\mathrm{DFT}}[\rho] = T_{\mathrm{s}}[\rho] + E_{\mathrm{ne}}[\rho] + E_{\mathrm{ee}}[\rho] + E_{\mathrm{nn}} = T_{\mathrm{s}}[\rho] + E_{\mathrm{ne}}[\rho] + E_{\mathrm{H}}[\rho] + E_{\mathrm{xc}}[\rho] + E_{\mathrm{nn}}, \tag{2.6}$$

where T_{s} is the kinetic energy functional for a system of non-interacting electrons computed as a function of Kohn-Sham molecular orbitals ϕ_i, E_{ne} and E_{ee} are the electon-nuclei attraction and electron-electron repulsion, respectively, E_{H} is the Hartree energy and E_{xc} contains all remaining quantities, which take into account the difference between the exact total energy and the DFT one. All these functionals are summarized below

$$T_{\mathrm{s}}[\rho] = -\frac{1}{2}\sum_{i=1}^{N} n_i \langle \phi_i | \nabla^2 | \phi_i \rangle \,; \tag{2.7}$$

$$E_{\mathrm{ne}}[\rho] = -\sum_{A=1}^{M} \int \frac{\rho(\mathbf{r}) Z_A}{|\mathbf{R}_A - \mathbf{r}|} d\mathbf{r}; \tag{2.8}$$

$$E_{\mathrm{H}}[\rho] = \frac{1}{2}\int\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \tag{2.9}$$

$$E_{\mathrm{xc}}[\rho] = (T[\rho] - T_{\mathrm{s}}[\rho]) + (E_{\mathrm{ee}}[\rho] - E_{\mathrm{H}}[\rho]), \tag{2.10}$$

where the first difference in $E_{\mathrm{xc}}[\rho]$ can be considered as a kinetic correlation energy and the second one describes both potential correlation and exchange energies; $0 \leqslant n_i \leqslant 2$ is an occupation number of orbital i. Thus, the only unknown functional in Eqn. 2.6 is the exchange-correlation one, which describes the effects of the electron-electron interaction.

Now, armed with these definitions, one may start approximating the exchange-correlation functional forms. Derivation of an explicit exchange-correlation functional remains one of the major challenges in Quantum Chemistry. Nevertheless, certain assumptions have been proposed and developed within the DFT. For instance, (i) in the Local Density Approximation (LDA) density can be treated locally as a uniform electron gas or (ii) the first derivative of the density can be included as a variable, thus leading to the Generalized Gradient Approximation (GGA). More advanced techniques like (iii) meta-GGA or hyper-GGA; (iv) hybrid (e.g. B3LYP[107,108]) and (v) short-range/long-range separated[109] functionals are also commonly used in the litarure.

Once an exchange-correlation functional form has been selected, the ground state DFT energy can be calculated using a set of N orthogonal orbitals ϕ_i that minimize the energy *via* Lagrangian $L[\rho]$:

$$L[\rho] = E_{\mathrm{DFT}}[\rho] - \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_{ij}(\langle\phi_i|\phi_i\rangle - \delta_{ij}). \tag{2.11}$$

This boils down to the following eigenvalue problem (Kohn-Sham set of equations) involving an effective one-electron operator

$$\hat{H}_{\mathrm{KS}}\phi_i = \varepsilon_i\phi_i; \tag{2.12}$$

$$\hat{H}_{\mathrm{KS}} = -\frac{1}{2}\nabla^2 + \hat{V}_{\mathrm{eff}}; \tag{2.13}$$

$$\hat{V}_{\mathrm{eff}}[\rho(\mathbf{r})] = \hat{V}_{\mathrm{ne}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + \hat{V}_{\mathrm{xc}}[\rho(\mathbf{r})], \tag{2.14}$$

which has to be solved self-consistently due to the fact that the electron density $\rho(\mathbf{r}) = \sum_i n_i|\phi_i(\mathbf{r})|^2$. In the equation above, an exchange-correlation potential $\hat{V}_{\mathrm{xc}}[\rho] = \frac{\partial E_{\mathrm{xc}}}{\partial\rho}$ and $\hat{V}_{\mathrm{ne}}$ term is sometimes replaced in the literature by a more general external potential $\hat{V}_{\mathrm{ext}}$. Finally, the DFT total energy reads

$$E_{\mathrm{DFT}}[\rho] = \sum_{i=1}^{N} n_i \left\langle \phi_i \middle| -\frac{1}{2}\nabla^2 + \hat{V}_{\mathrm{ne}} + \frac{1}{2}\int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' \middle| \phi_i \right\rangle + E_{\mathrm{xc}}[\rho] + E_{\mathrm{nn}}. \tag{2.15}$$

In practice, Kohn-Sham (KS) molecular orbitals ϕ_i can be represented as a Linear Combination of Atomic Orbitals (LCAO)

$$\phi_i = \sum_{A=1}^{M}\sum_{\mu\in A} c_{\mu i}\chi_\mu. \tag{2.16}$$

From now on, the $\sum_A\sum_{\mu\in A}$ is denoted by $\sum_\mu$ for simplicity. Substituting Eqn. 2.16 into Eqn. 2.12, one can derive the following matrix equation

$$\sum_\nu c_{\nu i}(H_{\mu\nu} - \varepsilon_i S_{\mu\nu}) = 0; \quad \forall\mu, i \tag{2.17}$$

$$H_{\mu\nu} = \langle\chi_\mu|\hat{H}_{\mathrm{KS}}|\chi_\nu\rangle\,; \tag{2.18}$$

$$S_{\mu\nu} = \langle\chi_\mu|\chi_\nu\rangle\,. \tag{2.19}$$

Once matrix elements (2.18) and (2.19) have been computed, Eqn. 2.17 can be solved numerically using different matrix diagonalisation techniques. Some of them have been implemented

in the Intel®Math Kernel Library[110] for efficient parallelization on Central Processing Units (CPUs) and in the MAGMA library[111] for parallelization on hybrid architectures like modern multicore systems with Graphics Processing Units (GPUs). Recently, several variational quantum eigensolvers have been developed and benchmarked on quantum computers.[112,113]

2.1.2 Density Functional based Tight-Binding (DFTB)

There is no free lunch ...
but DFTB provides filling and
affordable lunch.

Professor Dr. Thomas A. Niehaus

The empirical tight-binding methods (based on the limited atomic orbital basis-set representation) have been developed in order to study large molecular systems.[114] Being roughly three orders of magnitude faster than DFT (see Figure 2.1), they contain a number of element-specific parameters. These parameters are generally transferrable (depending on the functional), but most importantly they allow tight-binding schemes to describe nearly all quantum effects accessible by DFT and Hartree-Fock methods.

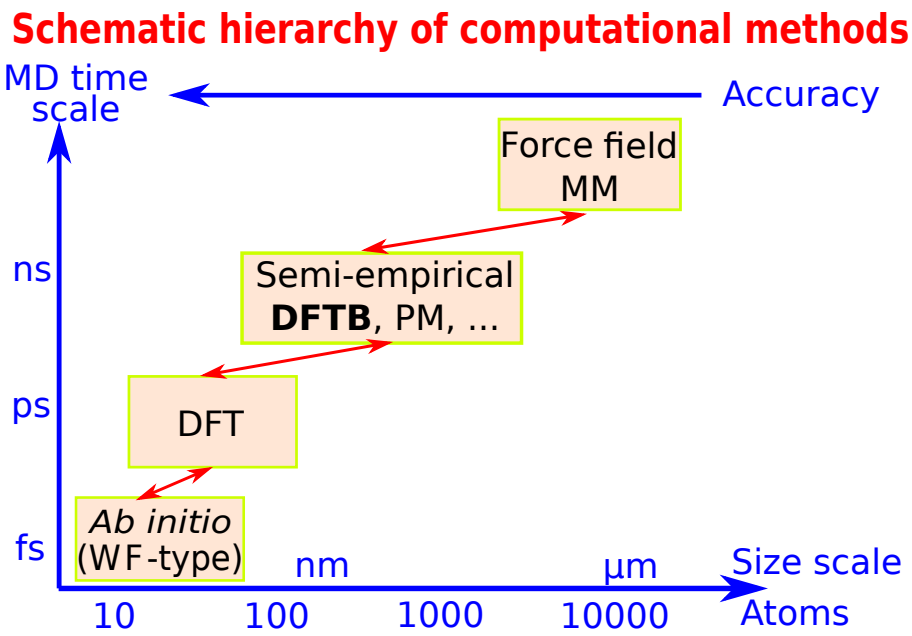


Figure 2.1: Schematic hierarchy of computational methods in Quantum Chemistry. WF stands for the wavefunction, PM is the parametric method, MD and MM stand for the molecular dynamics and molecular mechanics, respectively.

It was shown that a tight-binding scheme can be used as a basis for almost parameter-free approximate treatment within the DFT approach.[115] This led to the development of a Density Functional based Tight-Binding (DFTB) approach. The DFTB method was applied to study large molecules, clusters, nanoparticles and condensed-matter systems.[116,117]

First, one considers the electron density in Eqn. 2.6 as a sum of some reference density (e.g. sum of atomic densities) $\rho_0(\mathbf{r})$ and a small fluctuation $\delta\rho(\mathbf{r})$, i.e. $\rho(\mathbf{r}) = \rho_0(\mathbf{r}) + \delta\rho(\mathbf{r})$. The total energy (2.15) can be then expanded in a Taylor series up to a given order, leading to the following general equation

$$E[\rho_0 + \delta\rho] = E_0[\rho_0] + E_1[\rho_0, \delta\rho] + E_2[\rho_0, (\delta\rho)^2] + E_3[\rho_0, (\delta\rho)^3] + \dots \tag{2.20}$$

Finally, combining Eqn. 2.20 with Eqn. 2.15 and taking only terms with up to the second order of density fluctuations $\delta\rho$, one can derive to the following equation for the total energy[118]

$$E = \sum_{i=1}^{N} n_i \langle\phi_i|\hat{H}_0|\phi_i\rangle - \frac{1}{2}\int\int \frac{\rho_0(\mathbf{r})\rho_0(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{\text{xc}}[\rho_0] - \int \rho_0(\mathbf{r}) V_{\text{xc}}[\rho_0(\mathbf{r})] d\mathbf{r} + E_{\text{nn}} + \\ + \frac{1}{2}\int\int \left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \frac{\partial^2 E_{\text{xc}}}{\partial\rho\partial\rho'}\Big|_{\rho=\rho_0} \right) \delta\rho(\mathbf{r})\delta\rho(\mathbf{r}') d\mathbf{r}d\mathbf{r}', \tag{2.21}$$

where $\hat{H}_0 = -\nabla^2/2 + \hat{V}_{\text{eff}}[\rho_0]$.

Historically, DFTB models have been sequentially proposed and developed depending on the terms used in Eqn. 2.20, starting from the first-order non-self-consistent DFTB (DFTB0)[119,120], followed by the second-order Self-Consistent Charge DFTB (SCC-DFTB or DFTB2)[121] and finally leading to the third-order DFTB[122]. Different implementations and extensions of DFTB are available within the following software packages: DFTB+[123], deMon-Nano[99], ADF[124], Amber[125], Gromacs[126], Gaussian[127], DFTBaby[128] and CP2K[129].

First-order DFTB

This approach neglects the last term in Eqn. 2.21. The DFTB is based on a LCAO ansatz (2.16) where atomic orbitals are obtained from the atomic DFT calculations in a confined potential[118] and the basis is restricted to the valence shell to reduce the computational cost. The terms that depend on the reference density ρ_0 and on the nuclear-nuclear repulsion E_{nn} form the pairwise, repulsive and short-ranged quantity denoted by E_{rep}. The final energy expression for the DFTB0 then reads

$$E_{\text{DFTB0}} = \sum_{i=1}^{N} n_i \langle\phi_i|\hat{H}_0|\phi_i\rangle + E_{\text{rep}}, \tag{2.22}$$

Within the DFTB0, KS equations (2.17) transform into the following system of algebraic equations using the set of localized atomic orbitals $\chi_\mu = \chi_\mu(\mathbf{r} - \mathbf{R}_A)\ \forall\mu \in A$

$$\sum_{\nu} c_{\nu i}(H^0_{\mu\nu} - \varepsilon_i S_{\mu\nu}) = 0 \quad \forall \mu, i \tag{2.23}$$

Now, using the fact that

$$\sum_{i=1}^{N} n_i \langle \phi_i | \hat{H}_0 | \phi_i \rangle = \sum_{i=1}^{N} n_i \sum_{\mu\nu} c_{\mu i} c_{\nu i} H^0_{\mu\nu} = \sum_{i=1}^{N} n_i \varepsilon_i; \tag{2.24}$$

$$E_{\text{rep}} = \frac{1}{2} \sum_{A=1}^{M} \sum_{B=1}^{M} V^{\text{rep}}_{AB}, \tag{2.25}$$

where V^{rep}_{AB} denotes the repulsive potential between pair of atoms A and B, the total DFTB0 can be written as follows[118]

$$E_{\text{DFTB0}} = \sum_{i=1}^{N} n_i \varepsilon_i + \frac{1}{2} \sum_{A=1}^{M} \sum_{B=1}^{M} V^{\text{rep}}_{AB}. \tag{2.26}$$

The DFTB0 combined with the two-center approximation[115] allows to overcome the computational bottleneck of DFT due to the fact, that Hamiltonian and AO overlap matrix elements ($H^0_{\mu\nu}$ and $S_{\mu\nu}$, respectively) as well as repulsive potentials V^{rep} can be calculated only once for a set of interatomic distances between different elements. In practice, they are actually precomputed based on some reference electronic structure method (e.g. DFT with the LDA functional) and stored as an external datafile.

Second-order (Self-Consistent Charge) DFTB

The second-order DFTB scheme has been pioneered by Elstner *et al.* and published in 1998.[121] The derivations presented below are rather concise and based on this original work.

In order to improve the charge balance in heteronuclear systems, the second order term in Eqn. 2.21 has to be taken into account. First, the density fluctuation $\delta\rho$ is written as a superposition of atomic contributions $\delta\rho_A$, i.e. $\delta\rho = \sum_A \delta\rho_A$. The latter ones are further expanded in a series of radial and angular functions and this expansion is truncated after the monopole term[121]

$$\delta\rho_A(\mathbf{r}) \approx \Delta q_A F^A_{00}(|\mathbf{r} - \mathbf{R}_A|) Y_{00} \approx \Delta q_A \frac{F^A_{00}(|\mathbf{r} - \mathbf{R}_A|)}{\sqrt{4\pi}}, \tag{2.27}$$

where F^A_{00} is the normalized radial dependence of the charge fluctuation Δq_A for quantum numbers $m = l = 0$. Expression (2.27) preserves the total charge of the system, i.e. $\sum_A \Delta q_A = \int \delta\rho(\mathbf{r}) d\mathbf{r}$. Now, the second order term (denoted by E_{SCC}) can be rewritten in a simple form

$$E_{\mathrm{SCC}} = \frac{1}{2}\int\int\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \frac{\partial^2 E_{\mathrm{xc}}}{\partial\rho\partial\rho'}\Big|_{\rho=\rho_0}\right)\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')d\mathbf{r}d\mathbf{r}' =$$
$$= \frac{1}{2}\sum_{A=1}^{M}\sum_{B=1}^{M}\int\int\Gamma[\mathbf{r},\mathbf{r}',\rho_0]\delta\rho_A(\mathbf{r})\delta\rho_B(\mathbf{r}')d\mathbf{r}d\mathbf{r}' = \tag{2.28}$$
$$= \frac{1}{2}\sum_{A=1}^{M}\sum_{B=1}^{M}\Delta q_A\gamma_{AB}\Delta q_B;$$

$$\gamma_{AB} = \int\int\Gamma[\mathbf{r},\mathbf{r}',\rho_0]\frac{F_{00}^A(|\mathbf{r}-\mathbf{R}_A|)F_{00}^B(|\mathbf{r}'-\mathbf{R}_B|)}{4\pi}d\mathbf{r}d\mathbf{r}'. \tag{2.29}$$

In the limit of large interatomic distances, E_{SCC} describes a non-screened Coulomb-type interaction between two point charges Δq_A and Δq_B due to the vanishing exchange-correlation contribution within the LDA. On the contrary, when both charges are localised on the same atom B, $\gamma_{BB} \approx \mathrm{IP}_B - \mathrm{EA}_B \approx U_B$, where U_B is the Hubbard parameter, IP_B and EA_B denote the atomic ionization potential and the electron affinity, respectively.[130]

Finally, the total DFTB2 energy takes the following compact form

$$E_{\mathrm{DFTB2}} = \sum_{i=1}^{N} n_i\langle\phi_i|\hat{H}_0|\phi_i\rangle + \frac{1}{2}\sum_{A=1}^{M}\sum_{B=1}^{M}\Delta q_A\gamma_{AB}\Delta q_B + E_{\mathrm{rep}}. \tag{2.30}$$

Solving the secular KS equations (2.23) is no longer sufficient due to the fact that atomic charges depend on the molecular orbitals ϕ_i. Thus, a self-consistent procedure is required to minimize the energy (2.30). But first, Mulliken charge analysis has to be applied to charge fluctuations[118] $\Delta q_A = q_A - Z_A$, i.e.

$$q_A = \frac{1}{2}\sum_{i=1}^{N} n_i\sum_{\mu\in A}\sum_{B=1}^{M}\sum_{\nu\in B}(c_{\mu i}^* c_{\nu i}S_{\mu\nu} + c_{\mu i}c_{\nu i}^* S_{\nu\mu}), \tag{2.31}$$

leading to the following set of KS equations for the Self-Consistent Charge (SCC) DFTB

$$\sum_{\nu} c_{\nu i}(H_{\mu\nu} - \varepsilon_i S_{\mu\nu}) = 0; \quad \forall\mu, i \tag{2.32}$$

$$H_{\mu\nu} = \langle\chi_\mu|\hat{H}_0|\chi_\nu\rangle + \frac{1}{2}S_{\mu\nu}\sum_{C=1}^{M}(\gamma_{AC}+\gamma_{BC})\Delta q_C = H_{\mu\nu}^0 + H_{\mu\nu}^1; \tag{2.33}$$

$$S_{\mu\nu} = \langle\chi_\mu|\chi_\nu\rangle \qquad \forall\mu\in A; \quad \forall\nu\in B \tag{2.34}$$

An analytic expression has been derived for the interatomic forces by Elstner *et al.*[121]

$$\mathbf{F}_A = -\sum_{i=1}^{N} n_i\sum_{\mu\nu} c_{\mu i}c_{\nu i}\left[\frac{\partial H_{\mu\nu}^0}{\partial\mathbf{R}_A} - \left(\varepsilon_i - \frac{H_{\mu\nu}^1}{S_{\mu\nu}}\right)\frac{\partial S_{\mu\nu}}{\partial\mathbf{R}_A}\right] - \Delta q_A\sum_{C=1}^{M}\frac{\partial\gamma_{AC}}{\partial\mathbf{R}_A}\Delta q_C - \frac{\partial E_{\mathrm{rep}}}{\partial\mathbf{R}_A}. \tag{2.35}$$

The DFTB sets of parameters can be downloaded free of charge on www.dftb.org. Among them, the matsci[131] (originally developed for Materials Science applications) and mio[121] (initially parametrized for O, N, C and H) sets of parameters have been actively used for practical applications within both DFTB0 and SCC-DFTB approaches. However, several techniques have been recently proposed in order to improve the DFTB parametrization. In particular, the TANGO approach based on the SCC-DFTB coupled to the genetic algorithm has been developed for an enhanced global optimization.[132] Alternatively, unsupervised Machine Learning methods have been used to improve the fitting of repulsive potentials.[133]

Some limitations and extensions of the conventional DFTB

The DFTB approach has inherited all of the well-known drawbacks[134] of the conventional DFT. In particular,

- Poor description of weak interactions due to the lack of an explicit electron correlation. Dispersion correction has been originally proposed for the DFT total energy by Stefan Grimme[135] and has been later on integrated into some DFTB codes;
- Absence of a wavefunction makes it problematic to access excited states that have the same symmetry as the ground state. However, excited state properties like energies, forces and transition dipole moments can be calculated within the linear response Time-Dependent DFT (TD-DFT) method, which was originally developed by Mark Casida[136] in 1995 and further adapted for DFTB by Niehaus et al.[86];
- Self-interaction error occurs due to the double counting of an electron-electron interaction in the DFT total energy expression, i.e.

$$\frac{1}{2}\int\int\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + E_{\mathrm{xc}}[\rho] \neq 0. \tag{2.36}$$

  In particular, the long-range $1/r$ behaviour is not restored by conventional functionals (including LDA ones). However, this can be fixed by separation of the Coulomb interaction into the long-range and short-range contributions. This correction also improves the description of excited Rydberg states corresponding to the excitation of an electron into a diffuse orbital and charge-transfer states, where an electron is transferred over a large distance (e.g. intermolecular charge transfer). Two long-range corrected DFTB schemes have been proposed independently by Lutsker *et al.*[137] and Humeniuk and Mitrić[138] in 2015.

2.2 Time-Dependent DFT(B)

So far, all approximations and derivations concerned only ground state properties. The time has come to move towards the description of excited states, which is essential to describe many

relevant problems of molecular optics and electronic spectroscopy.

The Time-Dependent Density Functional Theory (TD-DFT) is an extension of the conventional DFT to the situation where a system in its ground stationary state is exposed to a time-dependent perturbation, which modifies its external potential.[136] A good overview of the TD-DFT fundamentals can be found in Ref. [139].

Similarly to the original proof by Hohenberg and Kohn[41] for the ground state, Runge and Gross have demonstrated in 1984 the existence of a unique correspondence between the time-dependent electron density $\rho(\mathbf{r}, t)$ and the time-dependent potential $\hat{V}_{\text{eff}}[\rho(\mathbf{r}), t]$.[140] This leads to the following set of time-dependent KS equations

$$i\frac{\partial \phi_j(\mathbf{r}, t)}{\partial t} = \left[-\frac{1}{2}\nabla^2 + \hat{V}_{\text{eff}}[\rho(\mathbf{r}), t]\right] \phi_j(\mathbf{r}, t). \tag{2.37}$$

In practice, Eqn. 2.37 has to be solved if the time-dependent potential is strong. It can be done within the Ehrenfest approach and is particularly useful for the real-time atomistic simulations of transient absorption spectroscopy.[141] On the other hand, in the limit of a weak perturbation, the linear response approximation can be applied to simplify the solution.[136] In the next section, I will introduce some basic linear response TD-DFT equations and will focus more on the tight-binding extension.

2.2.1 Linear Response TD-DFT(B)

The linear response only takes into account components of a density fluctuation $\delta\rho(\mathbf{r}, \omega)$ that depend linearly on the external perturbation, i.e.

$$\delta\rho(\mathbf{r}, \omega) = \int \chi(\mathbf{r}, \mathbf{r}', \omega)\delta V_{\text{ext}}(\mathbf{r}', \omega)d\mathbf{r}', \tag{2.38}$$

where $\delta V_{\text{ext}}(\mathbf{r}', \omega)$ is the linearized time-dependent KS potential from (2.14). The response function χ of the non-interacting KS system is computed as a function of the stationary Kohn-Sham orbitals (complex conjugation is further suppressed for simplicity)

$$\chi(\mathbf{r}, \mathbf{r}', \omega) = \lim_{\eta \to 0^+} \sum_{k=1}^{N}\sum_{l=1}^{N}(f_k - f_l)\frac{\phi_l(\mathbf{r})\phi_k(\mathbf{r})\phi_l(\mathbf{r}')\phi_k(\mathbf{r}')}{\omega - (\varepsilon_l - \varepsilon_k) + i\eta}, \tag{2.39}$$

where f_k is the Fermi occupation number of the KS molecular orbital k. Substituting (2.39) into (2.38) and following the derivations from Ref. [136], one arrives to the following non-Hermitian eigenvalue problem (Casida's equation)[85]

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X}^I \\ \mathbf{Y}^I \end{pmatrix} = \Omega_I \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \begin{pmatrix} \mathbf{X}^I \\ \mathbf{Y}^I \end{pmatrix}, \tag{2.40}$$

where $\Omega_I = E_I - E_0$ is the excitation energy corresponding to the vertical transition between ground state E_0 and excited state E_I, $\mathbb{1}$ is the identity matrix, $\mathbf{A}$ and $\mathbf{B}$ are matrices with the elements given by

$$A_{ia,jb} = \omega_{ia}\delta_{ij}\delta_{ab} + 2K_{ia,jb}; \tag{2.41}$$

$$B_{ia,jb} = 2K_{ia,jb}; \tag{2.42}$$

and indices i, j and a, b denoting the occupied and virtual Kohn-Sham orbitals, respectively; $\omega_{ia} = \varepsilon_a - \varepsilon_i$. Alternatively, the Tamm-Dancoff approximation has been introduced[142], which neglects the $\mathbf{B}$ matrix, thus leading to the following equation

$$\mathbf{A}\mathbf{X}^I = \Omega_I \mathbf{X}^I. \tag{2.43}$$

The coupling matrix elements $K_{ia,jb}$ for singlet excited states are calculated in the adiabatic approximation as follows[85,136]

$$K^S_{ia,jb} = \int\int \phi_i(\mathbf{r})\phi_a(\mathbf{r}) \left[\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \frac{\partial^2 E_{\text{xc}}}{\partial\rho(\mathbf{r})\partial\rho(\mathbf{r}')}\right] \phi_j(\mathbf{r}')\phi_b(\mathbf{r}')d\mathbf{r}d\mathbf{r}'. \tag{2.44}$$

The TD-DFT approach was coupled to the DFTB by Niehaus *et al.* in 2001.[86] It required some additional approximations beyond the exchange-correlation functional in the coupling matrix $K_{ia,jb}$. In the DFTB, both MOs and MO coefficients are real-valued, so I further suppress the complex conjugation in order to be consistent with the literature. Similarly to the SCC-DFTB derivations, one has to decompose the transition density $\rho^{ia} = \phi_i\phi_a$ into the atom-centered contributions ρ^{ia}_A, which are further expanded in series of radial and angular functions and truncated after the monopole term[86]

$$\rho^{ia}_A(\mathbf{r}) \approx q^{ia}_A F_A(\mathbf{r}) \tag{2.45}$$

where $F_A(\mathbf{r})$ is the normalized spherical density distribution and q^{ia}_A are the Mulliken atomic transition charges, which can be computed as follows

$$q^{ia}_A = \frac{1}{2}\sum_{\mu\in A}\sum_{B=1}^{M}\sum_{\nu\in B}(c_{\mu i}c_{\nu a}S_{\mu\nu} + c_{\mu a}c_{\nu i}S_{\nu\mu}). \tag{2.46}$$

The coupling matrix (2.44) for singlet TD-DFTB excited states then reads[86]

$$K^S_{ia,jb} = \sum_{A=1}^{M}\sum_{B=1}^{M} q^{ia}_A \gamma_{AB} q^{jb}_B. \tag{2.47}$$

For practical purposes, one can apply the following transformation $(\mathbf{A}-\mathbf{B})^{-1/2}(\mathbf{X}^I+\mathbf{Y}^I)=\mathbf{F}^I$ in Eqn. 2.40 in order to derive the commonly used TD-DFT(B) set of equations

$$\sum_{j\to b}^{N_{\mathrm{tr}}}(\omega_{ia}^2\delta_{ij}\delta_{ab}+4\sqrt{\omega_{ia}}K_{ia,jb}\sqrt{\omega_{jb}})F_{jb}^I=\Omega_I^2F_{ia}^I. \tag{2.48}$$

Since left-hand side in Eqn. 2.48 corresponds to a Hermitian matrix, eigenvectors $\mathbf{F}^I$ form, in principle, an orthonormal basis set, i.e.

$$\sum_I \mathbf{F}^I\mathbf{F}^I=1, \tag{2.49}$$

which does not hold true for the eigenvectors of Eqn. 2.48. In particular, one can demonstrate that $\mathbf{X}^I\mathbf{X}^I-\mathbf{Y}^I\mathbf{Y}^I=1$.[85]

Singlet oscillator strengths[136] are important quantities of the optical spectra and can be easily computed within the TD-DFTB scheme[143]

$$f_I=\frac{4}{3}\Omega_I\sum_{x,y,z}\left|\sum_{i\to a}^{N_{\mathrm{tr}}}\sum_{A=1}^{M}\mathbf{R}_Aq_A^{ia}\sqrt{\frac{\omega_{ia}}{\Omega_I}}F_{ia}^I\right|^2 \tag{2.50}$$

Triplet excited states can be computed as well within the TD-DFTB approach. This is achieved by introducing the following coupling matrix for triplet excited states[143]

$$K_{ia,jb}^T=\sum_{A=1}^{M}q_A^{ia}M_Aq_A^{jb}, \tag{2.51}$$

where M_A is the magnetic Hubbard parameter that can be obtained from the atomic DFT calculations.[86]

The first application of the linear response TD-DFTB was reported by Niehaus *et al.* in Ref. [86]. Absorption spectra have been computed for neutral polyacenes ranging in size from naphthalene to heptacene. Excitation energies of the lowest and brightest excited singlet states are in good agreement with experimental and TD-DFT data. Later on, vibrationally resolved UV/Vis spectra of various aromatic and polar molecules were calculated using TD-DFTB excitation energies and analytical gradients.[144] The results of TD-DFTB are also in good agreement with the TD-DFT calculations using local functionals.

Some limitations and extensions of the conventional TD-DFT(B)

As has been shown before, the poor description of charge-transfer and Rydberg states is one of the major drawbacks of the linear response TD-DFT(B). However, there are some other limitations:

- Eqn. 2.48 is valid only for closed-shell systems. Spin-unrestricted TD-DFTB has been developed in order to study the absorption spectra of open-shell complexes[145,146];
- The TD-DFT(B) is based on the linear response of the ground state, which may lead to an inappropriate description of the S_1/S_0 non-adiabatic coupling[128] and PES topology[147]. Thus, a proper description of dynamical processes that occur in the vicinity of the S_1/S_0 conical intersection (i.e. internal conversion to the ground state) remains challenging for methods based on the TD-DFT(B).

An extension of the DFTB (so-called DFTB-CI) has been developed by Rapacioli *et al.* for cationic molecular clusters.[148] It goes beyond the linear response formalism, thus eliminating all of the drawbacks presented above. This approach is based on the idea that a cluster can be described *via* a multiconfigurational wavefunction Ψ_0^+ expanded on a basis of charge-localized configurations Ψ_ξ^+ that correspond to removal of a single electron from the Highest Occupied Molecular Orbital (HOMO) of the charged fragment ξ.

Even though it was initially developed to describe the ground state of charged molecular clusters, excited states can be extracted as well. A better description of the ionic excited states has been achieved through an extension of the original formalism. The expansion of a wavefunction is patched with additional terms[149] that correspond to removal of an electron from any occupied orbital of a charged fragment, i.e.

$$\Psi_0^+ = \sum_{\xi=1}^{N_{\mathrm{fr}}} \sum_{i}^{\mathrm{occ}} c_{\xi i} \Psi_{\xi i}^+. \tag{2.52}$$

This modification becomes particularly important for clusters or stacks of large molecules with a small separation of occupied orbitals. Moreover, it allows to incorporate excited states that correspond to local excitations on the charged fragment and their coupling. This scheme has been shown to yield satisfactory PES of excited states in the full geometry range up to intermolecular dissociation of cationic PAH clusters.[149]

2.2.2 Analytical gradients for TD-DFT(B)

Molecular dynamics is a powerful tool that can be used to describe numerous physical and chemical processes. In order to propagate a classical trajectory on a given PES using Newton's equations of motion, excited state energy gradients have to be developed for the linear response TD-DFT(B). Their derivation relies on the so-called Z-vector method, which was initially applied by Furche and Ahlrichs[150,151] to compute analytical forces within the TD-DFT approach. It was further adapted by Heringer *et al.*[152,153] to compute the TD-DFTB gradients. Alternatively, one can follow derivations of TD-DFTB gradients with the long-range correction from Ref. [128].

Since ground state forces have been already developed by Elstner *et al.* for SCC-DFTB,[121], one has to calculate only the gradient of the TD-DFTB excitation energy Ω_I. However, this is not a trivial task due to the fact that a straightforward calculation with finite differences is roughly $2N_f$ times more expensive than a single point DFTB calculation (here N_f denotes the number of degrees of freedom).[152] One can also use a variational formulation of the TD-DFTB, which would require derivatives of the MO coefficients (roughly N_f times more expensive than a single DFTB calculation), thus limiting the practical application of such gradients.[152] In order to avoid this computational bottleneck, an auxiliary Lagrangian has to be used:

$$L[X, Y, \Omega, C, Z, W] = G[X, Y, \Omega] + \sum_{ia} Z_{ia} H_{ia} - \sum_{p,q,p\leq q} W_{pq}(S_{pq} - \delta_{pq}), \tag{2.53}$$

where $G[X, Y, \Omega]$ is a TD-DFT(B) functional[150] that leads to Eqn. (2.48). In this section, indices i, j, k and a, b, c denote occupied and virtual MOs, respectively, and p, q, r, s denote general MOs. In the equation above, quantities W and Z are Lagrange multipliers and their stationary conditions set restrictions on KS orbitals (e.g. that overlap $S_{pq} = \delta_{pq}$). The stationarity of L with respect to MO coefficients C ($\partial L/\partial C_{\mu p} = 0$) leads to the so-called Z-vector equation for an excited state I

$$\sum_{jb} (A + B)^S_{ia,jb} Z^I_{jb} = Q^I_{ai} - Q^I_{ia}. \tag{2.54}$$

Blocks of the Q^I matrix has to be computed according to Appendix B from Ref. [152] and the corresponding erratum published in Ref. [153]. Once Eqn. (2.54) is solved, the W^I matrix can be computed as well:

$$W^I_{ij} = \frac{1}{1 + \delta_{ij}} \left(Q^I_{ij} + \sum_{kb} 4K^S_{ij,kb} Z^I_{kb} \right); \tag{2.55}$$

$$W^I_{ab} = \frac{1}{1 + \delta_{ab}} Q^I_{ab}; \tag{2.56}$$

$$W^I_{ia} = W^I_{ai} = Q^I_{ai} + \varepsilon_i Z^I_{ia}. \tag{2.57}$$

It is worth mentioning that the Z-vector itself contains valuable information about the nature of excited state. In particular, it has been used to investigate detachment/attachment processes characterizing a hole/particle pair generation upon an electronic transition in dyes.[154]

At the stationary point of L

$$L[X, Y, \Omega, C, Z, W] = \Omega \Rightarrow \frac{dL}{dR} = \frac{d\Omega}{dR}. \tag{2.58}$$

Since L is variational in all parameters, $dL/dR = \partial L/\partial R$. Thus, the gradient of the excitation energy becomes

$$\begin{aligned}\frac{d\Omega_I}{d\mathbf{R}_A} &= \frac{\partial G}{\partial \mathbf{R}_A} + \sum_{ia} Z^I_{ia}\frac{\partial H_{ia}}{\partial \mathbf{R}_A} - \sum_{p,q,p\leq q} W^I_{pq}\frac{\partial S_{pq}}{\partial \mathbf{R}_A} = \\ &= \sum_{ij} T^I_{ij}\frac{\partial H_{ij}}{\partial \mathbf{R}_A} + \sum_{ab} T^I_{ab}\frac{\partial H_{ab}}{\partial \mathbf{R}_A} + 2\sum_{ia}\sum_{jb}(X+Y)^I_{ia}(X+Y)^I_{jb}\frac{\partial K_{ia,jb}}{\partial \mathbf{R}_A} + \\ &\quad + \sum_{ia} Z^I_{ia}\frac{\partial H_{ia}}{\partial \mathbf{R}_A} - \sum_{p,q,p\leq q} W^I_{pq}\frac{\partial S_{pq}}{\partial \mathbf{R}_A} \end{aligned} \tag{2.59}$$

Blocks of the T^I matrix has to be computed according to Appendix B from Ref. [152] and the corresponding erratum published in Ref. [153]. In the equation above, vector elements $(X+Y)^I_{ia}$ are computed as follows

$$(X+Y)^I_{ia} = \sqrt{\frac{\varepsilon_a - \varepsilon_i}{\Omega_I}}F^I_{ia}, \tag{2.60}$$

where F^I_{ia} is the eigenvector of (2.48). At this point, one can switch to the AO basis in order to take advantage of derivative terms that have been already computed for the ground state forces (2.35). The AO to MO transformation can be done as follows

$$\frac{\partial S_{pq}}{\partial \mathbf{R}_A} = \sum_{\mu\nu} C_{\mu p}C_{\nu q}\frac{\partial S_{\mu\nu}}{\partial \mathbf{R}_A}. \tag{2.61}$$

Finally, the gradient of the excitation energy on atom A reads

$$\frac{d\Omega_I}{d\mathbf{R}_A} = \sum_{\mu\nu} P^I_{\mu\nu}\frac{\partial H_{\mu\nu}}{\partial \mathbf{R}_A} - \sum_{\mu\nu} W^I_{\mu\nu}\frac{\partial S_{\mu\nu}}{\partial \mathbf{R}_A} + 2\sum_{\mu\nu\lambda\sigma}(X+Y)^I_{\mu\nu}(X+Y)^I_{\lambda\sigma}\frac{\partial K_{\mu\nu,\lambda\sigma}}{\partial \mathbf{R}_A}, \tag{2.62}$$

where $P^I = T^I + Z^I$ is the one-particle difference density matrix and $H_{\mu\nu}$ is the SCC-DFTB Hamiltonian matrix element from (2.33).

2.3 Non-adiabatic molecular dynamics

Standard molecular dynamics usually relies on the Born-Oppenheimer approximation, which retains a single adiabatic electronic state in the expansion of a wavefunction of the system. Several approaches have been developed in recent years to incorporate the non-adiabatic effects both at the *ab initio*[61] and mixed quantum-classical[62–64,73] levels of theory. Nowadays, one of the most commonly used quantum-classical method is the Tully's Fewest-Switches trajectory Surface Hopping (FSSH)[83,84], which is based on the assumption that the nuclear wavepacket

motion can be simulated by an ensemble of independent classical trajectories, each one evolving on a single electronic state at a given time with a certain probability to switch from one state to another (see Figure 2.2).

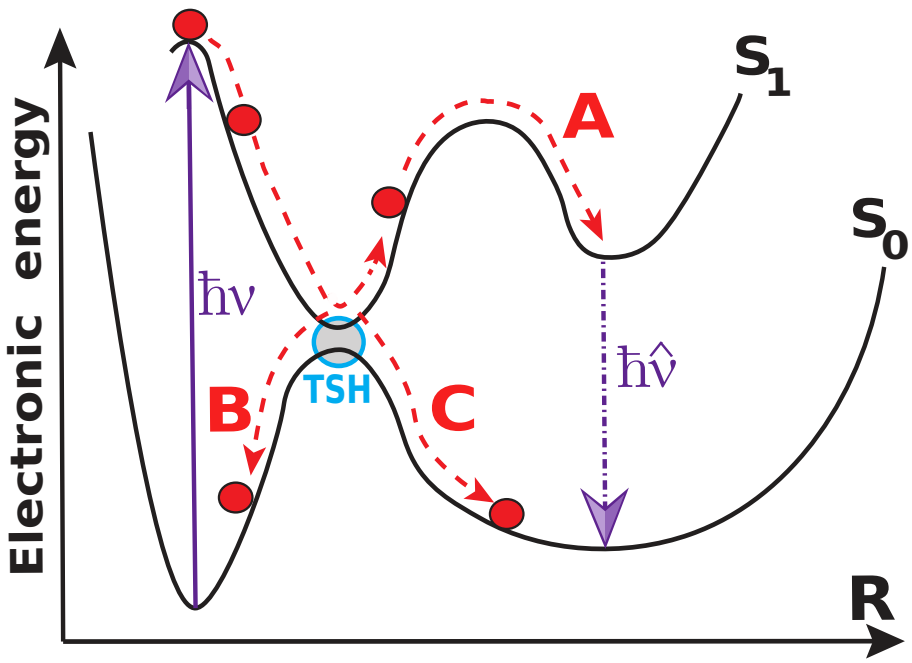


Figure 2.2: Schematic representation of some FSSH-trajectory scenarios following vertical transition from the ground state to an excited singlet state: A) system remains in a given excited state and eventually emits a photon with an excess of energy $\hbar\hat{\nu}$ (photoemission); B) transition between two states occurs in the avoided crossing region (or at the conical intersection) leading to the relaxation towards an equilibrium ground state configuration (radiationless decay); C) same as previous case but the system ends up in a different local minimum of the ground state PES (photoisomerization).

The FSSH propagation can be activated in the deMon-Nano code by the TSH keyword combined with keyword NSTATES=XXX, where XXX has to be replaced by the total number of excited states to be included in the propagation. Initial state (to which the vertical transition from the ground state occurs) has to be reflected in the EXST keyword (i.e. YYY in EXST=YYY has to be replaced by a corresponding index of the initial state). An example of the deMon-Nano input file with all necessary FSSH keywords can be found in Appendix A.

Some general FSSH equations and considerations are summarized in the next section, followed by more specific implementation details related to the DFTB formalism.

2.3.1 Fewest-Switches Trajectory Surface Hopping (FSSH)

In the Tully's scheme, the time-dependent electronic wavefunction Ψ_{el} is expanded on a basis of N_{st} adiabatic electronic states

$$\Psi_{\mathrm{el}}(\mathbf{r};\mathbf{R}(t)) = \sum_{J=1}^{N_{\mathrm{st}}} C_J(t)\psi_J(\mathbf{r};\mathbf{R}(t)), \tag{2.63}$$

where C_J is the complex expansion coefficient and $\psi_J(\mathbf{r};\mathbf{R}(t))$ is the adiabatic electronic wavefunction of J-th state, which depends explicitly on electronic coordinates $\mathbf{r}$ and parametrically on the nuclear coordinates $\mathbf{R}(t)$. I further denote $\psi_J(\mathbf{r};\mathbf{R}(t)) = \psi_J(t)$ for simplicity.

Substituting Eqn. 2.63 into the time-dependent electronic Schrödinger equation and taking into account the adiabatic nature of states, one derives the following equation for the propagation of expansion coefficients $C_J(t)$

$$i\frac{dC_J(t)}{dt} = C_J(t)E_J(t) - i\sum_{K\neq J} C_K(t)D_{JK}(t), \tag{2.64}$$

where E_J is the adiabatic energy of state J and $D_{JK} = \left\langle \psi_J \middle| \frac{\partial \psi_K}{\partial t} \right\rangle$ is the non-adiabatic coupling (NAC) between states J and K. The NAC can be calculated using a finite difference method[84]

$$D_{JK}(t+\Delta t/2) \approx \frac{1}{2\Delta t}[\langle\psi_J(t)|\psi_K(t+\Delta t)\rangle - \langle\psi_J(t+\Delta t)|\psi_K(t)\rangle]. \tag{2.65}$$

In practice, Eqn. 2.64 has to be solved numerically (e.g. using 4-th order Runge-Kutta approximation) with an electronic time step $\Delta\tau$ that is significantly (by several orders of magnitude) smaller than the nuclear time step Δt used for the propagation of a classical trajectory (see Figure 2.3). This does not lead to any computational bottlenecks since propagation of Eqn. 2.64 is still very fast compared to the cost of a single electronic structure calculation (even within the DFTB approach).

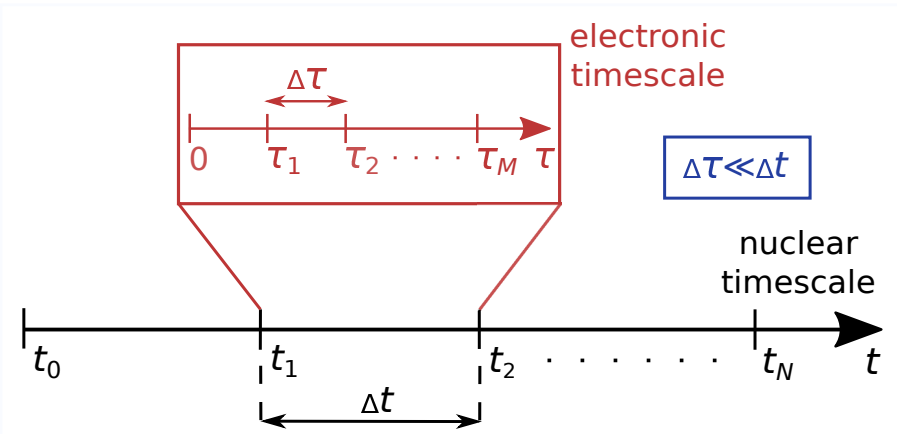


Figure 2.3: Relative timescales corresponding to the nuclear and the electronic propagation within the FSSH approach.

It is important to apply a decoherence correction on C_J since propagation of Eqn. 2.64 in FSSH is overcoherent, which means that electronic coherences $C_I C_J^*$ do not vanish after passing

through the region of strong NAC between states I and J. Decoherence corrections have been shown to be crucial in a number of applications.[72,155] Simplified decay of mixing[155,156] is a straightforward, yet efficient *ad hoc* correction for C_J coefficients

$$\widetilde{C}_J = C_J \exp(-\Delta t/\tau_{JI}) \qquad \forall J \neq I; \tag{2.66}$$

$$\widetilde{C}_I = \frac{C_I}{|C_I|}\sqrt{1 - \sum_{J\neq I}\left|\widetilde{C}_J\right|^2}. \tag{2.67}$$

In the equation above, index I denotes the active state, namely the one on which the trajectory is evolving at the considered moment of time, and τ_{JI} is the decoherence time, which is determined from the following equation

$$\tau_{JI} = \frac{1}{|E_J - E_I|}\left(1 + \frac{\alpha}{E_{\text{kin}}}\right), \tag{2.68}$$

where E_{kin} is the nuclear kinetic energy and α is the phenomenological parameter (recommended value is 0.1 hartree).[156] The tilde symbol is further suppressed, i.e. $\widetilde{C}_I = C_I$. This decoherence correction can be activated in the deMon-Nano code by the DECOH keyword.

Once both electrons and nuclei are propagated from t to $t + \Delta t$ and decoherence corrections applied, the probability to switch from the active state I to another state K during the electronic time step $\Delta\tau$ is estimated from the following equation[84]

$$P_{I\to K}(\tau) = \max\left[0; -2\Delta\tau \frac{\text{Re}\left(C_I^*(\tau)C_K(\tau)\right)}{|C_I|^2} D_{IK}(\tau)\right], \tag{2.69}$$

where $|C_I|^2$ is the electronic population of a given state I. The hopping probability (2.69) is derived from a condition that minimizes the total number of hops, thus giving rise to fewest-switches in the FSSH acronym. Nevertheless, different formulas of hopping probability exist in the literature, e.g. the one developed for the field-induced surface hopping.[157] Once $P_{I\to K}(\tau)$ is computed for all $t \leqslant \tau \leqslant t + \Delta t$, the final hopping probability can be calculated *via* an integration over τ[147,158]

$$P_{I\to K}(t + \Delta t) \approx \int_t^{t+\Delta t} \frac{P_{I\to K}(\tau)}{\Delta\tau} d\tau. \tag{2.70}$$

In the present implementation, hops may only occur at the nuclear timescale since it is consistent with the mixed quantum-classical formulation of the problem. 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.[84] The hop is accepted if the following condition is fulfilled

$$\sum_{J=1}^{K-1} P_{I\to J} < \xi \leqslant \sum_{J=1}^{K} P_{I\to J} \qquad K \neq 1; \tag{2.71}$$

$$\xi \leqslant P_{I\to 1} \qquad \text{otherwise.} \tag{2.72}$$

Energies are computed only along the trajectory, thus the considered PES is a 1D cross-section of a multidimensional one. The FSSH cannot deal with Conical Intersections (CoIn) due to the singularity of NAC at the point of the CoIn. However, several techniques have been developed to cope with it. The most straightforward way is to switch states once the energy gap between them drops below a given threshold, thus avoiding numerical instabilities. This threshold has to be particularly large (∼0.1 eV) for CoIns with the ground state due to the inappropriate S_1/S_0 PES topology in TD-DFT(B).[147] It should be mentioned, that this approach does not guarantee that an observed crossing point corresponds to a true CoIn (which is well defined in 2D space) due to high dimensionality of the PES. One can either terminate or continue the trajectory once the CoIn with the ground state occurs. Another alternative is to use the local diabatization technique, which eliminates numerical instabilities by propagating Eqn. 2.64 in the diabatic basis where $D_{JK} \equiv 0$.[159]

Initial conditions can be sampled either from the Wigner or from the thermal distribution of the ground state. In the latter case, a single NVT trajectory is equilibrated at a given temperature using a certain thermostat (e.g. Nosé-Hoover), and snapshots (coordinates and velocities) are taken to be further used as initial conditions for the FSSH trajectories. Notably, the Wigner sampling assigns kinetic energies that are often too high for hydrogen atoms,[160] which should be avoided, if possible, for such compounds like PAHs. For a more detailed discussion on the choice of initial conditions sampling in the excited-state dynamics see Refs. [73, 161].

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

$$E_I + E_{\text{kin}} = E_J + \beta^2 E_{\text{kin}}. \tag{2.73}$$

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 can adjust the velocities if the hop is accepted and reverse them if the hop is rejected based on the non-adiabatic coupling vector.[84] It has been shown by Bonhommeau *et al.*[162] that the choice of velocity rescaling is crucial in systems where the electronic orbital angular momentum is nonzero. It should be mentioned that commonly used velocity rescaling techniques do not conserve the total nuclear angular momentum.

The maximum statistical error depends on the total number of trajectories N_{traj} and for 95% confidence interval reads[73]

$$\varepsilon \approx \frac{0.98}{\sqrt{N_{\text{traj}}}}. \tag{2.74}$$

2.3.2 FSSH coupled to TD-DFTB

The first study describing methodological as well as development aspects of the FSSH coupled to the TD-DFTB for the electronic structure calculations was published by Mitrić *et al.*[87] NACs are key quantities of the FSSH since they drive the propagation of electronic populations and determine the hopping probabilities. The DFTB, as a density-based method, was not initially developed for the wavepacket calculations. Thus, it is not clear how to build electronic wavefunctions of excited states, which are used in Eqn. 2.65. Common practice is to use the Configuration Interaction Singles (CIS) approach to derive the electronic wavefunction of state K[87,89,136,158]

$$|\psi_K\rangle = \sum_{i\to a}^{N_{\text{tr}}} c_{ia}^K \left|\Phi_{ia}^0\right\rangle, \tag{2.75}$$

where c_{ia}^K are CIS expansion coefficients and $|\Phi_{ia}^0\rangle$ is the singly excited Slater determinant built from a given electronic transition $i \to a$. One should further follow the procedure, which is described in great details by Humeniuk and Mitrić (see Appendix C in Ref. [128]), to extract the wavefunctions overlap between t and $t + \Delta t$ based on the Slater determinants overlap

$$\langle\psi_J(t)|\psi_K(t+\Delta t)\rangle = \sum_{i\to a}^{N_{\text{tr}}}\sum_{j\to b}^{N_{\text{tr}}} c_{jb}^J(t) c_{ia}^K(t+\Delta t) \left\langle\Phi_{jb}^0(t)\middle|\Phi_{ia}^0(t+\Delta t)\right\rangle. \tag{2.76}$$

The total number of transitions N_{tr} in the CIS expansion (2.75) can be reduced using certain threshold. My benchmark calculations on medium-sized molecules demonstrate some numerical instabilities if all $i \to a$ transitions with $|c_{ia}^K| < 10^{-2}$ are excluded and provide reasonably accurate results already for $|c_{ia}^K| < 10^{-3}$. This truncation may reduce the number of transitions in Eqn. 2.75 by several orders of magnitude and it accelerates significantly the computation of NACs for large molecules. A similar approach has been proposed by Plasser *et al.*[163] for efficient computation of many-electron wavefunction overlaps. Nevertheless, even with this truncation, the calculation of NACs remains major computational bottleneck of the present FSSH/TD-DFTB implementation due to the fact that evaluation of singly excited Slater determinants scales exponentially with the number of electrons.

Two different ways to compute c_{ia}^K are commonly used in the literature:

1. $c_{ia}^K = (X+Y)_{ia}^K$ – was used by Tapavicza *et al.*[158] following Casida's assignment from Ref. [136] . Main drawback of this approach is that excited states are not orthogonal by default;

2. $c_{ia}^{K} = F_{ia}^{K}$ – was proposed by Werner *et al.*[164] Excited states are orthogonal due to the fact that F_{ia}^{K} is the eigenvector of the Hermitian eigenvalue problem (2.48).

The second set (current default option in the deMon-Nano code) has been used for all calculations presented in this thesis. However, one can activate the first set by adding CISCASIDA after the DECOH keyword in the input file, but further orthonormalization has to be implemented. It is worth mentioning, that certain choices of CIS coefficients may lead to an over- or underestimation of NACs and hopping probabilities. However, comparison of different techniques used to compute NACs and their impact on FSSH goes beyond the scope of this thesis.

The implementation and application of the combined FSSH/TD-DFTB approach was reported by Mitrić *et al.* in Ref. [87]. The applicability of the method was illustrated on the ultrafast excited-state dynamics of microsolvated adenine. The FSSH has been later on coupled to the long-range corrected TD-DFTB in the DFTBaby code[128] and used to investigate such processes as exciton localization in tetracene trimer[165], excimer formation in pyrene dimer[166], excitonic effects on the absorption spectra of pentacene aggregates[167] and relaxation of excited fluorene oligomers[128]. Other groups have developed interfaces between the DFTB+[123] code and external FSSH modules, namely Newton-X[89] and PYXAID[88]. The former one, enhanced by the transition density analysis, has been used by Stojanović *et al.*[89] to simulate the non-adiabatic molecular dynamics of two cycloparaphenylene molecules, while the latter one has been benchmarked based on electron-hole interactions in CdSe quantum dots and (10,5) semiconducting carbon nanotubes.[88]

The present implementation of the FSSH/TD-DFTB in the deMon-Nano code has been used to investigate relaxation mechanisms in neutral cata-condensed PAHs like polyacenes and phenacenes. More details can be found in Chapters 3 and 4 of this thesis. From the algorithmic point of view, my FSSH implementation is an internal module of the deMon-Nano code, i.e. all quantities are computed on the fly and stored in the memory. This is particularly useful for simulations of large molecular systems or for dynamics involving many states since no extensive data input/output is required as in case of an external FSSH propagator. Furthermore, this FSSH module can be interfaced with any other electronic structure method for excited states (e.g. DFTB-CI) that is available in the deMon-Nano code.

3

Non-adiabatic molecular dynamics of large cata-condensed PAHs[1]

3.1 Introduction

As has been mentioned in Chapter 1, the information about excited-state dynamics of large PAH molecules is generally missing. Not only it is relevant for the astrophysical applications, but it is also of great interest for the laboratory experiments, i.e. in the field of the atto- or femto-second spectroscopy.[25,26]

In the present Chapter, the first application of our FSSH/TD-DFTB implementation will be presented. It is devoted to the non-adiabatic molecular dynamics of highly excited monomers of cata-condensed PAHs. More precisely, the dynamical evolution of electronically excited polyacenes (see Figure 3.1) ranging in size from naphthalene to heptacene is investigated, which are systems of choice as they are known to present a closed-shell ground state electronic configuration[144,169] and their absorption spectra have been shown to be well described both at the TD-DFT[170] and the TD-DFTB[86,90] levels of theory. Moreover, a general discussion about possible contribution of polyacenes to DIBs has been published recently.[171] It is worth mentioning that polyacenes belong to the cata-condensed morphology with the zigzag shape of edges. Yet, the armchair-edge morphology characterizing phenacenes is also of interest and chrysene (see Figure 3.2) is the first in the series with four aromatic cycles. Furthermore, in the case of tetracene (see Figure 3.2) and chrysene isomers, it is worth mentioning that their respective experimental absorption spectra are rather similar, both showing a strong absorption around 270 nm. Given this similarity, the investigation of their respective relaxation dynamics should prove invaluable in unraveling shape- and/or symmetry-dependent effects.

Non-adiabatic dynamics of cata-condensed PAHs is an important and rather challenging process

[1] Published work, see:
[90] E. Posenitskiy *et al.*, Phys. Chem. Chem. Phys. **21**, 12139–12149 (2019)
[168] E. Posenitskiy *et al.*, J. Chem. Phys. **152(7)**, 074306 (2020)

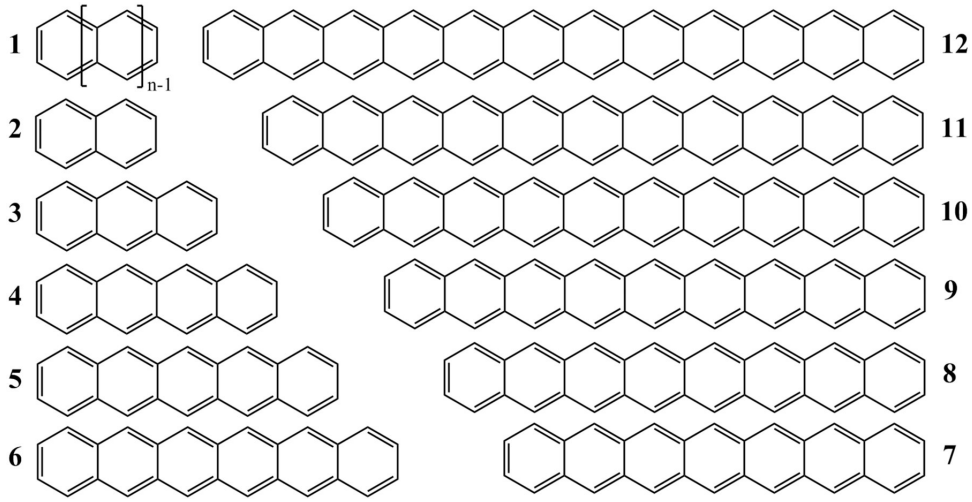


Figure 3.1: Structure of polyacenes in one Kekule resonance form from Ref. [169].

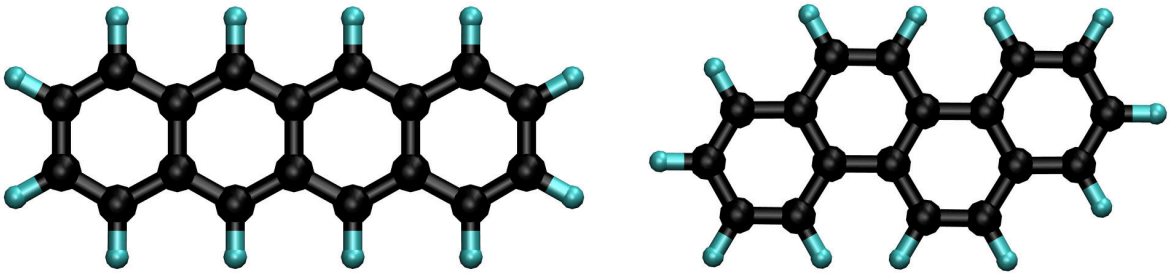


Figure 3.2: The balls and sticks representation of the tetracene (left panel) and chrysene (right panel) isomers. Black and cyan spheres are carbon and hydrogen atoms, respectively.

to investigate. It is particularly interesting to assess whether the commonly used Kasha's rule[172] holds for these compounds. This rule states that the photon emission occurs in a notable yield only from the lowest excited state. It is a common assumption in molecular dynamics simulation of UV-excited complexes. In particular, when there is no access to a reliable Quantum Chemistry method for excited states or when excited state forces are not available. However, deviations from this rule have been reported in several studies.[89,173] Non-adiabatic dynamics launched in the brightest excited states of polyacenes is a good test case due to the fact that their absorption intensity is several times larger than that of any other state. Furthermore, the nuclear dynamics is of great interest due to possible conformational changes and even fragmentation that has been actively studied over the years.[52,174–176] Such mixed quantum-classical simulations are still computationally demanding because excited and ground states have to be propagated along many trajectories with a small nuclear time step. This is only feasible with semi-empirical methods like DFTB, especially for large polyacenes. It is worth mentioning that experimental[25,26] and theoretical[22,23] studies on cationic polyacenes have indicated their possible contribution to DIBs and revealed certain trends in the excited-state dynamics of these compounds.

This chapter is organized as follows: in the next section, I first assess the validity of the TD-DFTB approach for calculating the absorption spectra of polyacenes ranging in size from naphthalene to heptacene, which is followed by more specific benchmark and comparison related to the chrysene and tetracene isomers. In Section 3.3, a detailed analysis of the non-adiabatic molecular dynamics of polyacenes is presented. This is followed by a discussion about an observed size effect on the ultrafast relaxation and the underlying mechanisms. Section 3.4 is dedicated to a similar study on the chrysene and tetracene isomers to emphasize the effect of armchair *versus* zigzag edges. Finally, the conclusions and perspectives are given.

Computational details

It was previously shown that the TD-DFTB scheme is able to represent key features of the absorption spectra of polyacenes.[86] However, these calculations have been repeated first with a different DFT functional and second with a different set of DFTB parameters, mainly to evaluate the accuracy of the latter since it will be used in the FSSH/TD-DFTB simulations. The TD-DFT absorption spectra were computed with the Gaussian 09 package[177] using BLYP functional and 6-31G(d,p) basis set while the TD-DFTB data were calculated with the deMon-Nano[99] code using the matsci-0-3[131] set of parameters. All geometries used in the TD-DFT were optimized with the B3LYP functional and 6-31G(d,p) basis set, TD-DFTB geometries were optimized using the matsci-0-3 parametrization. Our benchmark calculations also show that the DFT geometries optimized with 6-31G(d,p) or 6-311G* basis sets using the B3LYP functional do not differ significantly.

As far as I know, the TD-DFTB absorption spectrum of chrysene has not been published before. This is why it was decided to study it in more details and to compare two commonly used DFTB sets of parameters, namely MIO[121] and MAT[131]. The TD-DFT absorption spectrum of chrysene was computed using BLYP, B3LYP and CAM-B3LYP functionals and 6-31G(d,p) basis set.

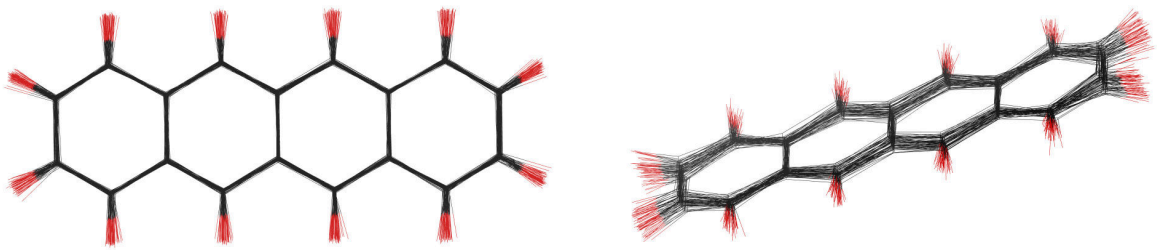
Absorbing bands (vertical sticks) were convoluted based on their oscillator strengths using Gaussian functions with a standard deviation equal to 0.1 eV.

Regarding the FSSH propagation, Eqn. 2.64 is integrated using a 4-th order Runge-Kutta algorithm with an electronic time step $\Delta\tau = 0.048$ as. Each classical trajectory is propagated with $\Delta t = 0.25$ fs during 300 fs. Adiabatic energies and NACs are only available at nuclear time steps t, $t+\Delta t$ and $t+\Delta t/2$, respectively. So in order to integrate Eqn. 2.64 with the smaller time step $\Delta\tau$ the energies are interpolated in the time interval $(t, t+\Delta t)$ and $D_{JK}(\tau) = D_{JK}(t+\Delta t/2)$ for $t \leqslant \tau \leqslant t + \Delta t$.

In the present simulations for polyacenes, gap thresholds of 10 meV have been used to track CoIns between excited singlet states and 100 meV for CoIns between excited singlet and ground states. For chrysene and tetracene, additional calculations have been performed with a 1 meV

threshold for excited states. It is worth mentioning that no CoIn with the ground state occurred during the FSSH simulations presented in this Chapter.

Initial conditions 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 5 Nosé-Hoover thermostats and a 0.5 fs time step. Snapshots (nuclear coordinates and velocities) were taken every 50 fs to be further used as initial conditions for the FSSH. The DFTB parameters were taken from the mio-set.[121] As an example, 63 initial configurations sampled for tetracene at $T = 300$ K are presented in Figure 3.3. An additional constraint has been applied by setting an overall angular momentum of the system to zero at each step. This is crucial for the FSSH since the velocity rescaling scheme (2.73) does not conserve the total angular momentum (unless it is equal to zero). This is not an issue if the energy gradient between two states that contribute to the hopping event is used to rescale the nuclear velocities.[162]



(a) Top view (b) Side view

Figure 3.3: 63 initial configurations of tetracene sampled from the thermal distribution at $T = 300$ K. Black color denotes carbon atoms, red color denotes hydrogen atoms. These figures have been produced with VMD.[178]

In the next Section, a simple exponent $\exp(-t/\tau)$ is used to fit the electronic population curves and to extract the decay times τ. However, this fit does not take into account the non-vanishing population at the end of the simulation (e.g. non-Kasha behaviour). In Section 3.4, a more advanced fitting has been used to extract the decay time τ of the initially excited state (IS):

$$f(t) = A + B\exp(-t/\tau), \tag{3.1}$$

where $A = \min[P_{\mathrm{IS}}(t)]/(1 + \min[P_{\mathrm{IS}}(t)])$; $B = 1/(1 + \min[P_{\mathrm{IS}}(t)])$ and $P_{\mathrm{IS}}(t) = |C_{\mathrm{IS}}(t)|^2$ is the electronic population of the initial state. This new fitting procedure, adapted from the work of Marciniak *et al.*[26], allows one to fit the initial population more accurately due to the shift A, which takes into account the non-vanishing population.

3.2 Analysis of the absorption spectra

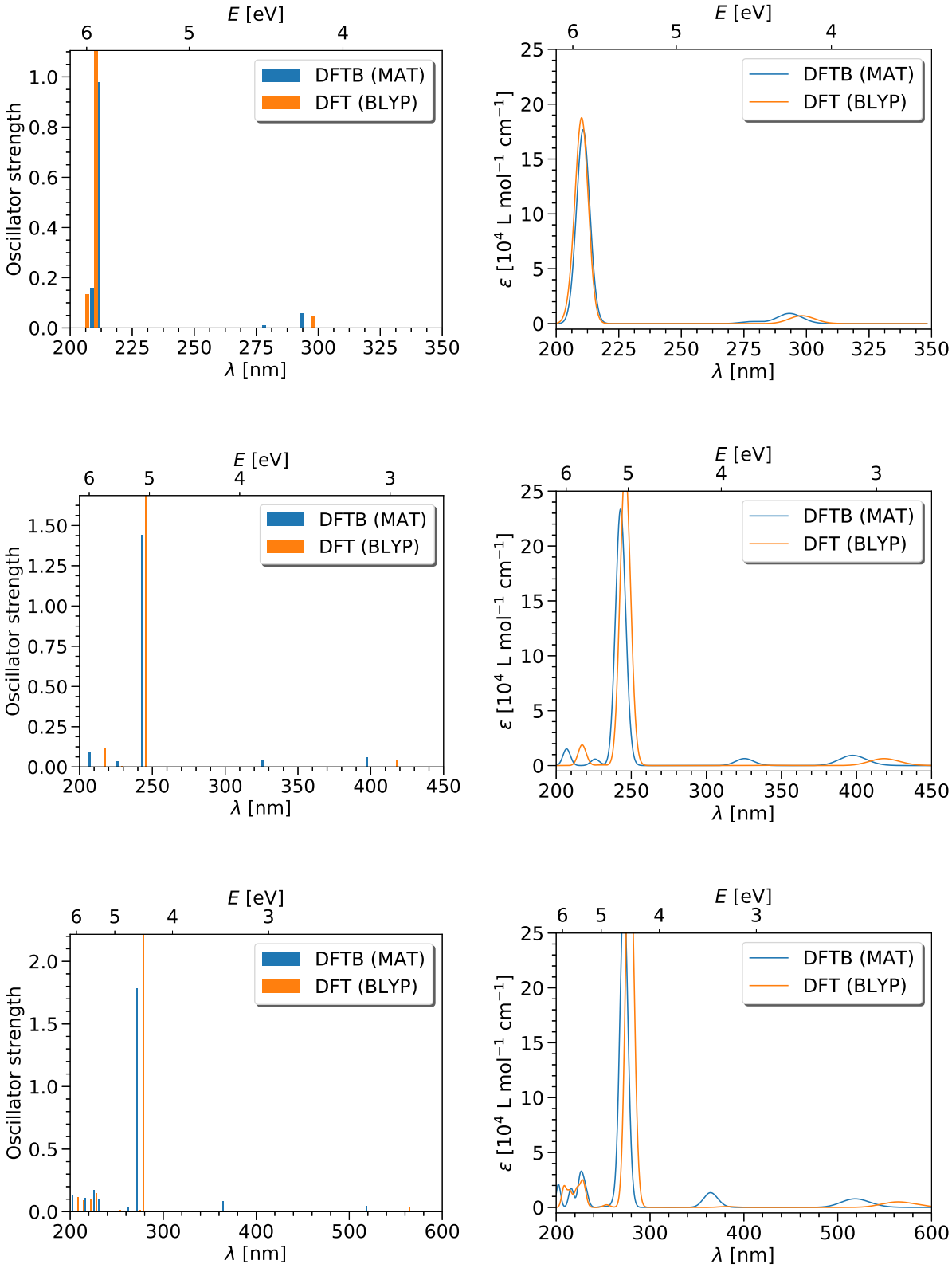


Figure 3.4: Absorption spectra of neutral naphthalene (top row), anthracene (middle row) and tetracene (bottom row) computed with TD-DFT (orange line) and TD-DFTB (blue line) at the equilibrium geometry. All spectra have been truncated at 200 nm.

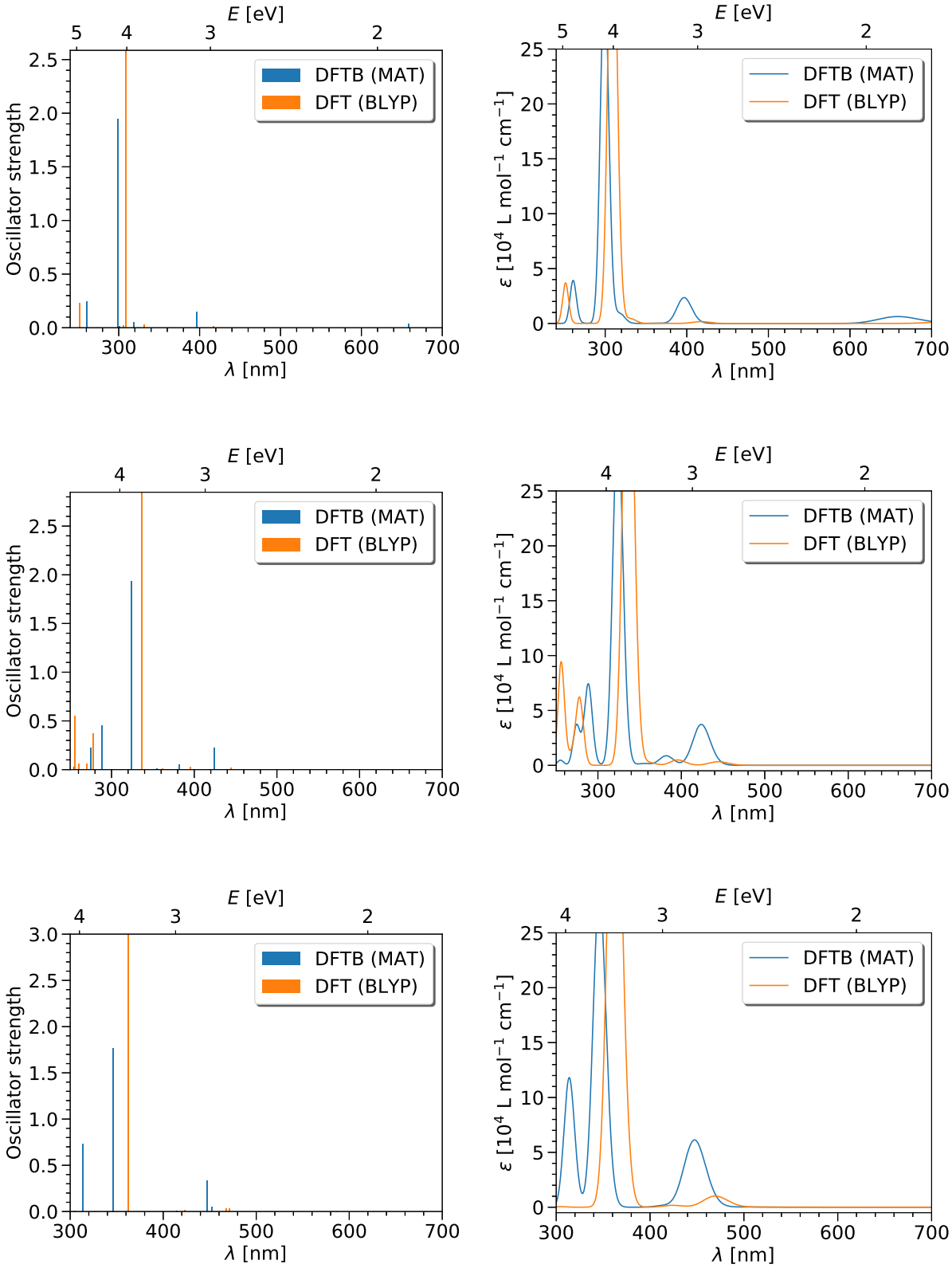


Figure 3.5: Absorption spectra of neutral pentacene (top row), hexacene (middle row) and heptacene (bottom row) computed with TD-DFT (orange line) and TD-DFTB (blue line) at the equilibrium geometry. Spectra have been truncated at 250 nm for pentacene and hexacene; at 300 nm for heptacene.

3.2.1 Polyacenes: from naphthalene to heptacene

Figures 3.4 and 3.5 presented above illustrate the TD-DFT and TD-DFTB absorption spectra of neutral polyacenes. The overall agreement between the two methods is reasonably good, especially taking into account that BLYP is a GGA-based functional and the 6-31G basis set is larger than the one of DFTB. It is worth mentioning that the agreement is particularly good for the brightest excited states, namely that with the largest oscillator strengths. However, the relative position of these states in TD-DFT becomes more and more red-shifted as the size of the compound increases. This is particularly pronounced in the case of heptacene. Notably, TD-DFTB underestimates (compared to the reference TD-DFT results) the oscillator strength of the brightest state for all considered polyacenes.

Polyacenes have the D_{2h} point group symmetry and a ground state belonging to the A_g irreducible representation. Thus, their brightest excited states belong to the B_{3u} irreducible representation if the considered molecules are aligned along x axis. Furthermore, they are dominated by the $\pi \rightarrow \pi^*$ transitions, which are predicted reasonably well by the TD-DFTB approach as has been shown by Niehaus *et al.* in Ref. [86]. Nevetheless, the same group has noticed that TD-DFTB fails to accurately describe $n \rightarrow \pi^*$ or $\sigma \rightarrow \pi^*$ transitions due to an artifical zero coupling induced by the Mulliken atomic transition charges. This is not an issue if the onsite correction of Domínguez *et al.*[179] is used.

The theoretical calculations are further validated based on comparison with available experimental and higher-level theoretical (CASPT2) data. In particular, excitation energies of the lowest and the brightest excited singlet states of polyacenes are summarized in Figure 3.6. Positions of experimental bands for naphthalene, anthracene, tetracene and pentacene are taken from Ref. [181] and references therein, for hexacene and heptacene from Ref. [182], for octacene from Ref. [183].

It is clear that the brightest and the first singlet excited states in TD-DFTB are in good agreement with experimental and CASPT2 values. This is consistent with the previously reported TD-DFTB results computed with a different set of parameters up to heptacene.[86] However, some departure grows with an increasing size of the molecule, i.e. for octacene the differences are 0.53 eV and 0.39 eV for the first and brightest singlet excited states, respectively. Thus, this study is limited to heptacene since the brightest singlet excited state is still predicted reasonably well and underestimates the experimental value by 0.22 eV only.

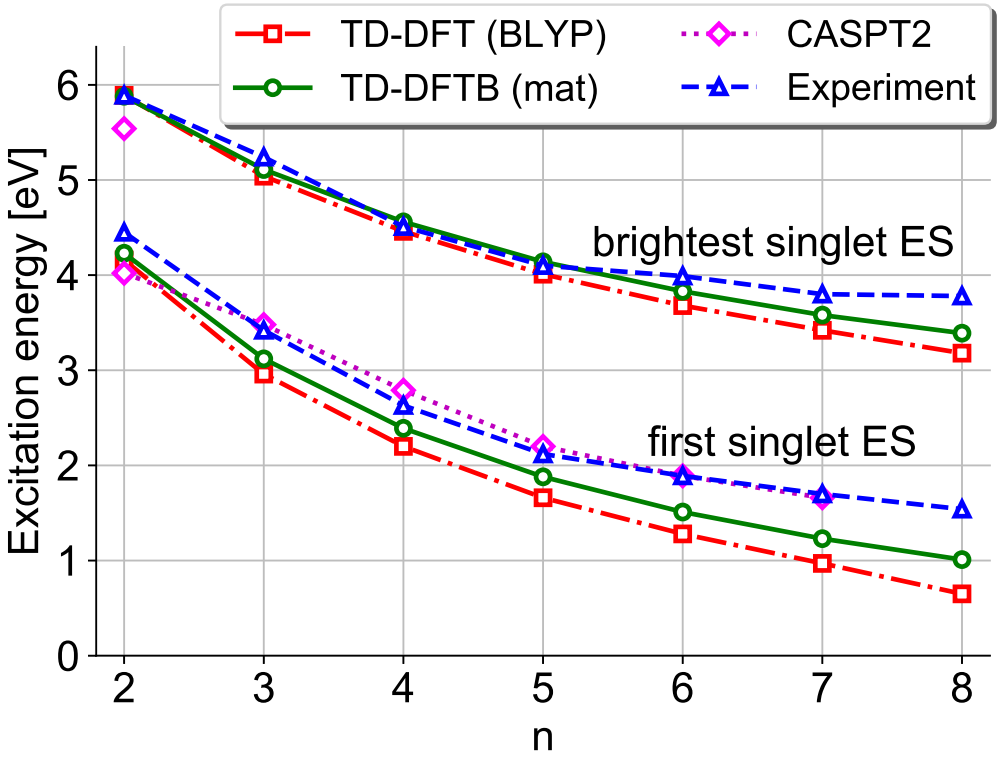


Figure 3.6: Energies of the brightest and first singlet excited states (ES) in the theoretical CASPT2[100,180] (magenta diamonds, dotted line), TD-DFT (red squares, dash-dotted line), TD-DFTB (green circles, solid line) and experimental[181–183] (blue triangles, dashed line) absorption spectra of polyacenes (n is the number of aromatic cycles in the compound). The energy marks are connected by lines to guide the eye. All plotted values can be found in Table S1 (see Appendix B).

3.2.2 Tetracene *versus* chrysene

A qualitative assignment of the absorption spectra of chrysene and tetracene has been presented quite a long time ago.[184,185] I first assess the accuracy of the TD-DFTB absorption spectra computed for chrysene and tetracene with the MAT and MIO sets of parameters. Both systems are considered at their respective DFTB equilibrium geometries, namely with D_{2h} symmetry for tetracene and C_{2h} for chrysene.

In the case of chrysene, the results obtained with MIO and MAT parameters are relatively similar (see Figure 3.7). However, the bands computed with MAT are blue-shifted as compared to the TD-DFT ones. Thus, the MIO set predicts better the positions of some TD-DFT absorbing bands, but the brightest band calculated at 4.61 eV (269 nm) with the MAT set is closer to the experimental one at 270 nm[186]. This value is in good agreement with other experimental measurements: 272 nm in a boric acid matrix[187] and 269 nm in a tetrahydrofuran solution[188]. Both sets of parameters show that the brightest excited state (S_8) is dominated (with a weight of 0.59 and 0.62 for MIO and MAT sets, respectively) by the HOMO–1 → LUMO+1 transition

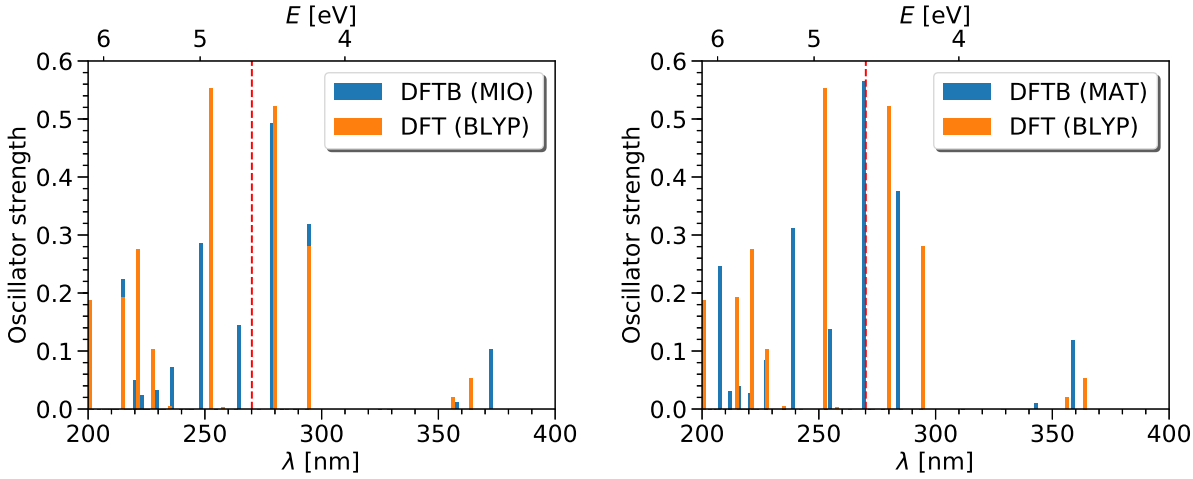


Figure 3.7: Absorption spectra (oscillator strengths) of chrysene computed with TD-DFT (orange sticks) and TD-DFTB (blue sticks) using MIO (left panel) and MAT parameters set (right panel) at the equilibrium geometry. Red vertical dashed line at 270 nm indicates the brightest broad band in chrysene (from the experiment in a hexane solution[186]).

(see Figures 3.8b and 3.8d) with oscillator strengths equal to 0.49 and 0.56 for MIO and MAT sets, respectively. Additional TD-DFT calculations have been performed for chrysene with B3LYP and CAM-B3LYP functionals (see Tables S5 and S6, respectively, in Appendix B). Indeed, TD-DFTB shows poor agreement with these functionals and one can see that both B3LYP and CAM-B3LYP spectra are strongly blue-shifted compared to the TD-DFTB ones. Notably, CAM-B3LYP fails to detect two bright states with relatively close oscillator strengths that are present in TD-DFTB and experimental[186–188] data. It shows a single very bright state with 1.66 oscillator strength at 5.24 eV.

There is a significant change when considering the absorption spectrum of tetracene (see Figure 3.9). The brightest excited states computed with MAT and MIO basis sets are S_7 and S_{10}, respectively, even though the dominant transition (with a weight of 0.55) is HOMO $\rightarrow$ LUMO+2 (see Figures 3.8a and 3.8c) in both cases. This is an indication that the states' splitting induced by the MIO parameters and the TD-DFTB coupling matrix is stronger in tetracene. This results in a shift of some higher-lying excited states below the brightest one. TD-DFT predicts the brightest absorbing band to be S_7 with the same dominant HOMO $\rightarrow$ LUMO+2 transition as in TD-DFTB. The trend of MIO set lifting up the brightest state has been also observed for larger polyacenes. Thus, the MAT parametrization is more reliable for the simulation of the electronic relaxation from the brightest excited state of polyacenes, which is likely to be affected by a wrong topology of excited PES.

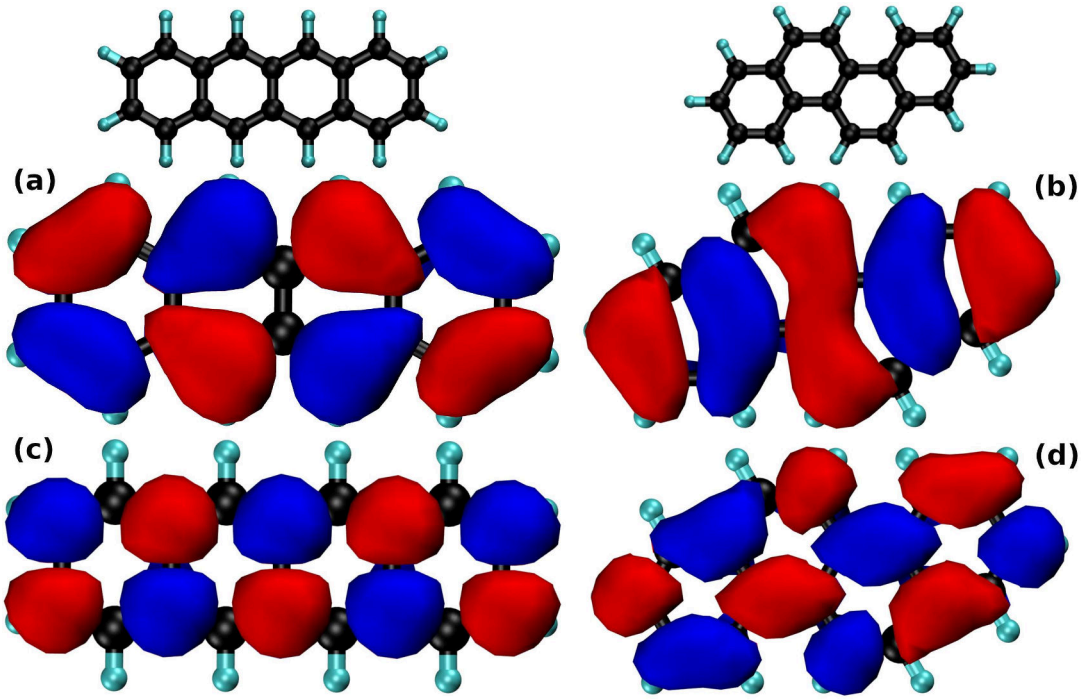


Figure 3.8: Isosurfaces of (a) HOMO and (c) LUMO+2 of tetracene (left column); (b) HOMO–1 and (d) LUMO+1 of chrysene (right column). Black and cyan spheres are carbon and hydrogen atoms, respectively. This figure has been produced with VMD.[178]

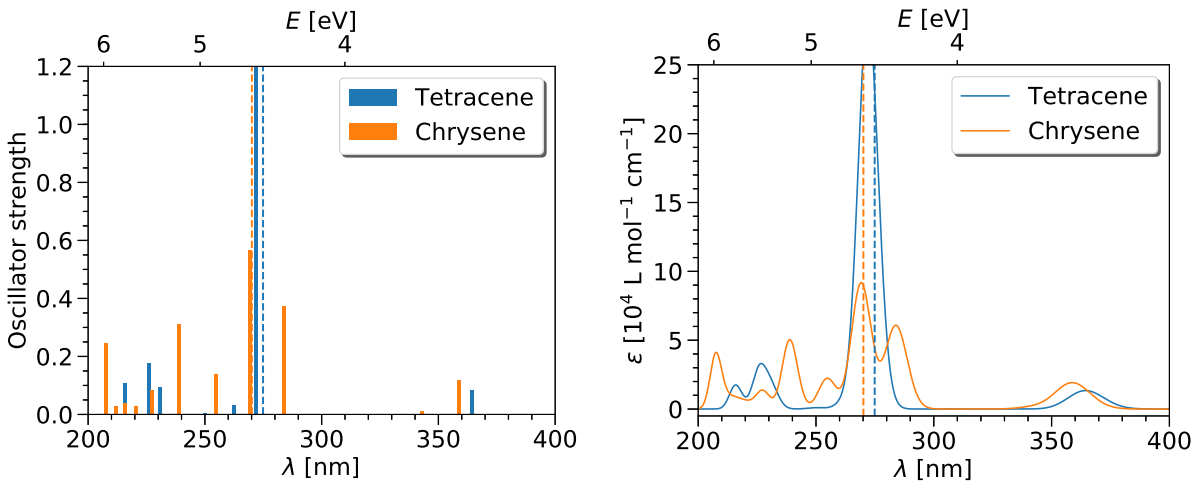


Figure 3.9: Oscillator strengths (left panel) and molar absorption coefficients ε (right panel, convoluted) from the UV absorption spectra of chrysene and tetracene computed with TD-DFTB (MAT parameters set) at the equilibrium geometry. The vertical dashed lines indicate positions of the brightest bands for the corresponding molecule from the experiments.[181,186]

3.3 Size dependence of the ultrafast relaxation

I distinguish between polyacenes with odd and even number of aromatic cycles based on their dynamical behaviour during the relaxation.

3.3.1 Polyacenes: odd number of aromatic cycles

In this section, the results of the FSSH/TD-DFTB calculations are presented for polyacenes with odd number of aromatic cycles. The brightest singlet excited states in the TD-DFTB absorption spectra of anthracene, pentacene and heptacene are S_7, S_{10} and S_{11}, respectively. I start with the analysis of FSSH/TD-DFTB dynamics in anthracene computed with different number of trajectories.

It is concluded from Figure 3.10 that 63 trajectories are enough to fit the decay of the initial state and that increase in the number of trajectories mainly leads to smoothing of the population curves. The general trends of the population transfer (e.g. fast decay of an initial state and the corresponding transfer to lowest singlets) are already represented with 63 trajectories. The decay times extracted from the exponential fits are close to 22 fs in both cases.

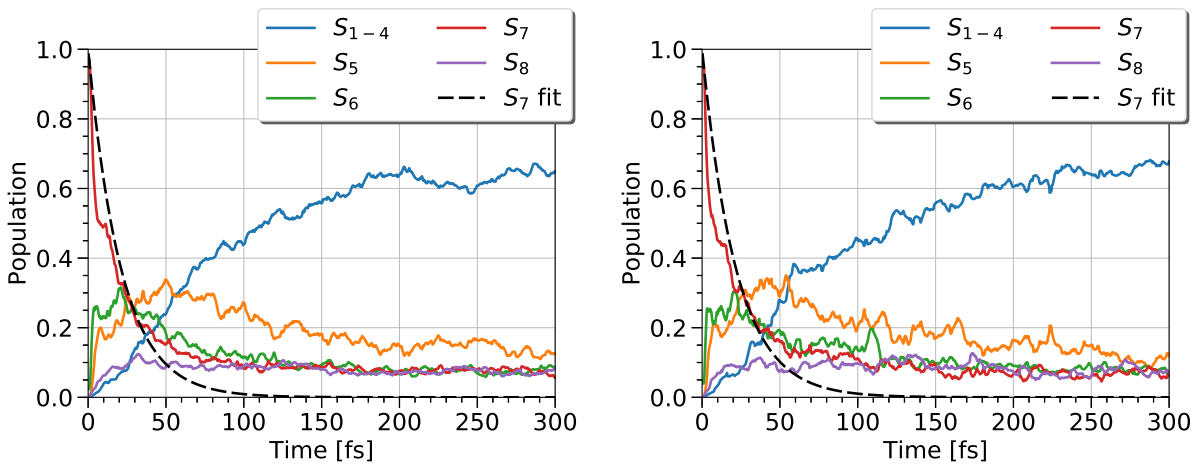
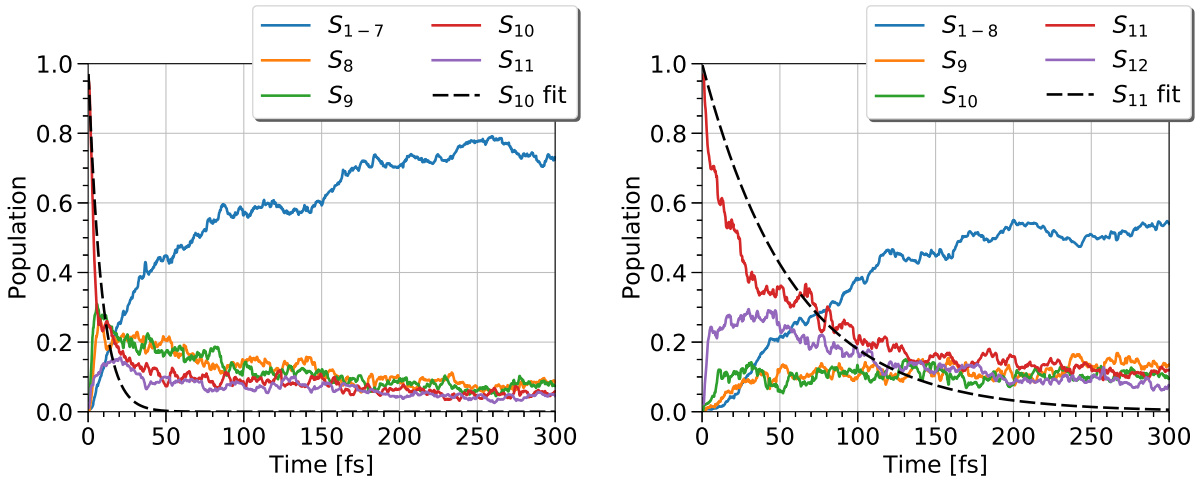


Figure 3.10: Population analysis of the first 8 singlet excited states in anthracene for an ensemble of 127 (left panel) and 63 (right panel) trajectories. The initial state is S_7. The blue curves hereafter represent the sums of averaged populations in the lower-lying singlet states. For instance, the S_{1-4} curve is $|C_1|^2 + |C_2|^2 + |C_3|^2 + |C_4|^2$.

The population transfer in pentacene (see left panel in Figure 3.11) is very similar to the one in anthracene, showing ultrafast decay (within approximately 8 fs) of the initial state and rapid growth of the population in the lower-lying singlets. However, in heptacene (see right panel in Figure 3.11), the latter trend is less pronounced due to the partial transfer of population to S_{12} (see the purple curve during first 50 fs). Thus, a part of the population is trapped for some time in the upper states resulting in a delay of the initial state decay.

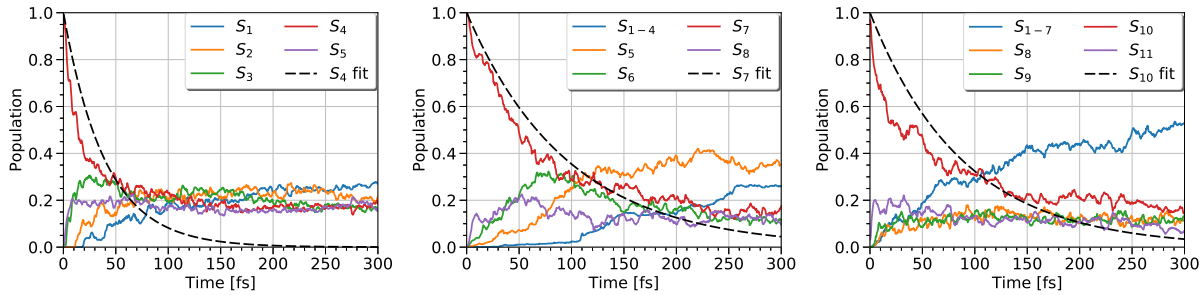


a) Pentacene (S_{10}) b) Heptacene (S_{11})

Figure 3.11: Population analysis of the first 11 and 12 singlet excited states in pentacene and heptacene, respectively, for an ensemble of 63 trajectories. The initial state is indicated between parentheses.

3.3.2 Polyacenes: even number of aromatic cycles

In this section, the results of the FSSH/TD-DFTB calculations are presented for polyacenes with even number of aromatic cycles (see Figure 3.12). The brightest singlet excited states in the TD-DFTB absorption spectra of naphthalene, tetracene and hexacene are S_4, S_7 and S_{10}, respectively.



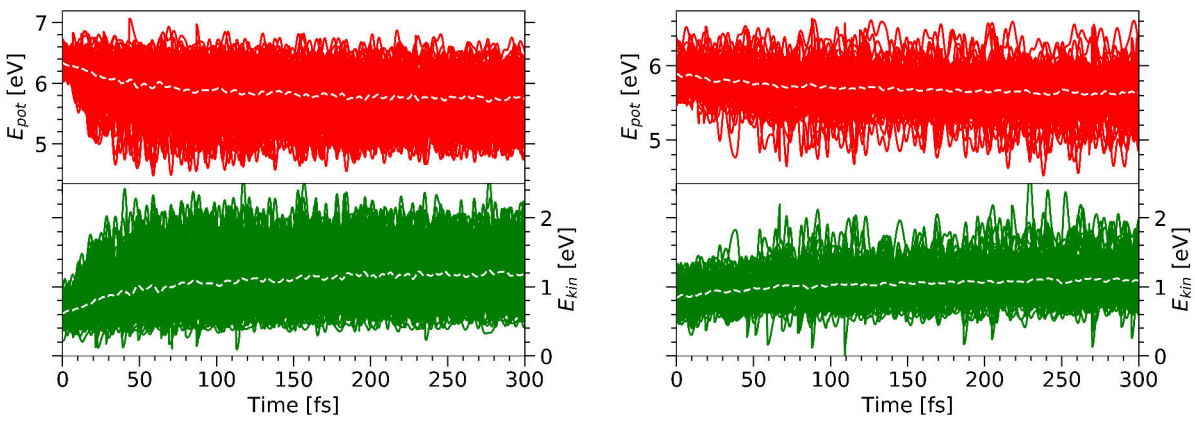
a) Naphthalene (S_4) b) Tetracene (S_7) c) Hexacene (S_{10})

Figure 3.12: Population analysis of the first 5, 8 and 11 singlet excited states in naphthalene, tetracene and hexacene, respectively, for an ensemble of 127 trajectories (naphthalene) and 63 trajectories (tetracene, hexacene). The initial state is indicated between parentheses.

The electronic population in tetracene is mainly shared between S_5, S_6, S_7 and S_8 during the first 70 fs and then S_5 (orange line in Figure 3.12b) is monotonously populated until it reaches approximately 40% of population. In hexacene, the main relaxation channel goes through the lower-lying singlet states (blue line in Figure 3.12c) after 50 fs, while population of each of the neighbouring states (S_8, S_9, S_{11}) amounts to about 10%. Nevertheless, both tetracene and

hexacene demonstrate slow decay (96 and 89 fs, respectively) of the initial singlet excited state for which the population drops below 20% after 200 fs of propagation.

Analysis for naphthalene (see Figure 3.12a) is slightly more complicated since the populations of all excited states involved in the propagation reach a plateau around 20%. The main difference is that S_1 and S_2 start sharing approximately 50% of the population after 200 fs. Thus, naphthalene has the highest energy transfer rate among all considered polyacenes. This can be also seen if one plots the decrease of the electronic potential energy *versus* the corresponding increase of the total nuclear kinetic energy along the trajectory. On average, 0.6 eV is transferred from electrons to nuclei in naphthalene after 300 fs of propagation, while in anthracene only 0.25 eV is deposited into the nuclear degrees of freedom (see Figure 3.13).



a) Naphthalene b) Anthracene

Figure 3.13: Electronic potential energy (red color, upper panel) and total nuclear kinetic energy (green color, bottom panel) along a swarm of 127 trajectories in naphthalene and anthracene. The white dashed line is the average value for each panel.

3.3.3 Discussion

Figure 3.14 displays the decay time of the brightest singlet excited state as a function of the number of aromatic cycles. All values were extracted from Figures 3.10–3.12. Although the decay is not fully exponential, our approximate fits on the full time window allow for a quantitative comparison between polyacenes. It is worth mentioning that decay times plotted in Figure 3.14 are within the same order of magnitude (except for naphthalene) as the ones reported for cationic species by Reddy *et al.* in Ref. [23], even though cationic states crossings could be qualitatively different from those in neutral molecules.[18,19] However, the present results exhibit an unexpected and striking alternation of the decay times for an even-odd number of cycles in the range of $n = 3$–6. From the organic photochemistry perspective[5], one could expect a monotonous decrease of the relaxation time as a function of size governed by energy gap configurations and increasing density of states.

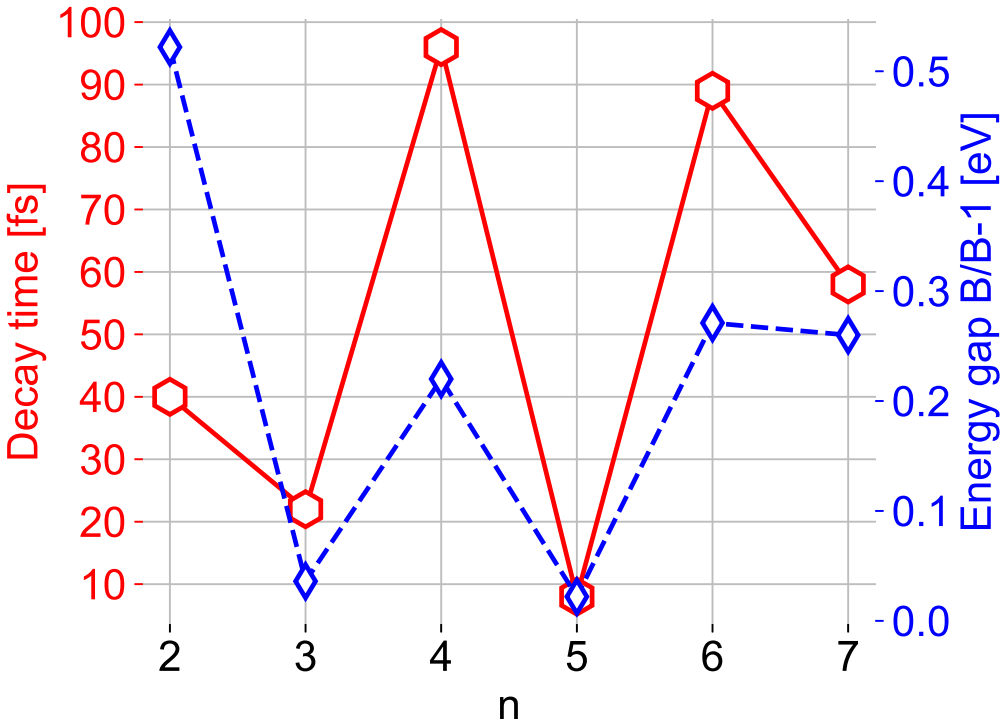


Figure 3.14: Decay time of the brightest (B) singlet excited state (red hexagons) and the energy gap (blue diamonds) between the B and B–1 (the one with lower energy in the TD-DFTB absorption spectra) states computed at the equilibrium geometry as a function of number of aromatic cycles in the polyacene.

Table 3.1: Comparison of decay times extracted from two different exponential fits of electronic populations plotted in Figures 3.10–3.12.

n-acene	τ, fs	τ_0, fs	State
2	22	40	S_4
3	17	22	S_7
4	65	96	S_7
5	8	8	S_{10}
6	51	89	S_{10}
7	40	58	S_{11}

Initially, a simple exponent $\exp(-t/\tau)$ has been used to fit the population curves and to compute the decay times for polyacenes. These results have been presented in Ref. [90]. This fitting procedure does not take into account the non-vanishing part of the population at the end of the simulation. Table 3.1 contains a new set of decay times τ computed using a shifted exponent according to Eqn. 3.1 and a set τ_0 that has been computed with a simple exponent $\exp(-t/\tau_0)$ (see Figure 3.14). Both fitting procedures have been applied to the populations presented

in Figures 3.10–3.12. The alternation of the decay times remains present, however with less variation observed for tetracene and hexacene. The improvement brought by the new fitting procedure is especially pronounced in the case of tetracene, as can be seen in Figure 3.18 (see next Section).

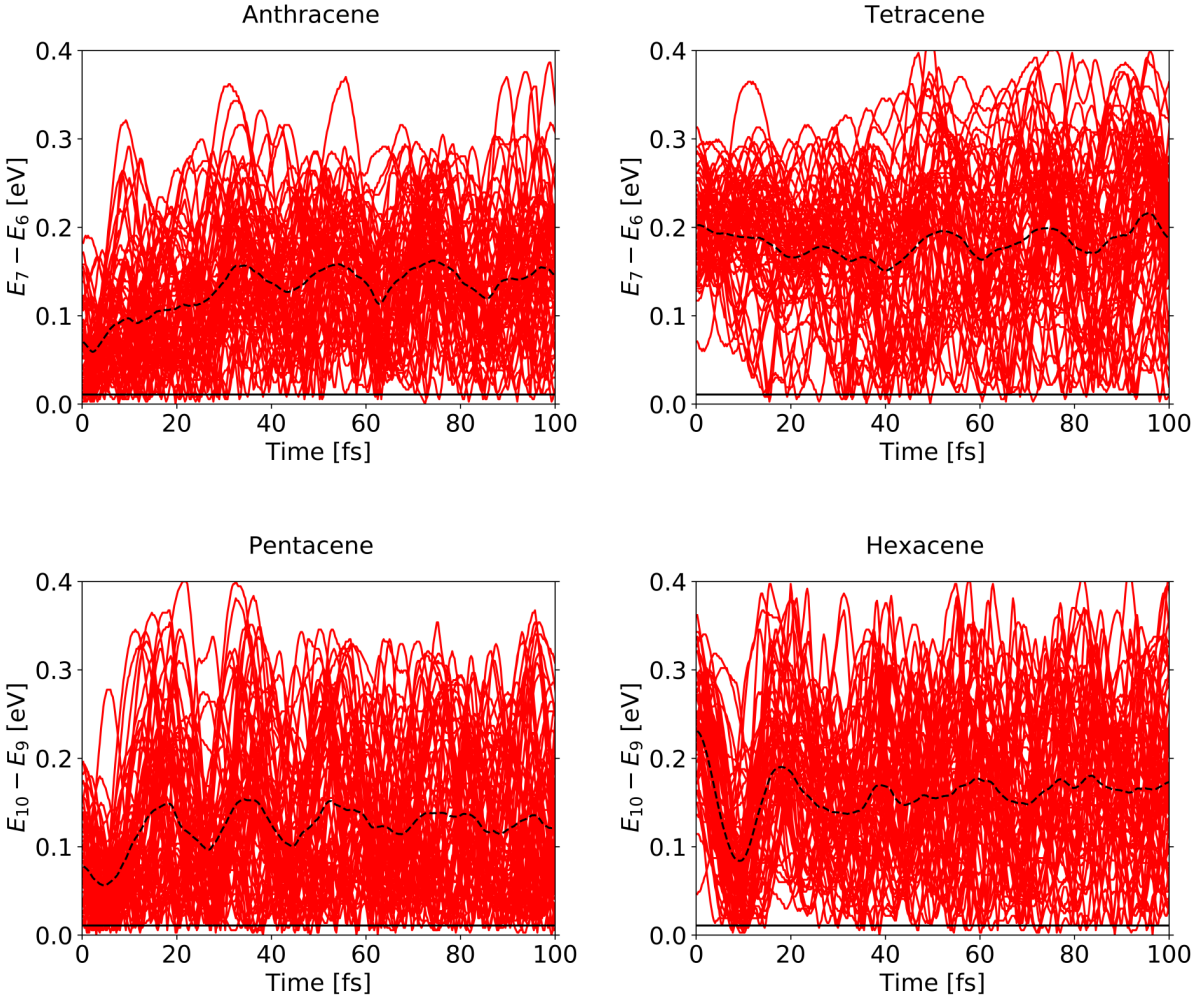


Figure 3.15: Energy gap between the B and B–1 states plotted along each trajectory (red lines) and averaged over the ensemble of trajectories (black dashed line) for a set of polyacenes. The horizontal black line at 0.01 eV is the CoIn threshold.

In order to explain this alternation, the energy gap between the initial state and the neighbouring state with lower energy has been computed for polyacenes with $n = 3$–6 aromatic cycles. It provides a good understanding of how the initial state population transfer occurs due to the fact that during relaxation most of the trajectories either undergo surface hopping in the region of avoided crossing with the state below or reach the CoIn with this state. It is clear from Figure 3.15 that during the first 20 fs a large fraction of trajectories in anthracene and pentacene either go to the region of small gaps or through the CoIn. Thus, there is a funnel in both polyacenes, which drives the rapid population transfer from the initial state. On the contrary, in tetracene, very few trajectories reach the region of small gaps over 20 fs and the averaged value is almost

two times larger than the one in anthracene. The latter is also true for hexacene, for which there is a clear drop of the energy gap around 10 fs, yet remaining larger than in pentacene.

The TD-DFTB calculations of the absorption spectra at the equilibrium geometry have been performed for nonacene and decacene (see Figure 3.16) to see whether the energy gap variation between the brightest excited state and the one below still occurs for larger polyacenes. Notably, the alternation of decay times cannot be explained with a simple single particle excitation model. This can be seen from the red curve in Figure 3.16, which represents the energy gap $\omega_{ia} = \varepsilon_a - \varepsilon_i$ of the dominant transition as a function of size before the TD-DFTB coupling matrix shuffles all single particle excitations. The general trend is a monotonous growth (except for anthracene and octacene) with the number of aromatic cycles. However, the energy gaps after TD-DFTB shuffling (blue curve in Figure 3.16) significantly differ between heptacene and octacene. Thus, one can expect a faster decay of the brightest singlet excited states for $n = 8, 9, 10$ as compared to $n = 6, 7$.

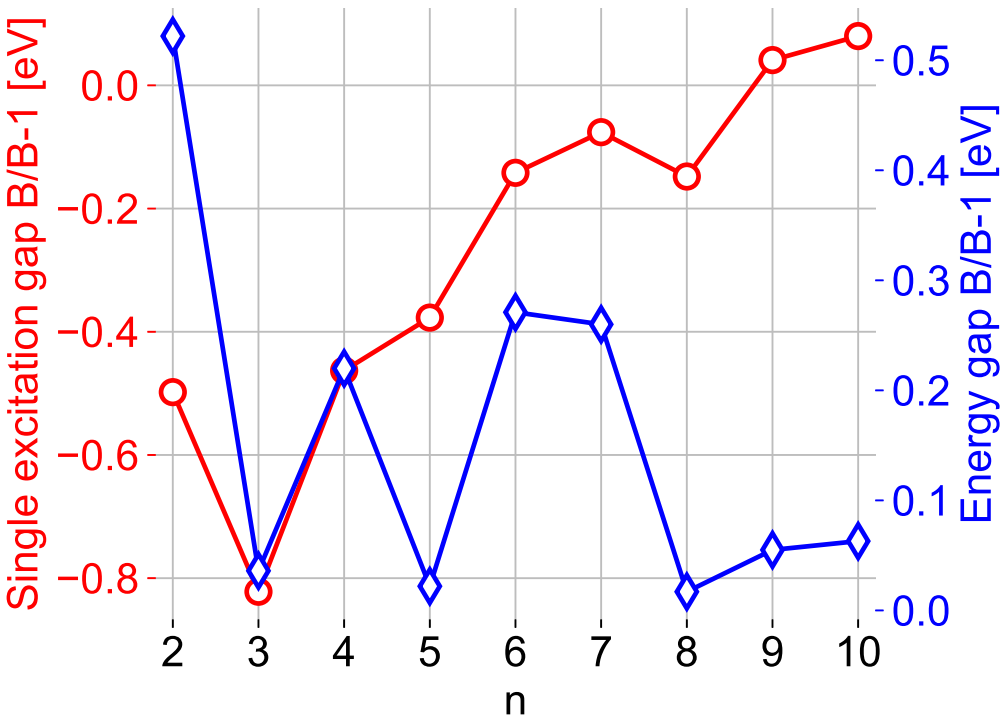


Figure 3.16: Energy gaps between the B and B–1 singlet excited states at the equilibrium geometry of polyacenes before (red circles) and after (blue diamonds) shuffling induced by the TD-DFTB coupling matrix.

Our results for polyacenes violate the Kasha's rule for the ultrafast decay to the S_1 state. There are several possible reasons for this: (i) for isolated molecules, the lack of dissipative modes to the environment results in the possibility of population back-transfer to higher-lying excited states from the underlying ones;[89] (ii) 300 fs is still a rather short timescale for the

complete decay of higher-lying excited states, thus some trajectories remain in the states above S_1 at the end of the simulation; (iii) conjugated systems like polyacenes have some low-lying excited states that are dominated by double excitations[189], which are missing in the linear response TD-DFTB. Decoherence correction could be thought of as another reason behind the non-vanishing populations above S_1. The calculations for naphthalene have been repeated without the decoherence correction (see Figure 3.17). Not only the trend remains the same but furthermore the S_1 population that ends as the largest state population in Figure 3.12a, remains the smallest one without correction, thereby enhancing the violation of Kasha's rule. Last, it is worth mentioning that experimental evidences of non-Kasha behaviors are well documented, in particular for small polyatomic molecules including naphthalene.[173]

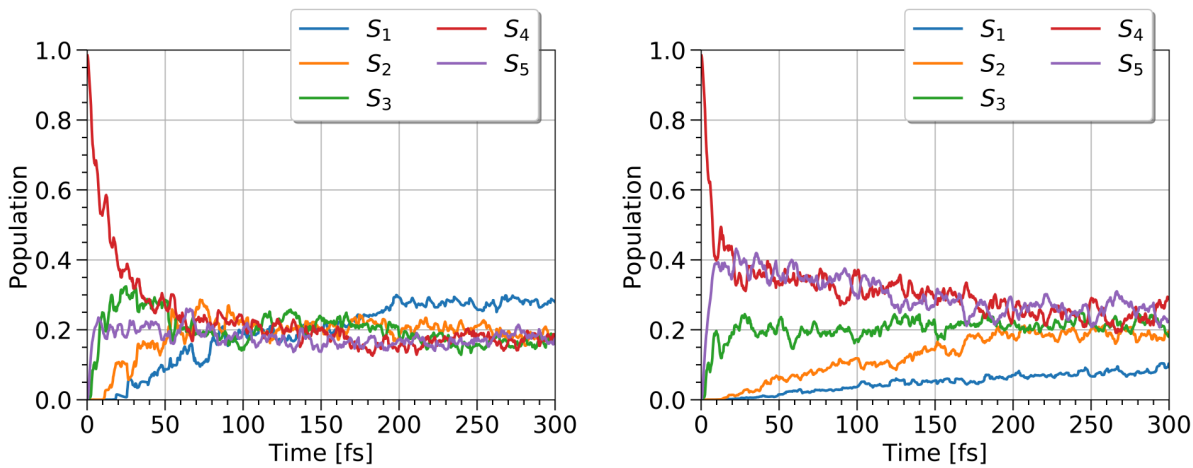


Figure 3.17: Population analysis of the first 5 singlet excited states in naphthalene for an ensemble of 63 trajectories with (left panel) and without (right panel) the decoherence correction. The initial (brightest) state is S_4.

Intersystem crossings (ISC) and the corresponding dynamics in excited triplet states have not been studied here. The efficiency of ISC in pure aromatic compounds has been debated in the literature.[190,191] Spin-orbit coupling is relatively small in molecules involving light atoms only. Yet, it has been shown that ISC could be significantly enhanced, not only in the presence of heavy atoms[192], but also upon departure from planarity in aromatic compounds associated with radicals.[193] Addressing ISC would require the coupling of both triplet and singlet TD-DFTB states in the dynamics *via* the spin-orbit coupling, which is not yet implemented in our code and would be computationally much more demanding. It is worth mentioning the SHARC software package[194] that is capable of running FSSH simulations for both singlet and triplet excited states and incorporates all types of interstate transitions. Last, the ISC timescale is most likely to be larger than our simulation timescale by several orders of magnitude.

3.4 Shape dependence of the ultrafast relaxation

It is worth mentioning that the brightest absorbing band is located close to 4.6 eV in both chrysene and tetracene (see Figure 3.9). The results of FSSH/TD-DFTB simulations launched in these particular states will be presented in order to compare the non-adiabatic dynamics as well as computed decay times. In particular, the shape of edges may play a certain role in the deactivation mechanisms.

3.4.1 FSSH dynamics of highly excited tetracene *versus* chrysene

The FSSH results obtained with 63 trajectories launched in the brightest excited state of tetracene have been reported in the previous section. These calculations have been repeated with 100 trajectories to assess the convergence as well as the accuracy of our results. Population curves averaged over 100 trajectories (see Figure 3.18) are qualitatively similar to the ones reported before for 63 trajectories and the computed decay time of 64 fs is in very good agreement with 65 fs calculated using a fitting procedure from Eqn. 3.1.

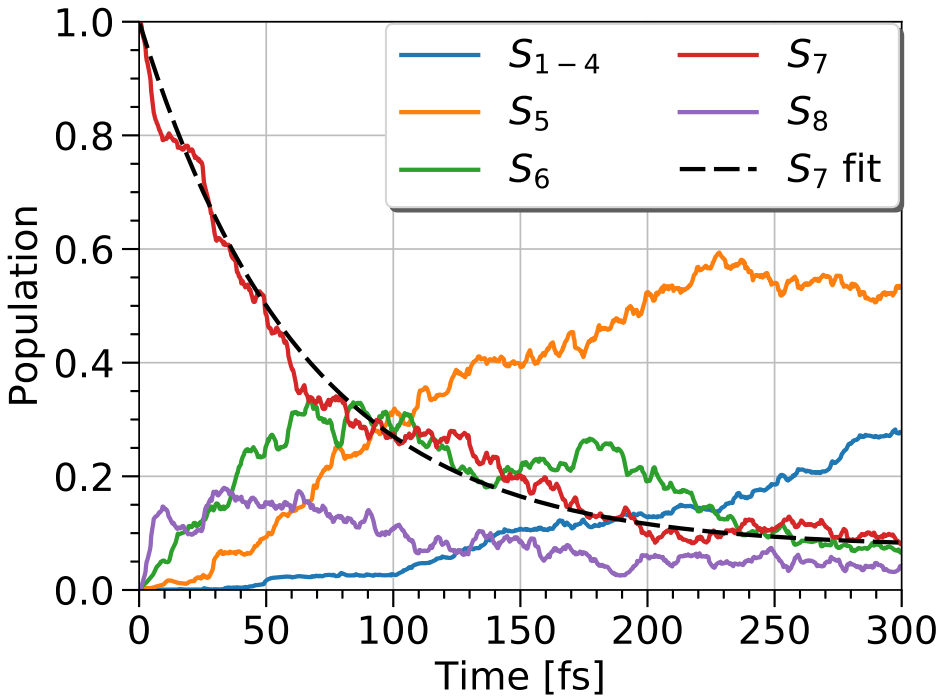
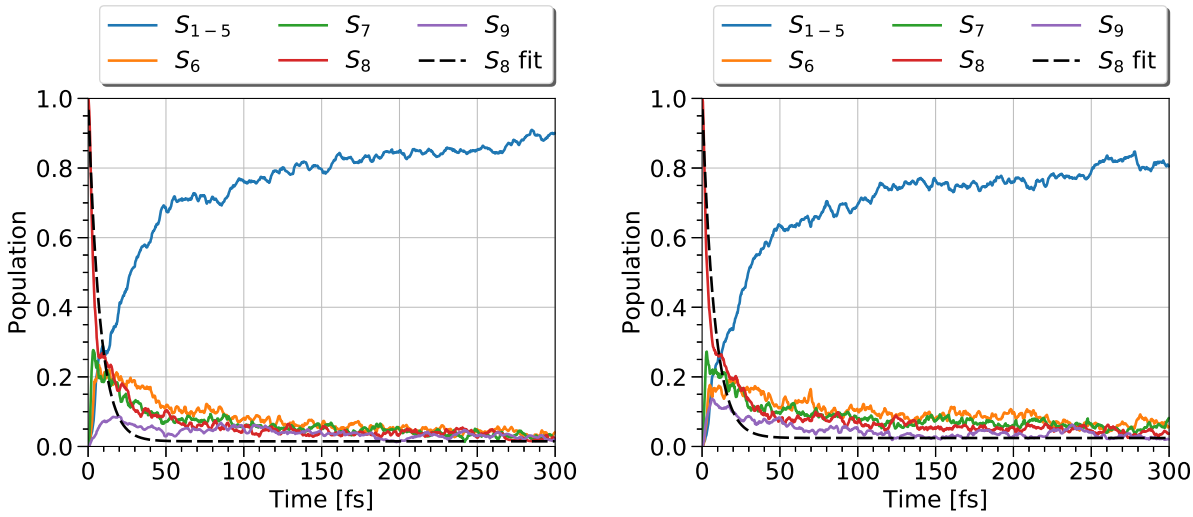


Figure 3.18: Population analysis of the first 8 singlet excited states in tetracene for an ensemble of 100 trajectories. The initial state is S_7.

Relaxation from the brightest excited state of chrysene (see Figure 3.19a) differs drastically from the one in tetracene. First of all, the computed decay time is 7 fs, which is an order-of-magnitude smaller than in tetracene. This is consistent with the rapid transfer of the initial population to the bunch of low-lying excited states, which ends up about 2% at the end of propagation

compared to 8% in tetracene. In order to demonstrate that an observed ultrafast relaxation in chrysene is not a direct consequence of the selected energy gap threshold, additional calculations have been performed under the same conditions, yet changing the threshold to 0.1 or 10 meV (see Figure S1 in Appendix C). The population curves are not drastically affected by these changes and the computed decay time of the S_8 state is approximately 7 fs in both cases, which is in very good agreement with the value computed for the 1 meV threshold. Calculations have been also performed for chrysene with the MIO set of DFTB parameters, using an ensemble of 89 trajectories launched in the brightest excited S_8 state and propagated during 300 fs (see Figure 3.19b). The computed decay time is approximately 8 fs, which is in good agreement with 7 fs calculated with the MAT parameters.



a) MAT set

b) MIO set

Figure 3.19: Population analysis of the first 9 singlet excited states in chrysene for an ensemble of a) 100 trajectories propagated with MAT parameters; b) 89 trajectories propagared with MIO parameters. The initial (brightest) state is S_8 in both cases.

It is worth mentioning that the fractional occupation (percentage of trajectories running in a given state at a given time) of the state below the brightest one is 7% in tetracene and 3% in chrysene after 300 fs of propagation. This can be due to a partial trapping of the population in some higher-lying excited states of tetracene, which is also consistent with the delayed decay of the initial state.

3.4.2 Discussion

The results of our calculations for chrysene and tetracene are summarized in Table 3.2.

In order to explain the difference in decay times and underlying relaxation mechanisms, the energy gaps between the initial state and the neighbouring one with lower energy have been computed along each trajectory for both tetracene and chrysene (see Figure 3.20). It is clear

Table 3.2: Summary of the FSSH simulations. Energies and positions of excited states are taken from the TD-DFTB spectra computed with the MAT set of parameters at the equilibrium geometry. IS stands for the initally excited (also the brightest) state.

Molecule	IS	Ω_{IS}, eV	N_{traj}	Decay time of IS, fs	Energy gap IS//IS–1, eV
Chrysene	S_8	4.61	100	7	0.05
Tetracene	S_7	4.56	100	64	0.22
Tetracene[a]	S_7	4.56	63	65	0.22

[a] Adapted from Ref. [90]

that at the start of dynamics, all trajectories in chrysene either go to the region of small gaps or through the CoIn. Thus, there is a funnel, which drives the rapid population transfer from the initial state. On the contrary, in tetracene, very few trajectories reach the region of small gaps and the average gap remains two times larger than the one in chrysene. This is consistent with our previous conclusion for polyacenes. It also correlates well with the energy gap computed at the equilibrium geometry (see Table 3.2). This initial gap turns out to be roughly conserved during the short time phase of the trajectory and can be used as a first order approximation for estimating the decay times. Consistently, the number of surface hops is larger in chrysene than in tetracene. In addition to the notably-differing relative gaps between excited states, the symmetry of the initial chrysene geometry (C_{2h}) is lower than that for tetracene (D_{2h}), and therefore yields more non-zero non-adiabatic coupling elements, which may also enhance the electronic relaxation and thus contribute to the much faster decay.

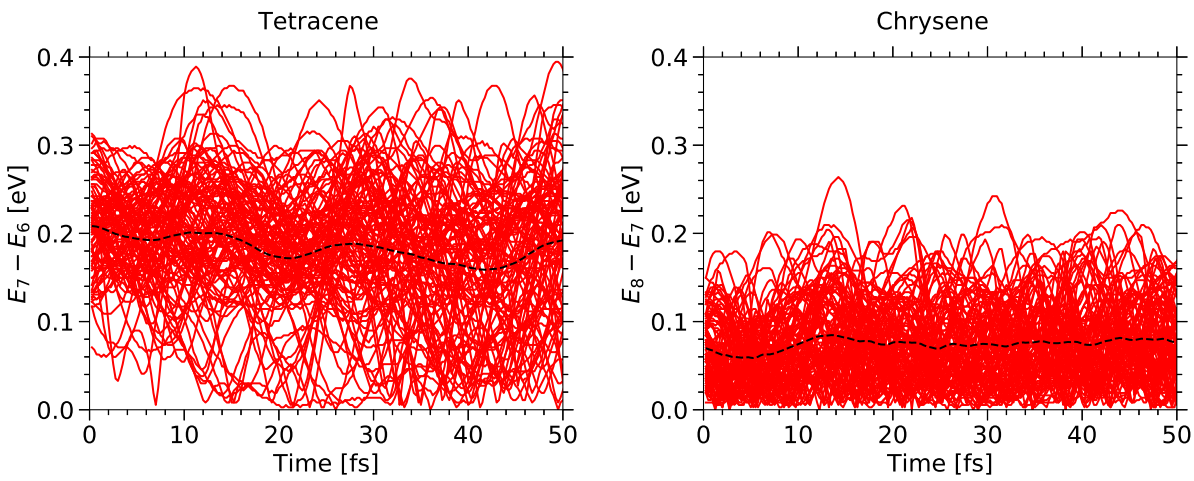


Figure 3.20: Energy gaps between the B and B–1 states plotted along 100 trajectories (red lines) and averaged over the entire ensemble (black dashed line) for tetracene and chrysene.

Results of our simulations can also be used to qualitatively analyze the emission of photoexcited chrysene and tetracene. Fractional occupation of each excited singlet state at the end of the simulation is summarized in Figure 3.21. In tetracene, the major fraction (53%) of trajectories

ends up in excited state S_5 with 24% shared equally between states S_2 and S_3 after 300 fs of propagation. This partial trapping in S_5 state is pronounced and can lead to a significant delay in the non-adiabatic relaxation or even induce a photon emission. However, in chrysene the situation changes drastically and 36% of trajectories have been observed in S_1, 24% shared equally between S_2 and S_3 and 27% in S_4 at the end of simulation.

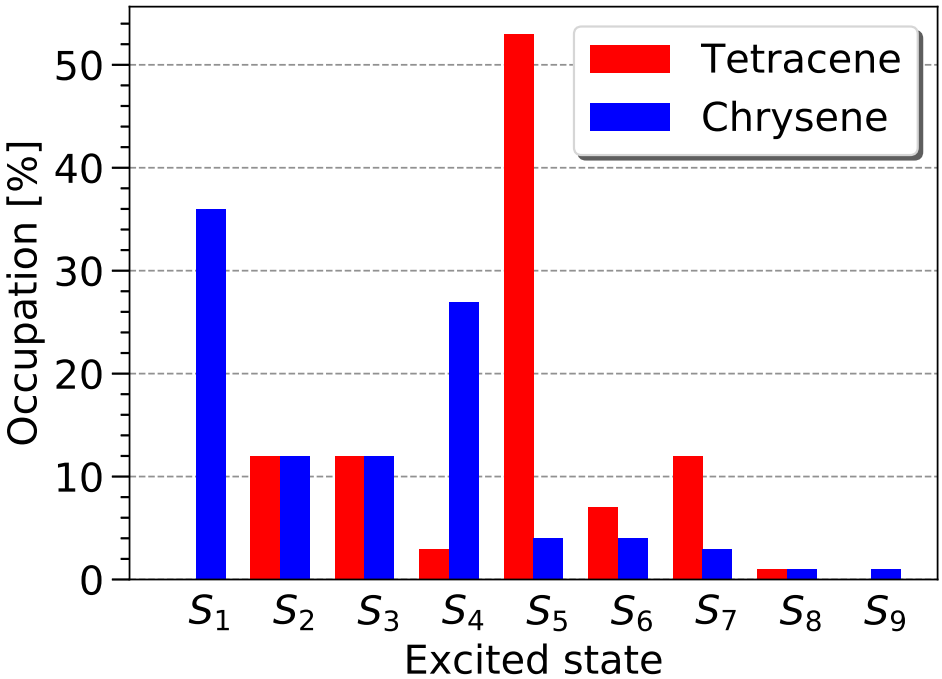


Figure 3.21: Occupation of each excited singlet state in tetracene and chrysene after 300 fs of FSSH propagation.

Finally, the role that the nuclear motion plays in the relaxation of excited chrysene and tetracene can be analyzed. To address this question, the geometrical configurations have been extracted at the end of each of 100 FSSH-trajectory simulations and compiled in Figure 3.22 for both isomers.

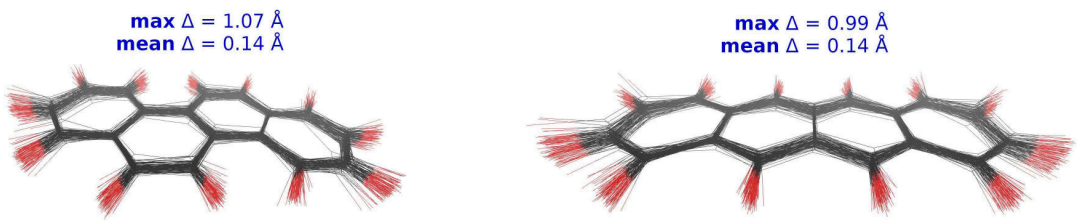


Figure 3.22: 100 configurations of chrysene (left panel) and tetracene (right panel) at the end of the FSSH simulations. The maximum and mean displacements (Δ) along the z axis are indicated with respect to the DFTB ground state equilibrium geometry. Black and red colors denote carbon and hydrogen atoms, respectively. These figures have been produced with VMD.[178]

It is clear that both chrysene and tetracene experience strong out-of-plane motion, which is the major deactivation channel during the relaxation. This motion breaks the corresponding symmetry of the molecule, thereby increasing the coupling between excited states. Nuclear displacements along the x, y and z axes have been computed for each geometry with respect to the DFTB ground state equilibrium configuration. The mean (maximum) displacements along x and y (in the molecular plane) are respectively 0.04 (0.31) Å and 0.05 (0.31) Å in chrysene and 0.04 (0.33) Å and 0.04 (0.30) Å in tetracene. Notably, the mean (maximum) displacements along the perpendicular axis z are 0.14 (1.07) Å and 0.14 (0.99) Å in chrysene and tetracene, respectively. These values are significantly larger than the ones computed along x and y, which is consistent with the strong out-of-plane perturbations visible in Figure 3.22. Considering polyacenes and phenacenes as rather rigid molecules (at least concerning their carbonaceous skeleton), the out-of-plane motion is expected to remain the leading deactivation channel during the relaxation of larger cata-condensed PAHs.

3.5 Conclusion

In this Chapter, a detailed theoretical study of the non-adiabatic molecular dynamics launched in the brightest singlet state has been presented for a set of polyacenes and for chrysene. The results have been obtained using Tully's FSSH scheme coupled to the TD-DFTB formalism for the electronic structure calculations within the deMon-Nano code.

In Section 3.2, the validity of TD-DFTB has been assessed by calculating the absorption spectra of polyacenes and upon comparison with TD-DFT, CASPT2 and available experimental data. The agreement is very good, especially taking into account the number of approximations behind the TD-DFTB formalism. This has been followed by a more detailed analysis dedicated to the tetracene and chrysene isomers. Furthermore, the accuracy of two sets of DFTB parameters, namely MAT[131] and commonly used MIO[121], has been evaluated. The conclusion is that the MAT set is more reliable, especially when the brightest excited singlet state of cata-condensed PAHs has to be computed.

In Section 3.3, the relaxation from the brightest singlet excited state *via* the cascade of radiationless transitions is reported for polyacenes ranging in size from naphthalene to heptacene. The detailed analysis reveals a striking alternation of decay times computed for linear PAH species with $n = 3$–6 aromatic cycles. This alternation is correlated with the energy gap between the initial state and the neighbouring one with lower energy. Yet, an even-odd alternation in the population transfer cannot be extrapolated to larger systems, even though this effect is well pronounced for polyacenes with 3 to 6 aromatic cycles. Further investigation is possible, but requires some improvements in the DFTB parametrization to better match the absorption spectra of octacene and larger species. Also, the computational cost of FSSH calculations for larger polyacenes will increase significantly due to the rapidly increasing amount of states to be

taken into account.

In Section 3.4, the relaxation from the brightest singlet excited state is reported for armchair-shaped chrysene and zigzag-shaped tetracene. The detailed analysis reveals an order-of-magnitude difference in computed decay times between the two molecules. This difference is correlated with the energy gap between the initially excited state and the neighbouring one with lower energy, which is consistent with our previous study on polyacenes. The much faster decay rate for armchair-edge chrysene may also be enhanced by the lesser symmetry induced with respect to zigzag-edge tetracene.

One should precise the limits of the present conclusions. The TD-DFTB spectrum neither involves doubly excited states, nor Rydberg states. Also, interband transitions *via* the spin-orbit coupling have not been considered here. They are not supposed to be dominant since spin-orbit coupling is expected to be small for light elements, nevertheless they have been mentioned to be active in some specific cases.[190,191,193] However, the main conclusions are expected to hold for short relaxation times.

This study contributes to the understanding of photophysics of PAHs and sheds light on the photostability of considered cata-condensed compounds following absorption of a UV photon. It also provides insights into possible effects of the size and/or shape of the corresponding molecular complex on excited-state dynamics. As far as I know, this study is the first to investigate the electronic relaxation from the higher-lying singlet states of large neutral PAH species. These results are of interest for astrochemistry, atmospheric chemistry and laboratory experiments concerned with atto- or femto-second laser spectroscopy of these compounds. It is worth mentioning that larger polyacenes have low-lying excited states with excitation energies in the visible UV range. Thus, these compounds can be observed and assigned if they survive the strong UV radiation. Assessment of the Kasha's rule is important in this context. Referring more particularly to the species in the interstellar medium, electronic absorption band widths can be related to the relaxation timescales.[22] Both the observed size- or shape-dependent relaxation mechanisms and the different final state populations might contribute to the assignment of DIBs. Furthermore, studies with higher-level electronic structure methods will shed more light on the topology of excited states. Finally, it would certainly be interesting to investigate whether the observed differentiation of the ultrafast relaxation is an isolated case property or if it extends to larger cata-condensed PAHs.

4

Testing Machine Learning potentials for simplified non-adiabatic dynamics

4.1 Introduction

Up to date, the humanity has accumulated an enormous amount of data. In fact, some estimates indicate that more than 90% of currently available data have been produced in the last couple of years. While conventional statistical approaches work well on medium-sized datasets, Machine Learning (ML) represents a good alternative for analysing and exploiting large amounts of information. It is commonly associated with Artifical Intelligence (AI) applications.

There is a large variety of available ML approaches and of their applications. As indicated in Figure 4.1, ML methodologies may differ depending on the objective. There are three major branches: (i) supervised learning, (ii) unsupervised learning and (iii) reinforcement learning.

An idea of the supervised learning is to find optimal parameters for a function that maps an input to an output based on a set of input-output pairs (training set). Input is usually represented by a vector of labelled parameters (features), while output is typically a value. Ideally, the model should infer some knowledge from the training set in order to perform reasonably well on previously unseen data (test set). Furthermore, supervised learning can be divided in classification and regression problems. A common example of the former is detection of spam or fraud e-mails. This often requires finding a complex non-linear decision boundary that would separate data points into groups (classes). Somewhat more advanced application is related to Computer Vision, e.g. automatic detection of objects on a picture. The main objective of a regression task is to predict a real-valued output. One example that is commonly used for teaching purposes is prediction of housing prices based on the data that is available about the house, such as size, year of construction, number of bedrooms etc. This is usually achieved by solving the following minimization objective

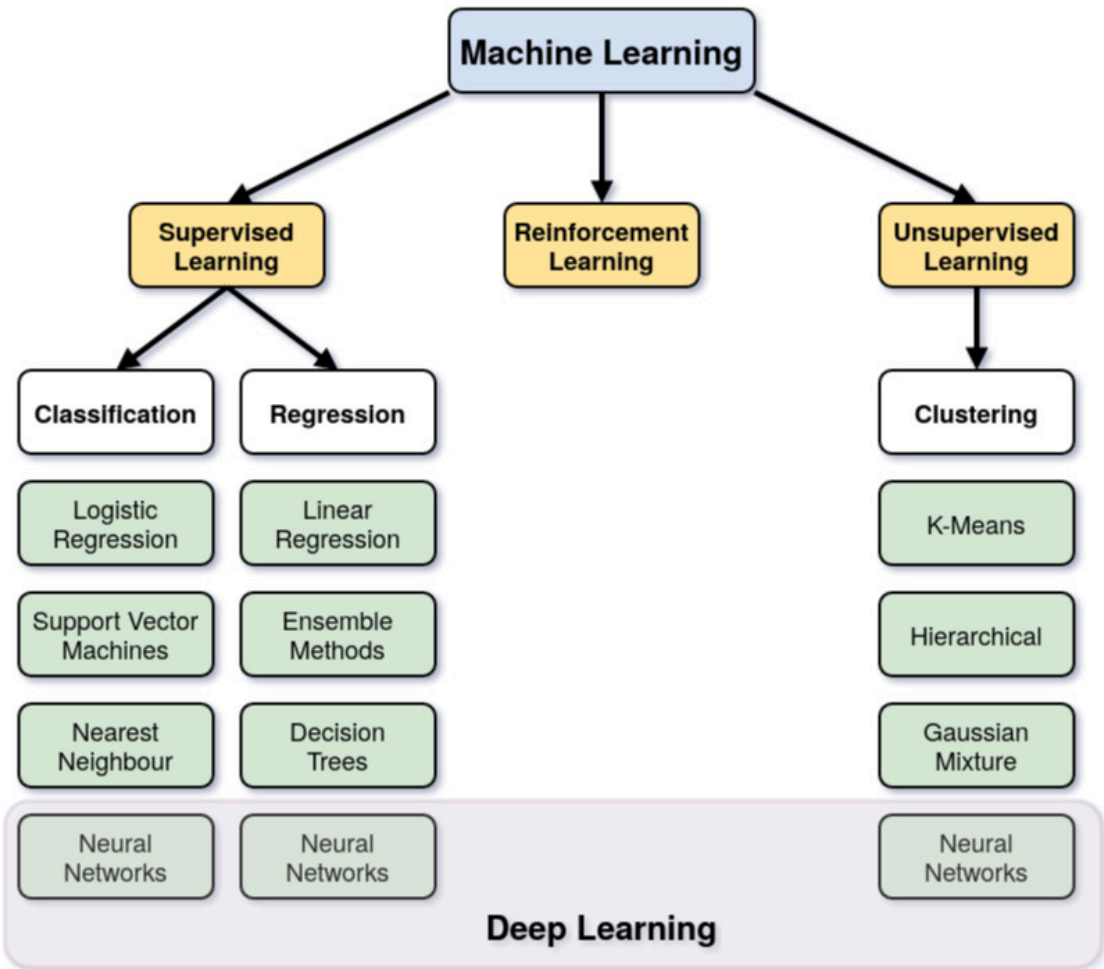


Figure 4.1: Schematic overview of some modern ML methodologies.

$$\min_{\theta_0,\theta_1,\ldots,\theta_n} J(\theta_0, \theta_1, \ldots, \theta_n), \tag{4.1}$$

where $\theta \in \mathbb{R}^{n+1}$ (n being the number of features x in the model) and $J(\theta)$ is the so-called cost function; θ_0 is called bias and it is introduced for practical purposes. Regression models may vary depending on the objective, thus rendering different model hypotheses $h_\theta(\mathbf{x})$. For instance, the linear regression hypothesis can be written as $h_\theta(\mathbf{x}) = \theta_0 + \sum_{i=1}^{n} \theta_i x_i = \theta^T \mathbf{x}$, while logistic regression (commonly used in classification tasks) hypothesis $h_\theta(\mathbf{x}) = g(\theta^T \mathbf{x})$ where $g(z) = 1/[1 + \exp(-z)]$ is the sigmoid function. An important algorithmic advantage of many ML schemes is that the cost function $J(\theta)$ is convex, i.e. it has a unique minimum. Thus, a set of parameters θ that corresponds to the minimization objective 4.1 can be found using one of the optimization schemes. This requires to iteratively update parameters θ based on some estimate of error between the model prediction and the provided data. One of the simplest optimization schemes is a gradient descent

$$\theta_i := \theta_i - \alpha \frac{\partial J(\theta)}{\partial \theta_i}, \tag{4.2}$$

where α is the learning rate. This parameter is very important to balance the computational efficiency of the ML algortihm. If α is too large, the gradient descent might diverge driving θ further and further away from the global minimum of $J(\theta)$. On the other hand, one would like to avoid α being too small since it slows down the optimization. More advanced optimization schemes can be used, for instance conjugated gradient or Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithms.[195]

The main objective of the unsupervised learning is to look for previously undetected patterns in a dataset. Cluster analysis or clustering (e.g. K-means) is a branch of the unsupervised learning that groups the data that has not been labelled, classified or categorized. It identifies common features in the dataset and responds based on the presence or absence of such features in each new data entry. This approach also helps to detect anomalous data points (outliers) that do not fit into either group.

Finally, reinforcement learning is a branch of the ML that studies how software agents ought to take actions in an environment in order to maximize the cumulative reward (see Figure 4.2). The major focus is given to finding a balance between exploration (of uncharted areas) and exploitation (of current knowledge).[196] A good example of reinforcement learning is AlphaGo[197] – a computer program that learnt to play the board game Go and won several best players in the world. More recently, AlphaZero[198] (AlphaGo's successor) managed to beat the state-of-the-art computer programs in chess, shogi and Go.

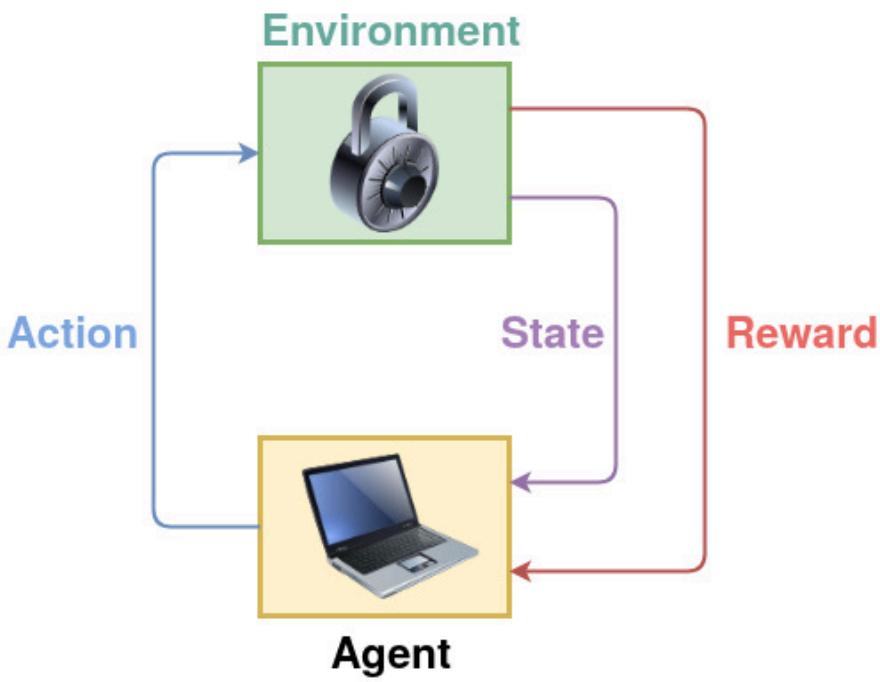


Figure 4.2: Schematic representation of the reinforcement learning approach.

There are some basic ML tricks (e.g. feature normalization, selection of initial parameters, shuffling of datasets) that I do not discuss here. For more practical details, see Ref. [199].

4.2 Neural Networks

...
How deep is your love?
Is it like the ocean?
...

From a song of Calvin Harris & Disciples

As can be seen in Figure 4.1, artifical Neural Networks (NNs) are universal and can be applied to any ML problem, thus forming a Deep Learning domain of the ML.[200,201] NNs have been introduced in 1943 to model and understand the signal processing in the brain.[202] In recent years, they have been actively studied and used in various fields of science.[203] There are many types of NNs with different functional forms. In particular, multilayer feed-forward (MLFF) neural networks (see Figure 4.3) have been used for the construction of complex functions or decision boundaries due to their ability to represent non-linear dependencies. It has been proven independently by a number of researchers that MLFF networks are "universal approximators",[204–207] i.e. they can approximate an unknown multidimensional function with an arbitrary accuracy based on a set of known function values. Alternatively, convolutional neural networks have been actively used for image and video analysis, recommender systems and natural language processing. This section is dedicated to NNs and their application to supervised learning.

It is worth mentioning, that the cost function of a NN is, in principal, non-convex. However, Deep Learning specialists claim that this is not an issue for practical applications. The more hidden layers and nodes are included, the more flexible is the functional form of the network. The entity of all layers including the input and output layers as well as the number of nodes per layer defines the architecture or topology of the NN. As indicated by the arrows in Figure 4.3, each node in each layer is connected to the nodes in the adjacent layers by a weight parameter. This can be be algorithmically rewritten using the symbol Θ^j_{ik} for the weight connecting node k in layer j with node i in the next layer. The input layer has the superscript 0 and the arrows label the direction of the flow of information through the network. Furthermore, there is a bias node connected to all nodes in the hidden layers and to the output node providing +1, which can be scaled by a bias weight b^j_i, where j is the layer of the target node and i is its number within this layer. The value y^j_i of any node i in any hidden layer j can then be calculated from the values of the N_{j-1} nodes in the previous layer and the weight parameters according to

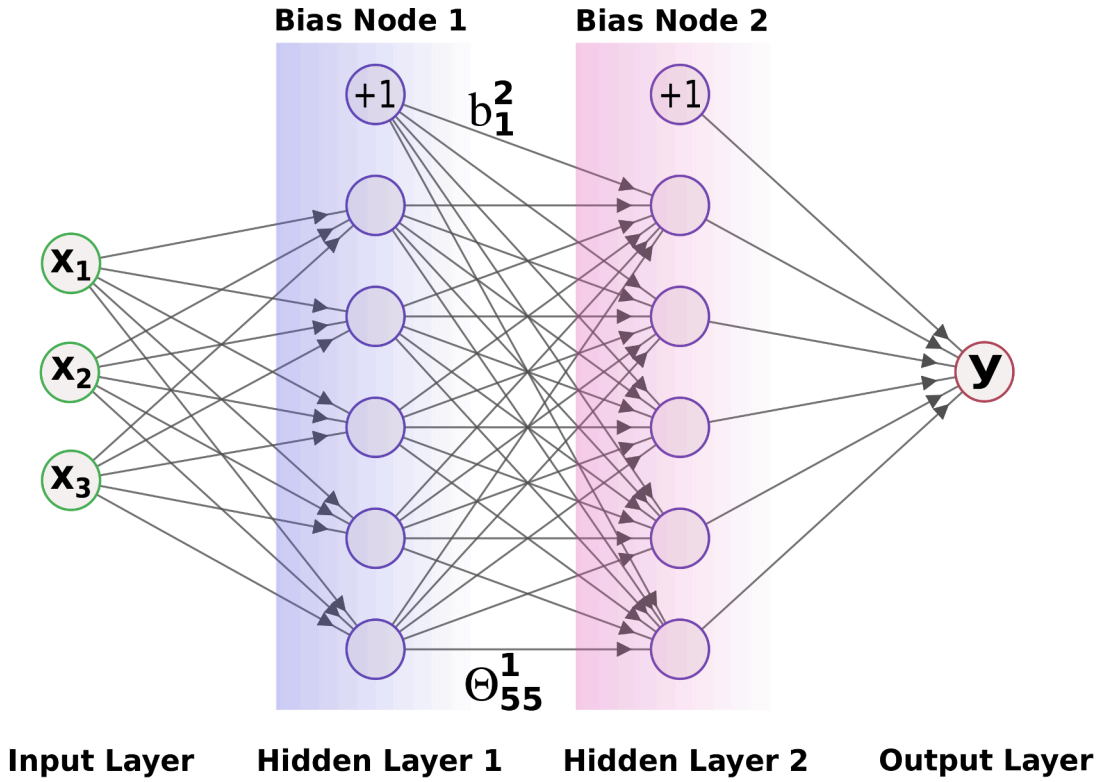


Figure 4.3: Schematic representation of a simple feed-forward NN with two hidden layers. The nodes (circles) are arranged in layers and the major goal is to establish a functional relation between a vector $\mathbf{x}$ of input parameters (features) and an output value y.

$$y_i^j = g_i^j \left(b_i^j + \sum_{k=1}^{N_{j-1}} \Theta_{ik}^{j-1} y_k^{j-1} \right), \tag{4.3}$$

where g_i^j is a non-linear function, which is usually called an activation function. It grants a capability to fit an arbitrary function since an output vector of layer j becomes an input (feature) vector for layer $j+1$. Thus, it recursively generates more and more complex features as propagation goes further (deeper) in the NN architecture. This helps to avoid issues that arise from manually chosen or crafted features. On the other hand, it makes the final output of a NN harder to interpret.

The choice of the activation function is non-trivial. In general, it should be non-linear and converge to –1, 0 or 1 in the limit of very large and very small arguments. Additional properties can be introduced based on a particular application. For example, NNs are commonly used in Quantum Chemistry and Materials Science to predict the total energy of a compound. Thus, activation functions must be differentiable with respect to the weight parameters in order to determine analytic forces needed for the gradient-based optimization of the weights. If forces are included in the NN training (e.g. to obtain more accurate PES or to use them for molecular dynamics), then activation functions should be at least twice differentiable. Common choises of activation functions include, but not limited to, the hyperbolic tangent, the sigmoid function

and Gaussian functions.

Once the information is propagated from the input to the output layer, the NN predictions can be derived. They are compared to the reference (training) dataset in order to estimate the cost function $J(\Theta)$. Similarly to the optimization objective 4.1, one can further compute partial derivatives $\frac{\partial J(\Theta)}{\partial \Theta^j_{ik}}$ of the cost function with respect to weight parameters Θ^j_{ik}. These quantities are required to perform the so-called back-propagation step in order to properly update the weights of the NN according to the minimization objective. Once the errors are computed for each node and the weights of the NN are updated, one arrives to the input layer and can perform the forward propagation again.

Even though supervised ML is a very powerfull tool, there are some practical issues that have to be kept in mind. One of them is the so-called overfitting, i.e. when the ML model performs well on the data from the training set but deviates from the remaining (validation or test) data. On the other hand, underfitting is another common issue associated with a low accuracy of the hypothesis on available training points. This usually happens when there is not enough features to describe the dataset. In terms of a NN architecture, using just a few (or too many) hidden layers as well as nodes per layer may lead to underfitting (overfitting).

Underfitting (overfitting) is commonly associated with a high bias (high variance). One way to balance them is *via* an additional regularization parameter λ that can be introduced in the model hypothesis. Regularization parameter belongs to a class of model hyperparameters, which means that its value is set before the learning process begins. For example, the learning rate and the NN topology are also hyperparameters. Several ways to avoid underfitting and overfitting in the ML applications are summarized in Table 4.1.

Table 4.1: Possible solutions to underfitting (high bias) and overfitting (high variance).

High bias	**High variance**
Try getting additional features	Get more training examples
Try adding more complex (e.g. polynomial) features	Try smaller sets of features
Try decreasing λ (if used)	Try increasing λ (if used)

Hyperparameter tuning is a challenging task for the ML applications that requires an additional fraction of the dataset to be reserved for this particular purpose. This fraction is usually called validation set since it is used for a cross-validation or model selection task. So far, only the training set (used to train the model) and the test set (used to estimate the model accuracy) have been discussed. Validation set is crucial to select optimal values of hyperparameters in order to keep the balance between an accuracy and a computational efficiency of the model. This can be done *via* a grid-like search where the model accuracy for a new set of hyperparameters is evaluated based on some metric such as a mean absolute error (MAE) or a root mean squared error (RMSE). The possible ratio between training, validation and test sets can be 60%/20%/20%, respectively. However, it varies depending on both the model (for instance,

models with a few hyperparameters require less validation data) and the training objective. In recent years, K-fold cross-validation (where K is an integer number) gained some popularity due to its relative algorithmic simplicity and remarkable efficiency. First, one has to split the total dataset into only training and test parts. Afterwards, the test set is kept aside and the remaining data is used to generate K splits of the training and validation sets. This drastically improves the performance of a ML model and allows to avoid overfitting. An example of a supervised ML workflow with the 3-fold cross-validation is illustrated in Figure 4.4.

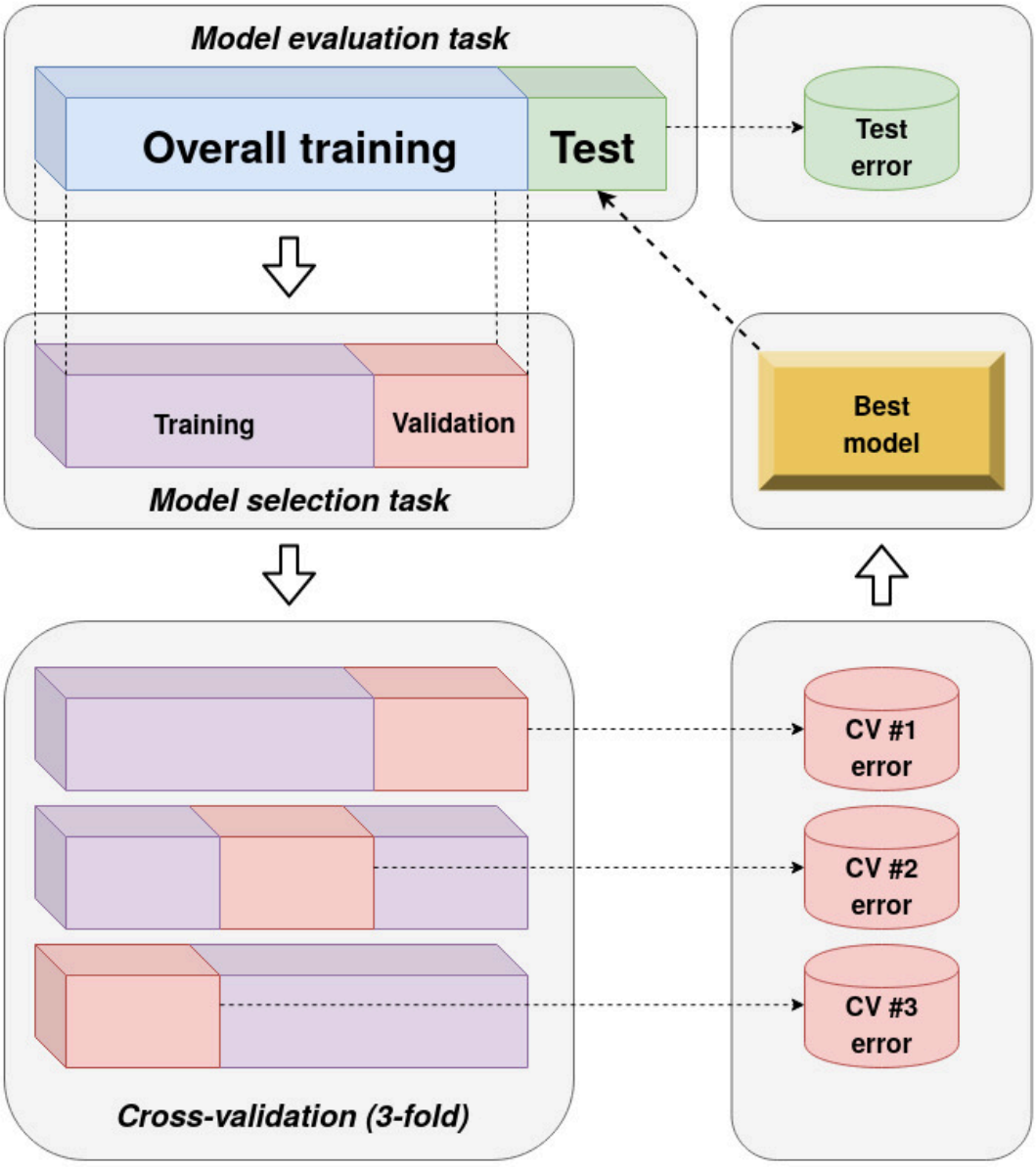


Figure 4.4: Schematic workflow of a supervised ML with the 3-fold cross-validation (CV).

Another way to monitor the validity of a ML model is to analyze the learning curves. They show how the cost function (error) depends on hyperparameters (training set size). Several learning curves as well as associated scenarios (overfitting/underfitting) are sketched in Figures 4.5 and 4.6.

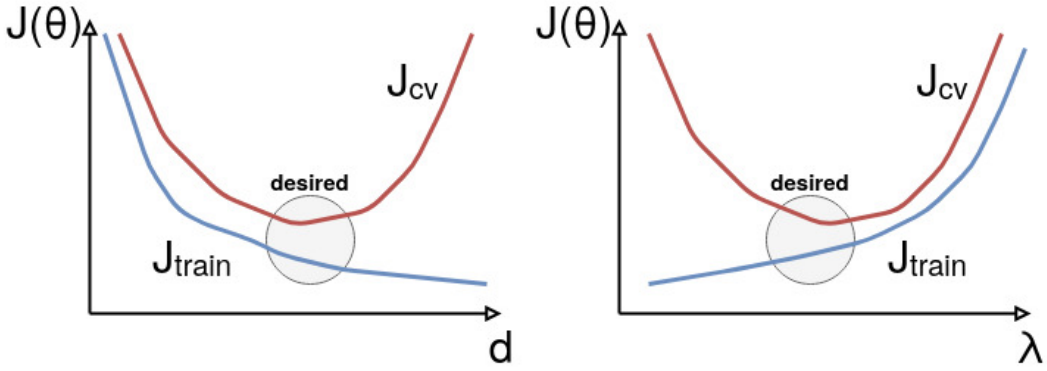


Figure 4.5: Schematic illustration of the learning curves as a function of model hyperparameters in case of a linear regression with polynomial features (left panel, d is the polynomial degree) or with the regularization parameter λ (right panel).

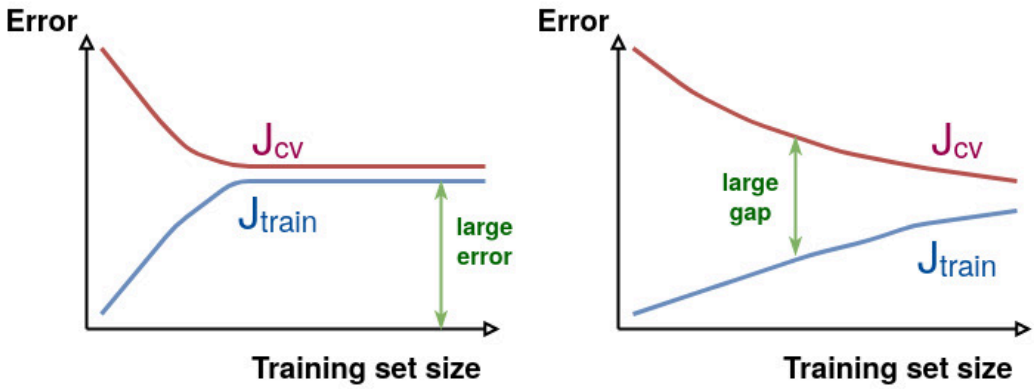


Figure 4.6: Schematic illustration of the learning curves as a function of the training set size in case of the high bias (left panel) or high variance (right panel).

Another class of supervised ML methods is based on kernels,[208] for instance non-linear support vector machines, kernel ridge regression or Gaussian process regression. Compared to NNs, kernel regression methods often use the input-to-output relationships directly. For example, given n output examples $\mathbf{y}$, Gaussian process regression provides a prediction model in the form of a linear combination of the given outputs y_i. Thus, the numerical complexity of accurate kernel-based ML models scales approximately from $O(n^2)$ to $O(n^3)$, thus becoming more demanding compared to the NN training for large values of n. On the other hand, kernel methods usually require less input-output pairs to make accurate predictions.

This chapter is dedicated to the application of NNs to excited-state dynamics. Recent studies have reported the non-adiabatic molecular dynamics using ML methods.[81,82,209–211] However, some of them require an *a priori* knowledge about the location of conical intersections in order to include more training points in this region[81,82], while others attempt to include the non-adiabatic couplings in the ML model[209–211], which is another challenging and computationally

demanding task. The major goal of the present study is to (i) apply NNs to obtain accurate machine-learned energies and forces for adiabatic excited states based on the TD-DFTB data; (ii) perform molecular dynamics and incorporate the non-adiabatic effects *via* the trajectory surface hopping (TSH) coupled to the Belyaev-Lebedev or Zhu-Nakamura approaches without any *a priori* knowledge about conical intersections or non-adiabatic couplings; (iii) evaluate the performance of these simplified schemes *versus* the Tully's fewest-switches surface hopping (FSSH) results and available experimental data.

4.3 Related work

In this section, I would like to present an overview of some ML applications to Quantum Chemistry and Material Science. This list is far from being complete since ML gained a lot of attention in recent years and new studies are published on a regular basis.

First, supervised ML models can be used to construct complex interatomic potentials. In 2007, Behler and Parinello[212] proposed a generalized NN representation of high-dimensional PESs that has been later on coupled to the metadynamics scheme.[213] A major advantage of this approach is a set of new molecular descriptors called symmetry functions. They are specifically designed to encode local atomic environments, i.e. two structures with different configurations yield different sets of symmetry function values, while identical local environments give rise to the same set. Furthermore, the symmetry function values must be rotationally and translationally invariant. For example, atomic Cartesian coordinates do not satisfy these requirements and cannot be used as input features. The total energy E of the system is computed as a sum of atomic contributions E_i. This model, also called the Behler-Parinello NN, has been used to compute the IR spectra from molecular dynamics simulations on a machine-learned PES[214] and to simulate the reactive gas-surface dynamics[215,216]. More details can be found in a tutorial review from Ref. [217].

One of the most straightforward, yet remarkably successful, applications of machine-learned PES is to perform extensive MD simulations. Ideally, this would correspond to an *ab initio* accuracy (depending on the reference method that was used to train the NN) at the computational cost of a force field. Some ML models include forces in the training process, which drastically improves the quality of a machine-learned PES.[218–221] This comes from additional $3M$ points (M is the number of atoms) that provide more information about the shape of a multidimensional surface. On the other hand, NNs have been also used to correct the DFT[222] and DFTB[223] results based on a Δ-Machine Learning[224] approach. Several Python-based software packages are freely available for the supervised ML applications. For example, see MLatom[225], DeePMD-kit[226] and SchNetPack[227]. More recently, the Deep Learning architecture SchNet[221] has been interfaced with SHARC[194] code for non-adiabatic molecular dynamics.[210] Alternatively, direct quantum dynamics can be performed using the

MCTDH approach on a machine-learned PES.[228]

On the other hand, a PES constructed by the ML can be used to accelerate the global optimization. In particular, the group of Prof. Bjørk Hammer from Aarhus University has succeeded to combine different types of the ML schemes with genetic (evolutionary) algortihms. The synergy of these approaches allows to perform simulations in an efficient data-driven way that keeps the balance between exploration and exploitation, i.e. the information from structures that have already been seen is used to make a next prediction (mutation). For example, see a model based on Gaussian processes[229], a GOFEE scheme[230] based on Bayesian statistics or a LEA algorithm[231] based on NNs. The same group has also applied the unsupervised learning[232,233] and, more recently, the reinforcement learning[234] techniques to improve the global optimization performance.

Applications of kernel-based methods to the supervised ML in quantum dynamics and materials science are discussed in Ref. [235] and Ref. [236], respectively. A good perspective of the unsupervised learning in atomistic simulations can be found in Ref. [237].

4.4 Computational details

As mentioned in Section 4.1, the current procedure consists of three consecutive steps: (i) train a NN model for energies and forces in excited states, (ii) perform the TSH simulations using results from the previous step and (iii) compare them with the FSSH and experimental data. Dynamics in low-lying excited states of phenanthrene ($C_{14}H_{10}$ – an armchair-shaped isomer of anthracene, see Figure 4.7) has been chosen as a test case due to its relatively large molecular size, due to diversity of available experimental results and due to some ambiguity[94] in their interpretation. Furthermore, phenanthrene is among the most abundant species observed in meteorites.[238]

For the first step, I rely on a recently developed Deep Learning architecture called SchNet[221], which is implemented in an open-source Python package SchNetPack[227]. There are several reasons why this particular implementation has been chosen. First of all, SchNet is a convolutional NN that automatically generates filters that map one hidden layer to another one based on pairwise interactions between a chosen atom and the surrounding ones. It has been shown to yield accurate molecular representations and energies, especially when forces are included in the training.[221] 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.[227] A shifted softplus is used as an activation function:

$$\mathrm{ssp}(x) = \ln\left(\frac{\mathrm{e}^x + 1}{2}\right), \tag{4.4}$$

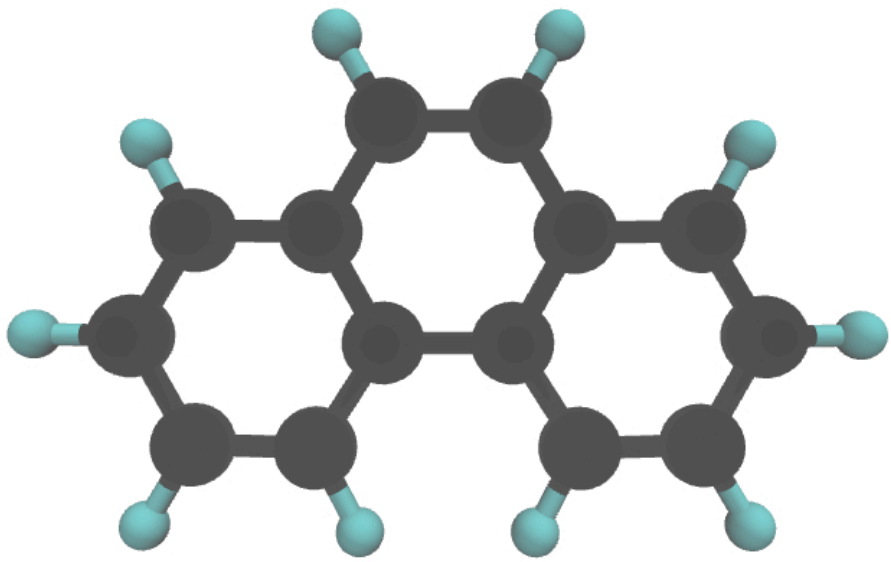


Figure 4.7: Balls and sticks representation of phenanthrene ($C_{14}H_{10}$) molecule. Black and cyan spheres are carbon and hydrogen atoms, respectively.

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

$$J\left([E, \mathbf{F}_1, \ldots, \mathbf{F}_M], [\widetilde{E}, \widetilde{\mathbf{F}}_1, \ldots, \widetilde{\mathbf{F}}_M]\right) = \rho||E - \widetilde{E}||^2 + \frac{1}{M}\sum_{A=1}^{M}\left|\left|\mathbf{F}_A - \left(-\frac{\partial \widetilde{E}}{\partial \mathbf{R}_A}\right)\right|\right|^2, \quad (4.5)$$

if both energies E and atomic forces $\mathbf{F}$ are available. In the equation above, ρ is the trade-off between the energy and force loss[239], M is the number of atoms, quantities with tilde denote the model (SchNet) predictions and plain symbols denote reference data. The second term in Eqn. 4.5 can be neglected if only energies are available in the dataset. All models have been trained with a mini-batch stochastic gradient descent (MBGD) and $\rho = 1$. The learning rate α is initially set to 10^{-4} and can decay exponentially with ratio 0.8 down to 10^{-6}. The validation set is used for an early stopping, which is a type of regularization that prevents overfitting.[217] More details about the SchNet architecture and practical aspects can be found in Ref. [240].

Furthermore, SchNetPack is natively interfaced with the Atomic Simulation Environment[241] (ASE) – an advanced open-source Python toolkit for atomistic simulations. I have implemented an interface between the ASE and the deMon-Nano code during a short stay in the group of Prof. Bjørk Hammer at Aarhus University. This interface (see Appendix D) has been included in the official stable release 3.19.0 of the ASE. Finally, the SchNet implementation is based on

a PyTorch[242] library for Deep Learning applications that can be efficiently executed on GPUs (highly recommended for NN training). At CALMIP High-Performance Computing center in Toulouse, up to 4 GPUs (NVIDIA®V100, Volta architecture) available per node together with 36 CPUs (2 processors with 18 cores each, Skylake®Intel architecture) and 384 gigabytes of RAM. Unless stated otherwise, a single GPU and 4 CPU cores have been used for the SchNet training.

Three TD-DFTB datasets have been generated independently for S_2, S_3 and S_4 adiabatic excited states of phenanthrene. These particular states have been chosen for the reasons described in the next Section. The DFTB parameters were taken from the mio-set.[121] The ability to reproduce energies and geometries corresponding to low-lying excited states of organic molecules has been critically assessed for differet sets of DFTB parameters in Refs. [243, 244]. Each generated dataset contains 10000 points that have been sampled from a single NVT trajectory equilibrated at $T = 500$ K during 50 ps using a chain of 5 Nosé-Hoover thermostats and 0.5 fs time step. 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. If needed, forces can be included in the dataset to enhance the training performance. Afterwards, the dataset is provided to SchNetPack with additional parameters that specify the NN topology, i.e. the number of hidden layers, the number of features per layer, the batch size (one of the hyperparameters of the MBGD scheme that determines how many training examples are fed into the NN to update the weights since feeding the entire dataset is not feasible) and the number of epochs (another hyperparameter of the MBGD that tells how many times the NN has to walk through the entire dataset). Additionally, one may activate the GPUs and/or include forces in the training. Finally, some modifications can be introduced in the derivation of the desired (output) quantity from the atomic contributions. The default option is a summation (e.g. the total energy of a molecule is obtained as a sum of atomic contributions provided by the NN). Some sets of parameters used in the training and the corresponding execution times are summarized in Table 4.2. The total number of model parameters is 455809 for model #1, 238465 for model #2 and 886529 for model #3. This indicates the importance of using GPUs for learning PES of large molecular complexes.

Table 4.2: Some examples of the SchNet model parameters and the corresponding execution times. These calculations have used forces and GPUs in the training.

		Model parameters			Dataset sizes			
#	State	Features	Hidden layers	Batch size	Training set	Validation set	Test set	Runtime, h
1	S_2	128	6	100	6000	2000	2000	14
2	S_2	128	3	100	6000	2000	2000	15
3	S_3	256	3	50	6000	2000	2000	15

The second step requires to incorporate the non-adiabatic effects in the TSH simulations with the SchNet-learned PESs. Two hopping probabilities (see below) have been integrated in the molecular dynamics driver of the ASE: (i) P^{BL} based on the Belyaev-Lebedev (BL) approach [76] adapted for adiabatic states and (ii) P^{ZN} based on the multidimensional extension of the Zhu-Nakamura (ZN) theory following Hanasaki *et al.* [80]

$$P_{IJ}^{\mathrm{BL}} = \exp\left(-\frac{\pi}{2}\sqrt{\frac{Z_{IJ}^3}{\ddot{Z}_{IJ}}}\right); \tag{4.6}$$

$$P_{IJ}^{\mathrm{ZN}} = \exp\left(-\frac{\pi}{4|a|}\sqrt{\frac{2}{b^2+\sqrt{|b^4\pm 1|}}}\right), \tag{4.7}$$

where a and b are the effective coupling and effective collision energy, respectively; $Z_{IJ} = |E_I - E_J|$ 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 equations above are applicable only at the diabatic curve-crossing point. The double sign in Eqn. 4.7 takes plus (minus) if diabatic forces have the same (opposite) sign at the crossing point. Parameters a and b are defined as

$$a^2 = \frac{1}{2\mu}\frac{\sqrt{|F_1F_2|}|F_2-F_1|}{(2V_{12})^3}; \tag{4.8}$$

$$b^2 = (E_T - E_X)\frac{|F_2-F_1|}{2V_{12}\sqrt{|F_1F_2|}}, \tag{4.9}$$

where μ is the reduced mass of the nuclei, F_1 and F_2 are diabatic forces (one-dimensional), V_{12} is the diabatic coupling, E_T is the total energy and E_X is the potential energy at the crossing point. A detailed derivation of diabatic gradients from multidimensional adiabatic ones can be found in Ref. [80].

In practical applications of the Landau-Zener approach (BL formulation), $\ddot{Z}_{IJ}$ at the crossing point t_c is usually computed using finite differences as follows

$$\ddot{Z}_{IJ}|_{t=t_c} = \frac{Z_{IJ}(t_c+\Delta t) - 2Z_{IJ}(t_c) + Z_{IJ}(t_c-\Delta t)}{(\Delta t)^2}. \tag{4.10}$$

Similarly to the FSSH scheme, each classical trajectory is propagated on a given excited PES with a fixed time step Δt (using the Velocity Verlet scheme) and hopping probabilities P_{IJ}^{BL} or P_{IJ}^{ZN} computed at each local minimum of the adiabatic energy gap Z_{IJ}. A switch from an active state I to state J is performed if P_{IJ}^{BL} or P_{IJ}^{ZN} is larger than a random number $0 < \xi < 1$ and the propagation continues on a different PES. However, the switch can be still rejected if there is not enough kinetic energy to compensate the difference in the potential energies (see description of frustrated or classically forbidden hops in Section 2.3). In order

to be consistent with the FSSH results, the same velocity rescaling procedure is used in the simplified (BL or ZN) TSH schemes. Alternatively, one may perform a more accurate velocity rescaling[79,162] based on computationally cheap gradients provided by the ML models. Due to the fact that both BL- and ZN-based schemes require data from three consecutive MD 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. A major advantage of conventional TSH approaches is that there is no time-dependent equation for electrons (like Eqn. 2.64 in FSSH) to propagate. Thus, BL- or ZN-based schemes naturally account for decoherence effects.[245] On the other hand, it also means that the electronic population $|C_I|^2$ of excited state S_I cannot be computed directly. However, it 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. Initial conditions for the TSH have been sampled from the thermal distribution of the ground state at $T = 300$ K following the procedure described for FSSH in Chapter 3.

Finally, coupled FSSH/TD-DFTB simulations have been performed using the mio-set (BIO) of DFTB parameters. In order to be consistent with the ML-based results, the same initial conditions have been used. Each classical trajectory has been propagated with $\Delta t = 0.25$ fs during 1 ps and the conical intersection threshold set to 10 meV for excited singlet states.

Experimental results for phenanthrene have been extracted from Refs. [96, 238, 246–248]. I am particularly interested in the full width at half maximum (FWHM) values from early studies performed in supersonic jets[238,246,247] and the relaxation times from pump-probe experiments of Blanchet *et al.*[96] The approximate decay time τ can be derived from the FWHM as follows[94]

$$\tau = \frac{1}{2\pi c \Gamma}, \tag{4.11}$$

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

4.5 Results

This section is dedicated to the results of several TSH models applied to the electronic relaxation in phenanthrene. It is organized as follows: (i) TD-DFT and TD-DFTB absorption spectra presented together with some CASPT2 and experimental results; (ii) the SchNet model applied to learn energies and forces based on the TD-DFTB reference datasets; (iii) results of the FSSH/TD-DFTB simulations analysed and compared with experimental findings; and (iv) the TSH simulations based on the Belyaev-Lebedev or Zhu-Nakamura approaches compared to the FSSH/TD-DFTB in order to assess the performance of non-adiabatic dynamics on the machine-learned PESs with no *a priori* knowledge about couplings and conical intersections.

4.5.1 Absorption spectra of phenanthrene

First, some experimental studies of UV spectra of phenanthrene and their conclusions are summarized. Positions of several low-lying singlet states of neutral phenanthrene together with the corresponding FWHM and decay times (where available) can be found in Table 4.3.

Table 4.3: Selected results from the experiments (in a chronological order) with neutral phenanthrene in different environments.

Source (year)	Environment	Attributed transition	Position, nm	FWHM, cm^{-1}	Decay time, ps
Amirav *et al.*[246] (1984)	supersonic jet	$S_0 \rightarrow S_1$	340.9	5.8	0.92
		$S_0 \rightarrow S_2$	282.6	11.1	0.48 (≈0.5)
Ohta and Baba[247] (1986)	supersonic jet	$S_0 \rightarrow S_1$	340.9	—	—
		$S_0 \rightarrow S_2$	282.6	12.6	0.42
Dick and Nickel[248] (1986)	isopentane glass at 4 K	$S_0 \rightarrow S_2$	—	67±7	0.08 (79±8 fs)
Brechignac and Hermine[238] (1994)	supersonic jet	$S_0 \rightarrow S_2$	—	12.0	0.44
Salama *et al.*[249] (1994)	Ne matrix at 4 K	$S_0 \rightarrow S_1$	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	—	—
Blanchet *et al.*[96] (2000)	gas phase	$S_0 \rightarrow S_1$	341.6	—	—
		$S_0 \rightarrow S_2$	283.7	—	0.52 (522±16 fs)
Huisken *et al.*[250] (2011)	supersonic jet	$S_0 \rightarrow S_1$	340.6	—	—
		$S_0 \rightarrow S_2$	282.4	—	—

Furthermore, absorption spectra of neutral phenanthrene have been computed at the TD-DFTB (mio-set) and the TD-DFT (BLYP functional with 6-31G(d,p) basis set) levels of theory. These results, together with selected experimental values, are compiled in Figure 4.8.

The agreement between the TD-DFTB and the TD-DFT spectra is reasonably good and both methods predict positions of two experimentally observed bright states around 284 nm and 266 nm with remarkable accuracy. It is worth mentioning that the band at 284 nm (4.37 eV) corresponds to $S_0 \rightarrow S_3$ transition in both TD-DFT and TD-DFTB, while available experimental data attribute it to $S_0 \rightarrow S_2$ 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,

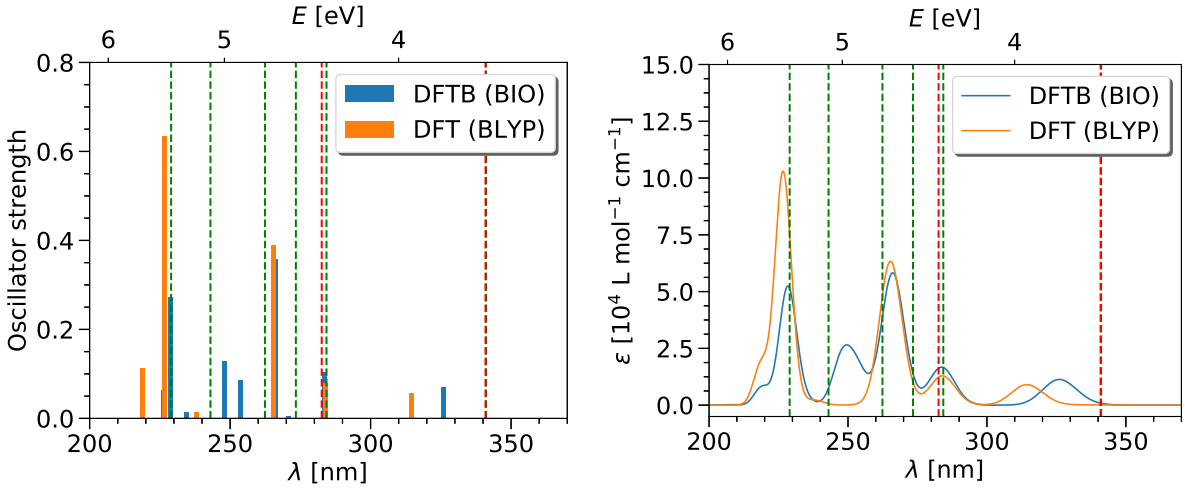


Figure 4.8: Absorption spectra (sticks in the left panel, convoluted in the right panel) of neutral phenanthrene computed with TD-DFT (orange) and TD-DFTB (blue) at the equilibrium geometry. All spectra have been truncated to 6.0 eV. Green and red vertical dashed lines indicate positions of some excited states from the experiments in Ne matrix[249] and in the supersonic jet[246,247], respectively (see Table 4.3).

see Table 4.4. Furthermore, the ratio κ of TD-DFTB oscillator strengths for S_3/S_2 and S_3/S_1 is 1.5 and 79.4, respectively. The latter value is in very good agreement with $\kappa = 84$ reported by Amirav *et al.*[246] 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.*[251] 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 eV and 4.26 eV, respectively. Notably, the brightest CASPT2 state corresponds to an excitation at 4.81 eV, which is rather close to the TD-DFTB one at 4.66 eV, but far from 5.47 eV computed at the TD-DFT level. On the other hand, since there are two bright states in TD-DFT with relatively close oscillator strengths (see Figure 4.8), the latter issue might be solved if a different combination of functional/basis set is used. In the same study, CASSCF calculations have demonstrated a significant overestimation of vertical excitation energies computed for low-lying singlet-singlet transitions in phenanthrene.[251] More recently, Nazari *et al.*[252] have performed a complex analysis of PESs and ultrafast dynamics of monomeric phenanthrene and its derivatives. The supporting TD-DFT and CASPT2 absorption spectra are in reasonably good agreement with previous studies. However, the ordering of the states can be somewhat misleading due to the fact that higher-lying excited states (S_3 and above) have been discarded in Ref. [252]. For instance, Nazari *et al.*[252] compare the S_2 state at 4.14 eV with the S_3 state at 4.37 eV from Ref. [251] (both values computed at the CASPT2 level of theory with different parameters).

Table 4.4: More details about several low-lying excited states of neutral phenanthrene extracted from the TD-DFTB and the TD-DFT absorption spectra. H and L stand for HOMO and LUMO, respectively.

Method	Excited state	Position, eV	Oscillator strength	Dominant transitions between MOs	Weight
TD-DFTB	S_1	3.70	$\approx 10^{-3}$	H–1→L H→L+1	0.57 0.43
	S_2	3.80	0.07	H→L H–1→L+1	0.70 0.20
	S_3	4.37	0.10	H→L+1 H–1→L	0.46 0.31
	S_4	4.58	$\approx 5 \cdot 10^{-3}$	H–2→L H→L+2	0.68 0.30
TD-DFT	S_1	3.69	$\approx 10^{-3}$	H→L+1 H–1→L	0.54 0.45
	S_2	3.94	0.06	H→L H–1→L+1	0.62 0.30
	S_3	4.37	0.08	H–1→L H→L+1	0.52 0.42
	S_4	4.61	$\approx 10^{-3}$	H→L+2 H–2→L	0.50 0.49

4.5.2 Training SchNet

The performance of several SchNet models is summarized in Table 4.5. Notably, all of them achieve chemical accuracy (errors smaller than 1 kcal/mol or 0.043 eV) on the TD-DFTB energies of phenanthrene. This is not surprising, taking into account the outstanding accuracy of SchNet for prediction of both ground state[221] and excited state[210] properties. However, it is worth mentioning that the real accuracy of the machine-learned quantities is correlated with the underlying electronic structure method that has been used for training. For example, SchNet errors may exceed 1 kcal/mol for some wavefunction-based methods like CASSCF.[210] In addition, errors are larger for the SchNet model #8 (trained without forces) compared to model #5, which is consistent with previous studies.[221,253] 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. Model #4 was trained without GPU and forces, resulting in only 463 training epochs over 48 hours. Rough extrapolation suggests that it would take more than 100 hours to perform a fully-converged training with 1000 epochs that would be significantly longer than learning the PES with both forces and GPUs available. The error

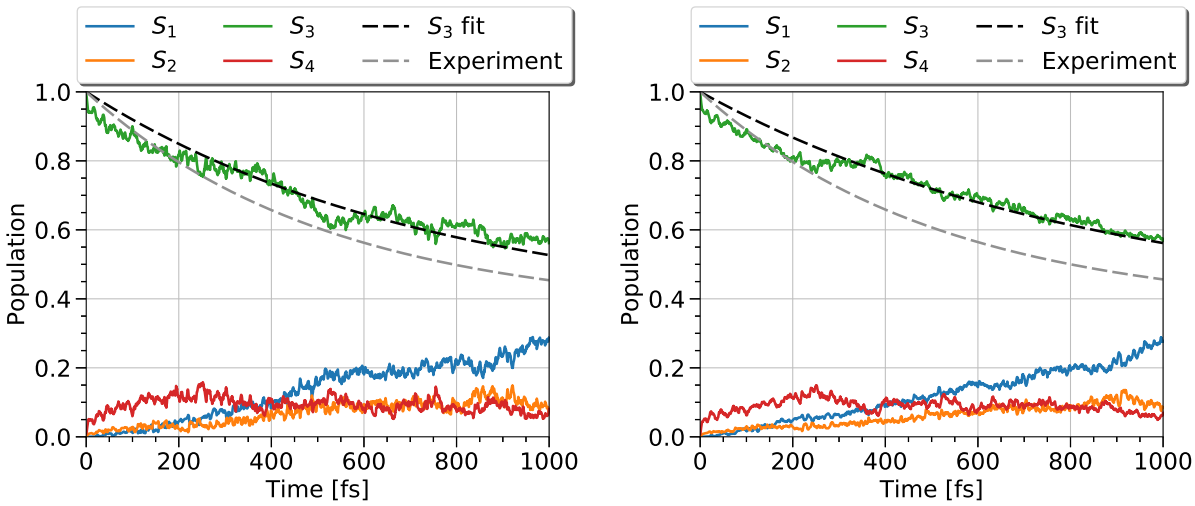
Table 4.5: Parameters of the considered SchNet models and the corresponding training times and errors. These models have used forces (except #4 and #8) and GPUs (except #4) in the training. The MAEs (RMSEs) for energies and forces are given in eV and eV/Å, respectively. The MAE and RMSE have been computed for a test set (2000 structures for all models except #3 and #7 where 6000 structures have been used).

		Model parameters		Dataset sizes				
#	State	Features	Hidden layers	Training set	Validation set	Runtime, h	Property	MAE (RMSE)
1	S_2	128	6	6000	2000	14	energy force	0.014 (0.018) 0.075 (0.126)
2	S_2	128	3	6000	2000	15	energy force	0.018 (0.023) 0.084 (0.138)
3	S_2	64	3	3000	1000	4	energy force	0.027 (0.035) 0.134 (0.213)
4	S_2	128	3	6000	2000	$\gtrsim$100 [no GPU]	energy force	— —
5	S_3	128	6	6000	2000	13	energy force	0.016 (0.022) 0.068 (0.113)
6	S_3	256	3	6000	2000	15	energy force	0.014 (0.018) 0.062 (0.106)
7	S_3	256	3	3000	1000	5	energy force	0.017 (0.022) 0.075 (0.124)
8	S_3	128	6	6000	2000	40	energy force	0.029 (0.038) —

does not vary significantly between models with 3 and 6 hidden layers (interactions), which has already been pointed out by Schütt *et al.*[221] for the QM9[254–256] dataset. Yet, the difference in accuracy of NNs with 1, 2 and 3 hidden layers can be more pronounced.[257] The number of features has a major effect as can be seen upon comparison of models #3 and #7. Using only 3000 training and 1000 validation points, model #7 with 256 features achieves similar accuracy as model #6 that has been trained with twice as much data. On the other hand, the performance of model #3 with 64 features is only comparable to that of model #8 that has not used forces at all. Overall, 128 features and 3 hidden layers 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 ML. In fact, thoroughly crafted NN can achieve remarkable accuracy even on relatively small training sets.[240]

4.5.3 FSSH/TD-DFTB simulations

Following the discussion in Subsection 4.5.1, the goal is to study the radiationless electronic relaxation from the TD-DFTB S_3 state of phenanthrene and to compare the computed decay times with experimental results from Table 4.3. An ensemble of 250 trajectories (see Figure 4.9) has been launched and propagated using the recently developed FSSH/TD-DFTB implementation in the deMon-Nano code.



100 trajectories 250 trajectories

Figure 4.9: Population analysis of the first 4 singlet excited states in phenanthrene for an ensemble of 100 (left panel) and 250 (right panel) trajectories propagated with BIO parameters. The grey dashed line corresponds to an exponential decay with $\tau = 522$ fs extracted from the experiments of Blanchet *et al.*[96] The initial state is S_3 in both cases.

First of all, the shapes of averaged population curves change slightly as the number of trajectories is increased from 100 to 250. This indicates that this ensemble may have not been fully converged yet, unlike it was for anthracene when going from 63 to 127 trajectories

(see Chapter 3). Furthermore, the computed decay times are 745 fs and 860 fs for 100 and 250 trajectories, respectively. Nevertheless, the S_3 population curve can be roughly divided in 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 towards upper-lying S_4 and lower-lying S_2 and S_1 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, which can be correlated with an increase of the S_1 population that becomes pronounced after 400 fs. The aforementioned results are in qualitative agreement with the complex multichannel relaxation, which is expected for phenanthrene according to Ref. [252]. As pointed out by Nazari *et al.*[252], 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 the FSSH/TD-DFT simulations in the two lowest excited singlet states of the phenanthrene derivative. The results are in qualitatively good agreement with experimental observations, despite neglecting the transitions towards higher-lying excited states. The major fraction (80%) of the population is rapidly (within less than 200 fs) transferred to the lowest excited state, in contrast with the FSSH/TD-DFTB (this work) simulations. Furthermore, the TD-DFT energy gaps (computed with ωB97XD and BHandHLYP functionals)[252] between the two lowest excited states are smaller than the ones in CASPT2[251,252] or TD-DFTB (this work) spectra. This explains faster and more complete (compared to the aforementioned FSSH/TD-DFTB results) relaxation reported by Nazari *et al.*[252] Overall, the FSSH/TD-DFTB approach applied to the electronic relaxation of neutral phenanthrene is in reasonably good agreement with experimental results.

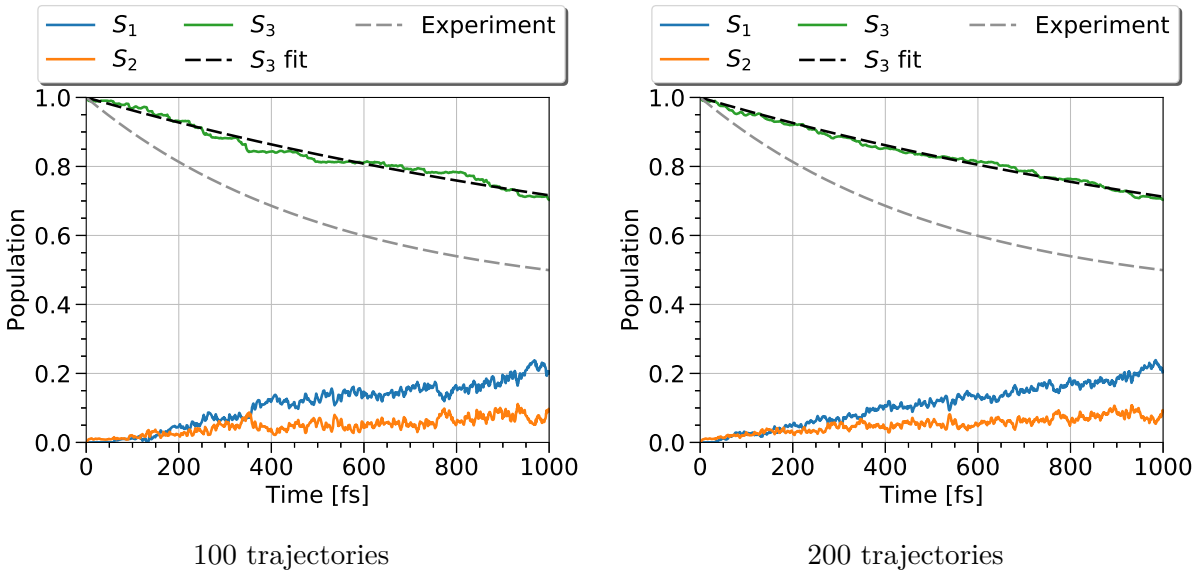


Figure 4.10: Population analysis of the first 3 singlet excited states in phenanthrene for an ensemble of 100 (left panel) and 200 (right panel) trajectories propagated with BIO parameters. The grey dashed line corresponds to an exponential decay with $\tau = 522$ fs extracted from the experiments of Blanchet *et al.*[96] The initial state is S_3 in both cases.

Also, downward transitions can be enforced by excluding all excited states that are energetically

higher than S_3 in the propagation. In theory, this could accelerate the population decay, but actually a significant delay (see Figure 4.10) has been observed. It takes more time to approach a funnel that drives the population transfer towards the lower-lying S_2 and S_1 states. This correlates with the relatively large gap between the initially excited S_3 and the neighbouring S_2 states in TD-DFTB (see Table 4.4). The computed decay times are 1515 fs and 1485 fs for 100 and 200 trajectories, respectively. Notably, the S_3 population curve looks rather linear when transitions to higher-lying states are neglected. These results also support the conclusion of Nazari *et al.*[252] that the photocycle of phenanthrene cannot be described *via* a simple two-state model.

It has been mentioned before that some *ad hoc* decoherence corrections like the simplified decay of mixing used in this work may induce a delay in the electronic relaxation, thus increasing the computed decay times.[245] Additional calculations have been performed with 100 trajectories propagated using $\Delta t = 0.25$ fs without the decoherence correction (see Figure 4.11). Indeed, the decay time is 149 fs, which is five times smaller than 745 fs computed with the simplified decay of mixing. Furthermore, there is a rapid transfer of the population to the upper-lying S_4 state, which carries away significant fraction of the population, thus accelerating the decay of the initally excited S_3. Finally, the computed relaxation without the decoherence correction is about three times faster than the experimentally observed ones.

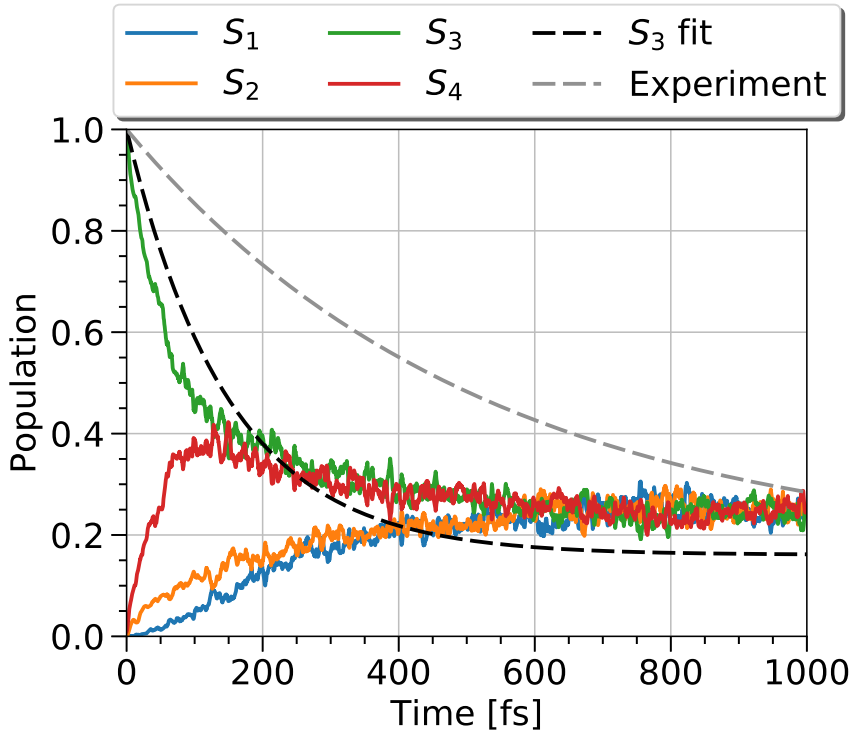
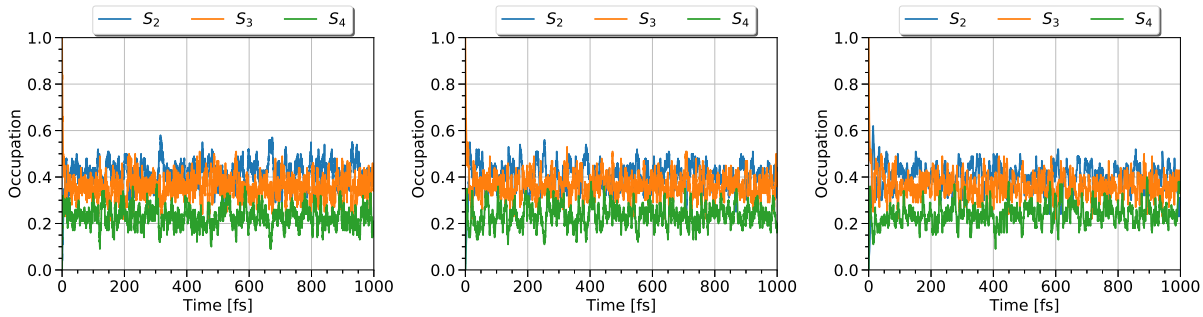


Figure 4.11: Population analysis of the first 4 singlet excited states in phenanthrene for an ensemble of 100 trajectories propagated without the decoherence correction. The grey dashed line corresponds to $\tau = 522$ fs extracted from the experiments of Blanchet *et al.*[96]

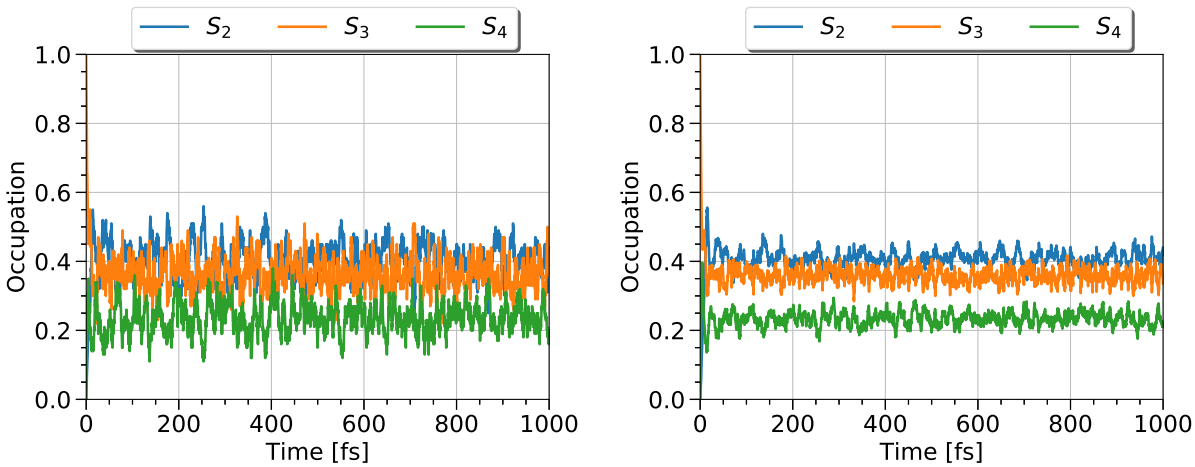
4.5.4 TSH/SchNet simulations

First of all, the TSH simulations with the ZN hopping approach and different propagation steps are analyzed (see Figures 4.12 and 4.13).



$\Delta t = 0.1$ fs $\Delta t = 0.25$ fs $\Delta t = 0.5$ fs

Figure 4.12: Occupations of the S_2, S_3 and S_4 excited states in phenanthrene averaged over an ensemble of 100 trajectories. Each trajectory has been propagated with an indicated time step Δt on machine-learned PESs with the Zhu-Nakamura hopping probability.



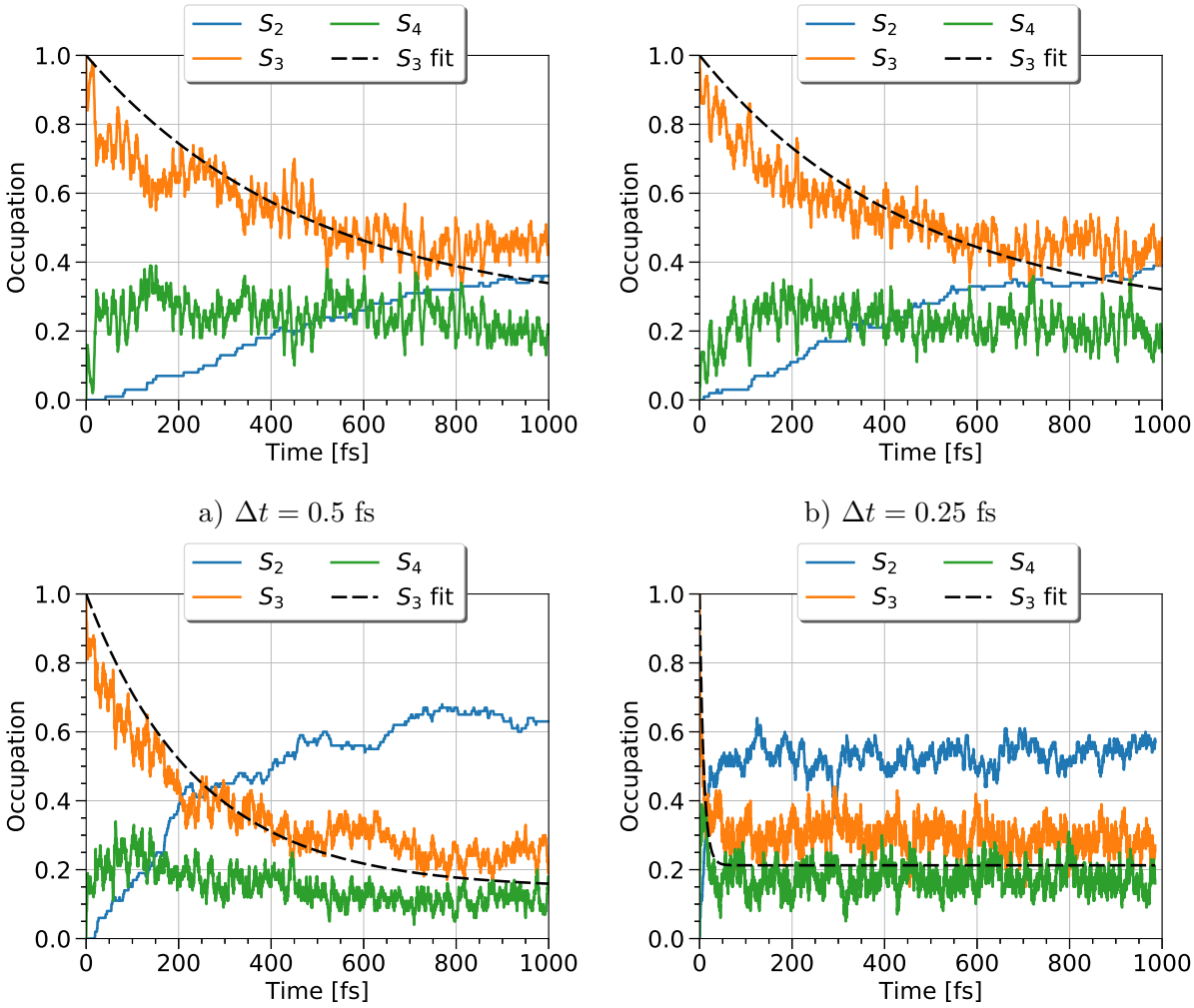
100 trajectories 500 trajectories

Figure 4.13: Occupations of the S_2, S_3 and S_4 excited states in phenanthrene averaged over an ensemble of 100 (left panel) and 500 (right panel) trajectories. Each trajectory has been propagated with the time step $\Delta t = 0.25$ fs on machine-learned PESs with the Zhu-Nakamura hopping probability.

It is clear that the ZN results are not significantly affected by changes in the time step. This is not surprising, taking into account the modified diabatic gradients proposed by Hanasaki *et al.*[80] The same group has pointed out that other multidimensional extensions of the ZN theory (like the one developed by Yu *et al.*[79]) may be affected by the propagation step due to the underlying assumptions about the topology of PESs. Overall, an ultrafast decay (within approximately 10 fs) is observed for the initially excited S_3 state. However, the electronic relaxation is not complete due to the residual oscillations in the occupation of each excited

state. Whether this is a consequence of the present implementation or a drawback of the multidimensional extension of the ZN approach is a question for future work. On the other hand, 100 trajectories is a relatively small ensemble size. Owing to the low computational cost of the ML predictions, one can further increase the number of trajectories (see Figure 4.13). The results are qualitatively similar despite the reduced magnitude of oscillations.

Similarly to the ZN case, the BL-based simulations performed with different time steps are analyzed (see Figure 4.14).



a) $\Delta t = 0.5$ fs

b) $\Delta t = 0.25$ fs

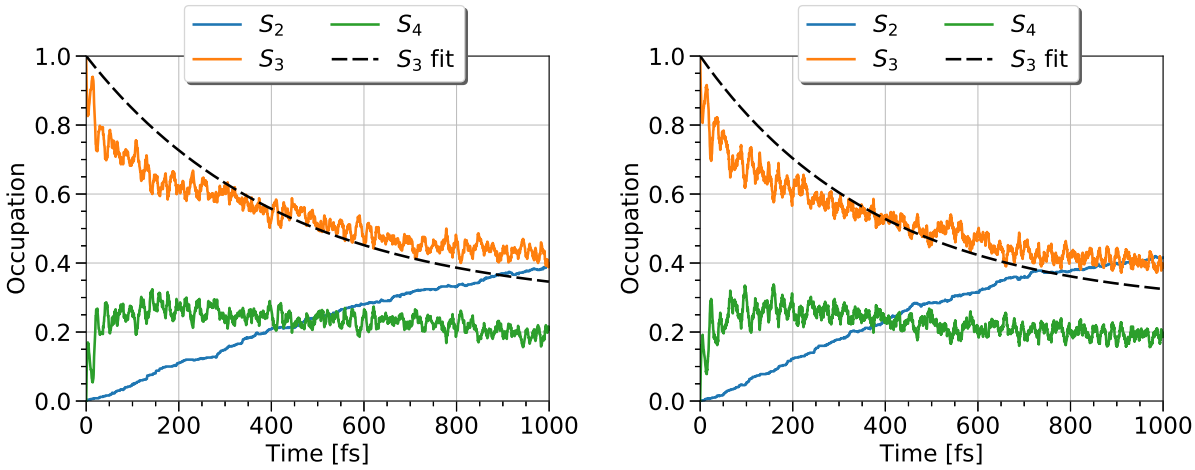
c) $\Delta t = 0.1$ fs

d) $\Delta t = 0.05$ fs

Figure 4.14: Occupations of the S_2, S_3 and S_4 excited states in phenanthrene averaged over an ensemble of 100 trajectories. Each trajectory has been propagated with an indicated time step Δt on machine-learned PESs with the Belyaev-Lebedev hopping probability. Black dashed line is an exponential fit.

Unlike before, there are significant differences in the electronic relaxation. The decay becomes an order-of-magnitude faster as Δt is decreased from 0.5 to 0.05 fs. One possible explanation is that the BL approach has been derived for a one-dimensional crossing in two-state models. Thus, it might not describe properly the non-adiabatic transitions between highly multidimensional

PESs. Furthermore, Smith and Akimov[245] pointed out that the BL-based approaches are well suited to study dynamics in manifolds of excited states with energy gaps below 0.1–0.2 eV, but may fail for states separated by gaps ⩾0.5 eV. Such relatively large gaps are observed between S_2 and S_3 excited states in the TD-DFTB absorption spectrum of phenanthrene (see Table 4.4). On the other hand, the second time derivative of the energy gap Z_{IJ} in Eqn. (4.6) is computed using finite differences, i.e. $\ddot{Z}_{IJ} \sim \mathcal{O}(\Delta t^2)$ and $P_{IJ}^{\mathrm{BL}} \propto \exp(-\Delta t)$. This issue, plus the fact that the number of hopping attemps increases as the time step is decreased (due to the larger number of MD steps that has to be calculated), may cause the observed variations. Yet, it has been shown for several systems that BL-based TSH simulations may deliver results at the same accuracy as the FSSH ones or even reproduce some trends from the quantum wavepacket calculations.[76,258] In fact, the BL-based simulations with $\Delta t = 0.25$ fs (see Figure 4.14b) are in good agreement with the FSSH calculations presented in Figure 4.9. Notably, the computed decay time for 100 trajectories is about 467 fs, which is right in between 0.44 and 0.48 ps derived from the works of Brechignac and Hermine[238] and Amirav *et al.*[246], respectively. However, the calculated value decreases to 383 fs for an ensemble of 500 trajectories (see Figure 4.15), which is still in good agreement with the experimental data. In the case of $\Delta t = 0.5$ fs, the results are qualitatively the same as the ones for $\Delta t = 0.25$ fs and the computed decay times are 486 and 419 fs for 100 and 500 trajectories, respectively. It is worth mentioning that the BL-based simulations with $\Delta t = 0.05$ fs (see Figure 4.14d) are similar to the ZN results, namely there is an ultrafast transfer of the electronic population from S_3 to S_2 and S_4 followed by oscillations in the occupation dynamics.



$\Delta t = 0.5$ fs $\Delta t = 0.25$ fs

Figure 4.15: Occupations of the S_2, S_3 and S_4 excited states in phenanthrene averaged over an ensemble 500 trajectories. Each trajectory has been propagated with the time step $\Delta t = 0.5$ fs (left panel) or $\Delta t = 0.25$ fs (left panel) on machine-learned PESs with the Belyaev-Lebedev hopping probability. Black dashed line is an exponential fit.

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 observed difference in

the electronic relaxation of phenanthrene is a direct consequence of using ML-based quantities or is due to the particular TSH implementations presented in this work. Figure 4.16 shows the occupation dynamics for an ensemble of 100 trajectories propagated using TD-DFTB energies and forces coupled to the BL or ZN approaches. The results are qualitatively similar to the previous ones obrained with machine-learned quantities. However, the decay in the BL case is slightly faster for the DFTB-based propagation compared to the ML-based one (395 fs for the former *versus* 467 fs for the latter). Nevertheless, the aforementioned dependence on the propagation time step seems to be a more important issue for practical applications of the Belyaev-Lebedev approach.

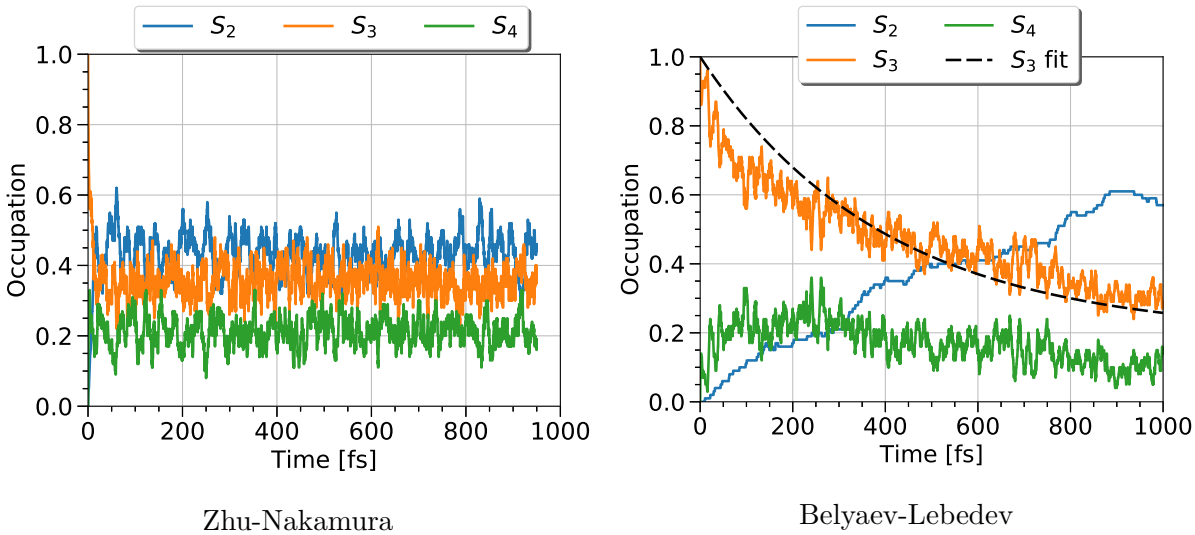


Figure 4.16: Occupations of the S_2, S_3 and S_4 excited states in phenanthrene averaged over an ensemble 100 trajectories. Each trajectory has been propagated with the time step $\Delta t = 0.25$ fs using the TD-DFTB energies and forces coupled to the Zhu-Nakamura (left panel) or Belyaev-Lebedev (right panel) hopping approaches.

Finally, 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 33 CPU hours when 4 states are included and 13 hours with 3 states (mean values averaged over 200 trajectories). Notably, the ML-based propagation of 3 states (see blue circles in Figure 4.17) takes only 30 minutes with $\Delta t = 0.25$ fs (4000 steps) and about 70 minutes with $\Delta t = 0.1$ fs (10000 steps), which is already a remarkable gain of the CPU time. However, one should also take into account the time consumed for the training step of the ML (SchNet) model, which varied between 4 and 15 hours depending on the number of training points (using one GPU, see Table 4.5). 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 (see Figure 4.17) while avoiding exponential scaling with respect to the number of excited states included in the propagation. The computational performances of BL/SchNet and ZN/SchNet approaches are almost the same

(10000 steps calculated in about 70 minutes), despite additional conversion of $3M$ adiabatic energy gradients into diabatic ones in the ZN case. However, there is a drastic difference when machine-learned quantities are replaced by the TD-DFTB ones. It takes approximately 108 or 163 minutes to perform 2000 propagation steps using BL/TD-DFTB or ZN/TD-DFTB schemes, respectively. The observed delay is mainly due to the extensive input/output required to compute energies (and forces in the ZN case) on the fly using an external deMon-Nano executable. As expected, the TSH/TD-DFTB computational performance is an intermediate case between the FSSH/TD-DFTB and approximate TSH/SchNet approaches.

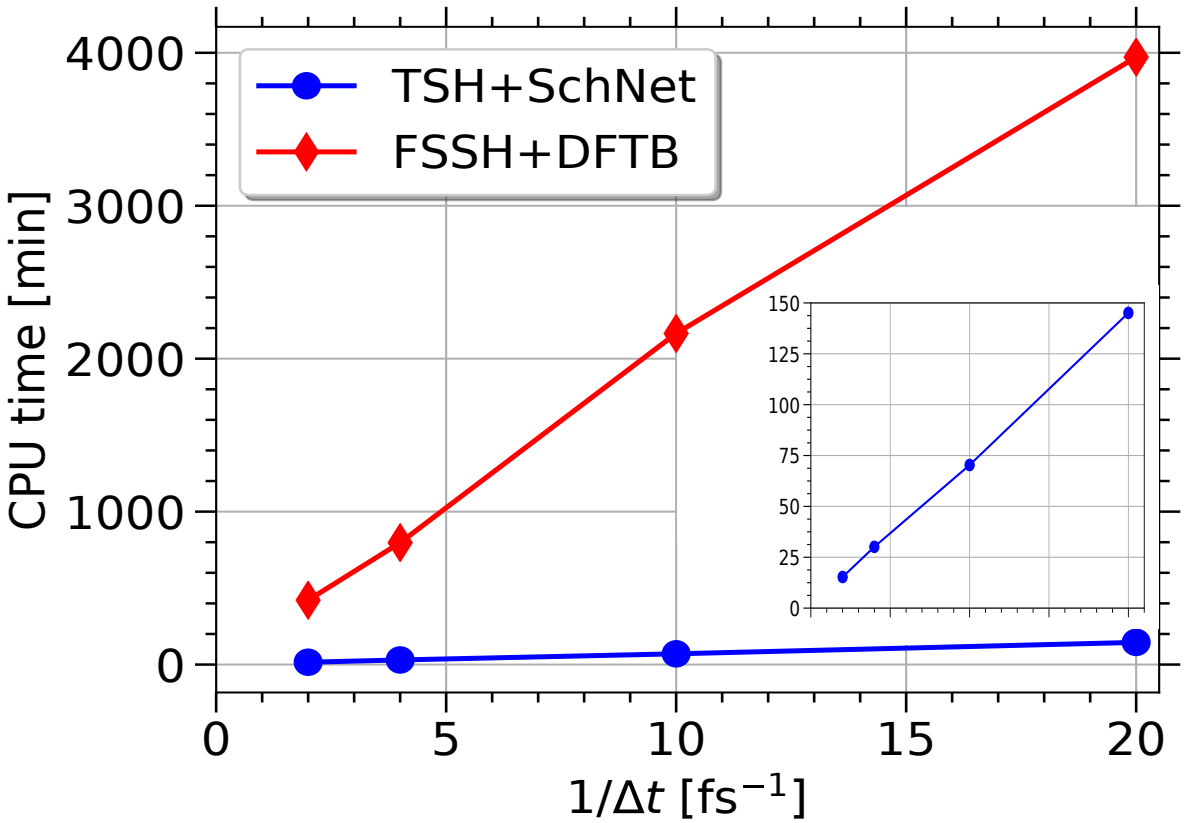


Figure 4.17: Time required to perform N=1000[fs]/Δt propagation steps for a classical trajectory of phenanthrene with 3 excited singlet states. Energies and forces are provided either by SchNet model (blue circles) or by TD-DFTB (red diamonds). Inset contains the results based on SchNet model only. The marks are connected by lines to guide the eye. Each trajectory ran on a single Skylake CPU at the CALMIP computing center in Toulouse.

4.6 Conclusion

In this Chapter, a detailed theoretical study dedicated to the non-adiabatic molecular dynamics of neutral phenanthrene has been presented. The results of several simplified TSH schemes coupled to the machine-learned PESs have been compared with those from Tully's FSSH/TD-DFTB.

In Section 4.5.1, the accuracy of the TD-DFTB method has been assessed by calculating the absorption spectrum of neutral phenanthrene and upon comparison with the TD-DFT, CASPT2 and available experimental data. The agreement is remarkably good, especially for the low-lying excited singlet states. In Section 4.5.2, the performance of several Machine Learning (SchNet) models trained on the TD-DFTB datasets have been evaluated. In summary, the SchNet architecture is able to accurately reproduce complex multidimensional PESs of phenanthrene, especially when forces are included in the training.

In Section 4.5.3, the electronic relaxation from the bright S_3 state *via* the cascade of radiationless transitions is investigated for phenanthrene. The detailed analysis reveals good agreement between the simulated decay and the experimetal results. Furthermore, the computed electronic relaxation in phenanthrene is significantly slower than what has been reported before for polyacenes. This correlates well with the large energy gap (0.57 eV) between the initially excited S_3 state and the one below in energy (S_2) in the TD-DFTB spectrum.

In Section 4.5.4, the results computed with simplified TSH schemes on SchNet-learned PESs are presented. The simulations have been performed with different time steps and hopping probabilities. First of all, the Zhu-Nakamura theory fails to reproduce the FSSH results or experimentally derived decay times. Indeed, it yields an ultrafast decay followed by oscillations in the occupation dynamics. On the contrary, the Belyaev-Lebedev approach is in better agreement with the reference data. Despite neglecting information about adiabatic energy gradients, the computed decay timescales are close the ones derived from the available experiments in supersonic jets. However, the BL-based results show significant differences for smaller time steps of the propagation. Finally, the computational performances have been compared for all approaches considered in this study.

One should precise the limits of the present conclusions. Firstly, the TD-DFTB absorption spectrum neither involves doubly excited states, nor Rydberg states. Also, interband transitions *via* the spin-orbit coupling have not been included in the FSSH propagation. Referring more particularly to the supervised Machine Learning methods, their real accuracy is determined by the training data. Thus, the aforementioned SchNet models naturally inherit all advantanges and disadvantages of the underlying TD-DFTB approach. The main drawback of some Machine Learning 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.

Overall, this study contributes to the understanding of photophysics of neutral phenanthrene. The main objective was to evaluate the performance of simplified TSH schemes and SchNet models for excited-state dynamics of large molecules. This work also aims to shed more light on the importance of an *a priori* knowledge about conical intersections or non-adiabatic couplings for Machine Learning applications. As far as I 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. These results are of interest for astrochemistry and laboratory experiments concerned with atto- or femto-second laser spectroscopy of PAHs. Referring more particularly to the species in the interstellar medium, electronic absorption band widths can be related to the relaxation timescales.[22,94] On the other hand, phenanthrene can be used as a benchmark system to study the photophysical processes due to diversity of available experimental and theoretical results.[252] Last but not least, it would be interesting to investigate the validity of Machine Learning models trained on data from the higher-level electronic structure methods and their application to excited-state dynamics of large compounds.

5
Conclusion

In this thesis, mixed quantum-classical approaches have been implemented and applied to simulate the non-adiabatic molecular dynamics of PAHs. In order to study large compounds, the TD-DFTB method has been chosen for the electronic structure calculations. The primary analysis was focused on the electronic relaxation from the higher-lying excited states and the corresponding decay times. Additional details about nuclear degrees of freedom were also discussed.

As a theoretical tool, the Tully's FSSH approach has been extensively used for non-adiabatic molecular dynamics calculations. In Chapter 2, one can find methodological and implementation details related to the FSSH scheme coupled to the TD-DFTB approach in the deMon-Nano[99] code. As shown in the following results chapters, the synergy of the two methods provided deep insights about all the processes of interest as well as about the underlying mechanisms.

The first part of the results was dedicated to excited-state dynamics of neutral cata-condensed PAHs. In Chapter 3, the validity of the TD-DFTB approach was assessed by calculating the absorption spectra of polyacenes and upon comparison with the TD-DFT, CASPT2 and available experimental data. The agreement is reasonably good, especially taking into account the number of approximations behind the TD-DFTB formalism. This was followed by a more detailed analysis dedicated to the tetracene and chrysene isomers. Furthermore, the accuracy of two sets of DFTB parameters, namely MAT and MIO, has been evaluated. The conclusion is that the MAT set is more reliable when the brightest excited singlet state of cata-condensed PAHs has to be computed. After validating the TD-DFTB approach, the first full-scale application was dedicated to the electronic relaxation from the brightest state of polyacenes ranging in size from naphthalene to heptacene. The detailed analysis revealed a striking alternation of decay times computed for linear PAH species with $n = 3$–6 aromatic cycles. This alternation was correlated with the energy gap between the initial state and the neighbouring one with lower energy. Yet, an even-odd alternation in the population transfer cannot be extrapolated to larger systems, even though this effect is well pronounced for polyacenes with 3 to 6 aromatic cycles. Further investigation is possible, but requires some improvements in the DFTB parametrization to better match the absorption spectra of octacene

and larger species. Also, the computational cost of the FSSH calculations for larger polyacenes will increase significantly due to the rapidly increasing amount of states to be taken into account. The size effect observed in polyacenes inspired me to continue towards comparative analysis. In particular, the electronic relaxation from the brightest excited state has been investigated for armchair-shaped chrysene and zigzag-shaped tetracene. Notably, the brightest singlet-singlet transition in the TD-DFTB spectra of both isomers was attributed to an excitation energy of approximately 4.6 eV. The detailed analysis revealed an order-of-magnitude difference in computed decay times between the two molecules. This difference was correlated with the energy gap between the initially excited state and the neighbouring one with lower energy, which is consistent with our previous study on polyacenes. The much faster decay rate for armchair-edge chrysene may also be enhanced by the lesser symmetry induced with respect to zigzag-edge tetracene.

The second part of the results was devoted to a more methodological study of different TSH approaches coupled to the TD-DFTB or Machine Learning for electronic structure calculations. The non-adiabatic dynamics in low-lying excited states of neutral phenanthrene was chosen as a test case due to its relatively large molecular size, due to diversity of available experimental results and due to some ambiguity in their interpretation. Firstly, the TD-DFTB absorption spectrum was analyzed and compared with the TD-DFT, CASPT2 and available experimental data. The agreement for the low-lying singlet-singlet transitions is remarkably good. Furthermore, the performance of several Machine Learning (SchNet) models trained on different TD-DFTB datasets was evaluated. The conclusion is that SchNet architecture for Deep Learning is able to accurately reproduce complex multidimensional PESs of phenanthrene, especially when forces are included in the training. Afterwards, the electronic relaxation from the bright S_3 state *via* the cascade of radiationless transitions was investigated for phenanthrene. The detailed analysis of the FSSH/TD-DFTB simulations revealed good agreement between the simulated decay and experimental results. Furthermore, the computed electronic relaxation in phenanthrene is significantly slower than what was reported before for polyacenes and chrysene. Once again, this correlates well with the relatively large TD-DFTB energy gap between the initially excited state and the one below in energy. The major focus was given to the results computed with the simplified TSH (Belyaev-Lebedev or Zhu-Nakamura) schemes on SchNet-learned PESs. The simulations were performed with different time steps and hopping probabilities. First of all, the Zhu-Nakamura theory failed to reproduce the FSSH results and experimentally derived relaxation timescales. The observed decay was ultrafast and followed by oscillations in the occupation dynamics. On the contrary, results of the Belyaev-Lebedev approach were in better agreement with the reference data. Despite neglecting information about adiabatic energy gradients, the computed decay timescales are close to the ones derived from the available experiments in supersonic jets. However, these results also showed significant differences for smaller time steps of the propagation. Finally, the computational performances were compared for all approaches considered in this study.

One should precise the limits of the present conclusions. First of all, the TD-DFTB absorption spectrum neither involved doubly excited states, nor Rydberg states. Also, interband transitions *via* the spin-orbit coupling were not included in the FSSH propagation. Referring more particularly to the supervised Machine Learning methods, their real accuracy is determined by the training data. Thus, the aforementioned SchNet models naturally inherited all advantanges and disadvantages of the underlying TD-DFTB approach.

Overall, this study contributes to the understanding of photophysics of large PAHs and sheds light on the photostability of the considered cata-condensed compounds following absorption of a UV photon. It also provides insights into possible effects of the size and/or shape of the corresponding molecular complex on the excited-state dynamics. These results are of interest for astrochemistry, atmospheric chemistry and laboratory experiments concerned with atto- or femto-second laser spectroscopy of PAHs. It is worth mentioning that higher polyacenes have low-lying excited states with excitation energies in the visible UV range. Thus, they can be observed and assigned if they survive the strong UV radiation. Assessment of the validity of the Kasha's rule for PAHs is also important in this context. Referring more particularly to the species in the interstellar medium, electronic absorption band widths can be related to the relaxation timescales.[22,94] Both the observed size- or shape-dependent relaxation mechanisms and the different final state populations might contribute to the assignment of DIBs. Finally, it is worth mentioning the correlation between the decay timescales and the energy gaps that was reported for all molecules considered in this work. The methodological part of this study is important as it aims to make a connection between the state-of-the-art Machine Learning techniques and the non-adiabatic dynamics. This is particularly promising due to a growing interest in excited-state dynamics of large molecular compounds.

Perspectives

Even though different aspects of the non-adiabatic molecular dynamics of neutral PAHs have been studied, the aforementioned results are still limited by the underlying electronic structure methods. Firstly, it is worth investigating the dynamics of charged compounds, in particular, the cationic ones due to diversity of available theoretical and experimental studies. This can be achieved with an extension of the existing spin-unrestricted TD-DFTB[145,146] formalism for open-shell systems. Secondly, the TD-DFTB coupled to the long-range correction[128,259] or enhanced by the transition density matrix analysis[89] can be used to study the charge transfer phenomena. On the other hand, the DFTB-CI approach is a promising tool that may be adapted to investigate the non-adiabatic dynamics of the cationic clusters of large PAHs.[149]

Moreover, the Tully's FSSH approach itself deserves certain attention. Firstly, the computation of non-adiabatic couplings remains a major challenge for the single-reference methods like the linear response TD-DFT(B). There are certain implementation ambiguities that should be

precised in the future. Moreover, the commonly used decoherence corrections are somewhat *ad hoc* and there are numeruous studies addressing this issue.[260,261] For the slow relaxation phenomena involving interband transitions, the triplet states have to be taken into account *via* the spin-orbit coupling[262] as implemented, for instance, in the SHARC software.[194]

Finally, future pump-probe experiments will shed more light on the ultrafast dynamics of large compounds. It would be interesting to investigate whether the theoretically computed differentiation of the electronic relaxation can be observed in the experiments. In order to support these studies, a more accurate description of the non-adiabatic processes will be helpful. For example, the MCTDH approach can be coupled to the Machine Learning potentials to perform the full quantum dynamics.[228] On the other hand, it has been shown that even the TSH propagation with machine-learned quantities may provide reliable results, especially if a higher-level electronic structure method is used to train a thoroughly crafted model.[81,82,210] Last but certainly not least, a critical assessment of the PAH hypothesis requires more information from the astronomical observations, laboratory experiments and theoretical calculations. Perhaps, the James Webb Space Telescope[263] with its remarkable spectroscopic precision will make a solid contribution to (or even it will wrap up) this long-lasting debate.

A

Example of the deMon-Nano input

One can find below a deMon-Nano input file that corresponds to a single FSSH-trajectory of neutral naphthalene that will be launched in S_4 (EXST=4) with a set of coordinates and velocities listed after GEOMETRY and MDYNAMICS READ, respectively. This trajectory and five corresponding electronically excited singlet states (NSTATES=5) will be propagated during 300 fs (1200 iterations with 0.25 fs nuclear timestep) with the decoherence correction (DECOH). The set of the SCC-DFTB parameters used in this example is matsci (PARAMETER PTYPE=MAT).[131]

```
# ==========================================

TSH NSTATES=5 DECOH

MDSTEPS MAX=1200
TIMESTEP 0.25
MDBATH NONE
MDYNAMICS READ
     1 -0.009538  0.001956 -0.015569
     2  0.002349  0.000772 -0.005799
     3 -0.000369 -0.004173  0.001722
     4 -0.010553  0.025768  0.040311
     5 -0.001188 -0.004119  0.005521
     6 -0.001331 -0.001191 -0.002196
     7 -0.003464  0.007135 -0.003150
     8 -0.000443 -0.000674  0.001227
     9  0.005298  0.020137  0.000709
    10  0.000358  0.008309 -0.000792
    11  0.002653 -0.006598 -0.000521
    12  0.010912  0.015617  0.006899
    13 -0.000560  0.002157  0.000564
```

```
14 -0.012375  0.021079  0.013728
15  0.001831 -0.003539 -0.006311
16 -0.000931 -0.014096  0.006531
17 -0.000156  0.001688  0.001197
18 -0.016829  0.010208  0.014758

DFTB SCC LRESP EXST=4
CHARGE 0.0
PARAMETER PTYPE=MAT
/home/deMonNano/basis

GEOMETRY
H  -1.310789 -2.534129  0.080092
C  -1.224067 -1.424951 -0.038956
C  -2.441200 -0.701885  0.125032
H  -3.370416 -1.253638  0.446435
C   0.029150 -0.712281 -0.167751
C   1.258212 -1.369851 -0.044249
H   1.196673 -2.498541 -0.137696
C   2.402380 -0.734269  0.124435
H   3.317460 -1.277610  0.242009
C   0.009471  0.738124 -0.101058
C   1.247506  1.405530 -0.045984
H   1.168816  2.530245 -0.080421
C   2.430833  0.706960  0.093188
H   3.359773  1.248540  0.324237
C  -1.241625  1.395886 -0.046832
H  -1.233189  2.526859  0.013122
C  -2.453124  0.693736  0.012098
H  -3.337282  1.294034  0.185612

# ==========================================
```

B

Supplementary spectroscopic data

In this Appendix, one can find more details regarding the absorption spectra of cata-condensed PAHs from Section 3.2.

Table S1: Experimental[181–183] and theoretical[90,100,180] excitation energies for first and brightest excited singlet states of polyacenes ranging in size from naphthalene to octacene. Theoretical values computed at CASPT2[100,180], TD-DFT (this work) and TD-DFTB (this work) levels of theory.

	Lowest singlet excitation energy, eV				Brightest singlet excitation energy, eV			
Number of cycles	CASPT2	TD-DFT	TD-DFTB	Exp.	CASPT2	TD-DFT	TD-DFTB	Exp.
2	4.02	4.16	4.23	4.45	5.54	5.89	5.88	5.89
3	3.48	2.96	3.12	3.42	–	5.04	5.11	5.24
4	2.79	2.20	2.39	2.63	–	4.46	4.56	4.51
5	2.20	1.66	1.88	2.12	–	4.01	4.14	4.10
6	1.89	1.28	1.51	1.89	–	3.68	3.83	3.99
7	1.66	0.97	1.23	1.70	–	3.42	3.58	3.80
8	–	0.65	1.01	1.54	–	3.18	3.39	3.78

Absorption spectra of chrysene computed with TD-DFTB (MIO/MAT sets of parameters) and TD-DFT (BLYP/B3LYP/CAM-B3LYP functionals) with 6-31G(d,p) basis set can be found on the next 2 pages. All spectra have been truncated to 6.2 eV (200 nm).

Table S2: TD-DFTB with MIO parameters.

Excitation energy, eV	Oscillator strength
3.328	0.1019013
3.464	0.0105935
3.813	0.0000000
4.210	0.3176991
4.299	0.0000000
4.351	0.0000000
4.383	0.0000000
4.447	0.4925214
4.541	0.0000000
4.686	0.1447922
4.712	0.0000000
4.865	0.0000000
4.992	0.2854493
5.073	0.0000000
5.091	0.0000000
5.170	0.0000000
5.254	0.0713305
5.330	0.0000000
5.360	0.0000000
5.403	0.0317962
5.494	0.0000000
5.509	0.0000000
5.526	0.0000000
5.558	0.0237825
5.637	0.0000000
5.637	0.0485070
5.721	0.0000000
5.727	0.0000000
5.763	0.2225047
5.833	0.0000000
5.884	0.0000000
5.959	0.0000000
5.974	0.0000000
6.059	0.0000000
6.146	0.0000000
6.148	0.0000000

Table S3: TD-DFTB with MAT parameters.

Excitation energy, eV	Oscillator strength
3.455	0.1177502
3.614	0.0101981
3.975	0.0000000
4.366	0.3742217
4.467	0.0000000
4.535	0.0000000
4.559	0.0000000
4.606	0.5653321
4.866	0.1375454
4.897	0.0000000
5.190	0.3112332
5.271	0.0000000
5.453	0.0000000
5.455	0.0835404
5.552	0.0000000
5.626	0.0275313
5.670	0.0000000
5.715	0.0000000
5.749	0.0392861
5.854	0.0000000
5.855	0.0305762
5.975	0.2462115
6.020	0.0000000
6.092	0.0000000
6.100	0.0000000
6.183	0.0000000

Table S4: TD-DFT with BLYP/6-31G(d,p).

Excitation energy, eV	Oscillator strength
3.4081	0.0535
3.4789	0.0193
3.7698	0.0000
4.2111	0.2799
4.2232	0.0000
4.3458	0.0000
4.4272	0.5212
4.5388	0.0000
4.7114	0.0000
4.8074	0.0027
4.9052	0.5520
5.1134	0.0000
5.2715	0.0044
5.4442	0.1022
5.5672	0.0000
5.5922	0.0000
5.5981	0.2754
5.6881	0.0000
5.7398	0.0000
5.7684	0.1923
5.9091	0.0000
5.9763	0.0000
5.9912	0.0000
6.1722	0.0000
6.1756	0.1869

Table S5: TD-DFT with B3LYP/6-31G(d,p).

Excitation energy, eV	Oscillator strength
3.7823	0.0359
3.8458	0.0726
4.4146	0.0000
4.7032	0.0000
4.7076	0.6210
4.7822	0.0000
4.8487	0.7252
5.2492	0.0000
5.3599	0.0000
5.4881	0.0005
5.6495	0.5533
5.7654	0.0155
5.9115	0.0000
6.0730	0.0000
6.1152	0.1302

Table S6: TD-DFT with CAM-B3LYP/6-31G(d,p).

Excitation energy, eV	Oscillator strength
4.1761	0.0061
4.3082	0.1321
4.9825	0.0000
5.2412	1.6569
5.2474	0.0000
5.4338	0.2433
5.5399	0.0000
6.0626	0.0000
6.1795	0.0045

C

Supplementary population analysis

In this Appendix, one can find some benchmark FSSH calculations.

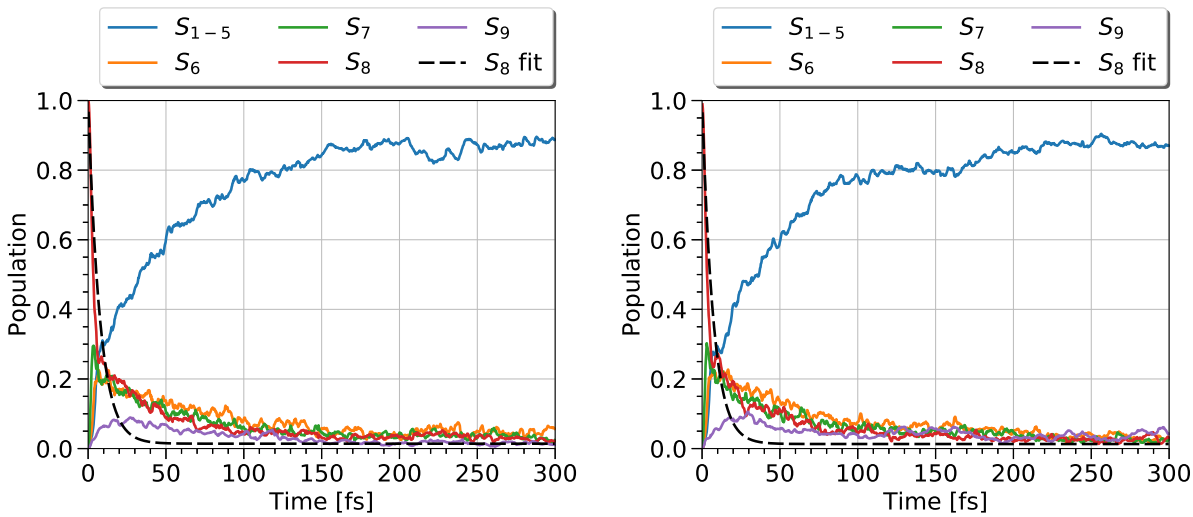


Figure S1: Populations averaged over 89 trajectories in chrysene following excitation to the brightest excited S_8 state with 0.1 meV (left panel) and 10 meV (right panel) energy gap thresholds. These simulations have been performed using the MAT set of DFTB parameters.

D

Interface with the Atomic Simulation Environment (ASE)

In this Appendix, one can find several snapshots and examples of my interface between the ASE Python package and the deMon-Nano code. ASE is a powerfull tool that contains many advanced computational methods (see Ref. [241] and references therein for more details), some of which have not been implemented yet in the deMon-Nano code (e.g. genetic algorithm for global optimization). Thus, I believe that my interface will be useful for future generations of the deMon-Nano users.

First of all, an installation procedure and requirements for the ASE can be found on the dedicated webpage: https://wiki.fysik.dtu.dk/ase/install.html. I highly recommend to install ASE and all required Python packages in the virtual environment such as miniconda. This helps to avoid issues that arise from accidentally updating one of the linked packages. Several environmental variables have to be set after installation. In particular, it is important to provide paths to the compiled deMon-Nano executable and to the folder with DFTB parameters (basis). Below is an example of two lines that can be introduced in the configuration file (e.g. in `.bashrc`) on the machine using ASE/deMon-Nano interface:

```
export DEMONNANO_BASIS_PATH='/path/deMonNano/basis/'
export ASE_DEMONNANO_COMMAND='/path/deMonNano/bin/deMon.x'
```

where `path` has to be replaced with a path to the deMon-Nano folder on your machine. Additionally, if a specific makefile is used (see `makefiles/sys/` subdirectory in the deMon-Nano folder), the resulting name of the deMon-Nano binary may differ from `deMon.x`. For example, I use a `./compile.sh evgeny` command to compile a deMon-Nano executable from the source code on one of my machines. It follows instructions provided in the `makefiles/sys/sys-evgeny.mak` file and builds a `deMon.evgeny.x` binary in the `bin/` directory. Thus, I have to use the following command

```
export ASE_DEMONNANO_COMMAND='/home/user/deMonNano/bin/deMon.evgeny.x'
```

To check that your version of the deMon-Nano contains the ASE interface, try adding the `PRINT ASE` line to the input file. If deMon-Nano execution with this line is successful, then `deMon.ase` file should appear in the same directory with the DFTB total energy written inside. You may have forces in the same file depending on the input parameters (e.g. if MDYNAMICS keyword is activated).

Once you are all set, try running the following line (note double hyphen '-' symbol in front of `calculators`):

```
ase test --calculators=demonnano
```

It should execute a number of custom ASE tests together with some additional benchmarks related to the deMon-Nano itself (see `demonnano/h2o.py` line in Figure S2). For the latter one, the ASE will compute energy, forces and optimized geometry for a water molecule in the ground state with the MAT set of parameters. The test is passed if results match the reference values that have been precomputed in advance (so there is no need to change them manually).

```
cp2k/cp2k_H2_LS.py                    0.00s SKIPPED: use --calculators=cp2k to enable
cp2k/cp2k_H2_None.py                  0.00s SKIPPED: use --calculators=cp2k to enable
cp2k/cp2k_H2_PBE.py                   0.00s SKIPPED: use --calculators=cp2k to enable
cp2k/cp2k_H2_libxc.py                 0.00s SKIPPED: use --calculators=cp2k to enable
cp2k/cp2k_MD.py                       0.00s SKIPPED: use --calculators=cp2k to enable
cp2k/cp2k_O2.py                       0.00s SKIPPED: use --calculators=cp2k to enable
cp2k/cp2k_dcd.py                      0.00s SKIPPED: cp2k command not available
cp2k/cp2k_restart.py                  0.00s SKIPPED: use --calculators=cp2k to enable
cp2k/cp2k_stress.py                   0.00s SKIPPED: use --calculators=cp2k to enable
crystal/bulk.py                       0.00s SKIPPED: use --calculators=crystal to enable
crystal/graphene.py                   0.00s SKIPPED: use --calculators=crystal to enable
crystal/molecule.py                   0.00s SKIPPED: use --calculators=crystal to enable
db/config.py                          0.00s OK
constraints/mirror.py                 2.67s OK
db/db2.py                             0.41s OK
db/db_web.py                          0.00s SKIPPED: no flask module
db/jsondb.py                          0.00s OK
db/metadata.py                        0.19s OK
db/mysql_tests.py                     0.00s SKIPPED: No MySQL module
db/o2b2o.py                           0.00s OK
db/sql_db_ext_tables.py               0.46s OK
db/update.py                          0.40s OK
demon/h2o.py                          0.00s SKIPPED: use --calculators=demon to enable
demon/h2o_xas_xes.py                  0.00s SKIPPED: use --calculators=demon to enable
demonnano/h2o.py                      2.64s OK
dft/bandgap.py                        0.00s OK
dft/dos.py                            0.00s OK
dft/hex.py                            0.20s OK
dft/interpolate.py                    0.02s OK
dft/monoclinic.py                     0.78s OK
dftb/h2o.py                           0.00s SKIPPED: use --calculators=dftb to enable
dftb/n2ni111.py                       0.01s SKIPPED: use --calculators=dftb to enable
```

Figure S2: A snapshot of the `ase test --calculators=demonnano` output in the command line.

One can find below an example of the geometry optimization performed with the ASE using a built-in BFGS optimizer and the deMon-Nano code for the electronic structure calculations.

```
# ==========================================================================

from ase.calculators.demonnano import DemonNano
from ase import Atoms
from ase.io import write
from ase.optimize import BFGS
import numpy as np

# create a molecule
d = 0.9775
t = np.pi/180*110.51
mol = Atoms('H2O',
            positions=[(d, 0, 0),
                       (d*np.cos(t), d*np.sin(t), 0),
                       (0, 0, 0)])

# set deMon-Nano input parameters
input_arguments = {'DFTB': 'SCC',
                   'CHARGE': '0.0',
                   'PARAM': 'PTYPE=BIO'}

# create and set ASE calculator
calc = DemonNano(label='rundir/',input_arguments=input_arguments)
mol.set_calculator(calc)

# optimize geometry using BFGS
dyn = BFGS(mol, trajectory='test.traj')
dyn.run(fmax=0.01)

# write output
write('h2o_optimized.xyz', mol)

# ==========================================================================
```

The aforementioned codes and detailed description of the deMon-Nano calculator class in ASE can be found on the dedicated webpage (see Figure S3 and https://wiki.fysik.dtu.dk/ase/ase/calculators/demonnano.html).

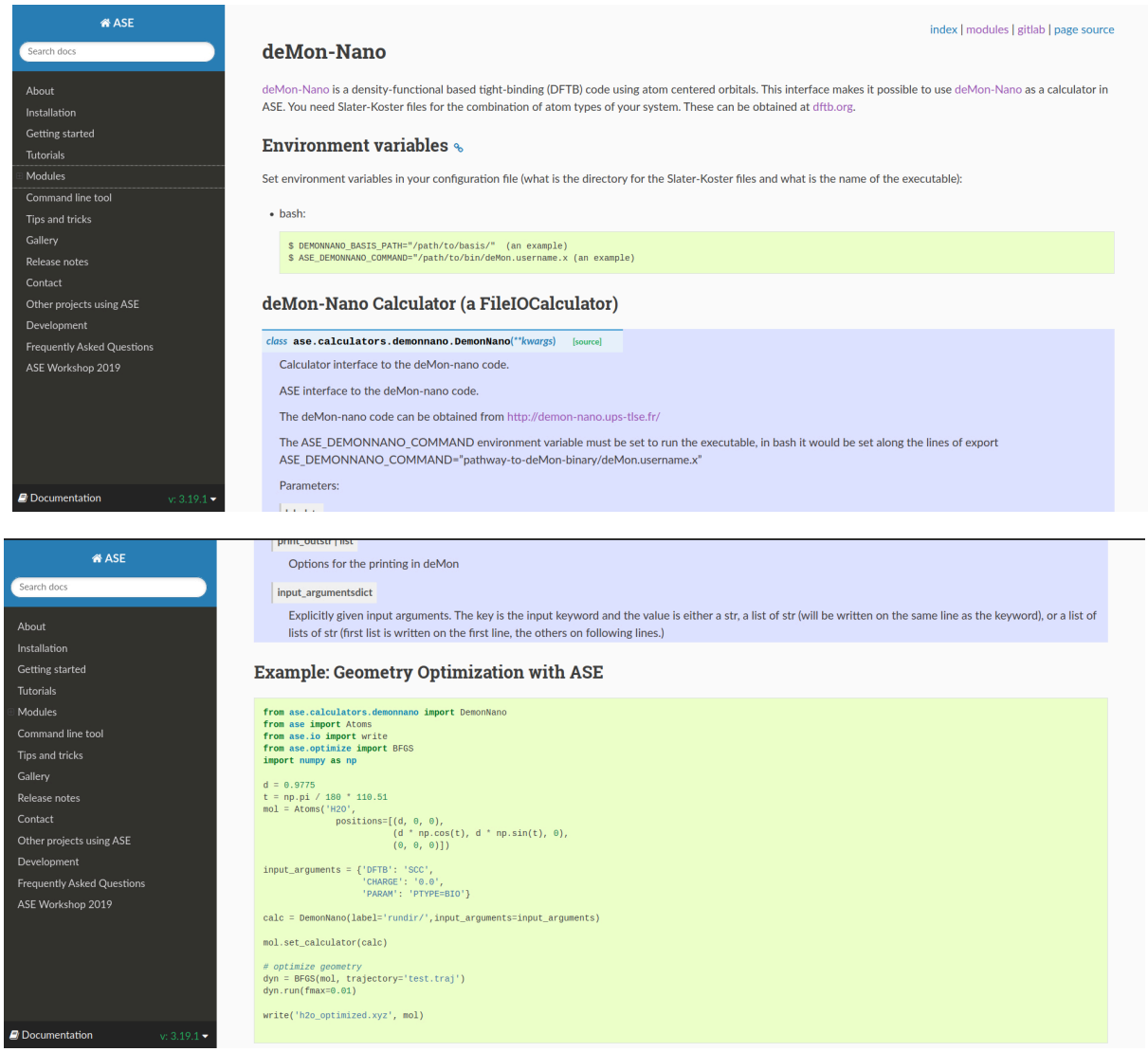


Figure S3: Screenshots from the page dedicated to the deMon-Nano calculator class on the official ASE website.

Bibliography

[1] F. Buonomo, G. Carraro, C. Chiosi, and C. Lia, Monthly Notices of the Royal Astronomical Society **312**, 371 (2000), ISSN 0035-8711, https://academic.oup.com/mnras/article-pdf/312/2/371/2949159/312-2-371.pdf, URL https://doi.org/10.1046/j.1365-8711.2000.03132.x.

[2] D. J. Hollenbach and A. G. G. M. Tielens, Rev. Mod. Phys. **71**, 173 (1999), URL https://link.aps.org/doi/10.1103/RevModPhys.71.173.

[3] M. W. Werner, T. L. Roellig, F. J. Low, G. H. Rieke, M. Rieke, W. F. Hoffmann, E. Young, J. R. Houck, B. Brandl, G. G. Fazio, et al., Astrophys. J. **154**, 1 (2004), astro-ph/0406223.

[4] N. Turro, *Modern Molecular Photochemistry* (University Science Books, 1991), ISBN 9780935702712, URL https://books.google.fr/books?id=KPlz1EEDiXsC.

[5] M. Klessinger and J. Michl, *Excited States and Photo-Chemistry of Organic Molecules* (Wiley, 1995).

[6] K. Sellgren, Astrophys. J. **277**, 623 (1984).

[7] A. Léger and J. L. Puget, Astron. Astrophys. **137**, L5 (1984).

[8] L. J. Allamandola, A. G. G. M. Tielens, and J. R. Barker, Astrophys. J. **290**, L25 (1985), ISSN 0004-637X.

[9] C. Joblin, A. G. G. M. Tielens, L. J. Allamandola, and T. R. Geballe, Astrophys. J. **458**, 610 (1996).

[10] A. Tielens, Annual Review of Astronomy and Astrophysics **46**, 289 (2008), https://doi.org/10.1146/annurev.astro.46.060407.145211, URL https://doi.org/10.1146/annurev.astro.46.060407.145211.

[11] C. Joblin and A. G. G. M. Tielens, eds., *Proceedings, PAHs and the Universe: A Symposium to Celebrate the 25th Anniversary of the PAH Hypothesis*, vol. 46. 46 of *EAS Publication Series*, EAS (EAS, 2011), ISBN 9782759806249, URL https://www.eas-journal.org/articles/eas/abs/2011/01/contents/contents.html.

[12] G. H. Herbig, Annual Review of Astronomy and Astrophysics **33**, 19 (1995), https://doi.org/10.1146/annurev.aa.33.090195.000315, URL https://doi.org/10.1146/annurev.aa.33.090195.000315.

[13] F. Salama, G. A. Galazutdinov, J. Krełowski, L. J. Allamandola, and F. A. Musaev, Astrophys. J. **526**, 265 (1999), URL https://ui.adsabs.harvard.edu/#abs/1999ApJ...526..265S.

[14] F. Salama, E. L. O. Bakes, L. J. Allamandola, and A. G. G. M. Tielens, Astrophys. J. **458**, 621 (1996), URL https://ui.adsabs.harvard.edu/#abs/1996ApJ...458..621S.

[15] G. P. van der Zwet and L. J. Allamandola, Astron. Astrophys. **146**, 76 (1985), URL https://ui.adsabs.harvard.edu/#abs/1985A&A...146...76V.

[16] A. Léger and L. D'Hendecourt, Astron. Astrophys. **146**, 81 (1985), URL https://ui.adsabs.harvard.edu/#abs/1985A&A...146...81L.

[17] M. K. Crawford, A. G. G. M. Tielens, and L. J. Allamandola, Astrophys. J. Lett. **293**, L45 (1985), URL https://ui.adsabs.harvard.edu/#abs/1985ApJ...293L..45C.

[18] K. F. Hall, M. Boggio-Pasqua, M. J. Bearpark, and M. A. Robb, The Journal of Physical Chemistry A **110**, 13591 (2006), https://doi.org/10.1021/jp064711g, URL https://doi.org/10.1021/jp064711g.

[19] A. M. Tokmachev, M. Boggio-Pasqua, M. J. Bearpark, and M. A. Robb, The Journal of Physical Chemistry A **112**, 10881 (2008), https://doi.org/10.1021/jp8044109, URL https://doi.org/10.1021/jp8044109.

[20] A. M. Tokmachev, M. Boggio-Pasqua, D. Mendive-Tapia, M. J. Bearpark, and M. A. Robb, The Journal of Chemical Physics **132**, 044306 (2010), https://doi.org/10.1063/1.3278545, URL https://doi.org/10.1063/1.3278545.

[21] S. Ghanta, V. S. Reddy, and S. Mahapatra, Phys. Chem. Chem. Phys. **13**, 14523 (2011), URL http://dx.doi.org/10.1039/C1CP21083A.

[22] V. S. Reddy, S. Ghanta, and S. Mahapatra, Physical Review Letters **104**, 111102 (2010), URL https://link.aps.org/doi/10.1103/PhysRevLett.104.111102.

[23] S. N. Reddy and S. Mahapatra, The Journal of Physical Chemistry A **117**, 8737 (2013), URL https://doi.org/10.1021/jp4033645.

[24] S. Ghanta, V. S. Reddy, and S. Mahapatra, Phys. Chem. Chem. Phys. **13**, 14531 (2011), URL http://dx.doi.org/10.1039/C1CP21084J.

[25] A. Marciniak, V. Despré, T. Barillot, A. Rouzée, M. C. E. Galbraith, J. Klei, C. H. Yang, C. T. L. Smeenk, V. Loriot, S. N. Reddy, et al., Nature Communications **6**, 7909 (2015), URL https://doi.org/10.1038/ncomms8909.

[26] A. Marciniak, V. Despré, V. Loriot, G. Karras, M. Hervé, L. Quintard, F. Catoire, C. Joblin, E. Constant, A. I. Kuleff, et al., Nature Communications **10**, 337 (2019), ISSN 2041-1723, URL https://doi.org/10.1038/s41467-018-08131-8.

[27] E. Peeters, S. Hony, C. Van Kerckhoven, A. G. G. M. Tielens, L. J. Allamandola, D. M. Hudgins, and C. W. Bauschlicher, Astron. Astrophys. **390**, 1089 (2002), astro-ph/0205400.

[28] H. Andrews, C. Boersma, M. W. Werner, J. Livingston, L. J. Allamandola, and A. G. G. M. Tielens, The Astrophysical Journal **807**, 99 (2015), URL https://doi.org/10.1088%2F0004-637x%2F807%2F1%2F99.

[29] Rosenberg, Marissa J. F., Berné, Olivier, and Boersma, Christiaan, A&A **566**, L4 (2014), URL https://doi.org/10.1051/0004-6361/201423953.

[30] E. K. Campbell, M. Holz, D. Gerlich, and J. P. Maier, Nature **523**, 322 (2015), ISSN 1476-4687, URL https://doi.org/10.1038/nature14566.

[31] J. Cami, J. Bernard-Salas, E. Peeters, and S. E. Malek, Science **329**, 1180 (2010).

[32] J. Palotás, J. Martens, G. Berden, and J. Oomens, Nature Astronomy **4**, 240 (2020), ISSN 2397-3366, URL https://doi.org/10.1038/s41550-019-0941-6.

[33] J. Soler, R. Sarkar, and M. Boggio-Pasqua, The Journal of Physical Chemistry A **123**, 1824 (2019), https://doi.org/10.1021/acs.jpca.8b11761, URL https://doi.org/10.1021/acs.jpca.8b11761.

[34] W. Karcher, R. Fordham, J. Dubois, P. Glaude, and J. Ligthart (1985).

[35] H.-H. Perkampus, *UV-vis atlas of organic compounds, 2nd ed.* (Weinheim: VCH, 1992), ISBN 3527285105.

[36] G. Malloci, C. Joblin, and G. Mulas, Chemical Physics **332**, 353 (2007), ISSN 0301-0104, URL http://www.sciencedirect.com/science/article/pii/S0301010407000031.

[37] C. W. Bauschlicher, A. Ricca, C. Boersma, and L. J. Allamandola, The Astrophysical Journal Supplement Series **234**, 32 (2018), URL https://doi.org/10.3847%2F1538-4365%2Faaa019.

[38] J. D. Thrower, E. E. Friis, A. L. Skov, B. Jørgensen, and L. Hornekær, Phys. Chem. Chem. Phys. **16**, 3381 (2014), URL http://dx.doi.org/10.1039/C3CP54073A.

[39] P. A. Jensen, M. Leccese, F. D. S. Simonsen, A. W. Skov, M. Bonfanti, J. D. Thrower, R. Martinazzo, and L. Hornekær, Monthly Notices of the Royal Astronomical Society **486**, 5492 (2019), ISSN 0035-8711, https://academic.oup.com/mnras/article-pdf/486/4/5492/28665904/stz1202.pdf, URL https://doi.org/10.1093/mnras/stz1202.

[40] D. Campisi, F. D. S. Simonsen, J. D. Thrower, R. Jaganathan, L. Hornekær, R. Martinazzo, and A. G. G. M. Tielens, Phys. Chem. Chem. Phys. **22**, 1557 (2020), URL http://dx.doi.org/10.1039/C9CP05440E.

[41] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964), URL https://link.aps.org/doi/10.1103/PhysRev.136.B864.

[42] Y. Garniron, Theses, Université Paul Sabatier - Toulouse III (2018), URL https://tel.archives-ouvertes.fr/tel-02089570.

[43] F. Pauzat, D. Talbi, M. D. Miller, D. J. DeFrees, and Y. Ellinger, The Journal of Physical Chemistry **96**, 7882 (1992), https://doi.org/10.1021/j100199a011, URL https://doi.org/10.1021/j100199a011.

[44] C. W. Bauschlicher and E. L. O. Bakes, Chemical Physics **262**, 285 (2000).

[45] Malloci, G., Mulas, G., and Joblin, C., A&A **426**, 105 (2004), URL https://doi.org/10.1051/0004-6361:20040541.

[46] A. Simon, M. Rapacioli, M. Lanza, B. Joalland, and F. Spiegelman, Phys. Chem. Chem. Phys. **13**, 3359 (2011), URL http://dx.doi.org/10.1039/C0CP00990C.

[47] B. Joalland, M. Rapacioli, A. Simon, C. Joblin, C. J. Marsden, and F. Spiegelman, The Journal of Physical Chemistry A **114**, 5846 (2010), https://doi.org/10.1021/jp911526n, URL https://doi.org/10.1021/jp911526n.

[48] Jolibois, F., Klotz, A., Gadéa, F. X., and Joblin, C., A&A **444**, 629 (2005), URL https://doi.org/10.1051/0004-6361:20053508.

[49] M. Rapacioli, S. Cazaux, N. Foley, A. Simon, R. Hoekstra, and T. Schlathölter, Phys. Chem. Chem. Phys. **20**, 22427 (2018), URL http://dx.doi.org/10.1039/C8CP03024C.

[50] A. Simon, M. Rapacioli, G. Rouaut, G. Trinquier, and F. X. Gadéa, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **375**, 20160195 (2017), https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2016.0195, URL https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2016.0195.

[51] A. Simon, J.-P. Champeaux, M. Rapacioli, P. Moretto-Capelle, F. X. Gadéa, and M. Sence, Theoretical Chemistry Accounts: Theory, Computation, and Modeling **137**, 106 (2018), URL https://hal.archives-ouvertes.fr/hal-01873276.

[52] P. Jusko, A. Simon, G. Wenzel, S. Brünken, S. Schlemmer, and C. Joblin, Chemical Physics Letters **698**, 206 (2018), ISSN 0009-2614, URL http://www.sciencedirect.com/science/article/pii/S0009261418302124.

[53] S. Rodriguez Castillo, A. Simon, and C. Joblin, International Journal of Mass Spectrometry **429**, 189 (2018), 1711.03046.

[54] M. Rapacioli, A. Simon, C. C. M. Marshall, J. Cuny, D. Kokkin, F. Spiegelman, and C. Joblin, The Journal of Physical Chemistry A **119**, 12845 (2015), https://doi.org/10.1021/acs.jpca.5b09494, URL https://doi.org/10.1021/acs.jpca.5b09494.

[55] C. Dubosq, C. Falvo, F. Calvo, M. Rapacioli, P. Parneix, T. PINO, and A. Simon, Astronomy and Astrophysics - A&A **625**, L11 (2019), URL https://hal.archives-ouvertes.fr/hal-02087987.

[56] C. Dubosq, F. Calvo, M. Rapacioli, E. Dartois, T. Pino, C. Falvo, and A. Simon, Astronomy and Astrophysics - A&A **634**, A62 (2020), URL https://hal.archives-ouvertes.fr/hal-02419276.

[57] L. Dontot, F. Spiegelman, and M. Rapacioli, Journal of Physical Chemistry A **123**, 9531 (2019), URL https://hal.archives-ouvertes.fr/hal-02385894.

[58] S. Zamith, M. Ji, J.-M. L'Hermite, C. Joblin, L. Dontot, M. Rapacioli, and F. Spiegelman, Journal of Chemical Physics **151**, 194303 (2019), URL https://hal.archives-ouvertes.fr/hal-02375800.

[59] C. Joblin, L. Dontot, G. Garcia, F. Spiegelman, M. Rapacioli, L. Nahon, P. Parneix, T. Pino, and P. Bréchignac, Journal of Physical Chemistry Letters **8**, 3697 (2017), URL https://hal.archives-ouvertes.fr/hal-01583599.

[60] M. Boggio-Pasqua, Habilitation à diriger des recherches, Université Toulouse III (2015), URL https://tel.archives-ouvertes.fr/tel-01184241.

[61] B. F. E. Curchod and T. J. Martínez, Chemical Reviews **118**, 3305 (2018), https://doi.org/10.1021/acs.chemrev.7b00423, URL https://doi.org/10.1021/acs.chemrev.7b00423.

[62] E. Fabiano, T. Keal, and W. Thiel, Chemical Physics **349**, 334 (2008), ISSN 0301-0104, URL http://www.sciencedirect.com/science/article/pii/S0301010408000785.

[63] M. Persico and G. Granucci, Theoretical Chemistry Accounts **133**, 1526 (2014).

[64] M. Barbatti, Wiley Interdisciplinary Reviews: Computational Molecular Science **1**, 620 (2011), https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.64, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/wcms.64.

[65] H.-D. Meyer, U. Manthe, and L. Cederbaum, Chemical Physics Letters **165**, 73 (1990), ISSN 0009-2614, URL http://www.sciencedirect.com/science/article/pii/000926149087014I.

[66] M. Beck, A. Jäckle, G. Worth, and H.-D. Meyer, Physics Reports **324**, 1 (2000), ISSN 0370-1573, URL http://www.sciencedirect.com/science/article/pii/S0370157399000472.

[67] D. Mendive-Tapia, B. Lasorne, G. A. Worth, M. A. Robb, and M. J. Bearpark, The Journal of Chemical Physics **137**, 22A548 (2012), https://doi.org/10.1063/1.4765087, URL https://doi.org/10.1063/1.4765087.

[68] P. Ehrenfest, Zeitschrift für Physik **45**, 455 (1927), ISSN 0044-3328, URL https://doi.org/10.1007/BF01329203.

[69] G. Worth, M. Robb, and B. Lasorne, Molecular Physics **106**, 2077 (2008), https://doi.org/10.1080/00268970802172503, URL https://doi.org/10.1080/00268970802172503.

[70] B. F. E. Curchod, F. Agostini, and I. Tavernelli, European Physical Journal B. **91**, 168 (2018), URL http://dro.dur.ac.uk/24991/.

[71] F. Agostini, S. K. Min, A. Abedi, and E. K. U. Gross, Journal of Chemical Theory and Computation **12**, 2127 (2016), https://doi.org/10.1021/acs.jctc.5b01180, URL https://doi.org/10.1021/acs.jctc.5b01180.

[72] J. E. Subotnik, W. Ouyang, and B. R. Landry, The Journal of Chemical Physics **139**, 214107 (2013), https://doi.org/10.1063/1.4829856, URL https://doi.org/10.1063/1.4829856.

[73] R. Crespo-Otero and M. Barbatti, Chemical Reviews **118**, 7026 (2018), https://doi.org/10.1021/acs.chemrev.7b00577, URL https://doi.org/10.1021/acs.chemrev.7b00577.

[74] in *Collected Papers of L.D. Landau*, edited by D. T. HAAR] (Pergamon, 1965), pp. 63 – 66, ISBN 978-0-08-010586-4, URL http://www.sciencedirect.com/science/article/pii/B9780080105864500146.

[75] C. Zener and R. H. Fowler, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **137**, 696 (1932), https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1932.0165, URL https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1932.0165.

[76] A. K. Belyaev and O. V. Lebedev, Phys. Rev. A **84**, 014701 (2011), URL https://link.aps.org/doi/10.1103/PhysRevA.84.014701.

[77] C. Zhu and H. Nakamura, The Journal of Chemical Physics **97**, 8497 (1992), https://doi.org/10.1063/1.463368, URL https://doi.org/10.1063/1.463368.

[78] C. Zhu and H. Nakamura, The Journal of Chemical Physics **98**, 6208 (1993), https://doi.org/10.1063/1.464814, URL https://doi.org/10.1063/1.464814.

[79] L. Yu, C. Xu, Y. Lei, C. Zhu, and Z. Wen, Phys. Chem. Chem. Phys. **16**, 25883 (2014), URL http://dx.doi.org/10.1039/C4CP03498H.

[80] K. Hanasaki, M. Kanno, T. A. Niehaus, and H. Kono, The Journal of Chemical Physics **149**, 244117 (2018), https://doi.org/10.1063/1.5046757, URL https://doi.org/10.1063/1.5046757.

[81] D. Hu, Y. Xie, X. Li, L. Li, and Z. Lan, The Journal of Physical Chemistry Letters **9**, 2725 (2018), https://doi.org/10.1021/acs.jpclett.8b00684, URL https://doi.org/10.1021/acs.jpclett.8b00684.

[82] W.-K. Chen, X.-Y. Liu, W.-H. Fang, P. O. Dral, and G. Cui, The Journal of Physical Chemistry Letters **9**, 6702 (2018), https://doi.org/10.1021/acs.jpclett.8b03026, URL https://doi.org/10.1021/acs.jpclett.8b03026.

[83] J. C. Tully, The Journal of Chemical Physics **93**, 1061 (1990), https://doi.org/10.1063/1.459170, URL https://doi.org/10.1063/1.459170.

[84] S. Hammes-Schiffer and J. C. Tully, The Journal of Chemical Physics **101**, 4657 (1994), https://doi.org/10.1063/1.467455, URL https://doi.org/10.1063/1.467455.

[85] M. E. Casida, Journal of Molecular Structure: THEOCHEM **914**, 3 (2009), ISSN 0166-1280, URL http://www.sciencedirect.com/science/article/pii/S0166128009005363.

[86] T. A. Niehaus, S. Suhai, F. Della Sala, P. Lugli, M. Elstner, G. Seifert, and T. Frauenheim, Phys. Rev. B **63**, 085108 (2001), URL https://link.aps.org/doi/10.1103/PhysRevB.63.085108.

[87] R. Mitrić, U. Werner, M. Wohlgemuth, G. Seifert, and V. Bonačić-Koutecký, The Journal of Physical Chemistry A **113**, 12700 (2009), https://doi.org/10.1021/jp905600w, URL https://doi.org/10.1021/jp905600w.

[88] S. Pal, D. J. Trivedi, A. V. Akimov, B. Aradi, T. Frauenheim, and O. V. Prezhdo, Journal of Chemical Theory and Computation **12**, 1436 (2016), https://doi.org/10.1021/acs.jctc.5b01231, URL https://doi.org/10.1021/acs.jctc.5b01231.

[89] L. Stojanović, S. G. Aziz, R. H. Hilal, F. Plasser, T. A. Niehaus, and M. Barbatti, Journal of Chemical Theory and Computation **13**, 5846 (2017), https://doi.org/10.1021/acs.jctc.7b01000, URL https://doi.org/10.1021/acs.jctc.7b01000.

[90] E. Posenitskiy, M. Rapacioli, B. Lepetit, D. Lemoine, and F. Spiegelman, Phys. Chem. Chem. Phys. **21**, 12139 (2019), URL http://dx.doi.org/10.1039/C9CP00603F.

[91] C. Joblin, F. Salama, and L. Allamandola, The Journal of Chemical Physics **102**, 9743 (1995), https://doi.org/10.1063/1.468793, URL https://doi.org/10.1063/1.468793.

[92] D. A. da Silva Filho, R. Friedlein, V. Coropceanu, G. Öhrwall, W. Osikowicz, C. Suess, S. L. Sorensen, S. Svensson, W. R. Salaneck, and J.-L. Brédas, Chem. Commun. pp. 1702–1703 (2004), URL http://dx.doi.org/10.1039/B403828B.

[93] R. S. Sánchez-Carrera, V. Coropceanu, D. A. da Silva Filho, R. Friedlein, W. Osikowicz, R. Murdey, C. Suess, W. R. Salaneck, and J.-L. Brédas, The Journal of Physical Chemistry B **110**, 18904 (2006), https://doi.org/10.1021/jp057462p, URL https://doi.org/10.1021/jp057462p.

[94] V. L. Ermolaev, Russian Chemical Reviews **70**, 471 (2001), URL https://doi.org/10.1070%2Frc2001v070n06abeh000657.

[95] J. A. Noble, C. Aupetit, D. Descamps, S. Petit, A. Simon, J. Mascetti, N. Ben Amor, and V. Blanchet, Phys. Chem. Chem. Phys. **21**, 14111 (2019), URL http://dx.doi.org/10.1039/C8CP06895J.

[96] V. Blanchet, S. Lochbrunner, M. Schmitt, J. P. Shaffer, J. J. Larsen, M. Z. Zgierski, T. Seideman, and A. Stolow, Faraday Discuss. **115**, 33 (2000), URL http://dx.doi.org/10.1039/B001138J.

[97] V. Blanchet, K. Raffael, G. Turri, B. Chatel, B. Girard, I. A. Garcia, I. Wilkinson, and B. J. Whitaker, The Journal of Chemical Physics **128**, 164318 (2008), https://doi.org/10.1063/1.2913167, URL https://doi.org/10.1063/1.2913167.

[98] M. Vinitha, A. M. Nair, A. S. Kumar, V. Blanchet, and U. R. Kadhane, Chemical Physics Letters **745**, 137250 (2020), ISSN 0009-2614, URL http://www.sciencedirect.com/science/article/pii/S0009261420301652.

[99] T. Heine, M. Rapacioli, S. Patchkovskii, J. Frenzel, A. M. Köster, P. Calaminici, H. A. Duarte, S. Escalante, R. Flores-Moreno, A. Goursot, et al., *demonnano*, 2009, URL http://demon-nano.ups-tlse.fr.

[100] F. Bettanin, L. F. A. Ferrão, M. Pinheiro, A. J. A. Aquino, H. Lischka, F. B. C. Machado, and D. Nachtigallova, Journal of Chemical Theory and Computation **13**, 4297 (2017), https://doi.org/10.1021/acs.jctc.7b00302, URL https://doi.org/10.1021/acs.jctc.7b00302.

[101] F. Talotta, J.-L. Heully, F. Alary, I. M. Dixon, L. González, and M. Boggio-Pasqua, Journal of Chemical Theory and Computation **13**, 6120 (2017), https://doi.org/10.1021/acs.jctc.7b00982, URL https://doi.org/10.1021/acs.jctc.7b00982.

[102] J. W. Park and T. Shiozaki, Journal of Chemical Theory and Computation **13**, 3676 (2017), https://doi.org/10.1021/acs.jctc.7b00559, URL https://doi.org/10.1021/acs.jctc.7b00559.

[103] D. Hollas, L. Šištík, E. G. Hohenstein, T. J. Martínez, and P. Slavíček, Journal of Chemical Theory and Computation **14**, 339 (2018), https://doi.org/10.1021/acs.jctc.7b00958, URL https://doi.org/10.1021/acs.jctc.7b00958.

[104] A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Books on Chemistry (Dover Publications, 1996), ISBN 9780486691862, URL https://books.google.fr/books?id=6mV9gYzEkgIC.

[105] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, Ltd, 2000), ISBN 9781119019572.

[106] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965), URL https://link.aps.org/doi/10.1103/PhysRev.140.A1133.

[107] K. Kim and K. D. Jordan, The Journal of Physical Chemistry **98**, 10089 (1994), https://doi.org/10.1021/j100091a024, URL https://doi.org/10.1021/j100091a024.

[108] K. Kim and K. D. Jordan, The Journal of Physical Chemistry **98**, 10089 (1994), https://doi.org/10.1021/j100091a024, URL https://doi.org/10.1021/j100091a024.

[109] H. Iikura, T. Tsuneda, T. Yanai, and K. Hirao, The Journal of Chemical Physics **115**, 3540 (2001), https://doi.org/10.1063/1.1383587, URL https://doi.org/10.1063/1.1383587.

[110] *Intel Math Kernel Library. Reference Manual* (Intel Corporation, Santa Clara, USA, 2009), iSBN 630813-054US.

[111] S. Tomov, J. Dongarra, and M. Baboulin, Parallel Computing **36**, 232 (2010), ISSN 0167-8191.

[112] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nature Communications **5**, 4213 (2014), https://doi.org/10.1038/ncomms5213, URL https://doi.org/10.1038/ncomms5213.

[113] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature **549**, 242 (2017), https://doi.org/10.1038/nature23879, URL https://doi.org/10.1038/nature23879.

[114] C. M. Goringe, D. R. Bowler, and E. Hernández, Reports on Progress in Physics **60**, 1447 (1997), URL https://doi.org/10.1088%2F0034-4885%2F60%2F12%2F001.

[115] G. Seifert, H. Eschrig, , and W. Bieger, Z. Phys. Chem. **267**, 529 (1986).

[116] F. Spiegelman, N. Tarrat, J. Cuny, L. Dontot, E. Posenitskiy, C. Martí, A. Simon, and M. Rapacioli, Advances in Physics: X **5**, 1710252 (2020), https://doi.org/10.1080/23746149.2019.1710252, URL https://doi.org/10.1080/23746149.2019.1710252.

[117] A. Di Carlo, Physica B: Condensed Matter **314**, 211 (2002), ISSN 0921-4526, proceedings of the Twelfth International Conference on Nonequilibrium Carrier Dynamics in Semiconductors, URL http://www.sciencedirect.com/science/article/pii/S0921452601014491.

[118] M. Elstner and G. Seifert, Phil. Trans. R. Soc. A **372**, 20120483 (2014).

[119] D. Porezag, T. Frauenheim, T. Köhler, G. Seifert, and R. Kaschner, Phys. Rev. B **51**, 12947 (1995).

[120] G. Seifert, D. Porezag, and T. Frauenheim, Int. J. Quantum Chem. **58**, 185 (1996).

[121] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, Phys. Rev. B **58**, 7260 (1998), URL https://link.aps.org/doi/10.1103/PhysRevB.58.7260.

[122] M. Gaus, Q. Cui, and M. Elstner, Journal of Chemical Theory and Computation **7**, 931 (2011), https://doi.org/10.1021/ct100684s, URL https://doi.org/10.1021/ct100684s.

[123] B. Aradi, B. Hourahine, and T. Frauenheim, The Journal of Physical Chemistry A **111**, 5678 (2007), https://doi.org/10.1021/jp070186p, URL https://doi.org/10.1021/jp070186p.

[124] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler, Journal of Computational Chemistry **22**, 931 (2001), https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.1056, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.1056.

[125] R. Salomon-Ferrer, D. A. Case, and R. C. Walker, WIREs Computational Molecular Science **3**, 198 (2013), https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1121, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/wcms.1121.

[126] H. Berendsen, D. van der Spoel, and R. van Drunen, Computer Physics Communications **91**, 43 (1995), ISSN 0010-4655, URL http://www.sciencedirect.com/science/article/pii/001046559500042E.

[127] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, et al., *Gaussian˜16 Revision C.01* (2016), gaussian Inc. Wallingford CT.

[128] A. Humeniuk and R. Mitrić, Computer Physics Communications **221**, 174 (2017), ISSN 0010-4655, URL http://www.sciencedirect.com/science/article/pii/S0010465517302576.

[129] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, WIREs Computational Molecular Science **4**, 15 (2014), https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1159, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/wcms.1159.

[130] R. Pariser, The Journal of Chemical Physics **24**, 250 (1956), https://doi.org/10.1063/1.1742461, URL https://doi.org/10.1063/1.1742461.

[131] J. Frenzel, A. F. Oliveira, N. Jardillier, T. Heine, and G. Seifert, *Semi-relativistic, self-consistent charge Slater-Koster tables for density-functional based tight-binding (DFTB) for materials science simulations* (TU Dresden, 2004-2009).

[132] M. Van den Bossche, H. Grönbeck, and B. Hammer, Journal of Chemical Theory and Computation **14**, 2797 (2018), https://doi.org/10.1021/acs.jctc.8b00039, URL https://doi.org/10.1021/acs.jctc.8b00039.

[133] J. J. Kranz, M. Kubillus, R. Ramakrishnan, O. A. von Lilienfeld, and M. Elstner, Journal of Chemical Theory and Computation **14**, 2341 (2018), https://doi.org/10.1021/acs.jctc.7b00933, URL https://doi.org/10.1021/acs.jctc.7b00933.

[134] F. Jensen, *Introduction to Computational Chemistry* (Wiley-Blackwell, 2011), ISBN 9780470011874.

[135] S. Grimme, Journal of Computational Chemistry **25**, 1463 (2004), https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.20078, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.20078.

[136] M. E. Casida, *Time-Dependent Density Functional Response Theory for Molecules* (1995), pp. 155–192, https://www.worldscientific.com/doi/pdf/10.1142/9789812830586_0005, URL https://www.worldscientific.com/doi/abs/10.1142/9789812830586_0005.

[137] V. Lutsker, B. Aradi, and T. A. Niehaus, The Journal of Chemical Physics **143**, 184107 (2015), https://doi.org/10.1063/1.4935095, URL https://doi.org/10.1063/1.4935095.

[138] A. Humeniuk and R. Mitrić, The Journal of Chemical Physics **143**, 134120 (2015), https://aip.scitation.org/doi/pdf/10.1063/1.4931179, URL https://aip.scitation.org/doi/abs/10.1063/1.4931179.

[139] M. A. Marques, N. T. Maitra, F. M. Nogueira, E. K. Gross, and A. Rubio, *Fundamentals of Time-Dependent Density Functional Theory*, vol. 837 of *Lecture Notes in Physics* (Springer-Verlag Berlin Heidelberg, 2012), ISBN 9783642235177.

[140] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984), URL https://link.aps.org/doi/10.1103/PhysRevLett.52.997.

[141] F. P. Bonafé, F. J. Hernández, B. Aradi, T. Frauenheim, and C. G. Sánchez, The Journal of Physical Chemistry Letters **9**, 4355 (2018), https://doi.org/10.1021/acs.jpclett.8b01659, URL https://doi.org/10.1021/acs.jpclett.8b01659.

[142] S. Hirata and M. Head-Gordon, Chemical Physics Letters **314**, 291 (1999), ISSN 0009-2614, URL http://www.sciencedirect.com/science/article/pii/S0009261499011495.

[143] T. Niehaus, Journal of Molecular Structure: THEOCHEM **914**, 38 (2009), ISSN 0166-1280, time-dependent density-functional theory for molecules and molecular solids, URL http://www.sciencedirect.com/science/article/pii/S0166128009003042.

[144] R. Rüger, T. Niehaus, E. van Lenthe, T. Heine, and L. Visscher, The Journal of Chemical Physics **145**, 184102 (2016), https://doi.org/10.1063/1.4966918, URL https://doi.org/10.1063/1.4966918.

[145] F. Trani, G. Scalmani, G. Zheng, I. Carnimeo, M. J. Frisch, and V. Barone, Journal of Chemical Theory and Computation **7**, 3304 (2011), URL https://doi.org/10.1021/ct200461y.

[146] A. Domínguez, B. Aradi, T. Frauenheim, V. Lutsker, and T. A. Niehaus, Journal of Chemical Theory and Computation **9**, 4901 (2013), https://doi.org/10.1021/ct400123t, URL https://doi.org/10.1021/ct400123t.

[147] E. Tapavicza, I. Tavernelli, U. Rothlisberger, C. Filippi, and M. E. Casida, The Journal of Chemical Physics **129**, 124108 (2008), https://aip.scitation.org/doi/pdf/10.1063/1.2978380, URL https://aip.scitation.org/doi/abs/10.1063/1.2978380.

[148] M. Rapacioli, F. Spiegelman, A. Scemama, and A. Mirtschink, Journal of Chemical Theory and Computation **7**, 44 (2011), URL http://dx.doi.org/10.1021/ct100412f.

[149] L. Dontot, N. Suaud, M. Rapacioli, and F. Spiegelman, Phys. Chem. Chem. Phys. **18**, 3545 (2016), URL http://dx.doi.org/10.1039/C5CP06344B.

[150] F. Furche and R. Ahlrichs, The Journal of Chemical Physics **117**, 7433 (2002), https://doi.org/10.1063/1.1508368, URL https://doi.org/10.1063/1.1508368.

[151] F. Furche and R. Ahlrichs, The Journal of Chemical Physics **121**, 12772 (2004), https://aip.scitation.org/doi/pdf/10.1063/1.1824903, URL https://aip.scitation.org/doi/abs/10.1063/1.1824903.

[152] D. Heringer, T. A. Niehaus, M. Wanko, and T. Frauenheim, Journal of Computational Chemistry **28**, 2589 (2007), https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.20697, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.20697.

[153] D. Heringer, T. A. Niehaus, M. Wanko, and T. Frauenheim, Journal of Computational Chemistry **33**, 593 (2011), https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.22916, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.22916.

[154] M. Pastore, X. Assfeld, E. Mosconi, A. Monari, and T. Etienne, The Journal of Chemical Physics **147**, 024108 (2017), https://doi.org/10.1063/1.4991561, URL https://doi.org/10.1063/1.4991561.

[155] G. Granucci and M. Persico, The Journal of Chemical Physics **126**, 134114 (2007), https://doi.org/10.1063/1.2715585, URL https://doi.org/10.1063/1.2715585.

[156] C. Zhu, S. Nangia, A. W. Jasper, and D. G. Truhlar, The Journal of Chemical Physics **121**, 7658 (2004), https://aip.scitation.org/doi/pdf/10.1063/1.1793991, URL https://aip.scitation.org/doi/abs/10.1063/1.1793991.

[157] P. G. Lisinetskaya and R. Mitrić, Phys. Rev. A **83**, 033408 (2011), URL https://link.aps.org/doi/10.1103/PhysRevA.83.033408.

[158] E. Tapavicza, I. Tavernelli, and U. Rothlisberger, Phys. Rev. Lett. **98**, 023001 (2007), URL https://link.aps.org/doi/10.1103/PhysRevLett.98.023001.

[159] F. Plasser, G. Granucci, J. Pittner, M. Barbatti, M. Persico, and H. Lischka, The Journal of Chemical Physics **137**, 22A514 (2012), https://doi.org/10.1063/1.4738960, URL https://doi.org/10.1063/1.4738960.

[160] R. Mitrić, Private Communication (2019).

[161] J. Suchan, D. Hollas, B. F. E. Curchod, and P. Slavíček, Faraday Discuss. **212**, 307 (2018), URL http://dx.doi.org/10.1039/C8FD00088C.

[162] D. Bonhommeau, A. Viel, and N. Halberstadt, The Journal of Chemical Physics **123**, 054316 (2005), https://doi.org/10.1063/1.1953530, URL https://doi.org/10.1063/1.1953530.

[163] F. Plasser, M. Ruckenbauer, S. Mai, M. Oppel, P. Marquetand, and L. González, Journal of Chemical Theory and Computation **12**, 1207 (2016), https://doi.org/10.1021/acs.jctc.5b01148, URL https://doi.org/10.1021/acs.jctc.5b01148.

[164] U. Werner, R. Mitrić, T. Suzuki, and V. Bonačić-Koutecký, Chemical Physics **349**, 319 (2008), ISSN 0301-0104, URL http://www.sciencedirect.com/science/article/pii/S030101040800133X.

[165] E. Titov, A. Humeniuk, and R. Mitrić, Phys. Chem. Chem. Phys. **20**, 25995 (2018), URL http://dx.doi.org/10.1039/C8CP05240A.

[166] J. Hoche, H.-C. Schmitt, A. Humeniuk, I. Fischer, R. Mitrić, and M. I. S. Röhr, Phys. Chem. Chem. Phys. **19**, 25002 (2017), URL http://dx.doi.org/10.1039/C7CP03990E.

[167] A. A. M. H. M. Darghouth, G. C. Correa, S. Juillard, M. E. Casida, A. Humeniuk, and R. Mitrić, The Journal of Chemical Physics **149**, 134111 (2018), https://doi.org/10.1063/1.5025624, URL https://doi.org/10.1063/1.5025624.

[168] E. Posenitskiy, M. Rapacioli, D. Lemoine, and F. Spiegelman, The Journal of Chemical Physics **152**, 074306 (2020), https://doi.org/10.1063/1.5135369, URL https://doi.org/10.1063/1.5135369.

[169] Y. Yang, E. R. Davidson, and W. Yang, Proceedings of the National Academy of Sciences **113**, E5098 (2016), ISSN 0027-8424, http://www.pnas.org/content/113/35/E5098.full.pdf, URL http://www.pnas.org/content/113/35/E5098.

[170] M. L. M. Rocco, M. Häming, C. E. V. de Moura, M. Barbatti, A. B. Rocha, A. Schöll, and E. Umbach, The Journal of Physical Chemistry C **122**, 28692 (2018), URL https://doi.org/10.1021/acs.jpcc.8b08945.

[171] Omont, A., Bettinger, H. F., and Tönshoff, C., A&A **625**, A41 (2019), URL https://doi.org/10.1051/0004-6361/201834953.

[172] M. Kasha, Discuss. Faraday Soc. **9**, 14 (1950), URL http://dx.doi.org/10.1039/DF9500900014.

[173] M. Chergui, Pure and Applied Chemistry **87**, 12. 525 (2015), URL http://infoscience.epfl.ch/record/212527.

[174] C. Joblin, G. Wenzel, S. Rodriguez Castillo, A. Simon, H. Sabbah, A. Bonnamy, D. Toublanc, G. Mulas, M. Ji, A. Giuliani, et al., Journal of Physics: Conference Series (2019), URL https://hal.archives-ouvertes.fr/hal-02396783.

[175] B. West, F. Useli-Bacchitta, H. Sabbah, V. Blanchet, A. Bodi, P. M. Mayer, and C. Joblin, The Journal of Physical Chemistry A **118**, 7824 (2014), https://doi.org/10.1021/jp506420u, URL https://doi.org/10.1021/jp506420u.

[176] J. Zhen, S. R. Castillo, C. Joblin, G. Mulas, H. Sabbah, A. Giuliani, L. Nahon, S. Martin, J.-P. Champeaux, and P. M. Mayer, The Astrophysical Journal **822**, 113 (2016), URL https://doi.org/10.3847%2F0004-637x%2F822%2F2%2F113.

[177] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al., Gaussian 09 (Revision D.01), Gaussian Inc., Wallingford, CT, 2009.

[178] W. Humphrey, A. Dalke, and K. Schulten, Journal of Molecular Graphics **14**, 33 (1996).

[179] A. Domínguez, B. Aradi, T. Frauenheim, V. Lutsker, and T. A. Niehaus, Journal of Chemical Theory and Computation **9**, 4901 (2013), https://doi.org/10.1021/ct400123t, URL https://doi.org/10.1021/ct400123t.

[180] M. Rubio, M. Merchán, E. Ortí, and B. O. Roos, Chemical Physics **179**, 395 (1994), ISSN 0301-0104, URL http://www.sciencedirect.com/science/article/pii/0301010494870160.

[181] H. H. Heinze, A. Görling, and N. Rösch, The Journal of Chemical Physics **113**, 2088 (2000), https://doi.org/10.1063/1.482020, URL https://doi.org/10.1063/1.482020.

[182] R. Mondal, C. Tönshoff, D. Khon, D. C. Neckers, and H. F. Bettinger, Journal of the American Chemical Society **131**, 14281 (2009), https://doi.org/10.1021/ja901841c, URL https://doi.org/10.1021/ja901841c.

[183] C. Tönshoff and H. F. Bettinger, Angewandte Chemie International Edition **49**, 4125 (2010), URL https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.200906355.

[184] E. Clar and W. Schmidt, Tetrahedron **31**, 2263 (1975), ISSN 0040-4020, URL http://www.sciencedirect.com/science/article/pii/0040402075802250.

[185] W. Schmidt, The Journal of Chemical Physics **66**, 828 (1977), https://doi.org/10.1063/1.433961, URL https://doi.org/10.1063/1.433961.

[186] E. Holiday and E. Jope, Spectrochimica Acta **4**, 157 (1950), ISSN 0371-1951, URL http://www.sciencedirect.com/science/article/pii/S0371195150800085.

[187] M. M. Husain, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **68**, 156 (2007), ISSN 1386-1425, URL http://www.sciencedirect.com/science/article/pii/S1386142506006743.

[188] O. Khorev, C. D. Bösch, M. Probst, and R. Häner, Chem. Sci. **5**, 1506 (2014), URL http://dx.doi.org/10.1039/C3SC53316F.

[189] S. Knippenberg, J. H. Starcke, M. Wormit, and A. Dreuw, Molecular Physics **108**, 2801 (2010), https://doi.org/10.1080/00268976.2010.526643, URL https://doi.org/10.1080/00268976.2010.526643.

[190] B. Stevens, M. F. Thomaz, and J. Jones, The Journal of Chemical Physics **46**, 405 (1967), https://doi.org/10.1063/1.1840417, URL https://doi.org/10.1063/1.1840417.

[191] V. Lawetz, G. Orlandi, and W. Siebrand, The Journal of Chemical Physics **56**, 4058 (1972), https://doi.org/10.1063/1.1677816, URL https://doi.org/10.1063/1.1677816.

[192] H. Morrison and A. Miller, Tetrahedron **37**, 3405 (1981), ISSN 0040-4020, URL http://www.sciencedirect.com/science/article/pii/S0040402001923890.

[193] K. Schmidt, S. Brovelli, V. Coropceanu, D. Beljonne, J. Cornil, C. Bazzini, T. Caronna, R. Tubino, F. Meinardi, Z. Shuai, et al., The Journal of Physical Chemistry A **111**, 10490 (2007), https://doi.org/10.1021/jp075248q, URL https://doi.org/10.1021/jp075248q.

[194] S. Mai, P. Marquetand, and L. González, WIREs Computational Molecular Science **8**, e1370 (2018), https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1370, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/wcms.1370.

[195] R. Fletcher, *Practical methods of optimization*, no. vol. 1 in Wiley-Interscience publication (Wiley, 1987), ISBN 9780471915478, URL https://books.google.fr/books?id=3EzvAAAAMAAJ.

[196] L. P. Kaelbling, M. L. Littman, and A. W. Moore, CoRR **cs.AI/9605103** (1996), URL https://arxiv.org/abs/cs/9605103.

[197] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al., Nature **529**, 484 (2016), URL http://www.nature.com/nature/journal/v529/n7587/full/nature16961.html.

[198] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, et al., Science **362**, 1140 (2018), ISSN 0036-8075, https://science.sciencemag.org/content/362/6419/1140.full.pdf, URL https://science.sciencemag.org/content/362/6419/1140.

[199] I. Witten, E. Frank, M. Hall, and C. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, The Morgan Kaufmann Series in Data Management Systems (Elsevier Science, 2016), ISBN 9780128043578, URL https://books.google.fr/books?id=1SylCgAAQBAJ.

[200] C. M. Bishop, *Neural Networks for Pattern Recognition* (Oxford University Press, Inc., USA, 1995), ISBN 0198538642.

[201] S. S. Haykin, *Neural networks and learning machines* (New York: Prentice Hall, 2009), ISBN 0131471392.

[202] W. S. McCulloch and W. Pitts, The bulletin of mathematical biophysics **5**, 115 (1943).

[203] J. Gasteiger and J. Zupan, Angewandte Chemie International Edition in English **32**, 503 (1993).

[204] F. Scarselli and A. C. Tsoi], Neural Networks **11**, 15 (1998), ISSN 0893-6080, URL http://www.sciencedirect.com/science/article/pii/S089360809700097X.

[205] J.-G. Attali and G. Pagès, Neural networks **10**, 1069 (1997).

[206] G. Cybenko, Mathematics of control, signals and systems **2**, 303 (1989).

[207] K. Hornik, M. Stinchcombe, and H. White, Neural networks **2**, 359 (1989).

[208] T. Hofmann, B. Schölkopf, and A. J. Smola, The annals of statistics pp. 1171–1220 (2008).

[209] P. O. Dral, M. Barbatti, and W. Thiel, The Journal of Physical Chemistry Letters **9**, 5660 (2018), https://doi.org/10.1021/acs.jpclett.8b02469, URL https://doi.org/10.1021/acs.jpclett.8b02469.

[210] J. Westermayr, M. Gastegger, and P. Marquetand, The Journal of Physical Chemistry Letters **11**, 3828 (0), https://doi.org/10.1021/acs.jpclett.0c00527, URL https://doi.org/10.1021/acs.jpclett.0c00527.

[211] J. Westermayr, M. Gastegger, M. F. S. J. Menger, S. Mai, L. González, and P. Marquetand, Chem. Sci. **10**, 8100 (2019), URL http://dx.doi.org/10.1039/C9SC01742A.

[212] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007), URL https://link.aps.org/doi/10.1103/PhysRevLett.98.146401.

[213] J. Behler, R. Martoňák, D. Donadio, and M. Parrinello, Phys. Rev. Lett. **100**, 185501 (2008), URL https://link.aps.org/doi/10.1103/PhysRevLett.100.185501.

[214] M. Gastegger, J. Behler, and P. Marquetand, Chem. Sci. **8**, 6924 (2017), URL http://dx.doi.org/10.1039/C7SC02267K.

[215] K. Shakouri, J. Behler, J. Meyer, and G.-J. Kroes, The Journal of Physical Chemistry Letters **8**, 2131 (2017), https://doi.org/10.1021/acs.jpclett.7b00784, URL https://doi.org/10.1021/acs.jpclett.7b00784.

[216] Y. Zhang, X. Zhou, and B. Jiang, The Journal of Physical Chemistry Letters **10**, 1185 (2019), https://doi.org/10.1021/acs.jpclett.9b00085, URL https://doi.org/10.1021/acs.jpclett.9b00085.

[217] J. Behler, International Journal of Quantum Chemistry **115**, 1032 (2015), https://onlinelibrary.wiley.com/doi/pdf/10.1002/qua.24890, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.24890.

[218] S. Chiriki, S. Jindal, and S. S. Bulusu, The Journal of Chemical Physics **146**, 084314 (2017), https://doi.org/10.1063/1.4977050, URL https://doi.org/10.1063/1.4977050.

[219] Z. Li, J. R. Kermode, and A. De Vita, Phys. Rev. Lett. **114**, 096405 (2015), URL https://link.aps.org/doi/10.1103/PhysRevLett.114.096405.

[220] L. Zhang, J. Han, H. Wang, R. Car, and W. E, Phys. Rev. Lett. **120**, 143001 (2018), URL https://link.aps.org/doi/10.1103/PhysRevLett.120.143001.

[221] K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, The Journal of Chemical Physics **148**, 241722 (2018), https://doi.org/10.1063/1.5019779, URL https://doi.org/10.1063/1.5019779.

[222] R. Ramakrishnan, M. Hartmann, E. Tapavicza, and O. A. von Lilienfeld, The Journal of Chemical Physics **143**, 084111 (2015), https://doi.org/10.1063/1.4928757, URL https://doi.org/10.1063/1.4928757.

[223] J. Zhu, V. Q. Vuong, B. G. Sumpter, and S. Irle, MRS Communications **9**, 867–873 (2019).

[224] R. Ramakrishnan, P. O. Dral, M. Rupp, and O. A. von Lilienfeld, Journal of Chemical Theory and Computation **11**, 2087 (2015), https://doi.org/10.1021/acs.jctc.5b00099, URL https://doi.org/10.1021/acs.jctc.5b00099.

[225] P. O. Dral, Journal of Computational Chemistry **40**, 2339 (2019), https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.26004, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.26004.

[226] H. Wang, L. Zhang, J. Han, and W. E, Computer Physics Communications **228**, 178 (2018), ISSN 0010-4655, URL http://www.sciencedirect.com/science/article/pii/S0010465518300882.

[227] K. T. Schütt, P. Kessel, M. Gastegger, K. A. Nicoli, A. Tkatchenko, and K.-R. Müller, Journal of Chemical Theory and Computation **15**, 448 (2019), https://doi.org/10.1021/acs.jctc.8b00908, URL https://doi.org/10.1021/acs.jctc.8b00908.

[228] G. W. Richings and S. Habershon, Journal of Chemical Theory and Computation **13**, 4012 (2017), https://doi.org/10.1021/acs.jctc.7b00507, URL https://doi.org/10.1021/acs.jctc.7b00507.

[229] T. L. Jacobsen, M. S. Jørgensen, and B. Hammer, Phys. Rev. Lett. **120**, 026102 (2018), URL https://link.aps.org/doi/10.1103/PhysRevLett.120.026102.

[230] M. K. Bisbo and B. Hammer, Phys. Rev. Lett. **124**, 086102 (2020), URL https://link.aps.org/doi/10.1103/PhysRevLett.124.086102.

[231] E. L. Kolsbjerg, A. A. Peterson, and B. Hammer, Phys. Rev. B **97**, 195424 (2018), URL https://link.aps.org/doi/10.1103/PhysRevB.97.195424.

[232] K. H. Sørensen, M. S. Jørgensen, A. Bruix, and B. Hammer, The Journal of Chemical Physics **148**, 241734 (2018), https://doi.org/10.1063/1.5023671, URL https://doi.org/10.1063/1.5023671.

[233] M. S. Jørgensen, M. N. Groves, and B. Hammer, Journal of Chemical Theory and Computation **13**, 1486 (2017), https://doi.org/10.1021/acs.jctc.6b01119, URL https://doi.org/10.1021/acs.jctc.6b01119.

[234] M. S. Jørgensen, H. L. Mortensen, S. A. Meldgaard, E. L. Kolsbjerg, T. L. Jacobsen, K. H. Sørensen, and B. Hammer, The Journal of Chemical Physics **151**, 054111 (2019), https://doi.org/10.1063/1.5108871, URL https://doi.org/10.1063/1.5108871.

[235] R. V. Krems, Phys. Chem. Chem. Phys. **21**, 13392 (2019), URL http://dx.doi.org/10.1039/C9CP01883B.

[236] V. L. Deringer, M. A. Caro, and G. Csányi, Advanced Materials **31**, 1902765 (2019), https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201902765, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201902765.

[237] M. Ceriotti, The Journal of Chemical Physics **150**, 150901 (2019), https://doi.org/10.1063/1.5091842, URL https://doi.org/10.1063/1.5091842.

[238] P. Brechignac and P. Hermine, AIP Conference Proceedings **312**, 613 (1994), https://aip.scitation.org/doi/pdf/10.1063/1.46586, URL https://aip.scitation.org/doi/abs/10.1063/1.46586.

[239] A. Pukrittayakamee, M. Malshe, M. Hagan, L. M. Raff, R. Narulkar, S. Bukkapatnum, and R. Komanduri, The Journal of Chemical Physics **130**, 134101 (2009), https://doi.org/10.1063/1.3095491, URL https://doi.org/10.1063/1.3095491.

[240] K. Schütt, P.-J. Kindermans, H. E. Sauceda Felix, S. Chmiela, A. Tkatchenko, and K.-R. Müller, in *Advances in Neural Information Processing Systems 30*, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc., 2017), pp. 991–1001.

[241] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, et al., Journal of Physics: Condensed Matter **29**, 273002 (2017), URL https://doi.org/10.1088%2F1361-648x%2Faa680e.

[242] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, et al., in *Advances in Neural Information Processing Systems 32*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019), pp. 8024–8035, URL http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

[243] M. Wanko, M. Garavelli, F. Bernardi, T. A. Niehaus, T. Frauenheim, and M. Elstner, The Journal of Chemical Physics **120**, 1674 (2004), https://doi.org/10.1063/1.1635798, URL https://doi.org/10.1063/1.1635798.

[244] A. Fihey and D. Jacquemin, Journal of Chemical Theory and Computation **15**, 6267 (2019), https://doi.org/10.1021/acs.jctc.9b00688, URL https://doi.org/10.1021/acs.jctc.9b00688.

[245] B. Smith and A. V. Akimov, The Journal of Physical Chemistry Letters **11**, 1456 (2020), https://doi.org/10.1021/acs.jpclett.9b03687, URL https://doi.org/10.1021/acs.jpclett.9b03687.

[246] A. Amirav, M. Sonnenschein, and J. Jortner, The Journal of Physical Chemistry **88**, 5593 (1984), https://doi.org/10.1021/j150667a027, URL https://doi.org/10.1021/j150667a027.

[247] N. Ohta and H. Baba, Molecular Physics **59**, 921 (1986), https://doi.org/10.1080/00268978600102481, URL https://doi.org/10.1080/00268978600102481.

[248] B. Dick and B. Nickel, Chemical Physics **110**, 131 (1986), ISSN 0301-0104, URL http://www.sciencedirect.com/science/article/pii/0301010486851515.

[249] F. Salama, C. Joblin, and L. J. Allamandola, The Journal of Chemical Physics **101**, 10252 (1994), https://doi.org/10.1063/1.467905, URL https://doi.org/10.1063/1.467905.

[250] F. Huisken, G. Rouillé, Y. Carpentier, M. Steglich, and T. Henning, AIP Conference Proceedings **1333**, 819 (2011), https://aip.scitation.org/doi/pdf/10.1063/1.3562747, URL https://aip.scitation.org/doi/abs/10.1063/1.3562747.

[251] R. González-Luque, L. Serrano-Andrés, M. Merchán, and M. P. Fülscher, Theoretical Chemistry Accounts **110**, 224 (2003), URL https://doi.org/10.1007/s00214-003-0470-2.

[252] M. Nazari, C. D. Bösch, A. Rondi, A. Francés-Monerris, M. Marazzi, E. Lognon, M. Gazzetto, S. M. Langenegger, R. Häner, T. Feurer, et al., Phys. Chem. Chem. Phys. **21**, 16981 (2019), URL http://dx.doi.org/10.1039/C9CP03147B.

[253] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, and K.-R. Müller, Science Advances **3** (2017), https://advances.sciencemag.org/content/3/5/e1603015.full.pdf, URL https://advances.sciencemag.org/content/3/5/e1603015.

[254] J.-L. Reymond, Accounts of Chemical Research **48**, 722 (2015), https://doi.org/10.1021/ar500432k, URL https://doi.org/10.1021/ar500432k.

[255] L. C. Blum and J.-L. Reymond, Journal of the American Chemical Society **131**, 8732 (2009), https://doi.org/10.1021/ja902302h, URL https://doi.org/10.1021/ja902302h.

[256] R. Ramakrishnan, P. Dral, M. Rupp, and A. von Lilienfeld, Scientific Data **1**, 140022 (2014).

[257] K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller, and A. Tkatchenko, Nature Communications **8**, 13890 (2017), https://doi.org/10.1038/ncomms13890, URL https://doi.org/10.1038/ncomms13890.

[258] W. Xie and W. Domcke, The Journal of Chemical Physics **147**, 184114 (2017), https://doi.org/10.1063/1.5006788, URL https://doi.org/10.1063/1.5006788.

[259] J. J. Kranz, M. Elstner, B. Aradi, T. Frauenheim, V. Lutsker, A. D. Garcia, and T. A. Niehaus, Journal of Chemical Theory and Computation **13**, 1737 (2017), https://doi.org/10.1021/acs.jctc.6b01243, URL https://doi.org/10.1021/acs.jctc.6b01243.

[260] J. E. Subotnik, A. Jain, B. Landry, A. Petit, W. Ouyang, and N. Bellonzi, Annual Review of Physical Chemistry **67**, 387 (2016), https://doi.org/10.1146/annurev-physchem-040215-112245, URL https://doi.org/10.1146/annurev-physchem-040215-112245.

[261] F. Plasser, S. Mai, M. Fumanal, E. Gindensperger, C. Daniel, and L. González, Journal of Chemical Theory and Computation **15**, 5031 (2019), https://doi.org/10.1021/acs.jctc.9b00525, URL https://doi.org/10.1021/acs.jctc.9b00525.

[262] X. Gao, S. Bai, D. Fazzi, T. Niehaus, M. Barbatti, and W. Thiel, Journal of Chemical Theory and Computation **13**, 515 (2017), URL https://doi.org/10.1021/acs.jctc.6b00915.

[263] J. P. Gardner, J. C. Mather, M. Clampin, R. Doyon, M. A. Greenhouse, H. B. Hammel, J. B. Hutchings, P. Jakobsen, S. J. Lilly, K. S. Long, et al., Space Science Reviews **123**, 485 (2006), URL https://doi.org/10.1007/s11214-006-8315-7.

List of Publications

2019	**Non-adiabatic molecular dynamics investigation of the size dependence of the electronic relaxation in polyacenes**. Evgeny Posenitskiy, Mathias Rapacioli, Bruno Lepetit, Didier Lemoine and Fernand Spiegelman. *Phys. Chem. Chem. Phys.* **21,** pp. 12139–12149; DOI: 10.1039/C9CP00603F
2020	**Density-functional tight-binding: basic concepts and applications to molecules and clusters**. Fernand Spiegelman, Nathalie Tarrat, Jérôme Cuny, Leo Dontot, Evgeny Posenitskiy, Carles Martí, Aude Simon and Mathias Rapacioli. *Advances in Physics: X* **5(1),** p. 1710252; DOI: 10.1080/23746149.2019.1710252
2020	**Theoretical investigation of the electronic relaxation in highly excited chrysene and tetracene: The effect of armchair vs zigzag edge**. Evgeny Posenitskiy, Mathias Rapacioli, Didier Lemoine and Fernand Spiegelman. *The Journal of Chemical Physics* **152(7),** p. 074306; DOI: 10.1063/1.5135369

List of Figures

1.1 Sketch of the various stages in the life cycle of stars and matter in the galaxy. Reproduced from Kulesa *et al.* (2012). Opportunities for Terahertz Facilities on the High Plateau. Proceedings of the International Astronomical Union, 8(S288), 256-263 with permission from Cambridge University Press. 2

1.2 Possible decay pathways for a photoinduced process according to Ref. [4] 3

1.3 Simplified Jablonski diagram representing different relaxation pathways following absorption of a UV photon. Intersystem crossing (ISC) is a radiationless transition between singlet and triplet states. 4

1.4 The ISO-SWS spectra of the nucleus of the galaxy NGC 4536, planetary nebula NGC 7027 and the photo-dissociation region at the Orion bar. Aromatic modes of the major PAH vibrations are indicated on the top axis. Reproduced from Els Peeters (2011). The PAH Hypothesis after 25 Years. Proceedings of the International Astronomical Union, 7(S280), 149-161 with permission from Cambridge University Press. 5

1.5 Several PAH molecules. Left and right columns correspond to cata- and peri-condensed morphologies, respectively. 6

1.6 Schematic potential energy profiles for relaxation pathways through (a) a conical intersection and (b) an avoided crossing. Adapted from the HDR thesis of Dr. Martial Boggio-Pasqua.[60] 9

1.7 Schematic relation between different approaches to non-adiabatic molecular dynamics. EF stands for Exact Factorization, GWP for Gaussian WavePacket and CT-MQC for Coupled-Trajectory Mixed Quantum-Classical approach. Adapted with permission from Rachel Crespo-Otero and Mario Barbatti (2018). Recent Advances and Perspectives on Nonadiabatic Mixed Quantum-Classical Dynamics. Chemical Reviews 118(15), 7026-7068). Copyright (2020) American Chemical Society. 10

2.1 Schematic hierarchy of computational methods in Quantum Chemistry. WF stands for the wavefunction, PM is the parametric method, MD and MM stand for the molecular dynamics and molecular mechanics, respectively. 17

2.2 Schematic representation of some FSSH-trajectory scenarios following vertical transition from the ground state to an excited singlet state: A) system remains in a given excited state and eventually emits a photon with an excess of energy $\hbar\hat{\nu}$ (photoemission); B) transition between two states occurs in the avoided crossing region (or at the conical intersection) leading to the relaxation towards an equilibrium ground state configuration (radiationless decay); C) same as previous case but the system ends up in a different local minimum of the ground state PES (photoisomerization). 28

2.3 Relative timescales corresponding to the nuclear and the electronic propagation within the FSSH approach. 29

3.1 Structure of polyacenes in one Kekule resonance form from Ref. [169] 36

3.2 The balls and sticks representation of the tetracene (left panel) and chrysene (right panel) isomers. Black and cyan spheres are carbon and hydrogen atoms, respectively. 36

3.3 63 initial configurations of tetracene sampled from the thermal distribution at $T = 300$ K. Black color denotes carbon atoms, red color denotes hydrogen atoms. These figures have been produced with VMD.[178] 38

3.4 Absorption spectra of neutral naphthalene (top row), anthracene (middle row) and tetracene (bottom row) computed with TD-DFT (orange line) and TD-DFTB (blue line) at the equilibrium geometry. All spectra have been truncated at 200 nm. 39

3.5 Absorption spectra of neutral pentacene (top row), hexacene (middle row) and heptacene (bottom row) computed with TD-DFT (orange line) and TD-DFTB (blue line) at the equilibrium geometry. Spectra have been truncated at 250 nm for pentacene and hexacene; at 300 nm for heptacene. 40

3.6 Energies of the brightest and first singlet excited states (ES) in the theoretical CASPT2[100,180] (magenta diamonds, dotted line), TD-DFT (red squares, dash-dotted line), TD-DFTB (green circles, solid line) and experimental[181–183] (blue triangles, dashed line) absorption spectra of polyacenes (n is the number of aromatic cycles in the compound). The energy marks are connected by lines to guide the eye. All plotted values can be found in Table S1 (see Appendix B). . 42

3.7 Absorption spectra (oscillator strengths) of chrysene computed with TD-DFT (orange sticks) and TD-DFTB (blue sticks) using MIO (left panel) and MAT parameters set (right panel) at the equilibrium geometry. Red vertical dashed line at 270 nm indicates the brightest broad band in chrysene (from the experiment in a hexane solution[186]). 43

3.8 Isosurfaces of (a) HOMO and (c) LUMO+2 of tetracene (left column); (b) HOMO–1 and (d) LUMO+1 of chrysene (right column). Black and cyan spheres are carbon and hydrogen atoms, respectively. This figure has been produced with VMD.[178] 44

3.9 Oscillator strengths (left panel) and molar absorption coefficients ε (right panel, convoluted) from the UV absorption spectra of chrysene and tetracene computed with TD-DFTB (MAT parameters set) at the equilibrium geometry. The vertical dashed lines indicate positions of the brightest bands for the corresponding molecule from the experiments.[181,186] 44

3.10 Population analysis of the first 8 singlet excited states in anthracene for an ensemble of 127 (left panel) and 63 (right panel) trajectories. The initial state is S_7. The blue curves hereafter represent the sums of averaged populations in the lower-lying singlet states. For instance, the S_{1-4} curve is $|C_1|^2 + |C_2|^2 + |C_3|^2 + |C_4|^2$. 45

3.11 Population analysis of the first 11 and 12 singlet excited states in pentacene and heptacene, respectively, for an ensemble of 63 trajectories. The initial state is indicated between parentheses. 46

3.12 Population analysis of the first 5, 8 and 11 singlet excited states in naphthalene, tetracene and hexacene, respectively, for an ensemble of 127 trajectories (naphthalene) and 63 trajectories (tetracene, hexacene). The initial state is indicated between parentheses. 46

3.13 Electronic potential energy (red color, upper panel) and total nuclear kinetic energy (green color, bottom panel) along a swarm of 127 trajectories in naphthalene and anthracene. The white dashed line is the average value for each panel. 47

3.14 Decay time of the brightest (B) singlet excited state (red hexagons) and the energy gap (blue diamonds) between the B and B–1 (the one with lower energy in the TD-DFTB absorption spectra) states computed at the equilibrium geometry as a function of number of aromatic cycles in the polyacene. 48

3.15 Energy gap between the B and B–1 states plotted along each trajectory (red lines) and averaged over the ensemble of trajectories (black dashed line) for a set of polyacenes. The horizontal black line at 0.01 eV is the CoIn threshold. 49

3.16 Energy gaps between the B and B–1 singlet excited states at the equilibrium geometry of polyacenes before (red circles) and after (blue diamonds) shuffling induced by the TD-DFTB coupling matrix. 50

3.17 Population analysis of the first 5 singlet excited states in naphthalene for an ensemble of 63 trajectories with (left panel) and without (right panel) the decoherence correction. The initial (brightest) state is S_4. 51

3.18 Population analysis of the first 8 singlet excited states in tetracene for an ensemble of 100 trajectories. The initial state is S_7. 52

3.19 Population analysis of the first 9 singlet excited states in chrysene for an ensemble of a) 100 trajectories propagated with MAT parameters; b) 89 trajectories propagared with MIO parameters. The initial (brightest) state is S_8 in both cases. 53
3.20 Energy gaps between the B and B–1 states plotted along 100 trajectories (red lines) and averaged over the entire ensemble (black dashed line) for tetracene and chrysene. 54
3.21 Occupation of each excited singlet state in tetracene and chrysene after 300 fs of FSSH propagation. 55
3.22 100 configurations of chrysene (left panel) and tetracene (right panel) at the end of the FSSH simulations. The maximum and mean displacements (Δ) along the z axis are indicated with respect to the DFTB ground state equilibrium geometry. Black and red colors denote carbon and hydrogen atoms, respectively. These figures have been produced with VMD.[178] 55

4.1 Schematic overview of some modern ML methodologies. 60
4.2 Schematic representation of the reinforcement learning approach. 61
4.3 Schematic representation of a simple feed-forward NN with two hidden layers. The nodes (circles) are arranged in layers and the major goal is to establish a functional relation between a vector $\mathbf{x}$ of input parameters (features) and an output value y. 63
4.4 Schematic workflow of a supervised ML with the 3-fold cross-validation (CV). . . 65
4.5 Schematic illustration of the learning curves as a function of model hyperparameters in case of a linear regression with polynomial features (left panel, d is the polynomial degree) or with the regularization parameter λ (right panel). 66
4.6 Schematic illustration of the learning curves as a function of the training set size in case of the high bias (left panel) or high variance (right panel). 66
4.7 Balls and sticks representation of phenanthrene ($C_{14}H_{10}$) molecule. Black and cyan spheres are carbon and hydrogen atoms, respectively. 69
4.8 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[249] and in the supersonic jet[246,247], respectively (see Table 4.3). 74
4.9 Population analysis of the first 4 singlet excited states in phenanthrene for an ensemble of 100 (left panel) and 250 (right panel) trajectories propagated with BIO parameters. The grey dashed line corresponds to an exponential decay with $\tau = 522$ fs extracted from the experiments of Blanchet *et al.*[96] The initial state is S_3 in both cases. 77

4.10 Population analysis of the first 3 singlet excited states in phenanthrene for an ensemble of 100 (left panel) and 200 (right panel) trajectories propagated with BIO parameters. The grey dashed line corresponds to an exponential decay with $\tau = 522$ fs extracted from the experiments of Blanchet *et al.*[96] The initial state is S_3 in both cases. 78

4.11 Population analysis of the first 4 singlet excited states in phenanthrene for an ensemble of 100 trajectories propagated without the decoherence correction. The grey dashed line corresponds to $\tau = 522$ fs extracted from the experiments of Blanchet *et al.*[96] 79

4.12 Occupations of the S_2, S_3 and S_4 excited states in phenanthrene averaged over an ensemble of 100 trajectories. Each trajectory has been propagated with an indicated time step Δt on machine-learned PESs with the Zhu-Nakamura hopping probability. 80

4.13 Occupations of the S_2, S_3 and S_4 excited states in phenanthrene averaged over an ensemble of 100 (left panel) and 500 (right panel) trajectories. Each trajectory has been propagated with the time step $\Delta t = 0.25$ fs on machine-learned PESs with the Zhu-Nakamura hopping probability. 80

4.14 Occupations of the S_2, S_3 and S_4 excited states in phenanthrene averaged over an ensemble of 100 trajectories. Each trajectory has been propagated with an indicated time step Δt on machine-learned PESs with the Belyaev-Lebedev hopping probability. Black dashed line is an exponential fit. 81

4.15 Occupations of the S_2, S_3 and S_4 excited states in phenanthrene averaged over an ensemble 500 trajectories. Each trajectory has been propagated with the time step $\Delta t = 0.5$ fs (left panel) or $\Delta t = 0.25$ fs (left panel) on machine-learned PESs with the Belyaev-Lebedev hopping probability. Black dashed line is an exponential fit. 82

4.16 Occupations of the S_2, S_3 and S_4 excited states in phenanthrene averaged over an ensemble 100 trajectories. Each trajectory has been propagated with the time step $\Delta t = 0.25$ fs using the TD-DFTB energies and forces coupled to the Zhu-Nakamura (left panel) or Belyaev-Lebedev (right panel) hopping approaches. 83

4.17 Time required to perform N=1000[fs]/Δt propagation steps for a classical trajectory of phenanthrene with 3 excited singlet states. Energies and forces are provided either by SchNet model (blue circles) or by TD-DFTB (red diamonds). Inset contains the results based on SchNet model only. The marks are connected by lines to guide the eye. Each trajectory ran on a single Skylake CPU at the CALMIP computing center in Toulouse. 84

S1 Populations averaged over 89 trajectories in chrysene following excitation to the brightest excited S_8 state with 0.1 meV (left panel) and 10 meV (right panel) energy gap thresholds. These simulations have been performed using the MAT set of DFTB parameters. 97

S2 A snapshot of the `ase test --calculators=demonnano` output in the command line. 100

S3 Screenshots from the page dedicated to the deMon-Nano calculator class on the official ASE website. 102

1 Exemples de molécules de PAH. Les colonnes de gauche et de droite correspondent respectivement aux morphologies cata- et péri-condensées. 136

2 Représentation schématique de scénarii typiques de dynamique FSSH après une transition verticale de l'état fondamental vers un état singulet excité: A) le système reste dans un état excité donné et émet éventuellement un photon avec un excès d'énergie $\hbar\nu$ (photoémission); B) la transition entre deux états se produit dans la région de croisement évité (ou autour d'une intersection conique) conduisant à la relaxation vers une configuration d'état fondamental d'équilibre (décroissance sans rayonnement); C) identique au cas précédent mais le système aboutit à un minimum local différent du minimum absolu de l'état fondamental PES (photoisomérisation). 139

3 Énergies des états excités (ES) les plus brillants et des premiers singulets (ES) obtenues avec les niveaux de théorie CASPT2[100,180] (losanges magenta, ligne pointillée), TD-DFT (carrés rouges, ligne pointillée), TD-DFTB (cercles verts, trait plein) et spectre d'absorption expérimental[181–183] (triangles bleus, ligne pointillée) des polyacènes (n est le nombre de cycles aromatiques dans le composé). Les valeurs des énergies sont reliées par des lignes pour guider l'œil. 141

4 Temps de décroissance de l'état excité singulet le plus brillant (B) (hexagones rouges) et de l'écart d'énergie (losanges bleus) entre les états B et B–1 (celui avec la plus faible énergie dans les spectres d'absorption TD-DFTB) calculés à la géométrie d'équilibre en fonction du nombre de cycles aromatiques dans le polyacène. 142

5 Écarts d'énergie entre les états B et B–1 le long des trajectoires (lignes rouges) et moyenne sur l'ensemble des trajectoires (ligne pointillée noire) pour le pentacène et l'hexacène. La ligne noire horizontale à 0.01 eV est le seuil de détection d'une intersection conique. 143

6 Ecarts d'énergie entre les états B et B–1 le long de 100 trajectoires (lignes rouges) et moyennés sur l'ensemble de l'ensemble (ligne pointillée noire) pour le tétracène et le chrysène. 145

7 Analyse de la population des 4 premiers états excités singulets dans le phénanthrène pour un ensemble de 100 (panneau de gauche) et 250 (panneau de droite) trajectoires propagées avec des paramètres BIO. La ligne pointillée grise correspond à une décroissance exponentielle avec $\tau = 522$ fs extraite des expériences de Blanchet *et al.*[96] L'état initial est S_3 dans les deux cas. 148

8 Evolution temporelle de l'occupation moyenne des états excités S_2, S_3 et S_4 dans le phénanthrène sur un ensemble de 100 trajectoires. Chaque trajectoire a été propagée avec un pas de temps indiqué Δt sur des PES apprises par machine avec la probabilité de saut de Zhu-Nakamura. 150

9 Evolution temporelle de l'occupation moyenne des états excités S_2, S_3 et S_4 dans le phénanthrène sur un ensemble de 100 trajectoires. Chaque trajectoire a été propagée avec un pas de temps indiqué Δt sur des PES apprises par machine avec la probabilité de saut de Belyaev-Lebedev. La ligne pointillée noire est un ajustement exponentiel. 151

10 Temps nécessaire pour effectuer les étapes de propagation N=1000[fs]/Δt pour une trajectoire classique de phénanthrène avec 3 états singulets excités. Les énergies et les forces sont fournies soit par le modèle SchNet (cercles bleus) soit par TD-DFTB (losanges rouges). L'encart contient les résultats basés sur le modèle SchNet uniquement. Les valeurs sont reliées par des lignes pour guider l'œil. Chaque trajectoire est propagée sur un seul processeur Skylake au centre de calcul CALMIP de Toulouse. 152

List of Tables

1.1 Important chemical reactions in the ISM. 3

3.1 Comparison of decay times extracted from two different exponential fits of electronic populations plotted in Figures 3.10–3.12. 48

3.2 Summary of the FSSH simulations. Energies and positions of excited states are taken from the TD-DFTB spectra computed with the MAT set of parameters at the equilibrium geometry. IS stands for the initally excited (also the brightest) state. 54

4.1 Possible solutions to underfitting (high bias) and overfitting (high variance). . . . 64

4.2 Some examples of the SchNet model parameters and the corresponding execution times. These calculations have used forces and GPUs in the training. 70

4.3 Selected results from the experiments (in a chronological order) with neutral phenanthrene in different environments. 73

4.4 More details about several low-lying excited states of neutral phenanthrene extracted from the TD-DFTB and the TD-DFT absorption spectra. H and L stand for HOMO and LUMO, respectively. 75

4.5 Parameters of the considered SchNet models and the corresponding training times and errors. These models have used forces (except #4 and #8) and GPUs (except #4) in the training. The MAEs (RMSEs) for energies and forces are given in eV and eV/Å, respectively. The MAE and RMSE have been computed for a test set (2000 structures for all models except #3 and #7 where 6000 structures have been used). 76

S1 Experimental[181–183] and theoretical[90,100,180] excitation energies for first and brightest excited singlet states of polyacenes ranging in size from naphthalene to octacene. Theoretical values computed at CASPT2[100,180], TD-DFT (this work) and TD-DFTB (this work) levels of theory. 93

S2 TD-DFTB with MIO parameters. 94

S3 TD-DFTB with MAT parameters. 94

S4 TD-DFT with BLYP/6-31G(d,p). 95

S5 TD-DFT with B3LYP/6-31G(d,p). 95

S6 TD-DFT with CAM-B3LYP/6-31G(d,p). 95

1 Résumé des simulations FSSH. Les énergies et les positions des états excités sont tirées des spectres TD-DFTB calculés avec l'ensemble MAT de paramètres à la géométrie d'équilibre. IS représente l'état d'excitation initiale (également le plus brillant). 144

2 Paramètres des modèles SchNet considérés, temps d'apprentissage et erreurs correspondants. Ces modèles ont utilisé des forces (sauf #4 et #8) et des GPU (sauf #4) pour la pahse d'apprentissage. Les MAE (RMSE) pour les énergies et les forces sont données respectivement en eV et eV/Å. Les MAE et RMSE ont été calculées pour un ensemble test (2000 structures pour tous les modèles sauf #3 et #7 où 6000 structures sont été utilisées). 147

Résumé étendu en français

Les hydrocarbures aromatiques polycycliques (PAH) ont été proposés comme porteurs principaux de bandes diffuses observées dans le milieu interstellaire, motivant des études approfondies de leur réponse photophysique et photochimique au rayonnement UV. Les mécanismes sous-jacents en compétition déterminent l'évolution du gaz dans le milieu interstellaire. L'objectif principal de cette thèse est de décrire et de comprendre les mécanismes de relaxation dans des PAHs de grande taille, par des simulations de dynamique moléculaire non adiabatique, couplées à l'approche de la réponse linéaire "Time-Dependent Density Functional based Tight Binding" (TD-DFTB) des états excités.

Des développements logiciels substantiels prérequis ont été effectués dans le code DFTB deMon-Nano (http://demon-nano.ups-tlse.fr), d'abord avec le calcul des gradients analytiques des surfaces d'énergie potentielle (PES) et des couplages non adiabatiques des états TD-DFTB. L'algorithme de trajectoire à sauts de surface minimaux (FSSH) de Tully a été ensuite adapté à l'approche TD-DFTB afin de prendre en compte les effets non adiabatiques. Après comparaison avec des méthodes de structure électronique de référence, la première application est dédiée à la dynamique non adiabatique de PAHs cata-condensés linéaires. La relaxation électronique de l'état excité le plus brillant a été simulée pour des polyacènes neutres constitués de 2 à 7 cycles aromatiques. Les résultats montrent une alternance marquée dans les temps de dépopulation de l'état initial pour les polyacènes contenant jusqu'à 6 cycles aromatiques, ce qui est corrélé avec une alternance des écarts d'énergie entre l'état excité initialement peuplé et l'état situé juste en dessous. Puis, l'influence de la géométrie sur la relaxation a été étudiée en comparant deux isomères, (i) le chrysène de type "armchair-edge" et (ii) le tétracène de type "zigzag-edge". Après évaluation des paramétrages de l'hamiltonien DFTB, la relaxation électronique à partir de l'état excité le plus brillant, situé autour de 270 nm pour les deux isomères, a été analysée. Les résultats montrent que la population électronique excitée du chrysène décroît un ordre de grandeur plus rapidement que celle du tétracène. Ceci est aussi corrélé à une différence significative des écarts d'énergie entre l'état initial et l'état situé juste dessous.

Un dernier développement majeur concerne l'utilisation d'algorithmes "Machine Learning" (ML) proposés comme un moyen d'éviter la plupart des calculs de structure électronique, très coûteux en temps calcul. Les performances d'algorithmes de réseaux de neurones appliqués à la dynamique des états excités ont été évaluées. Le cas de la relaxation électronique dans le phénanthrène neutre a été choisi comme test en raison de divers résultats expérimentaux disponibles. L'apprentissage de plusieurs réseaux de neurones a été effectué et leurs précision

et efficacité analysés. De plus, diverses déterminations approchées des probabilités de sauts de surface ont été interfacées avec l'approche ML, résultant en un coût négligeable des simulations de dynamique non adiabatique. L'efficacité des diverses approches simplifiées a été évaluée par rapport à la dynamique FSSH. Dans l'ensemble, la technique ML se révèle un outil très prometteur pour la dynamique dans les états excités à l'échelle de la nanoseconde.

Ce travail de thèse ouvre de nouvelles voies pour étudier la photophysique théorique de complexes moléculaires de grande taille. Enfin, les outils développés et implémentés dans le code deMon-Nano, de manière modulaire, peuvent être combinés avec d'autres approches DFTB sophistiquées (tel que "DFTB+Interaction de Configurations") plus adaptées aux études impliquant des états à transfert de charge.

Introduction (Chapitre 1)

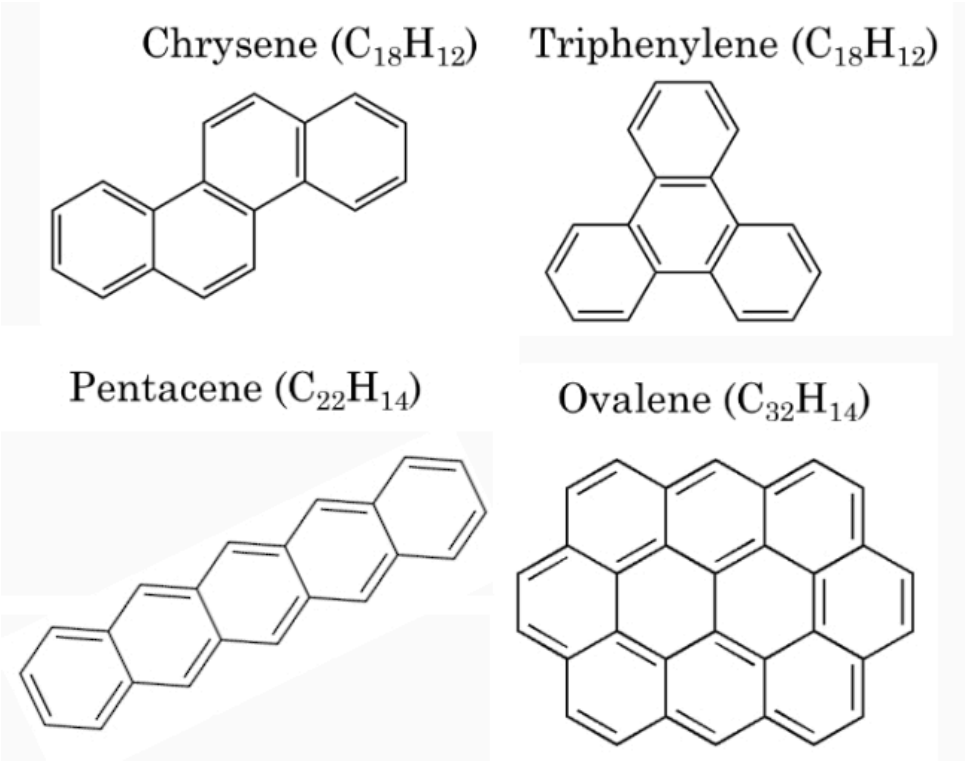


Figure 1: Exemples de molécules de PAH. Les colonnes de gauche et de droite correspondent respectivement aux morphologies cata- et péri-condensées.

Les PAH (voir Figure 1) sont des molécules organiques composées d'au moins deux hexagones carbonés (anneaux de carbone aromatiques) avec des hydrogènes périphériques. Ces anneaux résultent des liaisons sp^2 entre les atomes de carbone et confèrent l'aromaticité à la molécule correspondante. Les PAH péri-condensés contiennent des atomes de carbone partagés entre trois cycles alors que les cata-condensés comportent des cycles disposés suivant un alignement uni-dimensionnel. L'existence des PAH dans le Milieu Interstellaire est fondée sur l'hypothèse

que les molécules de PAH atteignent des températures élevées suite à l'absorption d'un photon UV, puis refroidissent par rayonnement *via* une émission IR. Ils peuvent survivre aux rayons UV grâce à leur remarquable stabilité. Plus de détails sur l'hypothèse PAH peuvent être trouvés dans des publications antécédantes Refs. [10, 11] et les bibliographies associées.

Méthodes (Chapitre 2)

Les noyaux sont plus lourds que les électrons et, en général, leur mouvement est beaucoup plus lent. Ainsi, on peut hiérarchiser le système moléculaire entre électrons décrits au niveau quantique et noyaux décrits par des trajectoires classiques sur les surfaces d'énergie potentielle électroniques. Ces hypothèses forment l'approximation de Born-Oppenheimer - l'un des concepts clés de la chimie quantique. Elles sont à la base des méthodes de dynamique moléculaire (MD) et également des méthodes mixtes quantique-classique de dynamique non-adiabatique.

La théorie "Density-Functional based Tight-Binding" (DFTB)

Les surfaces d'énergie potentielle et la structure électronique (états excités) sont calculés dans le cadre de la méthode DFTB. Les méthodes empiriques de liaison forte (basées sur la représentation limitée des bases orbitales atomiques) ont été développées afin d'étudier les grands systèmes moléculaires.[114] et sont environ trois ordres de grandeur plus rapides que les approches par la Théorie de la Fonctionnelle de la Densité (DFT). Elles nécessitent un paramètrage spécifique aux éléments. Ces paramètres sont généralement transférables, mais surtout, ils permettent à des schémas à liaison forte de décrire presque tous les effets quantiques accessibles par les méthodes DFT. Il a été montré qu'un schéma de liaison forte peut être dérivé d'un traitement approximatif presque sans paramètre de l'approche DFT.[115] Cela a conduit au développement de la méthode "Density-Functional based Tight-Binding" (DFTB). L'approche DFTB dans sa version autocohérente (SCC-DFTB) a été largement utilisée pour étudier de grandes molécules, des clusters, des nanoparticules et des systèmes de matière condensée.[116,117]

Les jeux de paramètres DFTB peuvent être téléchargés gratuitement sur www.dftb.org. Parmi eux, les ensembles de paramètres matsci[131] (initialement développés pour les applications de la science des matériaux) et mio[121] (initialement paramétrés pour O, N, C et H) ont été activement utilisés pour des applications pratiques.

La théorie fonctionnelle de la densité dépendante du temps (TD-DFT)

La théorie fonctionnelle de la densité dépendante du temps (TD-DFT) est une extension de la DFT conventionnelle à la situation où un système dans son état stationnaire fondamental est exposé à une perturbation dépendante du temps, qui modifie son potentiel externe.[136]

De manière analogue à la preuve originale de Hohenberg et Kohn[41] pour l'état fondamental, Runge et Gross ont démontré en 1984 l'existence d'une correspondance unique entre la densité d'électrons dépendante du temps $\rho(\mathbf{r}, t)$ et le potentiel dépendant du temps.[140] Cela conduit à l'ensemble de équations Kohn-Sham dépendantes du temps qui doivent être résolues explicitement si le potentiel dépendant du temps est fort. Par contre, dans la limite d'une faible perturbation, l'approximation de la réponse linéaire peut être appliquée pour simplifier la solution.[136]

Réponse linéaire TD-DFT(B)

En suivant les dérivations de Ref. [136] , on arrive au problème des valeurs propres non hermitiennes suivant (l'équation de Casida)[85]

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X}^I \\ \mathbf{Y}^I \end{pmatrix} = \Omega_I \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \begin{pmatrix} \mathbf{X}^I \\ \mathbf{Y}^I \end{pmatrix}, \tag{1}$$

où $\Omega_I = E_I - E_0$ est l'énergie d'excitation correspondant à la transition verticale entre l'état fondamental E_0 et l'état excité E_I, $\mathbb{1}$ est la matrice d'identité, $\mathbf{A}$ and $\mathbf{B}$ sont des matrices avec les éléments donnés par

$$A_{ia,jb} = \omega_{ia}\delta_{ij}\delta_{ab} + 2K_{ia,jb}; \tag{2}$$

$$B_{ia,jb} = 2K_{ia,jb}; \tag{3}$$

et les indices i, j et a, b désignant respectivement les orbitales Kohn-Sham occupées et virtuelles; $\mathbf{K}$ est la matrice de couplage; $\omega_{ia} = \varepsilon_a - \varepsilon_i$.

Gradients analytiques des surfaces d'énergie potentielle TD-DFTB

La dynamique moléculaire est un outil puissant qui peut être utilisé pour décrire de nombreux processus physiques et chimiques. Afin de propager une trajectoire classique sur une PES donnée en utilisant les équations de mouvement de Newton, des gradients de l'énergie des état excités doivent être calculés pour la réponse linéaire TD-DFTB. Leur dérivation repose sur la méthode dite du vecteur Z, initialement developpée par Furche et Ahlrichs[150,151] pour calculer les forces analytiques dans l'approche TD-DFT, ensuite adaptée par Heringer *et al.*[152,153] (ces gradients ont également été dérivés dans l'approche TD-DFTB à séparation de portée[128]).

Le gradient de l'énergie d'excitation sur l'atome A s'exprime

$$\frac{d\Omega_I}{d\mathbf{R}_A} = \sum_{\mu\nu} P^I_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial \mathbf{R}_A} - \sum_{\mu\nu} W^I_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial \mathbf{R}_A} + 2 \sum_{\mu\nu\lambda\sigma} (X+Y)^I_{\mu\nu} (X+Y)^I_{\lambda\sigma} \frac{\partial K_{\mu\nu,\lambda\sigma}}{\partial \mathbf{R}_A}, \tag{4}$$

où $P^I = T^I + Z^I$ est la matrice de densité de différence à une particule; $H_{\mu\nu}$ et $S_{\mu,\nu}$ sont les matrices hamiltonienne DFTB et le recouvrement.

Dynamique moléculaire non adiabatique

La dynamique moléculaire standard repose sur l'approximation de Born-Oppenheimer, qui conserve un seul état électronique adiabatique dans le développement de la fonction d'onde du système. Plusieurs approches ont été développées ces dernières années pour incorporer les effets non adiabatiques à la fois au niveau de la théorie *ab initio*[61] et mixte quantique-classique[62–64,73]. De nos jours, l'une des méthodes quantique-classique les plus couramment utilisées est l'approche "Fewest-Switches trajectory Surface Hopping" (FSSH) de Tully[83,84], fondée sur l'hypothèse que le mouvement du paquet d'ondes nucléaires peut être simulé par un ensemble de trajectoires classiques indépendantes, chacune évoluant sur un seul état électronique à un instant donné avec une certaine probabilité de passer d'un état à un autre (voir Figure 2) médiée par les couplages non-adiabatiques.

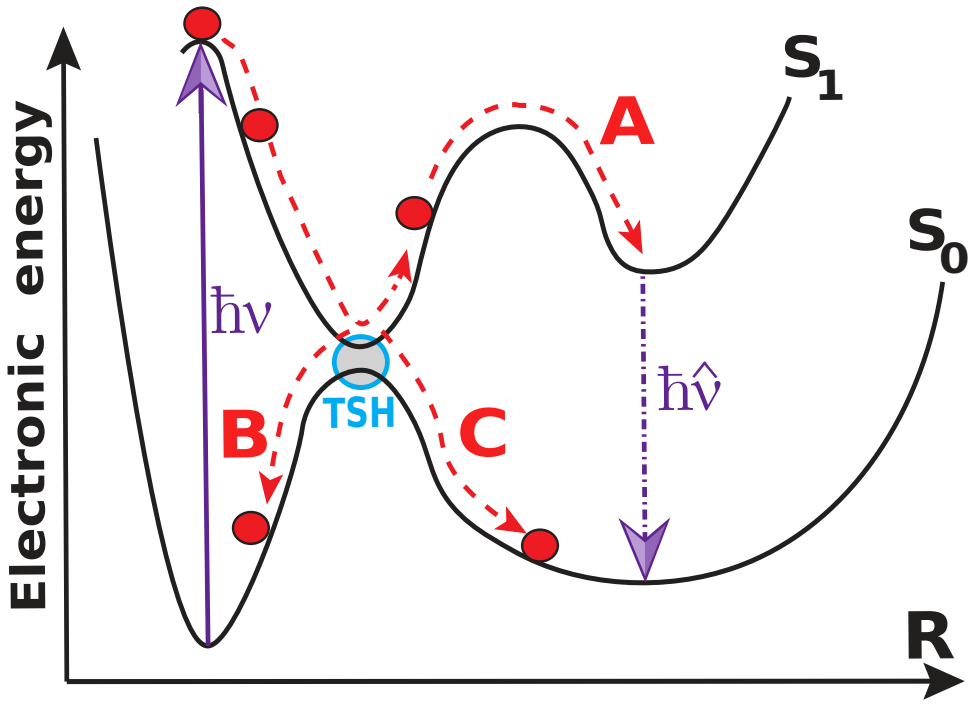


Figure 2: Représentation schématique de scénarii typiques de dynamique FSSH après une transition verticale de l'état fondamental vers un état singulet excité: A) le système reste dans un état excité donné et émet éventuellement un photon avec un excès d'énergie $\hbar\nu$ (photoémission); B) la transition entre deux états se produit dans la région de croisement évité (ou autour d'une intersection conique) conduisant à la relaxation vers une configuration d'état fondamental d'équilibre (décroissance sans rayonnement); C) identique au cas précédent mais le système aboutit à un minimum local différent du minimum absolu de l'état fondamental PES (photoisomérisation).

Au cours de cette thèse, le calcul des gradients de l'énergie, les couplages non adiabatiques et la propagation FSSH entre surfaces d'énergie potentielle TD-DFTB ont été implémentés dans le code deMon-Nano développé à Toulouse.

FSSH couplé à TD-DFTB

Les couplages non adiabatiques (NAC) sont des quantités clés de la FSSH car ils régissent la propagation des populations électroniques et déterminent les probabilités de saut. La DFTB, en tant que méthode fonctionnelle de densité, n'a pas été initialement développée pour les calculs de paquets d'ondes électroniques. Ainsi, il n'est pas clair de savoir comment construire des fonctions d'onde électroniques d'états excités. La pratique courante consiste à utiliser l'approche CIS (Interaction de Configuration des mono-excitations).[87,89,136,158]

Deux manières différentes de calculer les coefficients du développement CIS c_{ia}^K sont couramment utilisées dans la littérature:

1. $c_{ia}^K = (X+Y)_{ia}^K$ – a été utilisé par Tapavicza *et al.*[158] suite à la suggestion de Casida de Ref. [136]. Le principal inconvénient de cette approche est que les états excités ne sont pas orthogonaux par défaut;

2. $c_{ia}^K = F_{ia}^K$ – a été proposé par Werner *et al.*[164] Les états excités sont orthogonaux par défaut car F_{ia}^K est un vecteur propre d'une matrice hermitienne.

Au cours de cette thèse, la méthode FSSH/TD-DFTB a été implémentée, puis utilisée pour étudier les mécanismes de relaxation dans les PAH neutres cata-condensés comme les polyacènes et les phénacènes. Du point de vue algorithmique, mon implémentation FSSH est un module interne du code deMon-Nano, c'est-à-dire que toutes les propriétés dynamiques sont calculées "on the fly" (à la volée) le long des trajectoires et stockées en mémoire pour une exploitation statistique. Ceci est particulièrement utile pour les simulations de grands systèmes moléculaires ou pour la dynamique impliquant de nombreux états car aucune entrée/sortie de données étendue n'est requise comme dans le cas d'un propagateur FSSH externe. De plus, ce module FSSH peut être interfacé avec toute autre méthode de structure électronique pour les états excités (par exemple DFTB-CI) disponible dans le code deMon-Nano.

Dynamique moléculaire non adiabatique des grands PAH cata-condensés (Chapitre 3)[1]

[1] Travail publié, voir:
[90] E. Posenitskiy *et al.*, Phys. Chem. Chem. Phys. **21**, 12139–12149 (2019)
[168] E. Posenitskiy *et al.*, J. Chem. Phys. **152(7)**, 074306 (2020)

Analyse des spectres d'absorption

Les polyacènes ont la symétrie du groupe ponctuel D_{2h} et l'état fondamental appartient à la représentation irréductible A_g. Leur état excité les plus brillant appartiennent à la représentation B_{3u} irréductible si les molécules considérées sont alignées suivant l'axe x. Le spectre optique est dominé par les transitions $\pi \rightarrow \pi^*$, qui sont raisonnablement bien prédites par l'approche TD-DFTB comme l'a montré Niehaus *et al.* dansi la Ref. [86] . Néanmoins, le même groupe a remarqué que la TD-DFTB ne parvient pas à décrire avec précision les transitions $n \rightarrow \pi^*$ ou $\sigma \rightarrow \pi^*$ en raison d'un couplage artificiellement nul lié à l'utilisation des charges de transition atomiques de Mulliken. Ce problème peut être surmonté en utilisant la correction sur site de Domínguez *et al.*[179].

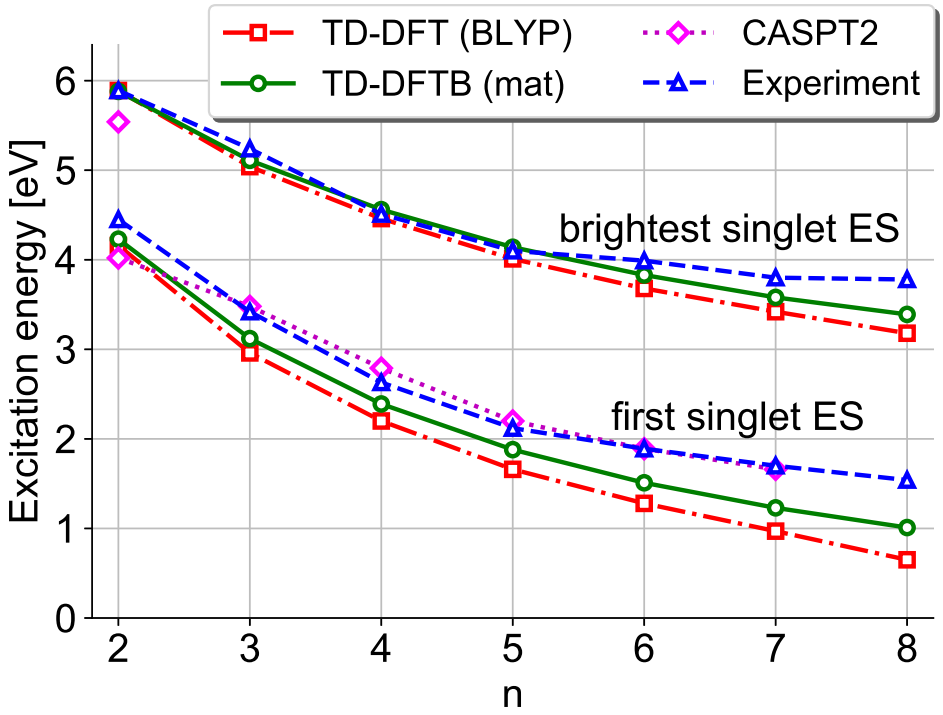


Figure 3: Énergies des états excités (ES) les plus brillants et des premiers singulets (ES) obtenues avec les niveaux de théorie CASPT2[100,180] (losanges magenta, ligne pointillée), TD-DFT (carrés rouges, ligne pointillée), TD-DFTB (cercles verts, trait plein) et spectre d'absorption expérimental[181–183] (triangles bleus, ligne pointillée) des polyacènes (n est le nombre de cycles aromatiques dans le composé). Les valeurs des énergies sont reliées par des lignes pour guider l'œil.

Les résultats théoriques TD-DFTB ont été validés sur la base d'une comparaison avec les données expérimentales et théoriques de haut niveau disponibles (CASPT2). En particulier, les énergies d'excitation des états singulets excités les plus bas et les plus brillants des polyacènes sont résumées dans la Figure 3. Les positions des bandes expérimentales pour le naphtalène, l'anthracène, le tétracène et le pentacène sont tirées de la Ref. [181] ainsi que de travaux

existants pour l'hexacène et l'heptacène[182] et également pour l'octacène[183].

Il est clair que les positions des états excités les plus brillants et des premiers singulets dans TD-DFTB sont en bon accord avec les valeurs expérimentales et CASPT2. Ceci est cohérent avec les résultats TD-DFTB précédemment publiés, calculés avec un ensemble de paramètres différent jusqu'à l'heptacène.[86] Cependant, l'erreur par rapport à l'expérience augmente avec la taille croissante du système, c'est-à-dire que pour l'octacène, les différences sont de 0.53 eV pour le premier singulet excité et 0.39 eV pour le plus brillant. Ainsi, nous avons limité notre étude jusqu'à l'heptacène pour lequel l'état excité singulet le plus brillant est encore raisonnablement bien prédit et la sous-estimation par rapport à la valeur expérimentale de 0.22 eV seulement.

Dépendance de la relaxation ultra-rapide à la taille

La Figure 4 montre le temps de décroissance de l'état excité singulet le plus brillant en fonction du nombre de cycles aromatiques dans le polyacène.

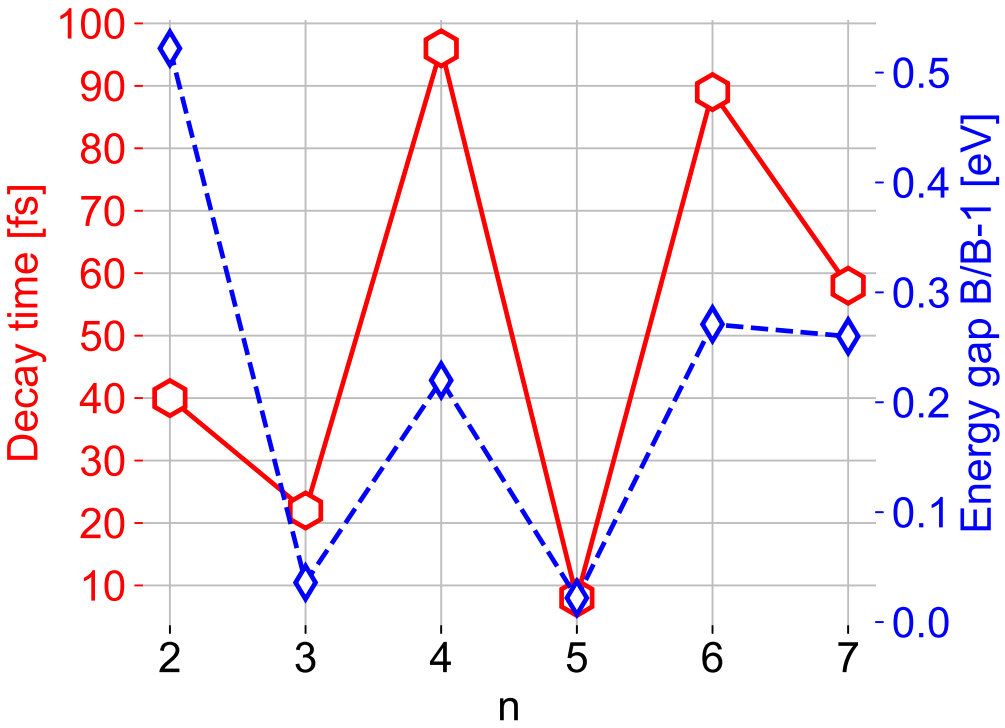


Figure 4: Temps de décroissance de l'état excité singulet le plus brillant (B) (hexagones rouges) et de l'écart d'énergie (losanges bleus) entre les états B et B–1 (celui avec la plus faible énergie dans les spectres d'absorption TD-DFTB) calculés à la géométrie d'équilibre en fonction du nombre de cycles aromatiques dans le polyacène.

Les temps de décroissance peuvent être extraits d'ajustements exponentiels des courbes de population calculées avec des conditions initiales à partir de l'état excité singulet le plus brillant.

Bien que la dépopulation ne soit pas totalement exponentielle, nos ajustements approximatifs sur la fenêtre de temps complet permettent une comparaison quantitative entre les polyacènes. Il est à noter que les temps de décroissance de la Figure 4 sont du même ordre de grandeur (sauf pour le naphtalène) que ceux publiés pour les espèces cationiques par Reddy *et al.* dans la Ref. [23] même si les croisements d'états cationiques pourraient être qualitativement différents de ceux des molécules neutres.[18,19] Cependant, les résultats actuels montrent une alternance inattendue et frappante des temps de décroissance pour un nombre pair-impair de cycles dans la plage de taille $n = 3$–6. Du point de vue de la photochimie organique[5], on pourrait s'attendre à la diminution monotone du temps de relaxation en fonction de la taille régie par ila diminution des écarts d'énergie et la densité croissante d'états électroniques.

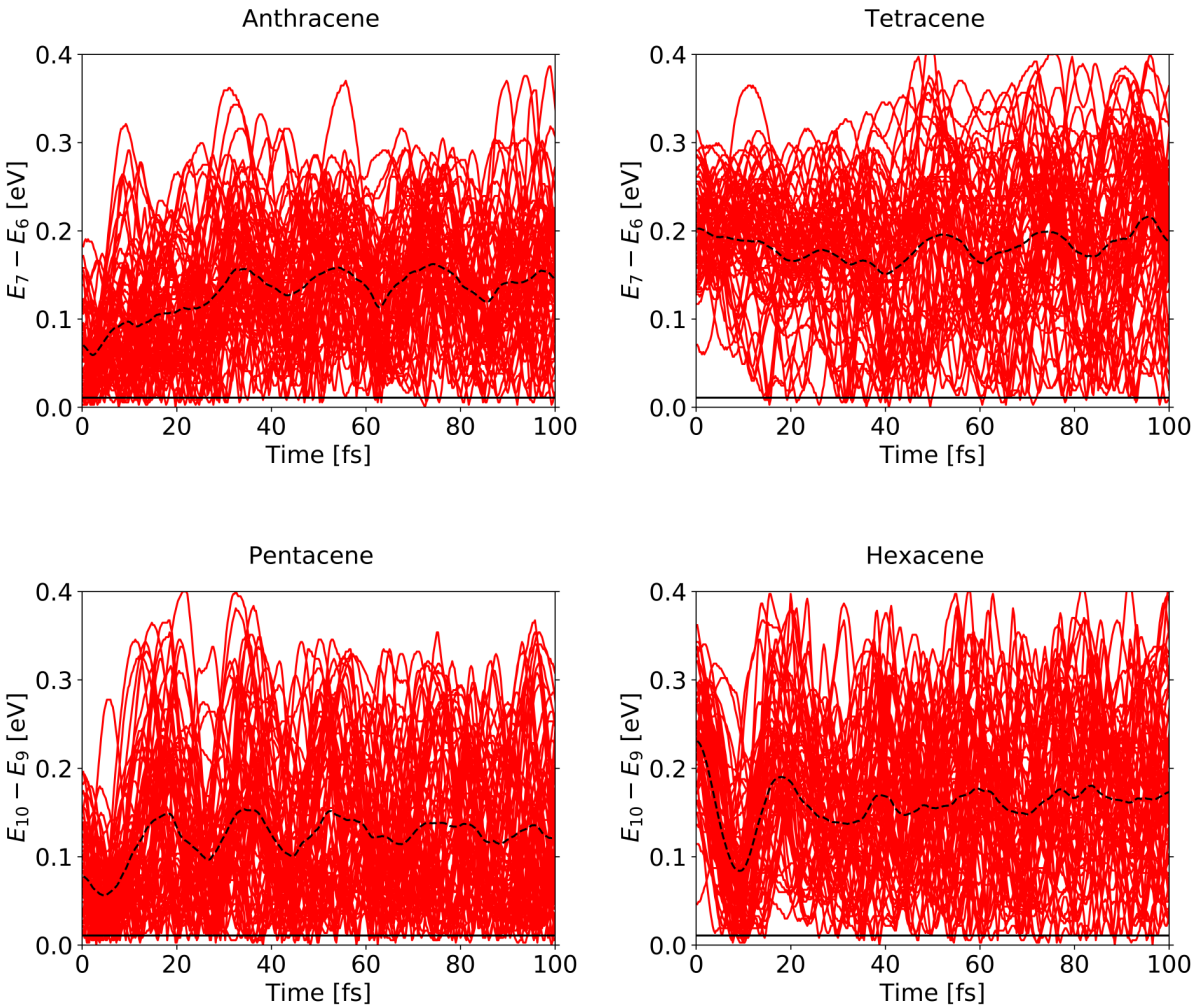


Figure 5: Écarts d'énergie entre les états B et B–1 le long des trajectoires (lignes rouges) et moyenne sur l'ensemble des trajectoires (ligne pointillée noire) pour le pentacène et l'hexacène. La ligne noire horizontale à 0.01 eV est le seuil de détection d'une intersection conique.

Afin d'expliquer cette alternance, l'écart d'énergie entre l'état initial B et l'état voisin d'energie inférieure B–1 a été déterminé pour les polyacènes comprenant $n = 3$–6 cycles aromatiques. Sa variation fournit une bonne compréhension de la façon dont se produit le transfert de population

depuis l'état excité initial : la plupart des trajectoires subissent un saut de surface dans la région où l'état initial B se rapproche énergétiquement de l'état adiabatique d'énergie immédiatement inférieure B-1 ou lorsqu'elles atteignent une intersection conique (CoIn). Il ressort clairement de la Figure 5 qu'au cours des 20 premières fs, une grande partie des trajectoires dans l'anthracène et le pentacène vont soit vers la région des faibles écarts, soit vers la CoIn. Ainsi, il existe un processus d'entonnoir dans ces deux polyacènes, qui entraîne le transfert rapide de population électronique de l'état initial. Au contraire, dans le tétracène, très peu de trajectoires atteignent la région des faibles écarts au delà de 20 fs et la valeur moyenne de l'écart est presque deux fois supérieure à celle de l'anthracène. Ceci est également vrai pour l'hexacène, cas pour lequel il y a une nette diminution de l'écart énergétique autour de 10 fs, tout en restant plus grande que dans le cas du pentacène.

Nos résultats pour les polyacènes enfreignent la règle de Kasha concernant la dépopulation ultrarapide à l'état S_1. Il y a plusieurs raisons possibles: (i) pour les molécules isolées, l'absence de modes dissipatifs vers l'environnement entraîne la possibilité d'un retour de population vers les états excités supérieurs à partir des états plus bas;[89] (ii) 300 fs est encore une échelle de temps assez courte pour la relaxation complète des états excités supérieurs, ainsi certaines trajectoires restent dans les états au-dessus de S_1 à la fin de la simulation; (iii) les systèmes conjugués comme les polyacènes ont certains états excités de basse energie dominés par des doubles excitations[189], qui manquent dans la réponse linéaire TD-DFTB. Enfin, il convient de mentionner que les preuves expérimentales de comportements non-Kasha sont bien documentées, en particulier pour les molécules polyatomiques de petite taille, y compris pour le naphtalène.[173]

Dépendance de la relaxation ultra-rapide à la morphologie

Nous avons également comparé l'influence de la morphologie pour des PAH de même taille, i.e. le chrysène et le tétracène. Les résultats de nos calculs sont résumés dans la Table 1.

Table 1: Résumé des simulations FSSH. Les énergies et les positions des états excités sont tirées des spectres TD-DFTB calculés avec l'ensemble MAT de paramètres à la géométrie d'équilibre. IS représente l'état d'excitation initiale (également le plus brillant).

Molécule	IS	Ω_{IS}, eV	N_{traj}	Temps de décroissance du SI, fs	Déficit énergétique IS//IS–1, eV
Chrysène	S_8	4.61	100	7	0.05
Tétracène	S_7	4.56	100	64	0.22
Tétracène[a]	S_7	4.56	63	65	0.22

[a] Adapté de Ref. [90]

Afin d'expliquer la différence entre les temps de décroissance et les mécanismes de relaxation sous-jacents, les écarts d'énergie entre l'état initial et l'état voisin avec une énergie plus faible ont été répertoriés le long de chaque trajectoire pour le tétracène et le chrysène (voir Figure 6). Il est clair qu'au début de la dynamique dans le cas du chrysène, toutes les trajectoires vont soit vers la région des petits écarts, soit vers une CoIn. Ainsi, il y a un entonnoir, qui entraîne une rapide dépopulation de l'état initial. Au contraire, dans le cas du tétracène, très peu de trajectoires atteignent la région des petits écarts et l'écart moyen reste deux fois plus grand que celui observé pour le chrysène. Ceci est cohérent avec notre conclusion précédente pour les polyacènes. La dépopulation est également bien corrélée avec le gap d'énergie calculé à la géométrie d'équilibre (voir Table 1). Cet écart initial s'avère être grossièrement conservé pendant la phase de temps court de la trajectoire et peut être utilisé comme une approximation du premier ordre pour estimer les temps de décroissance. De manière constante, le nombre de sauts de surfaces est plus important dans le chrysène que dans le tétracène. En plus des écarts relatifs sensiblement différents entre les états excités, la symétrie de la géométrie initiale du chrysène (C_{2h}) est plus basse que celle du tétracène (D_{2h}), ce qui se traduit par plus de valeurs non nulles des éléments de couplage non adiabatique et peut également acélérer la relaxation électronique.

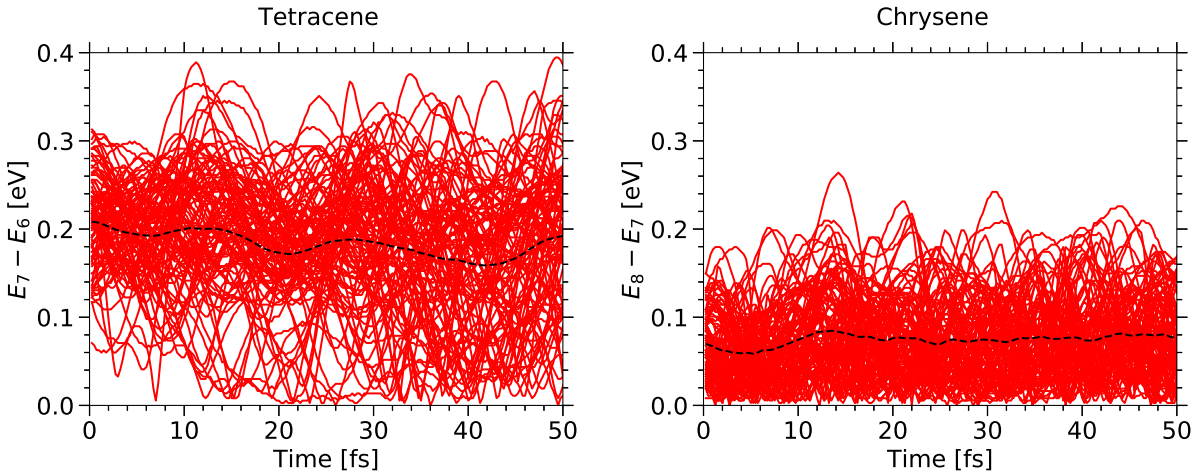


Figure 6: Ecarts d'énergie entre les états B et B–1 le long de 100 trajectoires (lignes rouges) et moyennés sur l'ensemble de l'ensemble (ligne pointillée noire) pour le tétracène et le chrysène.

Test des potentialités des techniques Machine Learning avec une dynamique non adiabatique simplifiée (Chapitre 4)

L'objectif principal de la présente étude est (i) d'appliquer des réseaux de neurones artificiels pour obtenir une bonne précision sur les énergies et les forces apprises par Machine Learning (ML) pour les états excités adiabatiques basés sur les données TD-DFTB; (ii) mettre en oeuvre des simulations de dynamique moléculaire et incorporer les effets non adiabatiques

via des trajectoires avec algorithmes de saut de surface (TSH) simplifiés correspondant aux approches Belyaev-Lebedev (BL) ou Zhu-Nakamura (ZN) sans aucune connaissance *a priori* des intersections coniques ou couplages non adiabatiques; (iii) évaluer les performances de ces schémas simplifiés *versus* les résultats de FSSH et les données expérimentales disponibles. La dynamique dans les états excités les plus bas du phénanthrène ($C_{14}H_{10}$ – un isomère en forme "armchair" de l'anthracène) a été choisie comme cas test en raison de sa taille moléculaire relativement grande, et de la diversité des résultats expérimentaux disponibles et en raison d'une certaine ambiguïté[94] dans leur interprétation.

Pour la première étape, je m'appuie sur une architecture de Deep Learning récemment développée appelée SchNet[221], qui est implémentée dans un package Python open-source SchNetPack[227]. L'implémentation SchNet est basée sur une bibliothèque PyTorch[242] pour les applications Deep Learning qui peuvent être exécutées efficacement sur les GPU. La seconde étape nécessite d'intégrer les effets non adiabatiques dans les simulations TSH. Deux probabilités de saut ont été intégrées dans le driver de dynamique moléculaire: (i) P^{BL} basé sur l'approche Belyaev-Lebedev (BL)[76] adapté pour les états adiabatiques et (ii) P^{ZN} basé sur l'extension multidimensionnelle de la théorie de Zhu-Nakamura (ZN) suivant Hanasaki *et al.*[80]

Apprentissage avec le modèle SchNet

Les performances de plusieurs modèles algorithmique de SchNet sont résumées dans le Table 2. Le MAE (RMSE) est l'erreur moyenne (l'erreur quadratique moyenne). L'entité de toutes les couches, y compris les couches d'entrée et de sortie ainsi que le nombre de nœuds (attributs) par couche, définit la topologie du réseau neuronal artificiel comme SchNet. Notamment, tous atteignent une précision chimique (erreurs inférieures à 1 kcal/mol ou 0.043 eV) sur les énergies TD-DFTB du phénanthrène. Cela n'est pas surprenant, compte tenu de la précision exceptionnelle de SchNet pour la prédiction des propriétés de l'état fondamental[221] et des états excités[210]. Cependant, il convient de mentionner que la précision réelle des quantités apprises par Machine Learning est corrélée avec la méthode de structure électronique sous-jacente qui a été utilisée pour l'apprentissage. Par exemple, les erreurs SchNet peuvent dépasser 1 kcal/mol pour certaines méthodes de structure électronique sophistiquées comme CASSCF.[210] De plus, les erreurs sont plus importantes pour le modèle SchNet #8 (apprentissage sans forces) par rapport au modèle #5, ce qui est cohérent avec des études précédentes.[221,253] De plus, l'inclusion des forces dans l'apprentissage SchNet permet d'optimiser le modèle beaucoup plus rapidement grâce à une quantité importante d'informations fournies par les gradients d'énergie. Le modèle #4 a été entraîné sans GPU ni forces, ce qui a donné lieu à seulement 463 périodes d'entraînement en 48 heures. Une extrapolation approximative suggère qu'il faudrait plus de 100 heures pour effectuer un entraînement entièrement convergé avec 1000 périodes d'entraînement, ce qui serait beaucoup plus long que l'apprentissage de la PES avec les forces et les GPU disponibles. L'erreur ne varie pas de manière significative entre les modèles à 3 et 6 couches

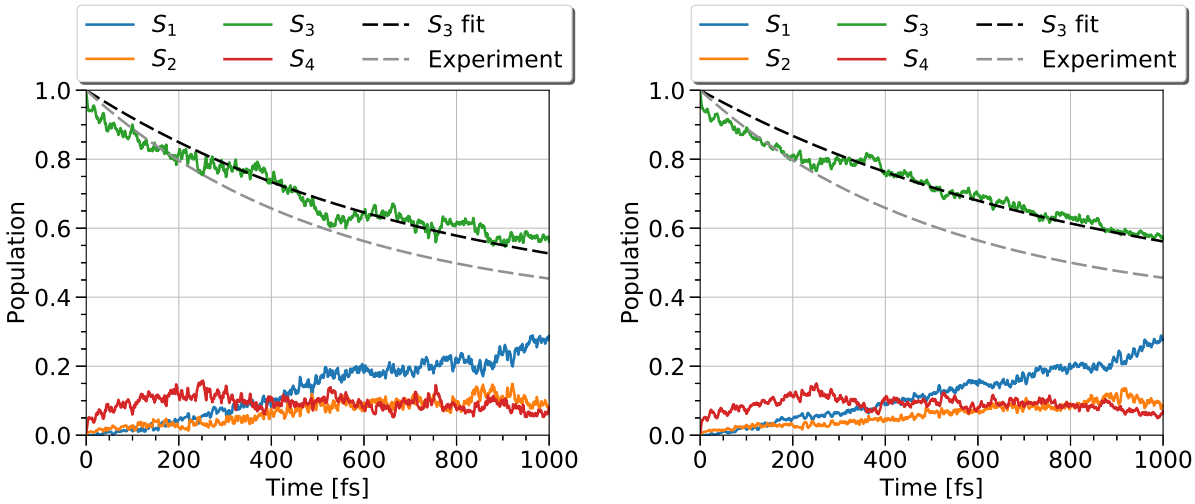
Table 2: Paramètres des modèles SchNet considérés, temps d'apprentissage et erreurs correspondants. Ces modèles ont utilisé des forces (sauf #4 et #8) et des GPU (sauf #4) pour la pahse d'apprentissage. Les MAE (RMSE) pour les énergies et les forces sont données respectivement en eV et eV/Å. Les MAE et RMSE ont été calculées pour un ensemble test (2000 structures pour tous les modèles sauf #3 et #7 où 6000 structures sont été utilisées).

#	Etat	Attributs	Calques cachés	Training set	Validation set	Durée, h	Propriété	MAE (RMSE)
1	S_2	128	6	6000	2000	14	énergie force	0.014 (0.018) 0.075 (0.126)
2	S_2	128	3	6000	2000	15	énergie force	0.018 (0.023) 0.084 (0.138)
3	S_2	64	3	3000	1000	4	énergie force	0.027 (0.035) 0.134 (0.213)
4	S_2	128	3	6000	2000	$\gtrsim$100 [no GPU]	énergie force	— —
5	S_3	128	6	6000	2000	13	énergie force	0.016 (0.022) 0.068 (0.113)
6	S_3	256	3	6000	2000	15	énergie force	0.014 (0.018) 0.062 (0.106)
7	S_3	256	3	3000	1000	5	énergie force	0.017 (0.022) 0.075 (0.124)
8	S_3	128	6	6000	2000	40	énergie force	0.029 (0.038) —

cachées (interactions), ce qui a déjà été souligné par Schütt *et al.*[221] pour l'ensemble de données QM9[254–256]. Pourtant, la différence de précision des réseaux de neurones artificiels avec 1, 2 et 3 couches masquées peut être plus prononcée.[257] Le nombre d'entités a un effet majeur comme on peut le voir en comparant les modèles #3 et #7. En utilisant seulement 3000 points d'apprentissage et 1000 points de validation, le modèle #7 avec 256 attributs atteint une précision similaire au modèle #6 qui a été formé avec deux fois plus de données. En revanche, les performances du modèle #3 avec 64 attributs ne sont comparables qu'à celles du modèle #8 qui n'a pas du tout utilisé les forces. Dans l'ensemble, 128 attributs et 3 couches cachées semblent être un choix raisonnable pour l'architecture SchNet. Toutes les conclusions susmentionnées indiquent l'importance du réglage des hyperparamètres pour les applications pratiques du ML. Une précision remarquable peut être obtenue même sur des ensembles d'apprentissage relativement petits.[240]

Simulations FSSH/TD-DFTB

Un ensemble de 250 trajectoires (voir Figure 7) a été propagé en utilisant l'implémentation FSSH/TD-DFTB récemment développée dans le code deMon-Nano. L'état excité initial est l'état S_3 à la géométrie de l'état fondamental.



100 trajectoires

250 trajectoires

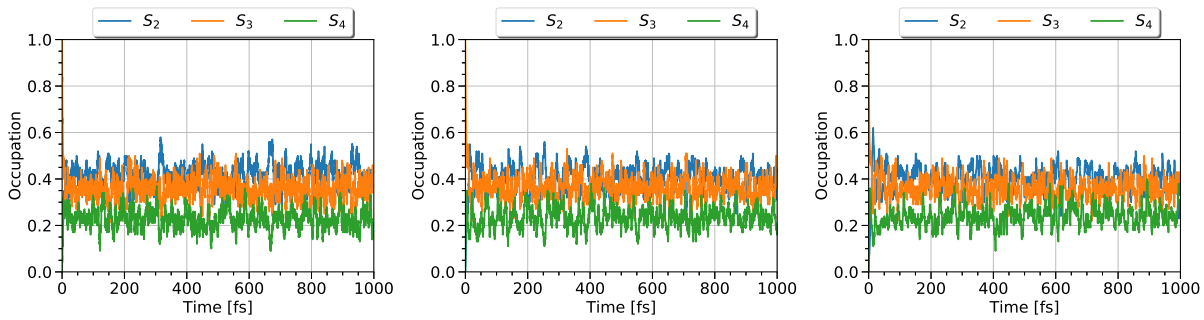
Figure 7: Analyse de la population des 4 premiers états excités singulets dans le phénanthrène pour un ensemble de 100 (panneau de gauche) et 250 (panneau de droite) trajectoires propagées avec des paramètres BIO. La ligne pointillée grise correspond à une décroissance exponentielle avec $\tau = 522$ fs extraite des expériences de Blanchet *et al.*[96] L'état initial est S_3 dans les deux cas.

Tout d'abord, les formes des courbes de population moyennées changent légèrement lorsque le nombre de trajectoires passe de 100 à 250. Cela indique que cet ensemble n'est peut être pas encore suffisant pour atteindre la convergence, contrairement à l'anthracène qui présente

des résultats stables en passant de 63 à 127 trajectoires. Les temps de décroissance τ calculés sont de 745 fs et 860 fs pour 100 et 250 trajectoires, respectivement. Néanmoins, la courbe de population S_3 peut être grossièrement divisée en deux parties. La première est située dans la fenêtre temporelle [0; 400] fs et correspond à une décroissance relativement rapide de la population initiale de S_3 vers les états supérieurs S_4 et inférieurs S_2 et S_1. Elle peut être approximée par un temps de décroissance exponentielle entre 550 à 600 fs. Le transfert de population après 400 fs semble être plus lent que le transfert initial, ce qui peut être corrélé à une augmentation de la population de l'état S_1 qui devient prononcée après 400 fs. Les résultats mentionnés ci-dessus sont en accord qualitatif avec la relaxation multi-états complexe attendue pour le phénanthrène selon Ref. [252] . Comme le soulignent Nazari *et al.*[252], la conversion interne à partir du premier état singulet brillant comporte des contributions rapides (100 fs) et lentes (600 fs). Le même groupe a également réalisé les simulations FSSH/TD-DFT dans les deux états singulets excités les plus bas d'un dérivé de phénanthrène. Les résultats sont en bon accord qualitatif avec les observations expérimentales, malgré le fait de négliger les transitions vers des états excités d'énergie supérieure à l'état initial. La majeure partie (80%) de la population est rapidement (en moins de 200 fs) transférée à l'état excité le plus bas, contrairement aux simulations FSSH/TD-DFTB (ce travail). De plus, les écarts d'énergie TD-DFT (calculés avec les fonctionnelles ωB97XD et BHandHLYP)[252] entre les deux états excités les plus bas sont plus petits que ceux obtenus avec la méthode CASPT2[251,252] ou TD-DFTB (ce travail). Ceci explique une relaxation plus rapide et plus complète (par rapport à nos résultats FSSH/TD-DFTB) rapportée par Nazari *et al.*[252] Dans l'ensemble, notre approche FSSH/TD-DFTB appliquée à la relaxation électronique du phénanthrène neutre est en assez bon accord avec les résultats expérimentaux.

Simulations TSH/SchNet

Nous analysons maintenant les simulations TSH, tout d'abord avec l'approche de sauts ZN et les différentes étapes de propagation (voir Figure 8). Il est clair que les résultats ZN ne sont pas significativement affectés par les changements de pas de temps. Cela n'est pas surprenant, compte tenu de l'utilisation des gradients diabatiques modifiés proposés par Hanasaki *et al.*[80] Le même groupe a fait remarquer que d'autres extensions multidimensionnelles de la théorie ZN (comme celle développée par Yu *et al.*[79]) peuvent être affectées par l'étape de propagation en raison des hypothèses sous-jacentes sur la topologie des PES. Dans l'ensemble, une décroissance ultra-rapide (à environ 10 fs) est observée pour l'état S_3 initialement excité. Cependant, la relaxation électronique n'est pas complète en raison des oscillations résiduelles dans l'occupation de chaque état excité. Que ce soit une conséquence de la mise en œuvre actuelle ou un inconvénient de l'extension multidimensionnelle, la validité de l'approche ZN est une question pour les travaux futurs. Notons également que 100 trajectoires représentent une taille d'ensemble relativement petite.



$\Delta t = 0.1$ fs $\Delta t = 0.25$ fs $\Delta t = 0.5$ fs

Figure 8: Evolution temporelle de l'occupation moyenne des états excités S_2, S_3 et S_4 dans le phénanthrène sur un ensemble de 100 trajectoires. Chaque trajectoire a été propagée avec un pas de temps indiqué Δt sur des PES apprises par machine avec la probabilité de saut de Zhu-Nakamura.

Nous analysons ensuite les simulations basées sur BL réalisées avec différents pas de temps (voir Figure 9). Contrairement à l'approche ZN il existe avec BL des différences significatives dans la relaxation électronique. La décroissance devient d'un ordre de grandeur plus rapide à mesure que Δt passe de 0.5 à 0.05 fs. Une explication possible est que l'approche BL a été dérivée pour un croisement unidimensionnel dans des modèles à deux états. Ainsi, il pourrait ne pas décrire correctement les transitions non adiabatiques entre des PES hautement multidimensionnelles. De plus, Smith et Akimov[245] ont souligné que les approches basées sur BL sont bien adaptées pour étudier la dynamique dans des variétés d'états excités avec des écarts d'énergie inférieurs à 0.1–0.2 eV, mais peuvent échouer pour des états séparés par des écarts $\geqslant$0.5 eV. De tels écarts relativement importants sont observés entre les états excités S_2 et S_3 dans le spectre d'absorption TD-DFTB du phénanthrène. Il faut noter que la seconde dérivée temporelle de l'écart d'énergie Z_{IJ} dans l'approche BL est calculée en utilisant des différences finies, c'est-à-dire $\ddot{Z}_{IJ} \sim \mathcal{O}(\Delta t^2)$ et $P_{IJ}^{\mathrm{BL}} \propto \exp(-\Delta t)$. Ce problème, ajouté au fait que le nombre de tentatives de sauts augmente à mesure que le pas de temps diminue (en raison du plus grand nombre de pas MD à calculer), peut provoquer les variations observées. Pourtant, il a été montré pour plusieurs systèmes que les simulations TSH basées sur BL peuvent fournir des résultats avec une précision similaire que celles de FSSH ou même reproduire certaines tendances des calculs de paquets d'ondes quantiques.[76,258] En fait, les simulations BL basées sur $\Delta t = 0.25$ fs sont en bon accord avec les calculs FSSH présentés dans la Figure 7. Notamment, le temps de décroissance calculé pour 100 trajectoires est d'environ 467 fs, ce qui est juste entre 0.44 et 0.48 ps dérivé des travaux de Brechignac et Hermine[238] et Amirav *et al.*[246], respectivement. Dans le cas de $\Delta t = 0.5$ fs, les résultats sont qualitativement les mêmes que ceux pour $\Delta t = 0.25$ fs et les temps de décroissance calculés sont respectivement de 486 et 419 fs pour 100 et 500 trajectoires. Il est à noter que les simulations basées sur BL avec $\Delta t = 0.05$ fs sont similaires aux résultats ZN, à savoir qu'il y a un transfert ultra-rapide de la population électronique de S_3 à S_2 et S_4 suivis d'oscillations dans la dynamique d'occupation.

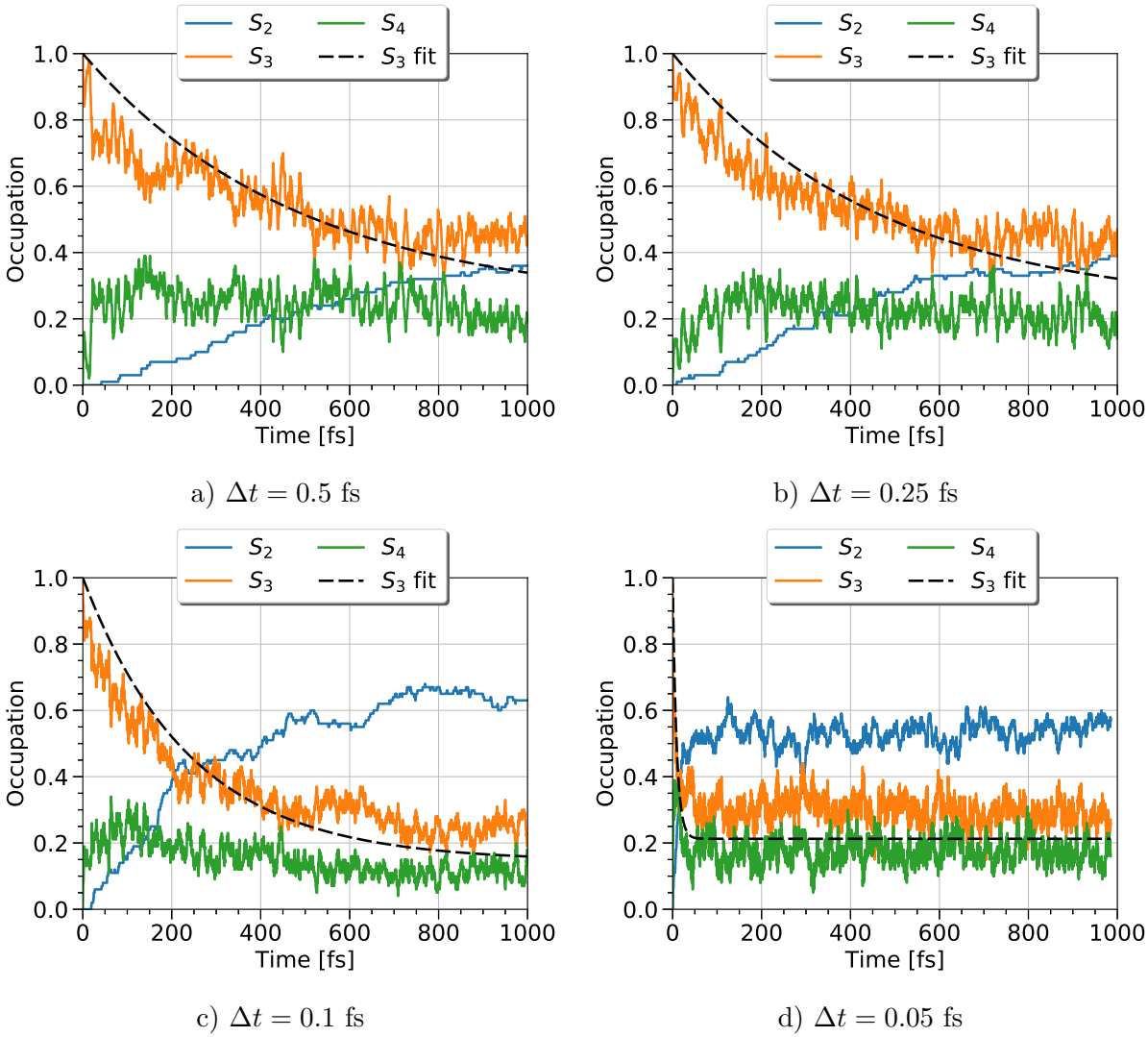


Figure 9: Evolution temporelle de l'occupation moyenne des états excités S_2, S_3 et S_4 dans le phénanthrène sur un ensemble de 100 trajectoires. Chaque trajectoire a été propagée avec un pas de temps indiqué Δt sur des PES apprises par machine avec la probabilité de saut de Belyaev-Lebedev. La ligne pointillée noire est un ajustement exponentiel.

Enfin, les coûts de calcul sont comparés pour les approches considérées dans cette étude. Chaque trajectoire est propagée sur un seul processeur Skylake au centre de calcul CALMIP de Toulouse. Une seule propagation de trajectoire FSSH avec $\Delta t = 0.25$ fs (4000 pas) prend 33 heures CPU lorsque 4 états sont inclus et 13 heures avec 3 états (valeurs moyennes sur 200 trajectoires). Notamment, la propagation basée sur le ML de 3 états (voir les cercles bleus dans la Figure 10) ne prend que 30 minutes avec $\Delta t = 0.25$ fs (4000 étapes) et environ 70 minutes avec $\Delta t = 0.1$ fs (10000 pas), ce qui est déjà un gain de temps CPU remarquable. Cependant, il faut également prendre en compte le temps consommé pour l'étape d'entraînement du modèle ML (SchNet), qui variait entre 4 et 15 heures sur un seul GPU en fonction du nombre de points d'entraînement (voir Table 2). D'autre part, l'apprentissage ne doit être effectué qu'une seule fois et le modèle SchNet résultant peut être utilisé en principe pour un nombre arbitraire de trajectoires. Ainsi,

les simulations TSH/SchNet peuvent être au moins un ordre de grandeur plus rapides que celles FSSH/TD-DFTB (voir Figure 10) tout en évitant une mise à l'échelle exponentielle par rapport au nombre d'états inclus dans la propagation. Les performances de calcul des approches BL/SchNet et ZN/SchNet sont presque les mêmes (le temps de calcul est d'environ 70 minutes pour 10000 pas), malgré la conversion supplémentaire des gradients d'énergie adiabatiques en gradients diabatiques dans le cas ZN. Cependant, il existe une différence drastique lorsque les quantités apprises par machine sont remplacées par celles de TD-DFTB. Il faut environ 108 ou 163 minutes pour effectuer 2000 étapes de propagation à l'aide des schémas BL/TD-DFTB ou ZN/TD-DFTB, respectivement. Cette augmentation du temps calcul observé est principalement dû à l'importance des entrées/sorties nécessaire pour calculer les énergies et les forces "on the fly" (à la volée) à l'aide de l'exécutable externe deMon-Nano. Comme prévu, les performances de calcul TSH/TD-DFTB sont un cas intermédiaire entre les approches FSSH/TD-DFTB et TSH/SchNet approximatives.

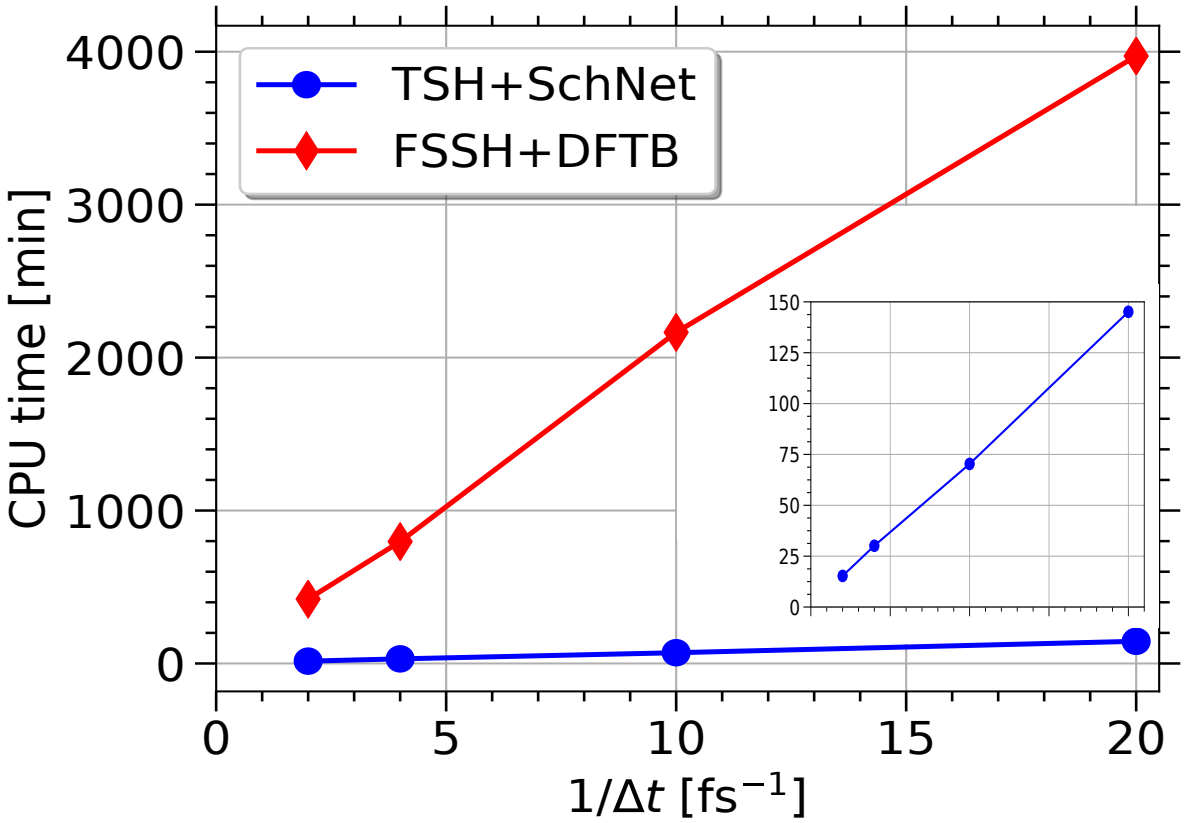


Figure 10: Temps nécessaire pour effectuer les étapes de propagation N=1000[fs]/Δt pour une trajectoire classique de phénanthrène avec 3 états singulets excités. Les énergies et les forces sont fournies soit par le modèle SchNet (cercles bleus) soit par TD-DFTB (losanges rouges). L'encart contient les résultats basés sur le modèle SchNet uniquement. Les valeurs sont reliées par des lignes pour guider l'œil. Chaque trajectoire est propagée sur un seul processeur Skylake au centre de calcul CALMIP de Toulouse.

Conclusion et perspectives (Chapitre 5)

Dans cette thèse, des approches mixtes classiques-quantiques ont été mises en œuvre et appliquées pour simuler la dynamique moléculaire non adiabatique des PAH. Afin d'étudier les grands composés PAH, la méthode TD-DFTB a été choisie pour les calculs de structure électronique.

En tant qu'outil théorique, l'approche FSSH de Tully a été largement utilisée pour les calculs de dynamique moléculaire non-adiabatique. Dans le chapitre 2, on peut trouver des détails méthodologiques et d'implémentation liés au schéma FSSH couplé à l'approche TD-DFTB dans le code deMon-Nano[99]. Comme le montrent les chapitres concernant les résultats, la synergie des deux méthodes a fourni des informations approfondies sur les processus et les mécanismes sous-jacents. En particulier une alternance pair-impair des temps de décroissance de l'état excité initialement peuplé (état le plus brillant) a été démontrée sur les polyacènes en fonction de leur taille entre $n = 3$ et 6 cycles. De même l'étude comparéee du chrysène et du tétracène a permis de démontrer un dépendance importante en fonction de la morphologie, pour deux composés qui ont pourtant des spectres d'excitation très similaire dans la région de l'état brillant.

La dernière partie des résultats est consacrée à une étude plus méthodologique concernant la possibilité de réaliser des simulations de dynamique non-adiabatique dans le cadre de techniques de Machine Learning, en utilisant des schémas simplifiés pour décrire les processus sauts de surfaces et la méthode TD-DFTB pour la structure électronique. Le phénanthrène neutre a été choisi comme test en raison de sa taille moléculaire relativement grande, de la diversité des résultats expérimentaux disponibles et en raison d'une certaine ambiguïté dans leur interprétation. Dans un premier temps, le spectre d'absorption TD-DFTB a été analysé et comparé aux TD-DFT, CASPT2 et aux données expérimentales disponibles. L'accord pour les transitions de singulet-singulet basse énergie est remarquablement bon. Les performances de plusieurs modèles d'apprentissage automatique (SchNet) à partir de différents ensembles de données TD-DFTB ont été évaluées. La conclusion est que l'architecture SchNet pour le Deep Learning est capable de reproduire avec précision les PES multidimensionnelles complexes du phénanthrène, en particulier lorsque des forces sont incluses dans l'étape d'apprentissage. Ensuite, la relaxation électronique de l'état S_3 (état brillant) *via* la cascade de transitions non radiatives a été étudiée. L'analyse détaillée des simulations FSSH/TD-DFTB révèle un bon accord entre la dépopulation simulée et les résultats expérimentaux. De plus, la relaxation électronique calculée dans le phénanthrène est significativement plus lente que ce qui a été rapporté auparavant pour les polyacènes et le chrysène. Encore une fois, cela correspond bien à l'écart d'énergie relativement important TD-DFTB entre l'état initialement excité et celui immédiatement en dessous en énergie. L'accent a été mis principalement sur les résultats calculés avec les schémas TSH simplifiés (Belyaev-Lebedev ou Zhu-Nakamura) avec des PES apprises par SchNet. Les simulations ont été réalisées avec différents pas de temps et expressions des probabilités de saut. Tout d'abord, la théorie de Zhu-Nakamura n'a pas réussi à reproduire

les résultats de la FSSH et les échelles de temps de relaxation dérivées expérimentalement. La décroissance ZN observée est trop rapide et suivie d'oscillations dans la dynamique d'occupation. Au contraire, les résultats de l'approche Belyaev-Lebedev concordent mieux avec les données de référence en dépit d'une absence d'information sur les gradients d'énergie adiabatique, les échelles de temps de décroissance calculées sont proches de celles dérivées des expériences disponibles dans les jets supersoniques. Cependant, ces résultats ont également montré une absence de convergence lorsque les pas de propagation deviennent petits, ce qui peut être génant. Enfin, les performances de calcul ont été comparées pour toutes les approches considérées dans cette étude.

Il faut préciser les limites des présentes conclusions. Tout d'abord, le spectre d'absorption TD-DFTB n'implique ni états doublement excités, ni états de Rydberg. De plus, les transitions interbandes *via* le couplage spin-orbite ne sont pas incluses dans la propagation FSSH. Pour la partie Machine Learning, notons que leur précision réelle est évidemment déterminée par les données d'apprentissage. Ainsi, les modèles SchNet susmentionnés ont naturellement hérité de tous les avantages et inconvénients de l'approche TD-DFTB sous-jacente.

Dans l'ensemble, cette étude contribue à la compréhension de la photophysique des grands PAH et jette un éclairage sur la photostabilité des composés cata-condensés après absorption d'un photon UV. Il fournit également des informations sur les effets possibles de la taille et/ou de la forme du complexe moléculaire correspondant sur la dynamique de l'état excité. Ces résultats présentent un intérêt pour l'astrochimie, la chimie atmosphérique et les expériences de laboratoire concernant la spectroscopie laser atto- ou femto-seconde des PAH. Il convient de mentionner que les grands polyacènes ont des états excités de basse énergie dans la gamme UV-visible. Ainsi, ils peuvent être observés et attribués s'ils survivent au fort rayonnement UV. L'évaluation de la validité de la règle de Kasha pour les PAH est également importante dans ce contexte. En se référant plus particulièrement aux espèces du milieu interstellaire, les largeurs des bandes d'absorption électronique peuvent être liées aux échelles de temps de relaxation.[22,94] Les mécanismes de relaxation observés en fonction de la taille ou de la forme et les différentes populations d'état final pourraient contribuer à l'assignation des bandes diffuses. Enfin, il convient de mentionner la corrélation entre les échelles de temps de dépopulation et les écarts entre niveaux d'énergies qui a été rapportée pour toutes les molécules considérées dans ce travail. La partie méthodologique de cette étude est importante car elle vise à établir la possibilité d'utiliser les techniques de Machine Learning de pointe pour décrire la dynamique non adiabatique. Ceci est particulièrement intéressant en raison d'un intérêt croissant pour la dynamique des états excités de grands composés moléculaires.

Ainsi, si différents aspects de la dynamique moléculaire non adiabatique des PAH neutres ont été étudiés, les résultats de cette thèse sont encore limités par les méthodes de structure électronique sous-jacentes. L'approche TD-DFTB avec séparation de portée[128,259] ou l'analyse matricielle des densité de transition[89] sont des améliorations possibles pour étudier les

processus impliquant des états excités à transfert de charge. Il serait bien sûr intéressant d'étudier la dynamique des composés chargés, en particulier les composés cationiques en raison de la diversité des études théoriques et expérimentales disponibles. Ceci peut être réalisé avec une extension du formalisme TD-DFTB[145,146] existant pour les systèmes à couches ouvertes. D'autre part, l'approche DFTB-CI est également un outil prometteur qui peut être adapté pour étudier la dynamique non adiabatique des agrégats cationiques de grands PAH.[149]

De plus, l'approche FSSH de Tully mérite une certaine attention. Premièrement, le calcul des couplages non adiabatiques reste un défi majeur pour les méthodes à référence unique comme la réponse linéaire TD-DFTB. Il existe certaines ambiguïtés de mise en œuvre qui devraient être précisées à l'avenir. De plus, les corrections de décohérence couramment utilisées sont quelque peu *ad hoc* et il existe de nombreuses études traitant de ce problème.[260,261] Pour les phénomènes de relaxation lente impliquant des transitions interbandes, les états des triplets doivent être pris en compte *via* le couplage spin-orbite[262] tel qu'implémenté, par exemple, dans le logiciel SHARC.[194]

Enfin, les futures expériences pompe-sonde éclaireront davantage la dynamique ultra-rapide de grands composés. Il serait intéressant de rechercher si la différenciation théoriquement calculée de la relaxation électronique peut être observée dans les expériences. Afin de soutenir ces études, une description plus précise des processus non adiabatiques sera utile. Par exemple, l'approche MCTDH peut être couplée aux potentiels d'apprentissage automatique pour effectuer la dynamique quantique complète.[228] D'un autre côté, il a été montré que même la propagation TSH avec des quantités apprises par Machine Leraning peut fournir des résultats fiables, surtout si une méthode de structure électronique de plus haut niveau est utilisée pour entraîner un modèle minutieusement conçu.[81,82,210] Dernier point mais non des moindres, une évaluation critique de l'hypothèse PAH a besoin de plus d'informations provenant des observations astronomiques, des expériences de laboratoire et des calculs théoriques. Finalement, il se pourrait que le télescope spatial James Webb[263] avec sa remarquable précision spectroscopique, apporte une contribution cruciale à ce débat de longue durée, voire le clôture.

Acknowledgements

Indeed, it would be much harder to accomplish this PhD without the following people that I was lucky to meet along my journey.

I would like to thank my supervisors Fernand Spiegelman and Didier Lemoine. Firstly, they granted me an opportunity to perform this research in a rather flexible way, but more importantly, they were open to listen, discuss and help if needed. Acknowledgements go to all members of the Théorie group (LCAR) and the MAD group (LCPQ) for providing a stimulating and healthy working environment. In particular, I owe a lot to my collaborators Bruno Lepetit and Mathias Rapacioli. Mathias introduced me to the deMon developers community and kindly shared his handy programming tools. Discussions with Bruno were always insightful and followed by great questions inspiring to dive into more details. I am grateful to Aude Simon, Eric Michoulier, Nadine Halberstadt, Jérôme Cuny and Jean-Marc L'Hermite for their friendly attitude and useful suggestions. I would also like to thank professor Bjørk Hammer and members of his research group for kind hospitality during my secondment in Aarhus. This list would not be complete without Carole Lecinana and Christine Soucasse, who were around to help me work through the administrative part of research, and Catherine Stasiulis, who was very kind to answer my questions in great details. Finally, I acknowledge David Sanchez and the CALMIP team for smooth operation of the supercomputing facilities.

Thanks to Daria, my parents and grand-parents for their unconditional support. Thanks to my friends for constantly reminding that there is life beyond the office. Thanks to all the EUROPAH members for lots of memorable and fun moments that we managed to squeeze in the hectic schedule of project-related activities. Thanks to my previous co-workers Vladimir Kashcheev, Roman Rezaev and professor Vladimir Fomin for sharing their research experience during my undergraduate studies.

I acknowledge the European Union (EU) and Horizon 2020 generous funding awarded under the Marie Skłodowska-Curie action to the EUROPAH consortium, grant number 722346. Part of the theoretical simulations presented in this work have been carried out thanks to the HPC resources of the CALMIP supercomputing center (Grant P18019) in Toulouse.