Enhancing Multi-Scale Simulations of Charge and Exciton Transfer with Machine Learning Methods

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

von der KIT-Fakultät für Chemie und Biowissenschaften

des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M.Sc. Philipp Michael Dohmen

aus Aachen

1. Referent: Prof. Dr. Marcus Elstner
2. Referent: Prof. Dr. Alexander Schug
Tag der mündlichen Prüfung: 18. Juli 2022

Abstract

The theoretical investigation of charge and energy transfer phenomena requires the use of computationally demanding multi-scale simulations, which combine a quantum and a classical description. This work aims at the integration of data-driven machine learning techniques into the workflow of such simulations to accelerate the simulations by substituting the costly quantum-chemical part.

It is shown that simple and compact kernel ridge regression models are able to drive non-adiabatic molecular dynamics simulations by predicting the elements of the transfer Hamiltonian. Reference data from a semiempirical method was closely reproduced. However, these models did not lead to an acceleration and training set sizes were highly limited, which may prohibit the training of more complex and larger molecules.

In contrast, neural network models provide a significant boost in efficiency compared to a semiempirical reference and an even larger speedup for higher levels of theory, while at the same time setting no limit to training set size. Additionally, they allow for the simultaneous prediction of transfer Hamiltonian elements and their derivatives, which is necessary for explicit treatment of the relaxation process and correct re-scaling of atomic momenta with non-adiabatic coupling vectors.

Further, the methodology was extended to exciton transfer. The influence of short range effects via supermolecular couplings was investigated and a diabatization scheme for more accurate and reliable computations was implemented. The application of neural networks to exciton transfer in anthracene could reproduce experimental diffusion constants and showed highly localized transfer.

Finally, the developments of this thesis were combined and culminated in the application of exciton transfer simulations in the light-harvesting complex II (LH2) from purple bacteria. This biological complex contains chromophores, which are arranged in two rings (B800, B850). To date, only one study was able to perform a single simulation of 300 fs length. This was due to the tremendous computational cost of such simulations, which are lifted by the developed data-driven approach shown in this work. The transfer in both rings was simulated for 10 ps in 1000 trajectories each, with a feasible amount of resources. Excitons in the B800 ring were highly localized and transferred in discreet hops, while the B850 chromophores induced coherent transport and a spread of the exciton on multiple molecules. Estimated exciton diffusion constants for both rings appeared much higher compared to those of organic semiconducting materials. Now, large scale simulations, aiming at a comprehensive picture of the complete light-harvesting process in photosynthetic organisms from absorption to charge separation, are in reach.

Zusammenfassung

Die theoretische Untersuchung von Ladungs- und Energietransferphänomenen erfordert rechenintensive Multi-Skalen-Simulationen, die eine quantenmechanische und eine klassische Beschreibung kombinieren. Diese Arbeit strebt eine Integration von datengesteuerten Methoden des maschinellen Lernens in den *workflow* solcher Simulationen an. Indem der kostspielige quantenchemische Teil ersetzt wird, können diese beschleunigt werden.

Es wird gezeigt, dass einfache und kompakte Kernel-Regressionsmodelle in der Lage sind, nicht-adiabatische Molekulardynamiksimulationen durch die Vorhersage der Elemente des Transfer-Hamiltonian voranzutreiben. Referenzdaten einer semiempirischen Methode wurden genau reproduziert. Diese Modelle führen jedoch nicht zu einer Beschleunigung und die Menge der Trainingsdaten ist deutlich eingeschränkt. Dies könnte das Training komplexerer und größerer Moleküle erschweren.

Im Gegensatz dazu bieten neuronale Netze eine erhebliche Effizienzsteigerung im Vergleich zu einer semiempirischen Referenzmethode und eine noch größere Beschleunigung für genauere quantenmechanische Methoden. Gleichzeitig besteht keine Begrenzung für die Größe der Trainingsdatenmenge. Außerdem ermöglichen die Modelle die gleichzeitige Vorhersage von Transfer-Hamiltonianelementen und deren Ableitungen, was für die explizite Behandlung des Relaxationsprozesses und die korrekte Neuskalierung der atomaren Impulse mit nicht-adiabatischen Kopplungsvektoren notwendig ist.

Darüber hinaus wurde die Methodik auf den Exzitonentransfer ausgeweitet. Der Einfluss von Kurzstreckeneffekten in supermolekularen Kopplungen wurde untersucht und ein Diabatisierungsschema für genauere und zuverlässigere Berechnungen wurde implementiert. Die Anwendung von neuronalen Netzen auf den Exzitonentransfer in Anthracen konnte die experimentell ermittelten Diffusionskonstanten reproduzieren und zeigte einen stark lokalisierten Transfer.

Schließlich wurden die Entwicklungen dieser Arbeit kombiniert und resultierten in der Anwendung von Exzitonentransfer-Simulationen auf den Lichtsammelkomplex II (LH2) von Purpurbakterien. Dieser biologische Komplex enthält Chromophore, die in zwei Ringen angeordnet sind (B800, B850). Bisher war es nur in einer Studie möglich, eine einzige Simulation von 300 fs Länge durchzuführen. Dies ist auf die enormen Rechenkosten solcher Simulationen zurückzuführen, die durch den in dieser Arbeit entwickelten datengesteuerten Ansatz verringert werden. Der Transfer in beiden Ringen wurde für 10 ps in jeweils 1000 Trajektorien simuliert, was mit einem vertretbaren Aufwand an Ressourcen realisierbar war. Die Exzitonen im B800-Ring waren stark lokalisiert und wurden in diskreten Sprüngen übertragen, während die B850-Chromophore einen kohärenten Transport und

eine Delokalisierung des Exzitons über mehrere Moleküle ermöglichen. Die abgeschätzten Diffusionskonstanten für den Transfer in beiden Ringen waren deutlich größer als die von organischen Halbleitermaterialien. Jetzt werden groß angelegte Simulationen mit dem Ziel möglich, den gesamten Lichtsammelprozess in photosynthetisch aktiven Organismen von der Absorption bis zur Ladungstrennung aufzuklären.

Acknowledgment

I would like to thank everyone who has contributed to the success of my doctoral thesis.

Firstly, thank you to Marcus Elstner for your scientific guidance, your encouragement and optimism throughout my PhD. I am deeply grateful to Tomáš Kubař, who lured me into this group with his interesting lecture on molecular dynamics. He was helpful in uncountable ways, had answers to every topic I came up with and is a full technical support team in one person.

It was a pleasure to closely collaborate with Mila Krämer, who introduced me to machine learning and had an open ear in case of repeated throwbacks. I always enjoyed our discussions, not only on scientific topics. Thank you to Monja Sokolov, Farhad Ghalami and Beatrix Bold for being great collaborators and friends, and for always sharing a good laugh.

My appreciation to Weiwei Xie and Daniel Wehl, who introduced me to this nasty thing we call propagation code. Without the kind soul of this group, Sabine Holthoff, I would have drowned in the depth of university bureaucracy.

I am extremely glad for the warm welcome at the beginning of my master's thesis by all members of the former "Girls' Tower", namely Beatrix Bold, Mila Krämer, Violetta Schneider and Franziska Wolff. I had a great time during our refreshing ice cream and coffee breaks as well as our sunset barbecues on the rooftop of the inorganic chemistry building.

Thank you to all TCB and RTG2450 members for your help, discussions and events during my PhD time.

My heartfelt gratitude to my family for their emotional and financial support, which allowed me to study and provided me with much needed breaks from science.

Finally, I am greatly indebted to my girlfriend Alexandra Friesen for her neverending love and support, especially during the long final days and nights of writing this thesis. You always had my back.

Contents

Part I.

Introduction

1. Introduction

Theoretical chemistry, computational chemistry and molecular modeling are related and overlapping subdomains in the overarching field of chemistry. They are not clearly delineated from each other and have in common that they deal with the development, improvement and application of theories and models for the description of chemical structures and their properties. These techniques may serve as a lens to gain insights into chemical and physical phenomena, that may not be easily accessible or accessible at all through experiments. In this way, computational chemistry can aid in the comprehension and interpretation of chemical processes.[1] Additionally, it may serve as a predictive tool to replace or shorten experimental efforts, e.g. by high-throughput screening of chemical structures with regard to specific properties on the quest for novel or optimized chemical compounds.[2,3]

The importance and usability of computational methods in chemistry has emerged and increased with the computing power and affordability of modern hardware, e.g. central processing units (CPUs), graphics processing units (GPUs) and the use of specialized high-performance computing (HPC) clusters. This was accompanied by progress in software development, for instance implementations of new and more accurate computational methods.[1]

In practice, calculations and simulations of chemical systems are limited by their computational cost. There is an interplay between accuracy and reachable time and length scales on one hand and affordability of calculations on the other hand. The more precise the method, the greater the computational demands. Molecular mechanics (MM) methods are able to describe huge chemical systems with thousands of atoms on long time scales efficiently. Even bigger system sizes and time scales can be reached with coarse-grained models, although information on the atomic scale is lost. However, classical physics may not be sufficiently accurate to describe the relevant phenomena or to precisely predict physical observables of the system. Additionally, these methods are not able to capture processes such as chemical reactions or the interaction of matter and light, which limits their applicability.[4]

Contrastingly, these phenomena can be captured by quantum mechanical (QM) methods, which are more exact. Semi-empirical QM methods solve the Schrödinger equation, but introduce approximations and empirically derived parameters. They can treat a few hundreds of atoms on short time scales. *Ab initio* methods calculate molecular systems from first principles, meaning the only input are physical constants and the coordinates of the system. Computational cost grows with the accuracy of the level of theory.[1]

Some chemical or physical phenomena involve different time and length scales, which need to be bridged in order to be accurately described. Charge (CT)[5] and exciton transfer (ET)[6] processes, as investigated in this work, require such multi-scale approaches. For this purpose, the system is divided into an active region of interest, where e.g. QM methods are employed and the transfer takes place, while the remainder is treated with a cheaper MM method. As in similar hybrid QM/MM approaches, the simulation is limited by the demanding QM calculations. The efficiency can be improved by introducing approximations and by employing semi-empirical QM methods. Still, the system sizes and simulation lengths are restricted.

Data-driven approaches, more specifically machine learning (ML) methods, can make use of the fact that the quantum chemical calculations are highly repetitive, i.e. the conformations sampled during such simulations are very similar. This motivates the substitution of the costly QM methods by more efficient ML models. By training an ML model on a small but representative data set, for which the relevant electronic parameters have been calculated beforehand, the computationally simpler ML model could then be used to drive the transfer simulations.[7,8]

This thesis aims at an integration of machine learning techniques into the multi-scale workflow of charge and exciton transfer simulations. Charge transfer in simple organic semiconductors (OSCs) has been studied with non-adiabatic molecular dynamics (NAMD) simulations[9–12] and serves as a test case for the application of ML models in this work. The ML/MM approach is examined for its accuracy and the desired increase in efficiency. Furthermore, the same methodology is extended to exciton transfer simulations, whose computational cost is several orders of magnitude higher compared to charge transfer simulations. To this end, organic semiconductors again function as proof of principle examples. Finally, with the tested methodology at hand, exciton transfer in biological light-harvesting complexes becomes feasible. It was not possible to investigate the latter systems in detail with the methods available so far.

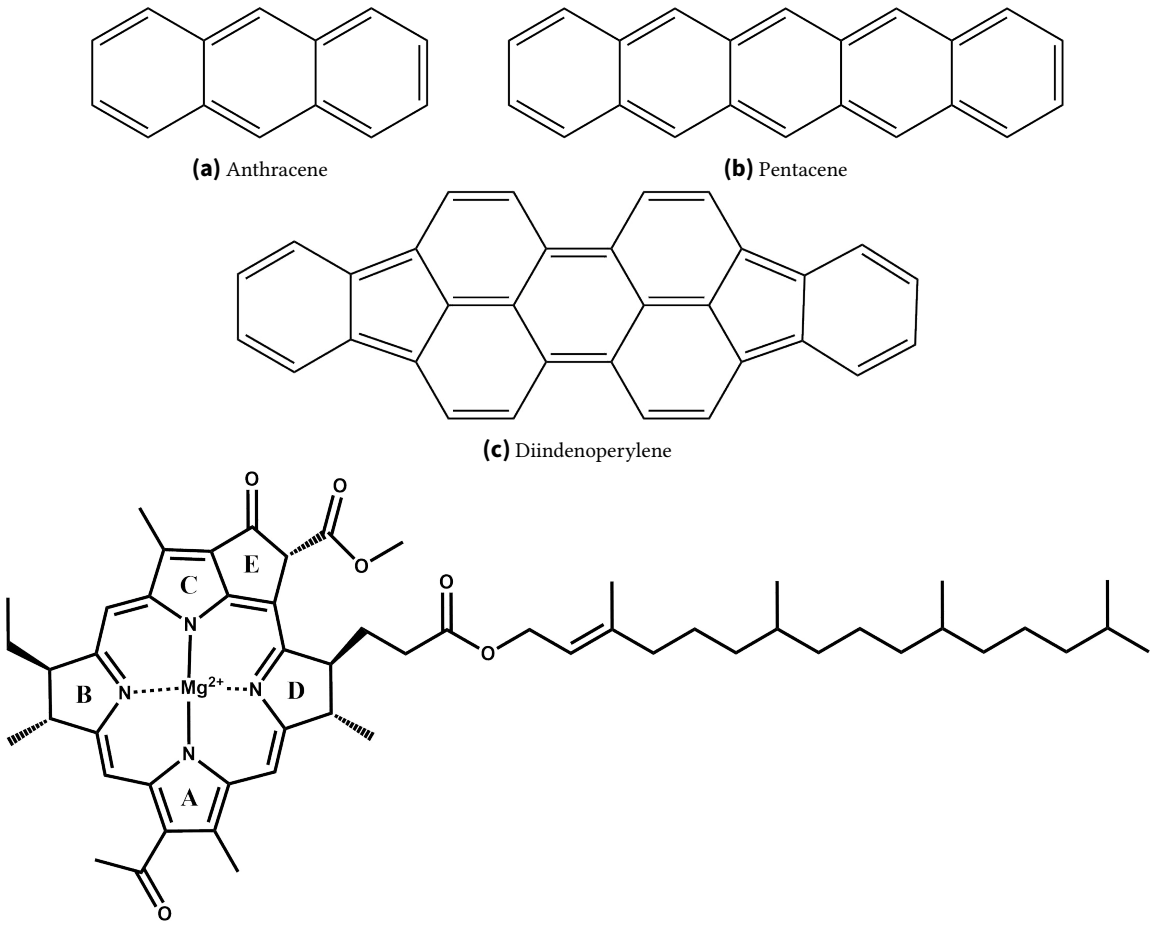
The outline of this thesis is as follows: At first, the theories and methods, that are relevant for this work are introduced. Among them are MM and QM methods for the calculation of energies and forces of molecular structures (chapter 2), as well as molecular dynamics (MD) methods for the simulation of ensemble and time dependent properties (chapter 3). This also includes the simulation approach for charge and exciton transfer. Additionally, ML models for the acceleration of MD simulations are introduced (chapter 4).

Chapter 5 covers the application of kernel ridge regression (KRR) models for charge and exciton transfer simulations in anthracene (Figure 1.1a), a well studied organic semiconductor. The elements of the transfer Hamiltonian are learned and used to drive the transfer. The observed restrictions when employing KRR models and their overcoming are addressed in chapter 6 for charge transfer. Here, neural network (NN) models are employed to learn transfer Hamiltonian elements and diagonal derivatives for anthracene and pentacene (Figure 1.1b).

The following chapters extend the integration of ML models into simulations of exciton transfer. Firstly, in chapter 7, the calculation of supermolecular excitonic couplings is

addressed and a more advanced scheme is implemented. Secondly, anthracene and diindenoperylene (Figure 1.1c) are candidates, on which NN driven exciton transfer simulations are examined. Moreover, the learning is extended to off-diagonal derivatives (chapter 8) of the transfer Hamiltonian. Eventually, the efficiency improvements are utilized in chapter 9 to perform exciton transfer simulations in the biological light-harvesting complex II (LH2), a huge molecular system containing several bacteriochlorophyll a (BChl a) molecules (Figure 1.1d).

Finally, the most important results and findings as well as open questions are summarized in chapter 10. This last chapter closes with an outlook on future work.



(a) Anthracene

(b) Pentacene

(c) Diindenoperylene

(d) Bacteriochlorophyll a

Figure 1.1.: Molecules investigated in this work.

Part II.

Theory and Methods

2. Calculation of Energies

This chapter introduces the various methods to describe the state of a molecule and to calculate the corresponding energy, forces and other properties of a particular molecular configuration. These methods are based on either classical or quantum mechanics. Because of differences in their nature, accuracy and computational cost, they are suited for different applications.

2.1. Molecular Mechanics

Molecular mechanics (MM) makes use of classical physics to model molecular systems. The smallest unit or particle to comprise such a molecular system in this approach is an atom. Within MM, the potential energy, which is related to the interactions between these particles, is given by an analytic function of atom coordinates, called the force field (FF). In the following, the all-atom force field is illustrated, as used in this work. There are also other types of force fields for applications to larger molecular complexes, e.g. the united-atom or coarse-grained approach, which merge hydrogen atoms to their heavy atomic partners or even combine several heavy atoms into one particle.[1,13]

2.1.1. All-Atom Force Fields

In an all-atom force field, a molecular structure is represented by a "ball and spring" model. Atoms are considered to be points having a defined position, mass and charge. Electrons are not taken into account as individual particles, but rather are merged into their respective nuclei. Bonds between atoms are modeled as springs with Hooke's law, assigned a length and a stiffness. Analyzing molecular structures of different kinds reveals the composition of structurally similar atoms and bonds (functional groups). For example all C–H single bonds have a similar length and a similar bond oscillation frequency, even in very different molecules. This leads to the principle of a force field parametrization and the introduction of atom types, depending on the respective element types and bonding partners.

The force field itself describes the potential energy of all atoms in the form of a sum over all atomic interactions (e.g. as in Equation 2.1). The interactions can be divided into bonding and non-bonding terms, the former being represented by bond stretching, angle torsion and dihedral torsion, the latter by electrostatic and van-der-Waals interactions.

Bond stretching and angle torsion are both modeled with a harmonic potential, while for the dihedral torsion a sum of periodic potentials (e.g. cosines) is needed. The electrostatic interaction (Coulomb potential) is a inverse first-power potential. The van-der-Waals interaction consists of two contributions, the (short-ranged) Pauli-repulsion, which is due to the Pauli exclusion principle, and the attractive (long-ranged) London-dispersion, caused by interactions of induced and instantaneous dipoles. It is modeled with the Lennard-Jones (12-6) potential.

$$
\begin{aligned}
V_{\text{tot}}^{\text{MM}} &= \underbrace{V_{\text{bond}} + V_{\text{angle}} + V_{\text{dihedral}}}_{\text{bonding}} + \underbrace{V_{\text{Coulomb}} + V_{\text{Lennard-Jones}}}_{\text{non-bonding}} \\
&= \frac{1}{2} \sum_i k_i^b \left(r_i - r_i^0\right)^2 \\
&+ \frac{1}{2} \sum_j k_j^a \left(\theta_j - \theta_j^0\right)^2 \\
&+ \frac{1}{2} \sum_k \sum_l V_k^{(l)} \left(1 + \cos\left[n\vartheta_k - \vartheta_k^{(0)n}\right]\right) \\
&+ \sum_m \sum_{n>m} \left(\frac{1}{4\pi\epsilon_0} \frac{q_m q_n}{r_{mn}} + 4\epsilon_{mn} \left[\left(\frac{\sigma_{mn}}{r_{mn}}\right)^{12} - \left(\frac{\sigma_{mn}}{r_{mn}}\right)^{6}\right]\right)
\end{aligned}
\tag{2.1}
$$

The parameters required within a force field are the force-constants (k) and equilibrium distances/angles (r^0, θ^0, ϑ^0) for the different bonding potentials, as well as atomic partial charges (q), depth (ϵ) and distance (σ), where the Lennard-Jones potential is zero. These parameters are empirically derived, e.g., from experiments (spectroscopy, calorimetry) or *ab initio* calculations. The atomic partial charges are tailor-made for specific molecules or classes of molecules e.g. DNA, proteins, lipids or small organic molecules, which limits their transferability to other classes of molecules. The functional form of the potentials and the parameters can differ between force fields.

The non-bonded interactions lead to the N^2-scaling of the computational cost with the number of atoms, which can be optimized further. This allows for simulation of huge molecular systems with up to millions of atoms and/or on long time scales of up to milliseconds. But the method also comes with limitations. The parameters are derived for a specific electronic state and drastic changes can not be represented within a single parametrization of a force field. Additionally, due to the harmonic potential used to describe the chemical bonds, the quantum nature inherent to electrons is neglected. As a consequence, formation and breakage of bonds, interactions with light or other phenomena, that involve the change of electronic states, cannot be simulated.[1,13]

2.2. Quantum Chemistry

The discovery of the wave-particle duality and the quantization of energy led to the development of a new theory for the description of matter: quantum mechanics (QM). The fundamental aspects of this theory as well as the methods relevant for this work are introduced in this chapter.

2.2.1. Foundations of Electronic Structure Theory

Quantum mechanics as a theory cannot be derived rigorously as it comes with a few postulates: The state of a quantum system or particle is described by the wave function, which does not have a physical meaning. Nevertheless, the square of the wave function has a physical meaning, which is the probability of finding the system in a given state. Any physical observable can be obtained by applying the corresponding operator to the wave function. The behavior of such a system in space and time is described by the time-dependent Schrödinger equation

$$\hat{H} \left| \Psi(\vec{r}, \vec{R}, t) \right\rangle = i\hbar \frac{\partial}{\partial t} \left| \Psi(\vec{r}, \vec{R}, t) \right\rangle , \tag{2.2}$$

with the Hamiltonian operator $\hat{H}$, the wave function for electrons and nuclei Ψ, with their positions $\vec{r}$ and $\vec{R}$. i is the imaginary unit and $\hbar$ the reduced Planck's constant. If the potential energy and thus the Hamiltonian of the system is not explicitly time-dependent, the wave function can be separated into parts depending on space or time variables only. The result of this separation is the time independent or stationary Schrödinger equation (Equation 2.3), and the wave function takes the form of a standing wave for a stationary state.

$$\hat{H} \left| \psi(\vec{r}, \vec{R}) \right\rangle = E \left| \psi(\vec{r}, \vec{R}) \right\rangle \tag{2.3}$$

This equation is an eigenvalue problem with the eigenvalues being the energies of the system (E_i) and the eigenvectors or eigenstates being the corresponding wave functions ($\psi_i(\vec{r}, \vec{R})$). If the Hamiltonian is known, solving the Schrödinger equation yields these energies and states. The only interaction relevant at the atomic or molecular level is the electromagnetic interaction between nuclei-nuclei, electron-electron and electron-nuclei, which can be described by the Coulomb potential. The Hamiltonian operator for a molecule (in atomic units) reads

$$\hat{H} = -\frac{1}{2} \sum_i \Delta_i - \frac{1}{2} \sum_k \frac{1}{M_k} \Delta_k - \sum_i \sum_k \frac{Z_k}{r_{ik}} + \sum_{k<l} \frac{Z_k Z_l}{r_{kl}} + \sum_{i<j} \frac{1}{r_{ij}} , \tag{2.4}$$

with the nuclear mass (M) and charge (Z), the distance between two particles i and j (r_{ij}), the Laplace operator (Δ) and indices i, j for electrons and k, l for nuclei, respectively. The

Hamiltonian consists of terms for the kinetic energy of electrons and nuclei as well as the potential energy of interactions between electrons-nuclei, nuclei-nuclei and electrons-electrons.

Assuming that the nuclei are infinitely heavier than the electrons, the motion of the latter is much faster compared to the former. In fact, the mass difference is high enough to expect an instantaneous adjustment of the electronic positions to the nuclear geometry. This is known as the Born-Oppenheimer approximation[14], which leads to a separation of nuclear and electronic wave functions and allows for the electronic Schrödinger equation (Equation 2.5) to be solved for a given set of stationary nuclei. Then, the electronic wave function only depends on the electronic coordinates.

$$\hat{H}_e \left|\varphi(\vec{r})\right\rangle = E_e \left|\varphi(\vec{r})\right\rangle \tag{2.5}$$

By solving this equation for a large set of nuclear arrangements, the electronic energy forms a hypersurface as a function of the nuclear coordinates, which is called potential energy surface (PES). As a consequence of the Born-Oppenheimer approximation, the kinetic energy of the nuclei can be neglected, while the nuclei-nuclei repulsion becomes a constant term. The Hamiltonian operator for the electronic structure of a molecule (Equation 2.6) can be simplified.

$$\hat{H}_e = \hat{T}_e + \hat{V}_{en} + \hat{V}_{ee} \tag{2.6}$$

There are a lot of different electronic structure methods, which try to solve the electronic Schrödinger equation (Equation 2.5). A basic principle that can be found in most methods is the variational principle (Equation 2.7). It states that an arbitrarily chosen trial wave function always leads to a higher (or equal) energy than that of the ground state. The wave function of the actual ground state ($\varphi_0(\vec{r})$) is the one associated with the lowest possible energy of the system (E_0). Hence, a trial wave function can be used as a starting point, which is subsequently improved through variation. Improvement in this case means, that the new wave function leads to a lower energy.

$$\left\langle \varphi_{\text{trial}}(\vec{r}) \middle| \hat{H}_e \middle| \varphi_{\text{trial}}(\vec{r}) \right\rangle = E_{\text{trial}} \geq E_0 = \left\langle \varphi_0(\vec{r}) \middle| \hat{H}_e \middle| \varphi_0(\vec{r}) \right\rangle \tag{2.7}$$

The construction of a trial wave function is usually done via the linear combination of atomic orbitals (LCAO) approach (Equation 2.8). The wave function is constructed in the shape of a Slater determinant, which contains one-electron functions, so called molecular orbitals (MO) Ψ.

$$\Phi_i = \sum_i c_i \chi_i \tag{2.8}$$

The latter are formed as linear combinations (superposition) of atomic orbitals (AO) χ, which are called a basis set. Solving the Schrödinger equation now becomes equivalent

to minimizing the energy of the trial wave function constructed from MOs by varying the AO-coefficients.[1,13] The electronic structure methods used in this work are described throughout the next chapters.

2.2.2. Density Functional Theory

In contrast to wave function based methods such as Hartree-Fock (HF) or Coupled Cluster (CC), Density Functional Theory (DFT) is based on the electron density (Equation 2.9) for the description of the electronic structure of a molecular system. The electron density is a probability density, stating the probability of finding an electron in a volume element. The dimensionality is reduced from $4N$ spatial and spin variables for the wave function to only 4 variables for the density. N refers to the number of electrons.

$$\rho(\vec{r}) = |\Phi(\vec{r})|^2 \tag{2.9}$$

Using the electron density as a possible representation of the system is intuitive because of its properties. Integrating the electron density yields the number of electrons in the system. The position of the nuclei can be identified by the cusps in the electron density. Their height contains information about the corresponding nuclear charge.

$$E = E\left[\rho\left(\vec{r}\right)\right] \quad \text{and} \quad E_{\text{trial}}\left[\rho_{\text{trial}}\left(\vec{r}\right)\right] \geq E_0\left[\rho_0\left(\vec{r}\right)\right] \tag{2.10}$$

Hohenberg and Kohn[15] provided proof for this in two theorems (Equation 2.10). The first demonstrates, that the energy of the ground state can be formulated as a functional of the electron density, while the second is similar to the variational principle, associating the ground state density with the lowest possible energy. The total energy in terms of functionals of the density consists of the kinetic energy of the electrons, as well as the interaction energies of electrons-nuclei and electrons-electrons.

$$E\left[\rho\right] = T_e\left[\rho\right] + V_{en}\left[\rho\right] + V_{ee}\left[\rho\right] \tag{2.11}$$

The last term can be split into the classical electrostatic Coulomb repulsion ($J\left[\rho\right]$) and the remainder of non-classical origin, the correction for self-interaction, exchange and correlation. Kohn and Sham[16] achieved a major breakthrough in the practical use of DFT by offering a way to effectively calculate the contribution of the kinetic energy. They considered a fictitious reference system of non-interacting electrons in an effective potential (V_{eff}) similar to HF, for which the kinetic energy ($T_{\text{ref}}\left[\rho\right]$) could be computed easily by reintroducing orbitals. The effective potential is set in a way that the density of the reference system is equal to that of the real interacting system

$$\rho = \rho_{\text{ref}} = \sum_i \left|\phi_i\right|^2 . \tag{2.12}$$

The (orthonormal) orbitals can be calculated within the variational principle using the Kohn-Sham equations (Equation 2.13). This has to be performed in an iterative process until self-convergence, because the external potential already depends on the density and thus on the orbitals.

$$\left(-\frac{1}{2}\Delta + V_{\text{eff}}\right)\phi_i = \epsilon_i \phi_i \tag{2.13}$$

The solution by Kohn and Sham describes the main part of the kinetic energy, which only misses a small contribution. The stated terms of classical origin and the main part of the kinetic energy can be formulated exactly, while all unknown contributions (correction for self-interaction and kinetic energy, exchange and correlation) are combined into the exchange-correlation functional ($E_{xc}[\rho]$).

$$E[\rho] = T_{\text{ref}}[\rho] + J[\rho] + V_{en}[\rho] + E_{xc}[\rho] \tag{2.14}$$

If the correct exchange-correlation functional was known, the exact ground state energy of a molecular system could be calculated. However, the explicit form of this functional is not known, so that approximations must be introduced for this otherwise exact theory. The crucial task in the development and advancement of DFT is to find better forms for the exchange-correlation functional, which can include empirically derived parameters.[1,13,17]

Simple formulations of functionals make use of the local density (LDA) or generalized gradient (GGA) approximation. The former is based on the uniform electron gas and the energy is assumed to only depend on the local value of the electron density. GGA class functionals additionally include the dependence of the energy on the gradient of the density at every point. Today, hybrid functionals are popular, as they have proven to be a versatile method for calculating many properties and often have a good balance between accuracy and efficiency. They contain a fraction of exact Hartree-Fock (HF) exchange in the exchange part of the functional.[13] Finally, long-range corrected (LC) or range-separated functionals have been developed. Here, exact exchange is added with a distance dependence and some severe problems in DFT are improved.[18]

2.2.3. Density Functional Tight Binding

Density functional tight binding (DFTB) is a semi-empirical QM method based on DFT. The term tight-binding comes from models of solid-state physics, where core electrons are assumed to be tightly bound to the atomic nucleus. Hence, they do not affect the electronic structure of their surroundings much and only valence electrons are treated explicitly. Additionally, a minimal basis for the valence electrons in an LCAO approach (Equation 2.15) is introduced.

$$|\Psi_i\rangle = \sum_{\mu} c_{\mu_i} |\phi_\mu\rangle \tag{2.15}$$

The atomic orbitals (ϕ) are obtained from single atom DFT calculations, which employ an additional harmonic potential for a confined basis. The total energy in DFTB is determined from a Taylor expansion of the total DFT energy (Equation 2.14) in terms of neutral atomic reference densities (ρ^0) and their fluctuation ($\delta\rho$) as a representation for the true density minimizing the energy (Equation 2.16).[19]

$$
\begin{aligned}
E^{\mathrm{DFTB}}\left[\rho^0+\delta\rho\right] &= E^1+E^0+E^2+E^3+\cdots \\
E^1 &= \sum_i n_i \left\langle \Psi_i \middle| -\frac{\Delta}{2} + V_{ne} + \int \frac{\rho^{0\prime}}{|r-r'|}\mathrm{d}r' + V_{\mathrm{XC}}\left[\rho^0\right] \middle| \Psi_i \right\rangle \\
E^0 &= -\frac{1}{2}\iint \frac{\rho^0\rho^{0\prime}}{|r-r'|}\mathrm{d}r\mathrm{d}r' - \int V_{\mathrm{XC}}\left[\rho^0\right]\rho^0\mathrm{d}r + E_{\mathrm{XC}}\left[\rho^0\right] + E_{\mathrm{nn}} \\
E^2 &= \frac{1}{2}\iint \left(\frac{1}{|r-r'|} + \left.\frac{\delta^2 E_{\mathrm{XC}}[\rho]}{\delta\rho\delta\rho'}\right|_{\rho^0,\rho^{0\prime}}\right)\delta\rho\delta\rho'\mathrm{d}r\mathrm{d}r' \\
E^3 &= \frac{1}{6}\iiint \left.\frac{\delta^3 E_{\mathrm{XC}}[\rho]}{\delta\rho\delta\rho'\delta\rho''}\right|_{\rho^0,\rho^{0\prime},\rho^{0\prime\prime}} \delta\rho\delta\rho'\delta\rho''\mathrm{d}r\mathrm{d}r'\mathrm{d}r''
\end{aligned}
\tag{2.16}
$$

Depending on the truncation of this Taylor series, a hierarchy of DFTB can be derived: First order DFTB (DFTB1)[20,21], second order self-consistent charge DFTB (SCC-DFTB2 or just DFTB2)[22] and third order DFTB (DFTB3)[23]. The energy of DFTB1 (zeroth and first order term) is further approximated as

$$
E_{\mathrm{DFTB1}} = E^1 + E^0 = \sum_i n_i\epsilon_i + \frac{1}{2}\sum_{ab} V_{ab}^{\mathrm{rep}}\,. \tag{2.17}
$$

The first term is often denoted the electronic energy. The eigenvalues (with occupation numbers n) are calculated in a non-self-consistent way from Kohn-Sham equations (Equation 2.18).

$$
\sum_\nu c_{\nu i}\left(H^0_{\mu\nu} - \epsilon_i S_{\mu\nu}\right) = 0 \tag{2.18}
$$

For the Hamiltonian and overlap matrices (Equation 2.19) only two-center integrals are evaluated, the elements of both matrices are computed once and then tabulated. Therefore, no integrals have to be evaluated during runtime, significantly speeding up DFTB calculations. DFTB is up to three orders of magnitude faster than conventional DFT.

$$
H^0_{\mu\nu} = \left\langle \phi_\mu \middle| H^0 \middle| \phi_\nu \right\rangle \qquad S_{\mu\nu} = \left\langle \phi_\mu \middle| \phi_\nu \right\rangle \tag{2.19}
$$

All other terms (third line in Equation 2.16) are merged into a repulsive energy term ($E^0 = E_{\mathrm{rep}}$), which is approximated as a sum of atomic pair potentials on atoms a and b and is obtained by fitting e.g. to reference DFT calculations.[24]

For molecular systems where charge transfer between atoms plays a role, e.g. heteroatomic molecules with significant difference in electronegativity, higher order terms must be included. The second order contribution is contained in DFTB2 as

$$E_{\mathrm{DFTB2}} = E_{\mathrm{DFTB1}} + \frac{1}{2}\sum_{ab}\Delta q_a \Delta q_b \gamma_{ab}\,, \tag{2.20}$$

incorporating charge fluctuations on atoms (Equation 2.21).

$$\Delta q_a = q_a - q_a^0 \tag{2.21}$$

The additional term essentially is a Coulomb interaction between atomic partial charges (Δq_a) at long distances and electron-electron interaction within one atom at short distances. The gross charges q_a are obtained via Mulliken population analysis from the MO coefficients and the charge of the core q_a^0 is subtracted. The MO coefficients are obtained from the Hamiltonian, which already involves the partial charges. Therefore, a self-consistent charge formalism must be introduced.[22] The performance of DFTB can be further improved by adding even higher order terms.

Standard DFTB is parameterized with a GGA functional. Thus, DFTB inherits some of the problems associated with GGA DFT, e.g. the self-interaction error, which severely affects the quality of several calculated properties. In Hartree-Fock, however, the interaction of an electron with itself does not occur. Therefore, long-range corrected functionals have been developed for DFT to reduce this error and have been adopted into the DFTB formalism. The electron-electron interaction is split into two parts, a short- and a long-range part (Equation 2.22).

$$\frac{1}{r_{12}} = \frac{1-\mathrm{erf}\,(\omega r_{12})}{r_{12}} + \frac{\mathrm{erf}\,(\omega r_{12})}{r_{12}} \tag{2.22}$$

The empirical range separation parameter ω has to be optimized through fitting to data from experiments or higher level of theory. It defines the splitting between the short- and long-range parts. The long-range part approaches exact HF exchange at long distances, while a pure density functional is employed at short distances. The correlation contribution is still evaluated by means of a GGA functional.

$$E_{\mathrm{XC}} = E_{\mathrm{C}}^{\mathrm{GGA}} + E_{\mathrm{X}}^{\omega,\mathrm{LDA}} + E_{\mathrm{X}}^{\omega,\mathrm{HF}} \tag{2.23}$$

The use of LC functionals leads to the correct $-\frac{1}{r}$ asymptotic behavior of the Kohn-Sham potential.[25,26]

For the calculation of excited states and their properties, the time-dependent linear response extension to DFTB (TD-DFTB) has to be used. This approach does not calculate the excited state directly, but rather properties of this state are obtained as a linear response. The

linear response in this case is the interaction with electromagnetic waves, mediating the transition from the ground state to an excited state. It can be seen as a small perturbation of the electronic structure.

Therefore, a TD-DFTB calculation is carried out on top of a DFTB ground state calculation. In a first step, the single particle energies and orbitals of the ground state are calculated. Afterwards, the pseudo-eigenvalue problem of the Casida equation (Equation 2.24)[27] is solved.

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{A} \end{pmatrix} \begin{pmatrix} \vec{X} \\ \vec{Y} \end{pmatrix} = \Omega \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \vec{X} \\ \vec{Y} \end{pmatrix} \tag{2.24}$$

Ω denotes the excitation energies, while A and B are matrices, whose elements are calculated from the energy differences of the Kohn-Sham orbitals and the linear response in terms of Coulomb and exchange-correlation contributions.[28,29]

2.3. Hybrid Approaches

The potential energy of a molecular system can be used to calculate forces on the nuclei, which are necessary to perform molecular dynamics simulations. Different methods, e.g. molecular mechanics force fields or one of the various quantum-chemical methods, may be employed to derive the potential energy. MM approaches are efficient, but less accurate and quantum effects, e.g. bond forming and breaking or transitions between electronic states, are not treated. Thus, large molecular systems can be simulated on long time scales, but no chemical reactions and interactions with electromagnetic waves can be described. In contrast, QM methods are able to simulate quantum effects. However, they come with high computational cost, which prohibits large system sizes and time scales. One way to overcome this is the application of hybrid QM/MM approaches, which combine both methods. The system is divided into different regions, where the energies and forces on nuclei are obtained by different methods. Usually, only a small region with a few nuclei is of primary interest and must be treated quantum-mechanically, while the entire environment is sufficiently treated by a force field.

$$E_{\text{tot}} = E_{\text{QM}} + E_{\text{MM}} + E_{\text{QM/MM}} \tag{2.25}$$

To account for the coupling between QM and MM regions (Equation 2.25), different embedding schemes have been developed. The easiest form in terms of computational effort is mechanical embedding, where the QM and MM densities are completely independent from each other. The interaction between them is treated purely by a non-bonding force field term (Equation 2.26), similar to the electrostatic and Van-der-Waals interaction in Equation 2.1. The partial charges of nuclei (i) in the QM region come directly from the QM calculation and are updated in every step. Meanwhile, the force field charges for atoms (m) in the MM region are kept fixed.

$$E_{\mathrm{QM/MM}} = \sum_{i}^{N_{\mathrm{QM}}} \sum_{m}^{N_{\mathrm{MM}}} \left(\frac{q_i q_j}{r_{im}} + 4\epsilon_{im} \left(\left(\frac{\sigma_{im}}{r_{im}} \right)^{12} - \left(\frac{\sigma_{im}}{r_{im}} \right)^{6} \right) \right) \tag{2.26}$$

This scheme completely neglects polarization of both regions. The electronic embedding scheme includes the polarization of the electronic structure of the QM region, which is caused by the MM region. Therefore, an additional term in the QM Hamiltonian is introduced (Equation 2.27), while the interaction of nuclei in the QM and MM region is kept within the QM/MM term.

$$H'_{\mathrm{QM/MM}} = H_{\mathrm{QM/MM}} - \sum_{j}^{n_{\mathrm{QM}}} \sum_{m}^{N_{\mathrm{MM}}} \left(\frac{q_m}{r_{jm}} \right) \tag{2.27}$$

In polarized embedding schemes, the mutual polarization of both regions is included via a polarizable force field. Therefore, both the QM and MM calculations have an iterative nature, which highly increases the computational cost.

Beside the mentioned QM/MM approaches, hybrid QM/QM approaches are also possible for small systems. Moreover, the system can be partitioned into more than two regions, e.g. QM1/QM2/MM, employing different levels of theory.[13] An implementation of this is the ONIOM (Our own n-layered Integrated molecular Orbital and Molecular mechanics) method.[30]

3. Molecular Dynamics

The previous chapter introduced theories and methods for the description of molecular configurations in terms of classical and quantum mechanics. Energies and forces from these methods can be used to evolve a molecular system in time, going beyond the picture of static molecules. This allows the dynamics of the system to be studied, which often goes hand in hand with its function. In addition, time-dependent and thermodynamic properties can be calculated.

3.1. Adiabatic Molecular Dynamics Simulations

In the standard case of molecular dynamics (MD) simulations, the Born-Oppenheimer approximation as mentioned in the previous chapter is considered valid. This approximation usually works well for heavy and slowly moving particles like atomic nuclei, whose movement can be decoupled from the movement of the electrons. Thus, the nuclei evolve on a single PES, which is associated with a single electronic state (an adiabatic state), usually the ground state. In standard MD simulations, the system is propagated classically, i.e. Newtons second law (Equation 3.1) is used as the equation of motion.

$$\vec{F}_i = m_i \vec{a}_i = m_i \frac{d^2 \vec{r}_i}{dt^2} \tag{3.1}$$

The force can be calculated as the negative gradient of the energy, which can be obtained e.g. from force fields or electronic structure calculations.[13]

$$\vec{F}_i = -\nabla V = -\frac{\partial V\left(\vec{r}_1, \cdots \vec{r}_n\right)}{\partial \vec{r}_i} \tag{3.2}$$

3.1.1. Integration

Solving the resulting differential equation, which contains derivatives of the nuclear coordinates with respect to time and space, reveals the time-dependent motion of the system. Unfortunately, only one- or two-body systems can be solved analytically, while many-body systems require a numerical solution. There are algorithms serving this purpose, e.g. the Verlet algorithm. It is based on a Taylor expansion of the nuclear coordinates as a function of time, which is truncated after the second order. In the

derivation, one step is performed in positive time ($t + \Delta t$) and one in negative time ($t - \Delta t$). Adding up both equations eliminates the first derivative of the coordinates with respect to time (velocity) and by rearrangements an equation is obtained, which can be used to propagate the system (Equation 3.3).

$$\begin{aligned} \vec{r}\,(t + \Delta t) &= \vec{r}\,(t) + \vec{v}\,(t)\,\Delta t + \frac{1}{2}\vec{a}\,(t)\,(\Delta t)^2 \\ \vec{r}\,(t - \Delta t) &= \vec{r}\,(t) - \vec{v}\,(t)\,\Delta t + \frac{1}{2}\vec{a}\,(t)\,(\Delta t)^2 \\ \vec{r}\,(t + \Delta t) &= 2\vec{r}\,(t) + \vec{a}\,(t)\,(\Delta t)^2 - \vec{r}\,(t - \Delta t) \end{aligned} \tag{3.3}$$

New nuclear positions at $t + \Delta t$ can be obtained from the current positions (t), the positions from the last step ($t - \Delta t$) and the current acceleration, which can be calculated from the force.

The movement of the system is completely deterministic. This means with given coordinates and velocities at $t = 0$ the positions and velocities at all future time points can be obtained. However, the number of integration steps of the equation of motion and thus the time scales, which can be reached within an MD simulation, are limited. This is due to the finite valued time step. The numerical integration of the equation of motion introduces an error, which depends on the length of Δt. The size of the error increases with the length of the time step. On the other hand, the computational effort of a simulation depends on the number of time steps to be calculated. Therefore, a trade-off between accuracy and computational cost must be made. Usually, simulations are stable, if the time step is chosen one order of magnitude smaller compared to the period of the fastest motion. Without special treatment the fastest motions in an atomic system are vibrations of bonds containing hydrogen, with a period on a time scale of 10 fs. Thus, the time step is normally set to 1 fs.[13]

3.1.2. Thermodynamic Ensembles

The concept of classical molecular dynamics simulation, carried out with energies and forces either from classical or quantum mechanics, is based on a microscopic theory. The system is assigned a microstate, which is defined by the coordinates ($\vec{r}$) and velocities ($\vec{v}$) or momenta ($\vec{p}$) of every particle of the system. The dynamic behavior of the system is obtained by solving the equation of motion, resulting in a sequence of microstates along time ($\vec{r}(t), \vec{p}(t)$). This is called a trajectory and it can be understood as a curve in the $6N$-dimensional phase space, which is the set of all possible microstates.

By contrast, a macrostate is defined by thermodynamic properties, such as the number of particles, volume, temperature and pressure. Statistical thermodynamics provides a link between the microscopic and macroscopic description. The partition function, which contains information about all microstates, allows the calculation of all thermodynamic

properties of a microscopic system. However, it cannot be derived analytically for a huge molecular system.

Instead, MD simulations can be employed to sample a representative part of the phase space, which is called an ensemble. For an ergodic system, this sampled time average equals the ensemble average, meaning that every point in phase space can be reached from any starting point within infinite time (Equation 3.4). Finally, this averaging enables thermodynamic properties to be obtained from simulations.

$$\langle X \rangle = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau X(t)\,\mathrm{d}t = \lim_{M \to \infty} \frac{1}{M} \sum_{i=1}^{M} X_i\left(\vec{r}, \vec{p}\right) \tag{3.4}$$

This is of importance, because most experimental measurements are carried out on a macroscopic sample. It contains a number of particles of the size of Avogadro's number and the measurement itself takes a finite amount of time. Therefore, properties are measured as averages over an ensemble of particles and averages over time.

Furthermore, experiments take place at defined thermodynamic properties such as temperature and pressure. Controlling these parameters is a key task in order to create comparable conditions to an experiment. This can be accomplished by means of special algorithms, e.g. so called thermostats and barostats, which emulate the coupling of the system to external heat or pressure baths. This allows sampling in an ensemble of choice with certain thermodynamic properties held constant, e.g. the number of particle N, the energy E, the temperature T, volume V, or the pressure P. The corresponding ensembles are the micro-canonical ensemble (NVE) in an isolated system, the canonical ensemble (NVT) or the isothermal-isobaric ensemble (NPT) in a closed system.[13]

3.2. Non-Adiabatic Molecular Dynamics Simulations

In adiabatic MD simulations, the heavy and slowly moving nuclei can be propagated classically. However, light and fast moving particles such as electrons, which show both particle and wave characteristics, must be propagated quantum mechanically.[13] The Born-Oppenheimer approximation uncouples nuclear and electronic motions and thus allows the classical propagation of nuclei on a single PES of one electronic state. This approximation breaks down for processes such as charge transfer, exciton transfer or photochemical reactions, which involve transitions between different electronic states. At (avoided) crossings, the PES from different electronic states come close and couple, which renders a separation of nuclear and electronic motion impossible. Transitions between these adiabatic states become possible. To describe the non-adiabatic phenomena in non-adiabatic molecular dynamics (NAMD) simulations, a quantum mechanical treatment of the dynamics of the system becomes necessary.[31]

3.2.1. The Semi-Classical Approach

Since a full quantum mechanical treatment of a molecular system with more than just a few degrees of freedom is not feasible, some simplifications have to be introduced in order to simulate larger systems. One category of methods, called mixed quantum-classical models, uses a semi-classical approach to describe non-adiabatic processes. The electronic degrees of freedom are treated quantum mechanically and are propagated according to the time-dependent Schrödinger equation (TDSE). All nuclear degrees of freedom are computed classically and their movement follows Newton's equation of motion. Additionally, the quantum propagation may be restricted to a subset of the system, which is called the QM region, similar to classical QM/MM MD simulations. But this QM region, other than in conventional QM/MM MD, only contains the electronic degrees of freedom, while no electronic degrees of freedom are calculated for the remainder of the system.[31] Since QM calculations involve much larger computational cost compared to force fields, they set the time limiting step in NAMD simulations.

Therefore, a subsystem-based approach can be applied to fragment the QM region into smaller parts. Organic semiconductors, the DNA base pairs or chromophores in biological complexes usually exhibit weak non-covalent intermolecular interactions. Thus, the electronic structure of two molecules in proximity can be assumed to be essentially identical to that of the isolated molecules. This allows for a conceptually simple coarse graining. The system can be divided into so called fragments or sites, which can be whole molecules or molecular parts and which are treated independently. Thus, the QM calculations are restricted to a reasonable size and they can be easily parallelized. With this fragmentation, a coarse-grained transfer Hamiltonian (Equation 3.5) for the electronic degrees of freedom can be defined, consisting of (on-)site energies (diagonal elements) for every fragment and pairwise couplings (off-diagonal elements) between them.

$$\boldsymbol{H}_{\text{transfer}} = \begin{bmatrix} \epsilon_1 & V_{12} & \cdots & V_{1N} \\ V_{21} & \epsilon_2 & \cdots & V_{2N} \\ \vdots & \vdots & \ddots & \\ V_{N1} & V_{N2} & & \epsilon_N \end{bmatrix} \tag{3.5}$$

The representation in this fragmented and localized basis is called diabatic. Diagonalizing the diabatic Hamiltonian matrix yields the potentially delocalized adiabatic states.[32]

3.2.2. Integration

The equations of motion for electronic and nuclear degrees of freedom in NAMD simulations are coupled and any propagator algorithm used has to be able to describe this coupling. The most prominent propagators are the mean-field Ehrenfest method[33] and variants of trajectory surface hopping (SH)[34,35]. In the Ehrenfest method, the nuclei are propagated on an effective PES, which is obtained as the average of all adiabatic states

weighted by their respective occupation. In surface hopping on the contrary, the trajectory evolves on a single adiabatic PES, but the current adiabatic state is allowed to change.

This work mainly uses the fewest-switches surface hopping (FSSH) method in the adiabatic representation. This means, the wave function is expressed as a linear combination of adiabatic states (Equation 3.6).

$$\Psi = \sum_i C_i^{\mathrm{ad}} \left|\psi_i\right\rangle \tag{3.6}$$

The occupation of a state is given as the square of the corresponding coefficient ($|C_i|^2$). The probability for a transition from adiabatic state m to n is obtained as

$$P_{\mathrm{FSSH}}\left(m \rightarrow n\right) = \max\left\{0, \frac{2\dot{\vec{R}}\vec{d}_{mn}\mathrm{Re}\left(C_n^{\mathrm{ad}*}C_m^{\mathrm{ad}}\right)}{\left|C_m^{\mathrm{ad}}\right|^2}\Delta t\right\}, \tag{3.7}$$

with the non-adiabatic coupling vector (NACV) d and the adiabatic expansion coefficients C (Re denotes the real part of a complex number). A uniform random number (ζ) is generated according to which a hop takes place or is rejected (Equation 3.8).

$$\sum_i^{n-1} P_{\mathrm{FSSH}}\left(m \rightarrow i\right) < \zeta \leq \sum_i^{n} P_{\mathrm{FSSH}}\left(m \rightarrow i\right) \tag{3.8}$$

The equation of motion for the electronic degrees of freedom is the TDSE, into which Equation 3.6 is inserted.

$$\dot{C}_m = C_m^{\mathrm{ad}} H_m^{\mathrm{ad}} - \sum_n C_n^{\mathrm{ad}} \dot{\vec{R}}\vec{d}_{mn} \tag{3.9}$$

This equation is used to update the occupation of the fragments by transformation back to the diabatic basis.[31,36]

When a transfer happens, a formerly unoccupied molecule will relax to a new equilibrium structure, which in turn leads to a decrease in its site energy. This response of a molecule to a change in its occupation can be taken into account by explicit relaxation (ER), where the QM forces weighted by the occupation drive the nuclear relaxation. The total energy of the system is a sum of energies of the MM environment, the QM region and an interaction term. The forces are obtained in a diabatic basis as

$$\vec{F}_k = m_k\ddot{\vec{R}}_k = -\frac{\partial E_{\mathrm{MM}}^0}{\partial \vec{R}_k} \underbrace{- \sum_i \sum_j a_i^* a_j \frac{\partial H_{ij}}{\partial \vec{R}_k}}_{\vec{F}_{\mathrm{QM}}} - \frac{\partial E_{\mathrm{QM/MM}}}{\partial \vec{R}_k}, \tag{3.10}$$

with the nuclear coordinates R of atoms k. i denotes the fragments of the QM region. The quantum forces can be expanded in terms of the derivatives of (diabatic) transfer Hamiltonian elements weighted by the diabatic coefficients a. Another strategy is to approximate the quantum forces with a Δ-SCF approach by calculating the differences of forces/gradients of the occupied (e.g. charged or excited) and unoccupied (e.g. neutral or ground) states Equation 3.11.

$$\vec{F}_{\Delta\text{-SCF}} = \frac{\partial}{\partial R_k} E^{\text{occupied}} - \frac{\partial}{\partial R_k} E^{\text{unoccupied}} \tag{3.11}$$

Alternatively, this feedback from electronic to nuclear degrees of freedom can be modeled via a phenomenological equation (Equation 3.12), where the site energy is reduced according to the occupation of a fragment by an empirically determined (internal) reorganization energy (λ), which is called implicit relaxation (IR).[37]

$$H'_{ii} = H^0_{ii} - |a_{ii}|^2 \lambda \tag{3.12}$$

When a transfer or hop takes place, the involved states do not necessarily have the same energy. To ensure the conservation of total energy for each trajectory, the nuclear momenta must be adjusted when a hop occurs. This is usually achieved by re-scaling the nuclear momenta along the direction of the non-adiabatic coupling vector (NACV) to compensate for the change in energy.[38]

$$\vec{p}' = \vec{p} + \alpha\, \vec{d}^{\text{ad}}_{mn} \tag{3.13}$$

Here m, n are are the involved adiabatic states and d is the NACV. The NACV can be related to the diabatic gradients as

$$\vec{d}^{\text{ad}}_{mn} = \frac{\left[U^\dagger \nabla H^0_{ij} U\right]_{mn}}{H^{\text{ad}}_n - H^{\text{ad}}_m}, \tag{3.14}$$

with the diabatic to adiabatic transformation matrix $\boldsymbol{U}$. If the adjustment of momenta is not sufficient to compensate the change in energy (an energy-forbidden or frustrated hop), the trajectory continues in the original electronic state j, and the momenta along the NACV are reversed. Alternatively, this momenta re-scaling scheme can be replaced by re-scaling the hopping probability with a Boltzmann factor and leaving the velocities unchanged.[39]

$$P'_{mn} = g_{\text{BC}}\, P_{mn} \quad \text{with} \quad g_{\text{BC}} = \begin{cases} \exp\left(-\frac{H^{\text{ad}}_n - H^{\text{ad}}_m}{k_B T}\right) & H^{\text{ad}}_n > H^{\text{ad}}_m \quad \text{(upwards hop)} \\ 1 & H^{\text{ad}}_n \leq H^{\text{ad}}_m \quad \text{(downwards hop)} \end{cases} \tag{3.15}$$

This so-called Boltzmann-corrected FSSH (BC-FSSH) method has the advantage that the computationally demanding NACVs need not to be evaluated, which is much more efficient.

3.2.3. Charge Transfer

Charge transfer can occur in two different scenarios, either the transfer of a negative excess charge in a system with an additional electron compared to the neutral state, or the transfer of a positive hole in an electron deficient system (Figure 3.1). In both cases, only one electron is transferred.

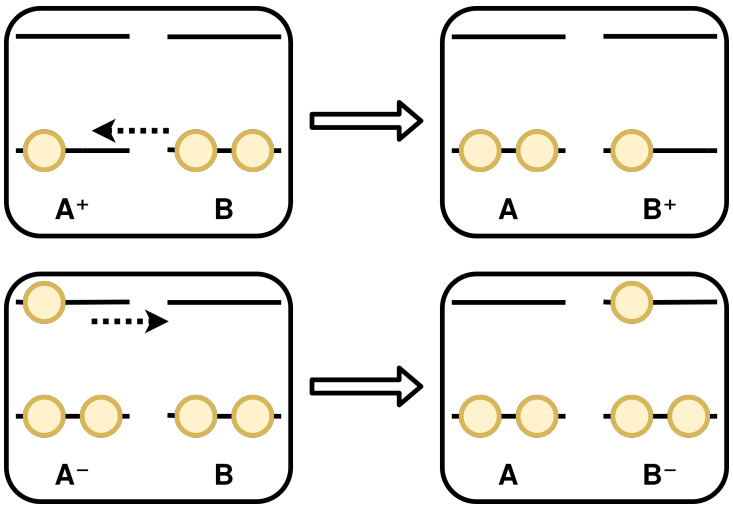


Figure 3.1.: Top: Transfer of a positive hole. Bottom: Transfer of a negative excess electron.

Since low-energy charge transfer typically occurs in a narrow energy window around the Fermi level, it is sufficient to consider only frontier orbitals on the fragments. Therefore, the size of the QM region can be reduced further by restricting it to a set of such frontier orbitals, e.g. the highest occupied molecular orbitals (HOMO) for hole transfer.[10] The hole wave function is then expressed as a linear combination of orthogonal molecular orbitals ($|\varphi_m\rangle$) localized on fragment molecules A (Equation 3.16).

$$|\Psi\rangle = \sum_A \sum_{m \in A} a_m |\varphi_m\rangle \tag{3.16}$$

This is called the fragment molecular orbital approach (FMO)[32,40]. The elements of the corresponding coarse grained Hamiltonian (Equation 3.5) are then calculated as orbital energies (site energies) and Hamiltonian elements between different orbitals (couplings) in neutral fragments (Equation 3.17).

$$H_{mn} = \langle \varphi_m | H[\rho_0] | \varphi_n \rangle \tag{3.17}$$

3.2.4. Exciton Transfer

The transfer of excitation energy can be described in terms of excitons, quasi-particles that can be understood as a bound excited electron and hole pair. In a simplified scheme an electron is promoted to an energetically higher orbital (e.g. the lowest unoccupied molecular orbital - LUMO) through interaction with light of suitable energy (Figure 3.2). The former occupied but now empty orbital (e.g. the highest occupied molecular orbital - HOMO) is called a hole. Excitons are classified according to the localization of the excited electron and hole: They can be localized on one fragment (Frenkel exciton), on two neighboring fragments (charge transfer exciton) or on distant fragments (Wannier-Mott exciton).[41] Frenkel excitons are of interest in this work.

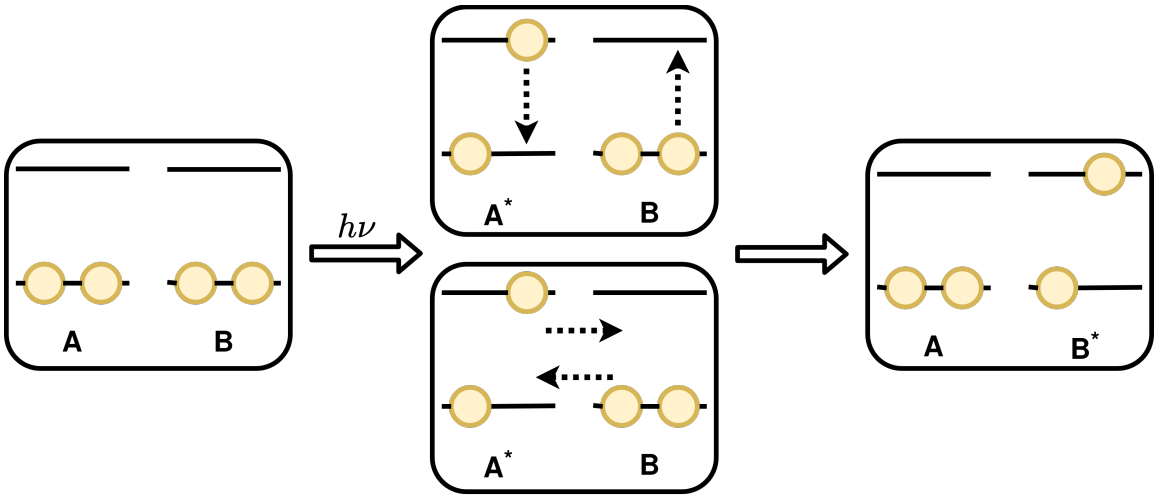


Figure 3.2.: Formation of an exciton by light absorption of fragment A, followed by Förster or Dexter transfer to a neighboring fragment B. Modified from [41]

The exciton and its energy can be transferred from one fragment to another via either of two possible mechanisms as depicted in Figure 3.2. In Förster transfer, the excited fragment is deexcited, while the neighboring fragment is excited. Both fragments interact through a long-range interaction. The same final state can be reached via the Dexter mechanism, where two electrons are exchanged between the involved fragments. This requires a sufficient overlap between the wave functions on the two fragments. For the transfer of Frenkel excitons, the Frenkel Hamiltonian (Equation 3.18)[42,43] can be used as a coarse-grained model of the electronic structure (Equation 3.5).[41]

$$H_{\text{Frenkel}} = \sum_{i}^{N} \epsilon_i \left|i\right\rangle \left\langle i\right| + \sum_{i \neq j}^{N} V_{ij} \left|i\right\rangle \left\langle j\right| \tag{3.18}$$

The first term in Equation 3.18 or the diagonal elements in Equation 3.5 are the site energies of individual fragments, into which the system is divided. They describe local excitations of sites, which can be obtained as excitation energies by any quantum-chemistry method that gives access to excited states, e.g. TD-DFT or TD-DFTB. Excitations of the

whole system are then described as linear combinations of single molecular excitons. The interactions between sites are the second term in Equation 3.18 or the off-diagonal elements of Equation 3.5.[41,44]

The coupling can be split into a short-range and a long-range contribution. The former contains interactions like overlap, exchange, and charge-transfer excitations. Thus, it depends on the spatial overlap of the wave functions of sites, which decreases rapidly (exponentially) with increasing distance between the sites. For fragments at longer distances, on the other hand, only the long-range interaction is usually considered, which only decays slowly with the distance. It can be described as a Coulomb interaction between transition densities (Equation 3.20).[45] The transition density has no classical analog like a charge density and is computed by integrating the product of ground and excited state wave functions over all spin and all but one spatial coordinate.[46]

$$\rho(\vec{r}) = \int \mathrm{d}s \int \mathrm{d}\vec{r}_2, \cdots, \mathrm{d}\vec{r}_N \phi_g(\vec{r}_1, \cdots, \vec{r}_N)\phi_e^*(\vec{r}_1, \cdots, \vec{r}_N) \tag{3.19}$$

$$V_{ij}^{\mathrm{Coul}} = \frac{1}{4\pi\epsilon_0\epsilon_e} \int\int \frac{\rho_i^{eg}(\vec{r}_i)\,\rho_j^{eg}(\vec{r}_j)}{\left|\vec{r}_i - \vec{r}_j\right|} \mathrm{d}\vec{r}_i \mathrm{d}\vec{r}_j \tag{3.20}$$

This can be understood as an interaction between charge distributions, which can be approximated in their representation by applying approaches from electrostatics. The most simple approach is called the point-dipole approximation (PDA) and only considers the first-order term of the multipole expansion of each transition density, the transition dipole. This approximation is only valid for large inter-site distances compared to the size of the sites themselves, yet it does not provide accurate results even for long distances. A more fine-grained representation of transition densities can be achieved by placing a number of transition charges (monopoles) on the sites. In the transition charges from electrostatic potentials (TrEsp) method[47], one transition charge is placed on every atom. These charges are then varied to fit the electrostatic potential of the reference transition density as closely as possible and provide significantly improved couplings compared to the PDA approximation. The transition density cube (TDC)[48] approach uses a grid in Cartesian coordinates to place the transition charges. This makes it a direct numerical integration of Equation 3.20, but it requires a large number of charges to be placed on a fine grid making it extremely demanding in terms of computational cost. The TrEsp and TDC methods are usually not applied to compute excitonic couplings on the fly, e.g. during molecular dynamics simulations, due to their computational demands. Instead, the respective transition charges are calculated once before the simulation and then kept fixed, which significantly decreases computation time during the simulation.[45]

When using TD-DFTB based methods, atomic transition charges can be obtained by a Mulliken population analysis. Due to the efficiency of the method, these charges can be computed on the fly. The coupling between transitions from ground state (0) to states a and b on fragments i and j can be calculated as Coulomb interaction (ζ) of atomic transition charges (Q)

$$V_{ij}^{\text{Coul}} = \sum_{X \in i} \sum_{Y \in j} Q_X^{0a} \zeta_{XY} Q_Y^{0b} \,, \tag{3.21}$$

where X, Y are atoms of the respective fragments.[11,49] The Coulomb interaction in the DFTB formalism is defined as

$$\zeta_{XY} = \int\int \frac{\Phi_X(r)\Phi_Y(r')}{|\vec{r} - \vec{r}'|} \mathrm{d}\vec{r}\mathrm{d}\vec{r}' \qquad \text{with} \qquad \Phi_X(\vec{r}) = \frac{1}{N_X} \sum_{\mu \in X} |\mu(\vec{r})|^2 \,. \tag{3.22}$$

$\mu(r)$ are the basis functions on atom X and N_X is the total number of basis functions.[26]

The total coupling with short and long range contributions can be obtained via supermolecular calculations, which include the whole dimer for which the coupling should be evaluated. Considering this coupled dimer, the two lowest adiabatic states $\{\Psi_1^A, \Psi_2^A\}$ with energies $\{E_1^A, E_2^A\}$ are linear combinations of the two interacting diabatic states $\{\Psi_1^D, \Psi_2^D\}$ of the monomers with energies $\{E_1^D, E_2^D\}$. The adiabatic states can be obtained by diagonalization of the diabatic Hamiltonian.[41]

$$\boldsymbol{H}^D = \begin{bmatrix} E_1^D & V_{12} \\ V_{12} & E_2^D \end{bmatrix} \quad \rightarrow \quad \boldsymbol{H}^A = \begin{bmatrix} E_1^A & 0 \\ 0 & E_2^A \end{bmatrix} \tag{3.23}$$

To calculate the supermolecular coupling, the opposite must be achieved: The adiabatic Hamiltonian matrix, which can be obtained from a dimer calculation, must be transformed back to the diabatic Hamiltonian matrix. A simple way to achieve this is by analytical diagonalization. For a dimer, the matrix diagonalization can be performed analytically. From the energetic gap of the adiabatic (dimer) and diabatic (monomers) states the supermolecular coupling can be computed as

$$V_{12} = \frac{1}{2} \sqrt{\left(\Delta E_{12}^A\right)^2 - \left(\Delta E_{12}^D\right)^2} \,. \tag{3.24}$$

Besides this analytical diagonalization method, there are more advanced approaches for the calculation of supermolecular couplings, e.g. approaches utilizing an algebraic diabatisation procedure based on molecular properties.[50] Here the diabatic and adiabatic Hamilton matrices are connected through an orthogonal transformation

$$\boldsymbol{H}^D = \boldsymbol{C}\boldsymbol{H}^A\boldsymbol{C}^T \,, \tag{3.25}$$

where $\boldsymbol{C}$ is the orthogonal transformation matrix and $\boldsymbol{C}^T$ is its transpose. To obtain the transformation matrix, the diabatization scheme makes use of a representation of the involved states in terms of a molecular property. The relation between diabatic ($\boldsymbol{p}^D$) and adiabatic states ($\boldsymbol{p}^A$) then reads

$$\boldsymbol{p}^D = \boldsymbol{p}^A \boldsymbol{C}^T \,. \tag{3.26}$$

Regarding the molecular property, the target diabatic states should be as close as possible to a set of well defined reference states $\boldsymbol{p}^{D,\mathrm{ref}}$. In other words, the applied transformation to $\boldsymbol{p}^A$ needs to result in a $\boldsymbol{p}^D$ that minimizes the distance to $\boldsymbol{p}^{D,\mathrm{ref}}$:

$$\boldsymbol{C}^T = \arg\min_{\boldsymbol{R}} \left|\left| \boldsymbol{p}^A \boldsymbol{R} - \boldsymbol{p}^{D,\mathrm{ref}} \right|\right| \,. \tag{3.27}$$

Here $||\cdot||$ is the Frobenius norm, the square root of the sum of all squared matrix elements. Equation 3.27 is a matrix approximation problem in linear algebra that is termed orthogonal Procrustes problem[51,52]. The solution to this is given by

$$\boldsymbol{C}^T = \boldsymbol{U}\boldsymbol{V}^T \,, \tag{3.28}$$

with the unitary matrices $\boldsymbol{U}$ and $\boldsymbol{V}^T$, which can be determined by a singular value decomposition:

$$\boldsymbol{M} = \left(\boldsymbol{p}^A\right)^T \boldsymbol{p}^{D,\mathrm{ref}} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T \tag{3.29}$$

The scheme described so far is based on a two-state approximation. In principle, this approach can be applied to any number of states involving the two fragments of the dimer. These can be multiple Frenkel states on single fragments, which are close in energy[53]. Indeed the diabatization is not limited to locally excited (LE) Frenkel states, but can also be adopted to states of a different character, e.g. charge transfer (CT) excitations.[54]

The molecular properties proposed by the authors[50,54] and used in this work are atomic transition charges (ATQ) and transition dipole moments (TDP). For ATQ, $\boldsymbol{p}^A$ is calculated on the dimer and corresponds to a matrix with one column of ATQ $\{q_1^A, q_2^A, q_3^A \cdots q_n^A\}$ for every excited state of interest. Likewise, $\boldsymbol{p}^{D,\mathrm{ref}}$ is derived depending on the character of the diabatic excited states. For locally excited states, the ATQ for the lowest excited states of interest on each individual monomer are calculated, while the ATQs on the remaining monomer are set to zero: $\{q_1^{D,\mathrm{ref}}, q_2^{D,\mathrm{ref}}, \cdots q_{n/2}^{D,\mathrm{ref}}, 0, 0, \cdots 0\}$ and $\{0, 0, \cdots 0, q_{n/2+1}^{D,\mathrm{ref}}, q_{n/2+2}^{D,\mathrm{ref}}, \cdots q_n^{D,\mathrm{ref}}, \}$. The calculations for TDP are performed similarly, with $\boldsymbol{p}^A$ and $\boldsymbol{p}^{D,\mathrm{ref}}$ being matrices made of the Cartesian elements of the corresponding TDP for each considered state.

3.2.5. Calculation of Observables

A measure for the distance of a propagated (pseudo-) particle from its initial position is the mean-square displacement (MSD). It describes the spatial deviation of an entity compared to a reference position (e.g. the starting position) along time. For an ensemble average of N_{traj} independent trajectories the MSD can be computed as[55]

$$\mathrm{MSD}(t) = \frac{1}{N_{\mathrm{traj}}} \sum_{l}^{N_{\mathrm{traj}}} \sum_{A} \left(\vec{x}_A(t)^{(l)} - \vec{x}_A(0)^{(l)} \right)^2 P_A^{(l)(t)}, \tag{3.30}$$

where $\vec{x}_A(t)^{(l)}$ and $P_A^{(l)}(t)$ are the center of mass of fragment A and the corresponding (diabatic) population along trajectory l, respectively. $\vec{x}_A(0)^{(l)}$ is the center of mass at $t = 0$. The MSD can be used to calculate the diffusion coefficient D[55] as

$$D = \frac{1}{2n} \lim_{t \to \infty} \frac{\mathrm{dMSD}(t)}{\mathrm{d}t}, \tag{3.31}$$

where n is the dimensionality (e.g. $n = 1$ for a one-dimensional chain). The derivative of the MSD with respect to time is usually estimated as the slope of the linear part when plotting the MSD versus the simulation time. Finally for charges (electrons or holes) the mobility μ can be obtained with the Einstein-Smoluchowski equation[56].

$$\mu = \frac{eD}{k_B T} \tag{3.32}$$

e is the elementary charge, k_B is the Boltzmann constant and T denotes the temperature. The inverse participation ratio (IPR) is an estimate for the delocalization of a particle in terms of a dimensionless number (of molecules), over which the particle is delocalized. It is calculated as

$$\mathrm{IPR} = \sum_{l}^{N_{\mathrm{traj}}} \frac{1}{\sum_k \left| U_{jk}^{(l)} \right|^4}. \tag{3.33}$$

4. Machine Learning

The previous two chapters introduced physics based theories and methods for the description of molecular structure and properties. Contrastingly, machine learning (ML) techniques offer a variety of non-physical but instead data-driven or statistical methods, which can be applied to problems in chemistry and physics for the very same purpose.

4.1. Foundations of Machine Learning

Any ML[57–60] model can be understood as a complex function, which maps some input variables (x_i), often called features, to some output variables (y_i), regularly referred to as labels or targets (Equation 4.1). This mapping is done by implicitly extracting patterns from sets of given reference data. Different learning techniques such as supervised, unsupervised or reinforcement learning can be distinguished. Using a data set consisting of pairs of features and associated labels refers to supervised learning. The target in supervised learning may consist of a continuous or discrete quantity, which should be predicted either quantitatively or qualitatively (categorized). These two cases refer to the two main classes of problems, which can be addressed with supervised machine learning, regression and classification. In this work only supervised ML approaches were employed to tackle regression problems and will be discuss in the following.

$$\text{Model: } x_i \rightarrow y_i \tag{4.1}$$

To generate such a ML model, the parameters of the underlying function or model must be adjusted in a way that the reference labels are best reproduced when given the corresponding features as inputs. Therefore, the deviation between predicted and true labels must be estimated. This can be measured with a cost or loss function, e.g. the mean squared error between prediction and reference. Minimizing the loss or cost function for a data set can be understood as training the model.

Beside the parameters necessary to calculate the targets, a model can have several other internal parameters, which for example define the architecture of the model. These so called hyperparameters may have a huge impact on the quality of the model, which is why they have to be optimized. Hyperparameter optimization is not performed while training a specific model. Instead, multiple different models are trained on the same data (training set) with varying hyperparameters, chosen *a priori* (cf. Figure 4.1). Afterwards, their performance is evaluated on additional and unseen data, the validation set, to select

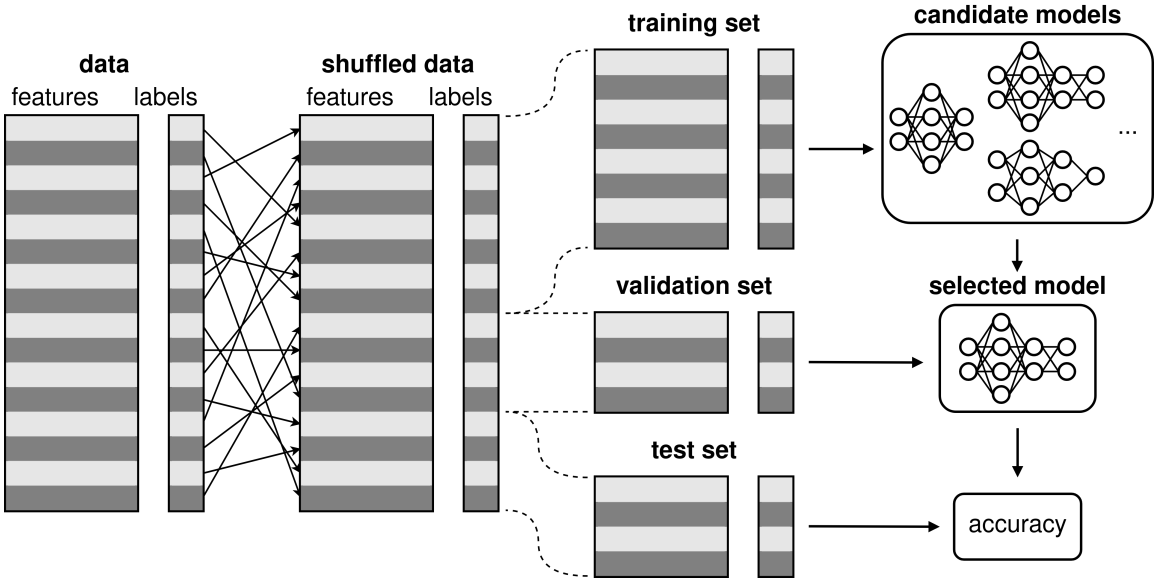


Figure 4.1.: Illustration of data sets in machine learning and their particular purpose in training, validation and evaluation. Modified from [61]

the model with optimal hyperparameters. This can be achieved with a simple but often costly grid-search or more specialized techniques to reduce computational costs.

A common issue that can appear in model training is overfitting. Here, the model is fitted very closely to the training data, including statistical noise. This can be reflected in the performance on different data sets: performance on the training set will be much better compared to the validation set. If beside the measurement of the loss or another suitable metric on the training set the very same metric is measured on the validation set during training, the training process can be stopped as soon as the two metrics deviate more than a chosen threshold (early stopping). This is often used in iterative training processes as they occur e.g. for neural networks. Another technique to prevent overfitting is regularization[62], which sets a preference for having simple models. For example, among all trained models, the one with the fewest variations or the smoothest one may be chosen. L1 regularization applies a penalty to the cost function, that should drive the model towards smaller weights:

$$C = C_0 + \frac{\lambda}{N} \sum_w |w| \ . \tag{4.2}$$

Here λ is the regularization strength and N is the number of data points.

The opposite to overfitting is underfitting, which can occur when there is insufficient training data, the hyperparameters are not well optimized, the relationship between features and labels is not valid or cannot be represented within the chosen model type.

Finally after training and validation, the model must be tested on a third held-out data set to judge the predictive performance (cf. Figure 4.1). In this context, various metrics can be

calculated. One frequently used metric to assess the quality of a trained and optimized model is the coefficient of determination (or R^2-score, Equation 4.3).

$$R^2 = 1 - \frac{\sum_i^N \left(y_i^{\text{ref}} - y_i^{\text{pred}}\right)^2}{\sum_i^N \left(y_i^{\text{ref}} - \overline{y}^{\text{ref}}\right)^2} \tag{4.3}$$

It describes the error of a trained model relative to a fictitious model, that would only ever predict the mean value of the data set. An R^2-score of 1, which is the highest possible value, describes a perfectly fitted model. A value of 0 means the model is as good as if it would only ever predict the mean value, while values below zero (there is no lower boundary) indicate worse models. Additionally, some common error metrics such as the mean absolute error (MAE, Equation 4.4), the mean relative error (MRE, Equation 4.5) and the maximum error (maxE) can be computed.

$$\text{MAE} = \frac{1}{N} \sum_i^N \left| y_i^{\text{ref}} - y_i^{\text{pred}} \right| \tag{4.4}$$

The MAE represents the distance of every reference label from the fitted function. Sometimes normalizing this metric is reasonable, because relative errors between models with different absolute values for their labels can be compared better. However, this is not always possible, as some distributions of labels have a mean near zero. Single bad predictions, so called outliers, cannot be captured with these metrics. Instead, maximum errors or scatter-plots, which are two-dimensional histograms of predicted versus reference labels, are better suited to detect outliers.

$$\text{MRE} = \frac{1}{N} \sum_i^N \frac{\left| y_i^{\text{ref}} - y_i^{\text{pred}} \right|}{\overline{y}^{\text{ref}}} \tag{4.5}$$

In chemistry and physics structure-property relationships are usually the subject of investigation. In order to suitably encode the information of a molecular structure and to make it machine readable, it is converted into a so called descriptor or representation. This representation should capture the physics of the underlying theory and at best be unique and obey certain invariances, such as permutational, translational and rotational invariance[63,64]. Shifting a molecule along or rotating it around an axis does not change the molecule's properties, thus the relationships between input and output should not change. Therefore, the descriptor should stay the same, which is not the case when using 3D Cartesian coordinates. Instead, these coordinates can be transformed to simple descriptors such as the inverse distance matrix[61] (Equation 4.6), where the diagonal does not contain any information, but the off-diagonal elements are the inverse distance between all atoms of the system.

$$X_{ij} = \frac{1}{r_{ij}} \tag{4.6}$$

To encode more of the underlying physics, e.g. the types and arrangement of nuclei, the Coulomb matrix (CM)[65] can be utilized (Equation 4.7). The diagonal elements represent a polynomial fit of atomic energies to nuclear charges, while the off-diagonal elements are similar to the inverse distance matrix. They are based on the Coulomb interaction between the nuclei. Both representations scale with N^2 or rather $\frac{N(N-1)}{2}$, if only the upper/lower triangle of the symmetric matrix is used.

$$X_{ij} = \begin{cases} 0.5 Z_i^{2.4} & \text{, if } i = j \\ \frac{Z_i Z_j}{r_{ij}} & \text{, if } i \neq j \end{cases} \tag{4.7}$$

More advanced descriptors are e.g. the symmetry functions developed by Behler et al.[66] or FCHL[67]. They also include information about the atomic environment, however, are much more costly to calculate compared to the inverse distance or Coulomb matrix. In the following, the two ML methods used in this work are introduced.

4.2. Kernel Ridge Regression

Kernel Ridge Regression (KRR) is a popular ML method, which employs the kernel trick. This allows for linear regression in a non-linear problem by implicitly transforming the features from the lower dimensional feature space into the higher dimensional kernel space, where the relationship between features and labels becomes linear. Here their relationship can be fitted with a linear function, which is much easier than with a non-linear function. Afterwards, the transformation is performed back to the original feature space. The advantage is, that the transformation into kernel space and thus the mapping in this space is not performed explicitly, but rather only implicitly by means of kernel functions. These non-linear functions measure the similarity of two data points (or rather of their representation) in kernel space. In essence, a kernel function is placed on every point in the training set to encode the information for the similarity of data points. A label can then be calculated as a sum of kernel functions, weighted by coefficients (w_m), which have to be trained:

$$y^{\text{pred}} = \sum_{m}^{N_{\text{train}}} k\left(\boldsymbol{X}, \boldsymbol{X}_m\right) w_m \,. \tag{4.8}$$

$\boldsymbol{X}$ is the representation of a query datum, that is to be predicted, and $\boldsymbol{X_m}$ are the representations of all training data points. Often the Gaussian (Equation 4.9) or Laplace (Equation 4.10) kernel (k) is used.

$$k\left(\boldsymbol{X}, \boldsymbol{X}'\right) = \exp\left(-\frac{||\boldsymbol{X} - \boldsymbol{X}'||_2^2}{2\sigma^2}\right) \tag{4.9}$$

$$k\left(\boldsymbol{X}, \boldsymbol{X}'\right) = \exp\left(-\frac{||\boldsymbol{X} - \boldsymbol{X}'||_1}{2\sigma^2}\right) \tag{4.10}$$

σ denotes the kernel width and is a hyperparameter to the model. It is related to the distance between training data points in kernel space. For the training the coefficients can be obtained in a closed form solution as

$$\boldsymbol{w} = \left(\boldsymbol{K} - \lambda \boldsymbol{I}\right)^{-1} \boldsymbol{y}^{\text{train}}, \tag{4.11}$$

with the kernel matrix $\boldsymbol{K}$, all coefficients $\boldsymbol{w}$ and the training labels $\boldsymbol{y}^{\text{train}}$. Training is directly combined with regularization, to avoid overfitting. The regularization strength λ also improves the invertability of the kernel matrix.

The only hyperparamters of a KRR model are the kernel width σ, the regularization strength λ and the kernel type, which makes KRR models easy to train compared to most other ML methods (e.g. neural networks). However, the disadvantage is that for every prediction the kernel between the query datum and all training data points has to be evaluated (Equation 4.8). This results in a dependency of the prediction time on the size of the training set. Thus, KRR is a method that works well with small data sets.[61,68]

4.3. Neural Networks

Neural networks (NN) are inspired by the design of the human brain for processing information. The smallest unit of a neural network is a neuron. For a set of inputs (x_i) a neuron holds the value (y) of the sum of linear functions of these inputs with some associated weights (w_i), which was passed though an activation function f (Equation 4.12).

$$y = f\left(\sum_i w_i x_i + b_i\right) \tag{4.12}$$

Multiple neurons are arranged in layers, which form the actual network (see Figure 4.2). The first layer (input layer) consists of a number of neurons, which receive as inputs the values of the representation. Every neuron in one of the following layers receives the outputs of the neurons from the previous layer as its inputs. The final layer combines the results into one or multiple outputs of the network. All layers after the input layer are called hidden layers, while the very last layer is called the output layer. Not all neurons are necessarily connected to all layers in the former and next layer, which would be called a fully connected network.

$$y_i^l = f\left(b_i^l + \sum_j^{N_{l-1}} w_{ij}^{l-1} y_j^{l-1}\right) \tag{4.13}$$

The output of a neuron i in layer l is calculated as the linear combination of outputs from neurons j in the previous layer $(l-1)$, of which there are N_{l-1}, passed through the activation function (Equation 4.13). A bias (b_i) can be used in every linear combination to achieve higher flexibility via shifting of the values of this specific contribution. The activation function usually is a nonlinear function, which makes it possible to fit and describe highly complex nonlinear behavior within a neural network. Commonly used examples are the hyperbolic tangent, the sigmoid function or modifications of the rectified linear unit (ReLU), e.g. the leaky softplus function (Equation 4.14).

$$\text{leaky-softplus}(v) = (1-\alpha)\log\left(\exp\left(v\right)+1\right) + \alpha v \tag{4.14}$$

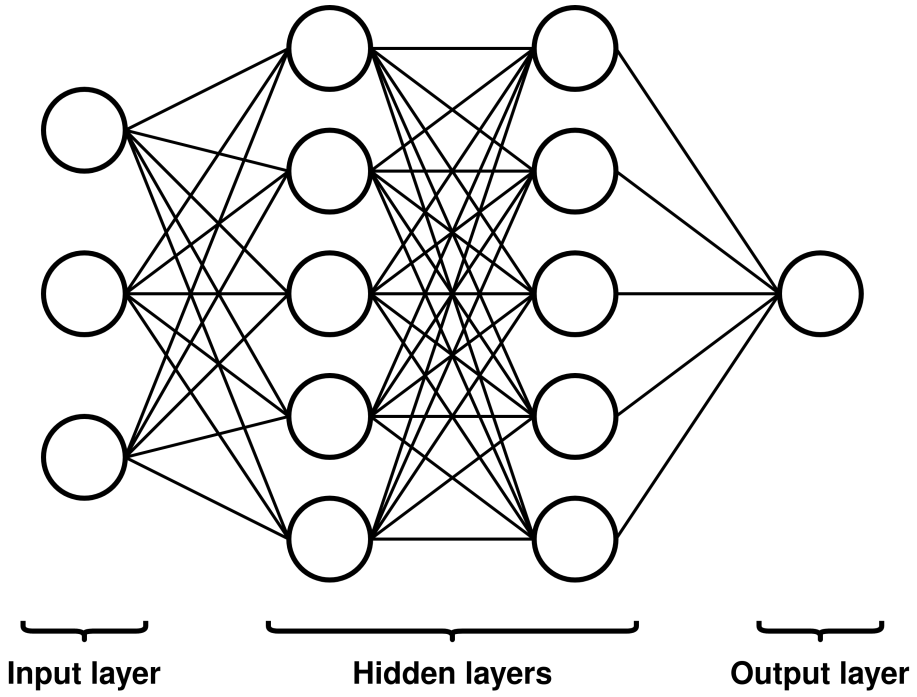


Figure 4.2.: Example architecture of a fully connected neural network: neurons (circles) arranged in layers. Modified from [62].

The weights (and biases) are the adjustable parameters of the model, which are randomly initialized and subjected to training. As described previously, the cost function is computed and needs to be minimized. This is realized using the backpropagation algorithm, which consists of the following steps. In a forward propagation step, the features of training data are fed to and processed by the model to generate a predicted label. Next, the cost is evaluated. This is followed by the actual backpropagation, which refers to the iterative calculation of the derivatives of the cost with respect to the weights and biases, starting

from the output layer. This is accomplished by applying the chain rule from the last to the first layer of the network. Afterwards, all parameters, which are subject to optimization, are updated e.g. according to the gradient descent method. These steps are applied recursively to train a NN model. An important (hyper-) parameter controlling the learning process is the learning rate. It can be adapted during model training and controls the step size that is taken towards the minimum of the cost function.[61,62,69]

Neural networks are highly flexible tools, which can fit extremely complex relationships of features and labels. This comes with the disadvantage of a huge number of hyperparameters, which makes such models prone to overfitting and the overall training and hyperparameter optimization potentially difficult and sophisticated. Despite these drawbacks, machine learning methods and neural networks in particular are highly promising for carrying out expensive multi-scale simulations that require quantum mechanical treatment at least partly and would not be easily possible without them.

Part III.

Contributions

5. Charge and Exciton Transfer Simulations in Organic Semiconductors driven by Machine Learned Hamiltonians

Reprinted in parts with permission from
Mila Krämer, Philipp M. Dohmen, Weiwei Xie, Daniel Holub, Anders S. Christensen and Marcus Elstner:
Charge and Exciton Transfer Simulations Using Machine-Learned Hamiltonians
***J. Chem. Theory Comput.* 2020, 16, 7, 4061–4070.**

Author Contributions:
The results presented in this chapter were generated in collaboration. Training data for the machine learned models was generated by Dr. Daniel Wehl (charge transfer) and me (exciton transfer). Training and evaluation of all models was done by Dr. Mila Krämer. All transfer simulations for charges and excitons were performed and analyzed by myself.

5.1. Introduction

Molecular organic semiconductors (OSCs) have become ubiquitous components of electronic devices, in the forms of organic light emitting diodes (LEDs) in modern display technologies[70–72], organic field effect transistors[73–75], or organic photovoltaic devices[76–78]. Compared to the silicon-based technologies for LEDs, transistors, or photovoltaic devices, OSCs are inexpensive and easier to manufacture and process[79]. With the near-limitless variety of organic compounds, searching for novel candidates with specific properties or optimizing known materials is a costly process. Here, theoretical approaches using simulation techniques can supplement or shorten experimental studies by efficiently screening large portions of chemical space for compounds with promising properties before committing to synthesizing them. This requires fast and robust methods for simulating charge and exciton transfer in large molecular systems which can reproduce experimentally observable quantities such as charge carrier mobilities or exciton diffusion lengths.[5]

Simple models used to describe the transfer of charges assume either a band-like or hopping-like regime. In the band model, the charge is strongly delocalized as the couplings between the charge carriers are large and the reorganization energy is small. The band theory holds for inorganic semiconductors, however, fails to reproduce e.g the charge mobility for OSCs. In contrast, hopping-like transfer occurs when the reorganization energy is large and the coupling is small, creating an activation barrier which is only overcome via thermal fluctuations. This results in a strong localization of the carrier on single molecules, and transfer occurs in discrete hops. The hopping model is widely used to compute the charge mobility for disordered organic systems. In organic molecular crystals, however, neither of these approximations hold.[5]

Novel approaches[9–12] to go beyond the limits of model theories use the direct simulation of charge or exciton transfer employing non-adiabatic molecular dynamics (NAMD) methods. These methods allow the explicit propagation of the electronic degrees of freedom and therefore do not depend on any prior assumptions on the underlying transfer mechanism. The coupled electron-nuclear motion is described with the time-dependent Schrödinger equation (TDSE) using a simplified electronic structure, while the nuclei are treated classically. The electronic structure is reduced to three parameters: site energies, reorganization energies, and coupling energies. While these values are not too costly to calculate on an *ab initio* or semiempirical level of theory for individual structures, extensive sampling along a molecular dynamics (MD) trajectory is necessary to obtain macroscopically observable materials properties. As the sampling process repeatedly passes through similar pairwise molecular conformations, the question naturally arises whether a machine learning (ML) model could substitute the repetitive calculations using a costly quantum mechanical (QM) method. An ML model trained on a limited set of quantum chemical data obtained before a simulation could then take over for the non-adiabatic molecular dynamics, greatly speeding up the prediction of materials properties in multi-scale workflows.

The application of machine learning methods to charge and exciton transfer properties has been investigated recently. Lederer et al.[80] used a kernel ridge regression model paired with a geometric representation to predict charge transfer couplings in pentacene crystals. As Musil et al.[81] showed, charge transfer couplings in both pentacene and azapentacene can be predicted using Gaussian process regression models with a simplified SOAP-Kernel to within a few meV with only a few thousand high-quality training examples. In another work, Caylak et al.[82] presented a neural network model, which is able to predict hole mobilities and charge transfer couplings for an amorphous aluminum complex. For excitonic properties, Häse et al.[83] used a modified Coulomb matrix combined with a neural network to predict excitation energies of bacteriochlorophyll in the Fenna-Matthews-Olson complex and study the exciton dynamics. There has been extensive work on ML-driven excited-state molecular dynamics [84–87] for small organic molecules. Finally, Wang et al.[88] systematically investigated the effects of representation and target value on model training and prediction performance. Most published models require large amounts of *ab initio*-level training data or extensive hyperparameter optimization, which can be prohibitive if a model is to be trained anew for every specific application. Also, it is unclear how the raw accuracy of the ML model affects a system's observable properties,

e.g. charge transfer mobilities and exciton diffusion constants, as this can only be examined by performing explicit simulations.

In this work, the semiempirical DFTB method[20–22,25,28,29,89] as a fast and accurate method for training data generation was used. This method has previously been shown to perform very well for the calculation of couplings[90,91] in conjunction with an empirically determined scaling factor. Kernel ridge regression (KRR) models were trained on this data and their accuracy was evaluated not just on static reference data points but also in explicit NAMD simulations of charge and exciton transfer. By keeping the model and the training procedure as simple as possible, a generic procedure that is easy to integrate into multi-scale workflows was obtained.

5.2. Computational Details

5.2.1. Generation of Training Data

Based on the reported crystal structure[92], an anthracene crystal was constructed containing $10 \times 40 \times 5$ molecules along the *a*-, *b*-, and *c*-crystal axes, respectively. Force field parameters were obtained from the general AMBER force field (GAFF)[93,94] and atomic charges were generated by the restrained electrostatic potential (RESP) fitting procedure[95,96] calculated at the HF/6-31G*[97,98] level of theory using `Gaussian` 09[99]. For system setup and classical MD simulations, the `GROMACS` 5.0.4 package[100,101] was used. After an initial energy minimization, the temperature was equilibrated for 1 ns at 300 K using the Nose-Hoover thermostat[102]. A productive simulation of 10 ns was performed with a time step of 2 fs to sample structures, which were saved every 1000th step.

A subset of 75 anthracene molecules ($5 \times 5 \times 3$) was used to generate the training data in three different data sets: the first data set (`full`) contained couplings between all possible anthracene pairs. The other data sets contained reduced data, including only couplings of molecule pairs within center of mass (COM) distance cutoffs of 0.75 (`short`) and 1.25 nm (`long`), respectively. These cutoffs were chosen to include specific neighbors of any individual fragment in the crystal. The first neighbors along the *a*- and *b*-crystal axes lied within the `short` cutoff, while the `long` cutoff also included the first neighbors along the *c*-axis as well as the second neighbors in the *b*-direction.

For training, the geometries and reference Hamiltonians of these data sets were used. The charge transfer Hamiltonians were calculated with the non-self-consistent variant of DFTB[20,21] as discussed in reference [10], considering only the HOMO (hole transfer). A scaling factor of 1.54 was used for electronic couplings to reach the accuracy of second-order coupled cluster (CC2) calculations.[90] For exciton transfer, excited states were computed using time-dependent long-range corrected density functional tight binding (TD-LC-DFTB2)[22,25,28,29] as implemented in `DFTB+`[89]. The lowest excitation of individual molecules was considered and couplings between these Frenkel states were calculated as

Coulomb couplings (Equation 3.21) with transition charges calculated for every individual structure.

5.2.2. Training and Evaluation of Machine Learned Models

Training and evaluation of all ML models was performed by Dr. Mila Krämer.

Separate models were used for the prediction of site energies and couplings, as their respective absolute values lie orders of magnitude apart and the dimensionality of the representation is different. For site energies only the coordinates of one individual molecule have to be considered, while the coupling is an intermolecular property and thus the coordinates of a molecular pair have to be taken into account. For the prediction of couplings, models used the absolute value as the target. The sign of the coupling is a consequence of the usually arbitrary signs of the wave function on the individual fragments. For simulation of transfer along a one-dimensional chain of molecules, taking the absolute value of the coupling is sufficient[103].

Models were trained for each of the coupling data sets `short`, `long`, and `full` and for site energies, along a learning curve (100, 1000, 5000, 10 000, and 25 000 training examples). The unsorted Coulomb matrix representation was used as implemented in the QML package[104].

Detailed information and all results can be found in reference [105].

5.2.3. Simulations of Charge and Exciton Transfer

A second anthracene crystal was constructed containing $40 \times 30 \times 14$ molecules along the crystal axes in the same manner as described above for training data generation. For this system, simulations of charge and exciton transfer using both DFTB and the trained ML models were performed. One dimensional lines along the *a*- and *b*-directions of 36 and 28 fragments in the middle of the crystal were chosen as a QM zone. No simulations were performed along the *c*-direction, as coupling and thus transfer in this direction is very small. Furthermore, the crystallographic *c*-direction and the experimental *c*-direction differ by about 35°, making direct comparison difficult.[106] Structures in equidistant time intervals were chosen as starting structures for subsequent simulations of charge and exciton transfer.

The mean-field Ehrenfest[33,107] (MFE) and Boltzmann-corrected fewest switches surface hopping (BC-FSSH)[35] methods were employed for non-adiabatic molecular dynamics simulations. The scheme of implicit relaxation (IR) was employed in both methods. As a systematic comparison of the performance of these propagation methods is beyond the scope of this work, the interested reader is referred to reference [108].

The hole/exciton wave function was initially localized on the first molecule, $\Psi(0) = \Psi_1(0)$. Time steps of 1 fs were used for the MFE simulations and 0.1 fs for the BC-FSSH simulations.

The TDSE was integrated numerically with the fourth-order Runge-Kutta algorithm with an integration time step of 0.01 fs. To calculate averages of the observables, swarms of 100 (MFE) and 500 (BC-FSSH) trajectories were propagated, simulated for 1 and 5 ps, respectively. The KRR, MFE, and BC-FSSH methods as well as DFTB were implemented in a local version of `GROMACS` 4.6[109].

The reorganization energy used for charge transfer was calculated with DFTB as 0.084 eV[10]. The corresponding value for exciton transfer was computed with TD-LC-DFTB2 as 0.563 eV, which is in good agreement with reported CC2 calculations.[110]. Charge transfer mobilities and exciton diffusion constants were obtained from the mean-square displacement (MSD) as described in subsection 3.2.5.

5.3. Results and Discussion

In the propagation simulations, only the models trained on the `short` data set were used, as the code for propagation only evaluates the couplings to the nearest neighbors along the chain. For the sake of brevity, only results for models trained on 100, 1000, and 25 000 examples of the `short` data set are presented.

5.3.1. Time Evolution of Couplings

To evaluate the performance of models for the prediction of charge and exciton transfer couplings more closely to actual transfer simulations, a time-series can be investigated. Therefore, predicted coupling values were compared with reference couplings along a trajectory of temporally sequential pair geometries. One single pair of first neighbors in *a*- and *b*-direction was arbitrarily chosen from the crystal for propagation. Coupling values were calculated along a 0.6 ps trajectory with a time step of 1 fs, as shown in Figure 5.1.

Predictions of all models appeared smooth and continuous, without jumps. However, some noise in the predictions was visible, although there were no strong outliers compared to the DFTB reference. Evidently, the models trained on the `short` data set started learning the couplings along the *a*-axis properly only at a training set size of 1000. The predictions for the *b*-direction were slightly better at lower training set sizes compared to the *a*-direction for both charge and exciton transfer. However, the model trained on 1000 examples already showed a close agreement with the reference for both crystal directions, indicating that it could be sufficiently accurate for use in propagation simulations.

5.3.2. Machine Learned Driven Simulations

In the following, the presented ML models were applied to perform non-adiabatic molecular dynamics simulations of hole and exciton transfer along the *a*- and *b*-crystallographic axes

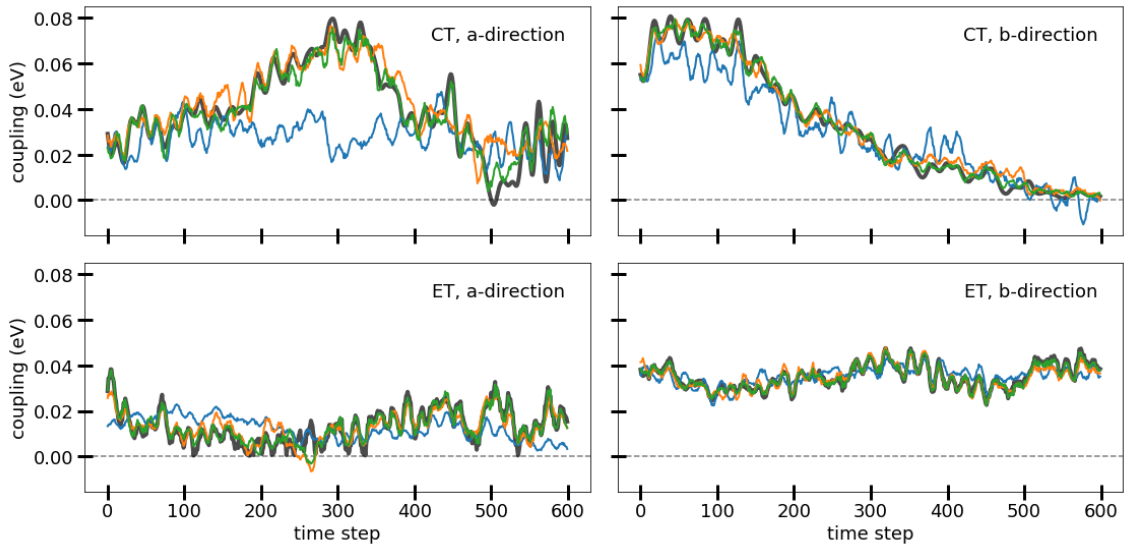


Figure 5.1.: Time evolution of the hole and excitonic couplings between a pair of first neighbors in *a*- and *b*-directions. The DFTB reference is shown in black; colored lines show predictions of models trained on the `short` data set with training set sizes of 100 (blue), 1000 (orange) and 25 000 (green).

of anthracene. All propagation simulations used models trained on the `short` data set, as only nearest-neighbor couplings were calculated.

Charge Transfer

Simulations of hole transfer, where the Hamiltonian was calculated by ML models trained on 100, 1000, and 25 000 examples, were compared to simulations using the corresponding DFTB reference. The averaged time-dependent MSD when employing the MFE as well as the BC-FSSH method is shown in Figure 5.2a and Figure 5.2b, with the corresponding mobilities given in Table 5.1. Results for all training set sizes can be found in the appendix (Appendix A).

Using the MFE method for the *a*-direction, the smallest model with only 100 training examples showed an overestimation of the MSD and hole mobility by a factor of two compared to the DFTB reference. This was not unexpected, as the model only learned the average coupling in this direction (Figure 5.1). The model with 1000 training examples came within $0.12\,\mathrm{cm^2\,V^{-1}\,s^{-1}}$ (8.5 %) of the DFTB value and within $0.39\,\mathrm{cm^2\,V^{-1}\,s^{-1}}$ (34 %) of the experiment. Further increase of the training set size showed only slight improvement, indicating that small errors for individual couplings only have a low impact on the mobilities obtained from simulations. In contrast to the *a*-direction, the *b*-direction seemed to be easier to learn, as all models performed well compared to both DFTB and the experiment. The medium sized model deviated by $0.36\,\mathrm{cm^2\,V^{-1}\,s^{-1}}$ (11 %) and $0.79\,\mathrm{cm^2\,V^{-1}\,s^{-1}}$ (27 %) from the reference method and the experiment, respectively.

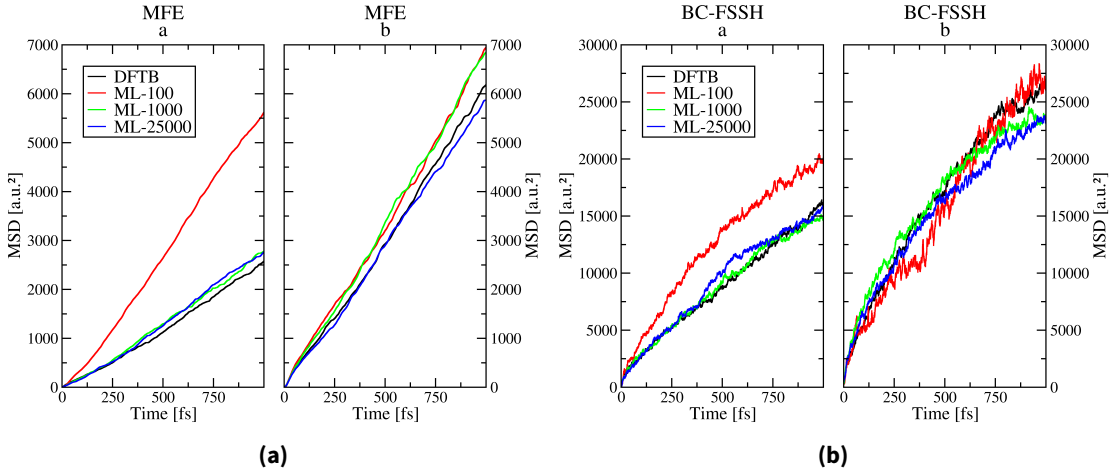


Figure 5.2.: Time evolution of the averaged MSD in *a*- and *b*-direction using the MFE (a) and BC-FSSH (b) methods for hole propagation with DFTB and ML models with various training-sizes.

For simulations using the BC-FSSH method, the performance of the machine learned models compared to DFTB was similar to the MFE simulations. For the *a*-direction, at least 1000 training examples were required, while 100 data points were sufficient for the *b*-direction to reproduce DFTB results. Again, the medium model with only 1000 training examples already returned quite robust and reasonable results. The obtained hole mobilities came within 0.27 $cm^2\,V^{-1}\,s^{-1}$ (3.4 %) and 2.24 $cm^2\,V^{-1}\,s^{-1}$ (17 %) of the DFTB reference for the two crystallographic directions, respectively. In contrast, the absolute hole mobilities obtained from all BC-FSSH simulations differed strongly from the experimental values, even when DFTB was used to calculate the Hamiltonian. Hole mobilities were overestimated by a factor of 4 to 8.

		Exp.[106]	DFTB	ML-100	ML-1000	ML-25000
MFE	*a*	1.14	1.41	3.20	1.53	1.58
	b	2.93	3.36	3.67	3.72	3.22
BC-FSSH	*a*	1.14	8.00	10.05	7.73	8.14
	b	2.93	13.32	14.30	11.08	11.02

Table 5.1.: Hole mobility in $cm^2\,V^{-1}\,s^{-1}$ as calculated from the averaged MSD in *a*- and *b*-direction using the MFE and BC-FSSH methods for hole propagation with DFTB and ML models with various training-sizes.

As shown in a recent study[36], an FSSH scheme can predict mobilities close to the experimental value if properly corrected, e.g. using non-adiabatic coupling vectors. The overestimation of mobilities observed here is rooted in two approximations that have to be applied to accommodate the machine learning model. First, in the BC-FSSH method, the velocities are not adjusted according to the non-adiabatic coupling vectors, as these are not obtained from the machine learning method. Instead, the hopping probabilities are re-scaled with a Boltzmann factor. This leads to an overestimation of hole mobilities[36]. Therefore, machine learning of non-adiabatic coupling vectors is a natural next step beyond the scope of the present work.

Second, the relaxation of the fragments in response to the charge was taken into account only implicitly (IR scheme). Relaxation describes the geometric response of a fragment to a change in its charge occupation. To reproduce this behavior explicitly, the machine learned models would need to explicitly learn the QM forces necessary to drive this phenomenon. The forces can be calculated e.g. as derivatives of the transfer Hamiltonian elements with respect to atomic coordinates. Explicit relaxation is currently impossible with the machine learning model as used here, since it would requires the additional prediction of these quantum forces.

To examine whether explicit relaxation alone could be sufficient for an accurate description, additional simulations with DFTB were performed using a different relaxation scheme including QM forces for the relaxation of every fragment. These forces were calculated as the derivatives of site energies and couplings with respect to atomic coordinates. Employing the explicit relaxation scheme decreased the MSD and slowed down the corresponding mobilities to 1.83 and 4.03 $cm^2\,V^{-1}\,s^{-1}$ for the *a*- and *b*-directions, respectively. These values are in much better agreement with the experiment, indicating that implicit relaxation accounts for the bulk of the error.

Exciton Transfer

In analogy to the application of machine learned models to charge transfer, their behavior in exciton transfer simulations along the *a*- and *b*-directions in the anthracene crystal was investigated. The ratio of average coupling and reorganization energy in the exciton transfer case is 10 times lower than that for charge transfer. This indicates that the exciton diffusion is in the hopping regime. The MFE method was not used in these simulations, as it is known to delocalize the wave function too quickly and thus cannot describe the expected hopping-like behavior.[11] As the program for explicit propagation of exciton transfer using TD-LC-DFTB2 is still under development, the diffusion constants calculated using the ML models were compared with those obtained by kinetic Monte Carlo simulations to solve the master equation (ME) within Marcus theory[111,112]. The time average over 5000 structures of one pair along each axis was used to obtain the average Coulomb couplings for the Marcus rate formula. The diffusion constants obtained from this reference and the ML models are shown in Table 5.2. The averaged time-dependent MSD for the propagation using one selected ML model can be found in the appendix (Figure A.3).

	ME	ML-1000 (BC-FSSH)
a	7.4×10^{-9}	2.8×10^{-7}
b	4.2×10^{-8}	1.9×10^{-6}

Table 5.2.: Diffusion constants in $m^2\,s^{-1}$ as calculated from the averaged MSD in *a*- and *b*-direction using the ME and BC-FSSH with an ML model for Coulomb couplings with different methods for propagation.

The BC-FSSH method again showed significant overestimation of the diffusion constants, here by a factor of 40 compared to ME results. While it is currently not possible to directly simulate exciton transfer using TD-LC-DFTB2 to analyze the same comparisons as for

charge transfer, it stands to reason that the implicit relaxation scheme and lack of non-adiabatic coupling vectors are the causes of the overestimation here as well. Therefore, machine learning models for non-adiabatic coupling vectors and quantum forces are highly desirable.

5.3.3. Timings

The primary motivation for the application of machine learned models in non-adiabatic molecular dynamics simulations of charge and exciton transfer is the replacement of expensive quantum mechanical methods with such models in order to increase the efficiency of these simulations. Therefore, durations were measured for the calculation of transfer Hamiltonian elements.

Calculation of the charge transfer Hamiltonian using the DFTB reference method took on average 0.57 ms per anthracene pair on one core of an Intel Xeon CPU E5-2630 v4 @ 2.20 GHz processor, while the calculation of the exciton transfer Hamiltonian required 55 ms per pair on the same processor. The timings for the different ML models varied strongly with the number of training examples, as the kernel calculation is the rate limiting step. For site energies, a model trained on 1000 examples was always used, while the number of training examples for the coupling model was varied.

As shown in Table 5.3, only machine learned models trained on 1000 or fewer examples could outperform DFTB for charge transfer couplings. This is a consequence of the simplicity of DFTB and the coarse-grained formalism. The costly calculations are done individually for the fragments, and assembling the Hamiltonian is an almost trivial operation once the fragment calculations are done. In contrast, the KRR model needs to calculate the comparatively costly kernel with every training representation. However, the more the individual fragments grow in the number of atoms, the more the models' N^2 scaling resulting from the Coulomb matrix will win out compared to that of DFTB (N^3). This improved scaling will make simulations of systems with large fragments (such as rubrene or phthalocyanine) feasible. Additionally, the DFTB calculations necessary for exciton transfer are far more costly than those for charge transfer, so the machine learned model for anthracene is at least one order of magnitude faster even with a huge training set size.

	ML	DFTB	
n_{train}		CT	ET
100	3.1×10^{-4} s		
1000	3.8×10^{-4} s	5.7×10^{-4} s	5.5×10^{-2} s
25000	3.1×10^{-3} s		

Table 5.3.: Comparison of timings (in seconds) for the calculation of couplings per pair.

The presented models can also provide significant gain in efficiency for simulations using the kinetic Monte Carlo model for hopping-like transfer. These simulations require an

efficient sampling of the average couplings in a system to calculate accurate transfer rates. This is straightforward for ordered structures such as crystalline anthracene. Here, the average couplings for only one fragment pair along each crystal axis must be computed from a sufficient amount of sampled structures (usually hundreds or thousands). In more disordered materials, every pair of fragments must be sampled in this way, leading to a combinatorial explosion of computational effort. In this type of simulation, using a machine learned model instead of QM methods for the calculation of the couplings thus results in a substantial decrease in computational cost.

5.4. Conclusion and Outlook

In this chapter, it was shown that it is possible to create compact and simple machine learning models that can drive charge and exciton transfer simulations. Even with a rudimentary representation and a low amount of training data, these models performed well enough to reproduce hole mobilities within 8.5 % of the DFTB reference and within 30 % of the experimental values in anthracene. This accuracy comes at a computational cost several orders of magnitude lower than *ab initio* methods. While the model was only a little faster than the semiempirical DFTB method in simulations of charge transfer in anthracene, it outperformed DFTB significantly for exciton transfer. With increasing size of the fragments, the improved scaling of the ML model will only increase this margin. While the simple Ehrenfest propagation scheme works well for materials with low reorganization energy[36], it tends to fail in other cases, resulting in fast artificial delocalization of the charge or exciton. As shown above, more sophisticated propagation methods require the calculation of additional properties such as non-adiabatic coupling vectors and molecular relaxation effects to give reliable estimates of mobilities or diffusion constants. Training machine learned models to also predict these properties is an essential step towards reliable and fast computational models.

For an application to biological systems like light-harvesting complexes[113], the influence of an electrostatic environment on electronic structure properties is essential. To this end, the representation can be modified and additional interaction terms such as those in reference [83] can be added.

While it is usually desirable to construct ML models that generalize across chemical space, this requires a large investment in training data and higher computational cost for more sophisticated representations. The use of a fast but accurate semiempirical model as a reference and the low requirements on training data enable the quick and easy training of an ML model for every specific system. For these models, specialization is a quality, and generalization can be limited to the conformational space of one specific system. This training step could be included in the setup of a multi-scale simulation approach, where the machine learning models trained using structures obtained during equilibration can then directly be used for propagation of charges or excitons.

6. Charge Transfer Simulations in Organic Semiconductors driven by Machine Learned Hamiltonians and Diagonal Forces

Reprinted in parts with permission from
Philipp M. Dohmen, Mila Krämer, Patrick Reiser, Pascal Friederich, Marcus Elstner and Weiwei Xie:
Modeling Charge Transport in Organic Semiconductors using Neural Network Based Hamiltonians and Forces
Unpublished Manuscript.

Author Contributions:
The results presented in this chapter were generated in collaboration. Training data for the machine learned models was generated by me, except for the FMO-DFTB1 data for anthracene, which was done by Dr. Weiwei Xie. Training and evaluation of all models for the prediction of gradients was done by Dr. Mila Krämer, while the other models were trained and evaluated by myself. I also performed and analyzed all time-series, transfer simulations and timings.

6.1. Introduction

Molecular organic semiconductors (OSCs) are of prime interest for a variety of different applications[70–78]. With the near-limitless variety of organic compounds, searching for novel candidates with specific properties or optimizing known materials is a costly process. Here, theoretical approaches using simulation techniques can supplement or shorten experimental studies by efficiently screening large portions of chemical space for compounds with promising properties before committing to synthesizing them. This requires fast and robust methods for simulating charge transfer in large molecular systems, which can reproduce experimentally observable quantities such as charge carrier mobilities. This can be done with non-adiabatic molecular dynamics (NAMD) methods by explicitly taking into account the electronic degrees of freedom by solving the time-dependent Schrödinger equation (TDSE)[9–12]. To mitigate the computational cost inherent in a quantum-chemical

description of large systems along extensive molecular dynamics (MD) trajectories, the nuclei are treated classically and the electronic structure of the individual molecules is simplified.

Using a conceptual coarse-graining of the transfer process, the electronic structure of the entire system can be compacted to two key parameters: the energies of frontier orbitals on each molecule (site energies) and the pairwise coupling terms which describe their interaction. Upon getting charged, a formerly uncharged molecule will relax to a new equilibrium structure, which in turn leads to a decrease in its site energy. This response of a molecule to a change in its charge occupation can be taken into account by explicit relaxation (ER), where the forces in the charged state weighted by the charge occupation drive the nuclear relaxation. Alternatively, this feedback from electronic to nuclear degrees of freedom can be modeled via a phenomenological equation, where the site energy is reduced according to the charge occupation of a molecule by an empirically determined reorganization energy, which is called implicit relaxation (IR)[114].

However, the cost of the *ab initio* or semi-empirical models used to describe the electronic structure can be prohibitively large, when the system size and trajectory lengths needed to accurately obtain observables are large. Data-driven approaches, more specifically machine learning (ML) methods, can make use of the fact that the quantum-chemical calculations are highly repetitive, i.e. the conformations sampled during such simulations are very similar. By training the ML model on a small but representative data set for which the relevant electronic parameters have been calculated using a quantum chemistry method, the computationally simpler ML method could then be used to drive the transfer simulations. Several works have previously investigated the applicability of ML models to charge and exciton transfer properties: Both kernel ridge regression (KRR)[80,88,105] and neural network (NN) models[81,82] have been used for predicting charge transfer couplings in organic and metal-organic systems. For excitonic properties, neural networks have been used to predict excitation energies in biological systems[83], and there has been extensive work on ML-driven excited-state molecular dynamics[84–87] for small organic molecules.

In the previous chapter, kernel ridge regression (KRR) models were employed for the prediction of site energies and couplings in anthracene for both charge and exciton transfer applications. By using the fast and accurate density functional tight binding (DFTB) method[20–22,28,29,89,115] to generate reference data cheaply and automating model training, a scheme was created which was easily applicable to new systems without requiring individual models to generalize across chemical space. These models were then applied in charge propagation simulations for anthracene, comparing the resulting observables with DFTB reference values and experimental data. The models were able to reproduce hole mobilities close to the reference and experimental values when using the mean-field Ehrenfest (MFE) propagator together with the implicit relaxation (IR) scheme with a fixed reorganization energy. When using the Boltzmann-corrected fewest-switches surface hopping (BC-FSSH) method for propagation, the ML method was close to the appropriate DFTB reference, while both were far off from the values obtained if molecular relaxation was explicitly (ER) taken into account. This is due to three reasons. Firstly, the employed scheme of implicit relaxation used the adiabatic surface occupation to weight

the reorganization energy. A recent study [114] showed, that using the diabatic wave function occupation of the charge instead leads to good agreement compared to explicit relaxation. Secondly, the reorganization energy was obtained from DFTB, which is known to underestimate this value[114]. Lastly, the Boltzmann-correction is an approximation that is known to cause overestimation of charge mobilities compared to the correct re-scaling of atomic momenta in the direction of the non-adiabatic coupling vectors (NACV)[36].

While the scheme of implicit relaxation yields good results for some materials, it is still an approximation. This may only have limited applicability, e.g. in systems where the assumption, that the molecules relax instantly, does not hold. Therefore, obtaining QM forces, which are required for explicit molecular relaxation, from a ML model is an obvious next step. These forces can be considered as a correction to the neutral state forces provided by the force field (FF) driving the nuclear dynamics in the simulation and can be obtained as derivatives of the Hamiltonian used for propagation with respect to the atomic positions. Unfortunately, obtaining a prediction of a property and its derivatives with respect to the inputs from a KRR model is problematic. Efforts to extend the KRR formalism[116] to this end were unsuitable for the application due to the high computational cost they incur. Furthermore, replacing DFTB with KRR models to drive simulations of charge transfer did not decrease the computational demands of such simulations[105].

Therefore, in this chapter neural network based models are presented, which are able to predict not just site energies and couplings but also the gradients of the site energy necessary for explicit relaxation. Neural networks come with the advantage that the computational cost for a prediction does not depend on the training set size, but only on the architecture of the network. Furthermore, there are efficient built-in routines for the calculation of derivatives of such models, e.g. when using `TensorFlow`[117]. The application was restricted to simulations of charge transfer and the models were trained on DFTB-quality data for anthracene and pentacene. Additionally, the investigations were extended to gradients from higher-level methods such as long-range corrected self-consistent DFTB and density functional theory (DFT). By using the obtained models in propagation simulations, their performance in reproducing hole mobilities in these materials in terms of both quality and computational cost is evaluated.

6.2. Computational Details

6.2.1. Generation of Training Data

It is crucial that gradient predictions are reliable across the entirety of the potential energy surface (PES) accessible throughout the simulation. To explore and train on a broader variety of structures geometries were sampled at a higher temperature (500 K) than the one used in the NAMD simulations for obtaining mobilities (300 K). In NAMD simulations using the IR scheme, only the PES of neutral molecules is explored. Contrastingly, when employing the ER approach, the PES of both neutral and charged molecules will be reached. Thus the training data set for site energy models applied in the IR scheme was constructed

from only neutral structures, while the analogue models for the ER scheme received sampled geometries along both the PES of the neutral and charged system in a 50:50 ratio. The models for couplings can be employed in both IR and ER type simulations, which is why only one model for anthracene and one for pentacene was trained on structures sampled from the neutral and charged state.

The procedure for generating training data was identical for both the anthracene and pentacene systems. For each system, a crystal based on experimental structures of the supercell was created. The anthracene crystal[92] contained $8 \times 8 \times 5$ molecules along the crystal axes (*a*, *b* and *c*), respectively. In pentacene, the size of the crystal[118] was $8 \times 16 \times 5$ molecules. Force field parameters for each system were derived from the general AMBER force field (GAFF)[93,94]. Atomic charges were generated from the restrained electrostatic potential (RESP) fitting procedure[95,96], calculated at HF/6-31G*[97,98] level of theory using `Gaussian 09`[99].

After an initial energy minimization the system was equilibrated for 1 ns with a time step of 2 fs using the Nose-Hoover thermostat[102] at 500 K in the NVT ensemble. Subsequently, a productive classical MD simulation was run at the same temperature for 1 ns with a time step of 1 fs, in which neutral structures were saved every 1 ps. For the sampling of charged structures multiple separate NAMD simulations were performed with the same MD settings, where the charge was constrained to one molecule. All classical and non-adiabatic MD simulations were performed with the `GROMACS` 4.6 software package[100,101].

Eventually, single molecule structures and pair structures sampled in the neutral and charged states were randomly selected from a $5 \times 5 \times 3$ molecular cube in the center of the crystal for subsequent calculations of site energies, their derivatives and couplings. The HOMO was chosen as the frontier orbital for hole transport in the studied systems. Site energies and electronic couplings were calculated using the FMO approach with non-self-consistent DFTB[20,21], often referred to as DFTB1. For the derivatives of the site energies, different data sets were constructed corresponding to the two approaches to obtain the gradients and the applied QM methods. FMO-gradients (see Equation 3.10) were computed with DFTB1 in the FMO approach, while Δ-gradients (see Equation 3.11) were obtained with self-consistent-charge DFTB2[22] with a long-range corrected functional[25] (LC-DFTB2) and the DFT functionals B3LYP[119] and ωB97X[120]. LC-DFTB2 calculations were carried out with the `dftb+`[89,115] program (version 19.1). LC-DFTB2 was used for this data set, as the long-range corrected functional can correct the underestimation of the reorganization energies observed when using the GGA derived non-self-consistent DFTB1[114]. The same holds for the ωB97X DFT-functional. DFT calculations employed the def2-TZVP basis set[121] together with the def2/J auxiliary basis[122] as implemented in `orca`[123,124] (version 5.0.1). It has to be noted, that all data sets for site energies include the same site energies from FMO-DFTB1 and only differ in the derivatives.

6.2.2. Training and Evaluation of Machine Learned Models

The neural network architecture previously presented by Li et al.[84] was adapted for this work, as it has been constructed in order to predict energies and gradients/forces for electronic states of small organic molecules. Either only energies (site energies or couplings) were learned, or site energies were trained together with their respective gradients in a single network. In contrast to the work[105] presented in the previous chapter chapter 5, the NN models were trained on the signed couplings and not on the absolute values. This is necessary to achieve a differentiable function in case of crossing zero. The models used the spatial coordinates of the atoms in the system as inputs, from which inverse interatomic distances between atoms were calculated as a translationally and rotationally invariant representation. Different models were trained for the prediction of site energies and couplings, as the dimensionality of the input was different: in models predicting the site energies (and their derivatives), all intramolecular interatomic distances were included in the representation, while for the coupling models only the intermolecular block of the matrix of inverse distances was used. This has been shown to be superior for the prediction of coupling elements[88] and to reduce the computational cost of each prediction. Initial experiments confirmed that the reduced representation gives lower errors at all training set sizes.

The trainable part of the network was a multi-layer dense feed-forward NN using the leaky softplus activation function (Equation 4.14)[125] with a slope of $\alpha = 0.03$. Models were trained using the `pyNNsMD` package[84] and `Tensorflow` 2.3[117] with the `Keras` API[126]. Weights were optimized using the Adam optimizer[127] and mean-squared-error loss. All models received training and validation data in a 9:1 ratio, and the loss on the validation set was monitored every epoch, so training could be stopped when it did not improve for more than 100 epochs (early stopping). Feature calculation and scaling were both implemented as layers of the neural network, reducing implementation overhead when using the models for prediction. While training models for only energies (site energies or couplings) was straightforward, during training of models for site energy and their gradients the loss was calculated on both the network's energy prediction and the gradients with respect to input coordinates. The relative weights α and β of these two parts of the loss function were considered as hyperparameters of the models Equation 6.1).

$$\text{loss} = \alpha \text{MSE}\left[E^{\text{ref}}, E^{\text{pred}}\right] + \beta \text{MSE}\left[(\nabla E)^{\text{ref}}, (\nabla E)^{\text{pred}}\right] \tag{6.1}$$

The Hyperband algorithm was used as an efficient and automatable way to find working combinations of hyperparameters[128]. Objective for this optimization was the R^2 score on the validation set in case of only training energies. For models trained in both energies and gradients, the objective was set to be a combination of validation R^2 scores for both properties:

$$R^2_{\text{merged}} = \frac{2 - \exp\left[1 - R^2_E\right] - \exp\left[1 - R^2_{\nabla E}\right]}{2}\,. \tag{6.2}$$

This should ensure, that models trained energies and gradients equally well. However, the quality of the resulting models was only weakly dependent on the chosen hyperparameters, indicating that a thorough hyperparameter search is unnecessary and can be omitted when applying the method to new systems. All models were obtained using a training set size of 30 000 structures, with 3333 additional structures used for validation and the optimal hyperparameter configurations. More details on the network configuration, training and hyperparameter search can be found in the appendix (section B).

6.2.3. Charge Transfer Simulations

Separate crystals were generated for the transfer simulations containing $40 \times 30 \times 14$ (36) molecules for anthracene in *a*-direction, $20 \times 40 \times 5$ (36) molecules for anthracene in *b*-direction and $42 \times 84 \times 3$ (73) molecules for pentacene in *T1*-direction (i.e. the diagonal direction in the *ab*-plane). One-dimensional lines of molecules were chosen along the respective directions in the middle of each crystal for the QM region, with the number of fragments specified in brackets above. After equilibration at 300 K, structures in equidistant time intervals of 10 ps were chosen as starting structures for subsequent simulations of charge transfer. The hole wave function was initially localized on the first molecule. A time step of 0.1 fs was used for the propagation of nuclei and the TDSE was integrated numerically with the fourth-order Runge-Kutta algorithm with an integration time step of 0.01 fs. The FSSH method with Boltzmann-correction (BC-FSSH) was used for the NAMD simulations for hole transport in anthracene and pentacene, either with the IR or ER scheme for relaxation. In the IR scheme, reorganization energies of 84 meV[10] and 37 meV[114] were used for anthracene and pentacene, respectively, which have been calculated with DFTB1. Averages of observables were calculated on 1000 trajectories, which were simulated for 1 ps each, while the first 350 fs were regarded as initial equilibration of the charge. All charge transfer simulations were performed within a local version of `GROMACS` 4.6 where DFTB, NN as well as the BC-FSSH methods were implemented.

6.3. Results and Discussion

6.3.1. Implicit Relaxation

As a first step neural network models were trained, evaluated and applied in NAMD simulations employing the implicit relaxation scheme. This means only the transfer Hamiltonian elements (site energies and couplings) were predicted by these models.

6.3.1.1. Model Training and Evaluation

After training the models as described in subsection 6.2.2, their predictions on 10 000 held-out test data from their respective reference data set were evaluated. All trained

models converged well within thousand epochs, with prediction errors shrinking as training set sizes increased. Overall, the hyperparameters included in the search had only a negligible influence on the quality of models obtained compared to the statistical noise from the initialization of the weights. While it is possible that specific hyperparameter configurations might also reduce the number of epochs needed until convergence, the added computational cost of the hyperparameter search far exceeds these gains. Therefore, it can be concluded that for this application a full hyperparameter search can be omitted. The full results of the hyperparameter search can be found in the appendix section B.

	ANT		PEN	
	site energy	coupling	site energy	coupling
MAE	2.652	1.772	2.802	5.520
max. err.	18.207	48.349	27.014	68.075
R^2	0.998	0.955	0.993	0.958

Table 6.1.: Quality metrics for models predicting Hamiltonian elements for anthracene and pentacene: mean absolute error (MAE), maximum error and coefficient of determination (R^2). All metrics except R^2-scores in meV.

Here, the evaluation results for the models, which were used in the propagation simulations with the IR scheme (subsubsection 6.3.1.3), are summarized. Corresponding quality metrics for models trained on the data obtained from FMO-DFTB1 are given in Table 6.1. Site energies for both anthracene and pentacene were learned accurately with models giving low mean absolute and maximum errors compared to the mean of the reference value, which is in the order of several eV. This was also reflected in the high R^2-scores (above 0.99) and the absence of outliers, as shown in the two-dimensional histograms in Figure 6.1 and Figure 6.2.

In contrast, the couplings were learned less well, with mean absolute errors of similar magnitude and even higher maximum errors compared to site energies, although the reference values are several orders of magnitude lower. The couplings along the different crystal directions in pentacene are quite dissimilar, leading to the bimodal distribution seen in the left of Figure 6.2 and a more difficult learning target for the ML model. As the mean magnitude of couplings in the pentacene data set is 28 meV compared to 8 meV in anthracene, the MAE of the coupling predictions in pentacene is greater. Still both coupling models showed R^2-scores higher than 0.95 and all investigated metrics were similar to the KRR models reported previously[105], which indicates that these models should be good enough for use in propagation simulations.

6.3.1.2. Time Evolution of Couplings

To investigate the quality of the coupling models closer to transfer simulation, the evolution of couplings along simulation time was computed. Therefore, reference coupling values from FMO-DFTB1 were calculated along 1 ps NAMD trajectories of arbitrary first neighbor

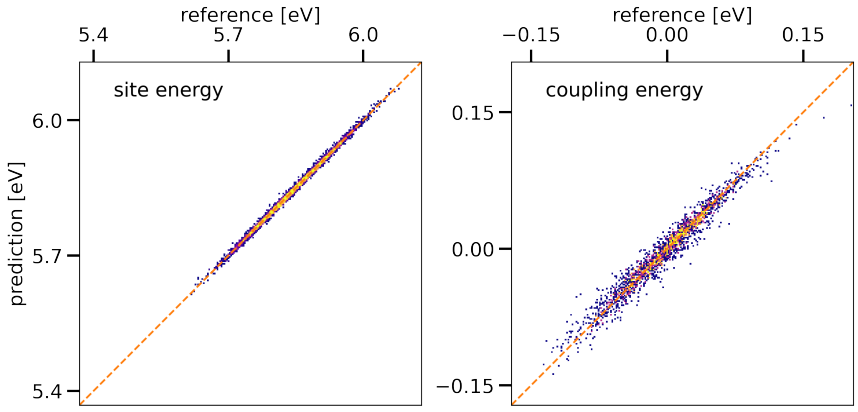


Figure 6.1.: Two-dimensional histograms of model predictions vs. FMO-DFTB1 targets for 10 000 structures in anthracene. Bright colors indicate high data density, unoccupied areas shown in white.

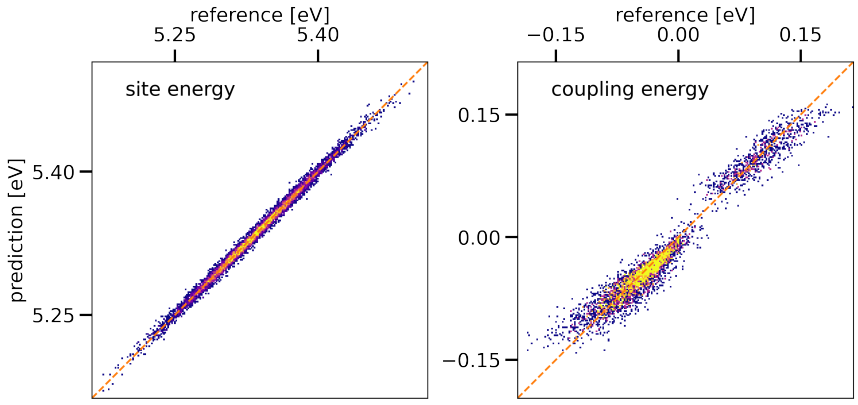


Figure 6.2.: Two-dimensional histograms of model predictions vs. FMO-DFTB1 targets for 10 000 structures in pentacene. Bright colors indicate high data density, unoccupied areas shown in white.

pairs of the investigate crystals and compared with predicted values of coupling models. The time-series of couplings are shown in Figure 6.3.

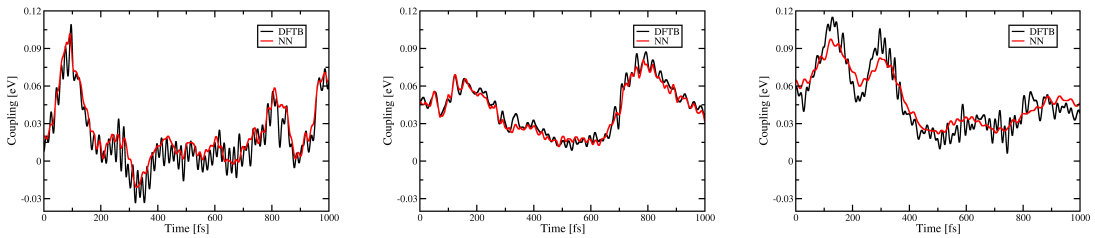


Figure 6.3.: Time evolution of hole transfer couplings between a pair of first neighbor anthracene molecules in *a*- (a) and *b*-direction (b) as well as an analogue pair of pentacene molecules in *T1*-direction (c). Couplings were calculated with FMO-DFTB1 and NN models.

Predicted coupling values were smooth and continuous along the simulation time without showing jumps and outliers. Compared to FMO-DFTB1 the predicted couplings for anthracene in *b*-direction were in close agreement along the whole trajectory. For anthracene in *a*-direction as well as pentacene in *T1*-direction the models captured the general trend of the reference coupling but failed to reproduce the fluctuations on short time scales. Whether the accuracy of the models is sufficient must therefore be investigated by application in charge transfer simulations.

6.3.1.3. Comparison of Obtained Observables

In the following, the trained neural network models were applied to perform NAMD simulations of hole transfer along linear chains of anthracene molecules in *a*- and *b*-direction as well as pentacene molecules in *T1*-direction. The elements of the transfer Hamiltonian were predicted by the models or calculated by the FMO-DFTB1 method as a reference. All simulations employed the implicit relaxation scheme with different values of reorganization energies obtained from various methods (c.f. Table B.3). All simulations were analyzed in terms of accuracy in reproducing hole mobilities compared to the reference method and experimentally determined values as well as inverse participation ratios (IPR) compared to the reference method. The latter is a quantity that expresses the degree of delocalization over a number of molecules. Calculated observables are displayed in Table 6.2, the corresponding plots of MSD vs time can be found in the appendix (section B).

H	λ	ANT-a		ANT-b		PEN-T1	
		μ	IPR	μ	IPR	μ	IPR
DFTB	DFTB1	2.4	2.0	6.6	2.5	11.6	4.0
NN		1.5	2.1	5.8	2.3	12.3	4.7
DFTB	LC-DFTB2	1.1	1.7	2.8	1.9	9.7	3.5
NN		0.8	1.7	2.3	1.8	7.3	4.0
DFTB	B3LYP	0.9	1.7	2.8	1.9	6.6	3.4
NN		0.5	1.7	2.0	1.7	7.9	3.9
DFTB	ωB97XD	0.4	1.3	0.5	1.3	2.7	2.6
NN		0.2	1.3	0.3	1.2	3.7	2.9
Experiment		1.1[106]	- -	2.9[106]	- -	10.5[129]	- -

Table 6.2.: Hole mobilities (in $cm^2\ V^{-1}\ s^{-1}$) and inverse participation ratios (IPR) for the investigated systems as obtained from simulations, where the transfer Hamiltonian (H) was obtained from the DFTB reference or NN models. In all simulations the BC-FSSH method with the IR scheme was employed. Different values for the reorganization energy λ were used.

The estimated mobilities from DFTB and NN driven simulations were in good agreement and were of similar quality compared to previously reported KRR models[105]. For both anthracene directions, the NN models constantly underestimated the mobilities by a small

factor, whereas NN driven simulations for pentacene employing the DFTB1 and B3LYP reorganization energies were slightly overestimated. The same picture was observed for the IPR values. Results from DFTB and NN models were in close agreement, while deviations appear larger for pentacene.

Both the reference and the ML mobilities were too large compared to experimentally determined hole mobilities when utilizing the DFTB1 reorganization energies. This is due to the strong underestimation of reorganization energies of OSCs by DFTB[114], leading to an overestimation of hole mobilities. A previous study[114] showed that the BC-FSSH mobility tends to be lower, when the reorganization energy as an input for the IR scheme gets higher, which is the case when turning to higher level of theory. Thus, results from simulations with LC-DFTB2 and B3LYP reorganization energies were close to experimental values, while the high reorganization energy from ωB97XD led to an underestimation of the mobility.

6.3.1.4. Comparison of Computational Cost

A primary motivation for the use of neural network models or machine learned models in general in non-adiabatic simulations of charge transfer is the replacement of computationally demanding QM methods to gain significant speedup. Here, the performance of the models was assessed in respect to simulation time. To compare timings with the different methods and models, 100 trajectories were run with only two molecules for 1000 steps of NAMD simulation with the exact same settings compared to the previous calculations to obtain observables. The timings were estimated as averages of time spent on single calculations of site energies and for couplings and are shown in Table 6.3. All computations were performed on single CPU cores of Intel Xeon Silver 4214 @ 2.2 GHz processors.

Table 6.3.: Comparison of timings (in ms) for the calculation of Hamiltonian elements (site energies and couplings) in anthracene and pentacene with DFTB and NN models.

	site energy		coupling	
	FMO-DFTB1	NN	FMO-DFTB1	NN
ANT	3.2	0.2	1.7	0.1
PEN	10.0	0.2	3.8	0.2

The computational cost for site energies of anthracene molecules was reduced by one order of magnitude when turning from DFTB to neural network models. For pentacene the models outperformed DFTB by one to two orders of magnitude, which shows the favorable N^2 scaling with the system size of the NN models compared to the N^3 scaling of DFTB (N being the number of atoms). Furthermore, the computational efforts for couplings were reduced by one order of magnitude for both anthracene and pentacene. This performance of neural network models is a clear advantage over KRR models, which were used in the previous chapter, and which could not speed up charge transfer simulations.

6.3.2. Explicit Relaxation

As a second step, additional neural network models were trained, evaluated and applied, which were able to predict the derivatives of site energies in addition to site energies within the same model. From these derivatives the forces necessary for explicit relaxation in NAMD simulations can be obtained. For couplings, the same models as in the IR based simulations were used.

6.3.2.1. Model Training and Evaluation

After training the models as described, they were evaluated in the same way on 10 000 held-out test data from their respective reference data set. These models also converged well within 1000 epochs with decreasing errors. The hyperparameter search had a minor influence on model quality. Full results of the hyperparameter search can be found in the appendix (Table B.2). Table 6.4 gives a few crucial metrics for models trained on the anthracene data obtained from DFTB methods. The corresponding metrics for models trained on DFT data are displayed in Table 6.5.

	FMO-DFTB1		Δ-LC-DFTB2	
	site energy	gradient	site energy	gradient
MAE	2.554	10.581	9.467	41.130
max err	18.207	257.772	50.091	533.835
R^2	0.998	0.998	0.975	0.986

Table 6.4.: Quality metrics for models predicting site energies and their derivatives for anthracene, trained on different DFTB data: mean absolute error (MAE), maximum error and coefficient of determination (R^2). All metrics except R^2-scores in meV (meV Å^{-1}).

	Δ-B3LYP		Δ-ωB97X	
	site energy	gradient	site energy	gradient
MAE	6.009	27.239	17.637	67.067
max err	40.308	308.923	87.480	5945.549
R^2	0.990	0.993	0.914	0.971

Table 6.5.: Quality metrics for models predicting site energies and their derivatives for anthracene, trained on DFT data: mean absolute error (MAE), maximum error and coefficient of determination (R^2). All metrics except R^2-scores in meV (meV Å^{-1}).

Most models gave excellent predictions with few outliers, as can be seen in Figure 6.4 and Figure 6.5. The only exception was the model trained on Δ-ωB97X data, where learning was slightly poorer and a few outliers can be seen. For the FMO-DFTB1 site energies and their gradients, the model reached a MAE of 2.55 meV for the energies and 10.58 meV Å^{-1} for the gradients, with R^2 scores above 0.99 for both. This is comparable to the equivalent

model without gradients in the previous section. The model trained on the Δ-LC-DFTB2 data gave slightly worse results, and did not perform equally well for both energies and gradients in contrast to the model for the FMO-DFTB1 data. This may be due to the construction of the LC-DFTB2 data set, where LC-DFTB2 is only used to calculate the gradients, and these are learned in conjunction with the FMO-DFTB1 site energies. The same holds for the models trained on DFT data. While the Δ-B3LYP model turned out excellent, the Δ-ωB97X model showed larger errors on site energies and gradients. Still the R^2 scores for the latter model were above 0.91.

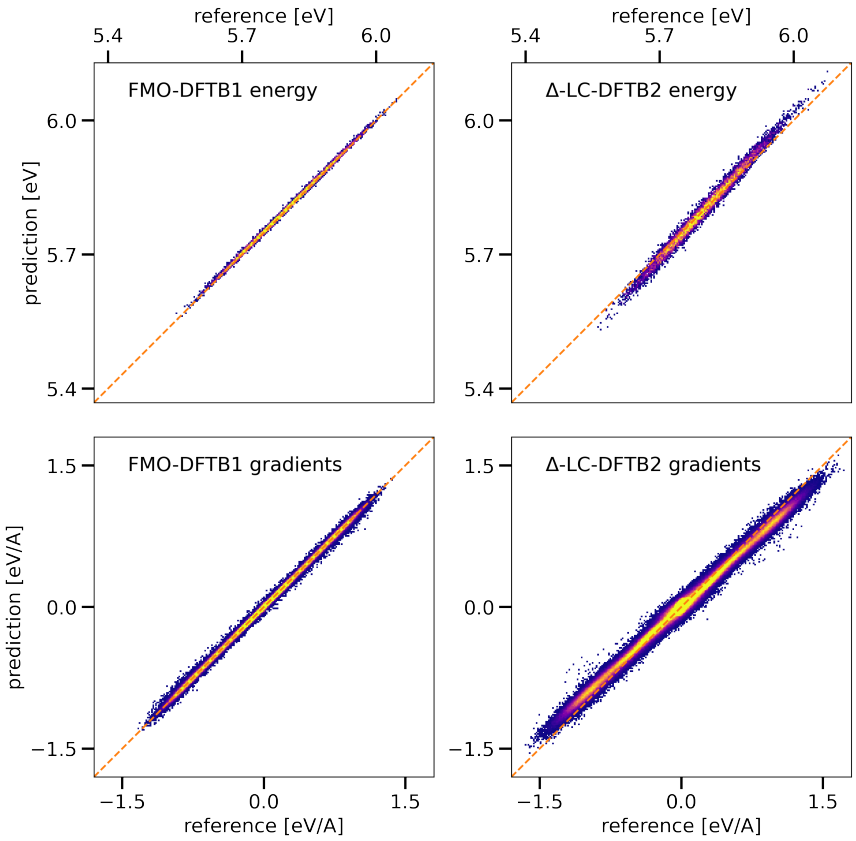


Figure 6.4.: Two-dimensional histograms of model predictions vs. DFTB targets for 10 000 structures in anthracene. Bright colors indicate high data density, unoccupied areas shown in white.

The errors for site energies and couplings were comparable to those obtained for similar training set sizes in the previously presented work (chapter 5)[105], indicating that the models should be sufficiently accurate to give good mobilities in simulations. The effects of the error on the forces cannot be easily quantified, but the maximum prediction errors can be an indication whether the predicted forces could impede the stability of the simulation. The maximum prediction errors for both the FMO-DFTB1 (0.26 eV Å^{-1}) and Δ-LC-DFTB2 gradients (0.53 eV Å^{-1}) as well as the Δ-B3LYP model (0.31 eV Å^{-1}) were well below the forces needed to break covalent bonds (≈1 eV Å^{-1} to 2 eV Å^{-1}[130]). For the Δ-ωB97X model, there were only few outliers exceeding these values. Overall, these large errors were only seen for a few individual outliers, as e.g. less than one percent of gradient predictions

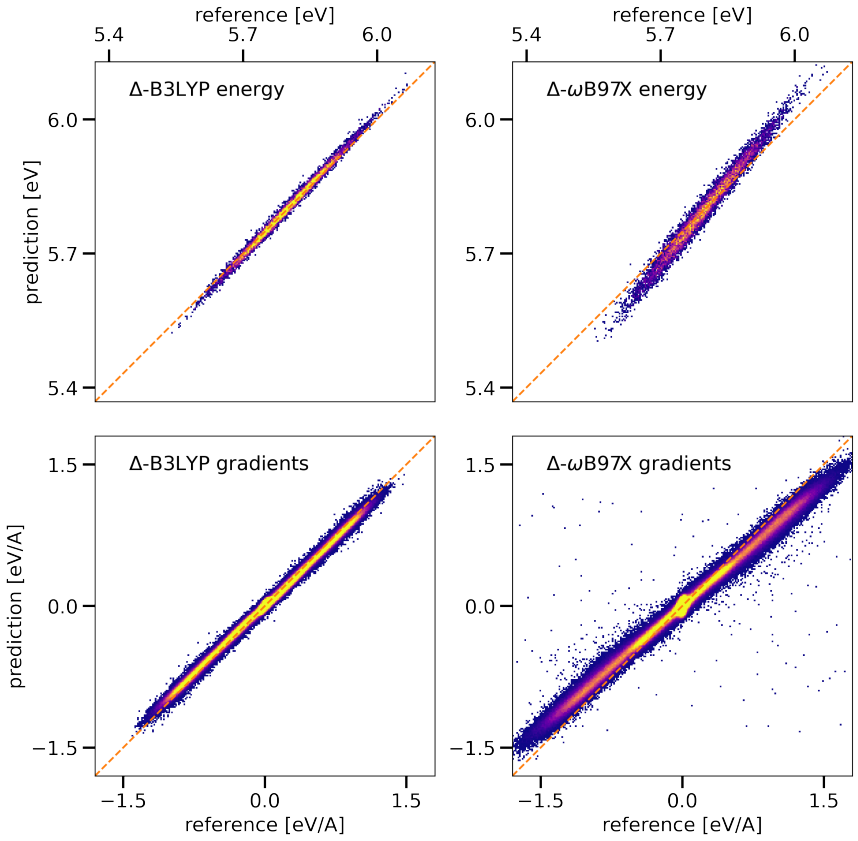


Figure 6.5.: Two-dimensional histograms of model predictions vs. DFT targets for 10 000 structures in anthracene. Bright colors indicate high data density, unoccupied areas shown in white.

showed errors above 40 meV $Å^{-1}$ (FMO-DFTB1) or 160 meV $Å^{-1}$ (Δ-LC-DFTB2). The metrics indicate that the gradient predictions should be sufficiently reliable for performing molecular dynamics simulations.

In pentacene, the results were quite similar. They are summarized in Table 6.6 and visualized in Figure 6.6. DFT data for pentacene were not computed due to their demands on resources. The prediction errors for site energies and their gradients in the pentacene system were slightly worse than for the anthracene models. Maximum prediction errors for the gradients were slightly lower than in anthracene and the error distribution remained narrow.

6.3.2.2.Time Evolution of Site Energies

To get a clearer estimate of the quality of the models for gradient prediction, time-series of site energies for single charged molecules of anthracene and pentacene were calculated. Trajectories of 1 ps were run, propagated with forces from the force-field and additional QM forces from FMO-DFTB1 or neural networks, which were trained on data from the same method. All simulations started from the same initial geometry. Unfortunately, this comparison was not possible for the rest of the models, because their reference methods

	FMO-DFTB1		Δ-LC-DFTB2	
	site energy	gradient	site energy	gradient
MAE	3.434	17.086	5.549	41.003
max err	22.456	239.181	31.647	414.127
R^2	0.990	0.988	0.974	0.959

Table 6.6.: Quality metrics for models predicting site energies and their derivatives for pentacene, trained on different DFTB data: mean absolute error (MAE), maximum error and coefficient of determination (R^2). All metrics except R^2-scores in meV (meV Å^{-1}).

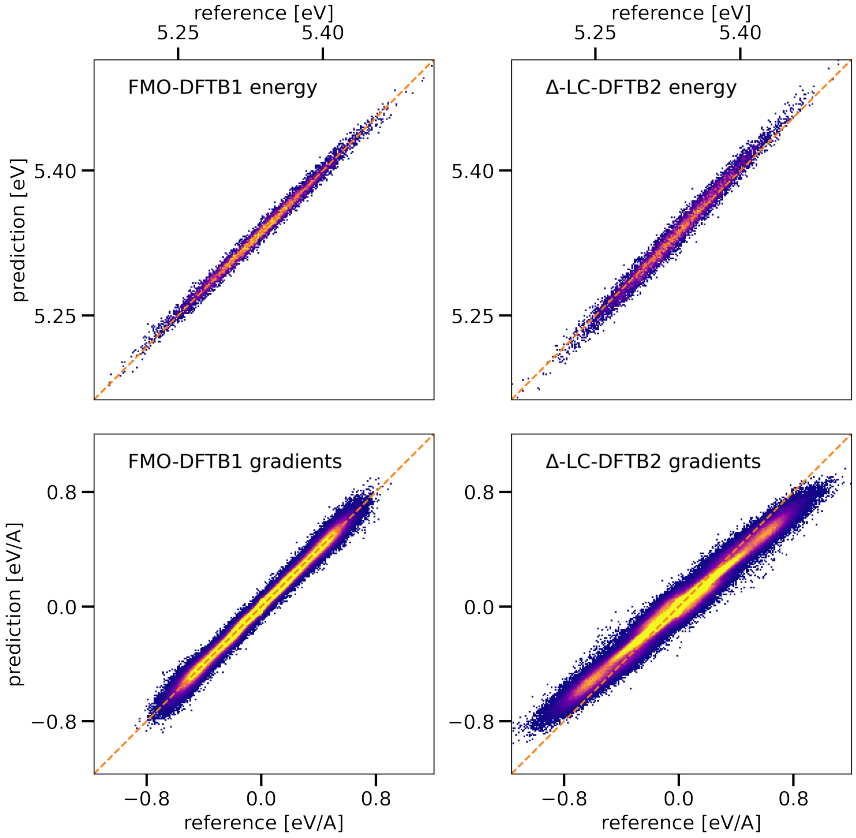


Figure 6.6.: Two-dimensional histograms of model predictions vs. DFTB targets for 10 000 structures in pentacene. Bright colors indicate high data density, unoccupied areas shown in white.

are not implemented for a use in NAMD simulations and would moreover be too time consuming. For a clearer view the site energy of the molecules are shown along a period of 500 fs only (from 250 to 750 fs of simulation time) in Figure 6.7.

The predicted site energies were in close agreement with those from the reference method at the beginning of the trajectories. After 300 fs slight deviations occurred, which increased in intensity over time. For anthracene the deviations stayed small, but became more pronounced along the trajectory for pentacene. This indicates a form of potential error accumulation for the latter molecule, but the error did not significantly increase until the

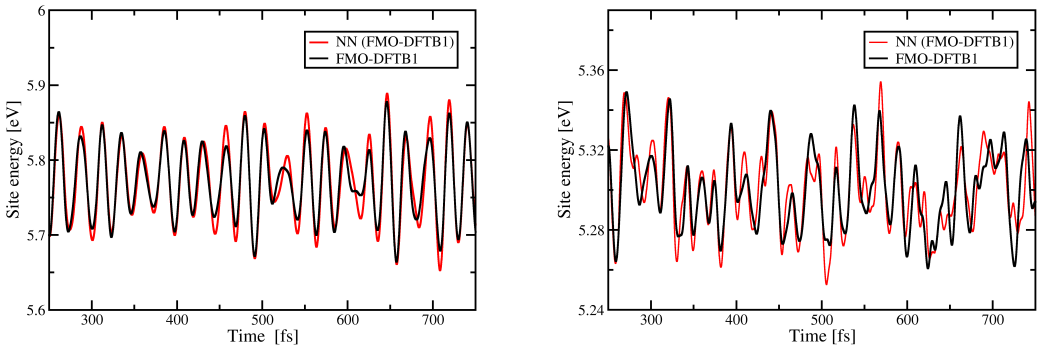


Figure 6.7.: Time evolution of site energies for a single charged molecule of anthracene (a) and pentacene (b). Site energies and corresponding forces were obtained from FMO-DFTB1 or a NN model.

end of the trajectory. The site energy is a property, that is very sensitive for changes in the geometry, and thus is a good measure for the quality of the applied forces. Deviations of site energies were small, although they were affected by error contributions (i) directly from the prediction of the site energy as well as (ii) indirectly from changes of the geometry due to the predicted forces. Thus, the time-series indicated that the model for the prediction of FMO-DFTB1 site energies and gradients should be sufficiently accurate for use in charge transfer simulations.

6.3.2.3.Comparison of Obtained Observables

In order to perform stable NAMD simulations, where the nuclear dynamics are partly driven by machine learned models, it turned out to be necessary to sample geometries for training data on the PES of both neutral and charged molecules. First tests with models trained on geometries in the neutral state only led to unstable and crashing simulations. Adding geometries that were sampled in the charged state gave robust models and stable simulations, so that no exceptionally bad predictions occurred in the more than 100 million simulation time steps performed in this work. As the QM forces are approximately one order of magnitude smaller than the forces from the MM force field, the margin of error that the ML models can produce before simulation stability is impacted is quite large and no further efforts (e.g. active learning) were necessary in this study.

The NAMD simulations presented here were driven by models, which were trained on the same site energies but different gradients for relaxation, calculated either with FMO-DFTB1, Δ-LC-DFTB2, Δ-B3LYP or Δ-ωB97X. Additionally, FMO-DFTB1 driven simulations were performed as the reference method, using FMO-gradients for the relaxation. The calculated hole mobilities and IPR values for all investigated systems are displayed in Table 6.7, the corresponding plots of MSD vs time can be found in the appendix (Figure B.4).

H	Force	ANT-a		ANT-b		PEN-T1	
		μ	IPR	μ	IPR	μ	IPR
DFTB	FMO-DFTB1	3.2	2.0	8.4	2.4	11.6	4.0
NN		2.9	2.1	8.7	2.3	14.6	4.8
NN	LC-DFTB2	2.7	1.9	8.1	2.1	14.2	4.5
NN	Δ-B3LYP	2.5	2.0	8.1	2.2	-	-
NN	Δ-ωB97X	1.9	1.7	5.8	1.8	-	-
Experiment		1.1[106]	-	2.9[106]	-	10.5[129]	-

Table 6.7.: Hole mobilities (in $cm^2\,V^{-1}\,s^{-1}$) and inverse participation ratios (IPR) for the investigated systems as obtained from simulations with the DFTB reference or various NN models. In all simulations the BC-FSSH method with the ER relaxation scheme was employed.

Comparing mobilities from NN-driven simulations with those from DFTB reference simulations (both using FMO forces), a good agreement for anthracene was found with deviations of 9 % and 4 % for *a*- and *b*-direction, respectively. The results for pentacene differed more, showing an overestimation of the mobility by roughly 26 % by the NN driven simulations. The exact same trend was observed for IPR values, where NN-driven simulations agree closely with the reference for anthracene but are overestimated for pentacene. The lower agreement with the DFTB reference for pentacene in contrast to anthracene was expected to be caused by the less accurate prediction of site energy gradients.

Again, both the DFTB reference and the ML values were too large compared to experimentally determined hole mobilities. This is expected due to the employment of forces, which lead to a low reorganization energy, similar to the previously discussed IR simulations. Therefore, it would be expect that the NAMD simulations driven by NN models trained on LC-DFTB2 or DFT gradients, would yield lower mobilities in better agreement with the experiments, just as observed when using higher reorganization energies from the same methods in the IR scheme (subsubsection 6.3.1.3). In fact, the expected trend of decreasing mobilities was observable, although the effect turned out to be smaller in magnitude compared to results utilizing the IR scheme.

For anthracene the simulations using Δ-LC-DFTB2 and Δ-B3LYP gradients only showed a slight decrease of the hole mobility, similarly to pentacene using Δ-LC-DFTB2. Results with Δ-ωB97X gradients for anthracene were improved, but the mobility was still overestimated by a factor of approximately 2 compared to the experiment. Moreover, all mobilities obtained by simulations within the ER scheme were smaller compared to their analog simulations within the IR scheme, also for FMO-DFTB1. In contrast, the delocalization in terms of IPR values were very similar for IR and ER simulations.

To investigate these findings, effective reorganization energies of anthracene were compared to "real" reorganization energies. An effective reorganization energy (λ_{eff}) can be estimated from the difference in site energies for a charged and uncharged molecule. The geometries for these two molecules can be generated in NAMD simulations in QM/MM

or ML/MM MD fashion, from pure QM MD simulations or from single QM optimized structures. Various combinations of structures from different approaches and different methods for the estimation of λ_{eff} are visualized in Figure 6.8 together with the actual reorganization energies (λ). All values are listed in the appendix (Table B.3).

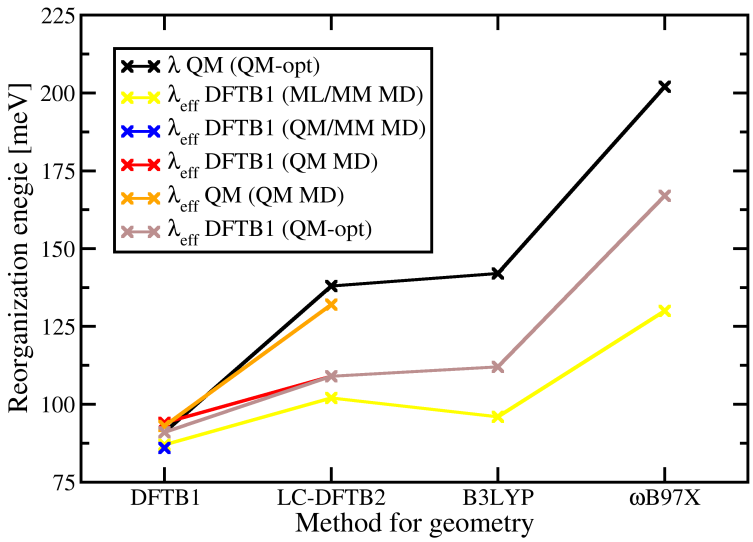


Figure 6.8.: Reorganization energies (λ) and effective reorganization energies (λ_{eff}) for anthracene. The latter were obtained with DFTB1 for structures from different approaches, which are given in brackets. For the orange curve, λ_{eff} was calculated with the same QM method, that was used to generate the structures.

The effective reorganization energies obtained from NN driven NAMD simulations (yellow curve) were significantly lower compared to the actual reorganization energies (black curve). The only exception was DFTB1, where both values agreed well with each other and with the DFTB1 NAMD reference (blue point).

To check, weather this result was rooted in the structures, effective reorganization energies were computed for structures from DFTB1 and LC-DFTB2 QM MD simulations, without involving the force field. The resulting values (red curve) were only slightly higher compared to those from NAMD simulations. Therefore, the structures do not seem to be the reason for the discrepancy. The corresponding DFT QM MD simulations were started, but are still running.

Until this point, all effective reorganization energies were obtained with DFTB1. Contrastingly, when the other QM methods were utilized to compute the site energy gap (orange curve), the effective reorganization energies were in good agreement with the actual λ values. To prove the opposite, DFTB1 was used to calculate the effective λ values on the optimized geometries (brown curve). Here, the values decreased significantly and were similar to their effective counterparts from NN driven NAMD simulations. This strongly indicates, that it is not sufficient to use only forces from more accurate methods.

In addition, it is necessary to calculate the site energies with the same method to reproduce the full effect of the QM reorganization.

The exceptional behavior of DFTB1, where all values match, is due to the formalism of the method. DFTB1 determines the energy as a sum of orbital energies and an additional repulsive energy term (Equation 2.17). The repulsive term depends only on the geometry and thus is the same for a neutral and charged molecule with the same structure. Hence, this term is canceled out and the reorganization energy is equal to the effective reorganization energy. This is not the case for all other QM methods.

6.3.2.4.Comparison of Computational Cost

Timings for simulations employing the ER scheme are shown in Table 6.8 and were generated in the same way as for simulations employing the IR scheme (shown in section 6.3). For the computation of gradients in the Δ-approach timings were averaged on the training-data calculations.

Table 6.8.: Comparison of timings (in ms) for the calculation of site energies plus gradients in anthracene and pentacene with DFTB, DFT and NN models.

	FMO-DFTB1	Δ-LC-DFTB2	Δ-B3LYP	Δ-ωB97X	NN
ANT	3.2	7.9×10^{2}	2.1×10^{6}	3.7×10^{6}	4.7×10^{-1}
PEN	1.0×10^{1}	2.8×10^{3}	- -	- -	8.6×10^{-1}

Timings for FMO-DFTB1 in the IR and ER scheme (without and with gradients) for both anthracene and pentacene were approximately the same. This shows that the time for the computation of FMO-DFTB1 gradients is negligible compared to the calculation of site energies, while this is not the case for NN models. Here, the prediction time for site energies and gradients (ER) was higher compared to the prediction of only site energies (IR). Still the computational cost for site energies and their gradients in the FMO scheme was reduced by about one order of magnitude when turning from DFTB to NN models. For the much more costly gradients calculated with LC-DFTB2, the respective neural network models outperformed DFTB by three orders of magnitude. Again, the favorable scaling with the system size becomes apparent when comparing timings for anthracene and pentacene. Switching to even more costly DFT gradients, the speedup for the NN models is about seven orders of magnitude for anthracene.

6.4. Conclusion and Outlook

In this chapter, non-adiabatic molecular dynamics simulations driven by neural network based Hamiltonian elements and diagonal gradients for hole transport in OSCs were

performed. In contrast to the KRR models[105] used in the previous chapter, the NN models allow for the learning of gradients and Hamiltonian elements simultaneously. Sampling both the neutral and charged state of the molecules was necessary to obtain models for stable simulations, which were partly driven by NN forces (explicit relaxation).

Hole mobilities and inverse participation ratios from the DFTB reference were reproduced with NN models in BC-FSSH simulations utilizing precalculated reorganization energies for implicit relaxation in anthracene and pentacene. Good agreement was also achieved in simulations, where the relaxation was accounted for explicitly by applying FMO-DFTB1 forces. However, the NAMD simulations driven by NN models trained on Δ-LC-DFTB2 gradients for anthracene and pentacene, as well as Δ-B3LYP gradients for anthracene, gave only a slight decrease of the mobilities. While the use of a Δ-ωB97X model for anthracene significantly reduced the mobility, mobilities from ER simulations were generally higher compared to IR simulations.

The estimated effective reorganization energies from NAMD simulations explained this finding, being lower compared to actual reorganization energies from LC-DFTB2, B3LYP and ωB97XD,. This issue could possibly be solved by employing the same method that was used to calculate the gradients for the computation of site energies as well. So far, it cannot be excluded that the error caused by the neural networks also contributes to this issue. However, switching to a consistent use of the same methods for the computation of site energies and gradients could further improve model predictions, as the gradients would then be the actual derivatives of the site energies. This is subject to ongoing work.

Moreover, employing the Boltzmann-correction is known to lead to an overestimation of the mobility[36]. Therefore, a next step will be to extend the approach to derivatives of couplings, which can be used to compute non-adiabatic coupling vectors. The computational cost of NN models was about one order of magnitude lower than for DFTB, two orders of magnitude lower than for LC-DFTB2 and seven orders of magnitude lower compared to DFT. These low computational demands allow the use of NN-driven NAMD simulations for charge transport in large-scale OSCs, such as charge transport in 2D crystals. Such data-driven approaches could also allow the simulation of exciton transfer in molecular materials, which can be described in a very similar way, but is significantly more challenging in terms of computational efforts.

7. Calculation of Supermolecular Excitonic Couplings

Reprinted in parts with permission from
Nils Schieschke, Beatrix M. Bold, Philipp M. Dohmen, Daniel Wehl, Marvin Hoffmann, Andreas Dreuw, Marcus Elstner and Sebastian Höfener:
Geometry Dependence of Excitonic Couplings and the Consequences for Configuration-Space Sampling
***J. Comput. Chem.* 2021, 42, 20, 1402–1418.**

7.1. Introduction

Excitonic coupling plays a key role for the understanding of excitonic energy transport, e.g. in organic photovoltaics[131,132]. However, the calculation of realistic systems is often beyond the applicability range of accurate wave function methods, so that lower scaling semi-empirical methods have to be used to model exciton transfer. The semi-empirical long-range corrected density functional tight binding method with the time-dependent linear response extension (TD-LC-DFTB2) has been benchmarked in comparison to more accurate methods such as coupled cluster singles and doubles (CC2) for the distance and angle dependence of supermolecular excitonic couplings. The test set consisted of dimers of selected organic molecules (acenes, guanine, purine and nitrogen-substituted naphthalenes). It was discovered, that TD-LC-DFTB2 yields accurate excitation energies, which are in agreement with the CC2 reference. However, for supermolecular couplings from an analytical diagonalization the standard confined parametrization of DFTB led to an underestimation of couplings compared to CC2.[133]

Therefore, a new parameter set was generated, which was based on the set used in Reference [29]. In the DFTB method, the atomic orbitals are calculated from atomic Kohn-Sham equations. Because free atomic orbitals are too diffuse, an additional harmonic potential is introduced, which confines the atomic orbitals. So called confinement radii for the wave function and the atomic density are usually set to reproduce atomization energies, geometries and vibrational frequencies. This leads to a very compact basis, resulting in accurate excitation energies and Coulomb couplings. Excitonic couplings in a supermolecular approach, however, require a more diffuse basis, as they largely depend on the exchange and overlap of separated fragments. The confinement radii of carbon were subject to the optimization. Different combinations of increased radii have been tested for

the calculation of couplings on a face-to-face anthracene dimer in comparison to CC2. A new parameter set was chosen leading to improved results.[133]

However, this benchmark used optimized face to face or tilted molecular homo-dimers at short distances. Because the two monomers comprising the dimer had the exact same structure, the excitation energies were degenerate, i.e. $\left(\Delta E_{12}^{D}\right)^2 = 0$ (c.f. Equation 3.24). Thus, the coupling was estimated as half the energy difference of the dimer states that were coupled. The underlying condition for this to be valid is that the two monomers have identical geometries and a separation over large distances yields degenerate energies for the coupling states.[133] These static model systems do not necessarily represent structures of real crystals. Therefore, in the following the influence of geometric fluctuations on supermolecular excitonic couplings of anthracene were investigated in comparison to Coulomb couplings. Additionally, occurring issues regarding the supermolecular approach were resolved by implementing a new diabatization method.

7.2. Computational Details

Sampled structures were taken from a simulation, which was performed for training data generation in chapter 5[105]. One anthracene pair in the crystallographic *b*-direction in the center of the crystal was extracted from the MD (5000 snapshots), as pairs along this direction show the highest coupling values. Supermolecular couplings were calculated with the analytical diagonalization approach (Equation 3.24). Excited state calculations were performed with TD-LC-DFTB2[22,25,26,28,29] as implemented in `DFTB+`[89,115] (version 19.1) and density functional theory (DFT) using the `ORCA` program package[123,124]. For DFT the ωB97X functional[120] with the Tamm-Dancoff (TDA) and the resolution of the identity (RI) approximation was used. The def2-TZVP basis set[121] was employed in combination with semi-numeric exact exchange, RI-J was used for the Coulomb contribution together with the def2/J auxiliary basis set[122]. Additionally, Coulomb couplings were calculated with TD-LC-DFTB2.

Exciton diffusion constants were calculated using a master equation (MEQ) approach, which was solved stochastically for a single exciton using kinetic Monte Carlo[5,134] simulations as described in [105]. Transfer rates were determined according to Marcus theory[111,112] as

$$k_{ij} = \frac{V_{ij}}{\hbar}\sqrt{\frac{\pi}{k_B T \lambda}}\exp\left(-\frac{\lambda}{4k_B T}\right), \tag{7.1}$$

with a reorganization energy λ of 0.56 eV[105] and the coupling V. Coupling values were obtained for the crystal structure (static) and are used beside the root mean square (RMS) value[135] or the sampled structures to incorporate fluctuations (dynamic).

$$\mathrm{RMS} = \sqrt{\langle V^2 \rangle} = \sqrt{\langle J \rangle^2 + \sigma^2} \tag{7.2}$$

A value to quantify the degree of influence is the coherence parameter[135]

$$C = \frac{\langle V \rangle^2}{\langle V^2 \rangle} = \frac{1}{1 + \frac{\sigma^2}{\langle J \rangle^2}} \tag{7.3}$$

reaching values near one or zero, when the coupling is defined by the average structure or by non-equilibrium conformations, respectively. The mean-square displacement (MSD) of the exciton averaged over 10 000 trajectories was used to calculate diffusion constants, as described in subsection 3.2.5.

7.3. Results and Discussion

7.3.1. Geometric Influences on Supermolecular Excitonic Couplings

Sampled structures of an anthracene crystal were utilized to assess the influence of geometry fluctuations on supermolecular couplings calculated with DFTB and DFT. In this case, the two monomers were not identical and a (hypothetical) separation to large distances in general did not yield degenerate excitation energies. This implies that in case of the MD simulation for each snapshot not only the dimer energies ΔE_{12}^{A} have to be calculated, but also the individual monomers to obtain the monomer energies ΔE_{12}^{D} (Equation 3.24).

However, it must be pointed out that numerical issues can occur. If the excitation energy gap of the dimer is smaller than the difference of monomer excitation energies, the square-root term becomes negative and the coupling turns imaginary. This problem may be rooted in the approximate nature of the overall approach, which assumes that the dimer states are a linear combination of the corresponding monomer states. Approximately 1 % of the couplings turned out to be imaginary and the corresponding snapshots were neglected for the analysis.

In Figure 7.1 the coupling distributions are shown. Calculated statistical measures can be found in Table 7.1. The results of the supermolecular approach showed only slightly increased mean values as well as slightly broader distributions of couplings compared to Coulomb couplings, indicating that effects of exchange and overlap are relevant for anthracene in *b*-direction to a certain extent. DFT couplings were only a little higher in energy, but at a comparable width of the distribution. The coherence parameters of approximately 0.97 indicated, that the coupling and thus the transfer is governed by the average structure in general.

To estimate the influence of couplings on the transfer, diffusion constants were estimated using a master equation (MEQ) approach, which was solved for coupling values calculated on the static crystal structure (static) and on the RMS value of sampled structures (dynamic). The latter contained fluctuations of couplings on top of the mean value. The calculated diffusion constants can be found in Table 7.2, the time derivative of the exciton MSD is shown in the appendix (Figure C.1).

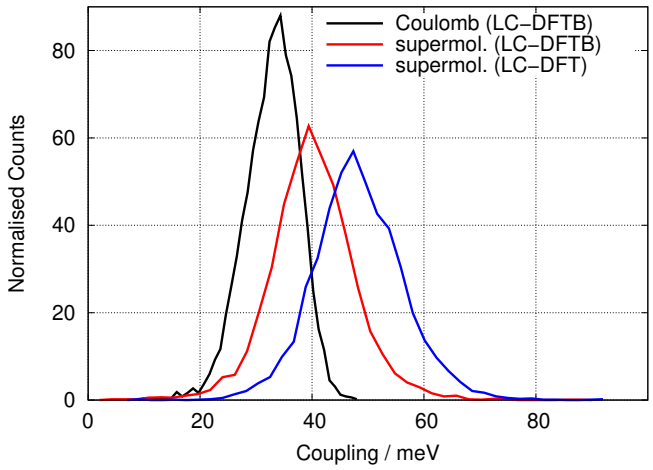


Figure 7.1.: Histograms of supermolecular excitonic couplings for sampled structures of an anthracene dimer in *b*-direction as calculated with DFTB and DFT from an analytical diagonalization approach. Additionally, DFTB Coulomb couplings are given.

Method	Type	mean	STD	RMS	C
LC-DFTB2	Coulomb	32	5	33	0.979
	supermol.	39	8	40	0.963
DFT	supermol.	47	8	48	0.972

Table 7.1.: Analysis of the histograms: mean values, standard deviations, root mean square values as in Equation 7.2 and coherence parameters as in Equation 7.3.

Differences between the diffusion constants were small and in agreement with the experimental value of 5×10^{-8} m^2 s^{-1}[136]. As expected, diffusion was faster with higher coupling values, but the influence of short range effects overall was low. Comparing values obtained by static and dynamic couplings, fluctuations turned out to have a minor impact on diffusion constants. The coupling was mostly defined by the average structure, while non-equilibrium structures seemed to account for a small additional enhancement of the coupling, leading to a slightly better agreement with the experiment.

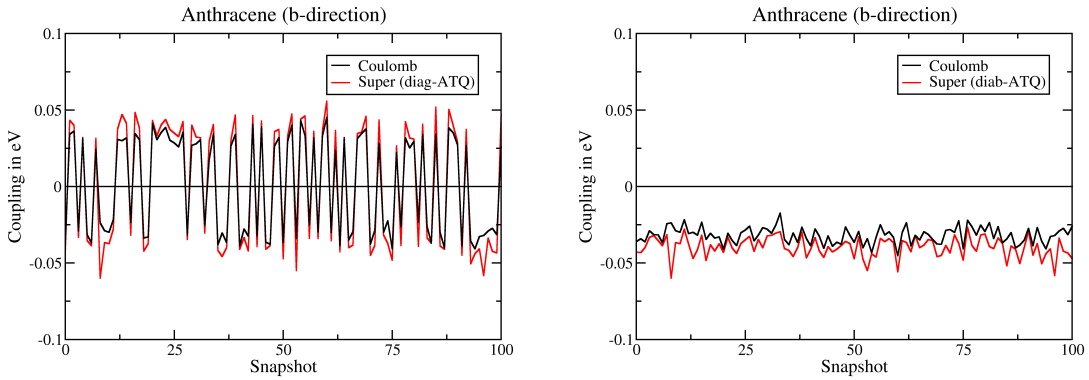
	Coulomb	Supermol.	
	LC-DFTB2	LC-DFTB2	LC-DFT
Static	2.6	4.4	5.9
Dynamic	3.5	5.2	7.6

Table 7.2.: Diffusion constants (in 10^{-8} m^2 s^{-1}) for exciton transfer along a linear chain of crystal anthracene in *b*-direction calculated with a MEQ approach with static and dynamic (RMS) coupling values. The experimental value is 5×10^{-8} m^2 s^{-1}[136].

7.3.2. An Improved Approach for Supermolecular Couplings

The simple approach for the calculation of supermolecular couplings employing the analytical diagonalization presents two issues. Firstly, if the excitation energy gap of the dimer is smaller than the difference of monomer excitation energies, the resulting coupling is imaginary and thus nonphysical. Secondly, only the magnitude of the coupling can be calculated with this approach, whereas information about the sign is lost completely. While this may not be an issue for dimers exhibiting a constantly large value of the coupling without sign changes, a continuous sign is necessary for use in transfer simulations. The relative sign must be consistent within a dimer along a trajectory and also between different pairs of fragments. These requirements hold for both Coulomb and supermolecular couplings. The sign of a set of atomic transition charges and thus the coupling is generated randomly. This is due to the phase of the orbitals (resp. the orbital coefficients) in DFTB being initialized randomly regarding their sign.

To address the issue of numerical instability, a more advanced method for the computation of supermolecular couplings was successfully implemented. It is based on a diabatization scheme and was developed by Aragó et al.[50,54]. The details are described in subsection 3.2.4. Comparing the new approach to the previously used analytical diagonalization, no additional calculations had to be run. This means both approaches are comparable regarding their computational cost.



(a) No sign tracking

(b) With sign tracking

Figure 7.2.: Coulomb and supermolecular couplings for anthracene in *b*-direction without and with the sign tracking procedure.

Until this point, the absolute value of the coupling was used for Coulomb as well as supermolecular couplings. This is a result of orbital coefficients and thus the atomic transition charges (ATQs) and transition dipoles (TDPs) fluctuating along a trajectory in an arbitrary manner, which directly transfers to nonphysical fluctuations of the coupling (cf. Figure 7.2). Test calculations showed, that artificially inverting the sign of ATQs/TDPs of one monomer inverts the sign of the coupling, while the magnitude is not affected. Performing the same for both monomers or for the dimers does not affect the sign of the coupling. Aragó et al.[50,54] proposed to set one transition charge with a large magnitude

to always be positive. If this particular charge is found to be negative, the sign of all elements of ATQs/TDPs for the specific monomer are inverted. In most of the cases this procedure led to a consistent sign for both Coulomb and supermolecular couplings from diabatization, as observed in Figure 7.2.

During excessive sampling however, some structures occurred, where this procedure could not be applied successfully. This issue occurred since the previously chosen transition charge reached values around zero and eventually were negative, although most other transition charges did not change their sign. To circumvent this problem a more general procedure was implemented. Interpreting the ATQs as an n-dimensional vector, where n is the number of atoms in the molecule, the scalar product is a measure of similarity of two different sets of ATQs. A positive scalar product indicates that both vectors point in the same direction, while a negative scalar product indicates opposing directions. Setting one set of ATQs as a reference from the beginning, all signs of other ATQs are inverted whenever the scalar product is negative. This procedure was successful in all encountered cases and is used for Coulomb and supermolecular couplings throughout the following.

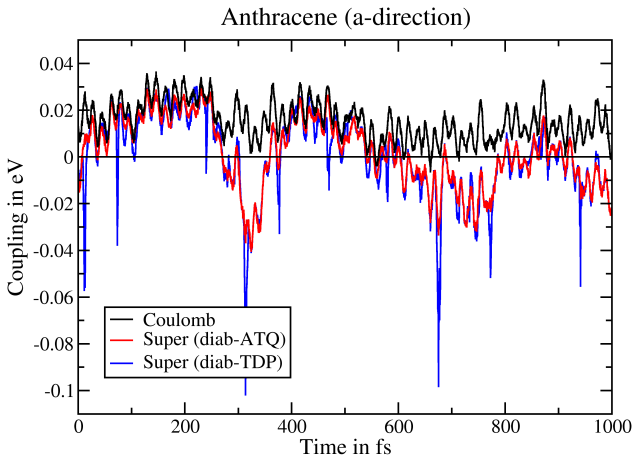


Figure 7.3.: Time series of excitonic couplings of an anthracene dimer in a-direction. Coulomb and supermolecular couplings from a diabatization with ATQs/TDPs are shown as obtained with TD-LC-DFTB2.

The distributions of couplings with corrected relative signs for anthracene in a-and b-direction are shown in the appendix (Figure C.2). For the a-direction the use of the diabatization scheme was crucial to get reasonable results, as the analytical diagonalization approach led to 53 % of imaginary values that had to be excluded. Figure 7.3 shows a time series of couplings for the a-direction in anthracene. For Coulomb and supermolecular couplings from ATQs, no observable discontinuities were observed, while some huge leaps for supermolecular couplings from TDPs occurred. This indicates that representing the adiabatic and diabatic states in terms of a three-dimensional vector is not sufficiently accurate for some geometries. The most accurate representation would be achieved through transition densities, to which atomic transition charges seem to be a reasonably fine-grained and reliable approximation.

7.4. Conclusion and Outlook

An MD simulation was used to investigate the influence of geometry fluctuations on supermolecular excitonic couplings for an anthracene crystal in comparison to Coulomb couplings. At short distances, exchange and overlap contributions can have a significant effect on excitonic couplings, enhancing the long range dominated Coulomb part. However, small shifts and tilts of the neighbors in the crystal structure, as observed during an MD simulation of a crystal, showed a decreasing effect on these contributions. Apparently, anthracene molecules are packed in such a way that short range effects are reduced. Short range molecular interactions could lead to strong Pauli repulsion in case of overlapping wave functions and could thus destabilize, e.g., the crystal structure. Therefore, stable crystal structures could exhibit conformations for which repulsion due to overlapping wave functions is minimized. This might explain the reduced importance of these effects for excitonic couplings of dimers occurring in organic crystals. Approximating the coupling in terms of Coulomb coupling and thus neglecting short range effects should therefore be justified in most cases. On the other hand, the benchmark[133] also showed an overestimation of LC-DFTB2 Coulomb couplings at long distances. Hence, a calculation of supermolecular couplings can be reasonable at least for comparison. Additionally, effects on the transfer in organic crystals cannot be completely excluded and may be interesting to investigate in direct transfer simulations with non-adiabatic MD in the future.

The diabatization scheme implemented and tested is clearly preferable to the previously used analytical diagonalization approach. Its flexibility allows for the inclusion of multiple locally excited and charge transfer states, while circumventing the numerical issues and keeping information about the sign of the coupling. Additionally, in the context of this scheme, a generally applicable procedure for generating a consistent relative sign for molecular pairs along and within trajectories was implemented, which works for both supermolecular and Coulomb couplings.

8. Exciton Transfer Simulations in Organic Semiconductors driven by Machine Learned Hamiltonians and Derivatives

8.1. Introduction

The chapters 5 and 6 mostly dealt with charge transfer. Nevertheless, within the approach used in this work, the differences between charge and exciton transfer in terms of the methodology are quite small. The most significant change is the calculation of the coarse-grained transfer Hamiltonian. For charge transfer, the orbitals of a molecule are assumed to be unaffected by the charge occupation. In practice this means, that only the orbitals of the neutral molecules have to be computed and the site energies and couplings can be obtained from the orbitals of interest (usually HOMO or LUMO).[10]

This approximation does not hold for excitons. The excitation energy is different from, e.g., the HOMO–LUMO gap, due to electron correlation. Hence, a calculation of the ground state is not sufficient and an excited state calculation becomes necessary, which comes at a higher computational cost compared to a ground state calculation.

Apart from the different origin of the transfer Hamiltonian elements, the description of the transfer itself is almost the same. The coupled equations of motion are solved with the same algorithms and there are only slight differences in the feedback from electronic to nuclear degrees of freedom. The quantum forces for explicit relaxation are calculated either from the gradients of the transfer Hamiltonian or via a Δ-SCF approach. In the latter, only the involved states are of a different nature.

In this chapter, neural network (NN) models are utilized for the prediction of exciton transfer Hamiltonian elements as well as their diagonal and off-diagonal derivatives for anthracene and diindenoperylene.

8.2. Computational Details

8.2.1. Generation of Training Data

An anthracene crystal was created based on an experimental crystal structure[92], containing $5 \times 30 \times 5$ molecules along the *a*-, *b*- and *c*-crystallographic axes, respectively. For diindenoperylene, the respective crystal[137] had a size of $15 \times 10 \times 5$ molecules. Force field parameters for both systems were derived from the general AMBER force field (GAFF)[93,94]. Atomic charges were generated from the restrained electrostatic potential (RESP) fitting procedure[95,96], calculated at HF/6-31G*[97,98] level of theory using `Gaussian` 09[99].

After an initial energy minimization the systems were equilibrated for 1 ns with a time step of 2 fs using the modified Berendsen thermostat[138] at 500 K in the NVT ensemble. Subsequently, a productive classical MD simulation in the ground state was run for every system at the same temperature for 10 ns with a time step of 2 fs. Here, the Nose-Hoover thermostat[102] was employed. Ground state structures were saved every 5 ps. Excited state structures were sampled in 38 QM/MM simulations with the same MD settings, with the exception that simulation time was reduced to 200 ps with a time step of 1 fs and structures were saved every 100 fs. The shorter simulation time was due to the computational cost of QM/MM simulations. In every QM/MM simulation only one molecule was placed inside the QM region, which was simulated in an excited state.

All MM MD simulations were performed with the `GROMACS` 2020.2 software package[100,101], while for QM/MM MD simulations a modified version of `GROMACS` 2021 was used. The latter version is combined with the `DFTB+`[89,115] program. As a QM method, self-consistent-charge density functional tight binding[22] utilizing a long-range corrected functional[25,26] and the time-dependent linear-response extension[28,29] (TD-LC-DFTB2) was used. The ob2-1-1/base[139] parametrization was employed.

This resulted in a set of 60 000 single molecule as well as 60 000 pair structures for both ground and excited state, which were taken from the center region of the crystals. This applied to both the anthracene and diindenoperylene systems. Only nearest neighbor pairs were considered. Subsequently, the elements of the transfer Hamiltonian and their derivatives were calculated. Site energies, atomic transition charges as well as ground and excited state forces for single molecules were obtained in vacuum from TD-LC-DFTB2 as implemented in `DFTB+`, taking the parametrization from reference [29]. Additionally, supermolecular calculations of dimers were performed with the same method and parameter set.

Derivatives of site energies were calculated with the Δ-SCF approach (Equation 3.11) as the difference of excited and ground state forces, similar to the strategy used in the previous chapter for charge transfer (chapter 6). Three different types of couplings were computed. Static Coulomb couplings were obtained with only one set of atomic transition charges, which was calculated on the unoptimized experimental crystal structure of a single molecule once and used for all subsequent calculations. In contrast, dynamic Coulomb couplings employ on-the-fly atomic transition charges, which change for every geometry.

Lastly, supermolecular couplings were computed with a diabatization scheme based on atomic transition charges of the monomers and pairs.

As mentioned, for static coulomb couplings the transition charges were kept constant for different geometries. Thus, it was possible to calculate the derivative analytically from the derivative of the Coulomb interaction ζ in Equation 3.22 with respect to atomic coordinates. The derivatives for the other coupling types were not obtained, as there are currently no implemented routines for derivatives of the atomic transition charges needed for derivatives of dynamic Coulomb couplings.

8.2.2. Training and Evaluation of Machine-Learned Models

The same neural network architecture[84] as presented in subsection 6.2.2 was used. Models were trained on site energies and their gradients, static Coulomb couplings and their gradients, or only dynamic Coulomb or supermolecular couplings, respectively. In contrast to chapter 5, the sign tracking scheme as tested in chapter 7 was employed and NN models were trained on the signed couplings and not on the absolute values. This is necessary to achieve a differentiable function in case of crossing zero, which is a prerequisite for the calculation of gradients. The models used the spatial coordinates of the atoms in the system as inputs, from which inverse interatomic distances between atoms were calculated as a translationally and rotationally invariant representation. For site energies and their gradients, all intramolecular atomic distances were used. The static Coulomb couplings and their respective gradients were represented by only the intermolecular part of the matrix, while for dynamic Coulomb and supermolecular couplings the full matrix (intra- and intermolecular part) was chosen.

All models received training and validation data in a 9:1 ratio, with 50 000 and 30 000 total data points for the training of site energies and couplings, respectively. The loss on the validation set was monitored every epoch and training was aborted if it did not improve for more than 20 epochs.

For the training of models including gradients, the loss was calculated on both the network's prediction for energies and gradients with respect to input coordinates. The values were multiplied with weighting coefficients, determining the importance of their contribution, before summing them up for the total loss function. The relative weights of these two parts were considered as hyperparameters of the models. Merged R^2 scores were set as the objective in hyperparameter optimization.

All details on the network configuration, training and hyperparameter search can be found in the appendix Table D.1.

8.2.3. Exciton Transfer Simulations

Separate crystals were generated for the transfer simulations containing $54 \times 10 \times 5$ (16) molecules for anthracene in *a*-direction and $10 \times 54 \times 5$ (16) molecules for anthracene

in *b*-direction. One-dimensional lines of molecules along the respective directions in the middle of each crystal were chosen for the QM region, with the number of fragments specified in brackets above. After equilibration at 300 K, 500 structures in equidistant time intervals of 2 ps were chosen as starting structures for subsequent simulations of exciton transfer. The wave function was initially localized on the first molecule of the linear chain. A time step of 0.1 fs was used for the propagation of nuclei and the TDSE was integrated numerically with the fourth-order Runge-Kutta algorithm with an integration time step of 0.01 fs. The fewest switches surface hopping method with the Boltzmann-correction (BC-SH) was used for most of the NAMD simulations with the explicit relaxation (ER) scheme. If derivatives of the couplings were available, the SH method was employed without the BC approximation, but employing the correct re-scaling scheme with non-adiabatic coupling vectors (NACVs). Averages of observables were calculated on 500 trajectories, which were simulated for 1 ps each, while the first 200 fs were regarded as initial equilibration of the exciton. All transfer simulations were performed within a local version of `GROMACS` 4.6 where the NN as well as the SH/BC-SH methods were implemented. Values of mean square displacement (MSD), inverse participation ratio (IPR) and exciton diffusion constants were obtained from the simulations as averaged over all run trajectories (c.f. subsection 3.2.5).

8.3. Results and Discussion

8.3.1. Analysis and Preparation of the Data-Set

Consistency of calculated properties and generated data sets is important, especially for the application in machine learning. If there are inconsistent or wrong data points, the models may be prevented from finding the desired patterns in the overall data set and from learning their data accurately. Therefore, the generated data should be carefully analyzed beforehand, which is demonstrated in the following.

For all investigated anthracene structures, the energetically lowest excitation was dominated by the HOMO to LUMO transition and additionally showed the highest oscillator strength. This state was separated from higher lying excitations by an energy gap of approximately 800 meV. Hence, this state was chosen as the state of interest for the generation of the data sets. For supermolecular calculations the diabatization was performed with the two lowest transitions of every monomer.

The distributions of site energies are depicted in Figure 8.1 and showed a Gaussian shape. Site energies for ground state structures had a mean of 3.03 eV with a standard deviation of 0.14 eV. While the mean value for excited state structures was shifted to a lower value of 2.80 eV, the distribution was broader with a standard deviation of 0.17 eV. This was expected due to the relaxation of the molecule in the excited state. The difference in site energies from ground and excited state is an estimate of the reorganization energy. Here, this value was not meaningful, because the ground state structures from the force field and the QM method deviate.

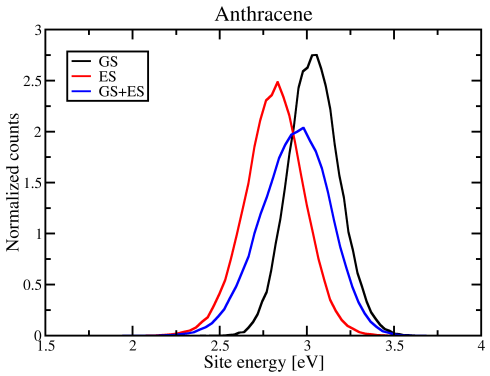
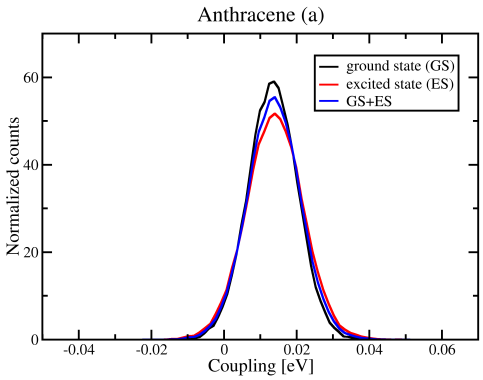


Figure 8.1.: Distribution of site energies for anthracene molecules. Geometries were sampled in ground and excited state.

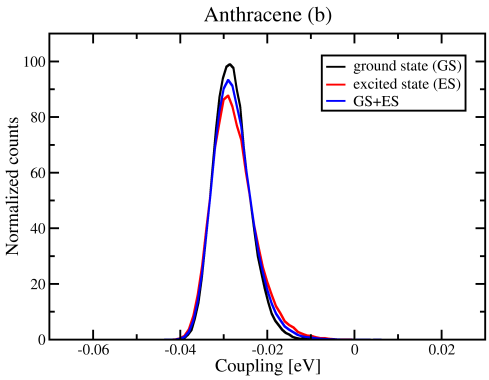
The distributions of couplings for both crystallographic directions and for all coupling types are shown in Figure 8.2. Again, all distributions exhibited a Gaussian shape. No significant differences between ground and excited state structures were observed. Mean values for static and dynamic Coulomb couplings for both *a*- and *b*-directions were approximately the same with 13 meV and 30 meV, respectively. Standard deviations slightly increased when including fluctuations of ATQs. Supermolecular couplings for the *b*-direction were comparable to Coulomb couplings, while for the *a*-direction, the supermolecular couplings were only half the size of the Coulomb couplings. Results for the *b*-direction were in agreement with reported values of 29 meV for DFT TrEsp couplings[140] , 25 meV for supermolecular couplings from simplified analytic diagonalization with SCS-CC2[110] and 35 meV from DFT supermolecular couplings employing the diabatization scheme. A reported value for supermolecular coupling (7 meV)[50] in the more distant and tilted *a*-direction agrees well with results results shown here, while Coulomb couplings are overestimated compared to reported DFT TrEsp couplings (5 meV)[140].

Choosing the correct states for diindenoperylene was more involved, as this molecule exhibited state swapping along the simulated trajectories. For most of the generated structures, the energetically lowest excitation was identified as the state of interest, showing the highest oscillator strength and being dominated by the HOMO to LUMO transition. The second excited state in these cases had an oscillator strength near zero and was dominated by the HOMO-1 to LUMO transition.

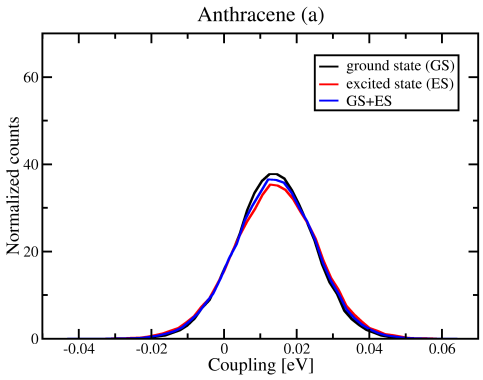
However, for roughly 10 % and 30 % of the ground and excited state structures, respectively, the first and second excited states swapped. This was also indicated by the shapes of the involved orbitals, which did not change significantly along the trajectory. DFT calculations with the ωB97X (def2-TZVP) functional supported these findings, both in terms of orbital shapes and composition of excited states. In fact, this seemed to be a DFTB issue, as no state swapping was present for a few representative DFT calculations. Distributions of site energies and couplings without the correction of state swapping are shown in the appendix (Figure D.1) along with DFTB and DFT orbitals (Figure D.3).



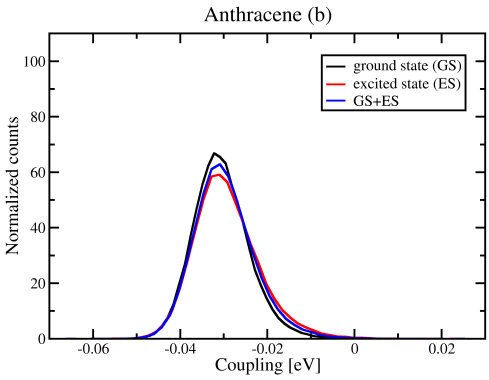
(a) Static Coulomb couplings



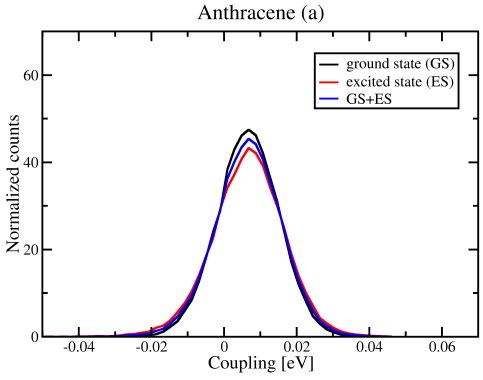
(b) Static Coulomb couplings



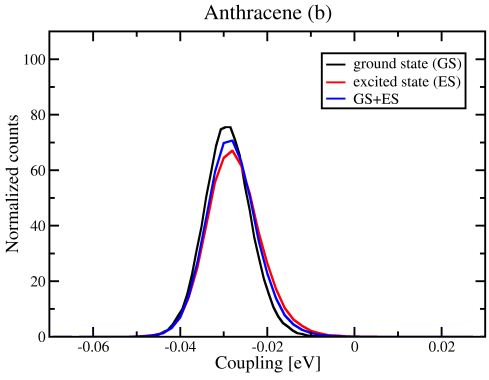
(c) Dynamic Coulomb couplings



(d) Dynamic Coulomb couplings



(e) Supermolecular couplings



(f) Supermolecular couplings

Figure 8.2.: Distributions of couplings for first neighbors of anthracene pairs in *a*- (left) and *b*-direction (right). Static (constant ATQs, top) and dynamic (varying ATQs, middle) Coulomb couplings as well as supermolecular couplings (bottom). Geometries were sampled in ground and excited state.

Additionally, some structures were excluded from the data set, because the first and second excited states were mixed and could not be ascribed with certainty. For instance, the first and second excited states were both found to be a mixture of transitions from HOMO to LUMO and HOMO-1 to LUMO. This behavior was observed in less than 2 % of the structures. Choosing the correct state was important to obtain consistent data and led to reasonable distributions of site energies and dynamic Coulomb couplings with fluctuating ATQs. Static Coulomb couplings with fixed ATQs do not exhibit this issue, as the same set of ATQs is used throughout the calculations and only the geometry changes.

Moreover, an issue was noticed when calculating supermolecular couplings for diindenoperylene. A two state diabatization with the correctly chosen states led to long tails in the distributions (Figure D.2). This is rooted in the low energetic separation of the first and second excited states of the molecule, which were found to swap for some structures as mentioned before. As a consequence, the adiabatic states are combinations of the two close lying diabatic states. The issue was solved by including the correctly chosen second excited states in a four state diabatization.

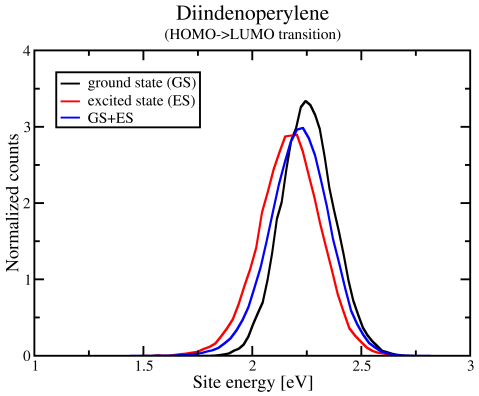
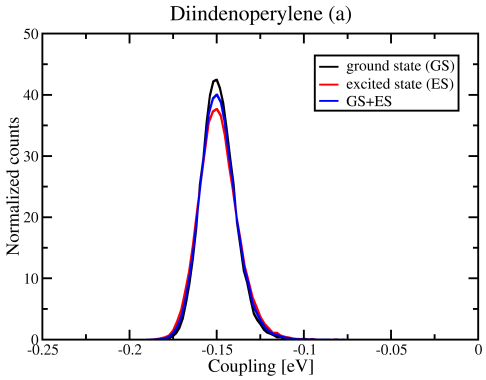


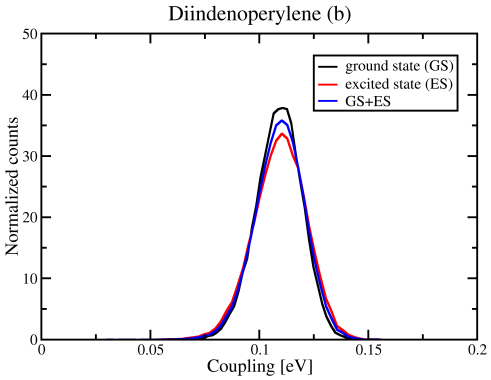
Figure 8.3.: Distribution of site energies for diindenoperylene molecules. Geometries were sampled in ground and excited state.

All distributions of corrected site energies and couplings can be found in Figure 8.3 and Figure 8.4. The difference of mean values of site energies from ground and excited state structures as well as the standard deviations were lower compared to anthracene, while couplings were generally much higher. Again, mean values for Coulomb couplings with constant and varying ATQs were similar, but the standard deviations for the latter were approximately doubled. This reflects the neglect of the fluctuations of the ATQs as influence on coupling values. Supermolecular couplings were significantly lower compared to Coulomb couplings, but agree well with a reported value (74 meV)[110] from an analytic diagonalization approach for the *a*-direction employing SCS-ADC(2).

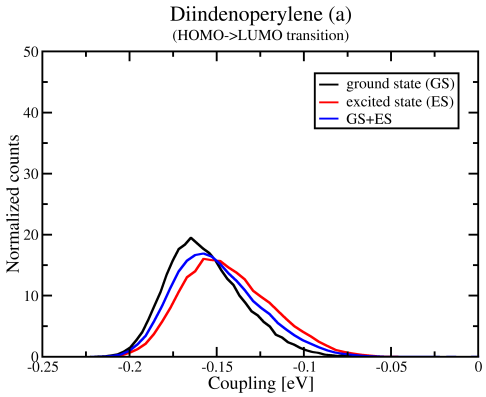
Couplings from supermolecular calculations encompass all short- and long-range contributions and should by definition be larger compared to Coulomb type couplings, which only include the long-range part. As indicated from previous applications of TD-LC-DFTB2 Coulomb couplings to organic molecules[133] and Bacteriochloriphyll a molecules[113],



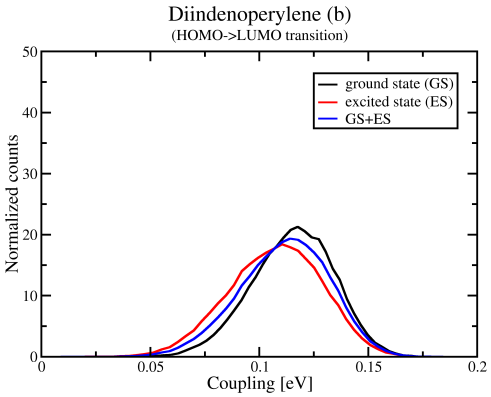
(a) Static Coulomb couplings



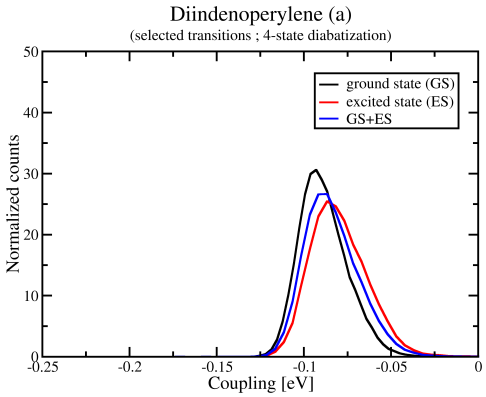
(b) Static Coulomb couplings



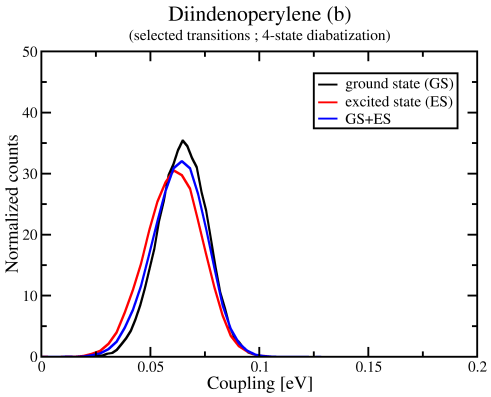
(c) Dynamic Coulomb couplings



(d) Dynamic Coulomb couplings



(e) Supermolecular couplings



(f) Supermolecular couplings

Figure 8.4.: Distributions of couplings for first neighbors of diindenoperylene pairs in *a*- (left) and *b*-direction (right). Static (constant ATQs, top) and dynamic (varying ATQs, middle) Coulomb couplings as well as supermolecular couplings (bottom). Geometries were sampled in ground and excited state.

Coulomb couplings were overestimated compared to supermolecular couplings. This is in line with the findings in this work.

8.3.2. Model Training and Evaluation

All trained models for anthracene converged within 1000 epochs. Similar to chapter 6, the hyperparamters seemed to have a minor impact on the quality of the models. Here, their predictions on 30 000 unseen training data points were used to evaluate the models. Error metrics as well as coefficients of determinations are summarized in Table 8.1 and Table 8.2.

	site energy		stat. Coulomb cpl.	
	energy	gradient	energy	gradient
MAE	4.692	14.802	0.137	0.040
max err	83.559	966.415	1.953	4.294
R^2	0.999	0.999	1.000	1.000

Table 8.1.: Quality metrics for models predicting site energies and their derivatives as well as static Coulomb couplings and their derivatives for anthracene: mean absolute error (MAE), maximum error and coefficient of determination (R^2). All metrics except R^2-scores in meV (meV Å^{-1}).

Site energies and their derivatives as well as static Coulomb couplings and their derivatives were learned almost perfectly. All R^2 values were near 1. Mean absolute errors were orders of magnitude lower compared to the mean values of the references. Additionally, no strong outliers can be seen in Figure 8.5.

	dyn. Coulomb cpl.	supermol. cpl.
MAE	3.055	2.716
max err	27.305	129.689
R^2	0.972	0.958

Table 8.2.: Quality metrics for models predicting dynamic Coulomb couplings and supermolecular couplings for anthracene: mean absolute error (MAE), maximum error and coefficient of determination (R^2). All metrics except R^2-scores in meV.

Usually, couplings for the application to charge and exciton transfer are represented by the intermolecular part of the inverse distance matrix, as mentioned and used in chapter 6. This was also successfully performed for static Coulomb couplings. For the two other coupling types, however, this representation led to a low accuracy of the model predictions. Switching to the full representation by inclusion of inter- and intramolecular parts of the inverse distance matrix, the quality of the models was improved significantly. This may be rooted in the internal geometric fluctuations, which affect these coupling types strongly and which are represented in the intramolecular blocks of the matrix, rather than in the intermolecular block. Still, the dynamic Coulomb couplings and supermolecular

couplings were learned less well compared to static Coulomb couplings. Mean absolute errors were one order of magnitude higher and R^2 values only reached 0.96 to 0.97. The distributions of prediction versus reference values shown in Figure 8.6 were broader and showed more deviations. For supermolecular couplings, a few outliers were noted. Nonetheless, both models were comparable in quality to the pentacene models for charge transfer couplings (c.f. subsubsection 6.3.1.1) in terms of quality and thus had to be tested in NAMD simulations to judge the influence of errors.

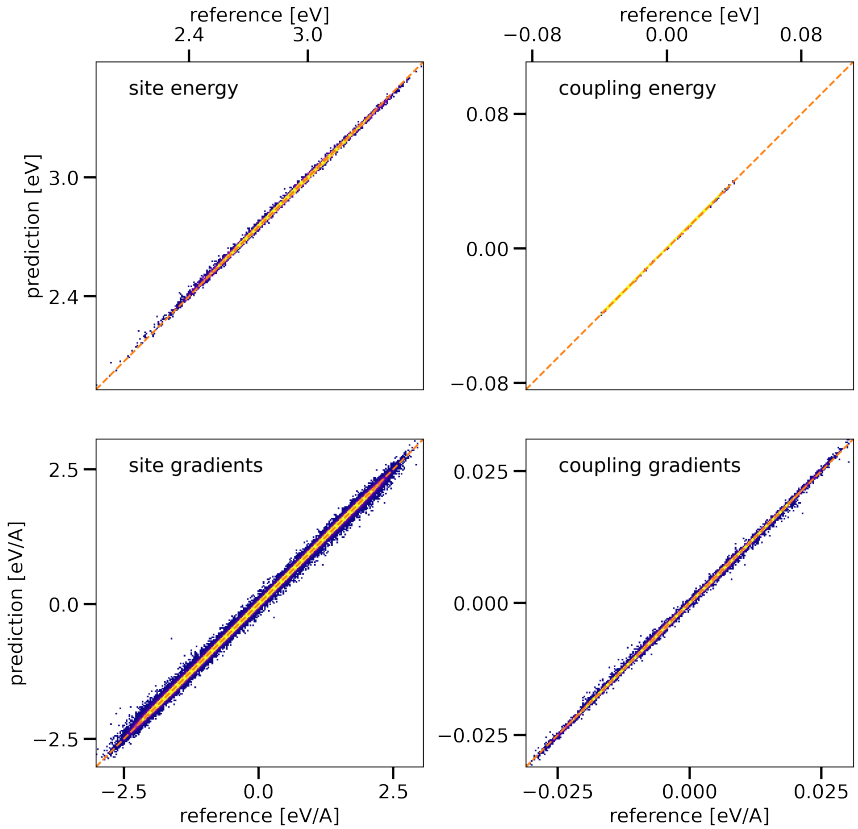


Figure 8.5.: Two-dimensional histograms of site energies, static Coulomb couplings and their gradients: model predictions vs. TD-LC-DFTB2 targets for 30 000 structures in anthracene. Bright colors indicate high data density, unoccupied areas shown in white.

First tests for the training of models for diindenoperylene were also performed and can be found in the appendix (section D). This larger molecule appeared to be more difficult to learn and models still have to be optimized further. Therefore, NAMD simulations are yet to be performed for this molecule.

8.3.3. Time Evolution of Hamiltonian Elements

A better estimate of model quality can be given by the investigation of the evolution of site energies and couplings along simulation time. Therefore, reference values were computed with TD-LC-DFTB2 along 1 ps NAMD trajectories of arbitrary first neighbor pairs from the

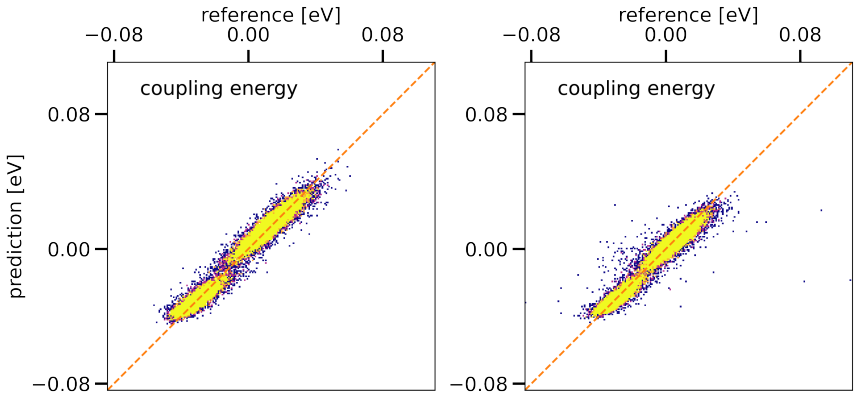


Figure 8.6.: Two-dimensional histograms of dynamic Coulomb couplings and supermolecular couplings: model predictions vs. TD-LC-DFTB2 targets for 30 000 structures in anthracene. Bright colors indicate high data density, unoccupied areas shown in white.

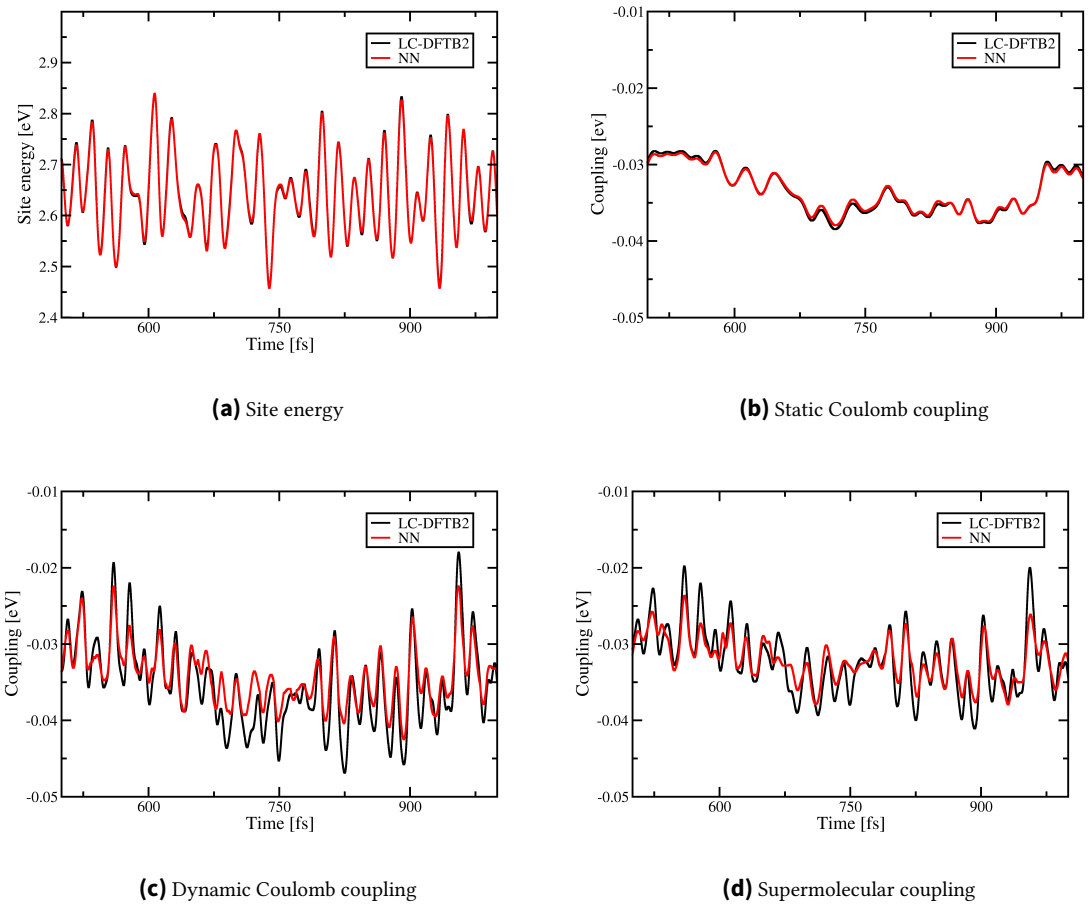
investigated anthracene crystals and compared with predicted values. Throughout these simulations, one molecule carried an exciton and was relaxed by quantum forces predicted by NN models, while the other was in the ground state. The time-series of site energies and couplings for anthracene in *b*-direction are shown in Figure 8.7. The corresponding plots for the *a*-direction can be found in the appendix (Figure D.4).

Predictions and reference values for site energies and static Coulomb couplings exhibited negligible deviations, which highlights the quality of the models further. Dynamic Coulomb and supermolecular couplings showed minor deviations, but were still able to reproduce most of the fluctuations of the reference. Similar results were obtained for anthracene in *a*-direction. These results are promising with regard to an application in transfer simulations.

8.3.4. Comparison of Obtained Observables

In the following, the trained NN models were applied to perform NAMD simulations of exciton transfer along linear chains of anthracene molecules in *a*- and *b*-direction. Simulations were performed either with the BC-SH method, in cases where no derivatives of couplings were available for the calculation of NACVs, or with the SH method, employing the correct re-scaling scheme for atomic momenta. Throughout all simulations, the explicit relaxation scheme was used, either with onsite or full relaxation. The former includes only quantum forces obtained from diagonal (site energy) derivatives, while the latter additionally includes off-diagonal (coupling) derivatives. Obtained exciton diffusion constants are shown in Table 8.3 with the corresponding curves of averaged MSDs in the appendix (section D).

For anthracene in *b*-direction, diffusion constants obtained from BC-SH simulations employing onsite relaxation and various coupling types showed rather low deviations. Higher values from dynamic Coulomb couplings as well as supermolecular couplings may be the result of reduced coupling fluctuations due to poorer predictions compared to static



(a) Site energy

(b) Static Coulomb coupling

(c) Dynamic Coulomb coupling

(d) Supermolecular coupling

Figure 8.7.: Time evolution of anthracene site energies as well as different coupling types for anthracene in *b*-direction. Values were calculated with the TD-LC-DFTB2 reference and NN models.

Method	Relaxation	coupling	ANT-b	ANT-a
BC-SH	onsite	stat. Coul.	8.9	2.1
		dyn. Coul.	9.5	1.4
		supermol.	10.1	0.8
SH	onsite	stat. Coul.	4.3	0.5
	full	stat. Coul.	4.8	0.4
Experiment[141]			5.0	1.8

Table 8.3.: Exciton diffusion constants (in 10^{-3} cm^2 s^{-1}) for anthracene.

Coulomb couplings. Results in the *a*-direction were significantly smaller, as the average coupling in this direction is lower. The diffusion constant obtained for supermolecular couplings is strongly reduced, compared to static Coulomb couplings. Here, the overestimation of Coulomb couplings became visible.

The Boltzmann-correction led to an overestimation of the diffusion coefficient for the *b*-direction and switching to the correct re-scaling scheme in SH simulations led to a reduction of the values by a factor of two. The latter were in close agreement with the experimental value. Moreover, the diffusion constants for the *a*-direction were decreased in SH simulations, but were underestimated compared to the experiment. It has to be noted, that the values for this direction were generally quite small compared to exciton diffusion constants of other materials[140], and thus are prone to errors. The inclusion of derivatives for couplings in the full ER scheme did not show a significant effect. Therefore, neglecting these terms is reasonable.

Inverse participation ratios were computed as 1.1 and 1.0 for the two directions *b* and *a*, respectively. This indicates a strongly localized hopping mechanism for the transfer. The obtained diffusion constants and IPR values were in agreement with recently published simulation results[140].

Effective reorganization energies were estimated as averages of site energies for a ground and excited state molecule in a 10 ps simulation without transfer. The obtained value of 410 meV is roughly 17 % smaller compared to the actual reorganization energy, which was calculated with TD-LC-DFTB2 via four-point calculation (494 meV). This supports the evidence from the analysis of the charge transfer simulations in chapter 6 and solidifies, that it is crucial to compute the site energies and gradients with the same method. Nevertheless, the deviations indicate either an introduction of small errors by the ML models or a conflict of quantum and classical forces from the force field.

8.4. Conclusion and Outlook

In this chapter, non-adiabatic molecular dynamics simulations of exciton transfer were successfully performed. The simulations were driven by neural network models, which predicted Hamiltonian elements and their gradients.

Generating data sets for the training of machine learned models was much more time consuming compared to charge transfer in terms of both computational effort and data engineering. While data generation for anthracene was comparatively easy, the choice of the correct transitions for diindenoperylene was more difficult. Additionally, the calculation of supermolecular couplings for the latter molecule required the inclusion of two diabatic states on every monomer and thus four adiabatic states on the dimer.

The trained models for anthracene showed high performance. Site energies, static Coulomb couplings and the gradients for both were predicted with near perfect accuracy. Models for dynamic Coulomb couplings and supermolecular couplings performed less well, but were still able to reproduce most of the coupling fluctuations along a time-series, which were far more pronounced in general.

For the *b*-direction of anthracene, utilizing different coupling types did not lead to significant changes in diffusion constants, while Coulomb couplings seemed to be overestimated

in the *a*-direction and supermolecular couplings led to a decrease of the diffusion constant. Employing the correct re-scaling scheme with NACVs decreased exciton diffusion constants. The obtained value for the *b*-direction was in close agreement with the experiment. Contributions of couplings to relaxation were negligible.

Anthracene showed a strongly localized and incoherent transfer mechanism, as indicated by the IPR values of 1 for both directions. Diffusion constants as well as IPR values were in agreement with reported simulation results. The effective reorganization energy was underestimated by 17 %.

For the calculation of TD-LC-DFTB2 reference values, the time limiting step for the Hamiltonian of the Coulomb type is the calculation of site energies and atomic transition charges, which was in the order of about 1 s for anthracene. Supermolecular couplings are far more costly, taking at least two computations of the dimer amounting to an increase of computational cost by one order of magnitude. Explicit quantum forces for relaxation require about 5 s to be calculated. The time limiting prediction in NN driven simulations is the coupling prediction, which takes roughly 0.1 ms independent of the coupling type. Overall, the NN models outperformed DFTB by 4 to 5 orders of magnitude for anthracene. The favorable scaling with the system size will only reinforce this performance.

Employing a fast data-driven approach is what makes these kind of cost intense simulations possible in the first place. Now, this methodology needs to be evaluated carefully on a bigger benchmark set to test its accuracy. Additionally, the relevance of charge transfer excitations, which are known to couple with pure locally excited Frenkel states, should be investigated. This may be relevent for organic molceules as well as biological light-harvesting complexes.[142,143]

9. Application: Exciton Transfer Simulations in the Light-Harvesting Complex II

Author Contributions:
The generation of Δ-charges was done by Dr. Sebastian Höfener and Dr. Monja Sokolov. All simulations and analyses were performed by myself.

9.1. Introduction

Photosynthesis is one of the most fundamental energy conversion processes. Through the absorption of sunlight by chromophore molecules, an excitation or exciton is created. This excitation and the contained energy is subsequently transferred to a reaction center, where it is utilized for a charge separation. Further electron transfer steps allow the synthesis of energy-rich chemical compounds, e.g. adenosinetriphosphate (ATP), which provide the energy for biochemical reactions in living organisms.[144]

The first step of photosynthesis, the collection and transfer of light, is called light-harvesting. Biological light-harvesting complexes operate at extremely high quantum efficiency, meaning the probability of an absorbed photon to reach a reaction center and drive charge separation is high.[144] To date, artificial appliances lack this highly desirable efficiency. Therefore, these so-called antennae are interesting compounds for an inspiration of a new generation of photovoltaic devices, concentrators or sensitizers for solar cells and other appliances.[44]

One of the best studied antennae is the major light-harvesting complex derived from anoxygenic purple bacteria, e.g. from the organism *Rhodospirillum molischianum* (*Rs. molischianum*) – light-harvesting complex II (LH2). In this α-helical transmembrane protein, the protein scaffold serves as an embedding for the photosynthetic pigments. The two bands in the absorption spectrum of this complex at 800 nm and 850 nm are caused by the absorption of Bacteriochlorophyll a (BChl a) chromophores, which are arranged in two rings of 8 (B800) and 16 (B850) molecules inside the protein complex (c.f. Figure 9.1).[145]

Numerous studies[146–149] have been conducted to investigate the properties of this and similar systems, which mostly relied on a ground state description of the complex. Thus, the dynamics of the excited state as arising from the coupled motion of electronic and

nuclear degrees of freedom could not be simulated. Only these kind of simulations would be able to reliably reveal transfer mechanisms and time scales. The size of the system, with 24 chromophores and more than 100 atoms each, renders the necessary quantum mechanical (QM) description and thus the direct dynamics challenging.

So far, only one study[150] managed to perform such a simulation, at the investment of tremendous computational resources. Sisto et al. employed an *ab initio* non-adiabatic molecular dynamics (NAMD) simulation of the LH2 complex from *Rhodoblastus acidophilus* with time-dependent density functional theory (TD-DFT) and the surface hopping (SH) approach. Due to the extreme computational demands of such simulations, only a single trajectory of 300 fs length was run via a highly parallellized GPU accelerated approach. Longer trajectories and more sampling of different initial conditions would be necessary to obtain observables reliably.

With the methodological progress presented throughout the previous chapters, it is now possible to perform on-the-fly exciton transfer simulations with NAMD methods for LH2 without having to invest an enormous amount of resources. This becomes feasible through the availability of neural networks for the efficient prediction of exciton transfer Hamiltonian elements. In the following, preliminary results are presented as a proof of principle application to light-harvesting complexes. 1000 trajectories with a length of 10 ps were run for each ring separately.

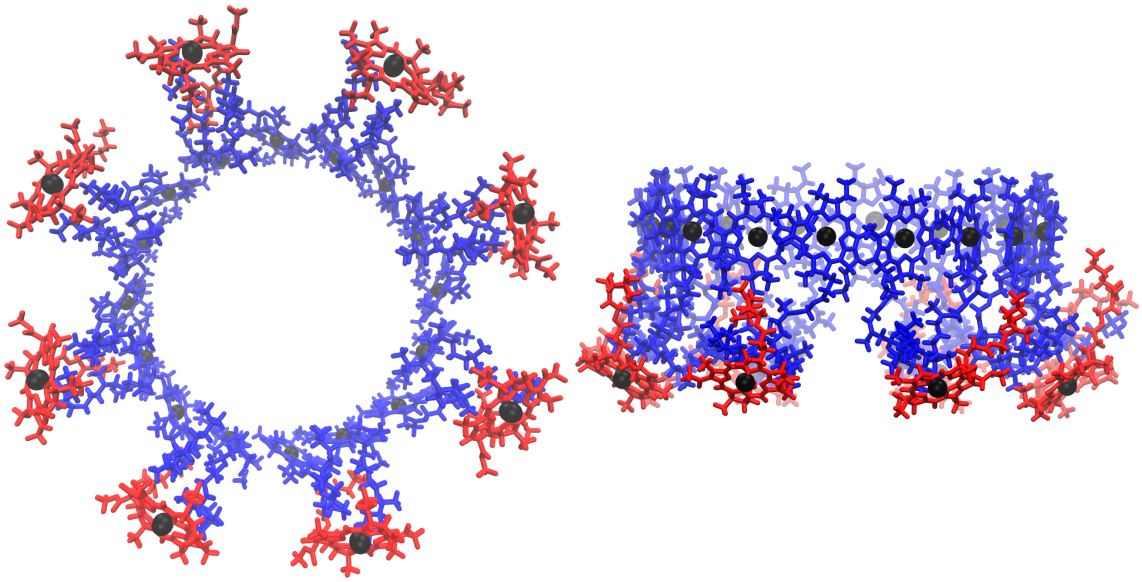


Figure 9.1.: Top (left) and side (right) view of BChl a chromophores in LH2. B800 depicted in red, B850 in blue. The magnesium ions are given as black spheres.

9.2. Computational Details

9.2.1. Generation of Training Data

Subject to this investigation is the LH2 complex from the purple bacterium *Rs. molischianum*. The simulated system was taken from reference [151] and is based on the crystal structure with the PDB code 1LGH[145]. The complex was inserted into a lipid bilayer (1-Palmitoyl-2-oleoylphosphatidylcholine) and solvated in water.

For classical MD simulations the `GROMACS` program package[100,101] (version 2020.2) with the CHARMM27[152] force field was used. After an initial energy minimization the temperature was equilibrated at 300 K for 5 ns using the Berendsen thermostat[153] and position restraints on the protein, chromophores and lipids. Afterwards, the pressure was equilibrated for 5 ns at 1 bar with the Parrinello-Rahman barostat[154]. Subsequently, the position restraints were removed in two simulation steps of 5 ns length each, first for the lipids, then for the protein and chromophores. To sample structures, productive simulations of 200 ns with the Nose-Hoover thermostat[102] and the Parrinello-Rahman barostat were performed. Atomic positions were saved every 40 ps, resulting in 5000 snapshots. The time step in all simulation steps was 1 fs.

Subsequently, the elements of the transfer Hamiltonian were calculated. Site energies and atomic transition charges were obtained in vacuum from long-range corrected self-consistent-charge density functional tight binding[22,25,26] with the time-dependent linear-response extension[28,29] (TD-LC-DFTB2) as implemented in `DFTB+`[89,115], taking the parametrization from reference [29]. Static Coulomb couplings were obtained with only one set of atomic transition charges, which was calculated on the unoptimized experimental crystal structure of a single BChl a molecule once and employed for all subsequent calculations. Only nearest neighbor pairs inside the two respective rings of BChl molecules were considered. To restrict the computational cost, the phytyl tail of the BChl a molecule is not considered in the QM calculations. This reduces the QM region from 140 to 85 atoms per molecule.

9.2.2. Training and Evaluation of Machine-Learned Models

The same neural network architecture[84] as presented in subsection 6.2.2 was used. One model was trained for the prediction of site energies for all BChl a molecules. A second model was trained on Coulomb couplings for all intra-ring couplings of only nearest neighbors. The models used the spatial coordinates of the atoms in the system as inputs, from which inverse interatomic distances between atoms were calculated as a translationally and rotationally invariant representation. For site energies, all intramolecular atomic distances were used, while for couplings only the intermolecular part was chosen.

All models received training and validation data in a 9:1 ratio, with 100 000 and 50 000 total data points for the training of site energies and couplings, respectively. The loss

on the validation set was monitored every epoch and training was aborted, if it did not improve for more than 500 epochs for site energies and 100 epochs for couplings.

All details on the network configuration, training and hyperparameter search can be found in the appendix section E.

9.2.3. Exciton Transfer Simulations

The same starting structure used for the sampling of training data was utilized to sample structures in equidistant time intervals of 1 ps as initial condition for subsequent NAMD simulations of exciton transfer. Both rings of chromophores (B800, B850) were simulated separately, because no inter-ring couplings were calculated so far. The wave function was initially localized on one molecule of the respective rings. A time step of 0.1 fs was used for the propagation of nuclei and the TDSE was integrated numerically with the fourth-order Runge-Kutta algorithm with an integration time step of 0.01 fs.

The fewest switches surface hopping method with the Boltzmann-correction (BC-SH) was used for the NAMD simulations with the implicit relaxation (IR) scheme. The reorganization energy of 65 meV was taken from reference [155], which depicted an experimental determination.

For an additional feedback from electronic to nuclear degrees of freedom, Δ-charges (Δq) were obtained from TD-DFT calculations as the difference of atomic partial charges in the excited and ground state. For the ground state structure of BChl a, optimized with B3LYP[119] (6-31G(d,p)[98]), charges in the ground and excited state were obtained via an ESP fit with CAM-B3LYP[156] within the `KOALA` program[157]. These charges were multiplied with the diabatic occupation and added to the force field charges (q):

$$q'_{ii} = q_{ii} + |a_{ii}|^2 \Delta q_{ii} \,. \tag{9.1}$$

1000 trajectories were run for 10 ps each for both of the two rings. All transfer simulations were performed within a local version of `GROMACS` 4.6, where the NN as well as the BC-SH methods were implemented. Values of mean square displacement (MSD) and inverse participation ratio (IPR) were obtained from the simulations, averaged over all run trajectories (c.f. subsection 3.2.5).

9.3. Results and Discussion

9.3.1. Model Training and Evaluation

Analyzing the obtained data set revealed the lowest excited state to be the state of interest, which was the transition with the highest oscillator strength and was dominated by the HOMO to LUMO transition. No state swapping occurred.

All trained models for BChl a converged within 1000 epochs. Here, the quality of their predictions was evaluated on 10 000 unseen training data points. Error metrics as well as coefficients of determinations are summarized in Table 9.1, two-dimensional histograms of model predictions versus TD-LC-DFTB2 reference values are given in Figure 9.2

	site energy	coupling
MAE	10.466	0.345
max err	112.009	4.939
R^2	0.876	1.000

Table 9.1.: Quality metrics for models predicting site energies and couplings for Bacteriochlorophyll a: mean absolute error (MAE), maximum error and coefficient of determination (R^2). All metrics except R^2-scores in meV.

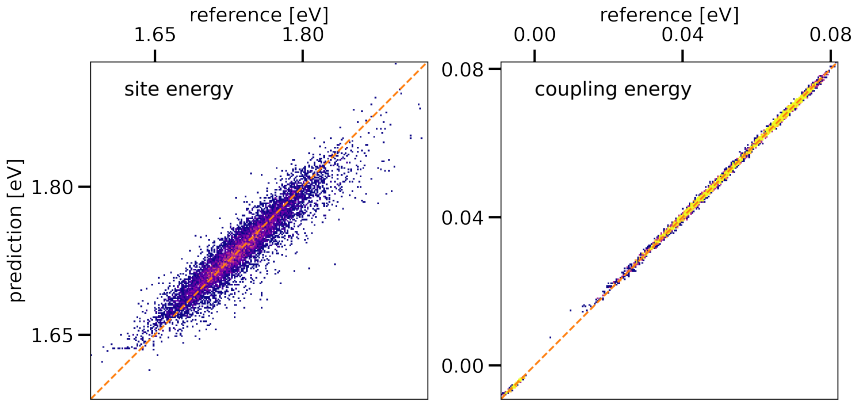


Figure 9.2.: Two-dimensional histograms of site energies and couplings: model predictions vs. TD-LC-DFTB2 targets for 10 000 structures of BChl a. Bright colors indicate high data density, unoccupied areas shown in white.

Models for BChl a site energies required significantly more data compared to anthracene in the previous chapter. 100 000 data points were necessary to reach a coefficient of determination of only 0.9. The mean absolute error of 10 meV was still more than two orders of magnitude lower compared to the mean of the reference values. Additionally, the two-dimensional histogram showed a high data density on the diagonal line, indicating that most of the data set is learned accurately. Moreover, no extreme outliers were visible.

Contrastingly, Coulomb couplings with static atomic transition charges were learned almost perfectly. This is in accordance with results for anthracene from the previous chapter.

9.3.2. Comparison of Obtained Observables

The trained NN models were applied to perform 1000 NAMD simulations of exciton transfer in each of the two rings of chromphores in LH2 independently. Single trajectories

with a length of 10 ps took a time of approximately 16 h and 23 h for B800 and B850 chromophores on single CPU cores, respectively.

Firstly, the degree of delocalization was investigated. The B850 chromophores exhibited an average coupling of 60 meV, which is in the order of the reorganization energy (65 meV). Therefore, transfer in this ring was expected to be highly delocalized and Marcus theory could not be applied. B800 chromophores on the other hand showed a small coupling of only 5 meV and the excitons were thus expected to be localized. These expectations were confirmed. Inverse participation ratios were calculated for both rings and are depicted in Figure 9.3. Excitons in the B800 ring were strongly localized and occupied 1.3 molecules on average. Contrastingly, excitons in the B850 ring were delocalized over approximately a third of the whole ring (4.7 molecules).

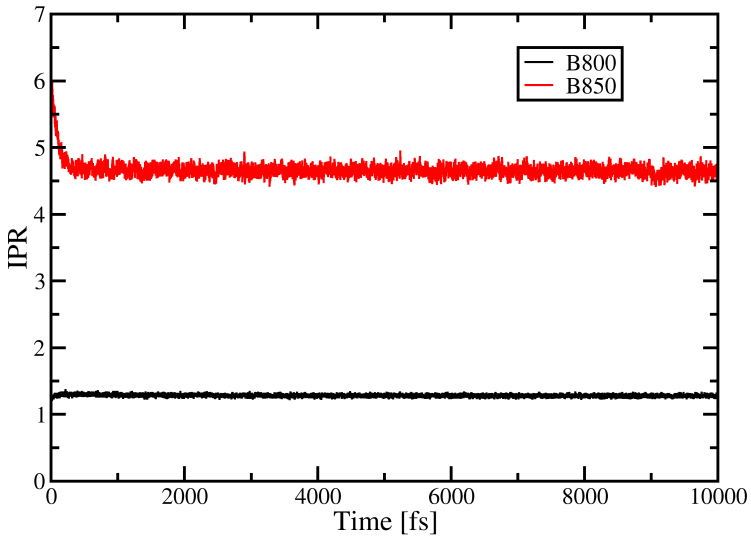


Figure 9.3.: Averaged inverse participation ratio versus simulation time for B800 and B850 chromophores.

For an estimate of the velocity of the transfer, diffusion constants can be calculated. In systems, where excitons are transferred along linear chains of molecules, this can be performed easily via the mean square displacement (MSD), which is a measure of how far an exciton traveled compared to its initial position. For an arrangement of molecules on a ring, a simple defined MSD would be meaningless. Hence, a modified definition of the MSD is desired, which must take into account the ring shape of the system, e.g. via transformation into a polar coordinate system.

Without consideration of the ring shape, diffusion constants of $82 \times 10^{-3}\,\mathrm{cm\,s^{-1}}$ (B800) and $227 \times 10^{-3}\,\mathrm{cm\,s^{-1}}$ (B850) were measured. Since the actually traveled distance on the ring path is greater than just the distance to the initial position, the diffusion constant is actually even greater in reality.

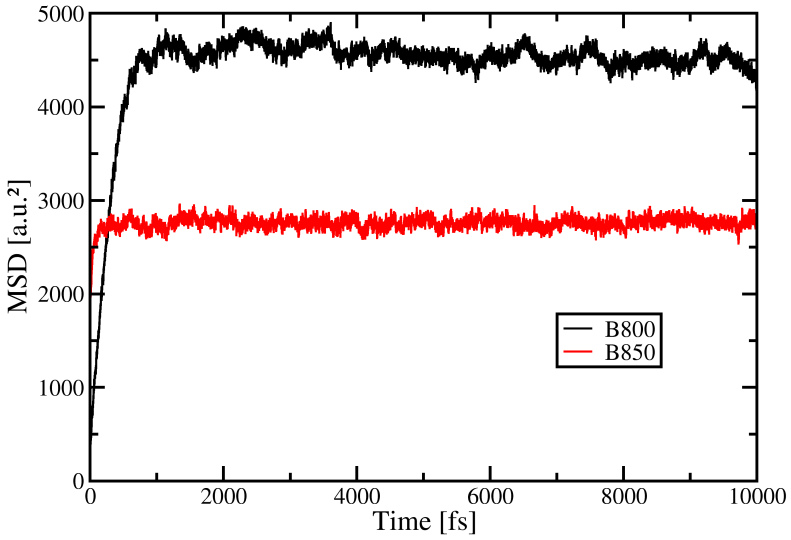


Figure 9.4.: Averaged mean square displacement versus simulation time for B800 and B850 chromophores.

From the plot of the standard MSD (Figure 9.4) it can be observed, that the MSD initially increased steadily until it was almost constant after a certain point. For all trajectories, the exciton was localized on the same molecule in the first step. From this point onward, the trajectories diverged and the exciton was transferred along the ring, either to the left or to the right side. The MSD increased with the distance to the initial position and the maximum was reached with the farthest molecule, which is the one opposing the initial molecule. Afterwards, the excitons in different trajectories are on average spread over the whole ring and the MSD stays constant. Assuming isotropic transfer due to the symmetry of the ring structure, the excitons should have traveled across one fourth of the ring on average. By measuring the diameter as the distance of two opposing molecules the traveled distance can be calculated as one fourth of the perimeter of each ring. This can be used for a rough estimate of the diffusion constant under consideration of the ring. The estimated diffusion constants were $140 \times 10^{-3}\,\mathrm{cm\,s^{-1}}$ (B800) and $435 \times 10^{-3}\,\mathrm{cm\,s^{-1}}$ (B850). These values were much larger compared to singlet exciton diffusion in organic semiconductors, which highlights the efficiency of biological light-harvesting complexes.[158]

For a visualization of the transfer in single trajectories, the diabatic occupation of the sites was plotted against the simulation time (Figure 9.5 and Figure 9.6). A transfer in discreet hops between neighboring molecules was observed for chromophores in the B800 ring, while the exciton was rapidly transferred in a delocalized state in the B850 ring.

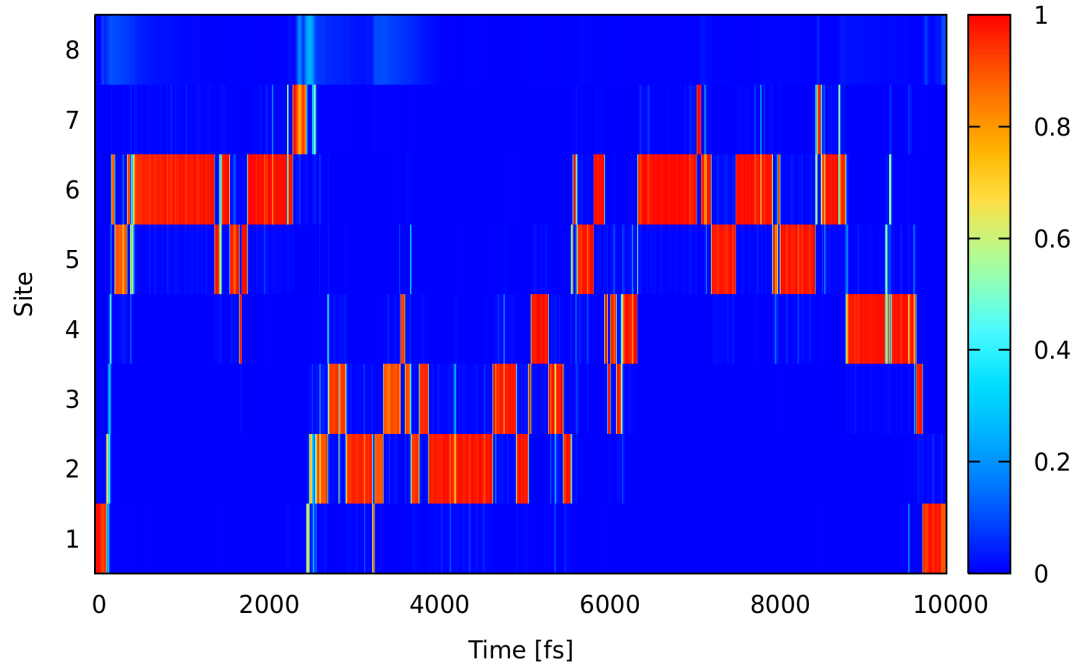


Figure 9.5.: Diabatic occupation of the sites in the B800 ring versus the simulation time for a single trajectory.

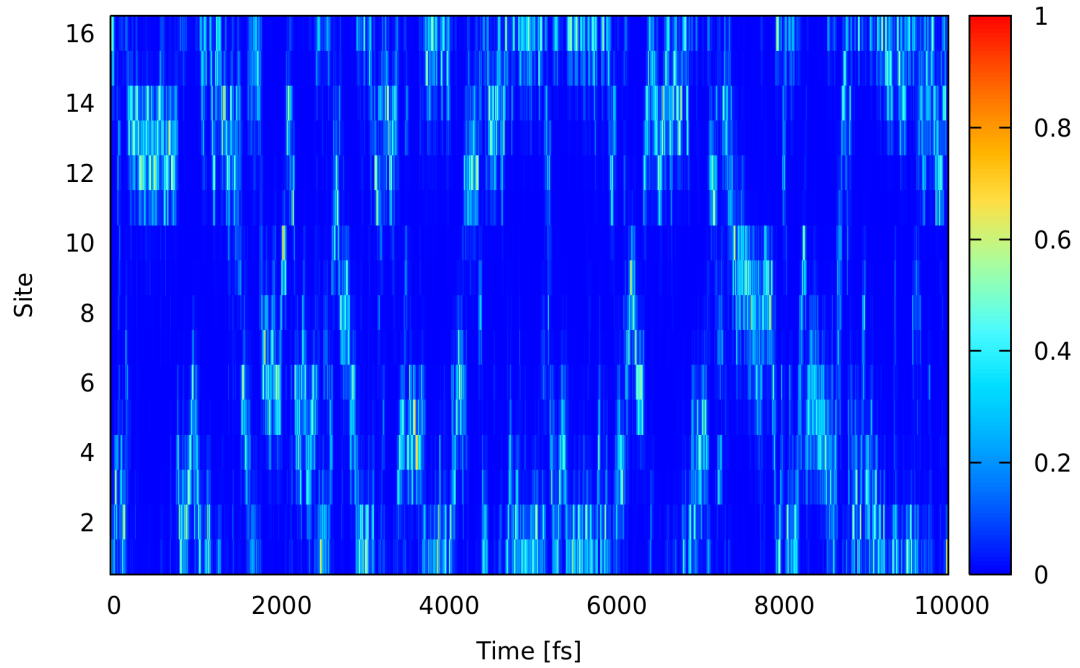


Figure 9.6.: Diabatic occupation of the sites in the B850 ring versus the simulation time for a single trajectory.

9.4. Conclusion and Outlook

In this chapter, non-adiabatic molecular dynamics simulations of exciton transfer in the LH2 complex were presented. With the aid of machine learning techniques, 1000 simulations were run for both rings of chromophores, with a length of 10 ps each, without the need to invest tremendous computational resources. This is important for sufficient sampling, when observables are to be obtained.

It was shown, that excitons within the B800 ring were strongly localized and transferred in discreet hops to neighboring molecules. In contrast, the exciton was delocalized over approximately 5 molecules in the B850 ring, where transfer was significantly faster and coherent. The diffusion constant as a measure for the velocity of the transfer was roughly estimated to be in the order of about $100 \times 10^{-3}\,\mathrm{cm\,s^{-1}}$ and $300 \times 10^{-3}\,\mathrm{cm\,s^{-1}}$ for B800 and B850, respectively. The transfer was much faster compared to singlet exciton transfer in organic semiconductors in general.

In future experiments, it will be the objective to include inter-ring couplings between B800 and B850 molecules, to also simulate the transfer between both rings. Additionally, diagonal and off-diagonal gradients can be learned together with the transfer Hamiltonian elements for the employment of more advanced propagator methods and the explicit treatment of exciton relaxation. Moreover, a strategy to account for the ring shape of the system, when calculating the MSD, is highly desirable to get more accurate results for diffusion constants.

So far, the site energies and couplings of the chromophores have been computed in vacuum without the consideration of the protein environment. Work ongoing already deals with the inclusion of these effects, which are known to be of importance to describe the differences of complexes from different organisms or mutants like the LH3 complex.[83,113,159]

This work represents a milestone in the development of efficient NAMD methods and enables the exploration of light-harvesting complexes through simulations in new levels of detail. Eventually, the goal is to obtain the full picture of exciton transfer from the initial absorption of the light to the final charge separation in the reaction center. To this end, transfer between multiple complexes, e.g. from LH2 to another LH2 and further to the LH1 complex, which is associated with the reaction center will be investigated.

Part IV.

Summary and Outlook

10. Summary and Outlook

This work aimed at the integration of machine learning (ML) techniques into the multi-scale workflow of charge and exciton transfer simulations.

In chapter 5, it was shown that simple and compact ML models can be utilized for the conduction of non-adiabatic molecular dynamics (NAMD) simulations of charge and exciton transfer in anthracene. Kernel ridge regression models were trained for the prediction of transfer Hamiltonian elements of anthracene and did not require much effort during the training process. The semiempirical DFTB reference method was reproduced closely in ML driven simulations with the mean-field Ehrenfest and surface hopping algorithms. However, the models did not accelerate the charge transfer simulations and training set sizes were highly limited, which may prohibit the training of more complex and larger molecules. Additionally, there was no feasible extension for the prediction of derivatives of the Hamiltonian elements, which are needed for the correct re-scaling of momenta and for explicit relaxation. Instead, approximations (implicit relaxation and Boltzmann-correction) had to be used, which introduce errors and may not be generally valid.

These issues were addressed in the next chapter (chapter 6). Neural network (NN) models provided a significant boost in efficiency compared to standard DFTB and an even larger speedup for higher levels of theory. Additionally, the favorable scaling of the models compared to QM methods amplified this advantage for larger molecules. Yet, the great performance of the NN models came with the downside of complex and difficult training procedures. The NN models were simultaneously trained on diagonal Hamiltonian derivatives in addition to the Hamiltonian elements, so the approximation of implicit relaxation could be dropped. This allowed for explicit relaxation without setting limits to the training set size. Simulations with the implicit relaxation scheme agreed well with the reference method, while for explicit relaxation effective reorganization energies were underestimated. This issue was traced back to an inconsistent use of methods for the calculation of site energies and gradients. However, an error contribution introduced by the ML models could not be excluded completely. This will be addressed in future work.

The next two chapters extended the methodology to exciton transfer. In chapter 7, the influence of short range effects via supermolecular couplings was investigated, which was indicated to be small in first tests. Nevertheless, it can be beneficial to calculate supermolecular couplings, as Coulomb couplings were shown to be overestimated. Furthermore, a diabatization scheme for more accurate and reliable computations of supermolecular couplings was implemented. In this course, a sign tracking procedure for all different coupling types was implemented as well.

The following chapter demonstrated the difficulties and importance of the generation of a correct training data set for exciton transfer in the case of diindenoperylene (chapter 8), although simulations with this compound are still pending. This chapter can also be seen as a continuation of chapter 6. Neural network models were expanded to the prediction of off-diagonal derivatives, which are necessary for the correct re-scaling of momenta with non-adiabatic coupling vectors, overcoming the Boltzmann-correction. The application to anthracene could reproduce experimental diffusion constants and showed highly localized transfer.

The final results chapter (chapter 9) combined the developments of this thesis and culminated in the application of on-the-fly NAMD simulations to exciton transfer in the light-harvesting complex II (LH2) from purple bacteria. This biological complex contains huge chromophores, which are arranged in two rings (B800, B850). To date, only one study was able to perform a single simulation of 300 fs length, which is insufficient for sampling. This was due to the tremendous computational cost of such simulations, which were lifted by the developed data-driven approach shown here. The transfer in both rings was simulated for 10 ps in 1000 trajectories each, with a feasible amount of resources. Excitons in the B800 ring were highly localized and transferred in discreet hops, while the B850 chromophores induced coherent transport and a spread of the exciton on multiple molecules. Estimated exciton diffusion constants for both rings appeared much higher compared to those of organic semiconducting materials.

The underestimation of explicit relaxation must be analyzed by retraining NN models on data, which was generated with a consistent employment of the same method for the calculation of site energies and gradients. Future work should focus on a thorough examination of the accuracy of the presented approach, e.g. by extending the work in chapter 8 to more organic semiconductor compounds in a benchmark study. The simulations of light-harvesting complexes should be refined by the use of more advanced couplings, explicit treatment of relaxation and the correct re-scaling scheme with non-adiabatic coupling vectors. Obtaining data for comparison with experimental measures, e.g. lifetimes and 2D electronic spectra, should validate the shown approach. A more accurate way of estimating the diffusion constant for such a nonlinear system should also be investigated. Moreover, the integration of further extensions for the method, e.g. the inclusion of electrostatic effects induced by the environment, are highly desired.

This work contributes to the methodological development of highly efficient methods for multi-scale simulations of charge and exciton transfer. Especially the latter enable the possibility for a refinement of the understanding of exciton transfer in terms of structural features, mechanisms and time scales. Now, large scale simulations, aiming at a comprehensive picture of the complete light-harvesting process in photosynthetic organisms from absorption to charge separation, are in reach.

Part V.

Appendix

A. Charge and Exciton Transfer Simulations in Organic Semiconductors driven by Machine Learned Hamiltonians

Detailed Propagation Results

Charge Transfer

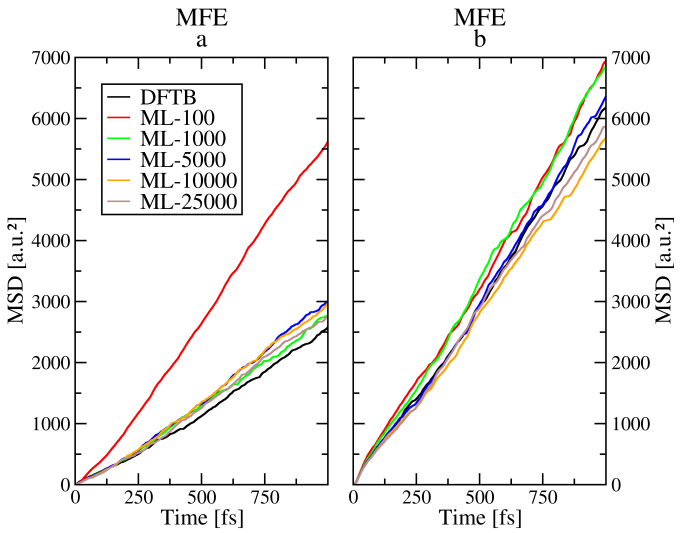


Figure A.1.: Time evolution of the averaged MSD in *a*- and *b*-direction using the MFE method for hole propagation with DFTB and ML-models with various training-size.

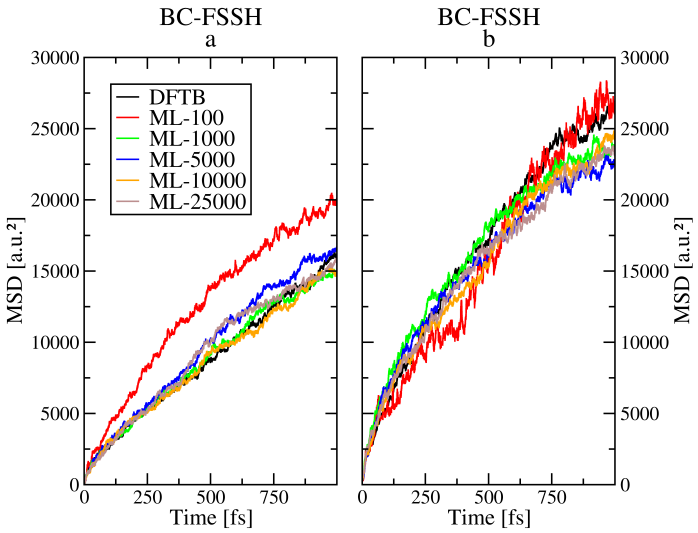


Figure A.2.: Time evolution of the averaged MSD in *a*- and *b*-direction using the BC-FSSH method for hole propagation with DFTB and ML-models with various training-size.

		Exp.	DFTB	ML-100	ML-1000	ML-5000	ML-10000	ML-25000
MFE	*a*	1.1	1.4	3.2	1.5	1.7	1.7	1.6
	b	2.9	3.4	3.7	3.7	3.5	3.1	3.2
BC-FSSH	*a*	1.1	8.0	10.1	7.7	8.7	7.6	8.1
	b	2.9	13.3	14.3	11.1	10.6	11.8	11.0

Table A.1.: Hole mobility in $cm^2\,V^{-1}\,s^{-1}$ as calculated from the averaged MSD in *a*- and *b*-direction using the MFE and BC-FSSH methods for charge propagation with DFTB and ML-models with various training-size.

Exciton Transfer

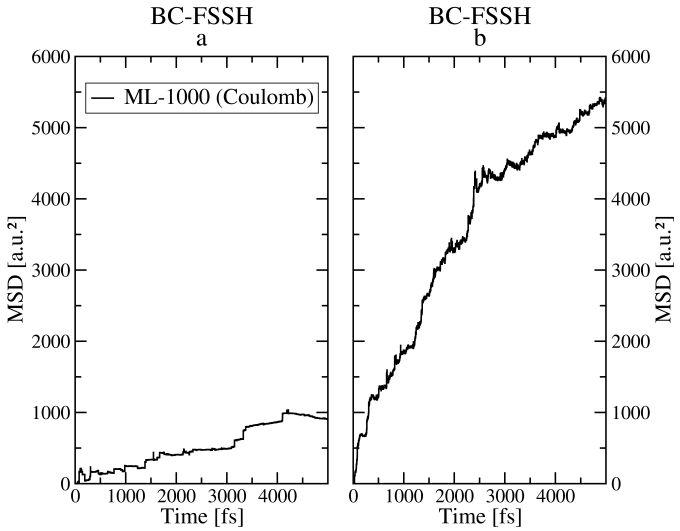


Figure A.3.: Time evolution of the averaged MSD in *a*- and *b*-direction using the BC-FSSH method for exciton propagation with a ML-model with a training-size of 1000.

B. Charge Transfer Simulations in Organic Semiconductors driven by Machine Learned Hamiltonians and Diagonal Forces

Training and Evaluation of Machine Learned Models

Models were trained using version 1.0.3 of the `pyNNsMD` code available from GitHub (`https://github.com/aimat-lab/NNsForMD`), using `TensorFlow` 2.4.1 [117] in Python 3.8.8 on NVidia GTX1080Ti cards. Visualizations were generated using Matplotlib version 3.3.4 [160]. All models used a batch size of 32 and were trained for at most 1000 epochs with an initial learning rate of 1×10^{-3}. The learning rate was reduced as training progressed to improve convergence: the models were trained for 20 steps at 1×10^{-3}, 100 steps at 1×10^{-4}, 400 steps at 1×10^{-5} and 480 steps at 1×10^{-6}.

Hyperparameter search was performed using the Hyperband algorithm as implemented in the `Keras Tuner` package (version 1.0.1) using a culling factor of 3. Optimizable hyperparameters and the optimal configurations for each model found in the hyperparameter search can be found in Table B.1 and Table B.2.

		NN depth	neurons/layer	weight reg.
	search space	3, 4 or 5	[20, 1000]	L1, L2 or None
ANT	site energies	4	70	None
	couplings	3	60	None
PEN	site energies	3	30	None
	couplings	3	290	L2

Table B.1.: Hyperparameter search space and parameters of best models obtained from hyperparameter search. These models were trained to predict only energies (site energies or couplings).

	search space	NN depth 3, 4 or 5	neurons/layer [20, 1000]	weight reg. L1, L2 or None	E loss weight [1, 5000]
ANT	FMO-DFTB1	3	80	L2	581
	Δ-LC-DFTB2	4	60	None	4421
	Δ-B3LYP	3	50	None	4981
	Δ-ωB97X	5	50	None	4731
PEN	FMO-DFTB	3	110	L2	4781
	Δ-LC-DFTB2	4	150	L2	4091

Table B.2.: Hyperparameter search space and parameters of best models obtained from hyperparameter search. These models were trained to predict site energies and their derivatives.

Reorganization Energies

It should be noted, that the gradients for the Δ-ωB97X model were generated with the ωB97X functional, while the reported value for λ^{114} was computed with the ωB97XD functional. We do not expect this to have a qualitative impact on the results.

Table B.3.: Reorganization energies (λ) as calcualted from QM and used for implicit relaxation. Effective reorganization energies (λ_{eff}) as estimated from NAMD simulations (DFTB or NN) employing explicit relaxation as well as QM-MD simulations with DFTB. All values in meV.

	Energies	Forces	ANT	PEN
λ	DFTB1	DFTB1	91	37^{114}
	LC-DFTB2	LC-DFTB2	138	91
	B3LYP	B3LYP	142^{114}	98^{114}
	ωB97XD	ωB97XD	202^{114}	170^{114}
λ_{eff}	FMO-DFTB1	FMO-DFTB1/MM	86	38
	NN (FMO-DFTB1)	NN (FMO-DFTB1)/MM	87	38
	NN (Δ-LC-DFTB2)	NN (Δ-LC-DFTB2)/MM	102	48
	NN (Δ-B3LYP)	NN (Δ-B3LYP)/MM	96	- -
	NN (Δ-ωB97X)	NN (Δ-ωB97X)/MM	130	- -
	DFTB1	DFTB1	94	- -
	LC-DFTB2	LC-DFTB2	109	- -
	DFTB1	DFTB1	93	- -
	LC-DFTB2	LC-DFTB2	132	- -
	DFTB1	DFTB1	91	- -
	DFTB1	LC-DFTB2	109	- -
	DFTB1	B3LYP	112	- -
	DFTB1	ωB97XD	167	- -

Detailed Propagation Results

Simulations Employing the IR Scheme

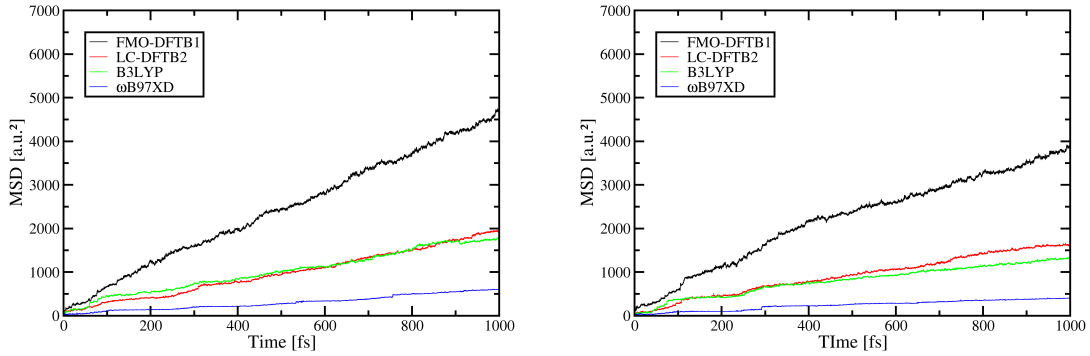


Figure B.1.: Mean suqare displacement (MSD) versus simulation time for anthracene *a*-direction. Simulations employed the implicit relaxation scheme with reorganization energies from various QM methods and the Hamiltonian elements were obtained from (a) FMO-DFTB1 or (b) NN models.

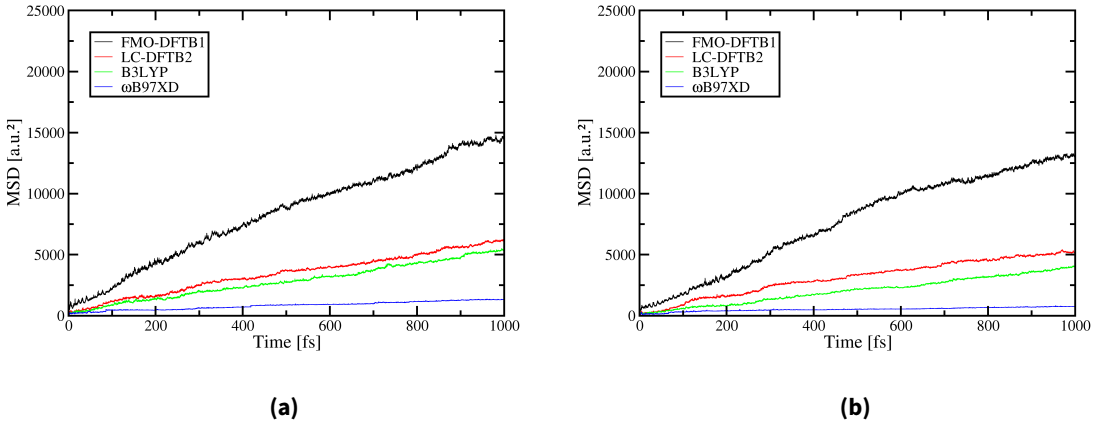


Figure B.2.: Mean suqare displacement (MSD) versus simulation time for anthracene *b*-direction. Simulations employed the implicit relaxation scheme with reorganization energies from various QM methods and the Hamiltonian elements were obtained from (a) FMO-DFTB1 or (b) NN models.

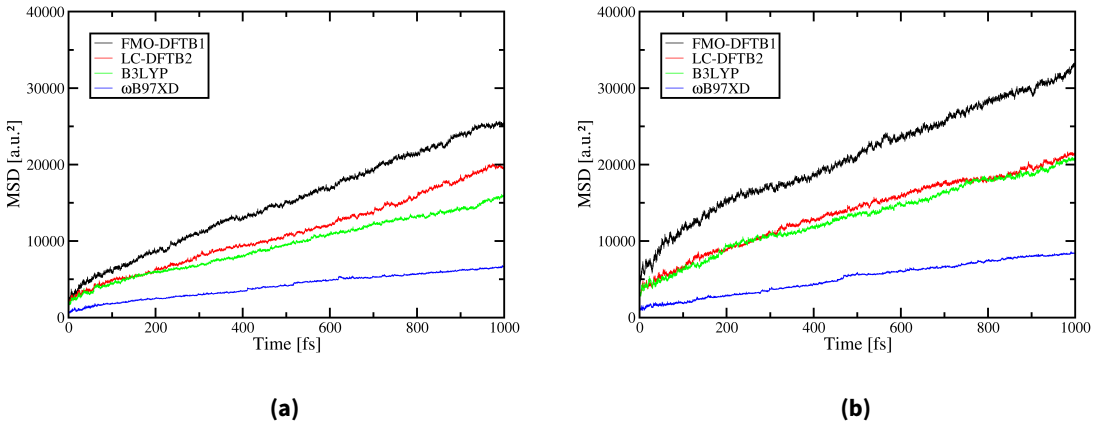


Figure B.3.: Mean suqare displacement (MSD) versus simulation time for pentacene *T1*-direction. Simulations employed the implicit relaxation scheme with reorganization energies from various QM methods and the Hamiltonian elements were obtained from (a) FMO-DFTB1 or (b) NN models.

Simulations Employing the ER Scheme

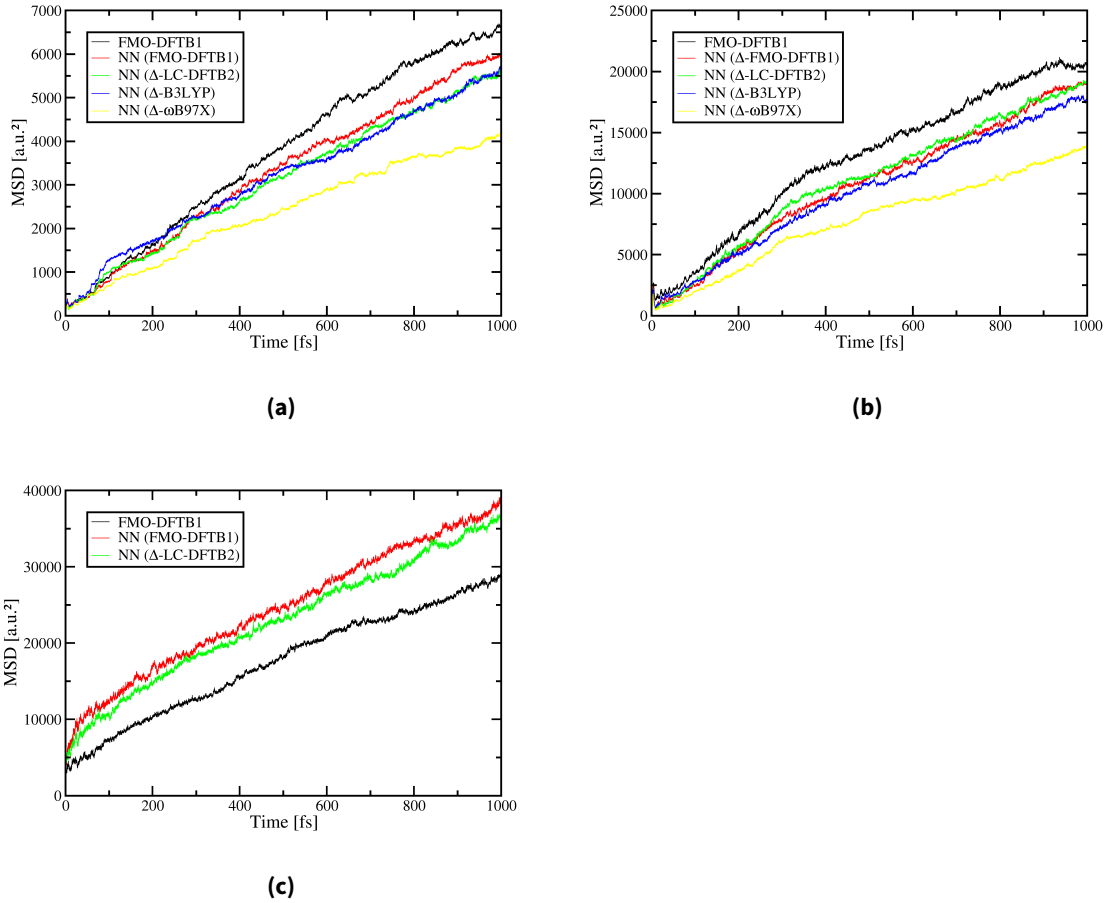


Figure B.4.: Mean suqare displacement (MSD) versus simulation time for anthracene in *a* (a) and *b*-direction (b) as well as petacene in *T1*-direction (c). Simulations employed the explicit relaxation scheme with forces from FMO-DFTB1 or NN models trained on various QM forces.

C. Calculation of Supermolecular Excitonic Couplings

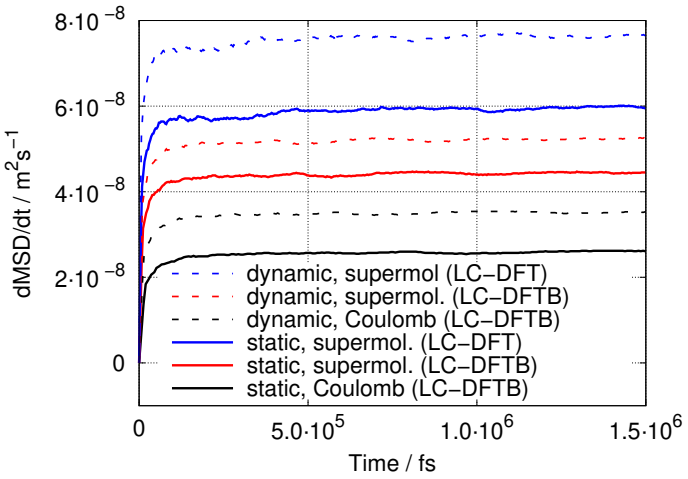


Figure C.1.: Time derivative of the exciton MSD from the MEQ approach for Coulomb and supermolecular excitonic couplings at TD-LC-DFTB2 and TD-LC-DFT level of theory, calculated on the crystal structure ("static", solid lines) and on the RMS of sampled structures ("dynamic", dashed lines).

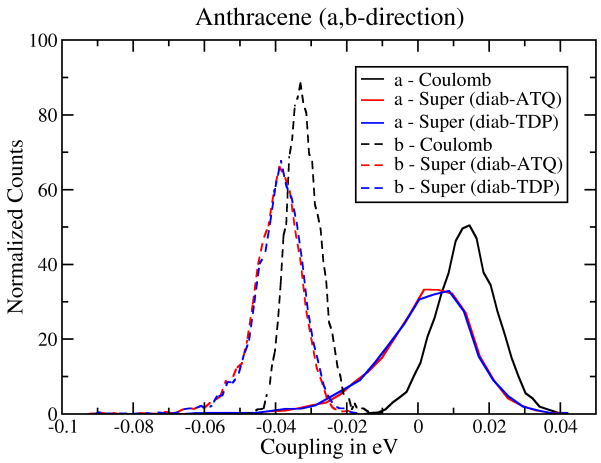


Figure C.2.: Histograms of excitonic couplings for sampled structures of an anthracene dimer in *a*- (solid lines) and *b*-direction (dashed lines). Coulomb and supermolecular couplings (diabatisation with ATQs/TDPs) are given.

D. Exciton Transfer Simulations in Organic Semiconductors driven by Machine-Learned Hamiltonians and Derivatives

Analysis and Preparation of the Data-Set

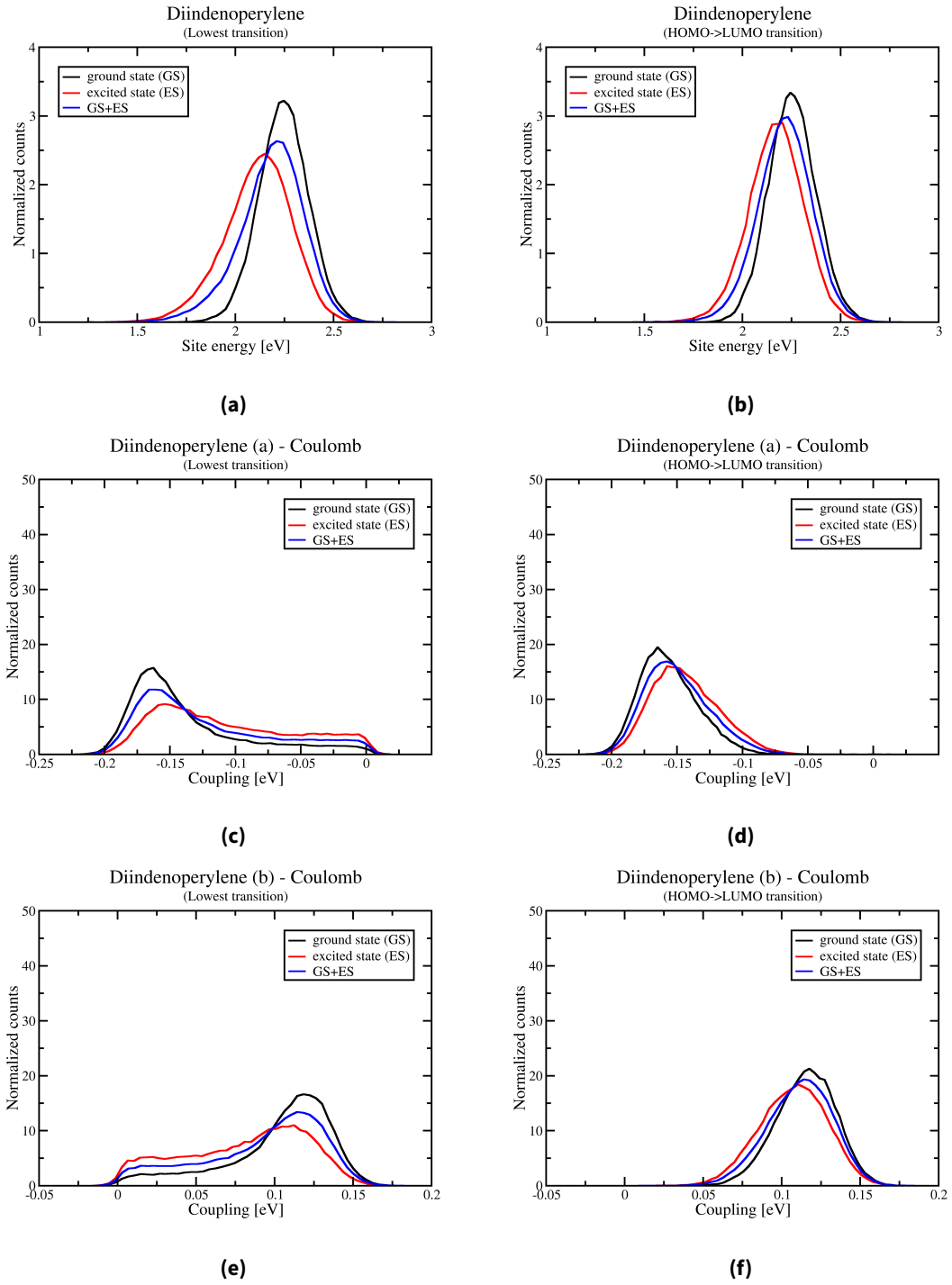


Figure D.1.: Site energies (top) and Coulomb couplings with fluctuating ATQs (middle and bottom) for diindeoperylene using the energetically lowest (left) versus the HOMO-LUMO transition (right), respectively.

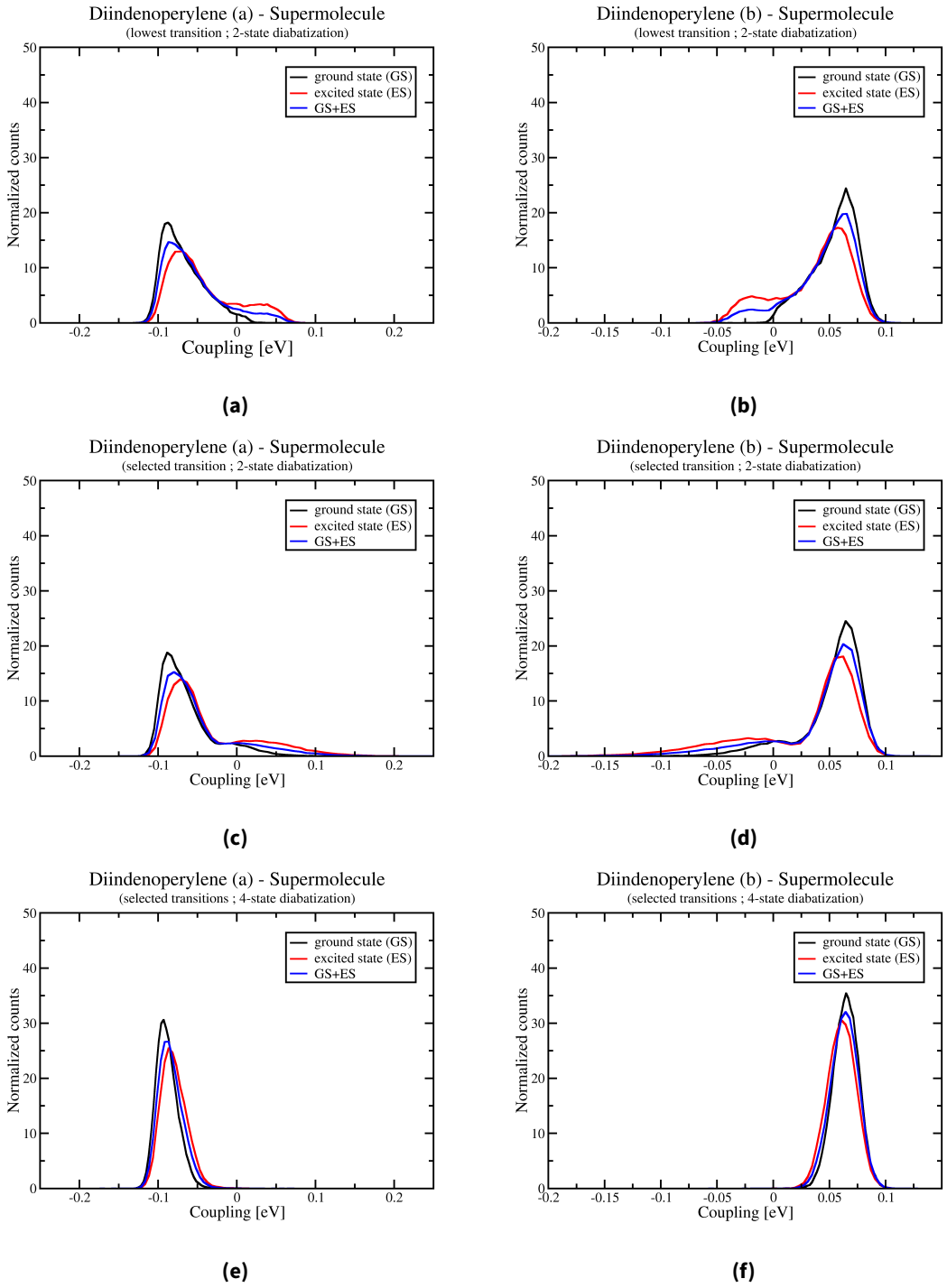
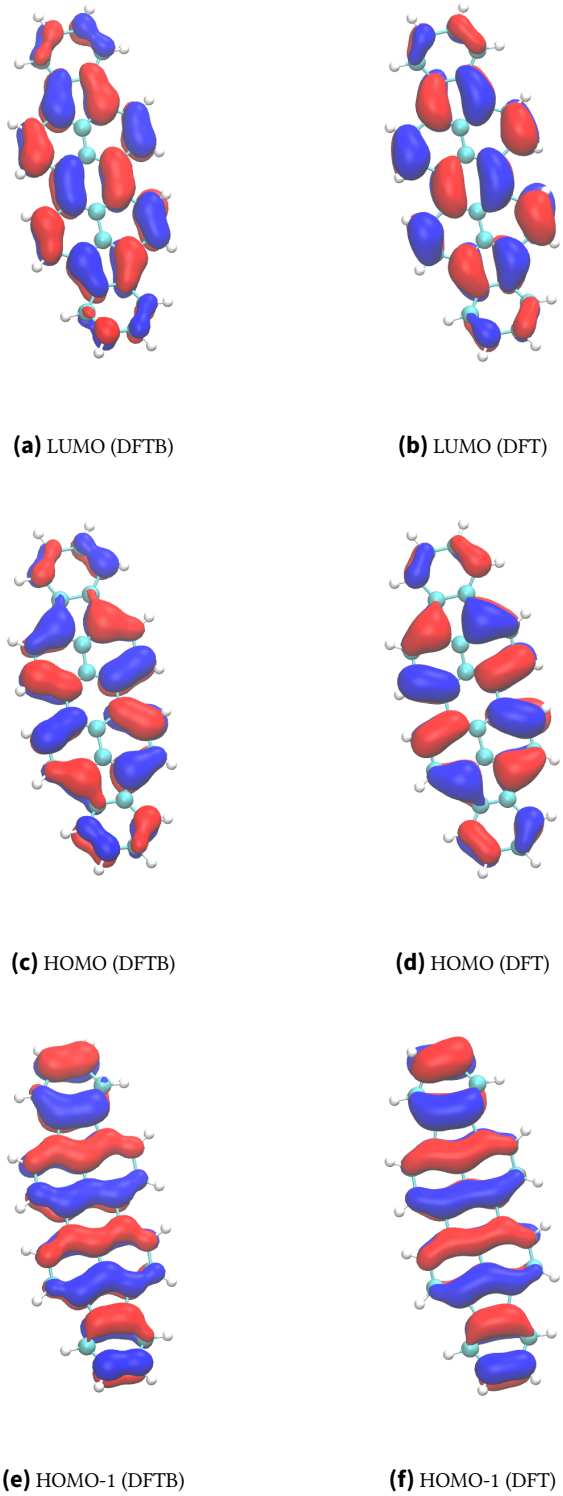


Figure D.2.: Supermolecular couplings for diindeoperylene using the energetically lowest versus the HOMO-LUMO transition in a 2- or 4-state diabatization, respectively.



(a) LUMO (DFTB)

(b) LUMO (DFT)

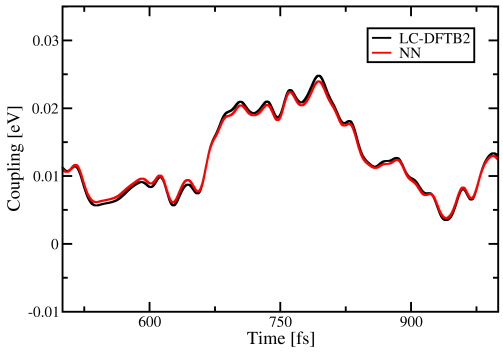
(c) HOMO (DFTB)

(d) HOMO (DFT)

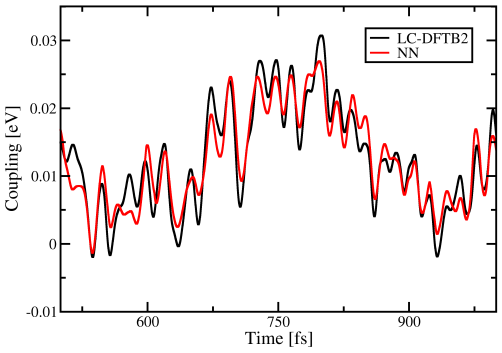
(e) HOMO-1 (DFTB)

(f) HOMO-1 (DFT)

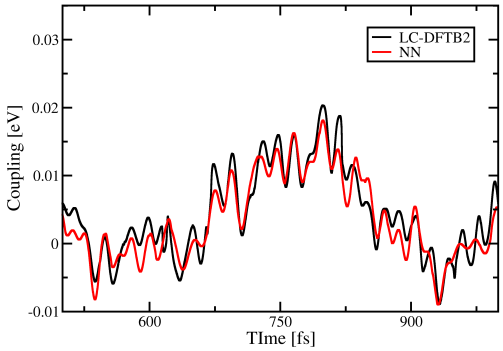
Figure D.3.: LUMO (top), HOMO (middle) and HOMO-1 (bottom) for diindeoperylene as calculated with DFTB (left) and DFT (right).



(a) Static Coulomb coupling



(b) Dynamic Coulomb coupling



(c) Supermolecular coupling

Figure D.4.: Time evolution of different coupling types for anthracene in *a*-direction. Values were calculated with the TD-LC-DFTB2 reference and NN models.

Model Training and Evaluation

parameter search space	depth 3, 4, 5	neur./layer [20, 1000]	reg. weight L1, L2	E loss weight [1, 1000]	∇E loss weight [1, 1000]
site energies	3	190	L1	51	911
stat. Coul. cpl.	3	90	L2	191	871
dyn. Coul cpl.	3	60	L2	- -	- -
supermol. cpl.	3	100	L2	- -	- -

Table D.1.: Hyperparameter search space and parameters of best models obtained from hyperparameter search.

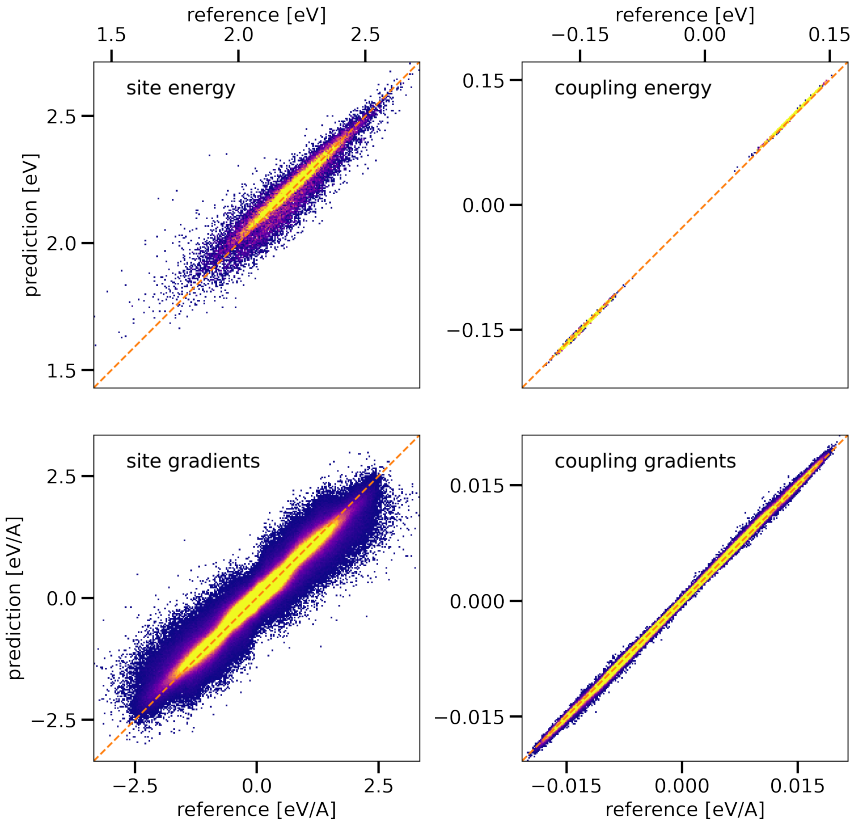


Figure D.5.: Two-dimensional histograms of site energies, static Coulomb couplings and their gradients: model predictions vs. LC-DFTB2 targets for 10 000 structures in diindenoperylene. Bright colors indicate high data density, unoccupied areas shown in white.

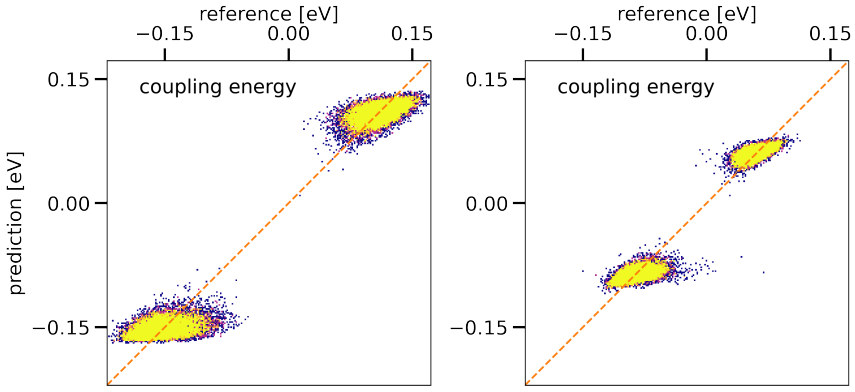


Figure D.6.: Two-dimensional histograms of dynamic Coulomb couplings and supermolecular couplings: model predictions vs. LC-DFTB2 targets for 10 000 structures in diindenoperylene. Bright colors indicate high data density, unoccupied areas shown in white.

Comparison of Obtained Observables

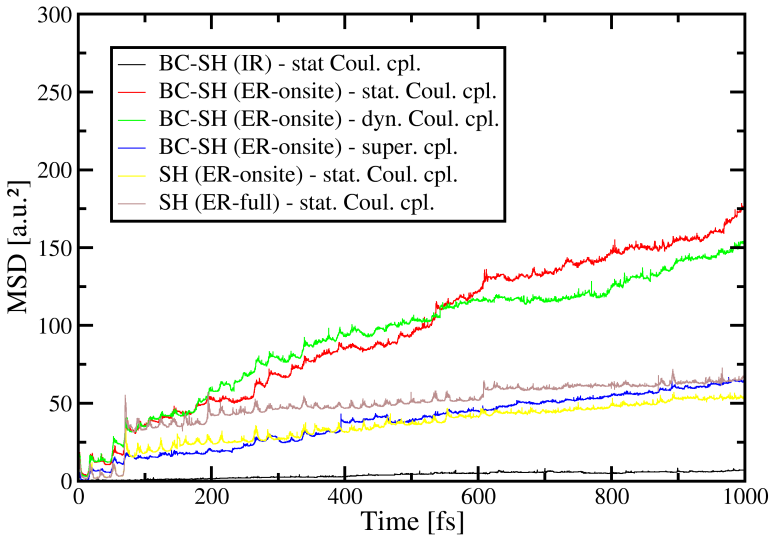


Figure D.7.: Time evolution of the averaged MSD in the *a*-direction of anthracene driven by NN models. Simulations employ the SH or BC-SH methods, respectively, with explicit relaxation (ER) and different variants of couplings: static and dynamic Coulomb couplings as well as supermolecular couplings.

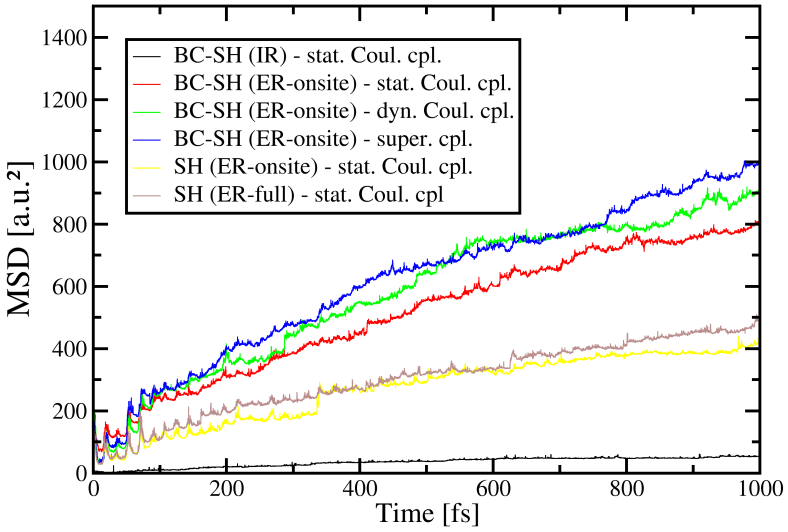


Figure D.8.: Time evolution of the averaged MSD in the *b*-direction of anthracene driven by NN models. Simulations employ the SH or BC-SH methods, respectively, with explicit relaxation (ER) and different variants of couplings: static and dynamic Coulomb couplings as well as supermolecular couplings.

E. Application: Exciton Transfer Simulations in the Light-Harvesting Complex II

Model Training and Evaluation

parameter search space	depth 3, 4, 5	neur./layer [20, 1000]	reg. weight L1, L2, None
site energies	3	320	None
couplings	5	50	None

Table E.1.: Hyperparameter search space and parameters of best models obtained from hyperparameter search. These models were trained to predict site energies and couplings.

List of Figures

List of Tables

Bibliography

[1] Christopher J. Cramer. *Essentials of Computational Chemistry: Theories and Models*. Wiley, 2004.

[2] W. Patrick Walters, Matthew T. Stahl, and Mark A. Murcko. "Virtual Screening – an Overview". In: *Drug Discov. Today* 3.4 (1998), pp. 160–178.

[3] Geoffroy Hautier. "Finding the Needle in the Haystack: Materials Discovery and Design Through Computational ab initio High-Throughput Screening". In: *Comp. Mater. Sci.* 163 (2019), pp. 108–116.

[4] Jerzy Leszczynski. *Handbook of Computational Chemistry*. Springer Science & Business Media, 2012.

[5] Harald Oberhofer, Karsten Reuter, and Jochen Blumberger. "Charge Transport in Molecular Materials: An Assessment of Computational Methods". In: *Chem. Rev.* 117.15 (2017), pp. 10319–10357.

[6] Michele Nottoli, Lorenzo Cupellini, Filippo Lipparini, Giovanni Granucci, and Benedetta Mennucci. "Multiscale Models for Light-Driven Processes". In: *Annu. Rev. Phys. Chem.* 72 (2021), pp. 489–513.

[7] O. Anatole Von Lilienfeld. "Quantum Machine Learning in Chemical Compound Space". In: *Angew. Chem. Int. Edit.* 57.16 (2018), pp. 4164–4169.

[8] Pavlo O. Dral. "Quantum Chemistry in the age of Machine Learning". In: *J. Physs Chem. Lett.* 11.6 (2020), pp. 2336–2347.

[9] Tomáš Kubař, Rafael Gutiérrez, Ulrich Kleinekathöfer, Gianaurelio Cuniberti, and Marcus Elstner. "Modeling Charge Transport in DNA Using Multi-Scale Methods". In: *Phys. Status Solidi B* 250.11 (2013), pp. 2277–2287.

[10] Alexander Heck, Julian J. Kranz, Tomas Kubar, and Marcus Elstner. "Multi-Scale Approach to Non-Adiabatic Charge Transport in High-Mobility Organic Semiconductors". In: *J. Chem. Theory Comput.* 11.11 (2015), pp. 5068–5082.

[11] Julian J. Kranz and Marcus Elstner. "Simulation of Singlet Exciton Diffusion in Bulk Organic Materials". In: *J. Chem. Theory Comput.* 12.9 (2016), pp. 4209–4221.

[12] Samuele Giannini, Antoine Carof, Matthew Ellis, Hui Yang, Orestis George Ziogos, Soumya Ghosh, and Jochen Blumberger. "Quantum Localization and Delocalization of Charge Carriers in Organic Semiconducting Crystals". In: *Nat. Commun.* 10.1 (2019), pp. 1–12.

[13] Frank Jensen. *Introduction to Computational Chemistry*. John Wiley & Sons, 2007.

[14] Max Born and Robert Oppenheimer. "Zur Quantentheorie der Molekeln". In: *Ann. Phys.* 389.20 (1927), pp. 457–484.

[15] Pierre Hohenberg and Walter Kohn. "Inhomogeneous Electron Gas". In: *Phys. Rev.* 136.3B (1964), B864.

[16] Walter Kohn and Lu Jeu Sham. "Self-Consistent Equations Including Exchange and Correlation Effects". In: *Phys. Rev.* 140.4A (1965), A1133.

[17] Wolfram Koch and Max C. Holthausen. *A Chemist's Guide to Density Functional Theory*. John Wiley & Sons, 2015.

[18] Takao Tsuneda and Kimihiko Hirao. "Long-Range Correction for Density Functional Theory". In: *WIREs Comput. Mol. Sci.* 4.4 (2014), pp. 375–390.

[19] Pekka Koskinen and Ville Mäkinen. "Density-Functional Tight-Binding for Beginners". In: *Comp. Mater. Sci.* 47.1 (2009), pp. 237–253.

[20] Dirk Porezag, Thomas Frauenheim, Thomas Köhler, Gotthard Seifert, and R. Kaschner. "Construction of Tight-Binding-Like Potentials on the Basis of Density-Functional Theory: Application to Carbon". In: *Phys. Rev. B* 51.19 (1995), p. 12947.

[21] Gotthard Seifert, Dirk Porezag, and Thomas Frauenheim. "Calculations of Molecules, Clusters, and Solids with a Simplified LCAO-DFT-LDA Scheme". In: *Int. J. Quantum Chem.* 58.2 (1996), pp. 185–192.

[22] Marcus Elstner, Dirk Porezag, G. Jungnickel, J. Elsner, M. Haugk, Th. Frauenheim, Sandor Suhai, and Gotthard Seifert. "Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties". In: *Phys. Rev. B* 58.11 (1998), p. 7260.

[23] Yang Yang, Haibo Yu, Darrin York, Qiang Cui, and Marcus Elstner. "Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method: Third-Order Expansion of the Density Functional Theory Total Energy and Introduction of a Modified Effective Coulomb Interaction". In: *J. Phys. Chem. A* 111.42 (2007), pp. 10861–10873.

[24] Marcus Elstner and Gotthard Seifert. "Density Functional Tight Binding". In: *Phil. Trans. R. Soc. A* 372.2011 (2014), p. 20120483.

[25] Thomas A. Niehaus and Fabio Della Sala. "Range Separated Functionals in the Density Functional Based Tight-Binding Method: Formalism". In: *Phys. Status Solidi B* 249.2 (2012), pp. 237–244.

[26] Vitalij Lutsker, Balint Aradi, and Thomas A. Niehaus. "Implementation and Benchmark of a Long-Range Corrected Functional in the Density Functional Based Tight-Binding Method". In: *TJ. Chem. Phys.* 143.18 (2015), p. 184107.

[27] Mark E. Casida. "Time-Dependent Density Functional Response Theory for Molecules". In: *Recent Advances In Density Functional Methods: (Part I)*. World Scientific, 1995, pp. 155–192.

[28] Thomas A. Niehaus, S. Suhai, F. Della Sala, P. Lugli, Marcus Elstner, Gotthard Seifert, and Thomas Frauenheim. "Tight-Binding Approach to Time-Dependent Density-Functional Response Theory". In: *Phys. Rev. B* 63.8 (2001), p. 085108.

[29] Julian J. Kranz, Marcus Elstner, Bálint Aradi, Thomas Frauenheim, Vitalij Lutsker, Adriel Dominguez Garcia, and Thomas A Niehaus. "Time-Dependent Extension of the Long-Range Corrected Density Functional Based Tight-Binding Method". In: *J. Chem. Theory Comput.* 13.4 (2017), pp. 1737–1747.

[30] Lung Wa Chung, W. M. C. Sameera, Romain Ramozzi, Alister J. Page, Miho Hatanaka, Galina P. Petrova, Travis V. Harris, Xin Li, Zhuofeng Ke, Fengyi Liu, et al. "The ONIOM Method and its Applications". In: *Chem. Rev.* 115.12 (2015), pp. 5678–5796.

[31] Nikos L. Doltsinis. "Nonadiabatic Dynamics: Mean-Field and Surface Hopping". In: *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*. NIC, 2002, pp. 377–397.

[32] Tomás Kubar, P. Benjamin Woiczikowski, Gianaurelio Cuniberti, and Marcus Elstner. "Efficient Calculation of Charge-Transfer Matrix Elements for Hole Transfer in DNA". In: *J. Phys. Chem. B* 112.26 (2008), pp. 7937–7947.

[33] Paul Ehrenfest. "Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik". In: *Z. Phys.* 45.7 (1927), pp. 455–457.

[34] John C. Tully and Richard K. Preston. "Trajectory Surface Hopping Approach to Nonadiabatic Molecular Collisions: The Reaction of H+ with D2". In: *J. Chem. Phys.* 55.2 (1971), pp. 562–572.

[35] John C. Tully. "Molecular Dynamics with Electronic Transitions". In: *J. Chem. Phys.* 93.2 (1990), pp. 1061–1071.

[36] Weiwei Xie, Daniel Holub, Tomáš Kubař, and Marcus Elstner. "Performance of Mixed Quantum-Classical Approaches on Modeling the Crossover from Hopping to Bandlike Charge Transport in Organic Semiconductors". In: *J. Chem. Theory Comput.* 16.4 (2020), pp. 2071–2084.

[37] Tomas Kubar and Marcus Elstner. "Coarse-Grained Time-Dependent Density Functional Simulation of Charge Transfer in Complex Systems: Application to Hole Transfer in DNA". In: *J. Phys. Chem. B* 114.34 (2010), pp. 11221–11240.

[38] Antoine Carof, Samuele Giannini, and Jochen Blumberger. "Detailed Balance, Internal Consistency, and Energy Conservation in Fragment Orbital-Based Surface Hopping". In: *J. Chem. Phys.* 147.21 (2017), p. 214113.

[39] Alexey V. Akimov and Oleg V. Prezhdo. "The PYXAID Program for Non-Adiabatic Molecular Dynamics in Condensed Matter Systems". In: *J. Chem. Theory Comput.* 9.11 (2013), pp. 4959–4972.

[40] Kazuo Kitaura, Eiji Ikeo, Toshio Asada, Tatsuya Nakano, and Masami Uebayasi. "Fragment Molecular Orbital Method: An Approximate Computational Method for Large Molecules". In: *Chem. Phys. Lett.* 313.3-4 (1999), pp. 701–706.

[41] Volkhard May and Oliver Kühn. *Charge and Energy Transfer Dynamics in Molecular Systems*. John Wiley & Sons, 2008.

[42] Jacov Frenkel. "On the Transformation of Light into Heat in Solids. I". In: *Phys. Rev.* 37.1 (1931), p. 17.

[43] Jacov Frenkel. “On the Transformation of Light into Heat in Solids. II”. In: *Phys. Rev.* 37.10 (1931), p. 1276.

[44] Carles Curutchet and Benedetta Mennucci. “Quantum Chemical Studies of Light Harvesting”. In: *Chem. Rev.* 117.2 (2017), pp. 294–343.

[45] Elise P. Kenny and Ivan Kassal. “Benchmarking Calculations of Excitonic Couplings Between Bacteriochlorophylls”. In: *J. Phys. Chem. B* 120.1 (2016), pp. 25–32.

[46] Thomas Renger. “Theory of Excitation Energy Transfer: From Structure to Function”. In: *Photosynth. Res.* 102.2 (2009), pp. 471–485.

[47] M. E. Madjet, A. Abdurahman, and Thomas Renger. “Intermolecular Coulomb couplings from ab initio Electrostatic Potentials: Application to Optical Transitions of Strongly Coupled Pigments in Photosynthetic Antennae and Reaction Centers”. In: *J. Phys. Chem. B* 110.34 (2006), pp. 17268–17281.

[48] Brent P. Krueger, Gregory D. Scholes, and Graham R. Fleming. “Calculation of Couplings and Energy-Transfer Pathways Between the Pigments of LH2 by the ab initio Transition Density Cube Method”. In: *J. Phys. Chem. B* 102.27 (1998), pp. 5378–5386.

[49] Per-Arno Plötz, Thomas Niehaus, and Oliver Kühn. “A New Efficient Method for Calculation of Frenkel Exciton Parameters in Molecular Aggregates”. In: *J. Chem. Phys.* 140.17 (2014), p. 174101.

[50] Juan Aragó and Alessandro Troisi. “Dynamics of the Excitonic Coupling in Organic Crystals”. In: *Phys. Rev. Lett.* 114.2 (2015), p. 026402.

[51] Peter H. Schönemann. “A Generalized Solution of the Orthogonal Procrustes Problem”. In: *Psychometrika* 31.1 (1966), pp. 1–10.

[52] John C. Gower and Garmt B. Dijksterhuis. *Procrustes Problems*. Vol. 30. OUP Oxford, 2004.

[53] Rocco P. Fornari, Juan Aragó, and Alessandro Troisi. “Exciton Dynamics in Phthalocyanine Molecular Crystals”. In: *J. Phys. Chem. C* 120.15 (2016), pp. 7987–7996.

[54] Juan Aragó and Alessandro Troisi. “Excitonic Couplings Between Molecular Crystal Pairs by a Multistate Approximation”. In: *J. Chem. Phys.* 142.16 (2015), p. 164107.

[55] Daan Frenkel and Berend Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Vol. 1. Elsevier, 2001.

[56] Ryogo Kubo, Morikazu Toda, and Natsuki Hashitsume. *Statistical Physics II: Nonequilibrium Statistical Mechanics*. Vol. 31. Springer Science & Business Media, 2012.

[57] Alexander Jung. *Machine Learning: The Basics*. Springer Nature, 2022.

[58] Trevor Hastie, Robert Tibshirani, Jerome H. Friedman, and Jerome H. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Vol. 2. Springer, 2009.

[59] Christopher M. Bishop and Nasser M. Nasrabadi. *Pattern Recognition and Machine Learning*. Vol. 4. 4. Springer, 2006.

[60] Anthony Yu-Tung Wang, Ryan J. Murdock, Steven K. Kauwe, Anton O. Oliynyk, Aleksander Gurlo, Jakoah Brgoch, Kristin A. Persson, and Taylor D. Sparks. "Machine Learning for Materials Scientists: An Introductory Guide Toward Best Practices". In: *Chem. Mat.* 32.12 (2020), pp. 4954–4965.

[61] Kristof T. Schütt, Stefan Chmiela, O. Anatole von Lilienfeld, Alexandre Tkatchenko, Koji Tsuda, and Klaus-Robert Müller. *Machine Learning Meets Quantum Physics*. Springer, 2020.

[62] Michael A. Nielsen. *Neural Networks and Deep Learning*. Vol. 25. Determination press San Francisco, CA, USA, 2015.

[63] Albert P. Bartók, Risi Kondor, and Gábor Csányi. "On Representing Chemical Environments". In: *Phys. Rev. B* 87.18 (2013), p. 184115.

[64] O. Anatole Von Lilienfeld, Raghunathan Ramakrishnan, Matthias Rupp, and Aaron Knoll. "Fourier Series of Atomic Radial Distribution Functions: A Molecular Fingerprint for Machine Learning Models of Quantum Chemical Properties". In: *Int. J. Quantum Chem.* 115.16 (2015), pp. 1084–1093.

[65] Matthias Rupp, Alexandre Tkatchenko, Klaus-Robert Müller, and O. Anatole Von Lilienfeld. "Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning". In: *Phys. Rev. Lett.* 108.5 (2012), p. 058301.

[66] Jörg Behler and Michele Parrinello. "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces". In: *Phys. Rev. Lett.* 98.14 (2007), p. 146401.

[67] Anders S. Christensen, Lars A. Bratholm, Felix A. Faber, and O. Anatole von Lilienfeld. "FCHL Revisited: Faster and More Accurate Quantum Machine Learning". In: *J. Chem. Phys.* 152.4 (2020), p. 044107.

[68] Matthias Rupp. "Machine Learning for Quantum Mechanics in a Nutshell". In: *Int. J. Quantum Chem.* 115.16 (2015), pp. 1058–1073.

[69] Jörg Behler. "Constructing High-Dimensional Neural Network Potentials: A Tutorial Review". In: *Int. J. Quantum Chem.* 115.16 (2015), pp. 1032–1050.

[70] Neal R. Armstrong, Weining Wang, Dana M. Alloway, Diogenes Placencia, Erin Ratcliff, and Michael Brumbach. "Organic/Organic Heterojunctions: Organic Light Emitting Diodes and Organic Photovoltaic Devices". In: *Macromol. Rapid Commun.* 30.9-10 (2009), pp. 717–731.

[71] Rico Meerheim, BjÖrn Lussem, and Karl Leo. "Efficiency and Stability of Pin Type Organic Light Emitting Diodes for Display and Lighting Applications". In: *Proc. IEEE* 97.9 (2009), pp. 1606–1626.

[72] Abhishek P. Kulkarni, Christopher J. Tonzola, Amit Babel, and Samson A. Jenekhe. "Electron Transport Materials for Organic Light-Emitting Diodes". In: *Chem. Mater.* 16.23 (2004), pp. 4556–4573.

[73] Christos D. Dimitrakopoulos and Patrick R. L. Malenfant. "Organic Thin Film Transistors for Large Area Electronics". In: *Adv. Mater.* 14.2 (2002), pp. 99–117.

[74] Christopher R. Newman, C. Daniel Frisbie, Demetrio A. da Silva Filho, Jean-Luc Brédas, Paul C. Ewbank, and Kent R. Mann. "Introduction to Organic Thin Film Transistors and Design of N-Channel Organic Semiconductors". In: *Chem. Mater.* 16.23 (2004), pp. 4436–4451.

[75] Antonio Facchetti. "Semiconductors for Organic Transistors". In: *Mater. Today* 10.3 (2007), pp. 28–37.

[76] Bernard Kippelen and Jean-Luc Brédas. "Organic Photovoltaics". In: *Energy Environ. Sci.* 2.3 (2009), pp. 251–261.

[77] Gilles Dennler, Markus C. Scharber, and Christoph J. Brabec. "Polymer–Fullerene Bulk-Heterojunction Solar Cells". In: *Adv. Mater.* 21.13 (2009), pp. 1323–1338.

[78] Carsten Deibel and Vladimir Dyakonov. "Polymer–Fullerene Bulk Heterojunction Solar Cells". In: *Rep. Prog. Phys.* 73.9 (2010), p. 096401.

[79] Anamika Dey, Ashish Singh, Dipjyoti Das, and Parameswar Krishnan Iyer. "Organic Semiconductors: A New Future of Nanodevices and Applications". In: *Thin Film Structures in Energy Applications*. Springer, 2015, pp. 97–128.

[80] Jonas Lederer, Waldemar Kaiser, Alessandro Mattoni, and Alessio Gagliardi. "Machine Learning–Based Charge Transport Computation for Pentacene". In: *Adv. Theory and Simul.* 2.2 (2019), p. 1800136.

[81] Félix Musil, Sandip De, Jack Yang, Joshua E. Campbell, Graeme M. Day, and Michele Ceriotti. "Machine Learning for the Structure–Energy–Property Landscapes of Molecular Crystals". In: *Chem. Sci.* 9.5 (2018), pp. 1289–1300.

[82] Onur Caylak, Anil Yaman, and Björn Baumeier. "Evolutionary Approach to Constructing a Deep Feedforward Neural Network for Prediction of Electronic Coupling Elements in Molecular Materials". In: *J. Chem. Theory Comput.* 15.3 (2019), pp. 1777–1784.

[83] Florian Häse, Stéphanie Valleau, Edward Pyzer-Knapp, and Alán Aspuru-Guzik. "Machine Learning Exciton Dynamics". In: *Chem. Sci.* 7.8 (2016), pp. 5139–5147.

[84] Jingbai Li, Patrick Reiser, Benjamin R Boswell, André Eberhard, Noah Z. Burns, Pascal Friederich, and Steven A. Lopez. "Automatic Discovery of Photoisomerization Mechanisms with Nanosecond Machine Learning Photodynamics Simulations". In: *Chem. Sci.* 12.14 (2021), pp. 5302–5314.

[85] Julia Westermayr, Michael Gastegger, Maximilian F. S. J. Menger, Sebastian Mai, Leticia González, and Philipp Marquetand. "Machine Learning Enables Long Time Scale Molecular Photodynamics Simulations". In: *Chem. Sci.* 10.35 (2019), pp. 8100–8107.

[86] Julia Westermayr and Philipp Marquetand. "Machine Learning and Excited-State Molecular Dynamics". In: *Mach. Learn.: Sci. Technol.* 1.4 (2020), p. 043001.

[87] Julia Westermayr, Michael Gastegger, and Philipp Marquetand. "Combining SchNet and SHARC: The SchNarc Machine Learning Approach for Excited-State Dynamics". In: *J. Phys. Chem. Lett.* 11.10 (2020), pp. 3828–3834.

[88] Chun-I Wang, Mac Kevin E. Braza, Gil C. Claudio, Ricky B. Nellas, and Chao-Ping Hsu. "Machine Learning for Predicting Electron Transfer Coupling". In: *J. Phys. Chem. A* 123.36 (2019), pp. 7792–7802.

[89] Balint Aradi, Ben Hourahine, and Thomas Frauenheim. "DFTB+, a Sparse Matrix-Based Implementation of the DFTB Method". In: *J. Phys. Chem. A* 111.26 (2007), pp. 5678–5684.

[90] Adam Kubas, Felix Hoffmann, Alexander Heck, Harald Oberhofer, Marcus Elstner, and Jochen Blumberger. "Electronic Couplings for Molecular Charge Transfer: Benchmarking CDFT, FODFT, and FODFTB Against High-Level ab initio Calculations". In: *J. Chem. Phys.* 140.10 (2014), p. 104105.

[91] Adam Kubas, Fruzsina Gajdos, Alexander Heck, Harald Oberhofer, Marcus Elstner, and Jochen Blumberger. "Electronic Couplings for Molecular Charge Transfer: Benchmarking CDFT, FODFT and FODFTB Against High-Level ab initio Calculations. II". In: *Phys. Chem. Chem. Phys.* 17.22 (2015), pp. 14342–14354.

[92] R. Mason. "The Crystallography of Anthracene at 95 K and 290 K". In: *Acta Crystallogr.* 17.5 (1964), pp. 547–555.

[93] Junmei Wang, Romain M. Wolf, James W. Caldwell, Peter A. Kollman, and David A. Case. "Development and Testing of a General Amber Force Field". In: *J. Comput. Chem.* 25.9 (2004), pp. 1157–1174.

[94] Junmei Wang, Wei Wang, Peter A. Kollman, and David A. Case. "Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations". In: *J. Mol. Graphics and Modell.* 25.2 (2006), pp. 247–260.

[95] U. Chandra Singh and Peter A. Kollman. "An Approach to Computing Electrostatic Charges for Molecules". In: *J. Comput. Chem.* 5.2 (1984), pp. 129–145.

[96] Brent H. Besler, Kenneth M. Merz Jr., and Peter A. Kollman. "Atomic Charges Derived from Semiempirical Methods". In: *J. Comput. Chem.* 11.4 (1990), pp. 431–439.

[97] G. A. Petersson, Andrew Bennett, Thomas G. Tensfeldt, Mohammad A. Al-Laham, William A. Shirley, and John Mantzaris. "A Complete Basis Set Model Chemistry. I. The Total Energies of Closed-Shell Atoms and Hydrides of the First-Row Elements". In: *J. Chem. Phys.* 89.4 (1988), pp. 2193–2218.

[98] G. A. Petersson and Mohammad A. Al-Laham. "A Complete Basis Set Model Chemistry. II. Open-Shell Systems and the Total Energies of the First-Row Atoms". In: *J. Chem. Phys.* 94.9 (1991), pp. 6081–6090.

[99] Michael J. Frisch, G. W. Trucks, H. Bernhard Schlegel, Gustavo E. Scuseria, Michael A. Robb, James R. Cheeseman, Giovanni Scalmani, Vincenzo Barone, Benedetta Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, Xiaosong Li, H. P. Hratchian, Artur F. Izmaylov, Julien Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, François Ogliaro, Michael J. Bearpark, Jochen Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, Rika Kobayashi, J.

Normand, Krishnan Raghavachari, Alistair P. Rendell, J. C. Burant, S. S. Iyengar, Jacopo Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, A Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and Douglas J. Fox. *Gaussian 09*. Wallingford, CT, USA: Gaussian, Inc., 2009.

[100] Herman J. C. Berendsen, David van der Spoel, and Rudi van Drunen. "GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation". In: *Comput. Phys. Commun.* 91.1-3 (1995), pp. 43–56.

[101] Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C. Smith, Berk Hess, and Erik Lindahl. "GROMACS: High Performance Molecular Simulations Through Multi-Level Parallelism from Laptops to Supercomputers". In: *SoftwareX* 1 (2015), pp. 19–25.

[102] Denis J. Evans and Brad Lee Holian. "The Nose–Hoover Thermostat". In: *J. Chem. Phys.* 83.8 (1985), pp. 4069–4074.

[103] Simone Fratini, S. Ciuchi, D. Mayou, G. Trambly De Laissardière, and A. Troisi. "A Map of High-Mobility Molecular Semiconductors". In: *Nat. Mater.* 16.10 (2017), pp. 998–1002.

[104] Anders S. Christensen, Lars A. Bratholm, Silvia Amabilino, Jimmy C. Kromann, Felix A. Faber, Bing Huang, A. Tkatchenko, and O. A. von Lilienfeld. *QML: A Python Toolkit for Quantum Machine Learning, 2019*.

[105] Mila Krämer, Philipp M. Dohmen, Weiwei Xie, Daniel Holub, Anders S. Christensen, and Marcus Elstner. "Charge and Exciton Transfer Simulations Using Machine-Learned Hamiltonians". In: *J. Chem. Theory Comput.* 16.7 (2020), pp. 4061–4070.

[106] Norbert Karl and Jörg Marktanner. "Electron and Hole Mobilities in High Purity Anthracene Single Crystals". In: *Mol. Cryst. Liq. Cryst. Sci. Technol., Sect. A.* 355.1 (2001), pp. 149–173.

[107] A. D. McLachlan. "A Variational Solution of the Time-Dependent Schrodinger Equation". In: *Mol. Phys.* 8.1 (1964), pp. 39–44.

[108] Weiwei Xie, Marin Sapunar, Nađa Došlić, Matthieu Sala, and Wolfgang Domcke. "Assessing the Performance of Trajectory Surface Hopping Methods: Ultrafast Internal Conversion in Pyrazine". In: *J. Chem. Phys.* 150.15 (2019), p. 154119.

[109] Szilárd Páll, Mark James Abraham, Carsten Kutzner, Berk Hess, and Erik Lindahl. "Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS". In: *International conference on exascale applications and software*. Springer. 2014, pp. 3–27.

[110] Vera Stehr, Reinhold F Fink, Bernd Engels, Jens Pflaum, and Carsten Deibel. "Singlet Exciton Diffusion in Organic Crystals Based on Marcus Transfer Rates". In: *J. Chem. Theory Comput.* 10.3 (2014), pp. 1242–1255.

[111] Rudolph A. Marcus. "On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. I". In: *J. Chem. Phys.* 24.5 (1956), pp. 966–978.

[112] Rudolph A. Marcus. "Electron Transfer Reactions in Chemistry. Theory and Experiment". In: *Rev. Mod. Phys.* 65.3 (1993), p. 599.

[113] Beatrix M. Bold, Monja Sokolov, Sayan Maity, Marius Wanko, Philipp M. Dohmen, Julian J. Kranz, Ulrich Kleinekathöfer, Sebastian Höfener, and Marcus Elstner. "Benchmark and Performance of Long-Range Corrected Time-Dependent Density Functional Tight Binding (LC-TD-DFTB) on Rhodopsins and Light-Harvesting Complexes". In: *Phys. Chem. Chem. Phys.* 22.19 (2020), pp. 10500–10518.

[114] Sara Roosta, Farhad Ghalami, Marcus Elstner, and Weiwei Xie. "Efficient Surface Hopping Approach for Modeling Charge Transport in Organic Semiconductors". In: *J. Chem. Theory Comput.* 18.3 (2022), pp. 1264–1274.

[115] Ben Hourahine, Bálint Aradi, Volker Blum, F. Bonafé, A. Buccheri, Cristopher Camacho, Caterina Cevallos, M. Y. Deshaye, T. Dumitrică, A. Dominguez, et al. "DFTB+, a Software Package for Efficient Approximate Density Functional Theory Based Atomistic Simulations". In: *J. Chem. Phys.* 152.12 (2020), p. 124101.

[116] Anders S. Christensen, Felix A. Faber, and O. Anatole von Lilienfeld. "Operators in Quantum Machine Learning: Response Properties in Chemical Space". In: *J. Chem. Phys.* 150.6 (2019), p. 064105.

[117] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems.* 2015.

[118] Christine C. Mattheus, Anne B. Dros, Jacob Baas, Auke Meetsma, Jan L. De Boer, and Thomas T. M. Palstra. "Polymorphism in Pentacene". In: *Acta Crystallogr. C.* 57.8 (2001), pp. 939–941.

[119] Axel D. Becke. "Becke's Three Parameter Hybrid Method Using the LYP Correlation Functional". In: *J. Chem. Phys.* 98.492 (1993), pp. 5648–5652.

[120] Jeng-Da Chai and Martin Head-Gordon. "Systematic Optimization of Long-Range Corrected Hybrid Density Functionals". In: *J. Chem. Phys.* 128.8 (2008), p. 084106.

[121] Florian Weigend and Reinhart Ahlrichs. "Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy". In: *Phys. Chem. Chem. Phys.* 7.18 (2005), pp. 3297–3305.

[122] Florian Weigend. "Accurate Coulomb-Fitting Basis Sets for H to Rn". In: *Phys. Chem. Chem. Phys.* 8.9 (2006), pp. 1057–1065.

[123] Frank Neese. “The ORCA Program System”. In: *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2.1 (2012), pp. 73–78.

[124] Frank Neese. “Software Update: The ORCA Program System, Version 4.0”. In: *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 8.1 (2018), e1327.

[125] Huizhen Zhao, Fuxian Liu, Longyue Li, and Chang Luo. “A Novel Softplus Linear Unit for Deep Convolutional Neural Networks”. In: *Appl Intell* 48.7 (2018), pp. 1707–1720.

[126] François Chollet et al. *Keras*. 2015. URL: `https://github.com/fchollet/keras`.

[127] Diederik P. Kingma and Jimmy Ba. “Adam: A Method for Stochastic Optimization”. In: *ArXiv14126980 Cs* (2017). arXiv: `1412.6980 [cs]`.

[128] Lisha Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar. “Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization”. In: *ArXiv160306560 Cs Stat* (2018). arXiv: `1603.06560 [cs, stat]`.

[129] Oana D. Jurchescu, Jacob Baas, and Thomas T. M. Palstra. “Effect of Impurities on the Mobility of Single Crystal Pentacene”. In: *Appl. Phys. Lett.* 84.16 (2004), pp. 3061–3063.

[130] Michel Grandbois, Martin Beyer, Matthias Rief, Hauke Clausen-Schaumann, and Hermann E. Gaub. “How Strong Is a Covalent Bond?” In: *Science* 283.5408 (1999), pp. 1727–1730.

[131] Stephen R. Forrest. “The Path to Ubiquitous and Low-Cost Organic Electronic Appliances on Plastic”. In: *Nature* 428.6986 (2004), pp. 911–918.

[132] John E. Anthony. “Functionalized Acenes and Heteroacenes for Organic Electronics”. In: *Chem. Rev.* 106.12 (2006), pp. 5028–5048.

[133] Nils Schieschke, Beatrix M. Bold, Philipp M. Dohmen, Daniel Wehl, Marvin Hoffmann, Andreas Dreuw, Marcus Elstner, and Sebastian Höfener. “Geometry Dependence of Excitonic Couplings and the Consequences for Configuration-Space Sampling”. In: *J. Comput. Chem.* 42.20 (2021), pp. 1402–1418.

[134] W. M. Young and E. W. Elcock. “Monte Carlo Studies of Vacancy Migration in Binary Ordered Alloys: I”. In: *Proc. Phys. Soc.* 89.3 (1966), p. 735.

[135] Ilya A. Balabin and José N. Onuchic. “Dynamically Controlled Protein Tunneling Paths in Photosynthetic Reaction Centers”. In: *Science* 290.5489 (2000), pp. 114–117.

[136] Jeffrey S. Meth, C. D. Marshall, and M. D. Fayer. “An Examination of Radiative and Nonradiative Excitation Transport in Thin Anthracene Crystals: Transient Grating Experiments”. In: *Solid State Commun.* 74.4 (1990), pp. 281–284.

[137] Michael A. Heinrich, Jens Pflaum, Ashutosh K. Tripathi, Wolfgang Frey, Michael L. Steigerwald, and Theo Siegrist. “Enantiotropic Polymorphism in Di-indenoperylene”. In: *J. Phys. Chem. C* 111.51 (2007), pp. 18878–18881.

[138] Giovanni Bussi, Davide Donadio, and Michele Parrinello. “Canonical Sampling Through Velocity Rescaling”. In: *J. Chem. Phys.* 126.1 (2007), p. 014101.

[139] Van Quan Vuong, Jissy Akkarapattiakal Kuriappan, Maximilian Kubillus, Julian J. Kranz, Thilo Mast, Thomas A. Niehaus, Stephan Irle, and Marcus Elstner. "Parametrization and Benchmark of Long-Range Corrected DFTB2 for Organic Molecules". In: *J. Chem. Theory Comput.* 14.1 (2018), pp. 115–125.

[140] Samuele Giannini, Wei-Tao Peng, Lorenzo Cupellini, Daniele Padula, Antoine Carof, and Jochen Blumberger. "Exciton Transport in Molecular Organic Semiconductors Boosted by Transient Quantum Delocalization". In: *Nat. Commun.* 13.2755 (2022), pp. 2041–1723.

[141] Juan Aragó and Alessandro Troisi. "Regimes of Exciton Transport in Molecular Crystals in the Presence of Dynamic Disorder". In: *Adv. Funct. Mater.* 26.14 (2016), pp. 2316–2325.

[142] Xin Li, Robert M. Parrish, Fang Liu, Sara I. L. Kokkila Schumacher, and Todd J. Martinez. "An ab initio Exciton Model Including Charge-Transfer Excited States". In: *J. Chem. Theory Comput.* 13.8 (2017), pp. 3493–3504.

[143] Michele Nottoli, Sandro Jurinovich, Lorenzo Cupellini, Alastair T. Gardiner, Richard Cogdell, and Benedetta Mennucci. "The Role of Charge-Transfer States in the Spectral Tuning of Antenna Complexes of Purple Bacteria". In: *Photosyn. Res.* 137.2 (2018), pp. 215–226.

[144] Lars Olof Björn. *Photobiology: The Science of Light and Life*. Springer, 2015.

[145] Juergen Koepke, Xiche Hu, Cornelia Muenke, Klaus Schulten, and Hartmut Michel. "The Crystal Structure of the Light-Harvesting Complex II (B800–850) from Rhodospirillum molischianum". In: *Structure* 4.5 (1996), pp. 581–597.

[146] Thorsten Ritz, Sanghyun Park, and Klaus Schulten. "Kinetics of Excitation Migration and Trapping in the Photosynthetic Unit of Purple Bacteria". In: *J. Phys. Chem. B* 105.34 (2001), pp. 8259–8267.

[147] Masoud Mohseni, Patrick Rebentrost, Seth Lloyd, and Alan Aspuru-Guzik. "Environment-Assisted Quantum Walks in Photosynthetic Energy Transfer". In: *J. Chem. Phys.* 129.17 (2008), 11B603.

[148] Carsten Olbrich and Ulrich Kleinekathöfer. "Time-Dependent Atomistic View on the Electronic Relaxation in Light-Harvesting System II". In: *J. Phys. Chem. B* 114.38 (2010), pp. 12427–12437.

[149] Thomas Renger and Frank Müh. "Understanding Photosynthetic Light-Harvesting: a bottom up Theoretical Approach". In: *Phys. Chem. Chem. Phys.* 15.10 (2013), pp. 3348–3371.

[150] Aaron Sisto, Clem Stross, Marc W. van der Kamp, Michael O'Connor, Simon McIntosh-Smith, Graham T. Johnson, Edward G. Hohenstein, Fred R. Manby, David R. Glowacki, and Todd J. Martinez. "Atomistic Non-Adiabatic Dynamics of the LH2 Complex with a GPU-accelerated ab initio Exciton Model". In: *Phys. Chem. Chem. Phys.* 19.23 (2017), pp. 14924–14936.

[151] Maria Ilaria Mallus, Yashoj Shakya, Jigneshkumar Dahyabhai Prajapati, and Ulrich Kleinekathöfer. "Environmental Effects on the Dynamics in the Light-Harvesting Complexes LH2 and LH3 based on Molecular Simulations". In: *Chem. Phys.* 515 (2018), pp. 141–151.

[152] Alex D. MacKerell Jr., Donald Bashford, M. L. D. R. Bellott, Roland Leslie Dunbrack Jr., Jeffrey D. Evanseck, Martin J. Field, Stefan Fischer, Jiali Gao, H. Guo, Sookhee Ha, et al. "All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins". In: *J. Phys. Chem. B* 102.18 (1998), pp. 3586–3616.

[153] Herman J. C. Berendsen, J. P. M. van Postma, Wilfred F. Van Gunsteren, A. R. H. J. DiNola, and Jan R. Haak. "Molecular Dynamics with Coupling to an External Bath". In: *J. Chem. Phys.* 81.8 (1984), pp. 3684–3690.

[154] Michele Parrinello and Aneesur Rahman. "Polymorphic Transitions in Single Crystals: A new Molecular Dynamics Method". In: *J. Appl. Phys.* 52.12 (1981), pp. 7182–7190.

[155] Margus Rätsep, Zheng-Li Cai, Jeffrey R. Reimers, and Arvi Freiberg. "Demonstration and Interpretation of Significant Asymmetry in the Low-Resolution and High-Resolution Qy Fluorescence and Absorption Spectra of Bacteriochlorophyll a". In: *J. Chem. Phys.* 134.2 (2011), 01B608.

[156] Takeshi Yanai, David P. Tew, and Nicholas C. Handy. "A new Hybrid Exchange–Correlation Functional using the Coulomb-Attenuating Method (CAM-B3LYP)". In: *Chem. Phys. Lett.* 393.1-3 (2004), pp. 51–57.

[157] Sebastian Höfener. "The KOALA Program: Wavefunction Frozen-Density Embedding". In: *Int. J. Quantum Chem.* 121.3 (2021), e26351.

[158] Oleksandr V. Mikhnenko, Paul W. M. Blom, and Thuc-Quyen Nguyen. "Exciton Diffusion in Organic Semiconductors". In: *Energy Environ. Sci.* 8.7 (2015), pp. 1867–1888.

[159] Edoardo Cignoni, Vladislav Slama, Lorenzo Cupellini, and Benedetta Mennucci. "The Atomistic Modeling of Light-Harvesting Complexes from the Physical Models to the Computational Protocol". In: *J. Chem. Phys.* 156.12 (2022), p. 120901.

[160] John D. Hunter. "Matplotlib: A 2D Graphics Environment". In: *Comput. Sci. Eng.* 9.3 (2007), pp. 90–95.

Publikationliste

Publikationen in Fachzeitschriften

Beatrix M. Bold, Monja Sokolov, Sayan Maity, Marius Wanko, **Philipp M. Dohmen**, Julian J. Kranz, Ulrich Kleinekathöfer, Sebastian Höfener, Marcus Elstner (2020). Benchmark and performance of long-range corrected time-dependent density functional tight binding (LC-TD-DFTB) on rhodopsins and light-harvesting complexes. *Phys. Chem. Chem. Phys.*, 22.19, 10500–10518

Mila Krämer, **Philipp M. Dohmen**, Weiwei Xie, Daniel Holub, Anders S. Christensen, Marcus Elstner (2020). Charge and Exciton Transfer Simulations Using Machine-Learned Hamiltonians. *J. Chem. Theory Comput.*, 16.7, 4061–4070.

Nils Schieschke, Beatrix M. Bold, **Philipp M. Dohmen**, Daniel Wehl, Marvin Hoffmann, Andreas Dreuw, Marcus Elstner, Sebastian Höfener (2021). Geometry dependence of excitonic couplings and the consequences for configuration-space sampling. *J. Comput. Chem.*, 42.20, 1402–1418.

Publikationen in Vorbereitung

Philipp M. Dohmen, Mila Krämer, Patrick Reiser, Pascal Friederich, Marcus Elstner, Weiwei Xie. Charge Transfer Simulations in Organic Semiconductors driven by Machine Learned Hamiltonians and Diagonal Forces.

Philipp M. Dohmen, Farhad Ghalami, Mila Krämer, Patrick Reiser, Pascal Friederich, Marcus Elstner, Weiwei Xie. Exciton Transfer Simulations in Organic Semiconductors driven by Machine Learned Hamiltonians and Gradients.

Philipp M. Dohmen, Monja Sokolov, David Hoffmann, Mila Krämer, Marcus Elstner. Exciton Dynamics in the Light-Harvesting Complex II with Non-Adiabatic Molecular Dynamics Simulations

Eidesstattliche Versicherung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Hilfsmittel verwendet habe. Darüber hinaus erkläre ich, dass die Arbeit nicht anderweitig als Prüfungsarbeit oder als Dissertation bei einer anderen Fakultät verwendet wird oder wurde.