Karlsruhe Institute of Technology

Modeling structural and electronic properties of nano-scale systems

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

von der Fakultät für Physik des
Karslruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

Dipl.-Phys. Tobias Neumann
aus Titisee-Neustadt

Tag der mündlichen Prüfung:	13.05.2016
Referent:	Prof. Dr. Wolfgang Wenzel
Korreferent:	Prof. Dr. Gerd Schön

Deutsche Zusammenfassung

Die anthropogene Erderwärmung ist eine der größten Herausforderungen der modernen Gesellschaft. Im Rahmen des Pariser Klimagipfels im Dezember 2015 einigten sich daher 195 Staaten auf eine Begrenzung der Erderwärmung auf „weit unter“ 2 °C im Vergleich zum vorindustriellen Zeitalter. Zwei mögliche Maßnahmen zur Reduktion der für die Erderwärmung verantwortlichen Treibhausgase sind die Erzeugung regenerativer Energie durch Photovoltaik und die Reduktion des Strombedarfs durch den Einsatz elektrischer Verbraucher mit hohem Wirkungsgrad, beispielsweise im Bereich der Beleuchtung.
Sowohl in der Photovoltaik als auch bei Leuchtmitteln sind bereits Halbleiter-basierte Bauteile verfügbar und weit verbreitet. Diese Solarzellen bestehen jedoch größtenteils aus Kristallen, deren Herstellung einer hohen Präzision und viel Energie bedarf, wodurch kostengünstige Produktion großflächiger Solarzellen begrenzt wird. Ebenso schreitet der Einzug von Leuchtdioden (LEDs) zur Raumbeleuchtung durch das begrenzte Farbspektrum des emittierten Lichts und die Punktförmigkeit der Lichtquelle nur langsam voran.
Ein Konzept, das die Probleme in diesen Anwendungsfeldern lösen könnte, sind Bauteile aus dünnen Schichten kleiner organischer Moleküle oder Polymerstrukturen (Organische Elektronik, OE). Die sehr große Anzahl möglicher chemischer Verbindungen, die per Aufdampfung oder aus der Flüssigphase kostengünstig und mittels einfacher Prozesse (bspw. Rolle-zu-Rolle Verfahren) zu großflächigen elektrischen Bauteilen verarbeitet werden, verspricht eine Vielzahl an möglichen Anwendungen im Bereich organischer Leuchtdioden (OLEDs) oder organischer Photovoltaik (OPV).
Gegenwärtig verfügbare Bauteile blieben bisher jedoch aufgrund geringer Wirkungsgrade und kurzer Lebensdauern hinter den Erwartungen zurück. Der wissenschaftliche Fortschritt ist dadurch behindert, dass die sukzessive experimentelle Synthese, Produktion und Vermessung kleiner Bauteile im Labormaßstab eine zeitlich und kostentechnisch ineffiziente Strategie zur Auswahl geeigneter chemischer Verbindungen darstellt. Die theoretische Beschreibung von OE-Bauteilen kann hier zweierlei Abhilfe schaffen: Zum einen ist ein tieferes Verständnis der fundamentalen quantenmechanischen Prozesse notwendig, um die Ursachen begrenzter Wirkungsgrade und schnellen Alterns der Materialien zu verstehen und zu beheben. Zum anderen ermöglicht die Simulation makroskopischer elektrischer Eigenschaften auf der Basis der elektronischen Struktur der einzelnen Moleküle eine effiziente Vorauswahl geeigneter Materialien, was den Zeit- und Kostenaufwand der Laborentwicklung der Bauteile drastisch senken kann.
In dieser Arbeit wird ein Konzept zur effizienten Modellierung der elektronischen Eigenschaften organischer Materialien vorgestellt. Durch die geschickte Kombination mehrerer Methoden auf verschiedenen physikalischen Zeit- und Größenordnungen können makroskopische Eigenschaften der Bauteile aus dem Aufbau der einzelnen Moleküle abgeleitet werden. Die Anwendung auf dünne Schichten kleiner organischer Moleküle, metall-organischer Oligomere

und metall-organischer Gitterverbindungen (Metal-Organic Frameworks, MOFs) in Kombination mit experimentellen Ergebnissen demonstriert die Verlässlichkeit und die Übertragbarkeit dieses multiskaligen Modellierungs-Ansatzes auch auf verwandte Anwendungsfelder. Die insgesamt gute Übereinstimmung mit experimentellen Daten erweckt Hoffnung auf computergestützte Entwicklung von OE-Bauteilen, wodurch sich Entwicklungszeit und -kosten senken lassen und somit die schnelle Etablierung der OE-Technologie ermöglicht wird.
Genaue Kenntnis über Position, Orientierung und Deformierung der Moleküle in realistisch großen Strukturen, die sich üblicherweise auf Zeitskalen im μs-Bereich formieren, ist essentiell für die korrekte Berechnung mikroskopischer Ladungsträgereigenschaften, aus denen sich makroskopische Größen, wie die Ladungsträgermobilität, ableiten lassen. Bisherige Verfahren sind durch methodische Eigenheiten in der Systemgröße und der simulierten Zeitskala beschränkt. In diesem Rahmen werden zwei Verfahren zur beschleunigten Simulation meso-skaliger (10 nm - 1 μm) Molekülstrukturen, einem der limitierenden Prozesse des oben genannten multiskaligen Modellierungsansatzes, vorgestellt. Die im Rahmen dieser Arbeit entwickelten Verfahren beinhalten die Konstruktion kollektiver, teilweise korrelierter, Systemänderungen in Monte-Carlo-Verfahren. Dadurch entstehen kurze Autokorrelationszeiten der Freiheitsgrade und eine hohe Systemdynamik, was die Simulation physikalisch langsamer Prozesse mit atomarer Auflösung ermöglicht. Die Anwendung auf verschiedene, zunehmend komplexe Modellsysteme und der Vergleich mit etablierten Verfahren zeigt, dass die vorgestellten Algorithmen zur Beschleunigung der Systemdynamik um bis zu zwei Größenordnungen führen. Dies ermöglicht prinzipiell die Simulation langsamer Prozesse in der mesoskaligen Strukturbildung und trägt somit signifikant zur computergestützten Erforschung und Entwicklung organischer Elektronik bei.

Contents

1. Introduction

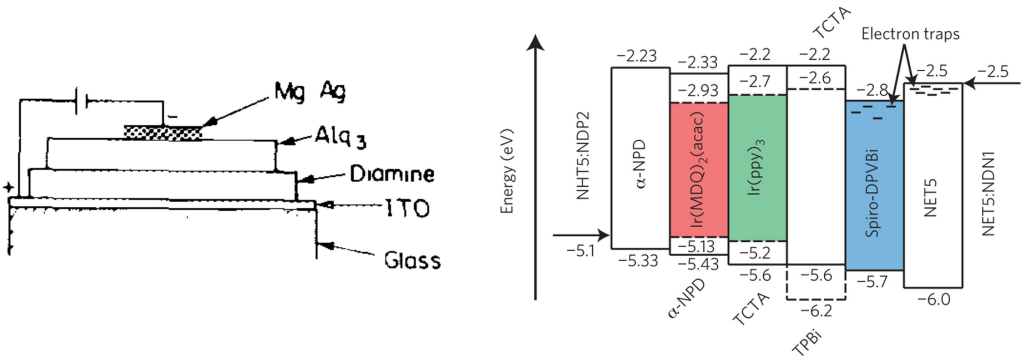
1.1. Electronic Devices Based on Organic Compounds

1.1.1. Efficient Technology against Global Warming

One of the key challenges of modern society is presented by global warming. While the role of human activities in the increase of global temperature was originally debated, it is now a widely accepted fact that global warming is an anthropogenic effect. Especially the emission of green house gases into the atmosphere warms the lower atmosphere of the planet and the oceans, causing a significant melt-down of perpetual ice at north and south poles and a rise of the water levels, threatening the existence of whole communities. Further, extreme meteorological perturbations such as frequent occurrences of hurricanes, floods and typhoons observed in the past decade are expected to be a direct consequence of anthropogenic global warming.
In context of the nuclear phase-out in Europe and the only slow switch towards renewable energy sources, coal-fired power plants remain a major source of electrical energy, leading to steady emissions of CO_2 into the atmosphere. In line with a continuous improvement of renewable energy resources, especially photovoltaics and storage devices for electrical power, an overall reduction of electrical consumption in households and industry remains one of the main challenges. The substitution of common light bulbs with power conversion efficiencies in the lower single digit range has thus received increasing attention over the past decades. Special energy saving lamps and silicon based light emitting diodes (LEDs) are widely available, but their radiation light spectrum is limited and often considered cold and unnatural and their disposal remains problematic. While inorganic semiconductors are widely applied in photovoltaic applications, the necessary manufacturing of sensitive crystalline layers prevents the large-scale, inexpensive production of durable, large-area devices.

1.1.2. Organic Light Emitting Diodes and Organic Photovoltaics

A different technology promising a wide range of devices for the application in lighting [1, 2], photovoltaics [3–5] and field-effect transistors [6] are organic electronic materials. Novel processing techniques such as solvent processing or printing of organic compounds promise the low-cost production of thin, flexible large-area devices manufactured in roll-to-roll processes [7, 8]. Since the discovery of electric luminescence in thin organic layers in 1987 by Tang and Van Slyke [9], substantial strides towards device efficiency [10–13] and improvements regarding processing techniques were made. The originally monolayer-structured device, displayed in Fig. 1.1.1a, was developed further towards multi-layered structures. These devices comprise multiple doped or pure materials [14] to allow white



(a) Figure taken from [9] (b) Figure taken from [14]

Figure 1.1.1.: Since the discovery of electrical luminescence in organic electronic layers in 1987, the Alq_3 OLED (a) originally proposed by Tang and Van Slyke was further developed into complex, multi-layered structures (b). The complex electronic structure of composite devices including bulk- and interface effects poses a major challenge in the systematic experimental analysis and characterization of device properties.

emission and optimized power conversion efficiencies (PCEs) by tuning bandgaps and level alignment of neighboring layers. The setup of one of these complex devices is illustrated in Fig. 1.1.1b.
The possibilities to synthesize materials for organic light emitting diodes (OLED) and organic photovoltaic (OPV) devices open many as-of-yet unexplored opportunities. In principle, this enables the specific modification of device properties on the molecular level. However, not all processes in OLEDs and OPV are fully understood. The mapping of single-molecule characteristics onto device properties is the consequence of several microscopic processes, rendering an analytical transfer of properties impossible. As a consequence, the identification of suitable organic molecules for the application in OLEDs or OPV is dominated by a trial-and-error procedure of experimentalists. Repetitive experimental synthesis, small scale production and device characterization is a costly search for the needle in the haystack, impeding the progress of organic electronic devices.

1.1.3. Metal-Organic Frameworks

A different class of materials that has attracted wide attention during the past decades are nano-porous metal-organic frameworks (MOFs). These ordered, three-dimensional structures consisting of metallic centers linked by organic ligands have been successfully applied in a variety of applications such as hydrogen storage [15], CO_2 capture [16, 17], catalysis [18, 19], sensing [20] and photovoltaics [21]. By mixing metal salts with organic ligands in salvatochemical reactions [22, 23], a variety of over 20000 different MOF structures have been experimentally realized to date. Furthermore, refined techniques on the basis of self-assembling monolayers allow the synthesis of single-crystal surface-mounted MOFs (SURMOFs) [24]. A

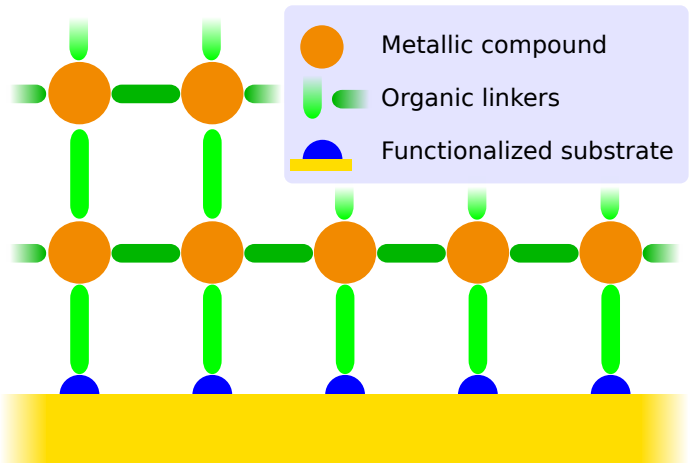


Figure 1.1.2.: Schematic illustration of metal-organic frameworks

schematic illustration of MOFs is given in Fig. 1.1.2.
Like in OLEDs and OPV, the vast possible number of organic linkers that determine intrinsic properties, such as pore size, electrical and optical excitation and diffusion behaviour of loaded guest molecules, open the prospect of specifically designed light-weight devices with tunable electronic properties. However, the lack of intrinsic electrical conductance, which results from the absence of significant band dispersion for most systems remains a major challenge in respect to application of MOFs in electrical devices. While modification of intrinsic properties by loading organic molecules into MOF pores was successfully demonstrated [25], the systematic *in vitro* synthesis, MOF growth and characterization remains a costly and time-consuming challenge. Moreover, experimental measurements fail to analyze processes on the atomistic scale within the pores, limiting the development of a systematic understanding of processes and characteristics of MOFs.

1.2. Modeling (metal-) Organic Electronics

1.2.1. Towards *in silico* Device Simulations

For metal-organic systems consisting of small organic molecules, polymers or metal-organic frameworks, the huge number of possible constituents is both advantage and disadvantage of the technology: Optimization starting at the molecular level by subsequent experimental synthesis, small-scale production and experimental device analysis, however, is slow and costly. Furthermore, as the macroscopic device properties are a result of several processes on the microscopic scale, a systematic understanding of all relevant processes responsible for e.g. limited efficiencies, short life times due to fast degradation or the insulating behavior of MOFs cannot be gained solely from experimental observations.
A different approach that may greatly reduce both cost and time as well as accelerate the design of (metal-) organic electronic devices is the application of predictive computational

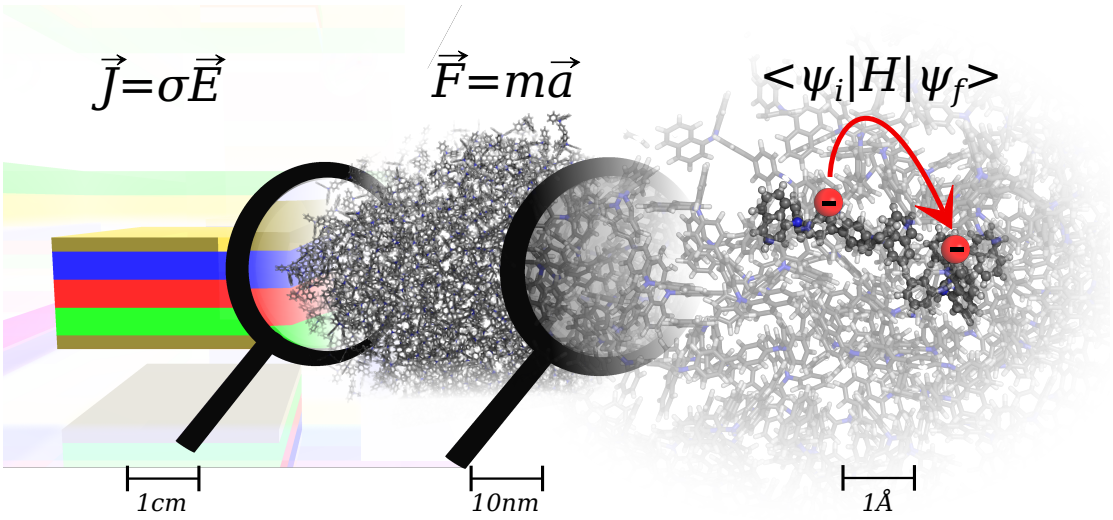


Figure 1.2.1.: Phenomena on different time and length scales pose a major challenge in the theoretical description of OE devices. The macroscopic properties on length scales of 1 cm and time scales of hours or days is determined by the mesoscopic molecular structure on the scale of 10 nm to 1 μm, formed on time scales of miliseconds, and microscopic quantum mechanical properties, such as electronic coupling between neighboring molecules on the scale of 1 Å.

modeling techniques. First, a bottom up simulation of devices starting on the quantum mechanical level allows the observation and analysis of the individual processes in organic electronic devices and is thus a systematic approach to tackle drawbacks such as degradation and low power conversion efficiencies. Secondly, as the cost for computational power has been steadily reduced over the past decade, established simulation approaches enable high-throughput pre-screening of materials at low cost in comparison to experimental approaches. Inclusion of electronic characteristics of bulk materials, organic-organic and inorganic-organic interfaces into the modeling process enables the description of full devices in the computer-chip, i.e. *in silico* design of new electronic devices.
However, while simulations contribute significantly to the development of devices in many industrial and academic areas such as automotive and aerospace industry or the design of silicon based electronics, molecular simulations play as yet only a minor role in the development of novel electronic organic materials and devices. In fact, the development of simulation methods for the description of full organic electronic devices has not yet been fully established, but is the focus of intense current research [3, 14, 27–30].
The slow development regarding application of simulations in organic material science is a result of phenomena occurring on multiple time and length scales in the device, as illustrated in Fig. 1.2.1: Configurational disorder resulting from either amorphous molecular packing of polymers or single molecules, or the atomic structure of loaded compounds in MOFs renders analytical solutions difficult. Further, inhomogeneous environment changes the electronic structure of the single constituents within the structure depending on their position. This

requires the individual calculation of the electronic structure of every single constituent. Although there are multiple methods readily available, the treatment of full devices on the micrometer length scale on the full quantum mechanical level is not feasible. Neglecting the quantum mechanical effects, on the other hand, will inevitably fail to capture relevant electronic effects necessary for the proper description of the device characteristics.

1.2.2. Multiscale Modeling for Full Devices Simulations

Efforts have been made to extend the capabilities of quantum chemistry methods towards the treatment of several thousand electrons [31], but it is apparent that present computational resource will fail to cover full devices on the nanometer-scale consisting of the order of 10^4 atoms. Further, investigation of different effects on different time and length scales is essential for the systematic understanding of microscopic processes and their influence on the overall device performance.
Hence, the predictive full device calculations via a single method is presently infeasible. An alternative approach lies in multiscale modeling techniques, where the treatment of different effects with specific methods is adapted to their specific length and time scales: First, the electronic structure of single molecules or pairs of molecules is calculated with high accuracy using quantum chemistry approaches with high accuracy. Subsequently the interaction between nuclei is mapped onto classical force-field functions [32–35], parametrized by the electronic structure, in order to generate mesoscopic models of realistic bulk morphologies and interfaces, resulting in structures containing several thousands of atoms. Microscopic parameters such as the electronic coupling between neighboring molecules are then extracted from quantum chemistry calculations and fed into macroscopic charge transport simulations, resulting in electronic device characteristics [27, 29, 36].
By applying this approach, device sample structures are calculated wile incorporating relevant effects on the quantum mechanical scale. Further, it is possible to systematically analyze the influence of microscopic effects on macroscopic device properties.

1.2.3. Meso-Scale Simulations: the Bottle-Neck of Efficient Device Simulations

One of the most challenging tasks in the multiscale modeling process is describing structure formation on the mesoscopic scale (10 nm - 1 μm), where functional substructures for OLEDs and OPV devices are formed. In the established molecular dynamics (MD) approach the inter- and intra-molecular interactions are modeled by classical forces. Newton's equation of motion are then discretely integrated by updating particle velocities according to atomic forces at subsequent time steps. Often, relevant processes in structural formation occur on millisecond time-scales and cannot be simulated due to the inherent limitation of MD to integration time steps on the femtosecond scale. Computational parallelization of available programs is limited to the evaluation of energy functions and gradients, scaling well for up to 100 CPUs. While millisecond simulations of proteins were achieved on specially designed architectures [37, 38], the main limitation of the algorithm, namely the inherent time step, cannot be overcome on generally available modern high-performance computing (HPC) architectures. In fact, important issues like domain formation or even degradation which occur

on time scales between 1 s to several hours remain inaccessible.
An alternative approach to the simulation of physical trajectories by means of MD is the sampling of states representing thermodynamic ensembles by statistical Monte Carlo methods. Many established protocols are based on a Markov-chain of states sampled using energy differences between subsequent steps [39, 40]. An acceptance criterion is applied between two successive steps, resulting in thermodynamically equilibrated distribution of states. The computational effort spent on the energy evaluation is comparable to the calculation of gradients in MD, whereas the construction of "moves" between two successive steps of the Markov-process does not underlie the fundamental restriction of an inherent time step. Development of methods that perform large configurational changes per energy evaluation may therefore result in speedups by orders of magnitude in comparison to MD. However, the widely used generic Metropolis Monte Carlo [40] protocols suffer from a lack of systematic construction of correlated system changes with high acceptance rates. While further improvements to Metropolis MC include collective system changes [41–44], these methods lack the essential feature of large, correlated moves with high acceptance rates. Therefore, the advantages of Monte Carlo methods regarding the formation of molecular structures have to date not been exploited to the fullest.

1.3. Structure of the Thesis

The aim of this thesis is twofold: In the first part I will present the multiscale modeling approach sketched above and apply this protocol to organic and metal-organic electronic materials. In a second, more method-related part, I will discuss methods to increase sampling efficiency in Monte Carlo based simulation protocols for the simulation of structural formation on the meso-scale. Hence, this work is structured as follows.
In chapter 2, the multiscale modeling protocol for the simulation of organic materials will be presented in detail. Fundamental physical concepts, algorithms and methods for the calculation of structures, effects and processes on different length- and time-scales are discussed and the assembling of different methods towards the theoretical description of full devices is described.
Subsequently, chapter 3 presents four applications of the multiscale modeling approach. First, I will analyze electronic transport properties in an amorphous thin layer consisting of small organic molecules as used in OLED devices. Thereafter, I present a study on the electronic properties of one-dimensional metal-organic olygomers and three-dimensional metal-organic frameworks. In these systems, organic molecules are coordinated to metallic centers, leading to intriguing characteristics. In each study, comparison to experimental data is provided to demonstrate the applicability of the multiscale modeling approach to extended (metal-)organic structures.
In chapter 4, a novel Monte Carlo (MC) based algorithm, the Acceptance Rate Optimized Monte Carlo (AROMoCa) is developed, which ameliorates limitations of both molecular dynamics and other established MC protocols. In the first version of the method I use gradient information to focus moves on regions of the configurational space that are far from the thermodynamic equilibrium, leading to large collective, yet uncorrelated moves with acceptance

rates close to unity. The second, refined version of the algorithm allows the construction of correlated moves with high acceptance probability, using a quadratic approximation to the energy function. The application to test systems with increasing complexity demonstrates the effectiveness of the AROMoCa approach. Concluding chapter 4, I investigate the Model-Hopping (MH) protocol [45] regarding its applicability to the simulation of structural formation of organic crystalline domains.
I summarize the results of all studies in chapter 5 and discuss possibilities and limitations of the multiscale-modeling approach and the AROMoCa protocol. This chapter is concluded by a sketch of further prospects, possible modifications and extensions to the presented methods which exceed the scope of this thesis, but hopefully lead to the establishment of *in silico* design of full organic electronic devices by predictive modeling.

2. Fundamental Methods and Concepts

2.1. Modeling Organic Electronic Devices

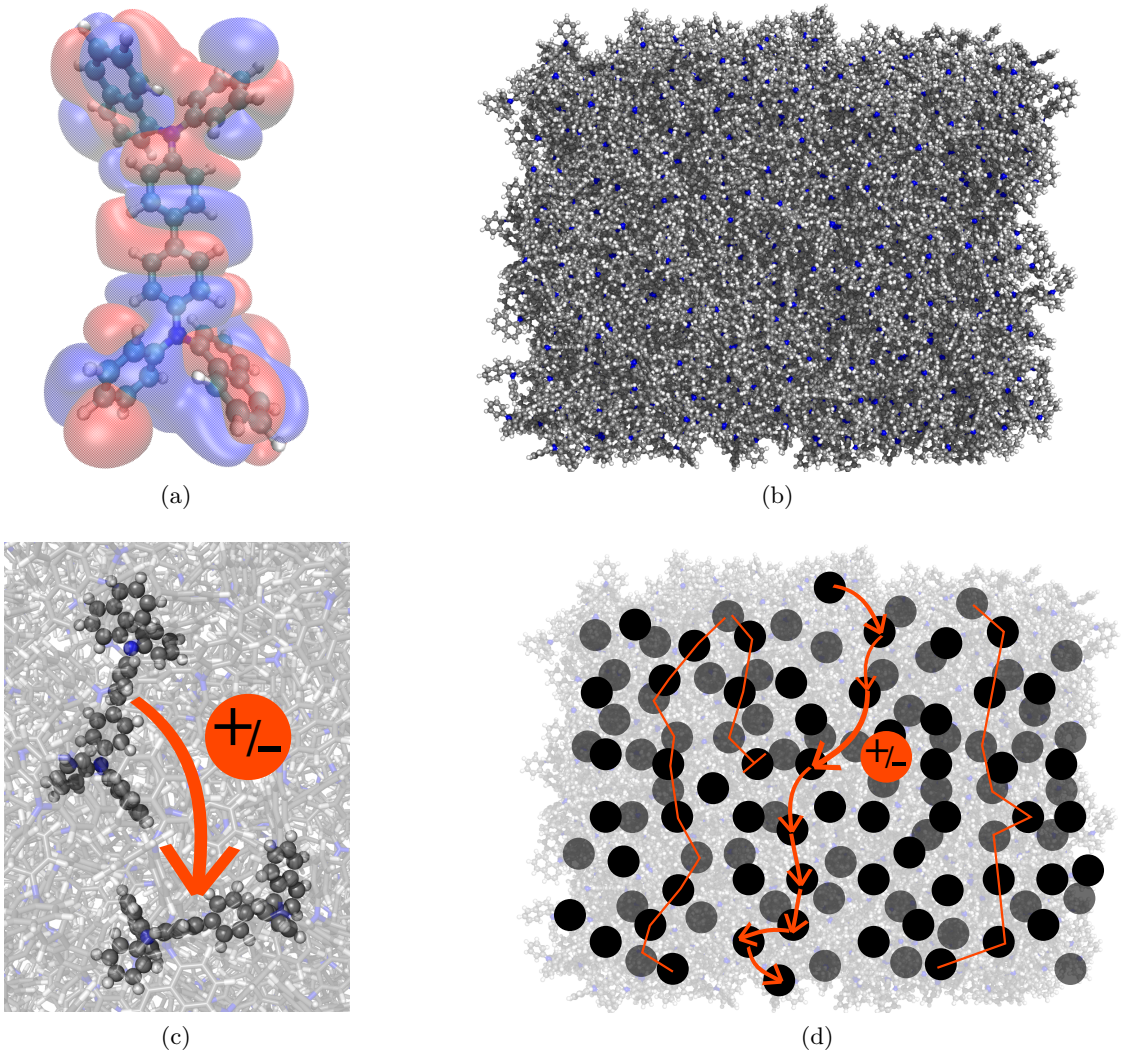
2.1.1. Organic Electronics for Efficient Lighting and Photovoltaics

The reduction of the emission of green house gases by efficient photovoltaic technology and power-saving electronic devices is one of the long term global environmental targets. In this context, the development of long-lived inorganic semiconductor devices, so far consisting of mostly silicon with mobilities of up to $1000\,\mathrm{cm}^2/\mathrm{Vs}$ [46] has been a crucial step. However, single crystalline semiconductors require refined production by e.g. vapor deposition, hampering the development of inexpensive, larg-area devices. Further, many of the crystalline materials with band-gaps in the optical range are indirect semiconductors. Therein, additional phonon processes are required for absorption and emission of photons, which results in reduced cross-sections and leads to rather thick layers and, consequently, to increased costs and long payback-times [46].
Another rising technology that has been intensely investigated over the last years, is organic electronics (OE) [46–48]. The large number of possible materials, the prospect of inexpensive, large-scale production and the promise of thin, flexible and efficient devices, raise the ambition of a variety of novel applications in addition or as a substitute to established light sources and PV technology. Since luminescence in thin amorphous organic layers was discovered in 1987 by C. W. Tang and S. A. VanSlyke [9], efforts were undertaken on the theoretical and experimental side to construct efficient organic light emitting diodes (OLEDs). The originally monolayer-structured device was developed further towards complex, multi-layered structures, consisting of multiple doped materials.
However, limited life times due to fast degradation and low efficiencies have limited the production of market-ready devices. Deeper understanding of the processes contributing to the emission of light in OLEDs is imperative for the establishment of durable devices [48–50]. Furthermore the wide range of possible molecules that can be synthesized and used in OLEDs turns the identification of suitable candidates into the search for the needle in a haystack. To fully understand degradation and reasons for limited efficiencies and thus facilitate the increasingly complex design of OLED materials, the development of an accurate and predictive theoretical description of the devices properties and processes is necessary.

2.1.2. The Multiscale Modeling Approach

Organic devices are often composed of amorphous layers of small organic molecules or polymers [51]. Unlike in crystalline semiconductors or metals, the carriers in many materials are not delocalized and thus cannot be described by extended wave packages. Moreover, due to

the amorphous packing of the molecules, each molecule has a unique electronic structure. The energies of electronic orbitals and hence the propagation of charges through the material depends strongly on the exact arrangement of the molecules. While there are methods available for the detailed calculation of electronic states of molecular structures (density functional theory (DFT), Hartree-Fock, see below), the description of a device on the nanometer-scale on the full quantum mechanical level is presently not feasible. On the other hand, a method that neglects the quantum mechanical nature of the electronic structure will fail to properly capture all effects necessary for the description of charge propagation in the device. Multiple length scales in the description of the full device (i.e. the electronic structure of a single molecule, the arrangement of several hundreds or thousands of molecules in a bulk and the charge-transport in the full device of the size of several nm) complicate the development of a single method for the full theoretical description of organic semiconductor devices.
A quite promising approach to tackle this obstacle is the development of a multiscale modeling protocol: Instead of simulating the whole device on the level of the smallest scale that is necessary to describe all relevant effects, the calculation is divided into multiple steps [49, 52–55]. Each step in the protocol deals with the system on a different scale, with methods designed to treat the effects of interest with appropriate accuracy. This multiscale method is exemplified in Fig. 2.1.1. In the case of a device consisting small molecules, the first step of a promising approach is the full-quantum mechanical calculation of the molecule via DFT- or Hartree-Fock-based approaches with single electron resolution. Molecular geometry, internal flexibility and charge density are calculated with high accuracy (section 2.3). As those methods are only applicable for systems consisting of up to 100-200 atoms, the calculation of the geometry of a full molecular layer requires different methods: a widely used possibility is to map charge density and intramolecular potential onto classical functions. Using those functions, Molecular modeling (MM) via molecular dynamics (MD) or Monte Carlo(MC)-based protocols is performed to generate sample structures consisting of up to several thousand molecules [14, 36, 56, 57]. Subsequently, quantum mechanical methods are again applied in order to calculate the energy levels of polarons and excitons on each molecule, as well as the electronic coupling between single molecules. Using these results as an input, the transfer rates for electrons hopping between molecules can be calculated using the Marcus theory of hopping transport. In the last step, Kinetic Monte Carlo (KMC) simulations are performed on a nanometer to micrometer scale to simulate the charge transport through the bulk material or device. In this step, the individual molecules are reduced to hopping sites and the propagation of charge carriers through the system is simulated based on previously calculated hopping rates (section 2.5). From the KMC simulations, charge carrier mobilities can be extracted. A change on the atomistic or electronic scale, e.g. deformation of molecules or substitution of atoms, feeds back into KMCs. This allows for example the study of doping (adding molecules of type B to a homogeneous structure of molecules of type A), defects and interfaces.
By applying different tools on different scales, this multiscale modeling approach fully exploits the advantages of all methods. Effects that can only be explained using a full quantum mechanical model are covered, while on the other hand no computational time is wasted by applying accuracy where it is not required. Similar approaches have been applied



(a) (b)

(c) (d)

Figure 2.1.1.: A multiscale modeling approach to the theoretical description of organic electronic devices: (a) The electronic structure of single molecules is calculated on a full quantum mechanical level. The atomic geometry and the orbitals, mapped onto point charges at the sites of the nuclei, are extracted as input for the classical force-field based calculations. (b) Using molecular modeling with molecular dynamics (MD) or Monte Carlo (MC) based methods, a sample morphology of several hundreds of molecules is generated, containing coordinates of all atoms. (c) For each pair of molecules in this sample structure, orbital overlap (or electronic coupling) and hopping rate of charge carriers (electrons or holes) between the molecules is calculated. (d) For the calculation of electronic transport properties, each molecule is reduced to a single hopping site with unique on-site energy. The hopping rates between each pair of molecules are fed into a Kinetic Monte Carlo (KMC) simulation. The KMC runs identify the most likely propagation paths through the bulk and yield values for electronic bulk properties, for example the charge carrier mobility.

for the study of organic electronic materials, reproducing correct material characteristics [3, 14, 27, 28, 58].

The structure of this chapter basically follows the development of the multiscale approach, starting at the electronic scale and progressing towards the mesoscopic scale of a full device: After an overview of the Born-Oppenheimer approximation in section 2.2, which is the basis of the methods used on the electronic level (DFT, Hartee-Fock) as well as on the atomistic level (MD, MC), two methods treating effects on the electronic level, DFT and Hartree-Fock, will be introduced in section 2.3. Moving to the next level of the multiscale approach, section 2.4 presents the approach of molecular modeling (MM), the simulation of atomistic structures based on classical energy functions. Subsequently, the calculation of hopping rates between molecules via Marcus theory and the kinetic Monte Carlo (KMC) approach for the simulation of charge carrier transport through organic layers is introduced in section 2.5, before summarizing the multiscale modeling approach in section 2.6.

2.2. Born-Oppenheimer-Approximation

The basis of the description and simulation of molecular systems is the stationary Schroedinger equation of the system, involving nuclei and electrons:

$$H\,|\Psi(\{\boldsymbol{r}_i\},\{\boldsymbol{R}_J\})\rangle = E\,|\Psi(\{\boldsymbol{r}_i\},\{\boldsymbol{R}_J\})\rangle \tag{2.2.1}$$

with the total energy E, the positions of electrons and nuclei, $\{\boldsymbol{r}_i\}$ and $\{\boldsymbol{R}_j\}$, and the multi-particle Hamiltonian H:

$$H = -\sum_J \frac{1}{2\,M_J}\nabla_J^2 + \sum_{J,L}\frac{Z_J\,Z_L}{R_{JL}} - \sum_i \frac{1}{2\,m_i}\nabla_i^2 + \sum_{i,k}\frac{1}{r_{ik}} - \sum_{i,J}\frac{Z_J}{r_{iJ}} \tag{2.2.2}$$

The first two expressions in the Hamiltonian represent the nuclei with the operators for the kinetic energy, $\frac{1}{2\,M_J}\nabla_J^2$, the charges Z_J and the distances between two nuclei, R_{JL}. The third and fourth expression is the Hamiltonian of a system of electrons with the kinetic operators $\frac{1}{2\,m_i}\nabla_i^2$ and the distance r_{ik} between two electrons. The last expression is the electrostatic interaction between electrons and nuclei, depending on the nuclei charge, Z_J, and the distance between electron and core, r_{iJ} in the electrostatic energy term.
The mass ratio between electrons and nuclei of approximately $\frac{m_i}{M_J} < \mathcal{O}(10^{-3})$ gave rise to an approximation to Eq. 2.2.2 that was proposed in 1927 by Max Born and Robert J. Oppenheimer: Due to the mass ratio, the velocity of nuclei is bound to be much smaller than the electron velocity and for the calculation of electronic states, the nuclei can be treated stationary (i.e. nuclei velocity set to zero). Then the nuclei induce a time-independent external potential for the electrons and the electronic Hamiltonian is reduced to:

$$H_{el} = -\sum_i \frac{1}{2\,m_i}\nabla_i^2 + \sum_{i,k}\frac{1}{r_{ik}} - \sum_{i,J}\frac{Z_J}{r_{iJ}} + c(\{R_J\}) \tag{2.2.3}$$

The constant c is the, in the case of stationary nuclei, time-independent electrostatic interaction between the nuclei and can be neglected.
The advantage of the Born-Oppenheimer-approximation is the decoupling of ionic and electronic degrees of freedom: While the solutions of Eq. 2.2.1 are dependent of all ionic and electronic degrees of freedom, the solutions of Eq. 2.2.3 depend only on the electronic degrees of freedom. The positions of the nuclei are merely parameters for the electronic states:

$$H_{el} \left| \Psi_{el,\{\boldsymbol{R}_J\}}(\{\boldsymbol{r_i}\}) \right\rangle = E_{el} \left| \Psi_{el,\{\boldsymbol{R}_J\}}(\{\boldsymbol{r_i}\}) \right\rangle \tag{2.2.4}$$

To put it in other terms, the total wave function of the system, Ψ_{total} is a product state of the nuclear wave function and the electronic wave function:

$$\Psi_{total} = \Psi_{nucl}(\{\boldsymbol{R}_J\}) \times \Psi_{el,\{\boldsymbol{R}_J\}}(\{\boldsymbol{r_i}\}) \tag{2.2.5}$$

The Born-Oppenheimer-approximation separates the two contributions and reduces the degrees of freedom drastically for the stationary solutions of the Schroedinger equation.

2.3. Electronic Structure Calculations

As explained in the introduction, the detailed modeling of a molecular system consisting of the order of $\mathcal{O}(10^4)$ atoms requires the application of methods on multiple scales. On the smallest scale, the calculation of the electronic structures of single molecules or small clusters is necessary to extract parameters for macroscopic calculations. Possible applications are the parametrization of electrostatic force fields by the electron density or the simulation of electronic hopping transport through organic layers.
In this section, two methods for the calculation of electronic structures of atoms and molecules will be presented. Both methods are applicable for systems consisting of approximately 100-200 atoms, depending on the required accuracy and the available computation time.
The first method, DFT, is based on the of a bijective mapping between Hamiltonian and wave function and accordingly Hamiltonian and electron density. The calculation of the ground state electron density is assumed to be sufficient for the accurate description of an atomic system. A functional of the ground state electron density is constructed, describing the energy of the system. Minimization of this functional with respect to the electron density then yields the ground state.
A different approach to the calculation of electronic structures is based on the Hartree-Fock-method. In this approach, the many-electron state is approximated by a product of single electron wave functions, and the Schroedinger equation is transformed into the Hartree-Fock equations. Further approximations, the NDDO and MNDO approximation, allow the treatment of larger systems and will be briefly presented.

2.3.1. Density Functional Theory

The fundamental principle of DFT is the connection between electronic ground state energy E, i.e. the Hamiltonian of the system, and the probability distribution function (also called

electron density), $\rho(\boldsymbol{r})$, that was established in 1964 by Hohenberg and Kohn [59]:

$$H_{el} \rightleftharpoons V_R(\{\boldsymbol{r}_i\}) \rightarrow \Psi_{el}\{\boldsymbol{r}_i\} \rightarrow \rho(\boldsymbol{r}) \tag{2.3.1}$$

The electronic Hamiltonian H_{el} is completely determined by the potential induced by the atom cores, $V_R(\{\boldsymbol{r}_i\})$, as the kinetic part and the electron-electron-interaction in Eq. 2.2.4 are the same for any electronic system of the same size. The parametrization $\{\boldsymbol{R}_J\}$ has been omitted for clarity. Based on this finding, Kohn and Sham showed one year later, that this relation is indeed bijective:

$$H \rightleftharpoons V_R(\{\boldsymbol{r}_i\}) \rightleftharpoons \Psi_{el}\{\boldsymbol{r}_i\} \rightleftharpoons \rho(\boldsymbol{r}) \tag{2.3.2}$$

This bijective mapping allows the expression of wave functions and ground-state energy as a functional of the ground state electron density:

$$\Psi_0 = \Psi_0[\rho_0] \tag{2.3.3}$$

$$E[\rho_0] = \langle \Psi_0[\rho_0]| - \sum_i \frac{1}{2\,m_i}\nabla_i^2 + \sum_{i,k}\frac{1}{r_{ik}} - \sum_{i,J}\frac{Z_J}{r_{iJ}}\,|\Psi_0[\rho_0]\rangle \tag{2.3.4}$$

$$= T[\rho_0] + V_{ee}[\rho_0] + \int dr\rho_0(\boldsymbol{r}) \sum_J \frac{Z_J}{|\boldsymbol{r} - \boldsymbol{R}_J|} \tag{2.3.5}$$

$$=: F[\rho_0] + \int dr\rho_0(\boldsymbol{r}) V_{ne} \tag{2.3.6}$$

Here $T[\rho_0]$ denotes the kinetic part of the electronic Hamiltonian, $V_{ee}[\rho_0]$ the electron-electron-interaction and the last expression describes the interaction between electrons and nuclei. As mentioned above, the electronic Hamiltonian of different molecular systems is defined solely by the potential induced the nuclei, while kinetic expression T and the interaction between the electrons V_{ee} is system independent. It is then reasonable to combine those parts into the system independent functional $F[\rho]$ in Eq. 2.3.6.
The ground state density is then derived by minimization of $E[\rho_0]$ with respect to ρ_0 with the variational principle

$$\frac{\delta}{\delta\rho(\boldsymbol{r})}\left[E_0[\rho(\boldsymbol{r})] - \mu\int dr\rho(\boldsymbol{r})\right] = 0 \tag{2.3.7}$$

with the condition that the electron density is normalized to the number of electrons:

$$\int \rho(\boldsymbol{r})dr = N \tag{2.3.8}$$

Construction of the functional $F[\rho]$

The key challenge now lies in the construction of the functional, as $F[\rho]$ cannot be derived analytically for a many-particle system. Following the Ansatz proposed by Kohn and Sham, an approximative functional is constructed for a system that can be solved and adding a correction. The first step is to approximate the real orbital wave functions by a many

particle wave function of non-interacting particles:

$$\Psi_s(\boldsymbol{r}_1, \boldsymbol{r}_2, ...\boldsymbol{r}_N) = \begin{vmatrix} \psi_1(\boldsymbol{r}_1) & \dots & \psi_N(\boldsymbol{r}_1) \\ \vdots & \ddots & \vdots \\ \psi_1(\boldsymbol{r}_N) & \dots & \psi_N(\boldsymbol{r}_N) \end{vmatrix} \tag{2.3.9}$$

In this approximation, the non-interacting electrons move in an effective position dependent potential:

$$V_{eff}(\boldsymbol{r}) = \int dr' \frac{\rho(r')}{|\boldsymbol{r} - \boldsymbol{r}'|} + V_{xc}(\boldsymbol{r}) \tag{2.3.10}$$

where the first part is the classical Coulomb interaction between the electrons and V_{xc} is the correction mentioned above that accounts for the exchange correlation of the electrons. This Ansatz is in fact first order perturbation theory: The set of decoupled wave functions in Eq. 2.3.9 is the exact solution to the Hamiltonian for non-interacting particles in the Coulomb potential induced by other electrons and nuclei. This unperturbated wave function is used to apply the first order perturbation in energy induced by the exchange interaction, U_x. Consequently, the functional is constructed as follows:

$$F[\rho] = T_s[\rho] + \int dr' \int dr \frac{\rho(r')}{|\boldsymbol{r} - \boldsymbol{r}'|} \rho(\boldsymbol{r}) + V_{xc} \tag{2.3.11}$$

The first part, T_s, is the kinetic functional for the system of non-interacting electrons and the second part is the classical electron-electron interaction. The last part, the exchange-correlation functional V_{xc} is the remainder that cannot be constructed from scratch. In an ideal case, it accounts exactly for the difference between the real system and the model system of non-interacting particles: The difference between the functionals T and T_s and the difference between V_{ee} and the classical Coulomb-contribution:

$$V_{xc} = T[\rho] - T_s[\rho] + V_{ee} - \int dr' \int dr \frac{\rho(r')}{|\boldsymbol{r} - \boldsymbol{r}'|} \rho(\boldsymbol{r}) \tag{2.3.12}$$

V_{xc} is called the exchange-correlation functional.

Exchange Correlation Functionals

The separation of the functional in the parts for the decoupled electrons and parts that are not known lead to the necessity of suitable exchange correlation functionals. To date, no explicit formulation of the exchange correlation functional has been presented and good approximations remain a key challenge of DFT. Three widely used exchange correlation functionals are the local density approximation, LDA, the local spin density approximation, LSDA, and the global gradient approximation GGA.
The LDA is based on the exchange correlation of an electron gas in the environment of a homogeneous positive charge distribution. The charge distribution of the background has to be set such that the system is neutral. For a uniform electron gas with electron density $\rho(\boldsymbol{r})$, the exchange correlation energy per particle, $\epsilon_{xc}(\rho(\boldsymbol{r}))$, can be calculated exactly. This was

done by Paul Dirac and Felix Bloch in 1929 way before the development of DFT [60]. Using this approximation, the exchange correlation functional in the LDA approximation is

$$V_{xc}^{LDA} = \int dr \rho(\boldsymbol{r}) \epsilon_{xc}(\rho(\boldsymbol{r})) \tag{2.3.13}$$

where $\epsilon_{xc}(\rho(\boldsymbol{r}))$ comprises the contributions of exchange and correlation. For the electron gas, the exchange energy is known exactly while the latter can only be determined approximately [61–64].
A first refinement of the LDA includes the electron spin: the local spin density approximation LSDA. The electron density is split into two separate spin densities, $\rho(\boldsymbol{r}) = \rho_\alpha(\boldsymbol{r}) + \rho_\beta(\boldsymbol{r})$, leading to an exchange-correlation functional:

$$V_x^{LSDA} = \int dr \rho(\boldsymbol{r}) \epsilon_{xc}(\rho_\alpha(\boldsymbol{r}), \rho_\beta(\boldsymbol{r})) \tag{2.3.14}$$

Again, $\epsilon_{xc}(\rho_\alpha(\boldsymbol{r}), \rho_\beta(\boldsymbol{r}))$ can be split into the exchange and the correlation contribution. While the exchange contribution to the functional can be derived directly from the LDA approach, the correlation is again just an approximation [61, 65].
As the correlation effects depend on all electrons in the system and cannot be represented by a local approximation, LDA and LSDA may not refined enough for many applications [66, 67]. An approach circumventing this strong limitation is the general gradient approximation, GGA [68].
In addition to the spin densities at the positions $\boldsymbol{r}$, the gradients of the spin densities $\nabla\rho_{\alpha,\beta}(\boldsymbol{r})$ are taken into account:

$$V_x^{GGA} = \int dr \rho(\boldsymbol{r}) \epsilon_{xc}(\rho_\alpha(\boldsymbol{r}), \rho_\beta(\boldsymbol{r}), \nabla\rho_\alpha(\boldsymbol{r}), \nabla\rho_\beta(\boldsymbol{r})) \tag{2.3.15}$$

Despite immense efforts to refine the exchange-correlation functionals, certain effects such as the Van-der-Waals interaction, are not covered by existing functionals and semi-empirical dispersion corrections need to be applied in addition. For a good overview of exchange-correlation functionals, their capabilities and limitations, see [69].

Implementations

There are multiple implementations of DFT available. Two of the most widely distributed packages for the accurate treatment of small systems are Gaussian and TURBOMOLE [70]. A package that applies more approximations and thus allows the calculation of larger systems is BigDFT [31]. The Vienna *ab initio* simulation package VASP [71, 72] is based on the expansion of wave functions into plane waves and can be applied to the calculation of periodic structures.
In the work presented in this thesis, TURBOMOLE was used for the treatment of small molecular entities. For the calculation of periodic systems such as metal-organic frameworks (MOFs), VASP was applied.

2.3.2. Semi-Empirical Hartree-Fock Methods

The Hartree-Fock method

Another approach to the prediction of the electronic structure of molecules is based on the Hartree-Fock method. For the treatment of atoms and ions, Hartree proposed to approximate of the state of all electrons by the product of single-electron wave functions in the 1920's. Later refinement by Fock in 1930 included the anti-symmetric nature of electrons by calculating the all-electron state as the Slater determinant of the single-electron wave functions:

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, ...\boldsymbol{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\boldsymbol{r}_1)\alpha_1 & \psi_1(\boldsymbol{r}_1)\beta_1 & \dots & \psi_{N/2}(\boldsymbol{r}_1)\beta_1 \\ \psi_1(\boldsymbol{r}_2)\alpha_2 & \psi_1(\boldsymbol{r}_2)\beta_2 & \dots & \psi_{N/2}(\boldsymbol{r}_2)\beta_2 \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\boldsymbol{r}_N)\alpha_N & \psi_1(\boldsymbol{r}_N)\beta_N & \dots & \psi_{N/2}(\boldsymbol{r}_N)\beta_N \end{vmatrix} \tag{2.3.16}$$

where α_i and β_i are spin functions [73]. The individual wave functions $\psi_i(\boldsymbol{r}_j)$ are orthogonal such that the whole wave function Ψ span an orthonormal basis.
The basis of the Hartree-Fock method is the application of the variational principle, to identify the ground state wave functions $\psi_i(\boldsymbol{r}_j)$ for which the energy is minimal:

$$\delta E = \delta \int d^3\Psi H \Psi = 0 \tag{2.3.17}$$

H is the Hamiltonian of the Born-Oppenheimer approximation, Eq. 2.2.3. Approximating the Hamiltonian by single-electron operators yields the Fock-operator:

$$F = H_{nucl} + \sum_{j=1}^{N/2} [2J_j - K_j] \tag{2.3.18}$$

A qualitative interpretation of this operator will be given in the following. For a detailed derivation and interpretation see Refs. [73, 74]. The operator H_{nucl} represents the energy of a single electron in the field of all nuclei. The operator J is called the Coulomb-term, defining the Coulomb-repulsion of the charge density of two electrons. K_j is a direct consequence of the antisymmetric definition of the product wave function as a Slater determinant. It represents the exchange energy of the electrons.
Inserting the Fock operator into Eq. 2.3.17, the solution to the ground state problem is then reduced to the Hartree-Fock equations:

$$F\psi_i = \epsilon_i S \psi_i \tag{2.3.19}$$

Here, F is the Fock-matrix (the representation of the Fock-Operator in a certain coordinate system). The vector ψ_i contains the coefficients for the wave function of electron i in the coordinate system with $2N$ dimensions. The overlap matrix S is introduced to couple the electronic functions, accounting for orbital overlap.

The Fock-matrix depends on the states ψ_i and the Hartree-Fock equations can be solved iteratively using the variational principle.
In most applications of the Hartree-Fock-method, the linear combination of atomic orbitals, LCAO, is applied. The molecular orbitals are set up as a linear combination of the (known) orbitals of the single atoms. The vectors ψ_i contain the coefficients of the single atomic orbitals.

NDDO

A simplification to the Hartree-Fock-method is to neglect the exchange energy of atoms centered around different atoms, called the neglect of diatomic differential overlap (NDDO) [75]. As the exchange between the orbitals of different atoms is represented by the matrix S, this approximation is in fact the substitution of S by the unit matrix. Specifically in the case of a system containing many atoms, large speedups can be achieved.

MNDO

A modification of the NDDO method, the modified neglect of diatomic differential overlap, MNDO, is a semi-empirical approximation to Hartree-Fock: Instead of calculating S, as in the Hartree-Fock-method, or setting S to unity, as in the NDDO approach, the overlap matrix and some integrals in the Fock-matrix are replaced by a parametric model [76]. The parametrization is either performed using experimental data directly or by the application of semi-empirical expressions. In the latter approach, the numerical parameters of the method are adjusted so that results of MNDO calculations fit well to experimental data.
The MNDO-method, as well as generalized methods based on MNDO, such as AM1, PM3, PM6 and PM7, are implemented in the molecular orbital package MOPAC [77, 78], that was used in parts of this thesis.

2.4. Molecular Modeling: Molecular Dynamics and Monte Carlo

For the treatment of systems larger than $\mathcal{O}(10^2 - 10^3)$ atoms, application of quantum mechanical methods as described in the previous section is not feasible. In DFT as well as the semi-empirical Hartree-Fock approaches, the calculation time for a system increases with $t \propto N_b^3$, where N_b is the number of basis functions.
However, for many systems containing chemically stable compounds, the focus lies on the ionic movement, i.e. the propagation of atoms comprising nuclei and electrons. Examples include the folding and unfolding of proteins and polymers, the simulation of atomic fluids or the structure prediction of amorphous layers of small organic molecules [1, 28, 79–86].
According to the Born-Oppenheimer approximation, the electronic structure will virtually adapt instantaneously to the movement of the nuclei, resulting in a $3N$-dimensional potential energy surface for a 3D-system consisting of N atoms. This potential energy surface induced by the electronic structure is parametrized by classical functions called "force fields" that depend solely on the positions of the nuclei or, for the case on non-bonded interactions, on the nuclei distances.

As the computation time for the calculation of the interaction via position-dependent pair-wise potentials scales quadratically with the number of atoms, $t \propto N^2$, larger systems can be treated in comparison to the QM methods described above. The description of molecular configurations via force fields is called molecular modeling (MM) and will be presented in this section.
There are two complementary methods for molecular modeling: Molecular dynamics (MD) and Monte Carlo (MC) based approaches. In MD, Newton's equation of motion are solved numerically. At each time step, the gradient of the energy functions, i.e. the forces on the atoms in the system, are calculated [1]. The N coupled differential equations are then solved by updating particle velocities at each time step. Atom positions are constructed by integrating accelerations and velocities discretely using a time step δt. As will be discussed later, this time step has to be chosen sufficiently small in order to avoid deviations from the physical trajectory. This imposes the main limitation to the MD method. Despite those limitation, MD is widely used for molecular modeling. Several implementations are available [88–92] and MD has been applied successfully to a large set of systems including the morphology prediction of thin films or glassy materials [28, 79–85]. Recently, specially designed hardware such as ANTON [37, 38] allowed the simulation of folding and unfolding of proteins on a millisecond scale.
The second approach is based on MC methods. In MC, a Markov-chain is constructed by random changes, so called "MC-moves" or just "moves" [39, 40]. The transition between to subsequent states $q(t_0)$ and $q(t_1)$ of a Markov-process depends only on the states at simulation time t_0 and t_1, but are independent of previous configurations, i.e. at $t < t_0$. Therefore, the configurations of each step are accepted or rejected based on energetic considerations regarding the configurations $q(t_0)$ and $q(t_1)$. The criterion for the acceptance has to be designed in a way that, in the limit of infinite simulation time, the states are distributed according to a thermodynamic distribution function. In the widely used Metropolis-MC scheme [40], the principle of detailed balance is applied in order to generate Boltzmann-distributed configurations. One of the main advantages of MC based methods is, that there is, in principle, no limitation to the construction of the moves. Specifically, unlike in MD, the step size (e.g. of displacements or rotations) can be chosen arbitrarily large. For small systems with only few degrees of freedom, such as the formation of small atomic clusters, MC methods are a promising approach [93]. However, the construction of correlated moves remains challenging.
After a more detailed description of the modeling of atomic interactions with force fields, both methods, MD and MC, will be presented in detail.

2.4.1. Modeling Atomic Interaction with Force-Fields

In molecular modeling, the interaction between atoms, that is based on the configuration electrons and cores of the atoms, is approximated by semi-empirical energy functions [32–35]. These force-fields depend on all ionic degrees of freedom of the system that is simulated

[1] Usually, the classical functions parametrizing the multi-dimensional potential energy surface of the nuclei are referred to as "force-fields" also if the forces are not evaluated and only the energy of the system is calculated, e.g. in MC-methods.

(e.g. all atom positions in the 3D space). They are characterized by numerous element- or compound-specific parameters. The parametrization can be performed on the basis of empirical values, i.e. by fitting the force-field functions to experimental results. Another approach for the generation of force-field parameters via *ab initio* QM calculations, such as DFT.
The optimal parametrization of a force-field is often a compromise between accuracy and simulation efficiency: the more approximations are made in the construction of a force-field, the faster the evaluation of the energy of a system. This allows not only the treatment of much larger systems, it also makes time-dependent propagation of a system affordable, where of the order of $\mathcal{O}(10^6)$ energy evaluations have to be performed. On the other hand, the approximations that are made omit the explicit treatment of quantum mechanical effects and limit the applicability of force-field based approaches to systems, where the change in the electronic structure throughout the process is negligibly small. Nevertheless, force-field based simulations are a promising approach to the simulation of molecular systems.
In terms of force-fields, the total energy of a system comprises three parts:

$$E_{\text{tot}} = E_{\text{intra}} + E_{\text{inter}} + E_{\text{extern}} \tag{2.4.1}$$

Here, E_{intra} is the internal interaction of a bonded unit, i.e. the interaction between atoms induced by covalent bonds. E_{inter} is the parametrization of the part of the non-bonded atom-atom interaction as a consequence of the interactions between cores and charge distributions. Additional potentials, E_{extern}, for modeling the influence of external quantities such as pressure, gravity or solvents (implicit treatment) can be applied in addition.
In the following, the different contributions to the total energy and their force-field functions will be presented individually.

Intramolecular Interaction: Bonded Force-Fields

In molecules, dislocated electrons form covalent bonds between single atoms. For the description of this internal contribution to the total energy, E_{intra}, the relative positions of the atoms of a covalently bonded unit are mapped onto the internal degrees of freedom: bond lengths, angles between bonds and dihedral angles. This is illustrated in Fig. 2.4.1.
The contributions of the individual parts to the internal energy are:

$$U_{bond}(r_{jk}) = \frac{1}{2}\, k_{jk}\, \left(r_{jk} - r_{jk}^0\right) \tag{2.4.2}$$

for each bond, where r_{jk}^0 is the equilibirum distance and k_{jk} the force constant,

$$U_{angle}(\alpha_{ijk}) = \frac{1}{2}\, k_{ijk}\, \left(\alpha_{ijk} - \alpha_{ijk}^0\right) \tag{2.4.3}$$

for each angle, with the force constant k_{ijk} and the equilibrium angle α_{ijk}^0 and

$$U_{dihedral}(\phi_k) = k_k\, (1 + \cos(n_k\, \phi_k - \delta)) \tag{2.4.4}$$

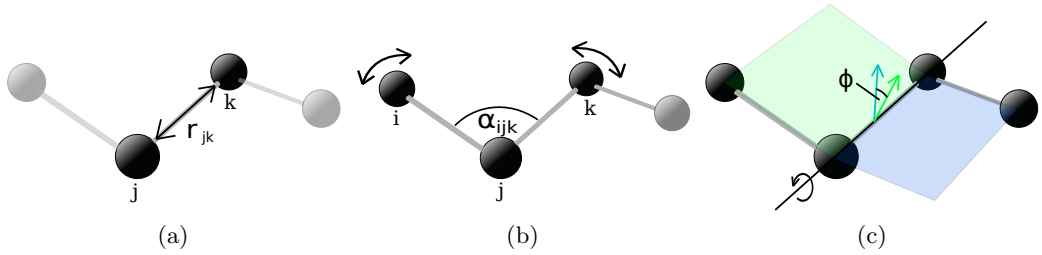


Figure 2.4.1.: Internal degrees of freedoms are used to calculate the contribution of covalent bonds to the total energy. (a) stretching of bonds, r_{jk} (b) the angles between two neighboring bonds α_{ijk} and (c) the relative orientation of the plains spanned by the positions of three neighboring atoms, ϕ_k.

for the dihedrals with the respective force constant k_k and two additional parameters n_k and δ (see below). It should be stated at this point, that there are different forms for the description of single-bond-dihedrals that are approximately equivalent. However, most of the established force-fields were derived and parametrized for a specific class of molecular systems, mostly proteins and might thus not be applicable for any system. The specific form as well as the parameters for dihedral potentials then need to be derived from quantum mechanical calculations for the systems explicitly.
The form of the potentials is chosen to represent the characteristics of the energy landscape that induced by the electronic structure: The equilibrium values of bonds and angles, r^0_{jk} and α^0_{ijk}, can be extracted from experimental data or QM calculations and set accordingly. The equilibrium angles for example are a result of the form of the atomic orbitals and lead to different values for each element. As the stretching of a bond and the bending of an angle (that is the deformation of electronic orbitals) requires large amounts of energy, only small changes are possible with the kinetic energy available at relevant temperatures ($T < 1000\,K$). In simulations that do not include chemical reactions (i.e. the breaking and reformation of covalent bonds), this harmonic approximation of bonded and angle potentials proved sufficient.
This is different for the dihedral angles: The cost to twist a molecular chain, e.g. perform rotations around single bonds, is energetically cheap compared to the distortion of bonds and angles. The approximation by a harmonic potential would be inaccurate and a periodic potential is needed to cover the potential landscape for a full rotation around $2\,\pi$, is given in Eq. 2.4.4. The parameter δ is chosen to generate minima of the potential at correct values of ϕ_k. The parameter n_k is called the multiplicity of the dihedral and is set to $n_k = 1, 2, 3, 4 \ldots$. It reproduces the appropriate number of minima upon rotation around $2\,\pi$ that are the result of the electronic the orbital structure.
By properly parametrizing the classical force-field functions, the electronic structure which that is the basis of the atomic interaction is taken into account properly.

Intermolecular Interaction: Non-Bonded Force Fields

For the interaction between the atoms of different molecules or single atoms, non-bonded force-fields are applied. Like the bonded force-fields, the potentials are chosen in a way to reproduce the realistic behaviour induced by the electronic structure. This covers the electrostatic interaction between the charge distributions of molecules, the induced dipole-dipole attraction (van-der-Waals-attraction) of atoms and the Pauli-repulsion between electrons in the outer shells of atoms.
The interaction between charge distributions can be modeled by mapping the electrostatic potential to point charges on the atom positions and the Lennard-Jones-potential is a good approximation for van-der-Waals attraction and Pauli-repulsion.

Coulomb electrostatics Upon the formation of chemical bonds, the local electron density of the atoms is shifted from one atom towards another. While the system in total remains neutral, an inhomogeneous charge distribution is induced. One approach for an efficient evaluation of the interaction between the charge distributions of molecules is the mapping of the charge distributions onto fractional charges, usually placed at atom centers ("partial charges"). The electrostatic contribution to the interaction between two molecules A and B is then modeled using the Coulomb interaction between the partial charges q_i:

$$U_{Coulomb} = \sum_{i}^{N_A} q_i^A \, \Phi_B(\boldsymbol{r}_i) = \frac{1}{4\,\pi\,\epsilon_r\,\epsilon_0} \sum_{i}^{N_A} \sum_{j}^{N_B} \frac{q_i^A \, q_j^B}{r_{ij}} \tag{2.4.5}$$

with the atom-atom distance r_{ij} and the electrostatic potential Φ_B:

$$\Phi_B(\boldsymbol{r}) = \frac{1}{4\,\pi\,\epsilon_r\,\epsilon_0} \sum_{j}^{N_B} \frac{q_j}{|\boldsymbol{r_j} - \boldsymbol{r}|} \tag{2.4.6}$$

There are multiple methods to derive partial charges. For a couple of systems that are widely studied, such as proteins in aqueous solutions, partial charges were constructed semi-empirically and are tabulated in special force fields [94–96]. Another more generic approach is the electrostatic potential fit (ESP-fit) [97, 98] fractional charges on the atom sites are constructed in a way that (a) the electrostatic potential of the electronic charge distribution is reproduced with minimal deviation on one or multiple surfaces around the molecule and that (b) the total charge of the system is preserved.
In praxis, the potential on the surface(s) around the molecule is calculated using quantum mechanical methods (e.g. DTF, Hartree-Fock, see sec. 2.3). Partial charges are then placed on the atom positions and varied until the deviation between the QM-potential and potential induced by the point charges is minimal. The ESP-method is implemented in various QM-packages such as MOPAC and TURBOMOLE [70, 77, 78].
Other methods for the calculation of the electrostatic interaction between molecules, such as the multipole-expansion of single-atom charge distributions, are computationally more extensive without increasing the accuracy significantly.

The main draw back of using fixed (static) point charges or static multipoles for the calculation of the electrostatic interaction is the neglect of polarization effects. The charge distribution of a molecule in the vicinity of other molecules differs from the charge distribution of the same molecule in vacuum. Partial charges, originally calculated for the molecule in vacuum, would need to be re-calculated on the fly during the simulation. While this is done in form of a combination of quantum mechanics and molecular dynamics [99–101], the application of such algorithms is limited in system size and not widely used. Different approaches, such as polarizable force-fields, have been established but are not readily applicable to any system. Classical simulations throughout this thesis were performed using static vacuum ESP-charges, if not stated otherwise.

Van-der-Waals-attraction and Pauli-repulsion: The Lennard-Jones-potential In addition to the electrostatic interaction, an additional force-field is needed to correctly reproduce the interaction between atoms. This additional interaction term is motivated by two physical observations: Firstly, neutral atoms of a noble gas, with a completely isotropic charge distribution, must have an attractive interaction. Otherwise, they would not condensate below a critical temperature. Secondly, without an additional energy term, atoms could come arbitrarily close to each other. This would lead to infinite energy for atoms that are not neutral. This is, of course, not physical.
The attractive interaction described above is the van-der-Waals interaction: fluctuations in the electron distribution of an atom, even a neutral atom, induce a temporary dipole $\boldsymbol{\mu}_1$. A second atom that is in the vicinity of the first atom is then polarized by the dipole field $\boldsymbol{D} = -\frac{1}{4\pi\epsilon_0}\frac{\boldsymbol{\mu}_1}{r^3}$, resulting in an induced mean-field dipole:

$$\mu_2 = c'\boldsymbol{D} = -c'\frac{1}{4\pi\epsilon_0}\frac{\boldsymbol{\mu}_1}{r^3} \tag{2.4.7}$$

with a distance of r between the atoms between atom 1 and 2. The interaction between the two dipoles is then

$$U_{\mathrm{VdW}}(r) = -\boldsymbol{D}\boldsymbol{\mu_2} = -c\frac{1}{r^6} \equiv -\left(\frac{\sigma}{r}\right)^6 \tag{2.4.8}$$

The repulsive potential is a result of spin statistics: According to the Pauli-principle, there cannot be two electrons in the same state, namely with the same energy and the same spin. If two fully occupied orbitals of two atoms with the same energy are brought close to each other, one electron has to be transferred into a higher orbital. This requires energy and results in a repulsive potential. As the thermodynamic occupation probability of higher orbitals decays exponentially with the difference in orbital energies, a reasonable form of the potential would be an exponential repulsion. In many cases, a different form is used:

$$U_{\mathrm{Pauli}}(r) = \left(\frac{\sigma}{r}\right)^{12} \tag{2.4.9}$$

This potential as well as an exponential repulsion increases very strongly for small distances. States with very small distances, where the difference between the exponential term and

Eq. 2.4.9 is significantly non-zero, are not occupied in thermodynamic equilibrium and the approximation of the exponential repulsion by an $1/r^{12}$-term is justified. Hence, a proper parametrization of Eqs. 2.4.8 and 2.4.9 is a good approximation to the physical Van-der-Waals attraction and Pauli-repulsion. The combination of the two equations is called the Lennard-Jones-, or 12-6-potential:

$$U_{\mathrm{LJ}}(r) = 4\epsilon \left\{ \left(\frac{\sigma}{r}\right)^{12} - 2\left(\frac{\sigma}{r}\right)^{6} \right\} \tag{2.4.10}$$

The form of the repulsive part in this potential is chosen for computational efficiency and the parameter σ can be extracted semi-empirically. Other potentials, e.g. Eq. 2.4.10 with different exponents or the Buckingham-potential

$$U_{\mathrm{Buckingham}}(r) = \epsilon \left\{ \exp\left(-\frac{r}{\rho}\right) - \left(\frac{\sigma}{r}\right)^{6} \right\} \tag{2.4.11}$$

are also used. The Lennard-Jones-potential is, however, computationally most efficient and widely used in the simulation of molecular structures.
The parameters ϵ and σ are usually fitted to experimental values and are tabulated for a variety of elements and their ions. Ideally the parameters should be fitted to atoms or ions of each pair of elements. Another, more efficient approach is to calculate the parameters for the interaction between to constituents of the same element and then use mixing rules to calculate the parameters for heterogeneous interaction between atoms of different types A and B:

$$\sigma_{AB} = \frac{1}{2}(\sigma_{AA} + \sigma_{BB}) \qquad \text{or} \qquad \sigma_{AB} = \sqrt{\sigma_{AA}\sigma_{BB}} \tag{2.4.12}$$

$$\epsilon_{AB} = \sqrt{\epsilon_{AA}\epsilon_{BB}} \tag{2.4.13}$$

This induces an additional error in the approximation of the full-quantum mechanical system by classical force-fields. However, to date, this approximation, including the mixing rules of the Lennard-Jones parameters, proved to be sufficient for molecular mechanics simulations. The parameters for the homogeneous interaction which are derived semi-empirically or by *ab initio* QM calculations for multiple applications are stored in large look-up tables that are accessed during the simulation by most of the widely used packages [94–96].

2.4.2. Molecular Dynamics

Integrating Newton's Equation of Motion

A straight forward solution of the dynamics of a molecular (or atomistic) system is called molecular dynamics. MD is based on the integration of Newton's equation of motion for each atom i:

$$\boldsymbol{F}_i = m_i\, \boldsymbol{a}_i \tag{2.4.14}$$

The force $\boldsymbol{F}_i$ acting on each particle is the negative gradient of the force-field potentials presented in the previous section. The force on each atom is calculated as the sum over

the internal forces, e.g. the bonds to the neighboring atoms in a molecule and the forces induced by angle and dihedral potential, the intermolecular forces, e.g. Lennard-Jones- and Coulomb-gradients, and external forces. In general, for a system consisting of N atoms, this leads to a set of $3N$ coupled differential equations. The analytic solution of such a set for a typical system with $N = \mathcal{O}(10^4)$ particles is not possible for obvious reasons. Instead, the equations of motion are solved discretely via step-by-step integration, as described in the following.
The Taylor-expansions of the trajectory of a single particle i is and its velocity are

$$\boldsymbol{r}_i(t+\delta t) = \boldsymbol{r}_i(t) + \delta t\, \boldsymbol{v}_i(t) + \frac{1}{2}\delta t^2\, \boldsymbol{a}_i(t) \tag{2.4.15}$$

$$\boldsymbol{v}_i(t+\delta t) = \boldsymbol{v}_i(t) + \delta t\, \boldsymbol{a}_i(t) \tag{2.4.16}$$

$$\boldsymbol{a}_i(t) = \frac{1}{m}\boldsymbol{F}_i(t) \tag{2.4.17}$$

Numerical step-by-step integration then yields the time resolved trajectory of the atomistic system: For a given initial configuration of atom positions $\{\boldsymbol{r}_i\}$, an initial random velocity vector $\{\boldsymbol{v}_i\}$ is generated and the forces for all atoms are calculated using the force-fields. After each time step, velocities and positions are adapted according to Eqs. 2.4.15 and 2.4.16 and the new accelerations calculated using Eq. 2.4.17. This results in the physical trajectory of the system. Thermodynamic expectation values of quantities, such as energy, pressure, volume or temperature are extracted by averaging over the simulated time interval. The time step δt is the critical parameter in each MD simulation. The smaller δt, the more steps need to be performed in order to reach a certain simulation time and propagation of the system. In each step, $\mathcal{O}(N^2)$ atom-atom interactions have to be calculated to generate the updated forces and energies. This leads to very large simulation times for arbitrarily small time steps. With large values for δt, on the other hand, the numerical step-by-step intergration in Eqs. 2.4.15 and 2.4.17 is a bad approximation, leading to unrealistic trajectories: Clashes of atom, breaking of bonds or strong deformation of molecules is usually the result of too large time step. Therefore, the time step used in MD simulations has been hovering on the femtosecond-scale for years, although efforts to increase the time step have been made [102, 103].

Coupling to Temperature and Pressure: Thermostats and Barostats

For the simulation of a system in a thermodynamic ensemble, e.g. the NPT or NVT ensemble, the movement of the particles has to be coupled to the macroscopic quantities, pressure or volume and temperature. In thermodynamic equilibrium, temperature and particle momentum are connected via

$$T = \frac{1}{3\,N\,k_\mathrm{T}} \sum_{i=1}^{N} \frac{|\boldsymbol{p}_i|^2}{m_i} \tag{2.4.18}$$

The pressure is related to particle momentum and particle-particle-forces via

$$P = \frac{1}{V}\left[\frac{1}{3}\sum_{i=1}^{N}\left(\sum_{j>i}^{N} F_{ij}|\boldsymbol{x}_i - \boldsymbol{x}_j| + \frac{|\boldsymbol{p}_i|}{m_i}\right)\right] \tag{2.4.19}$$

Barostats and thermostats are using the relations of Eqs. 2.4.18 and 2.4.19 to either rescale particle velocities after a certain simulation time in order to adjust pressure and temperature, or to add an extra term to Eq. 2.4.16 that adapts the momentum accordingly. Multiple barostats and thermostats are available [104, 105].

2.4.3. Metropolis Monte Carlo

Detailed Balance and Move Generation

An alternative approach to compute thermodynamic expectation values of the system is the sampling of the phase space by applying random changes to the system. This random sampling is a special form of an MC protocol, a technique that was established in the 1950's in line with the rise of computational resources. In the most general terms, Monte Carlo methods use random numbers to approach problems that are not solvable in an analytic fashion. One simple example is the calculation of definite integrals. Monte Carlo methods are widely used also in the fields of Mathematics and Physics, but also in other fields such as Economy [40, 106].
In the most widely used Monte Carlo scheme in molecular simulations, the Metropolis Monte Carlo approach, a Markov chain forms the basis of the algorithm: A sequence of simulation steps is generated, where the configuration at each step n, $\{\boldsymbol{x}_i\}_n$, depends exclusively on the system configuration at the prior step $n-1$:

$$\{\boldsymbol{x}_i\}_0 \rightarrow \{\boldsymbol{x}_i\}_1 \rightarrow \ldots \rightarrow \{\boldsymbol{x}_i\}_{n-1} \rightarrow \{\boldsymbol{x}_i\}_n \rightarrow \ldots \tag{2.4.20}$$

In order to calculate the system in a thermodynamic equilibrium, the Markov chain has to be constructed in a way, that, after an infinite number of simulation steps, $n_{max} \rightarrow \infty$, the occupation probabilities of all states $\{\boldsymbol{x}_i\}$, $P_{n_{max}}(\{\boldsymbol{x}_i\})$, reproduce a Boltzmann-distribution:

$$\lim_{n_{max}\rightarrow\infty} P_{n_{max}}(\{\boldsymbol{x}_i\}) = \frac{1}{Z}\exp\left(-\beta\, E(\{\boldsymbol{x}_i\})\right) \tag{2.4.21}$$

with $\beta = \frac{1}{k_\mathrm{B}T}$ and the partition function $Z = \sum_{\{\boldsymbol{x}_i\}} \exp\left(-\beta\, E(\{\boldsymbol{x}_i\})\right.$. This is achieved as follows:
In a system in thermodynamic equilibrium, the total rate of states moving from any state q into another state q', equals the rate of the inverse move, $q' \rightarrow q$. q is the set of coordinates describing the state of the system, e.g. the position vectors of all atoms, $q = \{\boldsymbol{x}_i\}$. For a system following the Boltzmann-distribution, the transition rate Γ from a state q into a state q' is the product of the transition probability W and the occupation probability of state q:

$$\Gamma(q \rightarrow q') = W(q \rightarrow q')\,\frac{1}{Z}\exp\left(-\beta\, E(q)\right) \tag{2.4.22}$$

With this, the condition of equal rates of transition and inverse transition, $\Gamma(q \to q') = \Gamma(q' \to q)$, results in

$$\frac{W(q \to q')}{W(q' \to q)} = \exp\left(-\beta \Delta E\right) \tag{2.4.23}$$

with the energy difference between states q and q', $\Delta E = E(q') - E(q)$. This is called the detailed balance criterion. The probability to perform a move, $W(q \to q')$, is the product of the probabilities π and ρ to propose and accept a move, respectively:

$$W(q \to q') = \pi(q \to q')\, \rho(q \to q') \tag{2.4.24}$$

In the most widely used Metropolis Monte Carlo, also called Generic Monte Carlo (GMC), the changes from one state into another are generated completely randomly and the move $q \to q'$ is constructed with the same probability as the reverse move, $q' \to q$, leading to $\pi(q \to q')= \pi(q' \to q)$. Any proposed move then need to be accepted or rejected with a probability $\rho(q \to q')$ that satisfies Eq. 2.4.23:

$$\frac{\rho(q \to q')}{\rho(q' \to q)} = \exp\left(-\beta \Delta E\right) \tag{2.4.25}$$

where $\Delta E = E(q') - E(q)$ is the change in energy induced by the move.
A possible choice of ρ that satisfies this relation is the Metropolis acceptance criterion [40]:

$$\rho(q \to q') = \begin{cases} \exp\left(-\beta \Delta E\right), & \Delta E > 0 \\ 1, & \Delta E \leq 0 \end{cases} \tag{2.4.26}$$

In practice, the Metropolis algorithm works as follows. Based on pseudo-random numbers, a change to the system, a so called "move" is applied, leading to a change in the coordinates: $q \to q'$. The energy difference between the states before and after the move, $\Delta E = E(q') - E(q)$, is calculated. The move is then accepted with the probability

$$\rho(q \to q') = \min\left\{1, \exp\left(-\beta \Delta E\right)\right\} \tag{2.4.27}$$

Advantages and Limitations of Metropolis Monte Carlo

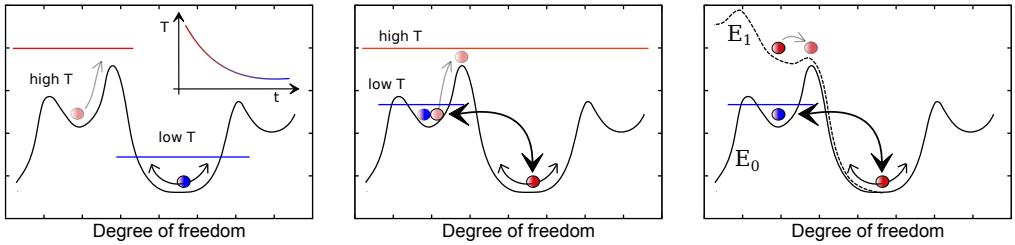
So far, the detailed construction of the single moves in a Metropolis Monte Carlo simulation was not discussed. In fact, there is no limitation to the changes that can be applied to the system in a single MC step. The change of any coordinate (such as atom positions, angles, bonds or dihedral angles) or even multiple coordinates can be applied. Therein lies the main advantage of MC over MD: Whereas in MD the propagation of the system is confined to the physical trajectory and significant changes are limited by a very small time step, arbitrarily large variations in the coordinates can be applied to the system in MC based protocols. Besides a possibly faster diffusion through the phase space, this induces significant advantages in systems with a rugged energy landscape: Two specific states q_0 and q_1, that are locally minimal in energy, are separated by an energy barrier, ΔE_b. If ΔE_b is larger than the intrinsic energy of the system, $\Delta E_b > k_\mathrm{B}\, T$, MD is very unlikely to propagate from q

to q'. The probability to perform the transition between q and q' decays exponentially with the barrier height. This holds also true for MC based methods with limited displacements in the coordinate separating the two states. If steps sizes in MC are on the other hand chosen to be of the order of $|q_1 - q_0|$, a direct transition between the minima is possible. This transition probability depends only on the energy difference $\Delta E = E(q_1) - E(q_0)$, not the energy barrier of the intermediate states.
There is, however, a drawback to the Metropolis-MC based methods: For systems with many degrees of freedom, such as a condensed layer of small organic molecules or a large protein, the change of a single coordinate in each MC step, however large, will not lead to an efficient sampling of the phase space. Furthermore, there are collective effects, such as the crystallization of molecules. Moving single coordinates at a time will fail to reproduce the collective behavior. As mentioned, there are no limitations to the construction of moves. While it is, in principle, possible to change an arbitrary large number of degrees of freedom in a single MC step, this primitive construction of collective movement is bound to fail: The probability to construct a move that leads to a strong increase in energy, $\Delta E > kT$, increases with the number m of coordinates that are changed simultaneously in a single move. However, the number of possible changes to the system leading to a strong decrease in energy, $\Delta E < k_\mathrm{B}\,T$, is limited in condensed systems and the higher probability to construct a bad move is not compensated. Roughly speaking, the acceptance probability of a move affecting m coordinates q_i is the product of the acceptance probabilities of all single changes to the system. In condensed systems, the acceptance rate for moves affecting of the order of $\mathcal{O}(1)$ coordinates lies typically in the range of $0.3 - 0.7$ and converges rapidly to 0 for increased m.
There are approaches to exploit the advantage of MC, the large step sizes, with acceptance rates significantly larger than 0. One possibility is the reduction of the number of degrees of freedom [57]. Another approach is the advanced construction of correlated moves, that will be discussed in detail in chapter 4.

2.4.4. Sophisticated Modeling Methods

Impeded Modeling in Rugged Energy Landscapes

Dense atomic systems, such as layers of small organic molecules or folded protein structures, are often characterized by very rugged energy landscape. Local minima in the energy are separated by energy barriers of various heights. Any barrier with $\Delta E_b \geq k_\mathrm{B}\,T$ is limiting propagation of the system across the energy landscape. There are several approaches to increase diffusion across energy barriers with $\Delta E_b > k_\mathrm{B}\,T$ in the multi-dimensional configuration phase, preventing trapping of the system in metastable configurations for long simulation times. Three methods will be presented in the following: basin hopping by simulated annealing (SA) [107], parallel tempering (PT) [45, 108–110] and Model Hopping (MH) [111].



(a) Simulated Annealing (SA) (b) Parallel Tempering (PT) (c) Model Hopping (MH)

Figure 2.4.2.: Increasing sampling efficiency in rugged energy landscapes. (a) In the simulated annealing (SA) protocol, the temperature is gradually reduced from a high temperature to the physical temperature. The hot system can cross energy barriers before equilibrating at the final temperature, and repeating multiple SA cycles leads to an efficient basing hopping between local minima. (b) In parallel tempering method (PT), also called replica exchange method, multiple replica of the system are simulated in parallel at different temperatures. Exchange between the models with an acceptance criterion asserting detailed balance allows increased propagation of the system simulated with the physical temperature. (c) In model hopping approach (MH), several replica are simulated using different energy models. At least one replica of the system with an energy model E_1 is biased towards a subset of the configuration space where the energy is estimated to be minimal. Exchange between the configurations of the biased energy model E_1 and the physical force field E_0 is performed deploying an acceptance criterion similar to PT in order to asserts detailed balance.

Simulated Annealing (SA)

A straight forward solution to the trapping of the system in metastable configurations is to increase the temperature to a high value of T_h, where $\Delta E_b \approx k_\mathrm{B}\, T_h$ for all energy barriers. In order to reproduce configurations that are equilibrated at any desired lower temperature, T_l, the temperature is gradually decreased throughout a simulation consisting of N simulations steps, resulting in a temperature at step n:

$$T_n = c^n\, T_h, \quad c = \left(\frac{T_l}{T_h}\right)^{1/N} \tag{2.4.28}$$

This simulated annealing protocol [107] is illustrated in Fig. 2.4.2a. Repetition of the SA protocol results in an efficient basing hopping [107] between local minima, enabling an fast system propagation and an efficient search for the global optimum in the energy landscape.

Parallel Tempering

Swendsen and Wang tackled the impeded sampling in rugged energy landscapes in 1986 by modeling multiple systems in parallel at different temperatures T_i leading to different values for $\beta_i = 1/k_\mathrm{B}\, T_i$. Originally developed for Monte Carlo simulations of spin glasses [108], the method was further developed [112] and transferred to MD simulations [45, 109, 110]. The main idea is illustrated in Fig. 2.4.2b and works as follows: A system of interest is modeled at multiple temperatures T_i, β_i simultaneously. After a certain time interval Δt the configurations C_i (or the temperatures T_i, respectively) between the models are exchanged with the probability:

$$\rho_{exchange} = \min\left\{1, \frac{\exp\left(-\beta_i E_j - \beta_j E_i\right)}{\exp\left(-\beta_i E_i - \beta_j E_j\right)}\right\} \tag{2.4.29}$$

The acceptance criterion can be derived by thermodynamic considerations originating in the occupation probabilities: In the thermodynamic Equilibrium, the configurations C_i with energies E_i are distributed according the Boltzmann-distribution:

$$\pi_{(i,i)} = \frac{1}{Z}\exp\left(-\beta_i\, E_i\right) \tag{2.4.30}$$

with the partition function $Z = \sum_q |\exp(-\beta E(q))|$. The probability to find configuration C_i in a state with the energy E_i at β_i and, at the same time, configuration C_j in a state with the energy E_j at β_j is then just the product of the individual occupation probabilities:

$$\pi_{(i,i),(j,j)} = \pi_{i,i} \times \pi_{j,j} = \frac{1}{Z}\exp\left(-\beta_i\, E_i - \beta_j\, E_j\right) \tag{2.4.31}$$

Another possible outcome of the same run would be to find exchanged configurations: configuration C_j with energy E_j at β_i and configuration C_i with energy E_i at β_j. The probability

for this simulation result is given by

$$\pi_{(i,j),(j,i)} = \pi_{i,j} \times \pi_{j,i} = \frac{1}{Z} \exp\left(-\beta_i E_j - \beta_j E_i\right) \tag{2.4.32}$$

If, after the time interval Δt, the configurations C_i in β_i and C_j in β_j are realized in the simulations (which happens with a probability according to Eq. 2.4.31) and the configurations are exchanged manually with a probability given by Eq. 2.4.29, the exchanged state occurs the with the total probability:

$$\hat{\pi}_{(i,j),(j,i)} = \pi_{(i,i),(j,j)} \times \rho_{exchange} \tag{2.4.33}$$

$$= \frac{1}{Z} \exp\left(-\beta_i E_i - \beta_j E_j\right) \times \frac{\exp\left(-\beta_i E_j - \beta_j E_i\right)}{\exp\left(-\beta_i E_i - \beta_j E_j\right)} \tag{2.4.34}$$

$$= \frac{1}{Z} \exp\left(-\beta_i E_j - \beta_j E_i\right) \tag{2.4.35}$$

$$= \pi_{(i,j),(j,i)} \tag{2.4.36}$$

Hence the exchange of the configurations using the probability of Eq. 2.4.29 results in a distribution of states similar to the thermodynamically equilibrated distribution $\pi_{(i,j),(j,i)}$.
In practice, more than two models are used for efficient sampling: On the one hand, the temperature of the hottest of the N systems needs to be chosen such that the internal energy is of the order of the largest energy barrier, $(k_\mathrm{B}\,T)_N \approx \Delta E_{max}$. On the other hand, the exchange probability between two models is only significantly non-zero if there is a overlap of the density of states of two models. This is only the case, if the difference in the β_i of "neighboring" systems i and $i+1$ are sufficiently small. The total temperature range between the highest temperature T_N and the temperature of interest, T_0, then needs to be bridged by several intermediate temperatures. The configuration of the hottest system can then be passed to the system at T_0 in several exchange attempts.
The PT algorithm is implemented in several MD simulation packages (e.g. LAMMPS, Gromacs) and was amongst other applications, applied successfully to polymers, proteins and solid state applications [113].

Model Hopping

In other cases efficient sampling is prevented by the fact that only a small fraction of the configuration space has significant thermodynamic occupation probability. In other terms, only this small part of the phase space is "relevant" for the measurement of the system's properties in thermodynamic equilibrium. In this case, information on the system can be used to make an educated guess about this region in the configuration space and bias the system towards it.
In PT, models are simulated at unphysical temperatures, and the exchange of configurations into a model at the physical temperature results in thermodynamically equilibrated structures. On the basis of the PT approach, Hansmann et. al proposed a protocol where multiple models are simulated using different energy models [45].

Exchange between the models with configurations C_i and energies E_i is proposed and accepted or rejected after a time interval Δt, or, in MC, after certain number of simulation steps. Simulation of one model with E_0 being the physical force-field, i.e. an unbiased model, asserts physical, thermodynamically equilibrated results. By applying the correct acceptance criterion, configurations that are exchanged into the energy model E_0 can be assumed to be physical. This is illustrated in Fig. 2.4.2c.
The derivation of the exchange acceptance probability follows the same Ansatz as in PT: Let C_i be the configuration of model i with the energy function E_i at a certain point in the simulation. The energy in this model is then determined by the configuration, $E_i(C_i)$. The probability to find two configurations C_i and C_j in the models using the energy functions E_i and E_j respectively is given by the product of the individual occupation probabilities:

$$\begin{aligned}\pi_{(i,i),(j,j)} =& \pi_{i,i} \times \pi_{j,j} && (2.4.37)\\ =& \frac{1}{Z} \exp\left(-\beta E_i(C_i)\right) \exp\left(-\beta E_j(C_j)\right) && (2.4.38)\\ =& \frac{1}{Z} \exp\left(-\beta (E_i(C_i) + E_j(C_j)\right) && (2.4.39)\end{aligned}$$

Analogously to the PT algorithm, the exchange acceptance probability between models i and j can be derived by considering a different possible outcome of the same simulation: configuration C_j in the model with the energy function E_i and the configuration C_i in the model with the energy function E_j. The thermodynamic occupation probability of finding the exchanged configurations is given by

$$\pi_{(j,i),(i,j)} = \frac{1}{Z} \exp\left(-\beta (E_i(C_j) + E_j(C_i)\right) \tag{2.4.40}$$

As the the occupation probability of configuration C_i in model i and configuration C_j in model j is given by Eq. 2.4.39, accepting exchange of the configurations between the models with the exchange probability of

$$\rho_{exchange} = \min\left\{1, \frac{\exp\left(-\beta\left(E_i(C_j) + E_j(C_i)\right)\right)}{\exp\left(-\beta\left(E_i(C_i) + E_j(C_j)\right)\right)}\right\} \tag{2.4.41}$$

results in the correct occupation probability of the exchanged states as given in Eq. 4.5.1.
The Model Hopping algorithm works as follows: Starting from the same initial configuration, multiple models with different energy functions are simulated. After each n steps, the simulation is paused and exchange of the configurations between models is proposed and accepted according Eq. 4.5.1. A reasonable approach is to generate the unphysical energy models E_i, $i \neq 0$ by adding a bias potential to the physical energy model and scaling the bias differently in each of the biased models. By proposing exchange between models i and $i+1$ after n steps, configurations are successively exchanged from the model with the largest bias into the unbiased model and can be deployed for further analysis.
Crucial to the successful exchange between biased and unbiased model is the construction of the biased energy functions E_i, $i \neq 0$. A detailed analysis of the algorithm and possible

ways to construct energy models E_i in 4.5 suggests that exchange between models is limited, if the biased models lead to configurations C_j with bad physical energy, i.e. $E_0(C_j) > k_\mathrm{B}\,T$.

2.5. Electronic Transport in Organic Materials

2.5.1. From Band Transport to Charge Hopping

The generation of a sample morphology of organic layers was discussed in the previous section. The subsequent and final step in the multiscale modeling approach is the transfer of properties of small entities such as atoms or molecules, calculated on the scale of single electrons, to macroscopic electronic transport properties of a full organic electronic layer or device.
In crystalline metals and semiconductors, electrons are delocalized, i.e. the single electron wave function is non-zero across large areas of the device. Transport of delocalized charge carriers such periodic systems can be described in terms of valence and conduction bands. This does not hold true in organic semiconductors: No conduction bands and valence bands are formed, on the basis of which the analytical calculation of electronic transport properties is possible. Fast decay of the electronic wave functions of the highest occupied molecular orbitals (HOMOs) or lowest unoccupied molecular orbitals (LUMOs) in amorphous organic materials leads to strong localization of the charge carriers on single molecules. The overlap of the wave functions of a single electron located on first one and then a neighboring molecule is usually quite small, inducing weak electronic coupling. For many molecular structures, this property results in low charge carrier mobilities. However, conjugation, e.g. the formation of π-orbitals, induces a delocalization of the HOMO and LUMO orbitals, at least across large parts of a molecule. This intramolecular delocalization results in increased transition rates throughout the bulk, inducing significant charge carrier mobilities.
A theory for the description of transfer in the general case of localized charge carriers was originally proposed by Marcus in 1956 [114]. Due to the large impact on the research of charge transfer processes in various fields, the Nobel Prize was awarded to Marcus in 1992. For the charge transfer between molecules, the semi-classical Marcus theory yields hopping rates between neighboring sites as a function of orbital overlap, the difference in the (free) energy between initial and final states and the reorganization energy.
Once the hopping rates between molecules are calculated, the macroscopic bulk properties need to be derived. Analytic approaches to this problem are available in literature [27, 115], their general applicability and parametrization, however, is subject to current discussion. A more generic approach that circumvents the usage of additional parameters is the numerical calculation of macroscopic properties: Using KMC, the time dependent trajectory of charge carriers through the system is generated. For the simulation of the hop of a single charge that is localized on a given initial site (i.e. the molecule where the charge is located at before the hop), all possible final configurations (i.e. the neighboring molecules where the charge may be localized after the hop) are considered. The final state as well as the physical time of the hopping process are drawn from a distribution on the basis of previously calculated hopping rates. Amongst other properties, charge carrier mobilities can then be extracted from the physical trajectories.

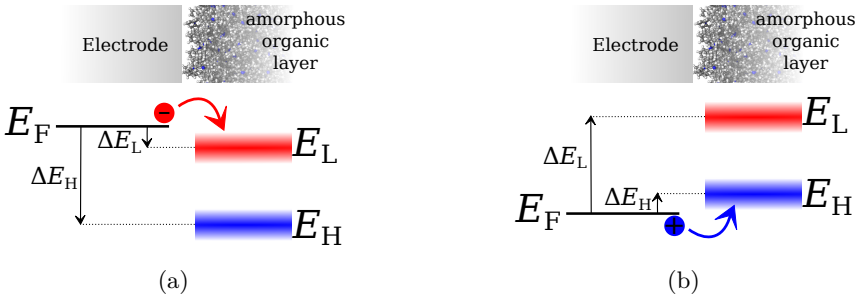


Figure 2.5.1.: The relative alignment of HOMO and LUMO energy levels, E_H and E_L, with the Fermi energy E_F, determine the charge carrier type: (a) If the energy level of the LUMO of the molecules is lower than the Fermi level, or the levels are at least closer than the thermal excitation, $|\Delta E_L| < k_B T$, electrons are injected into the system. Electrons are the preferred charge carrier, as long as $|\Delta E_H| >> k_B T$. (b) If the Fermi level is below the HOMO of the molecules, or, as for electrons, the difference between HOMO and Fermi level is within the thermal excitation, $|\Delta E_H| < k_B T$, the material is primarily hole conducting.

2.5.2. Marcus Theory of Hopping Transport

Molecular Energy Levels, Electron- and Hole-Conductors

Basis for the calculation of the charge transfer between molecules is the detailed description of the electronic structure of the molecules. The molecular orbitals can be calculated with different methods, e.g. DFT, Hartree-Fock or MNDO (see section 2.3). The orbitals that contribute to the transfer of an additional charge from one molecule to a neighboring molecule are the HOMO and the LUMO. Depending on the energy levels of electrons occupying these orbitals, the material can be considered electron- or hole-conducting: If the level of the LUMO is lower than (or at least close to) the Fermi-level of the injection material (i.e. the work function of the electrode), electrons are injected into the material, occupying LUMO-orbitals. The material is electron-conducting. If, on the other hand, the HOMO level is above (or close to) the Fermi-level, electrons are transferred from the HOMO of the molecules into the electrode. In other words, a hole is injected and the material is a hole-conductor. This is illustrated in Fig. 2.5.1. While the energy levels of HOMO and LUMO are well defined for a single molecule in vacuum, the charge distributions of surrounding molecules induces a shift in the orbital levels. Throughout the system, the energy levels of HOMO and LUMO have a smeared distribution rather than discrete, fixed values.

Eigenstates of Coupled Molecules

The transfer (or "hop") of a charge carrier from molecule A to a neighboring molecule B can be formulated in terms of a kinetic reaction. The reaction equations are

$$M_A^- + M_B \rightleftharpoons M_A + M_B^- \tag{2.5.1}$$

for electrons and

$$M_A^+ + M_B \rightleftharpoons M_A + M_B^+ \tag{2.5.2}$$

for holes. Independent of the charge carrier type, the initial (charge carrier localized on molecule A) and final state (charge carrier localized on molecule B) of the hopping process can be described using two wave functions, ϕ_i and ϕ_f. For a Hamiltonian of uncoupled molecules, the eigenvalue equations are

$$H_e^0 \left|\phi_i\right\rangle = \epsilon_i \left|\phi_i\right\rangle \tag{2.5.3}$$

$$H_e^0 \left|\phi_f\right\rangle = \epsilon_f \left|\phi_f\right\rangle \tag{2.5.4}$$

H_e^0 is the electronic Hamiltonian of the dimer system. If the molecules are coupled, however, the eigenstates of the coupled Hamiltonian in first order perturbation theory is a superposition of $|\phi_i\rangle$ and $|\phi_f\rangle$,

$$|\phi\rangle = c_i \left|\phi_i\right\rangle + c_f \left|\phi_f\right\rangle \tag{2.5.5}$$

and the Hamiltonian can be expressed as:

$$H_e = \epsilon_i \left|\phi_i\right\rangle \left\langle\phi_i\right| + \epsilon_f \left|\phi_f\right\rangle \left\langle\phi_f\right| + J_{if} \left(\left|\phi_f\right\rangle \left\langle\phi_i\right| + \left|\phi_i\right\rangle \left\langle\phi_f\right|\right) \tag{2.5.6}$$

In the limits where the charge carriers are completely localized, i.e. $c_i = 0$ or $c_f = 0$, the Schroedinger equation reproduces Eqs. 2.5.3 and 2.5.4. The mixing of the initial and final state by c_i and c_f is directly coupled to the overlap matrix element, J_{if}, and the energy difference $\Delta E = \epsilon_f - \epsilon_i$.
The Schroedinger equation of two coupled states can be written as

$$H_e \boldsymbol{c} = \epsilon \boldsymbol{c} \tag{2.5.7}$$

$$\begin{pmatrix} \epsilon_i & J_{if} \\ J_{if} & \epsilon_f \end{pmatrix} \begin{pmatrix} c_i \\ c_f \end{pmatrix} = \epsilon \begin{pmatrix} c_i \\ c_f \end{pmatrix} \tag{2.5.8}$$

This yields eigenvalues for the Hamiltonian of

$$\epsilon^{\pm} = \frac{1}{2}(\epsilon_i + \epsilon_f) \pm \frac{1}{2}\left((\epsilon_i - \epsilon_f)^2 + 4\,J_{ij}^2\right)^{1/2} \tag{2.5.9}$$

and eigenstates are determined by

$$(\epsilon_i - \epsilon^{\pm})\, c_i^{+-} + J_{if}\, c_f^{\pm} = 0 \tag{2.5.10}$$

$$J_{if}\, c_i^{\pm} + (\epsilon_f - \epsilon^{\pm})\, c_f^{\pm} = 0 \tag{2.5.11}$$

This relation underlines that the delocalization, namely the ratio of $c_i^{\pm}$ and $c_f^{\pm}$, is directly determined by J_{if} and ΔE.

Marcus Theory of Hopping

Using a semi-classical formulation, Marcus derived the hopping rate between two molecules with initial and final states ϕ_i and ϕ_f:

$$\Gamma_{if} = \frac{2\,\pi\,|J_{if}|^2}{\hbar}\,(4\,\pi\,\lambda\,k_\mathrm{B}\,T)^{-1/2}\exp\left(-\frac{(\Delta E+\lambda)^2}{4\lambda k_\mathrm{B}\,T}\right) \tag{2.5.12}$$

The derivation will be sketched in the following, starting from Fermi's Golden Rule:

$$\Gamma_{if} = \frac{2\,\pi}{\hbar}|\,\langle\Psi_i|\,H\,|\Psi_f\rangle\,|^2\,\delta(E_i - E_f) \tag{2.5.13}$$

H is the Hamiltonian of the full system, and Ψ_i and Ψ_f are initial and final states with energies E_i and E_f respectively.
According to the Born-Oppenheimer approximation, the total wave functions Ψ_i and Ψ_f factorize into electronic and ionic wave functions:

$$\Psi_\alpha = \phi_\alpha\,\chi_\alpha \tag{2.5.14}$$

with $\alpha = i,\ f$. Inserting the Hamiltonian of Eq. 2.5.6 results in

$$\Gamma_{if} = \frac{2\,\pi}{\hbar}|J_{if}\,\langle\chi_i|\chi_f\rangle\,|^2\delta(E_i - E_f) \tag{2.5.15}$$

The calculation of the overlap of the nuclear wave functions, $\langle\chi_i|\chi_f\rangle$, is following a semi-classical approach:
Putting an additional charge on a molecule is changing the molecular geometry. Bond lengths vary slightly between a neutral and charged molecule. The charge transfer from initial to final state needs to pass a transition state defined by electronic and nuclei configurations. Let $C = \{\boldsymbol{R}_k\}$ denote the state describing the positions of all nuclei of both molecules. Then, upon transferring the charge from one molecule to the other, the nuclei configuration is changing from an initial state C_i to a final state C_f. At the point of the charge transfer, with nuclei configuration C_t, two criterions have to be met: The first criterion is the conservation of energy, i.e. $\epsilon_i(C_t) = \epsilon_f(C_t)$. Secondly, according to the Franck-Condon principle, during the charge transfer, the positions of the nuclei do not change. The change of bond lengths of molecules can in first order approximation be modeled by harmonic functions. Hence, the energies of initial and final state are functions of the configuration of the nuclei:

$$\epsilon_\alpha = \frac{1}{2}K\,(C - C_\alpha)^2 \tag{2.5.16}$$

with $\alpha = i,\ f$. If the configuration is now considered a generalized reaction coordinate, the transition state C_t is the configuration C where the two parabolas meet, e.g. where $\epsilon_i(C) = \epsilon_f(C)$. Keeping in mind the quantum mechanical nature of the electrons, the existence of bound and anti-bound delocalized states with energies ϵ^+ and ϵ^- (Eq. 2.5.9), leads to a gap in the energy functions $\epsilon_\alpha(C)$. This is illustrated in Fig. 2.5.2.

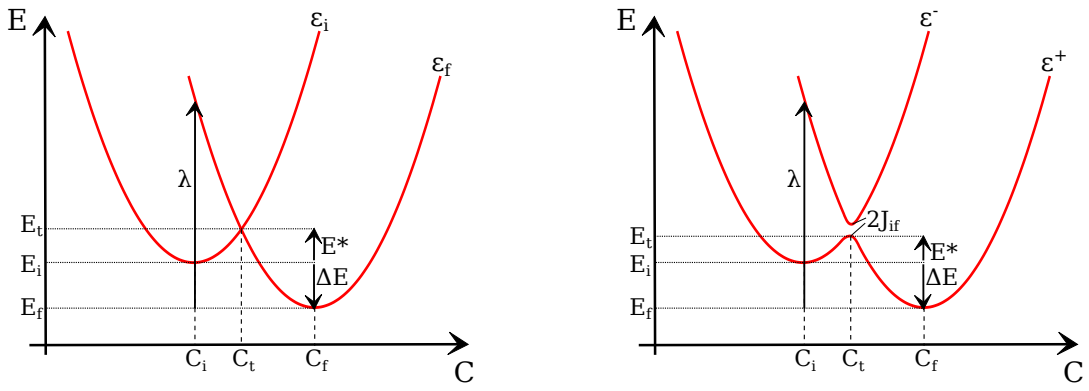


Figure 2.5.2.: Potential energy surfaces of the charge transfer in dependence of the nuclear configuration. (a) In a classical picture, where the charge is located either in the initial or the final state, the energies of initial and final states, ϵ_i and ϵ_f, are - in first order approximation - a harmonic function of the nuclear configuration. Due to conservation of energy, the transition configuration is the intersection of both parabolas. (b) The quantum mechanical nature of the electrons leads to a gap in the energy functions of $2\,J_{if}$. The eigenstates of the Hamiltonian are a bound and an anti-bound state, ϵ^+ and ϵ^-. The charge transfer is characterized by the energy difference ΔE between relaxed final and initial state, the reorganization energy λ and the matrix overlap element J_{if}.

Regarding the overlap of the nuclear wave functions, the $\delta-$distribution asserting the conservation of energy allows only states, where the energies of initial and final wave functions are equal. In the classical picture, this is the case when $\chi_i = \chi_f$. This is fulfilled at the transition point, and one finds $\langle\chi_i|\chi_f\rangle = 1$ in this special case. However, the energy at the transition point is E^* above the initial energy, E_i. In a classical picture, thermodynamic ensembles, the occupation probability then follows a Boltzmann distribution:

$$\rho(E^*) \propto \exp\left(-\frac{E^*}{k_\mathrm{B}\,T}\right) \tag{2.5.17}$$

Thus, the absolute square of the orbital overlap together with the conservation of energy can be substituted by the (normalized) Boltzmann occupation probability of the transition state:

$$\delta(E_i - E_f)\,|\,\langle\chi_i|\chi_f\rangle\,|^2 = \frac{1}{Z'}\exp\left(-\frac{E^*}{k_\mathrm{B}\,T}\right) \tag{2.5.18}$$

Where Z' is a renormalized partition function. It is convenient to express the energy of the transition state, E^*, in terms of energy difference ΔE and reorganization energy λ. For this, it is assumed that the harmonic functions describing the dependence of initial and final energy on the nuclear configuration, $\epsilon_i(C)$ and $\epsilon_f(C)$, have approximately the same force constant:

$$\epsilon_i(C) = \frac{1}{2}\,K\,(C_i - C)^2 \tag{2.5.19}$$

$$\epsilon_f(C) = \frac{1}{2}\,K\,(C_f - C)^2 + \Delta E \tag{2.5.20}$$

The configuration at the transition point, C_t, is where the functions intersect:

$$\epsilon_i(C_t) = \epsilon_f(C_t) \tag{2.5.21}$$

$$\frac{1}{2}\,K\,(C_i - C_t)^2 = \frac{1}{2}\,K\,(C_f - C_t)^2 + \Delta E \tag{2.5.22}$$

This results in an equation for the transition state C_t:

$$K\,C_t^2 + \frac{1}{2}\,K\,C_t\,(2\,C_f - 2\,C_i) + \frac{1}{2}\,K\,(C_i^2 - C_f^2) - \Delta E = 0 \tag{2.5.23}$$

Solving this equation and inserting into the expression for E^*:

$$E^* = \frac{1}{2}\,K\,(C_i - C_t)^2 \tag{2.5.24}$$

results in an expression for E^*:

$$E^* = \frac{(\Delta E + \lambda)^2}{4\,\lambda} \tag{2.5.25}$$

with the reorganization energy λ:

$$\lambda = \frac{1}{2} K \left(C_i^2 - C_f^2 \right) \tag{2.5.26}$$

This equation allows an interpretation of λ: Consider the case that the charge transfer would happen instantaneously without a change in the nuclear configuration, i.e. the charge suddenly sitting on the final molecule, but the nuclear configuration of the combined system remaining in the initial configuration, C_i. The energy of the system is then $\epsilon_f(C_i)$ and subsequently to the charge transfer, the system would relax to the minimum of $\epsilon_f(C)$ at $C = C_f$. The energy difference that is gained by this relaxation, i.e. the reorganization of the nuclear configuration, is $\epsilon_f(C_i) - \epsilon_f(C_f) = \lambda$.
Substitution of E^* in Eq. 2.5.18 and using the correct normalization results in [116]:

$$| \langle \chi_i | \chi_f \rangle |^2 = (4 \pi \lambda k_\mathrm{B} T)^{-1/2} \exp \left(- \frac{(\Delta E + \lambda)^2}{4 \lambda k_\mathrm{B} T} \right) \tag{2.5.27}$$

While the exponent is the Boltzmann weight of the transition state, $\exp(-E^*/k_\mathrm{B} T)$, the prefactor is not the classical partition function. The reasons for the deviation lie in the quantum mechanical nature of the vibrational states.
Inserting Eq. 2.5.27 this into Eq. 2.5.15 yields the expression of the hopping rates in Eq. 3.5.1.

Calculation of the Hopping Rate Parameters ΔE, J_{if} and λ

The energy difference between initial and final state of the hopping process, ΔE, the coupling matrix element J_{if} and the reorganization energy λ are the central quantities in the formula proposed by Marcus (Eq. 3.5.1). For each pair of molecules, i.e. each possible hopping process, they can be derived using quantum mechanical methods. Without going into detail, a qualitative description of the calculation of these quantities is given in the following.

Calculation of ΔE In principle there several ways to calculate ΔE for a hop from one molecule to a neighboring site. The easiest approach is to extract the orbital energies of the LUMOs (for electronic transport) or HOMOs (for hole transport) calculated for the molecules in vacuum and calculate ΔE by:

$$\Delta E = E_{\mathrm{HOMO/LUMO}}(1) - E_{\mathrm{HOMO/LUMO}}(2) \tag{2.5.28}$$

where $E_{\mathrm{HOMO/LUMO}}(1/2)$ denote the HOMO and LUMO orbital energies of inital and final state respectively. Calculation of the orbital energies in vacuum however does not account for the influence of the environment that is unique for each molecule in an amorphous organic structure. The resulting ΔE is then just based on deformation of the molecules. A slightly more sophisticated way to calculate ΔE is to include a shell of surrounding molecules into the calculations of HOMO or LUMO energies. While environmental effects are included, the change of orbital energies by additional charges occupying HOMO or LUMO is still neglected. This can be accounted for by placing an additional charge first on the initial

and then the final molecule and calculate ΔE as the difference in the ionization energies. Further refinement up to the point where the influence of the additional charge on the charge distribution of the neighboring molecules is taken into account can be found in literature [3, 56].

Calculation of J_{if} The calculation of the J_{if} is performed in first order perturbation theory: First, the wave functions defining the orbitals of the monomers (the molecule that is charged before the hop and the molecule that is charged after the hop) are calculated, i.e. the monomer wave functions. Secondly, the dimer Hamilton operator is extracted from the quantum mechanical calculations. This Hamiltonian comprises the monomer Hamiltonians on the block-diagonal elements and the overlap Hamiltonians J on the block-off-diagonal elements:

$$H_{dimer} = \begin{pmatrix} H_{mono,i} & J_{if} \\ J_{fi} & H_{mono,f} \end{pmatrix} \tag{2.5.29}$$

Each entry of this matrix is itself another matrix of the dimensionality of the basis describing the monomer wave functions. The orbital overlap is then calculated using the states $|\Psi_i, 0\rangle = |\Psi_i\rangle \oplus |0\rangle$ and $|0, \Psi_f\rangle = |0\rangle \oplus |\Psi_f\rangle$:

$$J_{if} = \langle \Psi_i, 0 | \, H_{dimer} \, |0, \Psi_f\rangle = \begin{pmatrix} \Psi_i \\ 0 \end{pmatrix} \begin{pmatrix} H_{mono,i} & J_{if} \\ J_{fi} & H_{mono,f} \end{pmatrix} \begin{pmatrix} 0 \\ \Psi_f \end{pmatrix} \tag{2.5.30}$$

Calculation of λ The reorganization energy λ usually comprises two parts: λ_i, the inner contribution by the ionic reorganization of donor and acceptor molecule and the energy λ_o, induced by the ionic relaxation of the environment. For organic systems, $\lambda_o << \lambda_i$ can be assumed. The inner contribution to the reorganization energy is calculated as the sum of the energy changes occurring due to the ionic relaxation of donor and acceptor:

$$\lambda_i = \lambda_{donor} + \lambda_{acceptor} \tag{2.5.31}$$

λ_{donor} is derived by removing the charge from the donor molecule, calculating the energy $E_{donor,0}$ of the neutral molecule, relaxing the geometry and calculating the energy of the final, relaxed, uncharged system, $E_{donor,1}$. The donor contribution to λ_i, λ_{donor} is then just the difference between $E_{donor,1}$ and $E_{donor,0}$. The calculation of the acceptor is performed accordingly: The acceptor is charged, the energy calculated, the donor molecule relaxed and $\lambda_{acceptor}$ calculated as the difference of the energies [117, 118].

2.5.3. Second Order Charge Transfer: Super-Exchange

The hopping processes within the original Marcus theory described above include direct hops between neighboring sites. Exemplified by the guest-MOF-System in section 3.5, there are systems where significant conductivity has been observed, but hopping processes are impeded by large energy differences between neighboring sites while the hopping between aligned orbitals is prevented by large distances.

In the context of charge transfer in DNA an extension to the established first-order processes presented above was proposed [119]. In this molecular super-exchange (SX), the transfer between second nearest neighbors occurs by virtual occupation of an intermediate state. The total electronic coupling between initial state i and final state f is then given by:

$$J_{if}^{tot} = J_{if} + J_{if}^{SX} = J_{if} + \sum_{v} \frac{J_{iv}\, J_{vf}}{\Delta E_{sx} + 0.5\, \lambda_v} \tag{2.5.32}$$

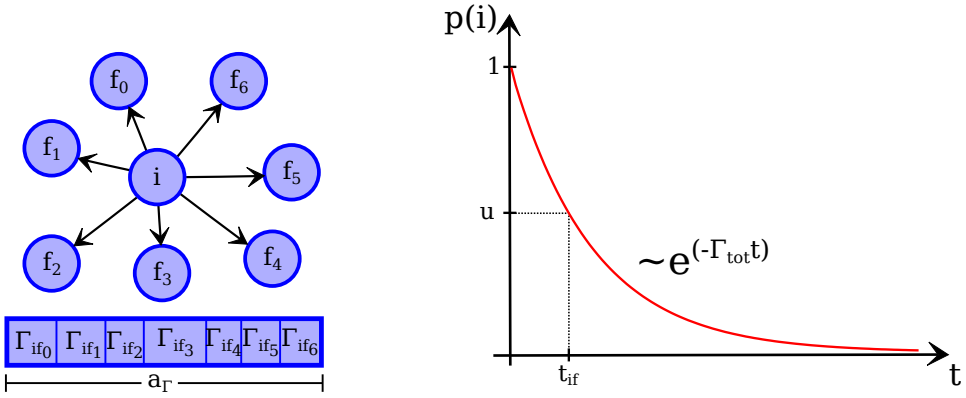
Here, the J_{if}, J_{iv} and J_{vf} are the direct nearest neighbor couplings of Eq. 2.5.30, between initial and final state, initial state and virtual state, and virtual state and final state respectively. The denominator of the super-exchange contribution comprises the difference $\Delta E_{sx} = E_v - \frac{1}{2}\,(E_i + E_f)$ between virtual site energy and averaged energy of initial and final states, and the reorganization energy of the virtual state λ_v. The reorganization energy is added to the energy difference to account for the fact that the virtual occupation does not result in an ionic relaxation of the virtual state, leading to a total energy of $E_v + 0.5\,\lambda_v$, with the on-site energy of the virtual state after ionic relaxation, E_v. For second nearest neighbors i and f that are aligned in energy, but experience small electronic coupling due to large distances, this increases the transfer rate Γ_{if} significantly: The exponential expression in Eq. 3.5.1 is close to 1 due to $\Delta E = 0$ and the coupling J_{if} is increased by the virtual occupation of the intermediate state.
This is an extension to existing charge transfer models inducing additional hopping processes. An illustration of the problem and possible hopping processes is presented for the case of guest-MOF systems in Fig. 3.5.1.

2.5.4. Kinetic Monte Carlo Charge Transport Simulations

Once the hopping rates between each pair of molecules are calculated, each molecule is reduced to a point-shaped hopping site. The propagation of charges through the system is simulated using Kinetic Monte Carlo (KMC).

The Kinetic Monte Carlo Protocol

In contrast to the MC methods described in 2.4, the quantities of interest cannot be extracted by simply sampling all possible configurations of the system. In addition to the configurations itself, a time dependence of the charge propagation is needed in order to link the propagation of the charge carriers with macroscopic electronic properties such as the charge carrier mobility. This is done by identifying a physical time step with each (randomly chosen) Monte Carlo step.
The main idea of the KMC approach is illustrated in Fig. 2.5.3. Starting from any initial sites i, transitions to all possible (neighboring) sites are considered as a next MC step. Each transition from i to f_k is associated with a transition rate Γ_{if_k}, calculated on the basis of Marcus formalism. The probability to perform this transition is given by the ratio of the



(a) (b)

Figure 2.5.3.: Schematic illustration of the a single KMC step. Starting from an initial state i, the hop to all possible final states f_0, f_1, ... f_N are considered as an MC step. The probability to perform a transition into a final state f_k depends linearly on the hopping rate Γ_{if_k}. A uniform random number $a_\Gamma \in (0, \sum_l \Gamma_{if_l})$ determines the final state k via the corresponding interval of the probability array constructed of all Γ_{if_k}. (b) For a given set of $\{\Gamma_{if_k}\}$, the probability of the charge to remain on the initial site i decays exponentially with the time. The time step $t_i f$ related to each hopping process is calculated by drawing a random number a_t from a probability distribution $\rho \propto \exp(-\Gamma_{tot}\, t)$ where the total rate Γ_{tot} is the sum of all individual transition rates.

transition rate to the sum of the transition rates to all neighbors:

$$\pi_{if_k} = \frac{\Gamma_{if_k}}{\sum_l \Gamma_{if_l}} \tag{2.5.33}$$

To randomly draw the transitions with the correct probabilities, the Γ_{if_k} are combined into an array. The total length of this array is then the sum of all Γ_{if_k}. A random number a_Γ is drawn from a uniform distribution with $a_\Gamma \in (0, \sum_l \Gamma_{if_l})$. The transition $i \to f_k$ that is to be performed is then identified by

$$\sum_{l=0}^{k-1} \Gamma_{if_l} < a_\Gamma < \sum_{l=0}^{k} \Gamma_{if_l} \tag{2.5.34}$$

The association of an appropriate time step with the transition $i \to f_k$ is based on statistical assumptions. The total rate by which the state i is left, is the sum of the individual rates of all possible transitions:

$$\Gamma_{tot} = \sum_l \Gamma_{if_l} \tag{2.5.35}$$

The change of the occupation of state i at a certain time t is linear in the rate and the current occupation. Instead of using absolute occupation numbers, this can directly be transferred to the occupation probabilities $\rho_i(t)$:

$$\frac{\partial \rho_i(t)}{\partial t} = -\rho_i(t)\, \Gamma_{tot} \tag{2.5.36}$$

The occupation probability of state i after a time t is then

$$\rho_i(t) = \exp\left(-\Gamma_{tot}\, t\right) \tag{2.5.37}$$

with the occupation of the state i starting at $t = 0$, i.e. with a probability $\rho_0 = \rho_i(t_0) = 1$. The time step associated with the hop off state i is then the time calculated following the distribution given by Eq. 2.5.37. This is illustrated in Fig. 2.5.3b. In practice, a random time step t_{if} following this distribution is generated by drawing a random number u from a uniform distribution with $u \in (0, 1)$ and calculating t_{if} via:

$$t_{if} = -\frac{\ln(u)}{\Gamma_{tot}} \tag{2.5.38}$$

Extracting the Charge Carrier Mobility

Running the KMC protocol for a large number of simulation steps yields trajectories containing positions of electrons and time steps. Using this information, mean charge carrier velocities can easily be calculated by averaging distance intervals over time steps. The averaged charge carrier velocities are then divided by the electric field that was applied during the KMC simulation, resulting in the charge carrier mobilities.
This concludes the principle approach of simulating charge propagation via KMC and an

implementation seems straight forward. In practice, however, several difficulties arise, like the fast evaluation of the Coulomb-interaction between all charge carriers in the system. The change of the position of a single charge carrier changes the hopping rates Γ_{if} of all charge carriers and all possible final sites f, rendering this interaction a crucial effect in KMC simulations.

2.6. Efficient Calculation of Macroscopic Electronic Parameters in Organic Semiconductors via the Multiscale Modeling Approach

On one hand, the charge transport depends strongly on quantum mechanical effects on the single molecule scale. On the other hand, large samples are needed for the simulation of hopping transport through bulk material in order to extract macroscopic parameters and account for such phenomena as charge percolation. As a full quantum mechanical description of devices on the nanometer scale is not feasible, a multi-scale modeling approach is applied. Quantum mechanical methods such as DFT or Hartree-Fock are used to calculate the electronic structure and the ionic geometry of molecules in vacuum. The quantum mechanical properties of the molecules are mapped onto classical functions depending solely on the position of the nuclei for the fast evaluation of the interaction between single molecules. Large molecular sample structures are generated using MD or MC-based protocols. On the basis of these sample morphologies, for each pair of molecules, the orbital overlap, the energy levels of molecular orbitals and the reorganization energies are calculated and fed into Marcus theory of hopping transport. The resulting rates are used as an input for KMC simulations of charge transport, yielding macroscopic quantities, such as charge carrier mobility.
Whereas the KMC simulation does not take into account the exact electronic structure of the single molecules, the hopping rates are a direct result of the orbital overlap and the orientation and deformation of the molecules in the sample morphology. The morphology in turn is influenced by the molecular geometry and the electronic structure via parametrization of the classical force-field functions. In this manner, properties of organic semiconductors on the smallest scale, the electronic structure of the single constituents, feed into KMC. Thus, the resulting macroscopic transport properties take into account the quantum mechanical nature of the molecules, without having calculated the whole sample on the full quantum mechanical level.
The methods presented in this chapter present an extensive toolkit for the modeling of electronic properties in organic semiconductors. This toolkit will be applied to the calculation of the electronic properties of four (metal-) organic materials in chapter 3.

3. Electronic Structure of Organic and Metal-Organic Materials

3.1. Organic and Metal-Organic Electronic Devices

3.1.1. Introducing Novel Properties by Organic Compounds

Metals and inorganic semiconductors are the basis of modern technology. Computer chips, transistors, photovoltaics (PV) and light emitting diodes (LEDs) are merely examples of the wide range of applications in everyday life. While it is possible to modify material properties using binary compounds (such as GaAs) or doped layers, the range of electrical device characteristics such as conductivity, band gaps and emission spectra is limited. Furthermore, the sensitive and frail production process of monocrystalline inorganic layers prevents inexpensive, large-scale manufacturing of durable, large-area devices.
A material class that offers a vast number of compounds with different properties is organic molecules. The incorporation of organic compounds into the concept of metallic or inorganic electronic devices promises a variety of novel materials with electrical properties tunable by molecular design. In the past decades, this prospect motivated two approaches in nanoelectronics: The construction of thin films of purely organic molecules (organic semiconductors) and the combination of metallic and organic compounds into metal-organic devices.

3.1.2. Organic Semiconductors

Since the discovery of electroluminescence in thin layers of small organic molecules in 1987 [9], organic semiconductors consisting of small molecules have been intensively studied and are widely applied in organic light emitting diodes (OLEDs) [1, 2], organic photovoltaics (OPV) [3–5] or organic field effect transistors (OFETs) [6]. The prospect of low-cost production, scalability to large area devices and control of device properties by specifically tuning single molecule properties has been the driving force in extensive efforts to increase device efficiency in the past years [10–13]. Organic semiconductors consist mostly of amorphous structures whereas organic crystals (e.g. acene-crystals) remain an exception. While amorphous layers can be realized on a large scale by vapor deposition or solution processing, to date, established materials are limited in stability, electric yields and life time.
Further, the large number of possible organic molecules render the identification of suitable molecules that result in the desired device properties a costly and time-consuming challenge to experimentalists. To circumvent this bottleneck in the development of organic materials, efforts have been made towards the calculation of device characteristics on the basis of the single-molecule properties [3, 14, 28, 36, 83, 120]. Computational resources are becoming

gradually cheaper, promising a full *in silico* ("in-the-computer-chip") design of devices by fast and low-cost pre-screening of materials. Additionally, the device simulation is anticipated to give insight on processes on the electronic scale, allowing a deeper understanding of the fundamental processes in organic electronics that are responsible for limited power conversion efficiencies (PCE) and device degradation.

3.1.3. From Purely Organic to Metal-Organic Structures

Another approach to overcome fast degradation and limited device stability is the combination of metallic and organic materials. The coordination between metallic centers and organic compounds via metal-organic bonds is known in natural processes: Fe^{II} centers coordinated to the nitrogen atoms of porphyrin, for example, play a major role in binding oxygen to hemoglobin, an essential process in physiological respiration. Another example is ferrocene ($Fe(C_5H_5)_2$), where two five-membered aromatic rings are coordinated by an Fe^{II} center. This "sandwich compound" and its derivatives are widely applied in pharmaceutical applications, as fuel additive and in material science [121–123].
In nanoelectronics, the prospect of combining outstanding properties of two material classes motivates the coordination of metals to organic compounds with the objective to realize ultra-thin, long term sustainable devices with tunable properties by molecular design.
It was shown that the encapsulation of molecular layers by carbon and copper electrodes leads to long-lived, temperature-stable devices with high PCE [124]. Another promising approach is the construction of extended molecular systems by combining organic linker molecules and metallic ions into metal-organic oligomers. As we will show in this chapter, this leads to rather ordered, tightly packed molecular structures that have demonstrated large life-times and high yields in experiment [9].
Further, the repetitive coordination between metallic centers and organic molecules enables the formation of three-dimensional periodic arrangements, in analogy to ionic crystals. The modular set-up of these porous metal-organic frameworks (MOFs) promises a multitude of applications with tunable properties and to date, MOFs were successfully applied for gas storage, molecule separation, molecular sensing and photovoltaics [126–128]. Hence, the vast number of combinations of metal-complexes with organic molecules opens the prospect of specific design of light weight, mono-crystalline devices of high purity, with tunable electronic properties such as band gap, optical excitation or conductivity. To date, however, conductivity in MOFs has been limited to values orders of magnitude below metallic conductivities and the application as electronic devices has yet to be demonstrated. Therefore, the electronic properties of MOFs have gained increasing interest over the recent years [129–136]. Amongst other things, it was demonstrated that loading of guest molecules into the MOF pores significantly changes the electronic properties of the hosting MOF [25].
While there are established approaches to separately study inorganic crystals or single organic compounds, the prediction of electronic properties of combined materials presents a formidable challenge to theory. A fundamental understanding of the electronic structure of metal-organic systems serves two purposes: first, the identification of microscopic properties and processes limiting electric yields and charge carrier mobilities and secondly, the subsequent *in silico* control of device properties by molecular design.

3.1.4. Structure of the Chapter

This chapter illustrates, how modeling tools can be applied to a variety of material classes based on organic compounds in order to calculate realistic device characteristics. Further, the studies intend to give a view on the electronic characteristics of (metal-) organic structures on the basis of microscopic properties of single constituents: In section 3.2 charge carrier mobilities in small-molecule organic layers are derived *ab inito* for nine different materials. After generating sample morphologies with atomic resolution using MD on the basis of established force-fields, electronic coupling and energy differences between pairs of molecules is calculated along with other microscopic properties. These are fed into an analytical model, yielding charge carrier mobilities. The dependence of the mobility on single-molecule properties is analyzed and a comparison to experimental data investigates the applicability of theoretical models to small-molecule organic electronics.
Subsequently, section 3.3 presents a study on metal centered molecular wires (MCMW). Repeated coordination of metallic redox centers to small organic molecules results in one-dimensional wire-like oligomers. Using Monte Carlo based modeling approaches we show that a set of oligomers arrange in a tightly packed manner that can be extended periodically to form an ordered structure. Using this sample morphology, the charge carrier mobility in an extended nanowire structure is calculated based on microscopic hopping parameters, analogous to the previous study of small molecules, and compared to experimental data.
The second half of this chapter presents two studies of a well know MOF, HKUST-1. MOFs comprise metallic centers and small organic molecules in form of a three dimensional periodic arrangement. The ordered geometry induces properties that differ significantly from organic single-molecule or polymeric structures. The electronic character of MOFs is investigated in two steps. In section 3.4, the intrinsic electronic properties of the Metal-organic framework HKUST-1 are studied. After a brief discussion of the suitability of various methods, band structures and density of states (DOS) are calculated. In order to investigate the influence of the unpaired electrons of the Cu-atoms in HKUST-1 on band structure formation and excitation energies, Cu is substituted by Zn in a second model for comparison. As the treatment of periodic structures goes at the expense of accuracy, excitation energies of small entities/fractions of the MOF-unitcell were analyzed using refined DFT-based methods with high accuracy. The resulting theoretical analysis is compared to experimental observations.
Subsequent to the study of intrinsic MOF properties, section 3.5 proposes a hopping mechanism for charge transfer through metal-organic frameworks as an alternative to band-like transport of metallic or inorganic crystalline structures. In order to analyze the influence of small molecules loaded into the MOF pores, the hopping rates between two MOF-orbitals, between two orbitals of guest-molecules and between MOF- and guest-orbitals are calculated. The role of second-order processes which were first presented in section 2.5.3 is investigated and kinetic Monte Carlo (KMC) simulations are applied to calculate charge carrier mobilities. Results presented in this chapter were partially published. The respective references are referred to accordingly.

3.2. Thin Small-Molecule Organic Layers

3.2.1. From Single Molecule Properties to Device Characteristics

As mentioned in the introduction to this chapter, organic semiconductors consisting of small molecules are used for a wide range of applications, such as OPV or OLEDs. The large number of theoretically synthesizable compounds with different electronic properties motivates efforts to specifically design device properties. To date, the practical application of small-molecule organic layers, e.g. in large area devices, has been limited by the low charge carrier mobility in amorphous thin films.
In the following, the multiscale modeling approach presented in chapter 2 is applied to nine widely used small molecule semiconductors. By disentangling the influences of different material specific properties on the charge carrier mobility, the molecular properties influencing charge carrier mobilities is identified. Information on the dependence of the mobility and single molecule properties enables targeted *in silico* design of small molecules for the application in organic electronic layers.
The nine materials that are studied in this section are displayed in the inset of Fig. 3.2.2a. Alq_3, α-NPD, DEPB, mBPD, NNP, pFFA and TPD form amorphous (disordered) layers, whereas TET- and PEN- (Tetracene and Pentacene) morphologies are crystalline. As explained in section 2.5, alignment of orbital energies with injecting electrodes leads injection or extraction of electrons into or from the system. The charge transfer between molecules then occurs either by moving electrons or holes, which have different mobilities in the same material. Alq_3 for example is usually employed as electron conductor while the alignment of the HOMO and LUMO orbitals of α-NPD molecules with state-of-the-art electrodes leads to hole transport.
This work is currently prepared for publication and is presented here in full consent with all contributing authors.

3.2.2. Generation of Atomistic Morphologies

Following the multiscale modeling approach presented in chapter 2, atomistic morphologies of the materials α-NPD, DEPB, mBPD, NNP, pFFA and TPD were generated via force-field based molecular dynamics (MD) simulations using the simulation package GROMACS [88, 89]. In order to compare with earlier work, an Alq_3 morphology was taken from [36] and the TET and PEN crystal structures from [137, 138]. Using preoptimized single molecule geometries, Van-der-Waals attraction and Pauli-repulsion were modeled using the Lennard-Jones potential as parametrized in the general AMBER force-field (GAFF) [34] which is known to reproduce thermodynamic properties of structures of a wide range of small molecules accurately [139]. The electrostatic interaction was approximated by the Coulomb potential with AM1-BCC partial charges. Periodic boundary conditions were applied in all three dimensions.
The initial dilute molecular structure was constructed by randomly filling a cubic box with 300 to 700 molecules. Subsequently, the box size was reduced to generate a density of $0.9\,\mathrm{g/cm^3}$, which is approximately 70 % of the Alq_3 density [140], to generate a condensed

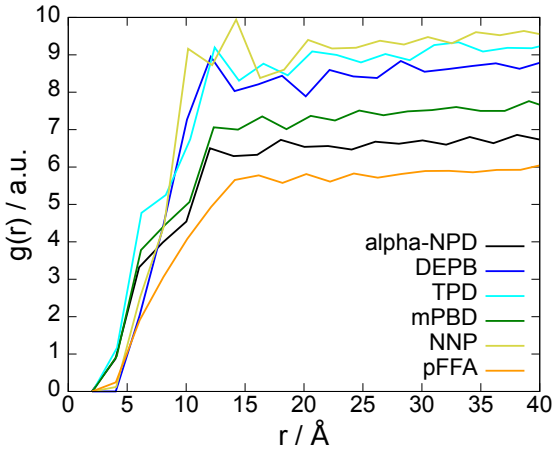


Figure 3.2.1.: Radial distribution functions $g(r)$ of the atomistic morphology structures generated by means of molecular dynamics. Neither of the structures shows ordering beyond the nearest neighbor.

starting structure. After smoothing atomic forces (resulting from low atomic distances) by energy minimization over 10^4 steps deploying the steepest-descend-algorithm as implemented in GROMACS, a short 10 ps NVT (constant number of particles, constant volume, constant temperature) run at $T = 800\,\mathrm{K}$, which is well above the glass temperature of all materials, was performed to distribute the molecules evenly in the simulation box and to equilibrate the velocity distribution. To relax the sample at the correct pressure, a 1 ns NPT (constant pressure instead of constant volume) run with $T = 800\,\mathrm{K}$ and $p = 1\,\mathrm{bar}$ was performed. Equilibration of the pressure to 1 bar was verified and mean square displacement of the molecules throughout the 1 ns run was surveyed to verify that the system was in the liquid state (in which the mean square displacement should be larger than the characteristic length scale of the molecule). The equilibration in the liquid phase was followed by a 5 ns NPT run, cooling the sample to room temperature at a rate of $\frac{dT}{dt} = 100\,\mathrm{K/ns}$. A final 2 ns NPT simulation at $T = 300\,\mathrm{K}$ was performed to ensure equilibration and to collect data. The mean square displacement in the final run was close to zero, indicating a condensed solid phase where movement is limited to thermal vibrations.
In order to characterize order and conformation in the sample (amorphous or crystalline), the radial distribution function of the six materials was calculated and is displayed in Fig. 3.2.1. No short- or long-range order is visible, clearly indicating amorphous structures.

3.2.3. Charge Carrier Mobilities

[1]For each pair of molecules in the morphologies that was generated as described above, electron coupling matrix (J_{if}), reorganization energy (λ_{if}) and Gibbs free energy difference (ΔG_{if}) including polarization effects and conformational disorder were calculated. The

[1]The calculation of the charge carrier mobilities was performed by Pascal Friederich and a brief overview is presented in the following. Full results have been submitted for publication.

calculation of the coupling matrix elements was based on the Loewdin orthogonalization procedure [141, 142] and the pairwise reorganization energies were derived using Nelson's four point method [118, 143]. Both calculations were performed using DFT as implemented in TURBOMOLE [70] with the B3-LYP functional [144] and the basis sets def2-TZVP [145], def-SV(P) and SV(P) [146] for λ_{if}, J_{if} and ΔG_{if} respectively. To account for environmental effects, i.e. the electrostatic influence of the surrounding amorphous phase on the electronic structure of the single molecules, a self consistent iterative approach was applied [3]. Values were obtained for enough pairs in order to reduce statistical fluctuations to an acceptable level.
Using the generalized effective medium model (GEMM)[27], an analytical alternative approach to KMC simulations (as described in chapter 2.5.4), the charge carrier mobilities were calculated according to:

$$\mu = \frac{e\,\beta\,M\,<J^2\,r^2>}{n\,\hbar\,\lambda^{1/2}} \left(\frac{\pi\,\beta}{1+\frac{\beta\,\sigma^2}{\lambda}}\right)^{1/2} \exp\left[-C\left((\beta\,\sigma)^2-\beta\,\lambda\right)\right] \tag{3.2.1}$$

Here, $<J^2\,r^2>$ and λ are the statistical averages over the squared product of microscopic hopping matrix elements J_{if} and the distance of the hopping sites i and f and reorganization energies respectively, as derived above. $\beta = k_\mathrm{B}\,T^{-1}$ the inverse temperature ($T = 300\,\mathrm{K}$ was used), n is the dimensionality, M is the mean number of nearest neighbor molecules and e and $\hbar$ are elementary charge and reduced Planck constant respectively. $\sigma = \sigma(E)$, the energy disorder, is the width of the local density of states, calculated as standard deviation of the energy differences between initial and final hopping states, $\sigma(\Delta E)$. The parameter of the effective medium model was set to $C = 0.25$, as previously used in literature [27, 147, 148].
Fig. 3.2.2 presents an overview of the charge carrier mobility calculations. In Fig. 3.2.2a, the theoretically derived mobilities are compared to experimental values from literature. Very good agreement (within one order of magnitude) is observed over several orders of magnitude for the hole mobility of both amorphous and crystalline systems. Fig. 3.2.2b shows the local density of states for the two materials NNP and pFFA as exemplification. As the mobility depends exponentially on the width of this distribution, it is considered a crucial parameter for the mobility calculation. Accordingly, the material NNP, with the disorder of $\sigma = 0.135\,\mathrm{eV}$ and thus 20 % larger than the disorder of pFFA with $\sigma = 0.112\,\mathrm{eV}$, has a mobility of about one order of magnitude lower than pFFA. Fig. 3.2.2c indicates the influence of the molecular shape on the charge carrier mobility: Alq_3 is of rather spherical shape, whereas the compound pFFA is an extended complex of phenyl rings. This leads to an overall smaller value of $<J^2\,r^2>$ for pFFA due to the wider distribution of the coupling matrix elements and the short nearest neighbor distances. However, the mobility of pFFA is orders of magnitude larger than the mobility of Alq_3, as the mobility depends only linearly on $<J^2\,r^2>$ but exponentially on σ. The better electronic coupling in Alq_3 is overcompensated by the two times larger disorder. The second parameter that enters the exponent of the charge carrier mobility is the reorganization energy λ. In a densely packed material, the ionic relaxation is limited by surrounding molecules. This can be taken into account by fixing the dihedral angles (the torsion of bonds), i.e. a "frozen dihedral approximation". The influence

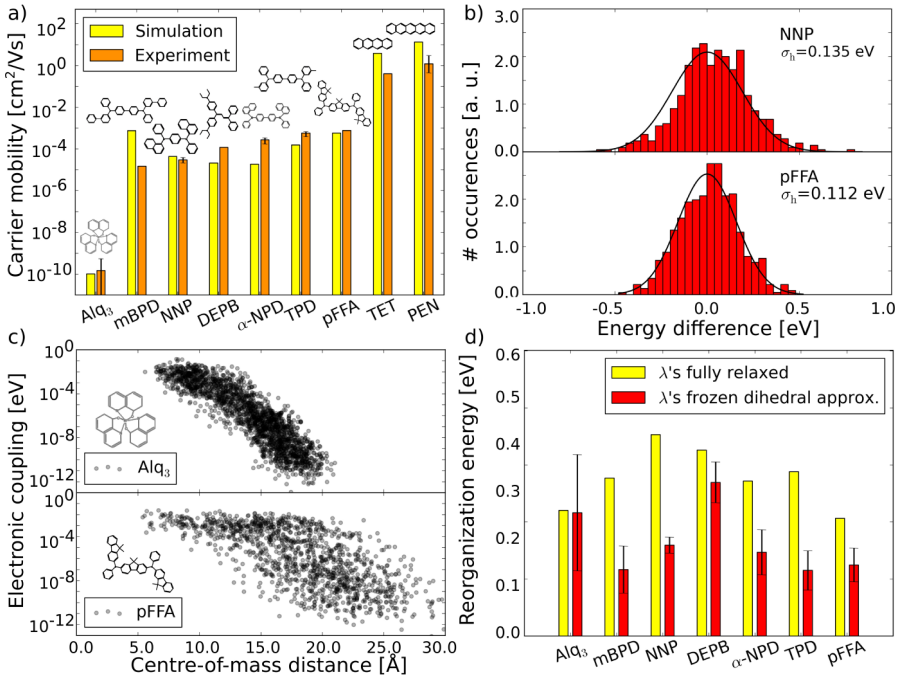


Figure 3.2.2.: a) Comparison of hole mobilities of 9 different organic semiconductors varying from poorly conducting Alq3 to highly conducting crystalline materials, such as Tetracene and Pentacene. Predictions from a parameter-free first-prinicples workflow are in agreement to one order of magnitude for all materials. b) Distribution of energy difference computed in the polaron model for NNP and pFFA. The width of the distribution yields the energy disorder, the main determinant of the mobility. c) Distance dependence of electronic couplings for a compact (Alq3) and a more extended molecule (pFFA). d) Reorganization energies in the fully-relaxed (yellow) and frozen-dihedral approximation (red) differ significantly for mBPD, NNP, α-NPD, TPD and pFFA. Figure and caption were provided by Pascal Friederich and will be published shortly.

Table 3.2.1.: Microscopic input parameters for and results of the analytical approach to the derivation of charge carrier mobilities. The hole mobilities are given for each of the nine materials and are compared to literature values in the last column.

	σ	$< J^2 r^2 >$	M	λ	μ	μ_{exp}
Alq_3	0.224	9.99×10-3	7.31	0.296	1.01×10-10	1.46×10-10 [149–153]
mBPD	0.110	1.52×10-3	8.52	0.143	7.38×10-4	1.49×10-5 [154]
NNP	0.135	1.64×10-3	7.65	0.160	4.31×10-5	2.99×10-5 [155]
DEPB	0.130	1.42×10-3	8.16	0.266	2.09×10-5	1.17×10-4 [156]
α-NPD	0.144	2.04×10-3	7.73	0.158	1.84×10-5	2.70×10-4 [150, 157]
TPD	0.129	1.56×10-3	8.49	0.110	1.52×10-4	5.74×10-4 [150, 152, 156, 158]
pFFA	0.112	1.46×10-3	7.70	0.134	5.70×10-4	7.60×10-4 [154]
TET	0.0	1.13×10-2	15.8	0.114	3.83	0.4 [159]
PEN	0.0	3.07×10-2	15.8	0.097	13.35	1.2 [160]

of this approximation on the values of λ is depicted in panel d of Fig. 3.2.2. Especially for mBPD, NNP, α-NPD and TPD, λ changes significantly under the confinement of neighboring molecules. This clearly underlines the role of the environment in the calculation of charge carrier mobilities.
The material dependent values for the microscopic parameters that feed into the mobility are given in Tab. 3.2.1 along with the resulting charge carrier mobilities. Experimental values for the charge carrier mobility are available in literature and are added in the last column for comparison.
The calculated mobility is in good agreement with experimental data within one order of magnitude, for material specific values ranging from $10^{-10}\,\mathrm{cm}^2/\mathrm{Vs}$ to $10^{-4}\mathrm{cm}^2/\mathrm{Vs}$. As there were no adjustable parameters in this study except for the single parameter C in the GEMM mobility formula, the observed agreement between between experiment and theory is encouraging. Our results suggest that the single molecule properties, such es electronic structure and geometry, determine the mobility to a large degree. By directly comparing the microscopic input parameters and the resulting mobilities of pFFA with NNP and Alq_3, the energetic disorder was identified as the most significant parameter. This opens the prospect of a full *in silico* prediction of organic semiconductor device characteristics on the basis of microscopic properties of single constituents.

3.3. Nanowires Constructed from Metal-Terpyridine Oligomers

3.3.1. From Amorphous Organic Layers to Ordered Metal-Organic Wire Structures

The encapsulation of functional organic materials into nanoscale electronic devices open the prospect of small-scale devices with properties tunable on the molecular level and lead to extensive studies over the past decades [161, 162]. Although test systems have been

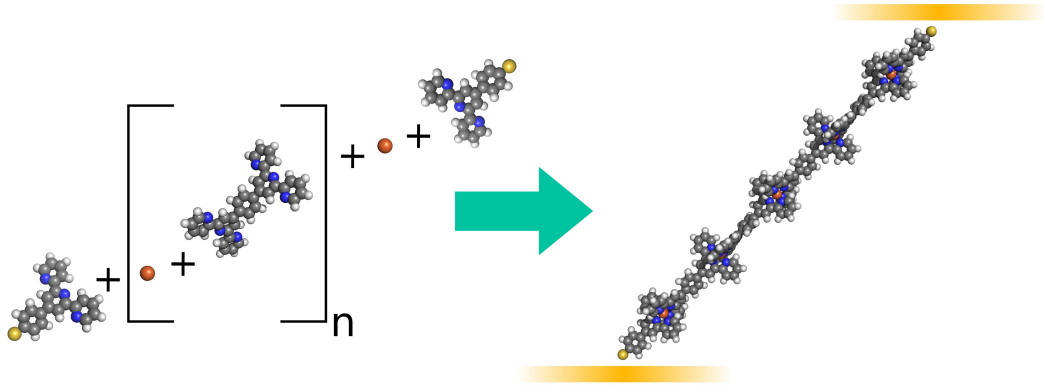


Figure 3.3.1.: Growth of Fe^{II}-terpyridine oligomers: By iteratively coordinating Fe^{II} redox centers to TPT units, extended metal centered molecular wires (MCMW) are formed. The wires are framed by MPTP groups for the attachment to functionalized substrates and electrodes. Note that the pyridine units flip during the coordination of the (M)TPT units to the Fe^{2+} ions.

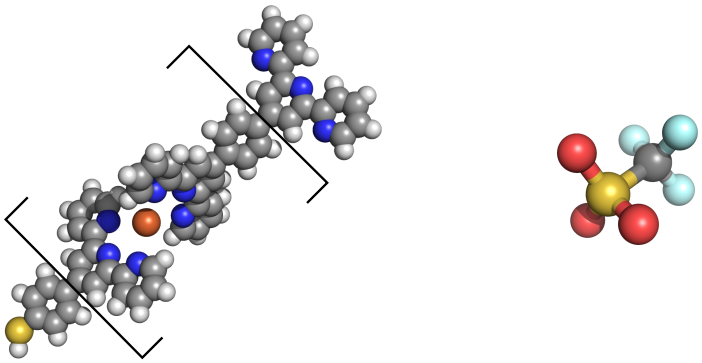
realized [163], the construction of stable, durable devices remains challenging. The damage of molecular structures by the deposition of metal electrodes is one of the major challenges in constructing thin film junctions [164, 165].
To overcome this hurdle, different electrode deposition methods were proposed [166–168] and, more recently, intermediate layers were introduced to protect the fragile molecular layers [169–171]. However, while preserving the molecular structure, both approaches prevent direct access to intrinsic material properties due to additional layers. A solution to this hurdle has been proposed where molecular layers are encapsulated between copper and carbon electrodes, producing stable films with high yields [172].
Another promising approach towards the construction of durable and sustainable devices with tunable electronic properties is to utilize metal-organic bonds to stabilize organic materials. This study aims at the construction of a theoretical model for an extended system consisting of Fe^{II}-terpyridine-oligomers to analyze microscopic properties and device stability. As illustrated in Fig. 3.3.1, the iterative coordination of Fe^{II} redox centers to 1,4-di(2;2';6';2"-terpyridine-4'-yl)benzene (TPT) leads to the formation of 1D-oligomers, referred to as "nanowires" or "metal centered molecular wires" (MCMW). 4'-(4-mercaptophenyl)terpyridine (MPTP) groups are used to attach the MCMWs to the electrodes. Triflate ions are added to the structure to compensate the positive charge of the Fe^{II} redox centers. Molecular mechanical simulations are performed in order to generate a sample morphology of stacked wires, on the basis of which electronic properties, i.e. charge carrier hopping matrix elements and charge carrier mobilities, are analyzed. The results are compared to experimental data.
The results of this section were already published in [9] and are presented here in consent with all contributing authors.

3.3.2. Parametrization of a Single Molecular Wire

In order to extract force-field parameters for the simulation of an extended wire structure on the basis of Metropolis MC, geometry optimization was performed on a molecular model of the wire chosen such that a full wire can be constructed by periodic extension: The 4'-(4-mercaptophenyl)terpyridine (MPTP) unit including a sulfur atom at the bottom (saturated by hydrogen), which is attached to a substrate in experimental setups, is coordinated to a Fe^{II} redox center, which is then linked to a 1,4-di(2;2';6';2"-terpyridine-4'-yl)benzene (TPT) nanowire unit. This unit is illustrated in Fig. 3.3.2a. The geometry of this molecular model and the geometry of the triflate counter ion depicted in Fig. 3.3.2b was optimzed using DFT with the BP86 exchange-correlation functional [173–175] and the def-SV(P) basis set [146] as implemented in the TURBOMOLE program package [70]. Using TURBOMOLE, the Kollman fit procedure [98] was applied to generate a set of point-charges located on the atom positions to reproduce the electrostatic potential of the moieties. For the construction of an extended wire unit, atom positions and partial charges were extracted from the atoms in the black brackets displayed in Fig. 3.3.2a. The distance between two redox sites, i.e. the length of one TPT-unit in the wire, was set to 1.55 nm according to the distance between the centers of the two phenyl rings of the optimized structure depicted in Fig. 3.3.2a. Extended structures were constructed by translation of 1.55 nm along the wire and rotation of the unit around the wire-axis to represent the alignment of the two phenyl rings of the optimized wire subset.
To determine relevant degrees of freedom for the classical simulations, the rotational flexibility of two adjacent wire units along the chain was analyzed. Therefore, the first TPT-unit in the MCMW was rotated in respect to the MTPT-moiety in steps of 9°. The geometry of each resulting configuration was optimized and the energy calculated using the semi-empirical Neglect of Diatomic Differential Overlap (NDDO) formalism as implemented in the package MOPAC [77, 78]. During the relaxation, the relative orientation of the groups was preserved by partially fixing the coordinates of three atoms nearest to the Fe^{II} center in both TPT and MTPT-unit, i.e. atoms spanning a plane perpendicular to the wire-axis. This is illustrated in Fig. 3.3.3a. The resulting energies are displayed in Fig. 3.3.3b. Even small distortions of less than 20° lead to energy penalties of approximately 0.1 eV. As $k_{\mathrm{B}}\,T \approx 0.025\,\mathrm{eV}$ at room temperature, a significant distortion is very unlikely and flexibility within single wires is neglected in the following classical simulations.

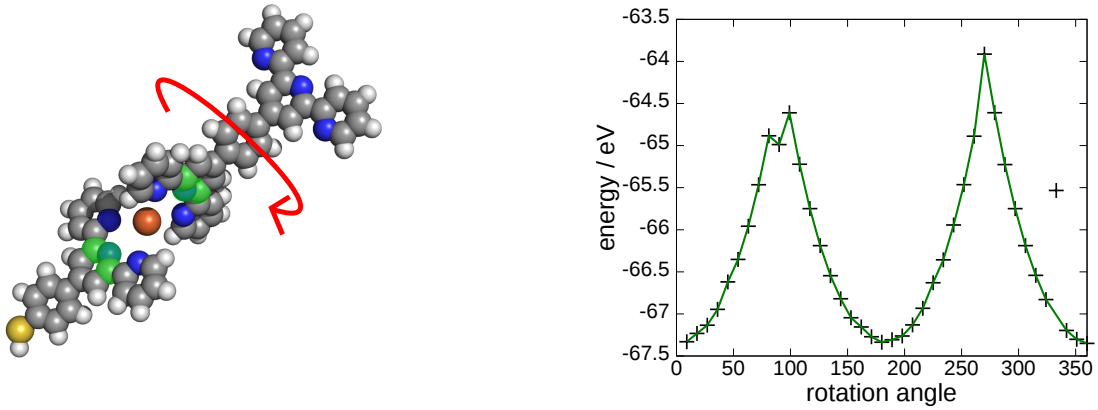
3.3.3. Wire Interlocking: Formation of a Robust Layer

The optimized units were replicated as described above in order to generate a sequence of four TPT-units (see Fig. 3.3.3b) with copies of MTPT-units saturating the nanowires. A triplet of MCMWs was constructed with an initial angle of 50° between wire-axis and z-axis. This angle was chosen to account for the influence of the top and bottom electrodes aligned to the x-y-plane, as the bonds electrode-sulphur and sulphur-carbon are expected to span an angle close to 130°. A total of 30 triflate counter ions was distributed amongst the wires to compensate the positive charge of the 3×5 Fe^{II} redox centers.
In order to produce a thermally equilibrated configuration, the triplet including the triflate



(a) (b)

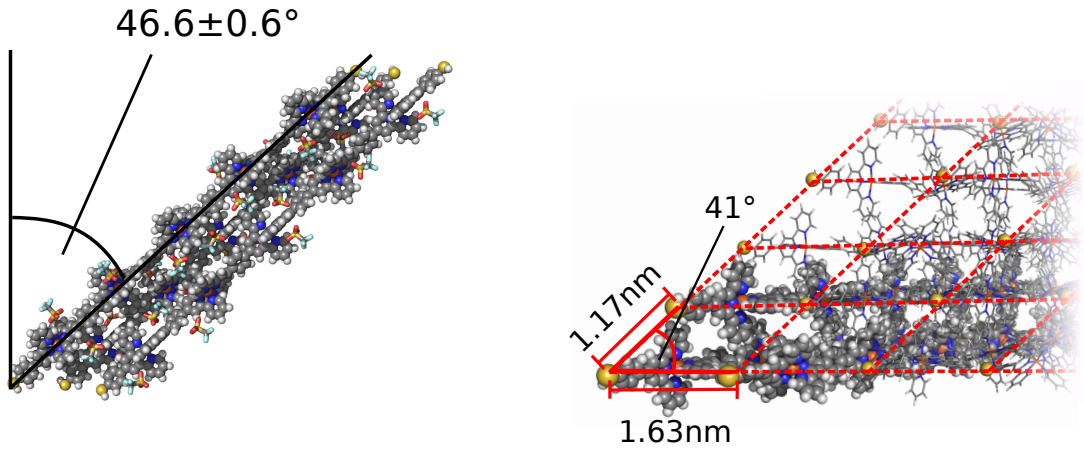
Figure 3.3.2.: Parametrization of nanowires and counter ions for the morphology generation: The geometry of the structures was optimized using DFT before extracting partial charges for the electrostatic force-field. (a) For the parametrization of the nanowires, a MPTP-unit (including a sulfur atom at the bottom) coordinated to a (TPT) nanowire unit via a Fe^{II} redox center was optimized in DFT. For the parametrization of extended wires, the atom positions and partial charges were extracted from the atoms within the black brackets. (b) Triflate counter-ion are added to the simulation for compensation of the Fe^{II} redox sites.



(a) (b)

Figure 3.3.3.: (a) To estimate the rigidity of two adjacent wire units, a TPT unit is rigidly rotated in steps of 9° in respect to the MTPT unit. During the subsequent geometry optimization using MOPAC, the relative orientation is preserved by partially fixing the atoms highlighted in green, i.e. allowing only relaxation parallel to the wire axis. The resulting energies in panel (b) show an increase of approximately 0.1 eV for a distortion of less than 20° from the rest position, indicating strong rigidity.

counter-ions was optimized using Metropolis MC as implemented in SIMONA [86]. To approximate the global optimum in the configurational space, basing hopping simulations where the optimization step was performed by simulated annealing (SA) was performed using ten SA-cycles consisting of 50000 MC steps each, annealing from $T_i = 4000\,\mathrm{K}$ to $T_f = 300\,\mathrm{K}$. The electrostatic interaction between the charge distributions of the constituents was modeled using the Coulomb potential with partial charges as derived above. The Lennard-Jones potential was applied to account for Van-der-Waals attraction and Pauli repulsion, using standard parameters of the GAFF force-field [34]. During the rigid body relaxation of wires and counter-ions, only the centers of geometry of the wires were fixed to the $x-y-$plane ($z = 0$) to account for the confinement between the electrodes. 3D rigid body rotation and 2D rigid body translation of the wires in the x−y−plane as well as 3D rigid body rotation and 3D rigid body translation of the triflate ions was performed during relaxation. In order to avoid the detachment of the triflate ions at high temperatures (e.g. at the beginning of the basing hopping cycles), the molecules were confined to a box of $70 \times 70 \times 70\,\mathrm{nm}^3$ using an attractive harmonic potential acting on molecules outside the volume. The box was chosen large enough in order to have no influence on the relaxed structure.
As the packing of the nanowires is expected to be determined by the strong cohesion energy induced by Coulomb interaction between Fe^{II} redox centers and counter-ions, the influence of the upper and lower electrode is assumed to be comparably small except for alignment in z-direction and is accounted for by initial placement of the wires as described above and constrained wire movement to the x-y-plane. The simulation resulted in a tightly packed



(a) (b)

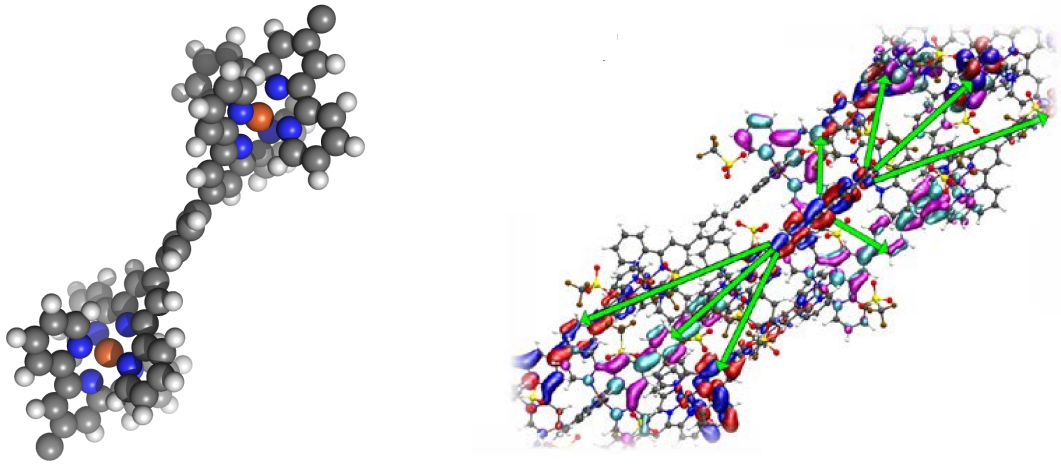
Figure 3.3.4.: (a) Side view of the optimized triplet. Optimized configuration of a triplet of nanowires consisting of four TPT-units enclosed by two MTPT-units. The center of geometry of the wires was fixed along the z-axis and the triplet was relaxed using 3D rigid body rotation in and rigid body translation in x-y-direction. Counter ions were included explicitly to compensate the Fe^{II} charges, and optimized without restriction. The optimization resulted in a tilt-angle of $46.6 \pm 0.6°$ between wire axis and z-axis. (b) Top view of the optimized triplet. In order get an estimate on the periodicity of an extended wire structure, the vectors between the centers of geometry of the triplet wires were calculated, resulting in a rombic unit-cell with lattice parameters of $a = 16.3\,\text{Å}$, $b = 16.3\,\text{Å}$ and an angle $\alpha = 41°$. Periodicity is illustrated by placing images of the triplet (wire-frame representation) on the lattice positions in the x-y-plane.

arrangement of the molecular structure, as displayed in Fig. 3.3.4a. The wires are arranged almost parallel (1.4°, 1.5° and 0.4° between each pair) and the molecular structure shows an orientation of $46.6 \pm 0.6°$ between molecular axis and z-axis. The latter angle almost reproduces the initial angle induced by the substrate, underlining the stability of the wires. More precisely, a difference between the angle resulting from the simulation of free wires and the angle induced by the substrate would decrease structural stability due to competing energy contributions of inter-wire-locking and of the covalent bond between the MTPT unit and the substrate. The vectors between the centers of geometry of the wires were calculated in order to estimate the prospect of periodic extension, resulting in unit cell-parameters of $a = 16.3\,\text{Å}$, $b = 16.3\,\text{Å}$ and a unit cell angle $\alpha = 41°$, as illustrated in Fig. 3.3.4b. In order to emphasize the role of the counter ions in the formation of the densely packed film, we calculated the average distance between the counter ions and the nearest Fe^{II} redox centers to be $6.4 \pm 0.5\,\text{Å}$. The narrow distribution indicates a rather symmetric distribution of the counter ions between the Fe^{II} centers, and a strong electrostatic interaction between wires and counter ions. In fact, the counter ions compensate the Coulomb repulsion between the wires, facilitating dense packing and mechanical stability by strong cohesion.

3.3.4. Electronic Structure of the Wires

[2]As the angle between two adjacent units in the wire connected via an Fe^{II} redox center is fairly rigid (see Fig. 3.3.3b), it is reasonable to assume that neighboring units stack with an angle of 90°. Furthermore, it is known that the delocalization of electrons in organic semiconductors is directly influenced by the (geometrical) alignment of π-orbitals of aromatic substructures. The orthogonal arrangement between neighboring aromatic substructures thus prevents delocalization of charge carriers across more than a single TPT-nanowire-unit and charge transport is expected to occur on the basis of charge carriers hopping between localized states on neighboring wires or adjacent units in a single wire.
In order to estimate the charge carrier mobilities for hopping transport through an extended nanowire thin film, the matrix coupling elements between different hopping sites in the molecular wire structure were calculated based on the relaxed molecular model. Each wire was cut into subsets containing a single TPT-nanowire-unit and half of both adjacent neighboring units, forming two "Cardan-joints" (i.e. connections between wire units via a Fe^{II} redox center). Based on these substructures displayed in Fig. 3.3.5a the coupling matrix elements between neighboring TPT-units were approximated by the coupling matrix element between the half-TPT-units and the central-TPT-unit. The orthogonal stacking of neighboring wire-units and the corresponding low electronic coupling form the bottle-neck for hopping processes along the wire, justifying this choice of atomic subsets. The coupling matrix elements were calculated by deploying the Löwding orthonormalization of neutral dimers. The underlying DFT calculations were performed using the DFT-package TURBOMOLE [70] with the B3-LYP exchange correlation functional [144] and the def2-SV(P) basis set [145]. For the DFT calculation, the Fe^{2+} ion substituted by an effective core potential (ECP, Zn

[2]The calculation of hopping matrix elements and mobilities was for the most part performed by Pascal Friederich.



(a) (b)

Figure 3.3.5.: (a) In order to calculate the coupling matrix elements for hopping processes between two adjacent TPT units, the wires were cut into single TPT-units including the neighboring half TPT-units. (b) Illustration of possible hopping processes along and between neighboring nanowires. Hopping between LUMOs (red, blue) but also between LUMOs and the second lowest unoccupied molecular orbitals, LUMOs+1 (cyan, pink) is considered. Figure (b) was provided by Pascal Friederich and was published in [9].

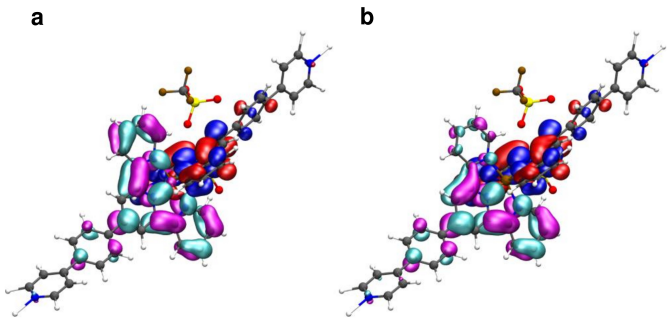


Figure 3.3.6.: Comparison of the LUMO (red/blue) and LUMO+1(pink/cyan) orbitals for an "Cardan-joint" (a) including the Fe^{2+} ion and (b) with the Fe^{2+} ion substituted by an effective core potential (Zn ECP) and a $q = 2+$ point charge. As the change in the orbitals is acceptably small, the substitution of Fe^{2+} by ECP and a point charge is reasonable. The figures are published in [9] and were provided by Pascal Friederich.

and K)[176] and a point charge of $q = 2+$ as the redox site has a vanishing contribution to the LUMO orbitals. This was verified by comparing LUMO and LUMO+1 orbitals for both structures including the Fe^{2+} ion and a Zn-core potential and a point charge $q = 2+$. As illustrated in Fig. 3.3.6, orbitals with Fe^{2+} and combined charge and core potential are comparable. The resulting coupling matrix elements are displayed in Tab. 3.3.1 and an illustration of hopping processes between wires and along a wire is presented in Fig. 3.3.5b. First, the choice of ECPs for the calculation of the hopping matrix elements along the wire influences the results by less than a factor two, indicating that the concept is sound. Second, the hopping matrix elements for hopping along the wire and off wire hopping processes are quite comparable. This indicates that, despite the robust 1D-arrangement of adjacent wire units, the transport through an extended layer occurs through both on and off wire hopping. Subsequently, an analytical approach to the calculation of charge carrier mobilities in the hopping regime [27] was applied for the case of negligible disorder:

$$\mu = \frac{e\,\beta\,f_M}{2\,n} \exp\left(-\beta\frac{\lambda}{4}\right) M \frac{< J^2\,r^2 >}{\lambda^2} \tag{3.3.1}$$

Here, e is the charge of an electron, β the inverse temperature, n the dimensionality, λ the reorganization energy (see section 2.5), M is the average number of hopping sites (i.e. the number of possible final hopping sites from any initial site) and $< J^2\,r^2 >$ the average of the squared hopping matrix elements weighed with the distance of the hopping sites. The factor $f_M = \frac{(\pi\,\beta)^{0.5}}{\hbar}\,\lambda^{3/2}$ contains several important system parameters. The reorganization energies were calculated to $\lambda_e = 0.214\,\text{eV}$ for electrons and $\lambda_h = 0.242\,\text{eV}$ for holes. In the periodically extended structure, displayed in Fig. 3.3.4b, each hopping site has $M = 24$ neighbors, as each wire is surrounded by six wires, and for each hopping site, four orbitals on every neighboring wire are within hopping range. Using the averaged matrix coupling elements in the last row of Tab. 3.3.1, the mobilities are calculated to be $\mu_e = 2.23\,10^{-2}\,\text{cm}^2/\text{Vs}$ for electrons

Table 3.3.1.: Coupling matrix elements for hopping processes along a single wire (on-wire) and between the wires (off-wire). HOMO coupling elements are included for sake of completeness. The on-wire hopping matrix elements are calculated using two different ECPs (K and Zn) and off-wire values are given for the largest and second largest coupling. The LUMO couplings are considerably larger than the HOMO couplings and the hopping matrix elements for on- and off-wire hopping are comparable.

hopping process	HOMO / eV	LUMO / eV
Along wire (K-ecp)		4.19×10^{-3}
Along wire (Zn-ecp)	6.85×10^{-8}	2.44×10^{-3}
Off wire (largest)	3.03×10^{-4}	4.62×10^{-3}
Off wire (2nd largest)	3.32×10^{-5}	1.03×10^{-3}
Off wire (average)	6.01×10^{-5}	2.42×10^{-4}
All coupling (average)	5.93×10^{-5}	3.03×10^{-4}

and $\mu_h = 3.06\,10^{-4}\,\mathrm{cm}^2/\mathrm{Vs}$ for holes. We note that the values for both hole and electron mobilities are an upper bound, as disorder was neglected completely and the GEMM model overestimates the charge carrier velocities [27].

3.3.5. Comparison to Experimental Mobilities

[3]In order to estimate the quality of the theoretical calculations, the results derived above were compared to experimental data. Only the major results will be given, for details on methods and further results, see [9].
The system of condensed metal center molecular wires (MCMW) was produced by alternately depositing TPT wire units and coordinating Fe^{II} redox centers to the TPT-units. Samples were produced using 15, 20, 30 and 40 deposition cycles. Using AFM (atomic force microscopy) the layer thickness of the four samples was measured, resulting in an average thickness of $43/40 = \mathrm{nm}$ per wire unit, as displayed in Fig. 3.3.7. With the length of a TPT unit in the nanowire of 1.55 nm (as derived above), this thickness dependence results in a tilt angle of the wires of $\theta \approx \arccos(1.08/1.55) = 46°$ from the surface normal, which is in perfect agreement with the results from the molecular simulations, $\theta = 46.6 \pm 0.6°$.
The J-V-characteristics of the devices was measured and the experimental mobility calculated using the model of the Richardson-Schottky emission. This model takes into account the injection barrier at the electrodes and results in a relation between current and mobility:

$$J = 4 \left(\frac{1}{f} + \frac{1}{f^{1/2}} - \frac{(1+f^{1/2})^{1/2}}{f} \right)^2 N_0\, E\, \mu\, \exp(-\phi_B / k_\mathrm{B}\, T) \exp(f^{1/2}) \tag{3.3.2}$$

Here, E is the electric field, $f = \frac{e^3\, E}{4\pi\epsilon\epsilon_0 (k_\mathrm{B}\, T)^2}$ the reduced electric field and N_0 the hopping site density. With an average of $N_0 = 1$ (one hopping site per cubic nm), this results in a

[3]Experiments were performed by Dr. Florian von Wrochem, MSL Sony Deutschland GmbH.

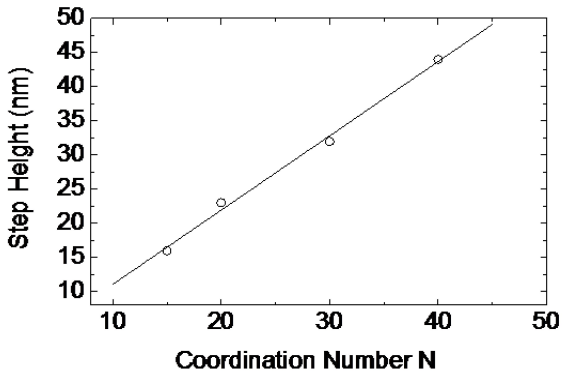


Figure 3.3.7.: Layer thickness in dependence of the coordination number, i.e. the number of deposition cycles. A linear fit to the data results in a specific layer thickness of 1.08 nm per deposition cycle. This figure was provided by Dr. Florian von Wrochem and is published in [9].

mobility of $\mu_{exp} = 0.1\,\mathrm{cm}^2/\mathrm{Vs}$, not far from the theoretically derived estimate for the electron mobility. For details on the derivation of the experimental mobility, see [9, 177].

3.4. The Electronic Structure of HKUST-1

3.4.1. From 1D to 3D: Metal-Organic Frameworks

Organic molecules have a variety of interesting electronic properties and their encapsulation into established inorganic materials presents a promising approach towards the development of novel materials. In this context, formation and electronic properties of one-dimensional metal-organic wire-like structures were investigated in the previous section. Another approach that was widely studied in the past decade is the construction of three-dimensional periodic structures by repetitively coordinating organic molecules to metallic clusters, similar to the formation of salt crystals via ionic bonds.

While these metal-organic frameworks (MOFs) are widely used, e.g. for gas storage or molecular separation [15–21], to date the binary combination of two material classes with completely different electronic properties resulted in insulating structures. The large number of possible material combinations and the modification of MOF properties by loading with guest molecules, however, open the prospect of light weight devices with tunable electrical properties. To overcome current limitations, fundamental understanding of the electronic properties of MOFs is essential.

In this section, the electronic structure of the well-known MOF Cu_3BTC_2 (HKUST-1) is analyzed. HKUST-1 consists of Cu-ions coordinated to BTC (1,3,5-benzene tricarboxylate) linker molecules, as depicted in Fig. 3.4.1. Monocrystalline HKUST-1 structures can be synthesized in form of surface mounted metal organic frameworks (SURMOFs) [24] and have been studied intensively in the past decade [178–181]. This work aims at a better understanding of the intrinsic electronic properties of HKUST-1 that are expected to play a

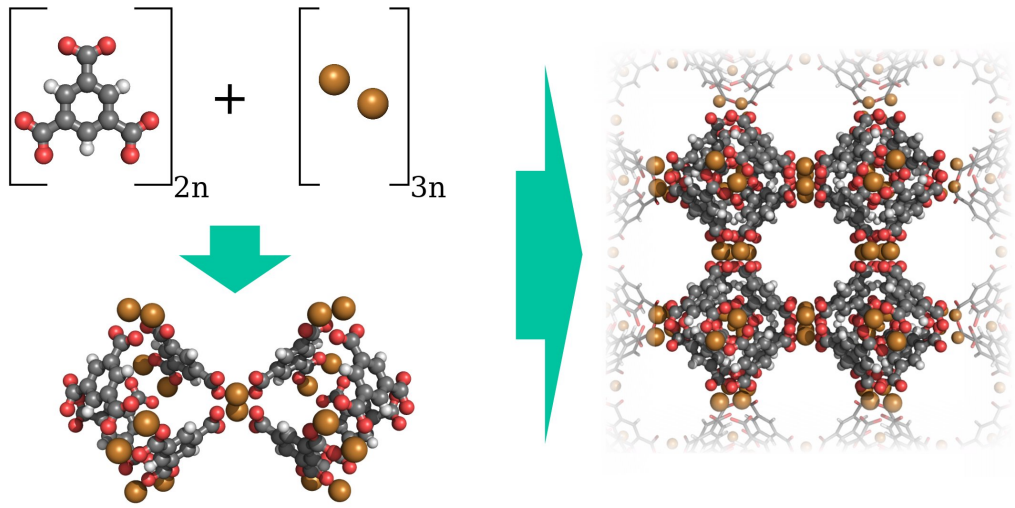


Figure 3.4.1.: Copper ions and $C_6O_6H_3$ (BTC) arrange in a paddle-wheel like arrangement. The combination of several paddle-wheel subunits leads to the formation of small cages, resulting in the metal organic framework Cu_3BTC_2, also known as HKUST-1.

role in the limitation of electric conductivity, such as band-gap, dispersion of bands and the role of open-shell orbitals.
Results presented in the following were already published in [182] and are presented here in full consent with all contributing authors.

3.4.2. Methods and Models - a Tricky Choice

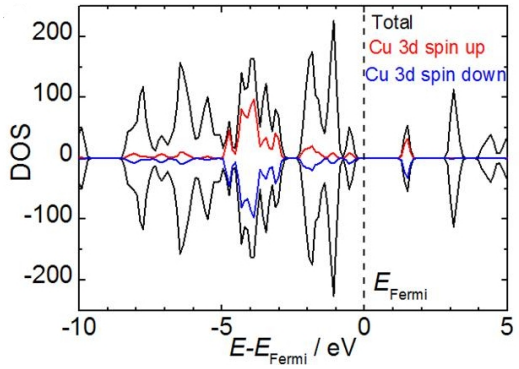
Density functional theory (DFT), presented in chapter 2, is an established method for the calculation of electronic properties of both organic molecules and metallic clusters. Depending on the system size and effects of interest, different DFT-based protocols can be applied. The plane wave approach with the projector augmented wave (PAW) method [183, 184] was developed for the calculation of periodic structures and can be applied to full MOF unit cells. However, such ground-state DFT methods using the local-density-approximation(LDA)- or generalized-gradient-approximation(GGA)-functionals are known to underestimate band-gaps because of the self-interaction error and it is difficult to treat open-shell orbitals accurately.
An extension to DFT for the calculation of excited states is time-dependent DFT (TD-DFT) [185]. While TD-DFT band-gaps are expected to be closer to experimental values than plane-wave results, computation of periodic structures is not possible and extensive computation times set a limit to the system size. This requires the restriction to a subset of the periodic system and it is unclear, if the calculated cluster excitation spectra are representative for the whole MOF unit cell.

In the following, we combine both approaches in order to minimize method dependent errors in the calculation of the electronic properties of HKUST-1. First, the complete MOF unit cell is studied using the plane wave approach with the PAW method for the calculation of band dispersion. Subsequently, excitation spectra of two different subsets of the HKUST-1 unit cell are calculated using TD-DFT at high level of accuracy to estimate the band-gap. The section is concluded by comparison of the theoretical results with experimental data.

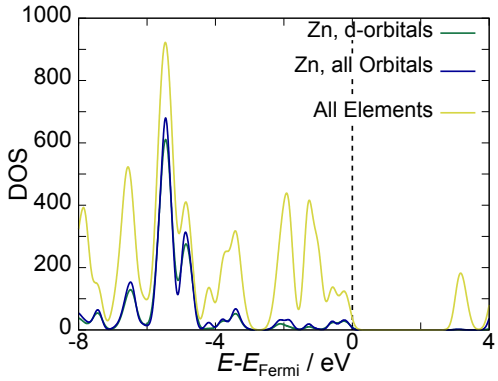
3.4.3. Density of States and Band Structures of Zn-HKUST-1 and HKUST-1

[4]In order to get an estimate of the electronic structure , the density of states (DOS) and the band structures along high symmetry lines in the FCC-unit-cell were calculated for two MOFs: the well-known HKUST-1 and a hypothetical HKUST-1 structure where Cu was substituted by Zn (Zn-HKUST-1 or $Zn_3(BTC)_2$). While Zn-HKUST-1 has not been realized experimentally yet, it is a useful model to highlight the influence of the open-shell character of the Cu-3d-orbitals on the electronic structure of HKUST-1.
The plane-wave package VASP [71, 72] with PAW-method [183, 184] was applied to calculate the electronic properties of periodic structures. The cutoff of the plane wave basis set was set to 550 meV. The calculation of the density of states was performed using a Γ-point calculation, i.e. a (1x1x1) Monkhorst and Pack k-point mesh [186]. The band structure was derived on 10 points along each line between the high symmetry points of the Brillouin zone of the Fm-3m-space group: the L-point (0.5, 0.5, 0.5), the Γ-point (0,0,0), the X-point (0,1,0) to the W-point (0.5,1,0). The Brillouine-zone and the points of high symmetry are displayed in the left inset of Fig. 3.4.2a. The Brillouin zone for the band-structure calculation was sampled on a (6x6x6) Monkhorst and Pack k-point mesh [186].
The total DOS of HKUST-1 as well as the contributions of the Cu-3d-orbitals are displayed in Fig. 3.4.2a. Spin-up and spin-down contributions are represented by opposite signs and different colors. The gap between the highest occupied orbital below the Fermi energy and the lowest unoccupied orbital above the Fermi level is approximately 2 eV. The highest occupied levels comprise states from both Cu-3d-orbitals and linker molecules, whereas the lowest unoccupied states are dominated by the empty Cu-3d orbitals. The second excitation at approximately 3.2 eV above the Fermi level comprises only contributions of the linker molecules.
In order to further analyze the role of the open-shell character of the Cu-3d-orbitals, the Cu-atoms were substituted by Zn. The resulting DOS is displayed in Fig. 3.4.2b. Again, the highest occupied levels just below the Fermi energy comprise both metal- and linker-states. The first excitation is approximately 3.2 eV above the Fermi level and shows no contribution of the Zn-orbitals. We note that DFT calculations are known to give inaccurate results for the band gap, but produce the right character of the electronic structure, allowing a direct comparison between HKUST-1 and Zn-HKUST-1 for the analysis of the role of the open-shell character of the Cu-3d-orbitals: The highest occupied states are quite comparable in both systems, i.e. they show contributions from both the linker molecules and the metal orbitals while being mainly dominated by the linker-orbitals. The second excitation of HKUST-1 and

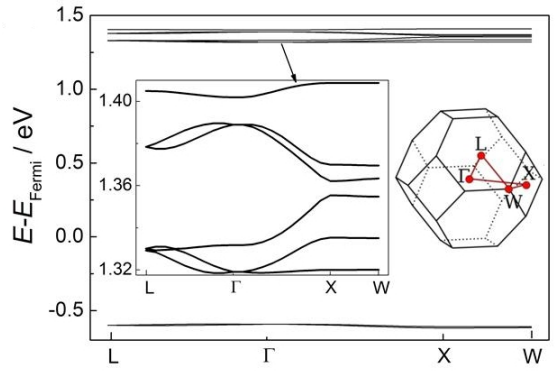
[4] Theoretical results of DOS and bandstructures were derived by Qiang Li (HKUST-1) and myself (Zn-HKUST-1).



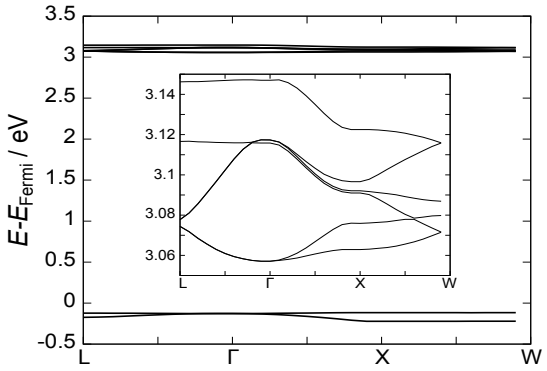
(a) DFT calculation of the total density of states and contribution of the Cu-3d orbitals of HKUST-1. The total DOS is displayed in black, the copper contributions are illustrated as red lines for spin-up and blue lines for spin-down. Spin-up and spin-down contributions are represented by opposite signs. Both contributions are virtually identical and show a bandgap of 2eV. The Fermi level was shifted to zero energy. The lowest-energy states in the conduction band are dominated by empty Cu-3d orbitals.



(b) DFT calculation of the total density of states and contribution of the Zn-orbitals of Zn-HKUST-1: The Zn-contribution to the DOS is displayed separately for all Zn-orbitals and the Zn-d-shells. A band gap of 3.2 eV is observed and the lowest excitation has no contributions from the Zn-orbitals, indicating that the Zn-3d-orbitals are fully occupied.



(c) Valence and conduction band of HKUST-1 along the high symmetry lines of the FCC-lattice. The Brillouin-zone and the high symmetry points are illustrated in the right inset. The bands show hardly any dispersion (left inset) due to the large size of the unit cell.



(d) The valence and conduction bands of Zn-HKUST-1 show a band-gap of 3.2 eV, in good agreement with the DOS in panel (b). The bands close to the fermi level are rather flat, as displayed in the inset.

Figure 3.4.2.: Calculation of the electronic structure of HKUST-1. In order to investigate the role of the open-shell character of the Cu-3d-orbitals, Cu was replaced by Zn in a hypothetical model system. Figure taken from [182].

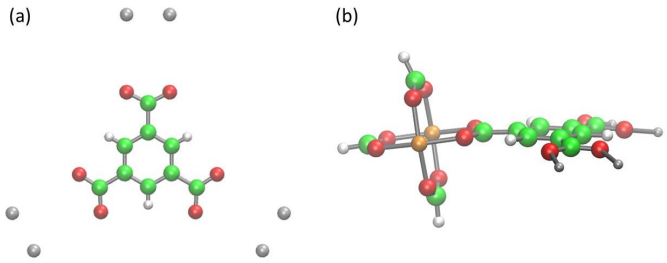


Figure 3.4.3.: As full TD-DFT treatment on a high level of accuracy is not feasible, two different clusters were extracted from the HKUST-1 unit-cell. (a) A BTC^{3+} linker ion, neutralized by three point charges, and (b) a Cu_2-O_8-paddle-wheel unit attached to a single BTC-linker. Figure taken from [182].

the first excitation of Zn-HKUST-1 basically coincide, indicating that the mid-gap states of HKUST-1 in Fig. 3.4.2a are indeed caused by the Cu-3d-open-shell orbitals. These orbitals are expected to play a major role in the electronic properties of HKUST-1 as they lower the band-gap by a factor two.
Subsequently, the band structures of both HKUST-1 and Zn-HKUST-1 were calculated along the high symmetry lines as described above and are displayed in Fig. 3.4.2c and Fig. 3.4.2d. HKUST-1 shows a bandgap of 2 eV in agreement with the calculated DOS. A close-up of the conduction bands is displayed in the inset. As can be expected for a very large unit cell of 26.3 Å, the bands show weak dispersion and a band-width of approx. 0.1 eV. The same holds true for the Zn-HKUST-1 system with a band-gap of approx. 3.2 eV and a band-width of 0.9 eV.

3.4.4. Excitation Spectrum of HKUST-1

[5]As the treatment of the full HKUST-1 unit cell using TD-DFT is not feasible with available resources, two different clusters were cut from the unit cell, as displayed in Fig. 3.4.3. The first cluster is the isolated BTC^{3-} linker-ion with six point charges of $q = +0.5$ elementary charges each to neutralize the system. The second cluster comprises the BTC-linker and a full Cu_2-O_8-paddle-wheel-unit.
The def2-TZVP basis set was used and different functionals were employed: the PBE functional was used for direct comparison to the plane wave calculations of the complete unit cell and, the B3-LYP and B3-LYP35 (with 35 % Hartree-Fock exchange) functionals [69] were applied for high accuracy.
For model (a), the excitation energies were calculated using the package TURBOMOLE [70] to be 3.41 eV (PBE), 4.41 eV (B3-LYP) and 4.93 eV (B3-LYP35)[69, 144]. While the PBE-result approximately matches the band-gap of the Zn-HKUST-1 and the second excitation of HKUST-1 of approx 3.2 eV, all three results exceed the band-gap value obtained for HKUST-1.

[5]Excitation spectra were calculated by Dr. Karin Fink.

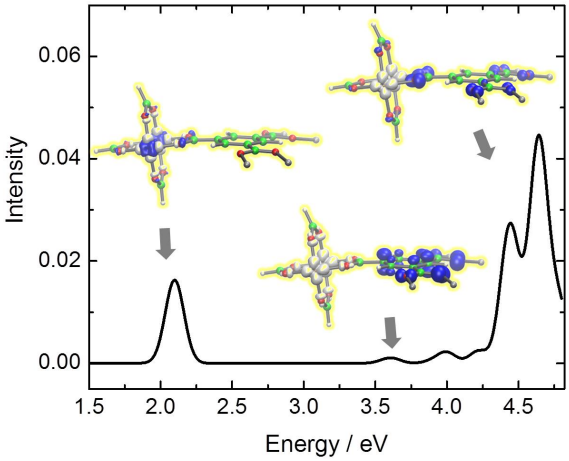


Figure 3.4.4.: Calculation of the excitation spectrum for a cluster extracted from HKUST-1, comprising a Cu_2-O_8-paddle wheel and a BTC-linker ion. The insets show the change of the electron density at the respective peaks in the excitation spectrum. Each excitation changes the electron density from the blue to the white regions. The first excitation is approximately 2 eV above the fermi level and is an excitation within the Cu-orbitals. In the excitations at 3.6 eV and 4.3 eV, the electron is excited from the BTC linker onto the Cu-atoms. The spectrum and density differences were visualized with the program PANAMA provided by Kühn and Weigend [187]. Figure taken from [182].

As the first excitation in the DOS-calculations of HKUST-1 is induced by the open-shell Cu-3d-orbitals (see Fig. 3.4.2a), Cu-atoms are included explicitly in the calculation of excitation energies in model (b), comprising paddle-wheel and BTC-linker. The resulting excitation spectrum is displayed in Fig. 3.4.4. The lowest excitation is at about 2 eV. The change in the electron density displayed in the insets shows that this excitation is a local d-d excitations at the individual Cu centers. However, these transitions are expected to have low intensities in optical measurements as d-d transitions in centrosymmetric systems are actually forbidden by the Laporte rule [188]. This excitation is thus not expected to be observed by ellipsometry measurements. The next excitations indicate energy gaps between 3.6 eV and 4.8 eV. They result from the transfer of electrons from the BTC linker molecule into the empty 3d orbitals of the Cu centers, as illustrated in the insets. This is in excellent agreement with the observations from the VASP calculations, where the first states above the Fermi level have the character of Cu-3d-orbitals.

3.4.5. Comparison with Experimental Measurements

[6]In order to verify the calculated electronic properties, the results are compared to ellipsometry measurements of HKUST-1 SURMOFs on a silicon substrate, performed with a spectral range from 0.7 to 5 eV, i.e. from 250 nm to 1800 nm. Ellipsometry characterizes the change of the polarization of light being reflected by a material using the amplitude ratio, Φ, and the phase shift, Δ. A Cauchy-model was fitted to the measured wave-length dependence of Φ and Δ, resulting in the dielectric function of the material. The imaginary part of the dielectric function then directly yields the wave-length dependent absorption of photons and thus the band gap. For background information, see [182].
The experimental results are displayed in Fig. 3.4.5. The imaginary part of the dielectric function in panel (c) shows an absorption of light with an energy above 3.6 eV.

In summary, we find that the calculation of the excitation spectrum for model (b) reproduces the measured band gap quite accurately with excitations starting from 3.6 eV. The lowest calculated excitation of Cu-d-d-states at 2 eV cannot be excited by light and is thus not visible in experiment. Calculation of model (a) resulted in excitations within the right order of magnitude between 3.41 eV and 4.93 eV. However, the result is strongly method dependent and experimental excitations cannot are not observed due to the lack of Cu-orbitals.
The periodic VASP-calculations yield excitations of 2 eV and 3.2 eV for HKUST-1 and 3.2 eV for the model Zn-HKUST-1. In HKUST-1, the lower excitation is, analogous to the cluster calculation, dominated by the Cu-3d-orbitals and thus not visible in the Zn-HKUST-1 calculation. While the absolute values differ by 20%, the general characteristics match experimental observations in terms of large band-gaps and flat bands. Furthermore, the electronic bands calculated for HKUST-1 and Zn-HKUST-1 are quite flat. This indicates that, even if electrons were excited into the valence band due to a smaller band-gap, conduction would be limited due to the comparably large effective mass of the quasi-particles.

3.5. HKUST-1 Loaded with TCNQ and F_4TCNQ

3.5.1. A Model for the Charge Transfer in Guest-MOF Systems

From Electronic Bands to Hopping Transport

The analysis of the intrinsic properties of HKUST-1 in the previous section showed a large bandgap of 3.2 eV and thus a basically insulating nature of the MOF. However, an increase of conductivity of HKUST-1 by orders of magnitude was observed upon loading the pores with small organic guest molecules [25]. The flat electronic bands around the Fermi-level of the intrinsic HKUST-1 suggest that the increased conductivity is induced by electrons or holes hopping between localized electronic states in the guest-MOF system. This is supported by prior experimental observations, which indicate that charge transport through HKUST-1 loaded with guest molecules is an activated process [25].

[6]Experiments were performed by Ovidiu Gordan, Zhigang Gu and Lars Heinke at the Institute für Funktionale Grenzflächen, KIT.

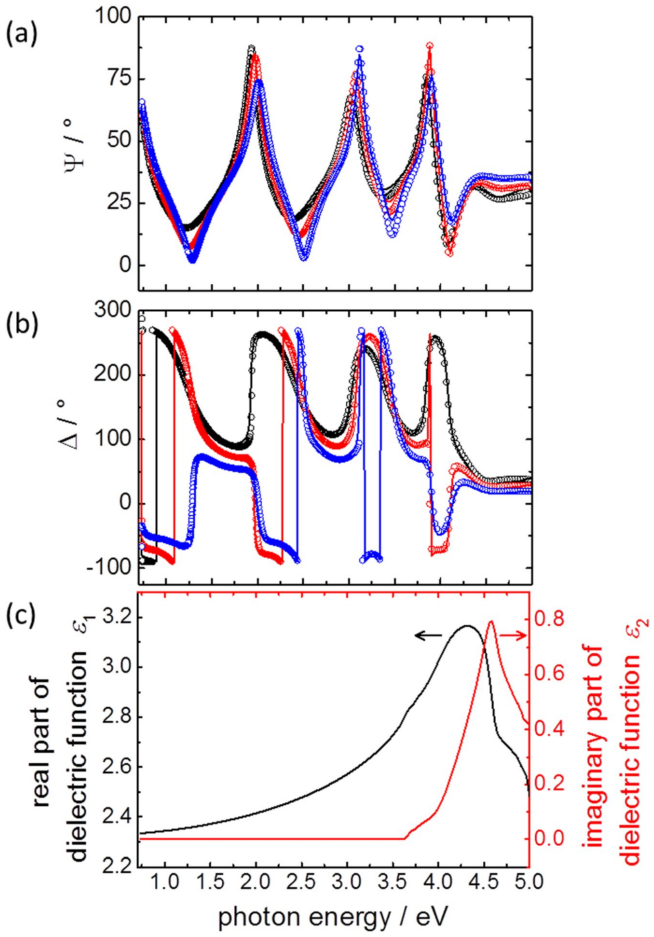


Figure 3.4.5.: Spectroscopic ellipsometry, the experimental data (circles) of Φ (a) and Δ (b) are fitted by a model (lines). Angles of incidence are 65 (black), 70 (red), and 75 (blue). (c) The real part (ϵ_1, black) and the imaginary (ϵ_2, red) part of the determined dielectric function of the HKUST-1 SURMOF. Figure and caption were provided by Ovidiu Gordan and Lars Heinke, as published in [182].

In this section, a model for the charge transfer through the guest-host systems HKUST-1 loaded with tetracyanochinodimethane (TCNQ) and tetrafluoro-tetracyanochinodimethane (F_4TCNQ) is presented. Marcus hopping rates are calculated and the influence of second order transfer integrals (molecular super-exchange) on charge carrier mobilities is analyzed using kinetic Monte Carlo (KMC) simulations.

Second Order Hopping Processes: Super-Exchange (SX)

In order to construct a model that explains the high measured current densities in HKUST-1 loaded with TCNQ and F_4TCNQ [25], possible hopping processes throughout the material need to be considered. In section 2.5, a fundamental model to describe hopping processes, the Marcus theory of charge transfer, was presented. In this theory, the hopping rate between initial state i and final state f is given by

$$\Gamma_{if} = \frac{2\,\pi\,|J_{if}|^2}{\hbar}\,(4\,\pi\,\lambda\,k_\mathrm{B}\,T)^{-1/2}\exp\left(-\frac{(\Delta E+\lambda)^2}{4\lambda k_\mathrm{B}\,T}\right) \tag{3.5.1}$$

J_{if} is the transfer integral between initial and final state, λ the reorganization energy and ΔE the difference between the molecular orbital energy levels of initial and final state. The transition rate is completely determined by those three parameters. For states that are not aligned in energy, the transfer rates decay exponentially with ΔE. The increase in conductivity of HKUST-1 upon loading [25] indicates either weak coupling between localized electronic states in the intrinsic MOF or level misalignment between HKUST-1 and injecting electrodes, whereas the guest states are close enough to the electrode work-functions to facilitate significant charge injection. In the first case, guest molecules induce additional hopping sites, increasing the overall electronic coupling. The second case indicates large gap between MOF and guest orbital energies and the charge transfer between guest molecules and HKUST-1 will be dampened exponentially by $\Delta E \gg k_\mathrm{B}\,T$. Transfer channels consisting of subsequent guest-MOF and MOF-guest hopping processes will thus not contribute significantly to the current. On the other hand, the nearest neighbor distance between guest molecules is approximately 1.3 nm (half size of the HKUST-1 unit-cell), leading to small electronic coupling J_{if} between guest molecules. Thus, direct guest-guest transfer processes cannot be responsible for large current densities either, and an additional transfer process is required to explain the experimentally observed conductivities [25]. This is illustrated in Fig. 3.5.1a.

One possible explanation for effective charge transfer between guest molecules is the relevance of second-order processes, where charge transfer between initial and final state occurs via virtual occupation of an intermediate state. As explained in section 2.5.3, the energy differences between initial and virtual state, $\Delta E_{i,v}$, and virtual and final state, $\Delta E_{f,v}$, modify the transfer integral J_{if}^{SX} as $1/\Delta E_{s,v}$ (for $s = i, f$), leading to $\Gamma \propto 1/\Delta E_{s,v}^2$, whereas only the energy difference between initial and final state enters the exponent in Eq. 3.5.1. Hence, good alignment between initial and final state, i.e. $\Delta E_{if} \approx 0$, may result high transfer rates. A detailed explanation of the super-exchange mechanism is presented in section 2.5.3.

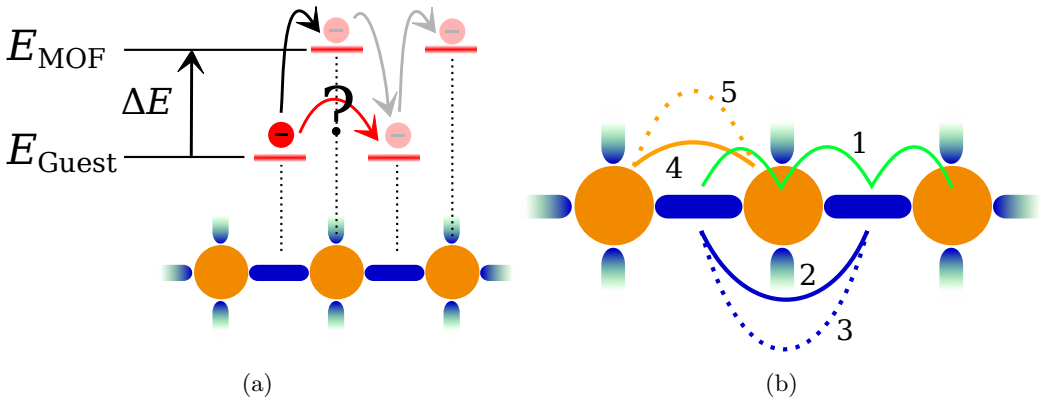


Figure 3.5.1.: Charge transfer processes in guest-MOF systems. (a) Large energy differences ΔE between MOF (orange discs) and guest (blue bars) orbitals and large distances between guest molecules induce limited first charge transfer between guest molecules (direct or as a subsequent guest-MOF-guest transfer). A different mechanism is necessary to explain observed current densities. (b) Overview of charge transfer processes in guest-MOF systems: First order transfer processes (solid lines) in a MOF-guest system are the subsequent hopping between guest and MOF sites (1, green) and the direct ransfer between guest-guest or MOF-MOF sites (2, blue; 4, orange). In order to explain high current densities in the MOF-guest system, additional second order-processes (super-exchange, dashed lines) are considered (3, blue; 5, orange).

3.5.2. LUMO and HOMO Alignment of TCNQ, F_4TCNQ and HKUST-1

In order to get a rough estimate on the characteristics of the hopping processes in the guest-MOF system, LUMO levels of TCNQ, F_4TCNQ and HKUST-1 were calculated using DFT as implemented in TURBOMOLE with the B3-LYP functional and the SV(P) basis set [144, 176]. Open shell calculations were performed to account for the antiferromagnetic nature of HKUST-1. As treatment of a complete HKUST-1 unitcell is not feasible with standard computational resources in a reasonable amount of time, calculations were performed on a subunit of the unit-cell, consisting of two paddle-wheel units (two Cu-atoms attached to eight O-atoms each) saturated by phenyl rings. The structures of TCNQ, F_4TCNQ and the MOF subunit are displayed in Fig. 3.5.2 alongside the orbital energy diagrams. Work functions of Pt and Al electrodes are included as an upper and lower bound for common metallic electrode work functions for reference.
The energies of the HOMOs of TCNQ, F_4TCNQ and the HKUST-1 subunit (each in vacuum) are $-7.40\,\text{eV}$, $-7.69\,\text{eV}$ and $-6.98\,\text{eV}$ respectively. The energies of the lowest unoccupied molecular orbitals (LUMOs) were determined to be $-4.89\,\text{eV}$, $-5.33\,\text{eV}$ and $-3.47\,\text{eV}$ for TCNQ, F_4TCNQ and the HKUST-1 subunit (each in vacuum) respectively. This trend is in rough agreement with values given in literature [189]. While the vacuum energy levels of molecular orbitals are in general shifted in electrostatic environment (i.e. the charge distribution of surrounding molecules), we assume that the order of the vacuum orbital levels of the guest molecules is maintained when loaded into the MOF.

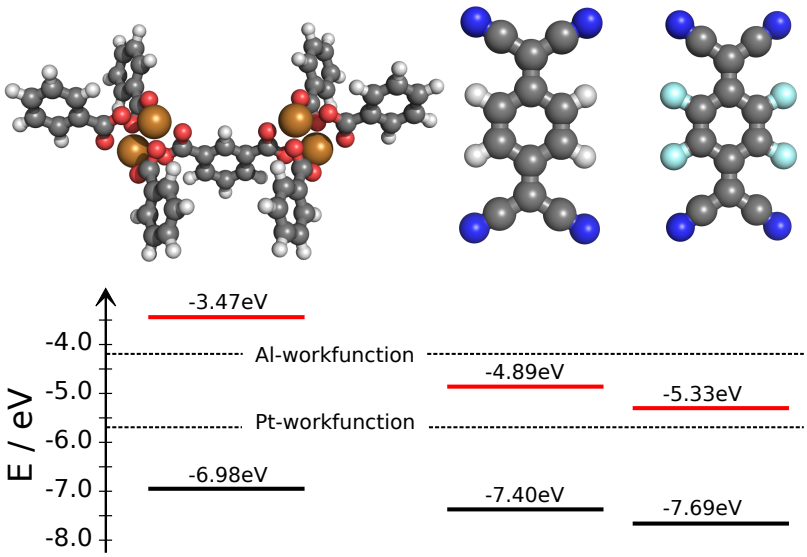


Figure 3.5.2.: Calculation of vacuum HOMO (black) and LUMO (red) energies for a HKUST-1 subunit (left), TCNQ (middle) and F4TCNQ (right). The work functions of commonly used electrode materials usually range from $-4.2\,\mathrm{eV}$ (Al) to $-5.7\,\mathrm{eV}$ (Pt), displayed by dashed lines for reference. The LUMO energies of TCNQ and F_4TCNQ are closest to the work functions, indicating electron transport via the guest LUMOs in the guest-MOF systems. Hole transport via the HOMOs is prevented by large injection barriers.

As work functions of commonly deployed electrode materials range from −4.2 eV (Al) to −5.7 eV (Pt) [190], hole transport is impeded by large energy differences between HOMOs and electrode work function. Further, the ΔE between HKUST-1-LUMO and guest-LUMO leads to an occupation of mainly guest orbitals, giving rise to the assumption that charge transport through the guest-MOF system is dominated by electrons transfer between guest molecules, using LUMOs of HKUST-1 as (virtual) intermediate state (see section 2.5 for details). Those assumptions are in agreement with observations reported in literature [189] [7].

3.5.3. Binding Mechanism of TCNQ and F_4TCNQ in HKUST-1

In order to accurately calculate microscopic parameters for charge transfer, namely the electronic coupling between the molecular orbitals, J, and the reorganization energies, λ, detailed information about the atomic structure of the guest-MOF system is essential. The configuration of the guest molecules in HKUST-1 was optimized using DFT as implemented in the package TURBOMOLE [70] using the B3-LYP functional and SV(P) basis set [144, 176]. Excitation calculations in the previous section indicated that limitation to a subset of the MOF unit cell provides representative results. Hence, only the relevant subset of the MOF unit-cell was used (Fig. 3.5.2). A molecular model of the HKUST-1 unit-cell was generated in accordance with experimental XRD-results. To account for the anti-ferromagnetic nature of HKUST-1, open shell calculations were performed. The anti-ferromagnetic state was verified using the resulting spin density. Three different optimized configurations as illustrated in Fig. 3.5.3 were analyzed, including previously proposed arrangements [25]. The first configuration (1) includes a guest molecule with all four nitrogen atoms are bound to a copper site. The second configuration (2), as proposed previously [25], comprises two covalent bonds between two nitrogen atoms of the guest molecules and HKUST-1. The third configuration (3) is an extension of configuration (2) and includes an additional molecule bound to the opposite side of the cavity, inducing pi-pi stacking between the two guest molecules near the cavity center.
The binding energy of each configuration was calculated according to

$$E_b = E_{bound} - E_{guest,vac} - E_{MOF,vac} \tag{3.5.2}$$

Here, E_{bound} is the energy of the optimized configurations (1),(2) and (3), and $E_{guest,vac}$ and $E_{MOF,vac}$ are the vacuum energies of the geometry optimized (F_4)TCNQ and the respective MOF-subset. For TCNQ in HKUST-1, we observed E_b of 1.97 eV, −0.99 eV and −1.25 eV per molecule for configurations (1), (2) and (3) respectively. The binding energies for F_4TCNQ in HKUST-1 were 1.84 eV, −0.91 eV and −1.18 eV per molecule for configurations (1), (2) and (3) respectively.
For both TCNQ and F_4TCNQ in HKUST-1, configuration (1), with bond formation between all nitrogen atoms and Cu-atoms, is energetically favorable by 0.72 eV and 0.65 eV. This

[7] Unfortunately, most recent studies imply that the charge transport is dominated by hole transfer between HOMOs [191]. The study presented in the following is based on the previous assumption of electron transport and the role of hole transport is currently being investigated.

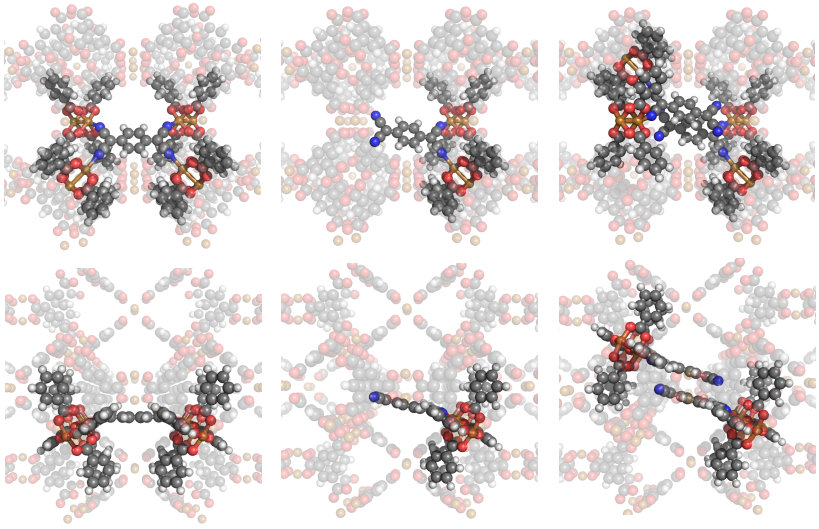


Figure 3.5.3.: Three possible binding mechanisms for TCNQ and F_4TCNQ in HKUST-1: Each nitrogen atom of the guest molecule is bound to a copper site (left panel, in the following denoted as configuration (1)), only two nitrogen atoms are bound to a copper site (middle panel, denoted as configuration (2)) and (previously proposed in literature [25]) guest molecules bound on opposite sides of the cavity induce pi-pi stacking in the middle of the cavity (right panel, denoted as configuration (3)). The geometries were optimized using DFT, for TCNQ and F_4TCNQ, resulting in binding energies for TCNQ in HKUST-1 of -1.97 eV, -0.99 eV and -1.25 eV per molecule for configurations (1), (2) and (3) respectively. Loading with F4TCNQ resulted in binding energies of -1.84eV, -0.91eV and -1.18eV per molecule for configurations (1), (2) and (3) respectively.

exceeds $k_B\,T$ at $T = 300\,\mathrm{K}$ by an order of magnitude and the system is primarily realized in configuration (1).

3.5.4. Electronic Coupling and Reorganization Energies

Subsequently, the electronic couplings J between individual hopping sites (LUMOs of guest molecules and HKUST-1) and the reorganization energies λ were calculated using configuration (1). The transfer integrals of LUMO orbitals was determined using the Löwdin orthogonalization method [141] based on orbitals calculated with the Quantum Patch method [3] using the def2-SV(P) basis and the BH-LYP functional, as implemented in TURBOMOLE [70, 145]. Nelson four point method was applied for the computation of λ [118, 143], using TURBOMOLE with a def2-TZVP basis and the B3-LYP functional [144, 145].
For the MOF-MOF transfer, values for J and λ between nearest neighbors and second nearest neighbors with distances of 0.93 nm and 1.32 nm, respectively, were calculated (MOF-LUMOs are centered on Cu-atoms of the paddle-wheel strucutres). For guest-guest molecules in adjacent cavities connected via a paddle-wheel, two arrangements were analyzed due to the symmetry of HKUST-1: one where the phenyl rings lie in plane ($\|$) and one where the phenyl

Table 3.5.1.: Electronic coupling between molecular orbitals in the systems TCNQ@HKUST-1 and F_4TCNQ@HKUST-1. Values for HOMO orbitals are included for comparison only. The coupling between the MOF-LUMO orbitals located on the copper sites were calculated for nearest and second nearest neighbors (1^{st} / 2^{nd}). Guest molecules in neighboring cavities connected to the same HKUST-1 paddlewheel can be aligned either parallel or orthogonal (defined by the normal vectors of the phenyl rings, $\perp/\parallel$).

J [eV]	HOMO	LUMO
TCNQ-HKUST-1	1.7·10-3	1.4·10-3
F_4TCNQ-HKUST-1	1.4·10-3	2.4·10-3
HKUST-1-HKUST-1 (1^{st} / 2^{nd})	6.0·10-4 / 7.3·10-6	5.7·10-2 / 3.0·10-5
TCNQ-TCNQ ($\perp/\parallel$)	2.0·10-6 / 4.3·10-6	1.7·10-8 / 4.1·10-9
F_4TCNQ-F_4TCNQ($\perp/\parallel$)	2.0·10-8 / 4.8·10-9	2.0·10-8 / 5.0·10-9

Table 3.5.2.: Reorganization energies λ for the hopping sites in HKUST-1 loaded with TCNQ and F_4TCNQ.

λ [eV]	HKUST-1	TCNQ	F_4TCNQ
HOMO	0.664	0.129	0.166
LUMO	0.580	0.266	0.271

rings are orthogonal to each other ($\perp$). The resulting values for J and λ are displayed in Tab. 3.5.1 and Tab. 3.5.2. Guest-MOF transfer was considered for nearest neighbors only. Couplings between LUMOs of guest molecules and HKUST-1 are comparable with $1.4\,10^{-3}$ for TCNQ and $2.4\,10^{-3}$ for F_4TCNQ. The direct coupling between guest molecules in neighboring cavities is orders of magnitude lower, $J = \mathcal{O}(10^{-8})$, and is thus expected to play a minor role in the charge transport of either system. The coupling between the LUMOs of HKUST-1, $5.7\,10^{-2}$, is quite large. However, level misalignment of the HKUST-1 LUMO with the electrode work functions and previous studies (see section 3.4) indicate injection barriers larger than $k_B\,T$ at room temperature and a weak contribution of MOF-MOF-hopping processes to the overall conductivity (except for very small energy differences between guest and MOF LUMOs). Values for HOMO orbitals are included only for reasons of completeness. The reorganization energies λ for the transfer between sites of the same type (i.e. MOF-MOF or guest-guest) are displayed in Tab. 3.5.2. For transfer between different types, the arithmetic mean is deployed. Notably, λ is of the same order of magnitude for both guest molecules but slightly larger for F_4TCNQ. Bearing in mind the dependence of Eq. 3.5.1 on J and λ, the higher value of λ for F_4TCNQ will be to some extent compensated by the higher coupling, J, indicating comparable mobilities for both guest-MOF systems.

3.5.5. Hopping Rates and Charge Carrier Mobilities

Using the electronic couplings J and reorganization energies λ derived above, transfer rates for charges hopping between neighboring guest sites were calculated using Marcus theory of

charge transfer (see section 2.5). The quantitative computation of HOMO-LUMO gaps with present state-of-the-art methods using (hybrid) (TD-)DFT approaches is difficult and derived electronic properties of MOFs have been shown to strongly depend on the applied method [182]. Hence, calculations of ΔE between MOF-states and guest-states with any of these methods result in rather large method, essentially XC-functional, dependent uncertainties. In order to demonstrate that our results are largely independent of this parameter we will treat ΔE as a free parameter in the following discussion. Accordingly, hopping rates and mobilities were calculated for a wide range $0\,\mathrm{eV} \leq \Delta E \leq 4\,\mathrm{eV}$.
As we assume the LUMOs of the guest molecules to be lower in energy than the HKUST-1 LUMO orbitals, the charge transfer is dominated by occupation of guest orbitals. The transfer rates of Marcus theory for hopping between guest molecules were calculated using the processes (1), (2) and (3) illustrated in the right panel of Fig. 3.5.1b. The rate of the successive process guest-MOF-guest is calculated via the inverse rates of the individual processes:

$$\Gamma^{-1}_{guest-MOF-guest} = \Gamma^{-1}_{guest-MOF} + \Gamma^{-1}_{MOF-guest} \tag{3.5.3}$$

The resulting rates calculated using Eq. 3.5.1 are displayed in the top panels of Fig. 3.5.4a for TCNQ in HKUST-1 and 3.5.4b for F_4TCNQ in HKUST-1. For small $\Delta E < k_\mathrm{B}\,T$ (for $T = 300\,\mathrm{K}$), there is significant probability of guest-MOF hopping and the successive transfer from guest to MOF and from MOF to guest is predominant due to the large couplings. The direct first order process between guest molecules is vanishingly small in comparison to the competing processes. For values of ΔE, where the process (1) is dampened exponentially by the energy difference between guest and MOF LUMOs, the super-exchange transfer between guest molecules (3) is mainly responsible for charge transfer through the system for both TCNQ and F_4TCNQin HKUST-1. In direct comparison, super-exchange transfer rates are about one order of magnitude larger for the F_4TCNQ-system than for TCNQ in HKUST-1, for the same ΔE.
For the calculation of the charge carrier mobilities, HKUST-1-supercells of $8 \times 8 \times 8$ unit cells were constructed. A single guest molecule can bind into the HKUST-1 pore in 12 different configurations of the same binding energy. To account for configurational disorder, each cavity was filled with two guest molecules with random arrangement, reproducing the configuration (1) displayed in Fig. 3.5.3. For the extended system, the Master equation was solved using kinetic Monte Carlo (KMC) simulations (see section 2.5) yielding values for the charge carrier mobility. Therefore, each pair of Cu-ions of a HKUST-1 paddle-wheel structure was mapped onto a single MOF-hopping site and the guest-hopping sites were placed in the middle of the phenyl rings of the guest molecule. Five independent model systems were constructed for both TCNQ and F_4TCNQ in HKUST-1 and the resulting mobilities were averaged. For each configuration, ten KMC simulations were performed with 20 electrons injected into the system and the electric field was set to $0.035\,\mathrm{V/nm}$. The electrons were distributed onto the orbital sites according to a fermi-distribution. The resulting mobilities in dependence of the energy difference ΔE between the LUMO orbitals of MOF and guest molecules are displayed in the bottom panels of Fig. 3.5.4a and Fig. 3.5.4b for TCNQ and F_4TCNQ respectively. KMC simulations were performed with and without the super-exchange process.

The resulting mobilities reflect the Marcus hopping rates, displayed in the respective upper panels: For small $\Delta E = 0\,\text{eV}$, the mobility is quite large with $\mu = \mathcal{O}(10^{-3})\,\text{eV}$ but drops quickly for energy differences of up to $\Delta E \approx 0.4\,\text{eV}$. This effect occurs in both KMC simulations with and without super-exchange, indicating that the large mobilities for small ΔE are a result of charge transport via guest-MOF-guest hopping and direct MOF-MOF-hopping. For $\Delta E > 0.4\,\text{eV}$, charge transfer is dominated by the super-exchange mechanism, as mobilities calculated with super-exchange ($\mu = \mathcal{O}(10^{-10})\,\text{eV}$) exceed mobilities without super exchange by three orders of magnitude ($\mu = \mathcal{O}(10^{-13})\,\text{eV}$).
In order to directly compare the results for both guest molecules, mobilities for F_4TCNQ were shifted by 0.6 eV along the x-axis, an estimate based on the energy difference of the vacuum LUMO orbitals of TCNQ and F_4TCNQ (Fig. 3.5.2). The results are displayed in the Fig. 3.5.4c. In the regime where the super-exchange transfer is the predominant process, mobilities of both systems are quite comparable. We note that a KMC simulation of the empty MOF was omitted as alignment of orbital energies and electrode work functions suggests that charges cannot be injected into the empty HKUST-1. This is supported by experimental findings of low current densities in pristine HKUST-1 structures.
Vacuum calculations suggest that the energy difference between LUMOs of MOF and guest molecules exceeds $\Delta E = 0.4\,\text{eV}$ and the calculation of rates and mobilities indicate that the super-exchange mechanism is the predominant charge transfer mechanism in HKUST-1 loaded with TCNQ and F_4TCNQ. Further, within the accuracy of the applied methods, Fig. 3.5.4c indicates comparable conductivity for HKUST-1 loaded with TCNQ and HKUST-1 loaded with F_4TCNQ.

3.5.6. Comparison to Experimental Data

[8]The comparable mobilities for both guest-MOF systems derived above contradict previously published experimental current densities [25], extensive experimental studies were performed. An overview of the results is presented in the following. Samples of single crystal SURMOF (surface mounted MOF) structures of HKUST-1 loaded with TCNQ and F_4TCNQ were prepared and I-V-characteristics were measured using the Hg-drop method. For each system, three different samples were prepared and currents measured at five different spots on the sample in dependence of the applied voltage. The resulting currents were averaged over all measurements and are displayed in Fig. 3.5.5. The loading of TCNQ and F_4TCNQ leads to an increase in the conductance of approx. four orders of magnitude. Notably, the TCNQ-HKUST-1 system and the F_4TCNQ-HKUST-1 system have similar experimental conductivity, which is in good agreement with the calculated mobilities.
In order to directly compare experimental data and theoretically derived electron mobilities, the following relation was applied:

$$\frac{J}{E} = \frac{J}{U/d} = \sigma = n\,e\,\mu \tag{3.5.4}$$

where σ is the conductivity, J the current density and the E the electric field that can be

[8]Experimental data was obtained by Jianxi Liu at the Institute für Funktionale Grenzflächen, KIT

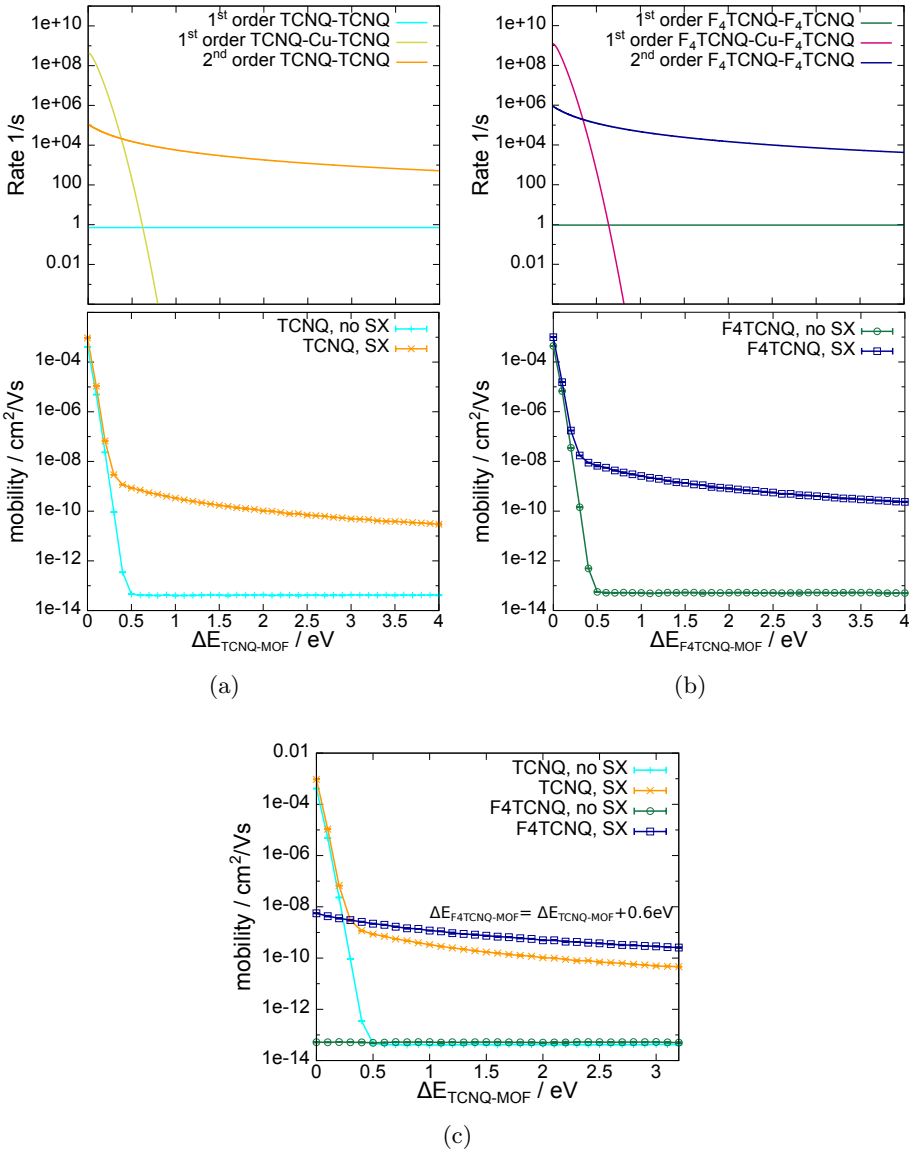
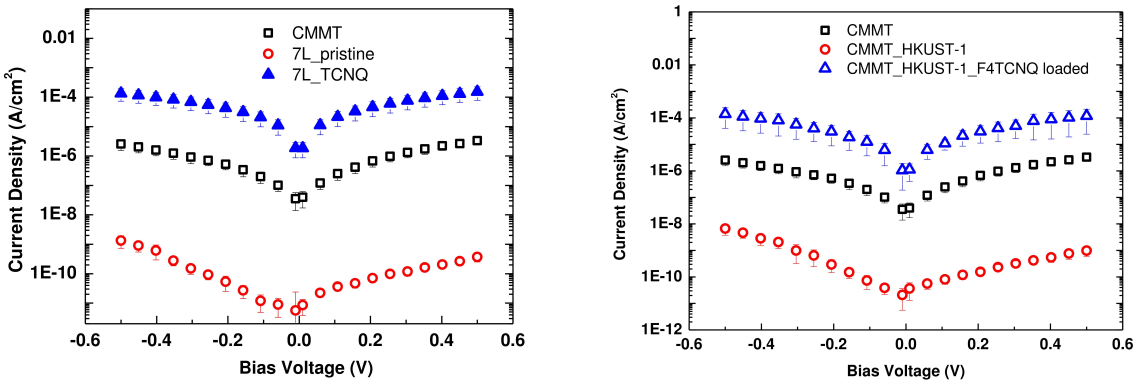


Figure 3.5.4.: Marcus transfer rates and electron mobilities of HKUST-1 loaded with TCNQ and F_4TCNQ as a function of the energy difference between guest and MOF LUMOs. Marcus rates are calculated for three different transfer processes between guest molecules: (1) the direct first order charge transfer between guest molecules, (2) the transfer rate of two successive first order charge transfers from guest molecule to HKUST-1 and from HKUST-1 onto the subsequent guest molecule and (3) the second order process (super-exchange) between guest molecules with virtual occupation of HKUST-1 orbitals. Marcus rates and mobilities for TCNQ (a) and F_4TCNQ (b) in HKUST-1 show similar characteristics: For low energy differences, rates and mobilities are dominated by the successive guest-MOF-guest transfer. For $\Delta E > 0.4\,\mathrm{eV}$, rates and mobilities including the super-exchange process exceed first order processes by orders of magnitude. (c) For direct comparison between both guest molecules, mobility of F_4TCNQ is shifted by 0.6 eV along the x-axis. Calculated Marcus rates and mobilities generated via KMC indicate that both guest-MOF-systems conduct equally well.



(a) (b)

Figure 3.5.5.: Experimental measured J-V-curves of (a) TCNQ and (b) F_4TCNQ loaded into HKUST-1. Red dots show the I-V of the pristine HKUST-1, the currents measured in the loaded systems are displayed as blue triangles. Both systems show an increase of approximately four orders of magnitude in the current upon loading. CMMT is used to connect HKUST-1 to the bottom electrode and pristine CMMT-curves are added for reasons of completeness. This figure was provided by Jianxi Liu.

calculated by voltage over device thickness. Conductivity σ and charge carrier mobility μ are connected via charge carrier density n and elementary charge e. At voltage of $U = 0.4V$, current density was measured to be $J = 10^{-4}\,\mathrm{A/cm^2}$ with a device thickness of $d \approx 50\,\mathrm{nm} = 50 \times 10^{-7}\,\mathrm{cm}$ (estimate from AFM measurements) for a 7-layer MOF-structure results in a conductivity of

$$\sigma_{exp} = 1.25 \times 10^{-7}\,\mathrm{S/m} \tag{3.5.5}$$

The maximum value of the electric field applied in the measurements is $E_{max} = U/d = 0.008\,\mathrm{V/nm}$ and of the same order of magnitude as the electric field in the KMC simulations ($0.035\,\mathrm{V/nm}$), allowing a direct comparison. Assuming one electron per unit cell, i.e. $n = 2.6343^{-1}\,\mathrm{nm^{-3}} = 3.8 \times 10^{20}\,\mathrm{cm^{-3}}$, Eq. 3.5.4 yields an experimental mobility of

$$\mu_{exp} = 1.42 \times 10^{-10}\,\mathrm{cm^2/Vs} \tag{3.5.6}$$

In comparison with the theoretically derived charge carrier mobility in Fig. 3.5.4 we observe the following: First, for $\Delta E_{guest-MOF} > 0.4\,\mathrm{eV}$, experimental mobility is of the same order of magnitude as the theoretically derived mobility including the super-exchange mechanism, exceeding the electron mobility without this additional second order process by three orders of magnitude. This indicates that super-exchange is the dominant process of charge transfer in the guest-MOF systems. An alternative explanation using first order processes only would require $\Delta E_{guest-MOF} \approx 0.3\,\mathrm{eV}$ for both TCNQ and F_4TCNQ, which can be excluded on the basis of the calculations above. Second, having identified super-exchange transfer as dominant charge transport mechanism in the combined guest-MOF systems, the experimentally

derived mobility in line with the $\Delta E_{guest-MOF}$ dependence of calculated mobilities indicates a lower bound for the disaligment of guest and MOF LUMOs of $1.0\,\mathrm{eV} < \Delta E_{guest-MOF}$. As TCNQ and F_4TCNQ are evidently aligned well enough with the work functions of the injecting electrodes to allow charge injection into the guest-MOF-systems, this indicates an injection barrier for an empty HKUST-1 of more than 1 eV. This barrier is consistent with the vacuum HOMO and LUMO energy calculations (see Fig. 3.5.2) and is potentially responsible for the insulating nature of HKUST-1.

3.6. Summary and Outlook

In this chapter, studies of three different types of (Metal-)organic electronic structures were presented. In section 3.2, the properties of amorphous organic semiconductors consisting of small molecules were analyzed. Atomistic morphologies of nine different materials were generated using classical force-field based methods and microscopic parameters were extracted in order to calculate charge carrier mobilities. For all nine materials, the resulting mobilities matched experimental values remarkably well. Furthermore, the energetic disorder, i.e. the width of the distribution of the orbital energies, was identified to be the limiting microscopic property towards high charge carrier mobilities.
In an approach to combine advantages of both organic and inorganic electronic structures, Fe^{II} redox centers were coordinated to organic terpyridine (TPT) units, resulting in extended, rigid one-dimensional metal-centered molecular wires (MCMW) in section 3.3. Force-field based Monte Carlo simulations of well-parametrized MCMWs resulted in a tightly packed arrangement that can be extended periodically. Charge carrier mobility was calculated on the basis of microscopic hopping parameteres derived *ab initio*. Within the approximations of the theroetical model, the calculated mobility showed very good agreement with experimental data.
Systems of even higher degree of symmetry were subject to the study of 3.4: The intrinsic electronic properties of the metal-organic framework (MOF) HKUST-1 were studied using DFT-based methods. Density of states, electronic bands and excitation energies were calculated, showing remarkable agreement with experimental data. In summary, flat electronic bands close to the Fermi-level due to the size of the unit cell and the large band-gap way beyond the thermal energy impede conductivity in pristine HKUST-1. Furthermore, the results highlight the impact of applied methods on accuracy of electronic properties. While correct characteristics can be reproduces using less accurate methods applied to full MOF unit-cells, quantitative predictions require computationally intensive TD-DFT calculations. However, expensive methods are limited to small entities of the unit-cells while yielding accurate quantities.
In order to increase conductivity in MOFs, HKUST-1 was loaded with TCNQ and F_4TCNQ molecules. Marcus transfer rates and charge carrier mobilities were calculated using geometry optimized guest-MOF structures. The impact of second-order processes (super-exchange) on the charge carrier mobility was studied by means of kinetic Monte Carlo. The results strongly suggest that super-exchange is the dominant charge transfer mechanism in the guest-MOF systems, leading to an increase of three orders of magnitude in the charge carrier mobility.

TCNQ and F_4TCNQ perform comparably well and absolute mobilities are in good agreement with experimental data for both materials [9].
In summary, charge carrier mobility in amorphous organic small-molecule layers is limited by large energetic disorder that is circumvented by constructing highly ordered structures such as MCMWs and MOFs. While MCMWs show remarkable mobilities that promise high PCE values, band-like conduction in MOFs is limited by flat bands and large band-gaps and charge hopping is prevented due to injection barriers. Loading of MOFs with small organic molecules can solve this issue by inducing additional hopping sites with energies close to the Fermi-level. Notably, we showed in this chapter that theoretical models are capable of predicting not only the correct characteristics of (metal-)organic materials, but also absolute charge carrier mobilities with a precision of up to one order of magnitude. Thus, the multiscale modeling approach is a promising toolkit towards *in silico* device simulations.

[9] Again we note that recent studies imply that the charge transport is dominated by hole transfer between HOMOs [191]. The presented results are based on the previous assumption of electron transport and the role of hole transport is currently being investigated.

4. Advanced Monte Carlo Based Algorithms

4.1. Limitations of Molecular Dynamics and Metropolis Monte Carlo

Computer simulations are a well established tool for the investigation of a variety of chemical and physical processes on the atomic, molecular or nanoscale [192]. Molecular dynamics (MD) or Monte Carlo (MC) simulations, where the interactions between the constituents of the system can be represented by classical force fields for many problems, can be applied for the simulation of time evolution of the system or its thermodynamic properties, respectively. Molecular simulations are often referred to as "virtual experiments" and are routinely applied to many problems, such as protein folding, morphology prediction of thin film or glassy materials [1, 28, 79–86]. With increasing computational power and continued efforts in the development of force fields and simulation protocols, molecular simulations have contributed significantly to the development of many scientific disciplines, ranging from astrophysics to biology and medicine.

Nevertheless there remain limitations to the system size and time-scale of the problems that can be sensibly studied with readily accessible computational hardware using these methods. Development of special-purpose computers, such as ANTON [37, 38], have shown that efforts to improve the range of applicability of molecular simulation methods can offer insights into complex phenomena that are presently difficult to access experimentally. For these reasons it remains both challenging and interesting to investigate the fundamental limitations of the present-day molecular simulation methods. The main limiting factor in molecular dynamics simulations is the inherently small time step that arises as a result of the discretization of Newton's equation of motion. For most generic molecular dynamics simulations with atomic resolution, the time steps of an individual integration step has been limited to the femtosecond-scale, despite extensive efforts in this field [193, 194].

Monte Carlo simulations are an alternative approach to compute thermodynamic expectation values for many problems. In the most widely used Metropolis MC approach [40], a Markov chain is generated by repeatedly applying a random modification to the system, usually called a "move", which is either accepted or rejected according to an acceptance criterion that generally satisfies detailed balance. As the random changes are not limited in size, MC simulations do not suffer from small configurational changes due to short time steps - their limitation lies in the move construction which severely limits MC based protocols in many applications. In principle, arbitrary modifications of the system can be proposed in a single "move", but as the acceptance rate typically falls exponentially with increasing energy (compared to the present state of the system) the acceptance rate for arbitrary modification of complex systems is typically exponentially small. As a result, MC protocols are generally

efficient for problems with smooth or discrete potentials or with weakly correlated degrees of freedom, but fail when collective effects play a role in the evolution of the system. It is generally possible to define “local” moves, which modify only a small fraction of the degrees of freedom of the system in a single step, but it is clear that such moves become inefficient for large system sizes.
Especially in condensed systems with thousands to hundreds of thousands of particles, the construction of collective moves with high acceptance rates is difficult. A notable exception are specialized Monte Carlo protocols, such as those proposed by Swendsen and Wang [195, 196] for discrete spin systems, which construct large-scale moves with zero energy change. For systems with continuous variables and complex potential functions (in particular with hard-core potentials), however, such methods are not generally available. In the absence of such moves the number of steps and computational effort required to generate an uncorrelated conformation, measured by the autocorrelation “time”, increases rapidly with the system size. This is the equivalent of the short time step in molecular dynamics, which results in a large number of energy/gradient evaluations to propagate the system as a whole along a relevant macroscopic reaction coordinate. In both MC and MD the computational time is dominated by the energy/gradient evaluation, and it should be noted that for most classical potentials the evaluation of the gradient is not significantly more expensive than the evaluation of the energy. As a result, the number of steps required to decorrelate a macroscopic variable is the measure of the computational efficiency of the method.
In MC methods one might therefore attempt to reduce the autocorrelation time by combining many uncorrelated “local” moves to a “collective” move before evaluating the energy. Unfortunately this straightforward approach fails, as the acceptance probability, which is then the product of the acceptance probabilities of the individual moves decreases rapidly, if there is an admixture of a few energetically unfavorable moves in a such a “collective” move (again typical for hard-core potentials). Generalized Monte Carlo algorithms, such as replica exchange [108], simulated annealing [107] or multiple try Monte Carlo [197, 198] increase the efficiency for specific applications but do not overcome the fundamental bottleneck described above.
Thus, both MD and MC based methods are limited: While MD is in principle capable of sampling relevant parts of the phase space, large-scale study of processes on the micrometer scale is not feasible on purchasable architectures. MC on the other hand is capable of arbitrary large step sizes and does not suffer from slow propagation. Large autocorrelation times, however, lead to low acceptance rates or uncorrelated moves, failing to capture relevant correlated effects.

4.1.1. Increased Acceptance and Correlated Moves

As described above, the construction of "local" or "correlated" moves with non-vanishing acceptance rates is one of the key challenges in the MC based simulation of condensed systems. On that account, there have been many propositions to increase the efficiency of Monte Carlo simulations since their introduction in the 1950s [199], for example the force biased move construction for the simulation of water [41, 200], cluster MC algorithms [201, 202] and approaches using feedback effects and generalized ensembles [203]. The fundamental principles

of many of these methods is exemplified by the famous Swendsen-Wang algorithm. Here, unit acceptance is achieved by constructing collective moves with a zero energy change that affect a large fraction of the total system. As this approach requires an efficient calculation of the exact energy change of the proposed move, it is only applicable for discrete systems with short-range interactions.
However, this special case of Monte Carlo simulations of discrete systems can be generalized by biasing the move construction with an energy estimator and changing all degrees of freedom simultaneously [41, 43, 44, 200]. This approach to efficient move construction was originally proposed by Rao and Pangali in 1979 [41] and later taken up in 1992 by Dereli et al. [42] and Timonova et al. for the modeling of diffusion and phase transitions [43]. The general idea was further investigated [44] and extended to associate a time scale with the MC algorithm [41, 204].
All these methods share the concept of biasing moves towards decreasing energy. Detailed balance is asserted by tuning the strength of this temperature-dependent bias. As the estimate of the energy change of the proposed move is not exact in general, the methods obey detailed balance only asymptotically in the limit of vanishing step size. Accuracy and efficiency of the method is dominated by the 'critical' step size that can be chosen for a particular system without a violation of detailed balance, that would result in inaccurate thermodynamic averages. In other words, the methods become computationally inefficient if the critical step size below which the violation of detailed balance may be tolerated is small compared to characteristic length scales. This problem is circumvented by other advanced MC methods such as the cluster algorithms proposed by Liu and Luijten [205] or Krauth [206]. However, these methods are not applicable to molecular simulations.
In the scope of this chapter, a generic Monte Carlo protocol [41, 200, 207], called Acceptance Rate Optimized Monte Carlo (AROMoCa) is developed. In AROMoCa the constructed moves are composed of many degrees of freedom and have acceptance rates close to unity while preserving detailed balance exactly. Two different approaches to overcome the limited step size in the biased move construction are analyzed.
The first approach is applicable to continuous systems and avoids the limitation of the 'critical' step size by applying an acceptance criterion that corrects the error of the estimate of the energy change. This way detailed balance is guaranteed for any step size. As explained in section 2.4, the application of an acceptance criterion impedes the over all acceptance probability with increasing number of degrees of freedom, resulting in vanishingly small acceptance rates for realistic systems with several thousand degrees of freedom. In order to limit the number of degrees of freedom without constructing uncorrelated moves, relevant parts of the phase space, i.e. degrees of freedom that are far from thermodynamic equilibrium, are identified. Changes are then applied to those degrees of freedom with higher probability, resulting in system changes with high acceptance rates.
In previously proposed protocols as well as in the first approach presented in the following, the change in energy induced by a MC step is estimated by the first order Taylor expansion, i.e. the atomic forces. The second approach includes second order Taylor expansion, thus improving the energy estimator quality and increasing acceptance rates significantly. Similar to the first approach, detailed balance is preserved by an acceptance criterion. The improved

energy estimator, however, allows the simultaneous change of all degrees of freedom in a single Monte Carlo step. Further, the second derivatives of the energy in different coordinates lead to direct coupling of degrees of freedom. This results in the construction of real correlated moves. While this approach is basically applicable to any type of system, this study focuses on the application to macro-molecular structures. Especially in multidimensional bonded systems such as peptides, proteins or polymers, simultaneous, correlated movement of several degrees of freedom is necessary in order to sample the relevant parts of the configuration space. Application of generic Metropolis MC requires the limitation to $\mathcal{O}(1)$ degrees of freedom that are changed simultaneously in a single MC step to achieve significant acceptance rates. While small proteins have been studied successfully using Metropolis MC based protocols [86, 208] applying changes to only a single degree of freedom per MC step, movement was limited to dihedral angles in these simulations. Furthermore, to avoid huge translations of the end of the chain upon the change of a single dihedral, step sizes have to be chosen vanishingly small. In order to efficiently model bonded units, correlated moves affecting also bonds and angles are necessary. Consider for example the case where only a small fraction of the back bone of a protein is flexible. A configurational change of this fraction that does not affect the positions of the initial and final atom requires the specific combination of changes in multiple coordinates. The construction of such a move via completely random generation of a set of independent coordinate changes occurs with vanishing probability. Hence, the acceptance rate would in fact be zero for generic Monte Carlo protocols. However, the application of the second order AROMoCa protocol allows the construction of correlated changes to all degrees of freedom while promising high acceptance rates and preserving detailed balance.

4.1.2. Biasing Molecular Simulations by Artificial Potentials

In other applications, such as the nucleation of molecular crystals, efficient simulation is not prevented by inefficient construction of single moves, but merely the fact that the area of the global minimum of the energy landscape occupies only a small fraction of the configuration space that is usually delimited by an energy barrier. The simulation time needed for a system to converge towards this subspace by mere diffusion in MD is again limited by the femtosecond time step, while large autocorrelation times in MC based protocols mentioned above prevent collective convergence of multiple degrees of freedom towards the global optimum.
However, there are systems where information about the structure is available (e.g. the arrangement of previously deposited molecules or the substrate configuration), on the basis of which the subspace of configurations for which the global potential is minimal can be estimated. By exploiting this information and proposing an educated guess, the propagation of the system can be biased towards configurations that are estimated to be energetically favorable, reducing the computational effort spent on the sampling of local minima and the diffusion towards the global optimum significantly. This can be exemplified by molecular crystals: for the formation of new crystalline layers on existing surfaces, information on preferred positions and orientations can be extracted by analyzing the geometry of the existing structure. A newly deposited particle can then be biased towards similar configurations. Especially in the case of molecular crystals, organic-anorganic or anorganig-anorganic inter-

faces, or the interface between crystalline and amorphous regions of a homogenous molecular structure, the phase space is increased by a factor 2^N (for a system of N particles) due to the rotational degrees of freedom. Therefore, geometry optimization upon diffusion is impeded. Obviously, pushing to system towards a desired final configuration via brute force is bound to result in unphysical structures for arbitrary bias potentials. In fact, the final configurations of a biased and an infinite unbiased simulation have to coincide. In other words, the bias is to induce a (maybe infinite) speedup of the system, not influence the resulting structures. A method to check the realization of thermodynamically reasonable states *on the fly* throughout the biased simulation is the Model Hopping (MH) approach proposed by Hansmann et al. [111]. MH is a variant of the Parallel Tempering (PT) method [45, 109, 110], where multiple models are simulated in parallel at different temperatures and the models are interchanged after a certain simulation time. Analogously, Model Hopping deploys bias potentials of different strengths to the different models, of which one is simulated using the physical force field only. Exchange between the models is attempted after a certain number of simulation steps and is accepted or rejected based on thermodynamic considerations. This acceptance criterion asserts the realization of only thermodynamically stable configurations in the unbiased model, i.e. states that would be realized anyway in infinite simulation time. This method can be combined with both MD and MC based protocols. When combining MH with MD, random diffusion in the phase space is superimposed by a drift towards the global optimum, reducing the overall simulation time. In MC based methods, biasing multiple degrees of freedom reduces the autocorrelation time tremendously, leading to collective, correlated system changes.

4.1.3. Structure of the Chapter

In sections 4.2, 4.3 and 4.4, an algorithm is presented that uses an estimator for the energy change of possible moves to construct collective moves with near unit acceptance probability. Two variations of this AROMoCa protocol will be presented: In the first version, a first order approximation to changes in the system energy is used to identify regions far from thermodynamic equilibrium and to construct collective moves with high acceptance rate. The second version of the algorithm deploys a more refined second order approximation of the change in energy, resulting in real correlated simultaneous changes of all degrees of freedom.
First, the detailed derivation of the AROMoCa algorithm is presented in section 4.2. The physical principles of AROMoCa are discussed and the construction of moves and acceptance criteria obeying detailed balance are derived in detail for both first and second order AROMoCa. For the latter, a detailed derivation of first and second order estimators in internal degrees of freedom is presented for the application to bonded macro-molecular systems. After a detailed derivation of the protocol, the first order AROMoCa is applied to four systems with increasing complexity in section 4.3 in order to demonstrate the effectiveness in comparison with MC and MD simulations. Additionally, the algorithm is compared to previously proposed methods such as force-biased-MC (fbMC) [43, 44]. Section 4.4 presents the analysis of the second order AROMoCa approach. A polymer-like model system is introduced and the performance of AROMoCa in comparison with generic MC and MD will

be studied in detail.
In the final section 4.5, the MH approach is briefly presented and possible choices for energy models are discussed. In order to get an estimate of the functionality of the algorithm, the influence of two main parameters, namely the choice of bias potential and number of models on the exchange probability is analyzed. Subsequently, in order to check its applicability to molecular systems, a model mimicking the growth of a molecular crystal is developed. A study of the transferability to realistic systems using the example of pentacene crystals concludes this section.
The algorithm based on the first-order energy estimator presented in 4.2 and the corresponding results of section 4.3 were already published in [5]. They are presented here in full consent with all contributing authors.

4.2. Methodology of Acceptance Rate Optimized Monte Carlo

4.2.1. Detailed Balance and Biased Move Construction

As mentioned above, many MC algorithms are based on the construction of Markov chains describing the evolution of a fictitious system where each new element depends only on its predecessor: In each extension of the chain a trial change to the system of interest (move) is proposed which is either accepted or rejected according to a method- and system-specific acceptance criterion. A sufficient but not necessary condition to compute proper thermodynamic expectation values is the detailed balance condition which postulates that the move construction and acceptance has to be chosen in a way that the total rate of states moving from initial state q_i into final state q_f, $\Gamma((q_i) \to (q_f))$ equals the rate of the inverse move from q_f to q_i. For a system following a Boltzmann distribution, which we will describe without loss of generality in the following, the transition rate from a state q_i to a state q_f is the product of the transition probability, $W((q) \to (q'))$, and the occupation probability for state q:

$$\Gamma((q_i) \to (q_f)) = W((q_i) \to (q_f)) \times \frac{1}{Z} \exp(-\beta E(q_i)) \tag{4.2.1}$$

where $E(q_i)$ is the potential energy of the state q_i, $\beta = 1/(k_\mathrm{B}\, T)$ is the inverse temperature and $Z = \sum_{q_i} \exp(-\beta E(q_i))$ is the partition function. The transition probability is the product of the probability to construct the move $\pi((q_i) \to (q_f))$, and to accept the move, $\rho((q_i) \to (q_f))$:

$$W((q_i) \to (q_f)) = \pi((q_i) \to (q_f))\, \rho((q_i) \to (q_f)) \tag{4.2.2}$$

A sufficient condition to reach thermodynamic equilibrium ($\Gamma((q_i) \to (q_f)) = \Gamma((q_i) \to (q_f))$) is the detailed balance criterion:

$$\frac{\pi((q_i) \to (q_f))\, \rho((q_i) \to (q_f))}{\pi((q_f) \to (q_i))\, \rho((q_f) \to (q_i))} = \exp\left(-\beta \Delta E\right) \tag{4.2.3}$$

with the change in energy induced by the proposed move, $\Delta E = E(q_f) - E(q_i)$.
In the case of completely uncorrelated random moves drawn from a given distribution (e.g. the cartesian coordinates, the orientation of a non isotropic particle, or internal degrees of

freedom) where the probability to construct the move is equal to the probability to construct the inverse move, $\pi((q_i) \to (q_f)) = \pi((q_f) \to (q_i))$, the probabilities of the move construction in Eq. 4.2.3 cancel out. From this Metropolis et al. [40] derived the Metropolis acceptance criterion:

$$\rho((q_i) \to (q_f)) = \begin{cases} \exp\left(-\beta \Delta E\right), & \Delta E > 0 \\ 1, & \Delta E \leq 0 \end{cases} \tag{4.2.4}$$

which is widely used for the simulation of physical and chemical systems. While the construction of such moves is conceptually and technically simple, it often leads to inefficient simulation protocols. Consider a system consisting of N particles with a total of $3\,N$ degrees of freedom. If only one or $\mathcal{O}(1)$ degrees of freedom are changed in one move, the algorithm will become inefficient in the thermodynamic limit and underperforms in comparison to MD simulations which change all degrees of freedom simultaneously. The evolution of the system in the thermodynamically relevant ensemble often requires collective changes of degrees of freedom which are exponentially unlikely to be proposed as a sequence of "local" moves. This problem arises both when the moves are proposed in sequence (which is often better in practice) or as collective moves composed of uncorrelated "local" moves. For many relevant potentials the acceptance rate will decrease exponentially with the number of uncorrelated "local" moves in the collective move. In either scenario, the acceptance probability is low and the autocorrelation time high.

If the existence of steep energy rises in local moves is limiting the autocorrelation time of the method, one should aim to construct collective moves that reduce the fraction of such events in the move construction. At this point it is important to note that the Metropolis algorithm is only one particular choice to realize the more general condition of Eq. 4.2.3. An alternative interpretation of Eq. 4.2.3 is to postulate an ideal algorithm where each proposed change in the configuration will be accepted, i.e. $\rho((q_i) \to (q_f)) = \rho((q_f) \to (q_i)) = 1$. In such an algorithm detailed balance must be fulfilled by constructing the probabilities π to select a move from all possible moves. The moves then need to be constructed such that

$$\frac{\pi((q_i) \to (q_f))}{\pi((q_f) \to (q_i))} = \exp\left(-\beta \Delta E\right) \tag{4.2.5}$$

To exactly satisfy this criterion, the energy difference ΔE of all considered moves has to be known for all possible moves before each MC step in order to compute the probabilities $\pi((q_i) \to (q_f))$. In the Swendsen-Wang family of methods this is possible due to the local nature of the Hamiltonian, but this type of approach is unrealistic for most systems with continuous degrees of freedom and long-range potentials. For processes with a time independent energy function one can assume $\Delta E((q_f) \to (q_i)) = -\Delta E((q_i) \to (q_f))$ (which is not true for example for metadynamics simulations [210–212] where an additional bias is applied to the system depending on its prior trajectory and the energy model thus changes with simulation time) and the relationship in Eq. 4.2.5 is met by proposing moves with a probability

$$\pi((q_i) \to (q_f)) = \frac{1}{A} \exp\left(-\beta \frac{\Delta E}{2}\right) \tag{4.2.6}$$

where $\Delta E = E(q_f) - E(q_i)$ is the change in energy by moving from initial state q_i to final state q_f and the normalization factor $A = \sum_{q_f} \exp\left(-\beta\frac{\Delta E}{2}\right)$ is calculated by summing over the moves to all possible final states q_f.
In order to illustrate this idea, let us consider an idealized system where the energy change ΔE is known for every possible move. The system consists of many particles where every particle has only two possible states, e.g. the adsorption of a gas of non-interacting particles on a surface. In this case there are only two possible states and moves (from the adsorbed state into the gaseous state and vice versa) for which the energy change is known exactly prior the move. Let us assume that the adsorbed state is energetically favorable by an energy $\Delta E_0 > 0$. Then, according to Eq. 4.2.6, the probability to adsorb a particle from the vacuum onto the surface is

$$\pi_a = \frac{1}{A} N_v \exp\left(\beta\frac{\Delta E_0}{2}\right) \tag{4.2.7}$$

while the probability to do the inverse move, the release of a particle, is

$$\pi_v = \frac{1}{A} N_a \exp\left(-\beta\frac{\Delta E_0}{2}\right) \tag{4.2.8}$$

Here, N_a and N_v are the number of particles in the adsorbed and vacuum state respectively and $A = N_v \exp(\beta\Delta E_0/2) + N_a \exp(-\beta\Delta E_0/2)$ is the normalization factor.
In the standard Metropolis Monte Carlo protocol a particle is selected at random, and a move to the other state is proposed and accepted or rejected according to Eq. 4.2.4. This leads to the rejection of most moves of particles which are already adsorbed on the wall, since $\Delta E = \Delta E_0 > 0$ for this move. The biased protocol on the other hand picks a particle according to the Boltzmann probability of the energy change induced by its change of state and every move is accepted.
For realistic systems, the computational costs to compute the exact change in energy for every possible move is unacceptably high. However, a suitable estimator for the energy difference $\widetilde{\Delta E} \approx \Delta E$ can be used to construct the move probabilities and accelerate the performance of the method. This estimator of the energy change for possible moves has to meet two criteria: First, it should be cheap in comparison with the real energy function. Second, the difference between estimated energy change and real energy change has to be smaller than $k_\mathrm{B}\,T$, i.e.

$$|\widetilde{\Delta E} - \Delta E| < k_\mathrm{B}\,T \tag{4.2.9}$$

Using this estimator, moves $q_i \to q_f$ leading to an estimated energy change of $\widetilde{\Delta E}$ are constructed with a probability according to Eq. 4.2.6, i.e.

$$\pi((q_i) \to (q_f)) = \frac{1}{A} \exp\left(-\beta\frac{\widetilde{\Delta E}}{2}\right) \tag{4.2.10}$$

with $A = \sum_{q_f} \exp\left(-\beta \frac{\widetilde{\Delta E}}{2}\right)$ analogously to Eq. 4.2.6. Considering Eq. 4.2.3, this results in

$$\frac{\rho((q) \to (q'))}{\rho((q') \to (q))} = \frac{\pi((q') \to (q))}{\pi((q) \to (q'))} \exp\left(-\beta \Delta E\right) \tag{4.2.11}$$

$$= \frac{A}{A'} \exp\left(-\beta \left(\Delta E - \widetilde{\Delta E}\right)\right) \tag{4.2.12}$$

This relation for the acceptance probabilities of move and inverse move clarifies the importance of a good quality energy estimator $\widetilde{\Delta E}$. For a perfect estimator of the energy, the exponent of Eq. 4.2.12 vanishes and the acceptance rate is close to unity.

4.2.2. First Order Energy Approximation

For systems with classical interactions (e.g. molecular or atomistic systems), the obvious approximation of the change in energy is the first order term of the Taylor expansion of the energy function in the generalized coordinates q_j, where j denotes the degree of freedom:

$$\widetilde{\Delta E} = \sum_j \frac{\partial E}{\partial q_j} \Delta q_j \tag{4.2.13}$$

For a system with only translational degrees of freedom, the change in energy can be expressed by the atomic forces $\boldsymbol{F}_j$ and displacements $\Delta \boldsymbol{x}_j$:

$$\widetilde{\Delta E} = \sum_j \nabla_j E|_{\boldsymbol{x}_j} \Delta \boldsymbol{x}_j = -\sum_j \boldsymbol{F}_j(\boldsymbol{x}_j) \Delta \boldsymbol{x}_j \tag{4.2.14}$$

Here, ∇_j is the derivative of the energy function in the coordinates of atom j. For the inverse move with $\Delta \boldsymbol{x'}_j = -\Delta \boldsymbol{x}_j$, we find:

$$\widetilde{\Delta E'} \approx \sum_j \nabla_j E|_{\boldsymbol{x}'_j} \Delta \boldsymbol{x}'_j = -\sum_j \nabla_j E|_{\boldsymbol{x}'_j} \Delta \boldsymbol{x}_j = \sum_j \boldsymbol{F}_j(\boldsymbol{x}'_j) \Delta \boldsymbol{x}_j \tag{4.2.15}$$

Inserting this into Eqs. 4.2.6 and 4.2.3 results in the modified acceptance criterion for the change of a single degree of freedom x:

$$\frac{\rho}{\rho'} = \frac{A}{A'} \exp\left(\frac{\beta}{2} \left(\nabla E|_x + \nabla E|_{x'}\right) \Delta x\right) \exp\left((-\beta \Delta E\right) \tag{4.2.16}$$

with A and A' being the sums over all probabilities for the moves and inverse moves. The acceptance criterion is modified, if more than a single degree of freedom is changed during one Monte Carlo step:

$$\frac{\rho}{\rho'} = \left(\frac{A}{A'}\right)^m \left[\prod_j^m \exp\left(\frac{1}{2} \left(\nabla_j E|_{\boldsymbol{x}_j} + \nabla_j E|_{\boldsymbol{x}'_j}\right) \Delta \boldsymbol{x}_j\right)\right] \exp\left(-\Delta E\right) \tag{4.2.17}$$

where m is the number of simultaneous system changes per simulation step. As the estimated energy change that is taken into account during the move construction (Eqs. 4.2.6 and 4.2.14) is corrected by the real change in energy, this acceptance criterion asserts detailed balance exactly.

As explained in the introduction to this chapter, the simultaneous random change of all degrees of freedom in large systems inevitably leads to vanishingly low acceptance rates in standard MC simulations. In fact, the acceptance rate is expected to drop exponentially with increasing number of simultaneous random changes. The considerations above are thus used to construct an algorithm based on the first order energy approximation that concentrates on relevant parts of the system: First, the regions of the system that are far from thermodynamic equilibrium are identified. Subsequently, moves are proposed to such regions with higher probability and the individual displacements are constructed according to Eq. 4.2.10. Most MD/MC simulations are performed for molecular systems and the derivation of the protocol will focus on these applications. For simplicity, only translational degrees of freedom are considered, but without loss of generality (e.g. a system consisting of interacting spherical particles). Extension to other degrees of freedom, such as rigid rotations or changes applied to internal degrees of freedom of molecules are straight forward by using torques and internal forces. Each Monte Carlo step is divided into four parts: First, for each degree of freedom (in our case: every particle and every direction) a probability to perform a displacement of any size between $-\Delta x_{max}$ and Δx_{max} is calculated:

$$\pi_{k,j} = \frac{1}{A_{k,j}} \int_{-\Delta x_{max}}^{\Delta x_{max}} dx \, \exp\left(\frac{F_{k,j}\, x}{2\, k_\mathrm{B}\, T}\right) \tag{4.2.18}$$

$$= \frac{1}{A_{k,j}} \frac{2\, k_\mathrm{B}\, T}{F_{k,j}} \left[\exp\left(\frac{F_{k,j}\, \Delta x_{max}}{2\, k_\mathrm{B}\, T}\right) - \exp\left(-\frac{F_{k,j}\, \Delta x_{max}}{2\, k_\mathrm{B}\, T}\right)\right] \tag{4.2.19}$$

using a normalizing factor:

$$A_{k,j} = \sum_k \sum_j \int_{-\Delta x_{max}}^{\Delta x_{max}} dx \exp\left(\frac{F_{k,j}\, x}{2\, k_\mathrm{B}\, T}\right) \tag{4.2.20}$$

Here k denotes the particle index and $j = 0, 1, 2$ the direction index for movement in x, y and z direction. The larger the force acting on a particle in a certain direction the larger $\pi_{ik,j}$ for this specific move. In a second step, $m \in \mathbb{N}$ moves are picked from the probability array constructed of all $\pi_{k,j}$. Neither are moves picked completely randomly nor are all degrees of freedom changed simultaneously. On the contrary: degrees of freedom that are far from equilibrium (having a large force and large value of $\pi_{k,j}$) are more likely to be changed than those near local equilibrium. After identifying the m different particle-direction combinations the displacement $\Delta x_{k,j} = \Delta x_{max} \times \xi_{k,j}$ is calculated by drawing $\xi_{k,j}$ from the probability distribution:

$$\pi(\xi_{k,j}) = \frac{1}{B} \exp\left(\frac{F_{k,j}\, \xi_{k,j}\, \Delta x_{max}}{2\, k_\mathrm{B}\, T}\right) \tag{4.2.21}$$

which is the equivalent of Eq. 4.2.6. Again there is a normalizing factor

$$B = \int_{-\Delta x_{max}}^{\Delta x_{max}} dx \exp\left(\frac{F_{k,j}\, x}{2\, k_\mathrm{B}\, T}\right) \tag{4.2.22}$$

for the probability $\pi(\xi_{k,j})$. B also appears in Eq. 4.2.19 and cancels out when calculating the total move probability as the product of $\pi_{i,j}$ and $\pi(\xi_{i,j})$, leaving A as only relevant normalization factor. Numerically the displacement $\xi_{i,j}$ is derived by drawing a random number η between 0 and 1 from a uniform random distribution and calculating $\xi_{i,j}$ using:

$$\Delta x_{i,j} = \frac{2\, k_\mathrm{B}\, T}{F_{i,j}} \ln\left(\eta\, \left(\exp\left(|\gamma_{i,j}|\right) - \exp\left(-|\gamma_{i,j}|\right)\right) + \exp\left(-|\gamma_{i,j}|\right)\right) \tag{4.2.23}$$

with $\gamma_{i,j} = \frac{F_{i,j}\, \Delta x_{max}}{2\, k_\mathrm{B}\, T}$. It can be shown that this results in the correct probability distribution in Eq. 4.2.21 for the displacement [43]. After displacing each (i,j)-combination the forces $\boldsymbol{F'}_{i,j}$ and the normalization factor A in Eq. 4.2.20 of the new configuration are calculated. The new configuration is then accepted according to Eq. 4.2.17.

4.2.3. Second Order Energy Approximation

Towards Correlated Moves

Approximating the energy change by the first order Taylor expansion limits the range where the estimator can be applied. Hence, in rugged landscapes, the deviation from the real energy change in comparison to $k_\mathrm{B}\, T$ is only small for small step sizes, i.e. step sizes of the order of a fraction of the typical interaction length scale (e.g. σ of the Lennard-Jones force-field). Hence, the method described above leads to a limited number of changes that can be applied to the system in a single MC step while having non-vanishing acceptance rates.
To overcome this limitation, it is straight forward to increase the quality of the estimator by including the second term of the Taylor expansion:

$$\widetilde{\Delta E} = \sum_k \frac{\partial E}{\partial q_k}\, \Delta q_k + \sum_{k,l} \frac{1}{2} \frac{\partial^2 E}{\partial q_k \partial q_l}\, \Delta q_k\, \Delta q_l \tag{4.2.24}$$

where k and l are indices of degrees of freedom. While for the first order estimator given in Eqs. 4.2.14 and 4.2.13 the individual degrees of freedom are completely decoupled, the second order term in Eq. 4.2.24 leads to a correlation between the coordinates and allows the construction of real correlated moves. This is neither possible using generic Metropolis MC nor AROMoCa deploying a first order energy estimator. The construction of moves $\{\Delta q_k\}$ with a probability $\pi(\widetilde{\Delta E})$ according to Eq. 4.2.10 however is not as straight forward as in the case of the first order approximation described above. The solution to this hurdle is a coordinate transformation into the eigen modes of the system.

Decoupling the Degrees of Freedom

Let $\boldsymbol{q} = (q_0, q_1, \ldots, q_n)^T$ be the vector comprising all n generalized coordinates q_k. The estimated change in energy is then of the form:

$$\widetilde{\Delta E} = \Delta \boldsymbol{q}^T \, W \, \Delta \boldsymbol{q} + \boldsymbol{b}^T \Delta \boldsymbol{q} \tag{4.2.25}$$

$\Delta \boldsymbol{q} = \boldsymbol{q}_f - \boldsymbol{q}_i$ is the vector of all changes of the generalized coordinates, i.e. the difference between final and initial states, f and i. The matrix W contains the second order derivatives and the vector $\boldsymbol{b}$ the first order derivatives:

$$b_k = \frac{\partial E}{\partial q_k} \tag{4.2.26}$$

$$W_{kl} = \frac{1}{2} \frac{\partial^2 E}{\partial q_k \partial q_l} \tag{4.2.27}$$

In general, the matrix W is not of diagonal shape. While the off-diagonal entries in W are in fact responsible for the desired correlation between coordinates, the construction of the moves in the basis of $\{q_k\}$ is not possible.
In order to decouple the degrees of freedom, two subsequent transformations are performed. First, a translation of the vector $\Delta \boldsymbol{q}$:

$$\Delta \boldsymbol{u} = \Delta \boldsymbol{q} + \frac{1}{2} W^{-1} \boldsymbol{b} \tag{4.2.28}$$

where W^{-1} is the inverse of the matrix W. This leads to a form of ΔE that does not contain a linear term of $\Delta \boldsymbol{q}$

$$\widetilde{\Delta E} = \Delta \boldsymbol{u}^T \, W \, \Delta \boldsymbol{u} + d, \qquad d = -\frac{1}{4} \boldsymbol{b}^T \, W^{-1} \, \boldsymbol{b} \tag{4.2.29}$$

In this form, the internal degrees of freedom are still coupled via the matrix W. According to the eigen decomposition theorem, a possible representation for any square matrix W is

$$W = PDP^{-1} \tag{4.2.30}$$

Here P is the matrix composed of the eigenvectors of W and D is the diagonal matrix comprising from the corresponding eigenvalues λ_k. Inserting Eq. 4.2.30 into Eq. 4.2.29 results in

$$\widetilde{\Delta E} = \Delta \boldsymbol{u}^T \, PDP^T \, \Delta \boldsymbol{u} + d \tag{4.2.31}$$

A final coordinate transformation

$$\Delta \boldsymbol{\xi} = P^T \, \Delta \boldsymbol{u} \tag{4.2.32}$$

decouples the internal degrees of freedom $\Delta \boldsymbol{\xi}$:

$$\widetilde{\Delta E} = \Delta \boldsymbol{\xi}^T D \Delta \boldsymbol{\xi} + d = \sum_k \lambda_k \, \Delta \xi_k^2 + d \tag{4.2.33}$$

Move Construction

Once the correlated degrees of freedom are decoupled via the coordinate transformations, a move can be constructed analogously to the algorithm presented in the previous section: The acceptance rate of moves approaches unity, if a certain move with change in energy of $\widetilde{\Delta E}$ is proposed with a probability of

$$\pi(\widetilde{\Delta E}) \propto \exp\left(-\frac{\beta}{2}\widetilde{\Delta E}\right) \tag{4.2.34}$$

Applying the second order estimator in the decoupled space given by Eq. 4.2.33 leads to

$$\pi(\{\Delta\xi_k\}) \propto \prod_k \exp\left(-\frac{\beta}{2}\lambda_k\,\Delta\xi_k^2\right) \tag{4.2.35}$$

The constant d disappears due to normalization and the probability distribution π is just a decoupled multidimensional Gaussian function with $\sigma_k = \sqrt{k_\mathrm{B}\,T/\lambda_k}$. The displacements $\Delta\xi_k$ following this distribution can be constructed by generating pseudo-random numbers r_k from a standard normal distribution and multiplying with σ_k:

$$\Delta\xi_k = r_k\sqrt{\frac{k_\mathrm{B}\,T}{\lambda_k}} \tag{4.2.36}$$

Reverse transformation of Eq. 4.2.32 and subsequently 4.2.28 result in the moves for the internal degrees of freedom, $\Delta\boldsymbol{q}$, with the proper probability given in Eq. 4.2.34. For a perfect energy estimator, this leads to unit acceptance.
While based on the same approach as the first-order estimator AROMoCa protocol presented in the previous section, this algorithm differs significantly in two aspects: First, there is real correlation between the degrees of freedom in the mixed derivatives of the matrix W. The first AROMoCa approach is merely a collection of individual moves with high acceptance probabilities. Second, the decoupling of the degrees of freedom via the coordinate transformation requires the simultaneous change of all degrees of freedom in a single step. Detailed balance demands that the inverse step is possible. However, as the eigenvectors of the matrix W change from step to step, the basis in which the step $\Delta\xi$ is constructed is different for move and inverse move. Changing only a single decoupled degree of freedom per step thus violates detailed balance, as the inverse move is impossible, if the matrices W of two subsequent steps differ. This can be exemplified using a two-dimensional model system: Let w.l.o.g. the orthogonal eigenmodes $\boldsymbol{m}_0$ and $\boldsymbol{m}_1$ of the system be aligned with x- and y-axis before a move, $\boldsymbol{m}_0 = (1,0)^T$ and $\boldsymbol{m}_1 = (0,1)^T$. Consider a MC move during which the system is changed only along a single eigenmode, e.g. a step of Δx along eigenmode $\boldsymbol{m}_0$, $(\Delta x, 0)^T$. In general, the eigenmodes after the move, i.e. the eigenmodes for the inverse move, are rotated by an angle α around the normal to the $x-y-$plane in respect to the original eigenmodes, resulting in $\boldsymbol{m}_0' = (\cos(\alpha), \sin(\alpha))^T$ and $\boldsymbol{m}_1' = (\sin(\alpha), \cos(\alpha))^T$. Obiously, any change of a single degree of freedom along either of the new eigenmodes will fail to construct the inverse move, $(-\Delta x, 0)^T$, thus violating detailed balance.

Algorithm Overview

The algorithm then works as follows:

1. For a given configuration, construct vector $\boldsymbol{b}$ and matrix W via first and second derivatives of the energy.
2. Calculate eigenvalues and eigenvectors of W
3. Construct the displacements of the decoupled coordinates, $\Delta\xi_k$ using Eq. 4.2.36
4. Transform the vector $\Delta\boldsymbol{\xi}$ into $\Delta\boldsymbol{q}$ via Eqs. 4.2.32 and 4.2.28.
5. Calculate the real change in energy, ΔE and accept the move with a probability that accounts for the error of the estimator:

$$\rho = min\left\{1, \frac{A}{A'}\exp\left(-\frac{\Delta E - \widetilde{\Delta E}}{k_\mathrm{B}\,T}\right)\right\} \tag{4.2.37}$$

Here, A and A' are norming constants of the probability distributions for the move construction and $\widetilde{\Delta E}$ is the energy estimator. Notably, if the function of the energy estimator changes between two subsequent steps, the acceptance criterion has to be adapted, as $\widetilde{\Delta E}' \neq -\widetilde{\Delta E}$. Eq. 4.2.11 then results in an acceptance criterion:

$$\rho = min\left\{1, \frac{A}{A'}\exp\left(-\frac{\Delta E - 0.5\left(\widetilde{\Delta E} - \widetilde{\Delta E}'\right)}{k_\mathrm{B}\,T}\right)\right\} \tag{4.2.38}$$

4.2.4. Bonded Systems and Internal Degrees of Freedom in AROMoCa

Internal Degrees of Freedom

While atomic forces can be applied as the first order estimator, the construction of the second order estimator is far from trivial. More precisely, the second order estimator needs to be constructed for every system specifically. In the following, the treatment of bonded macro-molecular systems is presented and the estimator functions for the systems studied in section 4.4 are derived in detail.
For bonded units, if rigid translation and rotation of the whole complex are neglected, the degrees of freedom q_i of the system can be mapped onto the space spanned by internal coordinates $\{\eta_i\}$: bond lengths r_k, angles α_l, and dihedral angles (short: dihedrals) ϕ_m. Different variables for the indices, k, l, m are used for clarification. An illustration of the internal degrees of freedom is displayed in Section 2.4, Fig. 2.4.1. The system energy comprises an internal contribution of the bond, angle and dihedral force-fields and external potentials accounting e.g. for electrostatic and Lennard-Jones interactions[1]. Additionally, solvent models

[1]Usually, the expression "external potential" is used to describe non-intrinsic force fields, i.e. potentials other than the interaction between the systems constituents. However, we will in the course context of AROMoCa refer to any potential other than the intra-molecular interactions (bond, angle and dihedral force fields) as "external potentials".

may be applied, resulting in a total energy of:

$$E_{tot} = E_{int}(\{\eta_i\}) + E_{ext}(\{\boldsymbol{x}_i\}) + E_{sol} \tag{4.2.39}$$

While the external and solvent contributions are easily expressed in cartesian coordinates, the internal contributions are usually calculated in terms of angles, bonds and dihedrals. The potentials for deflection of the internal coordinates from their rest positions are given by:

$$V_{b,k} = \frac{1}{2} k_{b,k} \left(r_k - r_{k,0}\right)^2 \tag{4.2.40}$$

for bonds,

$$V_{a,l} = \frac{1}{2} k_{a,l} \left(\alpha_l - \alpha_{l,0}\right)^2 \tag{4.2.41}$$

for angles, and

$$V_{d,m}(\phi) = k_{d,m} \left[1 - \cos(n(\phi_m - \phi_{m,0}))\right] \tag{4.2.42}$$

for dihedrals. k, l and m denote the bond, angle and dihedral indices respectively, the index 0 the value of the degrees of freedom in the rest position. The parameter $n \in \mathbb{N}$ in the dihedral potential is called multiplicity and depends on the orbital configuration of the atoms defining the rotation axis. While the dihedral potential is often parametrized differently:

$$V_d(\phi) = k_d \left(1 + \cos(n\,\phi - \delta)\right) \tag{4.2.43}$$

we find the equivalent parametrization of Eq. 4.2.42 to be more convenient and intuitive. For the harmonic potentials of bond and angle contributions, the change in energy for a coordinate change $q \to q'$, $q = \eta - \eta_0$ is:

$$\begin{aligned}
V_{harmonic} =& \frac{1}{2}\, k\, q^2 && (4.2.44)\\
\Delta V_{harmonic} =& \frac{1}{2}\, k\, \left[q'^2 - q^2\right] && (4.2.45)\\
=& \frac{1}{2}\, k\, \left[(q + \Delta q)^2 - q^2\right] && (4.2.46)\\
=& \frac{1}{2}\, k\, \left[\Delta q^2 + 2\, q\, \Delta q\right] && (4.2.47)\\
=& \frac{1}{2} \frac{\partial^2 V_{harmonic}}{\partial q^2}\, \Delta q^2 + \frac{\partial V_{harmonic}}{\partial q}\, \Delta q && (4.2.48)\\
\equiv& \Delta q\, A_{harmonic} \Delta q + b_{harmonic}\, \Delta q && (4.2.49)
\end{aligned}$$

Notably, this change in energy equals the second order Taylor expansion and the estimator for the energy change as given in Eqs. 4.2.24 and 4.2.25 is exact, leading to unit acceptance for bonds and angles.

For the dihedral potential, the second order energy estimator is only an approximation. In the middle between a minimum and a maximum of the dihedral potential given in Eq. 4.2.42, i.e. at $\phi - \phi_0 = \frac{2\,\pi}{4\,n}$, the second derivative is zero, leading to singularities in the inverse

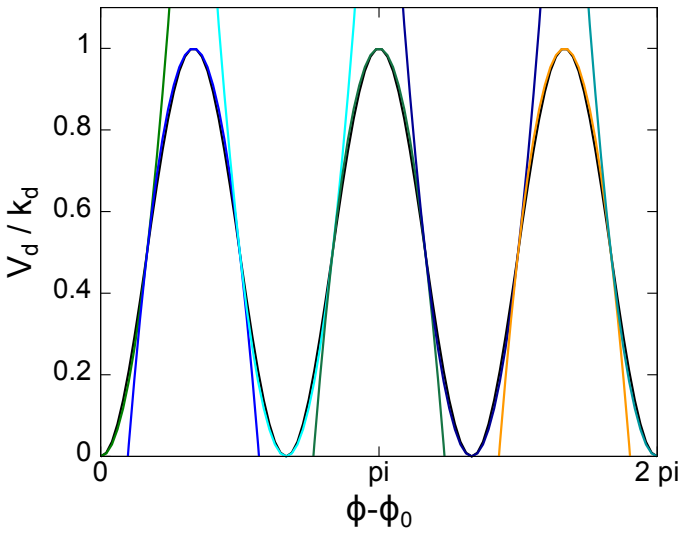


Figure 4.2.1.: Approximation of the cosine-dihedral potential by parabolas defined on intervals: The second derivative of the cosine-potential vanishes at the point of inflection, leading to problems in the move construction. This can be prevented by approximating the potential on parabolas defined on intervals between the points of inflection.

interaction matrix W^{-1} in Eq. 4.2.28. A different approach that circumvents singularities is the piece-wise harmonic approximation of the dihedral potential:

$$\widetilde{V_d} = \frac{1}{2}\,\tilde{k}_d\,(\phi - \widetilde{\phi_0})^2 + c\,k_d \tag{4.2.50}$$

Here, the approximated harmonic constant $\tilde{k}_d$ is given by

$$\tilde{k}_d = \pm\frac{8\,n^2}{\pi^2} \tag{4.2.51}$$

and $c = 0, 1$ depending on whether ϕ is in a segment of positive or negative curvature. $\widetilde{\phi_0}$ is the middle of the interval on which the approximation is defined, i.e. a minimum or a maximum of the dihedral potential. This is illustrated in Fig. 4.2.1. As the internal potential E_{int} is the sum over the contributions of the individual internal degrees of freedom, $\{\eta_i\}$, there is no coupling between the internal coordinates. The matrix W_{int} for the internal

potential is of diagonal form:

$$W_{int} = Diag(\{\alpha_i\}) = \begin{pmatrix} k_{b,0} & 0 & \dots & & & & & 0 \\ 0 & k_{b,1} & 0 & \dots & & & & \\ \vdots & 0 & \ddots & & & & & \\ & \vdots & 0 & k_{a,0} & 0 & & & \\ & & & 0 & k_{a,1} & 0 & \vdots & \\ & & & & & \ddots & 0 & \vdots \\ & & & & \dots & 0 & \tilde{k}_{d,N-4} & 0 \\ 0 & & & & & \dots & 0 & \tilde{k}_{d,N-3} \end{pmatrix} \tag{4.2.52}$$

For a system consisting of N atoms, the entries for the vector $\boldsymbol{b}$ are calculated according to

$$\boldsymbol{b}_{int,k} = k_{b,k}\,(r_k - r_{k,0}) \tag{4.2.53}$$

for $N-1$ bonds ($k = 0, \dots, N-2$),

$$\boldsymbol{b}_{int,l} = k_{a,l}\,(\alpha_l - \alpha_{l,0}) \tag{4.2.54}$$

for $N-2$ angles ($l = N-1, \dots, 2\,N-4$), and

$$\boldsymbol{b}_{int,m} = \tilde{k}_{d,m}\,(\phi_m - \widetilde{\phi_{m,0}}) \tag{4.2.55}$$

for the $N-3$ dihedrals ($m = 2\,N-3, \dots 3\,N-7$) with $\tilde{k}_{d,m}$ as defined in Eq. 4.2.51 and $\widetilde{\phi_{m,0}}$ being either a maximum or minimum in the potential, i.e. the middle of the interval for which the parabolic approximation of the dihedral potential is defined.

Treatment of External Potentials

External potentials such as non-bonded or solvent force fields are usually expressed by distance dependent functions, i.e. they are defined in cartesian coordinates. In order to expand the simulation beyond the bonded force fields, the change of such external potentials for a system with $3N-6$ internal degrees of freedom (i.e. a chain of N atoms connected via single bonds) needs to be expressed in terms of internal coordinates $\{\Delta\eta_i\}$:

$$\Delta E_{ext} \approx \widetilde{\Delta E}_{ext}\,(\{\boldsymbol{x}_k = \boldsymbol{x}_k(\{\eta_i\})\})) \tag{4.2.56}$$

$$= \sum_{i=1}^{3N-6} \left(\sum_{k=1}^{N} \sum_{d=1}^{3} \frac{\partial E_{ext}}{\partial x_{k,d}} \frac{\partial x_{k,d}}{\partial \eta_i} \right) \Delta\eta_i \tag{4.2.57}$$

$$+ \frac{1}{2} \sum_{i,j=1}^{3N-6} \left(\sum_{k=1}^{N} \sum_{d=1}^{3} \frac{\partial E_{ext}}{\partial x_{k,d}} \frac{\partial^2 x_{k,d}}{\partial \eta_i\, \partial \eta_j} + \sum_{k,k'=1}^{N} \sum_{d,d'=1}^{3} \frac{\partial^2 E_{ext}}{\partial x_{k,d}\, \partial x_{k',d'}} \frac{\partial x_{k,d}}{\partial \eta_i} \frac{\partial x_{k',d'}}{\partial \eta_j} \right) \Delta\eta_i\, \Delta\eta_j \tag{4.2.58}$$

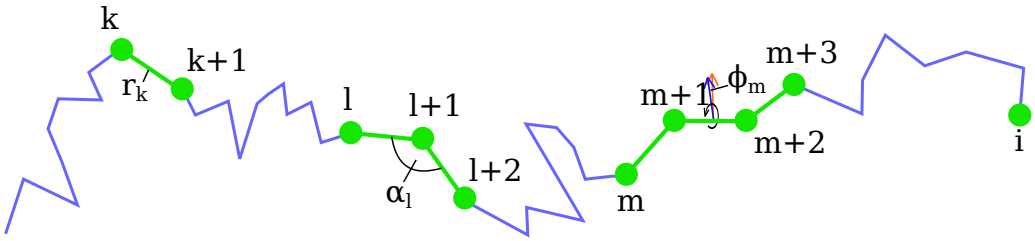


Figure 4.2.2.: The derivatives of the position of atom i with respect to internal coordinates is calculated by analysing the change of vectors from atom k to atom i, from atom $l+1$ to atom i, or atom $m+2$ to atom i for bonds, angles or dihedrals, respectively.

This results in the following entries for matrix W_{ext} and vector $\boldsymbol{b}_{ext}$ in Eq. 4.2.25:

$$W_{ext,i,j} = \frac{1}{2}\left(\sum_{k=1}^{N}\sum_{d=1}^{3}\frac{\partial E_{ext}}{\partial x_{k,d}}\frac{\partial^2 x_{k,d}}{\partial\eta_i\,\partial\eta_j} + \sum_{k,k'=1}^{N}\sum_{d,d'=1}^{3}\frac{\partial^2 E_{ext}}{\partial x_{k,d}\,\partial x_{k',d'}}\frac{\partial x_{k,d}}{\partial\eta_i}\frac{\partial x_{k',d'}}{\partial\eta_j}\right) \tag{4.2.59}$$

$$\boldsymbol{b}_{ext,i} = \sum_{k=1}^{N}\sum_{d=1}^{3}\frac{\partial E_{ext}}{\partial x_{k,d}}\frac{\partial x_{k,d}}{\partial\eta_i} \tag{4.2.60}$$

Notably, only by including external potentials, non-zero off-diagonal elements enter the matrix, leading to real correlated MC moves.
The derivatives of the cartesian coordinates with respect to internal coordinates, $\frac{\partial x_{k,d}}{\partial\eta_i}$, and $\frac{\partial^2 x_{k,d}}{\partial\eta_i\partial\eta_j}$, need to be derived specifically for any system of interest. In case of a chain of bonded atoms, the following notation will be used:

$\boldsymbol{x}_i$	position of atom i
$\Delta\boldsymbol{x}_i$	change of the position of atom i
r_k	length of bond k
$\boldsymbol{v}_k$	vector of bond k, between atoms k and $k+1$
$\hat{v}_k$	normal vector of bond k
$\boldsymbol{v}_{l+1,i}$	vector from atom $l+1$ at angle l to atom i
$\hat{n}_{l+1}$	normal vector of angle l at atom $l+1$
α_l	angle l
ϕ_m	dihedral angle m
$R(\hat{n}_{l+1},\alpha_l)$	Matrix for the rotation around normal vector of angle l by the angle α_l
$D(\hat{v}_{m+1},\phi_m)$	Matrix for the rotation around the bond $m+1$ by the dihedral angle ϕ_m

The indices are chosen such that ids for atoms, bonds, angles and dihedrals start with 0 each. E.g. the bond with id 0 is between atom 0 and atom 1, the angle with id 3 (i.e. the fourth angle) is at atom 4, i.e. the angle between the bonds 3 and 4. An illustration is given in Fig. 4.2.2. The derivatives can be extracted by considering the change of the position of any atom i upon a finite change of any other internal degree of freedom located at atom l.

Therefore, the position vector of atom i is expressed as the sum of the position vector of atom l and the vector between atoms l and i:

$$\boldsymbol{x}_i = \boldsymbol{x}_l + \boldsymbol{v}_{l+1,i} \tag{4.2.61}$$

which results in:

$$\Delta\boldsymbol{x}_i = \Delta\boldsymbol{v}_{l+1,i} \tag{4.2.62}$$

Let us now w.l.o.g. consider the change of angle l. The vector between the atom l and any subsequent atom i after the change is then given by:

$$\begin{aligned} \boldsymbol{v}'_{l+1,i} =& R(\hat{n}_{l+1}, \Delta\alpha_l)\, \boldsymbol{v}_{l+1,i} && (4.2.63)\\ =& \boldsymbol{v}_{l+1,i} + \frac{\partial}{\partial\alpha_l} R(\hat{n}_{l+1}, \alpha_l)|_{\alpha_l=0}\, \Delta\alpha_l\, \boldsymbol{v}_{l+1,i} && (4.2.64) \end{aligned}$$

resulting in the first derivative of the position of atom i in angle l:

$$\frac{\partial}{\partial\alpha_l}\boldsymbol{x}_i \approx \frac{\Delta\boldsymbol{x}_i}{\Delta\alpha_l} = \frac{\partial}{\partial\alpha_l} R(\hat{n}_{l+1}, \alpha_l)|_{\alpha_l=0}\, \boldsymbol{v}_{l+1,i} \qquad \forall \quad l < i-1 \tag{4.2.65}$$

The subsequent change of another angle m modifies Eq. 4.2.64 to be of the form:

$$\boldsymbol{v}'_{l+1,i} = \boldsymbol{v}_{l+1,i} + c_l\, \Delta\alpha_l + c_m\, \Delta\alpha_m + c_{lm}\, \Delta\alpha_l\, \Delta\alpha_m \tag{4.2.66}$$

and the first and second order derivatives can be extracted from the c_l, c_m and c_{lm}. Following this approach, first and second derivatives for a bonded system were calculated and are given in Tab. 4.2.1. A detailed derivation of the first order derivatives and the second order derivative for the case of angles is given in appendix C. Hence, the calculation of the derivatives is reduced to the calculation of rotation matrices and their derivatives:

$$\frac{\partial^a R(\hat{n}, \alpha)}{\partial\alpha^a} = (S(\hat{n}))^a\; R(\hat{n}, \alpha) \tag{4.2.67}$$

with the matrix S:

$$S\left(\hat{n} = (n_1, n_2, n_3)^T\right) = \begin{pmatrix} 0 & -n3 & n2 \\ n3 & 0 & -n1 \\ -n2 & n1 & 0 \end{pmatrix} \tag{4.2.68}$$

Where $R(\hat{n}, \alpha)$ is the matrix for a rotation around axis $\hat{n}$ by an angle α. While the computation of all derivatives leads to computational costs of $\mathcal{O}(N^3)$ for a system with N atoms, cutoffs can be applied to the estimator in order to reduce the number of derivative evaluations.

Table 4.2.1.: First and second order derivatives of the cartesian position of atom i in internal coordinates. Derivatives for indices not listed explicitly are zero.

coord(s)		derivative	index range
bond	$\frac{\partial \boldsymbol{x}_i}{\partial r_k}$	$= \hat{v}_k$	$\forall \quad k < i$
angle	$\frac{\partial}{\partial \alpha_l}\boldsymbol{x}_i$	$= \frac{\partial}{\partial \alpha_l} R(\hat{n}_{l+1}, \alpha_l)\|_{\alpha_l=0}\, \boldsymbol{v}_{l+1,i}$	$\forall \quad l < i-1$
dihedral	$\frac{\partial}{\partial \phi_m}\boldsymbol{x}_i$	$= \frac{\partial}{\partial \phi_m} D(\hat{v}_{m+1}, \phi_m)\|_{\phi_m=0}\, \boldsymbol{v}_{m+2,i}$	$\forall \quad m < i-2$
b.-b.	$\frac{\partial^2}{\partial r_k^2}\boldsymbol{x}_i$	$= 0$	$\forall \quad k$
a.-a.	$\frac{\partial^2}{\partial \alpha_l^2}\boldsymbol{x}_i$	$= \frac{\partial^2}{\partial \alpha_l^2} R(\hat{n}_{l+1}, \alpha_l)\|_{\alpha_l=0} \boldsymbol{v}_{l+1,i}$	$\forall \quad l = m < i-1$
	$\frac{\partial^2}{\partial \alpha_l \partial \alpha_m}\boldsymbol{x}_i$	$= \frac{\partial}{\partial \alpha_l} R(\hat{n}_{l+1}, \alpha_l)\|_{\alpha_l=0} \frac{\partial}{\partial \alpha_m} R(\hat{n}_{m+1}, \alpha_m)\|_{\alpha_m=0}\, \boldsymbol{v}_{m+1,i}$	$\forall \quad l < m < i-1$
d.-d.	$\frac{\partial^2}{\partial \phi_m^2}\boldsymbol{x}_i$	$= \frac{\partial^2}{\partial \phi_m^2} D(\hat{v}_{m+2}, \phi_m)\|_{\phi_m=0} \boldsymbol{v}_{m+2,i}$	$\forall \quad m < i-2$
	$\frac{\partial^2}{\partial \phi_l \phi_m^2}\boldsymbol{x}_i$	$= \frac{\partial}{\partial \phi_l} D(\hat{v}_{l+1}, \phi_l)\|_{\phi_l=0} \frac{\partial}{\partial \phi_m} D(\hat{v}_{m+1}, \phi_m)\|_{\phi_m=0}\, \boldsymbol{v}_{m+2,i}$	$\forall \quad l < m < i-2$
b.-a.	$\frac{\partial^2}{\partial \alpha_l \partial r_k}\boldsymbol{x}_i$	$= \frac{\partial}{\partial \alpha_l} R(\hat{n}_{l+1}, \alpha_l)\|_{\alpha_l=0}\, \hat{v}_k$	$\forall \quad l < k < i$
b.-d.	$\frac{\partial^2}{\partial \phi_m \partial r_k}\boldsymbol{x}_i$	$= \frac{\partial}{\partial \phi_m} D(\hat{v}_{m+1}, \phi_m)\|_{\phi_m=0}\, \hat{v}_k$	$\forall \quad m+1 < k < i$
a.-d.	$\frac{\partial^2}{\partial \alpha_m \partial \phi_m}\boldsymbol{x}_i$	$= \frac{\partial}{\partial \alpha_l} R(\hat{n}_{l+1}, \alpha_l)\|_{\alpha_l=0} \frac{\partial}{\partial \phi_m} D(\hat{v}_{m+1}, \phi_m)\|_{\phi_m=0}\, \boldsymbol{v}_{m+2,i}$	$\forall \quad l-1 < m < i-2$
	$\frac{\partial^2}{\partial \alpha_m \partial \phi_m}\boldsymbol{x}_i$	$= \frac{\partial}{\partial \phi_m} D(\hat{v}_{m+1}, \phi_m)\|_{\phi_m=0} \frac{\partial}{\partial \alpha_l} R(\hat{n}_{l+1}, \alpha_l)\|_{\alpha_l=0}\, \boldsymbol{v}_{l+1,i}$	$\forall \quad m < l-1 < i-2$

4.3. Application of the First Order AROMoCa to Model Systems

In the following a sequence of investigations of increasingly complex systems will be presented. Throughout these studies, advantages and limitations of the AROMoCa approach are analyzed and discussed in detail. The computational cost of a single step comprises different contributions depending on system parameters, and a comparison of the efficiencies of generic (Metropolis) Monte Carlo (GMC), molecular dynamics (MD) and AROMoCa in terms of computational cost proves difficult. Hence, the number of simulation steps will be used as a measure for efficiency, as this quantity is independent on system size or hardware specification. While the AROMoCa approach is expected to be slightly more expensive than GMC and MD in terms of a single simulation step, the major contribution to simulation time in molecular systems is the evaluation of the long-range particle-particle-interactions. Those interactions (forces and energies) have to be calculated for every step in all methods and is $\mathcal{O}(N^2)$, N being the number of particles in the system. The additional characteristic cost of the AROMoCa algorithm is a result of the evaluation of the estimated acceptance probabilities for the possible moves and is only $\mathcal{O}(N)$. As this is typically much smaller (<10%) than the cost of the energy/gradient evaluation, the number-of-steps comparison gives a good estimate of the performance of the AROMoCa approach. All simulations reported below were performed with a Python-based implementation of AROMoCa using the OpenMM package [92] for the calculation of forces between the atoms and system energies.

Non-Interacting Particles in a Double-Well Potential

In the first set of simulations the efficiency of AROMoCa with respect to the simultaneous movement of many particles at the same time was analyzed. As discussed above, the efficiency of MC as compared to MD decreases rapidly with the number of degrees of freedom for systems where the energy landscape is simple if only single particles are moved. Therefore, non-interacting particles in a double-well were investigated. Therefore $N = 10^4$ particles were initially placed with the positions uniformly distributed in an interval between $x_{min} = -10.0\,\mathrm{nm}$ and $x_{max} = 10.0\,\mathrm{nm}$ and the potential was given by

$$U_{DW} = c_0 \left(a\,x^4 + b\,x^2 + c\,x \right) \tag{4.3.1}$$

with $a = 0.0005\,/\mathrm{nm}^4$, $b = 0.04\,/\mathrm{nm}^2$ and $c = -0.05\,/\mathrm{nm}$ (see inset Fig.4.3.1a). The potential strength was set to $c_0 = 5\,k_\mathrm{B}\,T$

The system is a good application for a primary analysis of the AROMoCa performance, as the dynamics of the system comprises two phases: First, in the relaxation phase, each particle will quickly relax to its nearest local minimum, while the distribution of the particles amongst the minima (initially approx. 50:50) will remain far from the thermodynamic equilibrium of the system. Only in the second phase, on a much slower time scale, will the system equilibrate between the minima. This relaxation is characterized by the transfer rate of particles crossing the barrier from either left or right and the characteristic time scale is determined by the barrier height. In thermodynamic equilibrium the transition rates from right to left and from left to right will equal after equilibration for any method. However, methods differ in their absolute rates, by which efficiency can be measured. It is therefore a key goal of the development of accelerated simulation methods to increase the total rate, measured as function of the number of energy evaluations. This will control the accuracy of the determination of thermodynamic expectation values.

AROMoCa and GMC simulations were perfomed comprising 10×10^6 MC steps with a fixed displacement of 1.0 nm to either the right or left. The particles were equidistributed in the beginning between $x_{min} = -10.0$ and $x_{max} = 10.0$ and the temperature was chosen such that $k_\mathrm{B}\,T = 1.0\,\mathrm{kJ/mol}$. The height of the barrier in this system is $2.54\,k_\mathrm{B}\,T$ for particles in the shallower left minimum and $5.70\,k_\mathrm{B}\,T$ for particles in the deeper right minimum. The distance between the minima is approx. $\Delta x_{minima} = 13\,\mathrm{nm}$, hence 13 consecutive steps in the same direction are required to move a particle from one minimum to the other. AROMOCA simulations with collective moves for $m = 1$, 2, 4, 8, 16 and 32 independent particle displacements per step were performed. The average energy per particle as a function of the number of steps is displayed in Fig. 4.3.1a. The different slopes in the energy relaxation clearly indicate the two phases of the simulation. GMC and AROMoCa with $m = 1$ result in approximately equal energy relaxation, whereas for $m \geq 2$ AROMoCa converges much faster in energy (up to a factor of 20 for $m = 32$) than GMC.

The final distribution of particles was computed and compared to the Boltzmann distribution in the given potential. In the long-time limit both methods converge to the Boltzmann distribution (inset of Fig. 4.3.1b), but the AROMoCa simulations converge much faster, even though the particles are moved independently, i.e. there are no collective effects. Fig. 4.3.1b

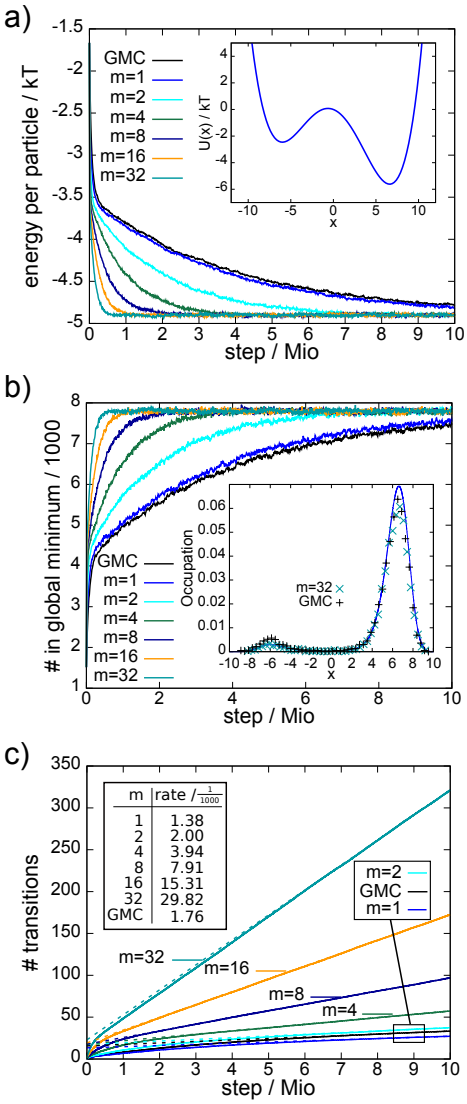


Figure 4.3.1.: Comparison of generic Metropolis MC (GMC) and AROMoCa simulations with $m = 1, 2, 4, 8, 16$ and 32 multiple displacements per MC step. a) Energy convergence in an one-dimensional double-well potential (inset). While AROMoCa with m=1 and GMC perform equally, equilibration accelerates with higher m and the equilibrium state is reached up to 20 times faster using AROMoCa with m=32 than in the GMC simulation. b) The number of particles that reach energies within $k_\mathrm{B}\,T$ of the global optimum converges faster in AROMoCa for larger values for m. The inset shows the final distribution function of the particles after the 10^7 MC steps in GMC and AROMoCa (with $m = 32$) simulations compared to the exact distribution function. c) As the particles were initially equidistributed between $x_{min} = -10.0$ and $x_{max} = 10.0$ the convergence is measured by the transition rate of particles across the barrier. The particle current across the energy barrier at $x = 0$ is shown as a function of the number of simulation steps. A linear function is fitted to the transition values of the last 2×10^6 MC steps to determine the transition rates (table inset). The results clearly show a strongly increased relaxation rate in the system for AROMoCa with $m \geq 2$ in comparison to GMC, leading to an overall increased performance. Figure and caption published in [5].

shows the fraction of particles that are within $k_B T$ of the global optimum. As for convergence of the energy per particle, plotted in Fig. 4.3.1a, AROMoCa with $m \geq 2$ outperforms GMC by far. As explained above, the total transition rate of particles crossing the barrier determines the sampling speed and hence the efficiency of the algorithm. The sum of the left-to-right and the right-to-left particle currents crossing the barrier at approximately $x = 0$ is shown in Fig. 4.3.1c. Linear fits to the last 2×10^6 MC steps result in the currents, displayed in the inset. As the transition rate of AROMoCa increases with nearly unit slope almost linearly with m, AROMoCa proves much more efficient for $m \geq 2$.
Recalling the AROMoCa algorithm, this result appears counter-intuitive at first: Particles that are not in equilibrium are much more likely to be selected and moved by AROMoCa than particles near the minimum. Furthermore, the move selection criterion (left or right, see Eq. 4.2.19) will favor relaxation to the nearest local minimum over crossing of barriers, which is in fact a sequecne of steps uphill. Nevertheless, AROMoCa performs much better than GMC because of its high acceptance rate (inset table Fig. 4.3.1c): On the one hand, the local relaxation (moving with the gradient of the potential) must take precedence over barrier crossing due to detailed balance being preserved by AROMoCa. On the other hand, relaxation dynamics is increased, as many particles are moved with near unit acceptance rate in every single step. This is demonstrated by direct comparison between AROMoCa with $m = 32$ with an MC protocol that moves 32 particles at random with the standard Metropolis acceptance rate: Here, the current across the barrier is approximately 30 times higher in AROMoCa in comparison to generic MC.

Liquid-Crystal Transition for Lennard-Jones Systems

Subsequently the analysis of AROMoCa performance was extended to a system of interacting particles. Therefore, AROMoCa was applied to a two dimensional Lennard-Jones model for Argon (Lennard-Jones parameters $\epsilon = 0.99601\,\mathrm{kJ/mol}$ and $\sigma = 0.3405\,\mathrm{nm}$) at $T = 50\,K$ which is below the melting point. As initial configuration, 526 particles were distributed randomly on a square surface of $9.5 \times 9.5\,\mathrm{nm}^2$. The system size was chosen based on initial estimates on the convergence time for GMC. The limitation to the two dimensional case was applied for easy visualization of domain growth and a 3D model is studied in the next section. Simulations of AROMoCa with $m = 1$, 2, 3, 4, and 5 were performed and compared to GMC. The maximal displacement was set to $\Delta x_{max} = 0.75\,\mathrm{nm}$ in all simulations. To evaluate the convergence of AROMoCa and GMC, the total system energy per step and the q_6 -bond-order parameter were monitored. The q_6-parameter is based on the parameter proposed by Steinhardt et al. [213], which is defined as:

$$q_l(a) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} \left| \frac{1}{n(a)} \sum_{NN(a)} f_c(r)\, Y_{lm}(\theta\phi) \right|^2} \tag{4.3.2}$$

Here, $\sum_{NN(a)}$ is the sum over the nearest neighbors of atom a, $n(a) = 6$ is the number of nearest neighbors taken into account, (r, θ, ϕ) are the spherical coordinates of the vectors from atom a to the nearest neighbors and Y_{lm} are the spherical harmonics. The cutoff

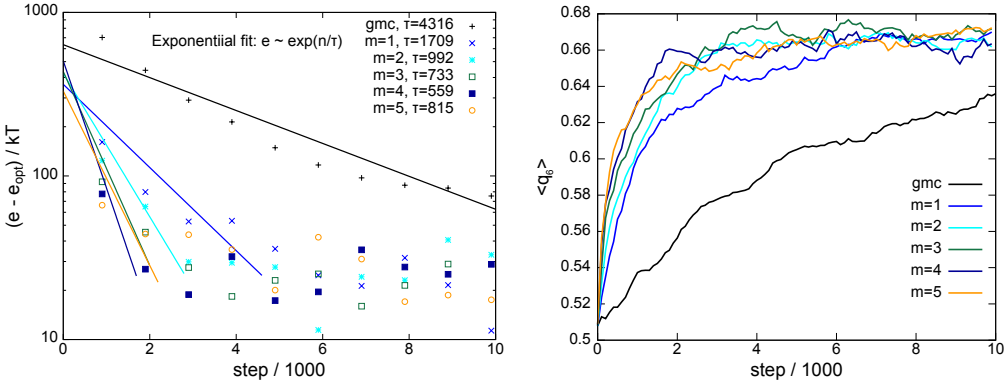


Figure 4.3.2.: Liquid-to-crystal transition of a two dimensional Lennard-Jones liquid. The convergence speed in terms of total system energy (left) and the q_6-bond-order parameter (see text) averaged over all particles (right) of generic Metropolis Monte Carlo is compared to AROMoCa using different m. Exponential fits of the decay in energy indicate that AROMoCa converges up to six times faster than generic MC. The rate of convergence increases with the number of simulataneous displacements per MC step m for up to $m = 4$. The trends of the q_6 -bond-order parameter verify that the total order of the system is increased througout the simulations faster in the AROMoCa simulation than using GMC. Figure and caption published in [5].

function:

$$f_c(r) = \left(1 + \exp\left(\frac{r - r_0}{t}\right)\right)^{-1} \tag{4.3.3}$$

with $t = 0.15\,\text{Å}$ and $r_0 = 5.451\,\text{Å}$ ($1\,\text{Å}$ larger than the equilibrium distance) was used to ensure that the six nearest neighbors contribute to the order parameter only if within a certain limit.

The dependence of the average system energy and bond order parameter on the simulation step is displayed in Fig. 4.3.2. The energy relaxation in the liquid-to-crystal transition is well represented by an exponential function, $E(step) \propto \exp(-step/\tau)$. Comparison between the fits for GMC and AROMoCa, also displayed in Fig. 4.3.2, results in an energy decay that is eight times faster in AROMoCa than in GMC ($\tau = 559$ vs. $\tau = 4316$). Measurement of the bond order parameter in the right panel of Fig. 4.3.2 indicates an increase in the local order of the system induced by crystal transition. As for the energy, the order parameter increases much faster using AROMoCa than GMC. Fig. 4.3.3 shows snapshots at step 0, 100, 1000, 2500 and 5000 of the GMC trajectory and the trajectory of AROMoCa with $m = 4$, respectively, to further illustrate the ordering of the system. This visualization clearly suggests a fast reduction of the fraction of disordered atoms, i.e. the particles not in crystalline environment, to a constant fraction that arises from thermal fluctuations within the first 2500 MC steps of the AROMoCa simulations. The same degree of order is not reached in the GMC simulations after 5000 MC steps.

Subsequently the growth of crystalline domains in the sample was investigated in order to

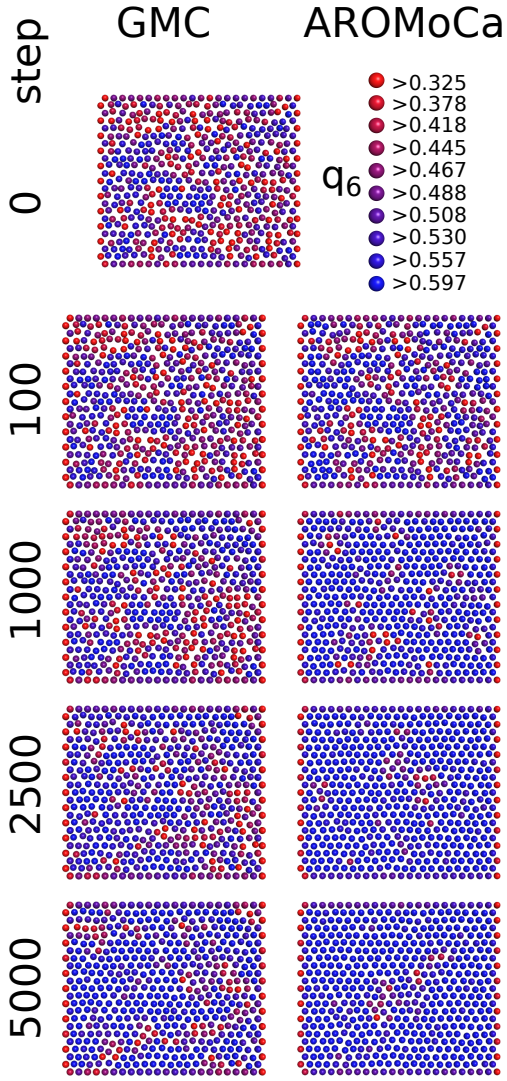


Figure 4.3.3.: The q_6-bond-order parameter of each atom and step was calculated to analyse the behaviour of the 2D Lennard-Jones model throughout the simulation. The color code indicates the degree of order for each particle as indicated in the top right. Snapshots for steps 0, 100, 1000, 2500 and 5000 are displayed. First in AROMoCa the system reaches a nearly crystalline configuration after less than 2500 steps and defects (red particles) occur afterwards to a constant fraction due to thermal fluctuations, while GMC does not reach the same degree of crystallinity after 5000 MC steps. Figure and caption published in [5].

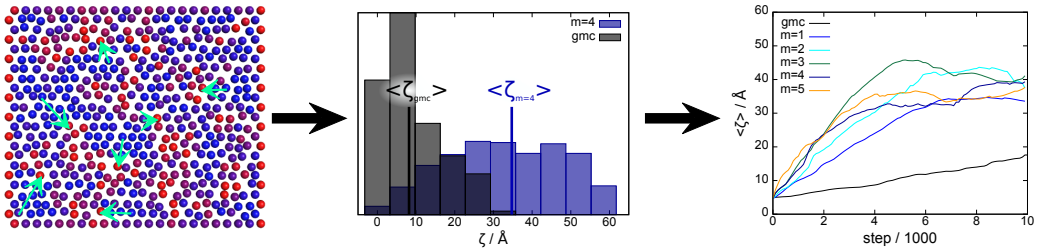


Figure 4.3.4.: Analysis of the growth of crystalline areas in the 2D system: For each particle and step, ζ, the distance to the nearest particle that is not in a crystalline environment (defined by $q_6 < 0.418$) is measured (left). The central panel shows the distribution of the GMC simulation and the AROMoCa simulation using $m = 4$ at step 5000 and the right panel the moving average of $< \zeta >$ for AROMoCa with $m = 1, ...5$ and GMC. $< \zeta >$ reaches an Eq.ilibrium value of 40Å (just below half the system size) much faster with ARMoCa even when using $m = 1$ than with GMC. Additionally AROMoCa performance improves with increasing m. Figure and caption published in [5].

further analyze the mechanism of the ordering process. The distribution of the distance ζ of each atom to the nearest "disordered" particle (defined by $q_6 < 0.418$), i.e. the nearest particle not in a crystalline environment, was measured on the basis of the q_6 values occurring in the simulation. ζ was calculated for snapshots taken every 100 simulation steps over a simulation of 10k MC steps. The definition of ζ, its distribution and the evolution of its mean value $< \zeta >$ are illustrated in Fig. 4.3.4. A moving average over up to 20 subsequent frames of the simulation (a lower number was averaged in the first 20 frames) was computed in order to compensate for the quite large fluctuations induced by a small size of the system. This trend again illustrates the advantage of AROMoCa over GMC: $< \zeta >$ reaches an equilibrium value of 40 Å , which is just below half the system size, much faster for all values of m in AROMoCa in comparison to GMC. Furthermore, the performance of AROMoCa improves with higher m.

Finally a three dimensional LJ liquid consisting of 216 particles was studied. The results of AROMoCa and GMC simulations were compared to a molecular dynamics (MD) simulation of the same system. AROMoCa simulations were performed using $m = 1$, 2 and 4 displacements per MC steps with a maximal step size of $\Delta x_{max} = 0.75$ nm. Again, Argon paramters were applied and a temperature of $T = 50$ K well below the melting point was used. The MD simulation was performed deploying the Langevin integrator as implemented in the OpenMM toolkit [92]. A time step of $\Delta t = 1.0$ fs was used. In all simulations periodic boundary conditions were applied in all three dimensions. The system energy was computed throughout the simulations and is plotted in Fig. 4.3.5 as a function of the number of simulation steps. Two plateaus in the energy are observed in the MD and AROMoCa simulations for $m = 4$, corresponding to the supercooled liquid and crystalline phase, respectively. To verify the phases, radial distribution functions were calculated and are displayed in Fig. 4.3.5b for the two plateaus in Fig. 4.3.5a. As indicated by the analysis of the radial distribution functions, all simulations first converge to a supercooled liquid state (encircled by dashed lines, dashed

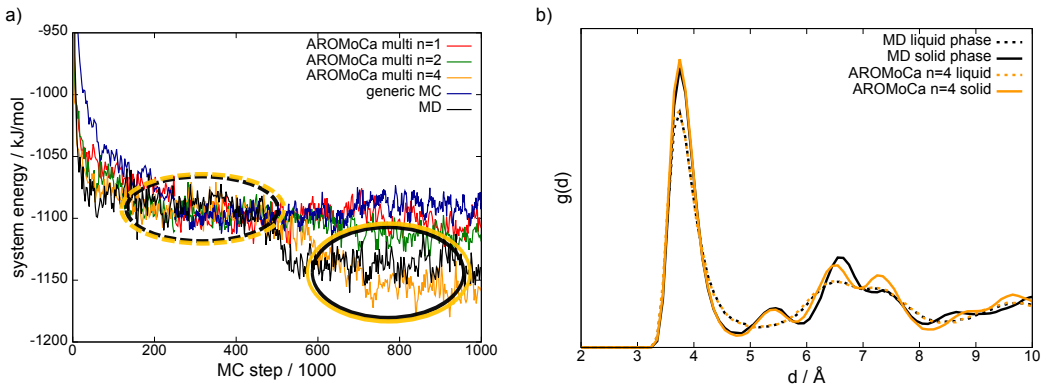


Figure 4.3.5.: Liquid-to-crystal transition of a three-dimensional Lennard-Jones liquid. The radial distribution functions show distinct peaks after the transition process (solid lines) in comparison to the super cooled liquid phase (dashed lines) indicating a crystalline phase. The transition to a crystal is not observed at all in the standard Metropolis Monte Carlo simulation. AROMoCa relaxes the system with approximately the same rate as the MD simulation. The right panel shows radial distribution functions for the encircled regions in Fig. (a) by averaging over 200 frames of the trajectories. The radial distribution functions were plotted for the MD run and the AROMoCa run with $m = 4$. Figure and caption published in [5].

RDF plots) before the crystal transition takes place in a subset of the simulations: Upon crystal transition the RDF change their shape and show additional distinct peaks (encircled in solid lines, solid RDF plots). Crystal transitions was observed for the MD simulation and for the AROMoCa simulation with $m = 4$ between the simulation steps 4×10^5 and 6×10^5 at approximately the same number of function evaluations. No phase transition occurred in the GMC runs and with AROMoCa using $m \leq 2$. This result indicates that some collective processes can be modeled using AROMoCa with comparable computational efficiency as by deploying MD simulations.

Detailed Balance and Acceptance Rates: Limits of fbMC

The following section aims at a direct comparison of AROMoCa with a similar approach to apply gradient-driven collective moves in MC methods, especially with force-bias MC methods (fbMC) [44]. In fbMC and other previously developed methods [41, 43, 44, 200] the gradient acting on single particles is used to determine the step size by which every single degree of freedom is to be changed in one simulation step. By constructing displacements such that the ratio between the probabilities to perform a move and the corresponding back move is approximately the Boltzmann weight, i.e. by using $\Delta E \approx \nabla E|_{\boldsymbol{x}} \Delta \boldsymbol{x} = -\boldsymbol{F}(\boldsymbol{x}) \Delta \boldsymbol{x}$, detailed balance is preserved approximately. Different versions of this approach have been proposed: Some methods apply an acceptance criterion that asserts detailed balance exactly, while in other approaches, all proposed steps are accepted under the assumption that de-

tailed balance is only weakly violated. In the latter approach, the acceptance rate is unity by definition, i.e. every move is accepted, while the computation of thermodynamic expectation values may be incorrect. This is not the case for methods deploying an acceptance criterion, for which, however, the efficiency depends on the acceptance rate for a particular system. In the following, both variants of the fbMC method are investigated and compared to AROMoCa for simple test system, in order to analyze the impact of an acceptance criterion on thermodynamic expectation values, such as the mean energy of the ensemble and on the acceptance rates.

A small system of Argon with 216 atoms with Lennard-Jones interactions was considered. To generate relaxed ensembles in gaseous, liquid and solid phase, the system was first equilibrated at three different temperatures ($T = 200\,\mathrm{K}$, $T = 80\,\mathrm{K}$ and $T = 40\,\mathrm{K}$). Subsequently, fbMC and AROMoCa simulations were performed at each temperature with 10^6 MC steps each. Random displacements in all three dimensions were applied with step sizes ranging from 0.005 nm to 0.075 nm. A 100 ps MD run at each temperature was performed for reference, resulting in averaged energies of $-1715.1 \pm 7.0\,\mathrm{kJ/mol}$, $-1192 \pm 11\,\mathrm{kJ/mol}$ and $-165.7 \pm 9.8\,\mathrm{kJ/mol}$ for solid, liquid and gaseous system respectively. For fbMC and AROMoCa, mean energy and standard deviation plotted over the step size are displayed in Fig. 4.3.6.

The energies of the ensemble simulated with the fbMC method resemble the reference energy for small steps in the gaseous and liquid systems. However, the deviation is increasing strongly with steps larger than 0.025 nm: While, in the gaseous system, the relative deviations from the energy averaged over the MD simulations are relatively small for the step sizes between 0.01 nm and 0.025 nm (deviations ranging from 1.2 % to 3.3 %), the deviations amount to 25 % and even 155 % for step sizes of 0.05 nm and 0.075 nm respectively. The liquid system shows similar behaviour, with a relative deviation of less than 2.5 % for step sizes of up to 0.025 nm but deviations of already 50 % for step sizes of 0.05 nm. In the simulation of the condensed system fbMC without acceptance criterion leads to a strong violation of detailed balance: the physical ensemble is not well represented and the expectation values of the energy increase drastically with the step size. Although the relative deviation is less than 4 % for step sizes of up to 0.025 nm, the absolute difference of about 7.166 kJ/mol even for the smallest step of 0.01 nm is more than 21 times $k_\mathrm{B}\,T$ at the simulation temperature of $T = 40\,\mathrm{K}$. This strongly hints at an incorrect computation of thermodynamic expectation values and points out the necessity to use an acceptance criterion in fbMC to obtain proper thermodynamic averages: Adequate results are obtained only for very small step sizes by approximating the change in energy by the product of gradient and displacement. In contrast, using AROMoCa with the acceptance criterion derived in Eq. 4.2.17, the physical ensemble is simulated without observable deviations, resulting in correct energies for all temperatures and step sizes up to 0.075 nm. Notably, in the crystalline system, the absolute deviation is only 1.55 kJ/mol for the largest tested step size of 0.075 nm and the highest deviation of all step sizes in all systems/temperatures is lower than 1 %.

Subsequently, the influence of an acceptance criterion applied in fbMC was analyzed by studying acceptance rates for a system consisting of 1728 Lennard-Jones particles (same parameters as above). Similar to the previous studies simulations consisting of 1000 MC

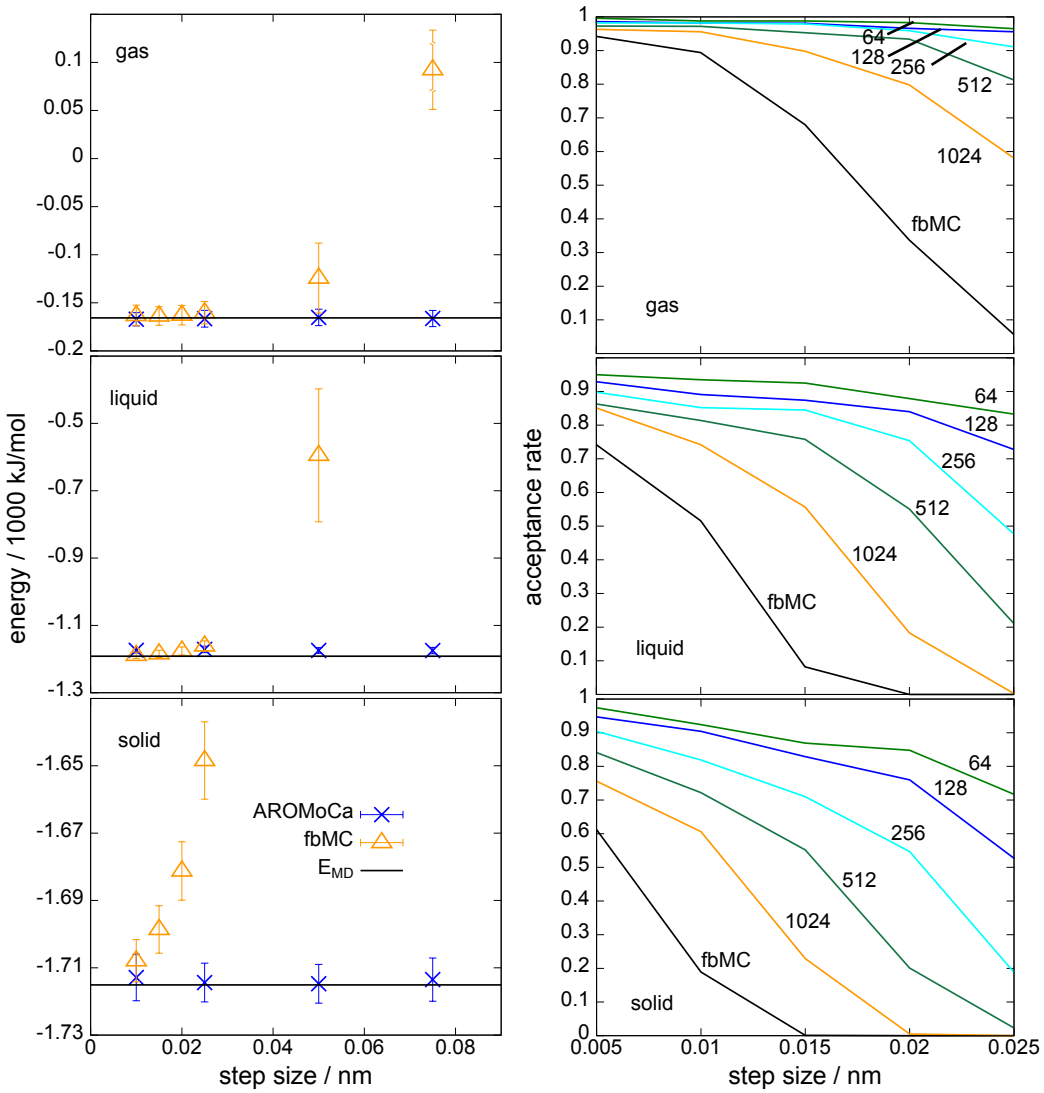


Figure 4.3.6.: Analysis of detailed balance and acceptance rates of fbMC and AROMoCa in different phases of the simulations. Left: Energy averaged over 10^5 snapshots of a 10^6 step simulations of an Argon system in the gaseous ($T = 200\,\mathrm{K}$, top), liquid ($T = 80\,\mathrm{K}$, middle) and solid state ($T = 40\,\mathrm{K}$, bottom) for different methods (AROMoCa and fbMC) as a function of the step size. No acceptance criterion was applied to the fbMC simulations. MD simulations of the same systems were performed for comparison. AROMoCa reproduces the MD energy distributions ($-1715.1 \pm 7.0\,\mathrm{kJ/mol}$, $-1192 \pm 11\,\mathrm{kJ/mol}$, and $-165.7 \pm 9.8\,\mathrm{kJ/mol}$ respectively) very well. Right: Acceptance rates of AROMoCa and fbMC runs with 1000 MC steps of gaseous, liquid and solid Argon (1728 particles at $T = 200\,\mathrm{K}$, $80\,\mathrm{K}$ and $40\,\mathrm{K}$ from top to bottom) for different step sizes. AROMoCa simulations were performed using $64 < m < 1024$ system changes per MC step. Especially for larger step sizes (> 0.1 nm) AROMoCa reaches much higher acceptance rates than fbMC. Figure and caption published in [5].

steps were performed for the gaseous, liquid and crystalline phase ($T = 200\,\mathrm{K}$, $T = 80\,\mathrm{K}$ and $T = 40\,\mathrm{K}$) after initial equilibration at each respective temperature. The acceptance criterion of Eq. 4.2.17 was applied for fbMC with the product over all particles and directions and acceptance rates were calculated for step sizes between 0.005 nm to 0.025 nm. AROMoCa simulations were performed for $m = 64$, 128, 256, 512 and 1024 parallel displacements per MC step. The resulting acceptance rates are displayed in the right hand side of Fig. 4.3.6. The AROMoCa acceptance rates are close to 1.0 for up to $m = 512$ parallel displacements for all step sizes in the gaseous state, where the fbMC acceptance rate becomes vanishingly small for step sizes close to $0.025\,nm$. In the liquid and solid phase acceptance rates larger than 0.5 can be achieved in AROMoCa: By reducing m, movement is basically limited to the most important regions of the system: As explained above, the smaller m, the larger the probability to only move those parts of the system that are far from the equilibrium. High acceptance rates are then achieved by increasing the probability to improve the total system energy. As in fbMC, there is no distinction between the regions of the system close to equilibrium and the regions farther away from equilibrium and the acceptance rate is reduced drastically in this approach.

4.4. Application of the Second Order AROMoCa

4.4.1. A Model for Macro-Molecular Systems

In the following, the study of model systems of bonded macro-molecules is presented in order to test the applicability of the AROMoCa version including the second order energy estimator to peptides, polymers and proteins. I will discuss a simplified model of a chain of atoms, resembling e.g. the back bone atoms of a protein, which nevertheless comprises all relevant internal degrees of freedom and can easily be adapted to test applicability and performance of AROMoCa to realistic systems. Additional features, such as side chains or hydrogen atoms, are neglected for reasons of simplicity as their influence on the generation of large-scale moves is expected to be negligible. The system energy is invariant under translation and rotation and the position of the initial atom and the direction of the bond between first and second atom were fixed in all simulations. As a simplified model for an external potential that induces the construction of correlated moves, the final atom of the chain was attached to the surface of a sphere using a harmonic potential:

$$U_{fix}(r) = 0.5\,c\,(r - r_0)^2 \tag{4.4.1}$$

Here, r is the distance to the center of the sphere located 3.1 Å in x-direction from the rest position of the last atom of the chain. The equilibrium radius was set to $r_0 = 3.0\,\text{Å}$ (in the range of realistic LJ length scales) and a strong force constant of $c = 100\,\mathrm{kJ/mol/Å^2}$ was chosen in order to investigate the AROMoCa performance under a strong constraint. An illustration of a setup comprising 20 atoms is given in Fig. 4.4.1. To avoid any systematic error in the observations, the structures were constructed randomly at the beginning of each simulation with realistic values for the rest positions, i.e. bond lengths of $r_k \in [1.35\,\text{Å}, 1.45\,\text{Å}]$, angles of $\alpha_l \in [1.35\,\text{Å}, 1.45\,\text{Å}]$ and values for the dihedral angles of $\phi_m \in [\frac{5}{6}\,\pi, \frac{10}{6}\,\pi]$. Internal

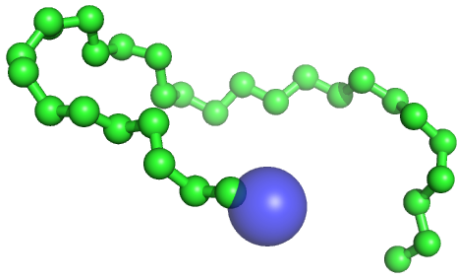


Figure 4.4.1.: A chain of atoms is used as a model to analyse the application of AROMoCa to polymeric or protein-like structures. Bond lengths, angles and dihedrals are used as coordinates. Throughout all simulations, the first atom (bottom right) is kept fixed. The last atom is attached to the surface of a sphere with radius 3.0 Å, depicted in blue, with an harmonic potential in order to study the application of AROMoCa to external potentials.

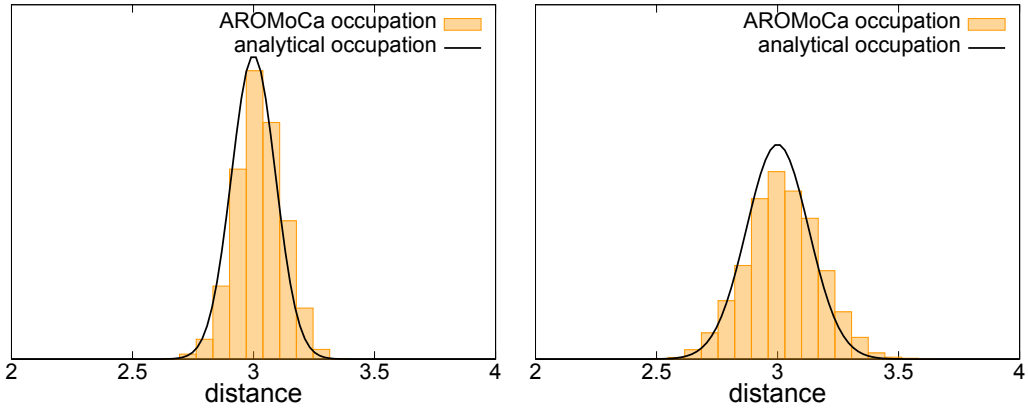
potentials were calculated according to Eqs. 4.2.40, 4.2.41 and 4.2.42. Force constants were set to $k_b = 50 \times 10^4\,\mathrm{kJ/mol/Å^2}$, $k_a = 500\,\mathrm{kJ/mol}$ and $k_d = 4.0\,\mathrm{kJ/mol}$ for bonds, angles and dihedrals respectively, similar to realistic force-field values.

4.4.2. Preservation of Detailed Balance

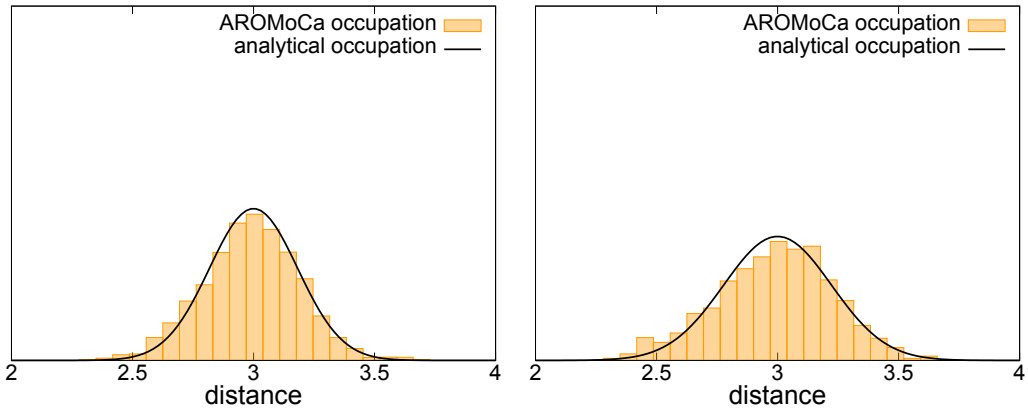
To verify that detailed balance is preserved by AROMoCa, simulations were performed at different temperatures $T = 100\,\mathrm{K}$, $T = 200\,\mathrm{K}$, $T = 400\,\mathrm{K}$ and $T = 600\,\mathrm{K}$ using a chain of $N = 30$ atoms. The distance r between the last atom and the center of the artificial harmonic potential was recorded over the course of 10^4 simulation steps. The distribution of r was compared to the analytical occupation in thermodynamic equilibrium, calculated according to:

$$\rho(r) = \frac{1}{Z} \exp\left(-\frac{U_{fix}(r)}{k_\mathrm{B}\,T}\right) \tag{4.4.2}$$

where Z is the partition function and $U_{fix}(r)$ the artificial potential as defined in Eq. 4.4.1. The results are displayed in Fig. 4.4.2. For each of the studied temperatures, the simulated occupation matches the analytical distribution almost perfectly. This indicates that detailed balance is indeed preserved by AROMoCa and that the correct thermodynamic ensemble is sampled. Restricting this analysis to the single quantity r is sufficient, as this value is not a system coordinate, i.e. it is not drawn by the AROMoCa algorithm directly. In fact, it is a function of all internal degrees of freedom that need to be distributed properly to achieve the correct distribution of r.

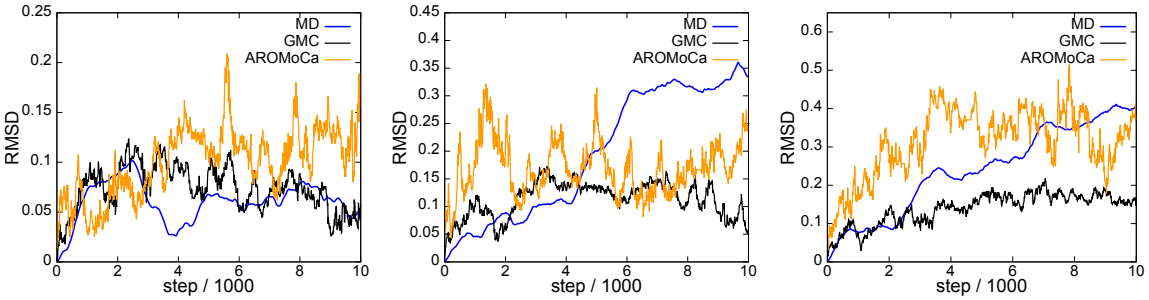


(a) $T = 100\,\mathrm{K}$ (b) $T = 200\,\mathrm{K}$



(c) $T = 400\,\mathrm{K}$ (d) $T = 600\,\mathrm{K}$

Figure 4.4.2.: To verify the preservation of detailed balance by AROMoCa, a chain of $N = 30$ atoms was simulated over 10^4 AROMoCa steps at different temperatures. The occurance of the distance of the last atom to the center of the artificial harmonic potential was recorded for different temperatures (bars) and compared to the analytic occupation distribution.



(a) $T = 100\,\mathrm{K}$ (b) $T = 300\,\mathrm{K}$ (c) $T = 500\,\mathrm{K}$

Figure 4.4.3.: Root-mean-square-deviation (RMSD) of the atoms in the system with $N = 30$ over the simulation steps for AROMoCa, GMC and MD at three different temperatures. The rapid changes in RMSD in both MC based simulations in comparison to MD indicate that the configurational change per simulation step is larger in MC methods. Further, AROMoCa reaches large RMSD values faster than MD, while the RMSD of the GMC simulations is limited to smaller values, especially in the case of $T = 500\,\mathrm{K}$.

4.4.3. Sampling Efficiency

Subsequently, the AROMoCa algorithm was compared to the well established protocols of MD and generic Monte Carlo (GMC). MD simulations were performed using a python code with OpenMM interface [92]. The Langevin integrator was used and the time step was set to the standard value of $\Delta t = 1\,\mathrm{fs}$. In GMC, only a single coordinate was changed per MC step to achieve significant acceptance rates and the step size was drawn from a random distribution between -0.1 and 0.1 (units in radian for angles and dihedrals or Å in the case of bonds). The root-mean-square-deviation (RMSD) from the initial rest position was calculated for every 10th frame using the gromacs toolkit (g_rms) [88, 89] and is displayed for three different temperatures $T = 100\,\mathrm{K}$, $T = 300\,\mathrm{K}$ and $T = 500\,\mathrm{K}$ for GMC, AROMoCa and MD in Fig. 4.4.3. GMC as well as AROMoCa show fast fluctuations on a small scale in the RMSD, while the course of the MD-RMSD is rather smooth. This is a not quite surprising consequence of the methods: in MC, large changes to the system are applied in each step (in the case of GMC only to a single coordinate), whereas MD is a sequence of small steps limited by the 1 fs time step. Moreover and more important, the RMSD of AROMoCa is superimposed by fluctuations on a larger scale of the order of half the maximal observed RMSD. These fluctuations are not observed in either MD or MC. This indicates a fast sampling of the phase space by AROMoCa, induced by large correlated changes to all degrees of freedom in every single simulation step. In fact, especially for $T = 300\,\mathrm{K}$, AROMoCa reaches the maximal RMSD (> 0.3) three times within 10^4 simulation steps, compared to a single occurrence in MD. GMC fails to reproduce high values of the RMSD within the simulation time for $T = 300\,\mathrm{K}$ and $T = 500\,\mathrm{K}$.

In order to quantitatively compare the sampling performance of AROMoCa and MD, the distribution of each single degree of freedom was measured. As illustrated in Fig. 4.4.4a,

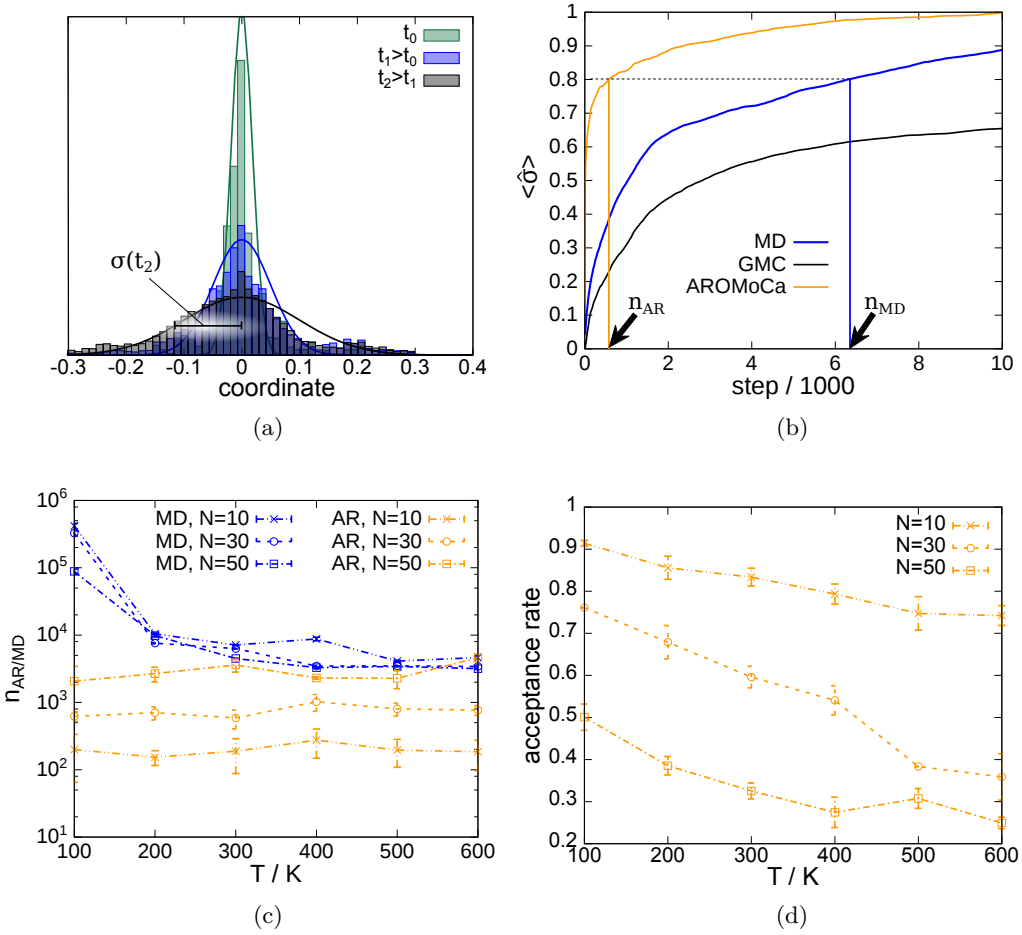


Figure 4.4.4.: Quantitative comparison of the sampling efficiency: (a) The occupation distribution of any internal coordinate broadens throughout the simulation. The width of the distribution is a measure for the sampling speed. Therefore, the width σ of each individual internal coordinate is normed to the value of the coordinate where the corresponding potential eqals $0.5\,k_\mathrm{B}\,T$, resulting in $\widehat{\sigma}$. The average of $\widehat{\sigma}$ over all internal coordinates of the system is plotted over the simulation step in panel (b). Each harmonic degree of freedom contributes $0.5\,k_\mathrm{B}\,T$ to the inner energy in thermodynamic equilibrium, and the number n of steps where $\langle\widehat{\sigma}\rangle$ equals 1.0 indicates thermalization of a system with only harmonic potentials. As the dihedral potential is not harmonic, we find $\langle\widehat{\sigma}\rangle < 1.0$ in many runs and the number n of steps where $\langle\widehat{\sigma}\rangle = 0.8$ is extracted as a measure for sampling efficiency. This number of steps for MD, n_{MD}, and AROMoCa, n_{AR}, is plotted over the temperature for different system sizes N in panel (c). Each datapoint is the average over three independent runs. For small systems or small temperatures, the speedup of AROMoCa compared to MD is two to three orders of magnitude. For high temperatures, the speedup remains of the order one to two orders of magnitude for $N = 10$ and $N = 30$, whereas MD and AROMoCa perform comparable for $N = 50$ at $T \leq 300\,\mathrm{K}$. (d) The AROMoCa acceptance rates are plotted in dependence of temperature for the respective systems.

the distribution of a single coordinate broadens throughout a simulation. The width of each coordinate can be used to measure sampling speed. In order to make the widths for bond lengths, angles and dihedrals comparable, the width σ_i recorded for every coordinate at every simulation step is normed to the value of the coordinate, where the corresponding potential equals $0.5\,k_\mathrm{B}\,T$. For angle l, for example, this would result in

$$\widehat{\sigma_l} = \frac{\sigma_l}{(\alpha_{l,k_\mathrm{B}\,T} - \alpha_{l,0})} \tag{4.4.3}$$

where $\alpha_{l,0}$ is the rest position and $\alpha_{l,k_\mathrm{B}\,T}$ is defined via:

$$k_a\,(\alpha_{l,k_\mathrm{B}\,T} - \alpha_{l,0})^2 = 0.5\,k_\mathrm{B}\,T \tag{4.4.4}$$

Here, k_a and $\alpha_{l,0}$ are the force constant and equilibrium angle, respectively. For every simulation step, the $\widehat{\sigma_i}$ are averaged over all coordinates. The resulting dependence of the averaged $\langle\widehat{\sigma}\rangle$ on the simulation step for $N = 30$ atoms at $T = 300\,\mathrm{K}$ is displayed in Fig. 4.4.4b. Subsequently, the number n of steps where $\langle\widehat{\sigma}\rangle = 0.8$ is extracted as a measure for sampling efficiency. This value is used as indicator for thermalization of the system: In thermodynamic equilibrium each harmonic degree of freedom corresponds to an inner energy of $0.5\,k_\mathrm{B}\,T$ but $\langle\widehat{\sigma}\rangle = 1.0$ is not always achieved as the dihedral potential is not harmonic. The corresponding values n_{MD} and n_{AR} for MD and AROMoCa respectively were recorded for different temperatures between $T = 100\,\mathrm{K}$ and $T = 600\,\mathrm{K}$ and three system sizes, $N = 10$, $N = 30$ and $N = 50$. To account for statistical fluctuations, three independent systems were sampled for each data point. As GMC failed to reach $\langle\widehat{\sigma}\rangle$ within 10^5 simulation steps for all setups, only MD and AROMoCa results are considered in Fig. 4.4.4c. At low temperatures, AROMoCa outperforms MD by several orders of magnitude. While the speedup of AROMoCa over MD is maintained at high temperatures for the smaller systems $N = 10$ and $N = 30$, AROMoCa and MD performance becomes comparable for $N = 50$ at high temperatures. Notably, AROMoCa performance is temperature independent but is decreasing with system size, while MD performance is almost system size independent but is decreasing towards low temperatures.

The tuning of MC simulation parameters usually presents a fine line between high acceptance rates and significant system changes of individual MC steps: Moves affecting a limited number of degrees of freedom per MC step allow high acceptance rates when the step size is sufficiently small, whereas large, collective system changes usually dampen the acceptance rate exponentially. The comparison to MD above however indicates, that AROMoCa indeed manages to circumvent this dilemma. In addition to the large configurational changes per MC step, indicated by the results above, the acceptance rate was recorded. For the three systems $N = 10$, $N = 30$ and $N = 50$, the acceptance rate averaged over three independent runs in dependence of the temperature is displayed in Fig. 4.4.4d. The acceptance rates drop with increasing temperature and with increasing system size, and are thus an explaining for the decreasing speedup of AROMoCa over MD for increasing N and T. We recall that the AROMoCa algorithm has exactly unit acceptance for systems where all potentials are harmonic in the degrees of freedom, as the estimator in Eq. 4.2.38 is exact for harmonic

potentials. However, the dependence of the external coordinate r on the internal coordinates is not linear and the estimator quality for the external potential decreases with increasing configurational changes. As the step size drawn from the AROMoCa algorithm is increasing with the temperatures, acceptance rates decrease consequently. However, regarding the results of Fig. 4.4.4c, this is compensated by larger configurational changes per simulation step. While the acceptance rate decreases with increasing number of degrees of freedom (i.e. with N), we note that acceptance rates are still remarkably large, considering that every single degree of freedom is changed simultaneously in every MC step.

4.5. Model Hopping

4.5.1. Controlled Biasing of Molecular Simulations: Model Hopping

A different approach to increase sampling in rugged energy landscapes is to bias the propagation of the system towards those parts of the phase space where the potential is expected to be minimal. While any infinite simulation at finite temperature will result in a major occupation of those regions, there are systems where the global optimum only occupies a small fraction of the phase space and may in addition be separated by a large energy barrier from other parts. As a result, a large fraction of the simulation time is spent on the propagation of the system into this region, before thermodynamic expectation values can be calculated. As explained in section 2.4.4, information about the system can be extracted e.g. from previously formed substructures in order to gently bias the system towards regions of the phase space where the potential is minimal. In order to preserve detailed balance, N systems are modeled in parallel using different energy functions E_0, E_1, ... E_i, ... E_j, ... E_N for every model. The correct density of states in thermodynamic equilibrium is reproduced, when exchange of two configurations C_i and C_j between models i and j is proposed after a certain number of simulation steps and accepted with a probability of

$$\rho_{exchange} = \min\left\{1, \frac{\exp\left(-\beta\left(E_i(C_j) + E_j(C_i)\right)\right)}{\exp\left(-\beta\left(E_i(C_i) + E_j(C_j)\right)\right)}\right\} \tag{4.5.1}$$

Here, E_i and E_j are the energy functions of models i and j, respectively. In general, these energy functions comprise the physical force field and a scaled bias potential. Unphysical results, possibly evoked by a badly chosen bias potential, are ruled out by scaling the bias potential to 0 in one of the models. The trajectory of the unbiased model can then be used for the calculation of thermodynamic properties, as exchange of thermodynamically unstable configurations into the unbiased model is prevented by applying Eq. 4.5.1. A detailed description of the MH approach and the derivation of Eq. 4.5.1 is given in section 2.4.4.

4.5.2. Choice of the Energy Functions E_i

The overall speedup gained by the Model Hopping approach is mainly dominated by the choice of the energy models E_i: While a strong bias leads to a fast propagation of the config-

urations in the biased energy models, the exchange probability between models is dampened by the strength of the bias, as we will show in the following. However, as thermodynamic expectation values can only be calculated on the basis of states realized in the unbiased model, the simulation only profits from the additional simulation time induced by the parallel modeling of multiple systems in the case of frequent exchange between the model. Otherwise, modeling multiple models results in increased computational cost without any benefit.
Hansmann et al. [111] originally proposed to add a bias potential E_B, scaled by a different factor $a_i \in [0,1]$ for every model i, to the physical Hamiltonian E_A:

$$E_i = E_A + a_i\, E_B \tag{4.5.2}$$

The exchange probability between two models i and j is then given by

$$\rho_{exchange} = \min\left\{1, \exp\left(\beta \Delta a \Delta E_B\right)\right\} \tag{4.5.3}$$

with $\Delta a = a_j - a_i$ and $\Delta E_B = E_B(C_j) - E_B(C_i)$. This exchange probability is independent of the physical energy model E_A. The impact of the bias strength on the exchange probability can be exemplified by considering a simulation of two models i and j: W.l.o.g., let $a_i = 0$, $a_j = a_0 > 0$, $E_B(C_k) > 0\ \forall k$. Model i is completely unbiased whereas the propagation of the configuration is biased towards a desired state in model j. After a certain simulation time the configuration C_j will be closer to the desired state than C_i, resulting in bias energies $E_B(C_j) < E_B(C_i)$ and thus $\Delta E_B < 0$. The exchange probability is then given by

$$\rho_{exchange} = \exp\left(-\beta\, a_0\, |\Delta E_B|\right) < 1 \tag{4.5.4}$$

In this choice of energy models E_i as given in Eq. 4.5.2, the acceptance probability between models is exponentially dampened by the bias strength, a_0. Unlike in the PT approach, this cannot be circumvented by increasing the number of models N: Indeed, decreasing $\Delta a = a_0/N$ between neighboring models leads to an increase in the exchange probability between each pair. However, the total number of exchange attempts to "hand down" a configuration from the model with the strongest bias into the unbiased model is increased in equal measure and the total exchange probability as a product of the single exchange probabilities remains that of Eq. 4.5.4.
Thus, a different choice of energy models will be used in the course of this work, where both the physical energy function and the bias potential are scaled in all models according to

$$E_i = (1 - a_i)\, E_A + a_i\, E_B \tag{4.5.5}$$

with $a_i \in [0, 1.0]$. For two systems i, j, inserting Eq. 4.5.5 into Eq. 4.5.1, results in

$$\rho_{exchange} = \min\left\{1, \exp\left(\beta \Delta a (\Delta E_B - \Delta E_A)\right)\right\} \tag{4.5.6}$$

with $\Delta a = a_j - a_i$, $\Delta E_B = E_B(C_j) - E_B(C_i)$ and $\Delta E_A = E_A(C_j) - E_A(C_i)$.
Accordingly, consider a simulation deploying two models with $\Delta a = a_0 > 0$ and assume $\Delta E_B < 0$ after a certain simulation time. According to Eq. 4.5.6 acceptance probability

of an exchange attempt is increased if $\Delta E_A < 0$. In other words, if the configuration of the biased system, C_j, is energetically favorable compared to the unbiased configuration, C_i, exchange is accepted with high probability. Furthermore, if the gain in physical energy, ΔE_A, upon exchange overcompensates the "loss" in bias energy, ΔE_B, exchange is accepted with unit probability, independent of a_0. This consideration highlights the role of the bias strength: An arbitrary large bias potential, resulting in $\Delta E_B \ll 0$, exponentially dampens exchange acceptance. Only if ΔE_A and ΔE_B are comparable in magnitude, the exchange probability differs significantly from zero. Hence, the strength of the bias potential needs to be tuned with consideration, as it is always a trade-off between fast convergence in the simulation periods between the exchange attempts and exchange acceptance probability.

4.5.3. A Model for the Nucleation of Molecular Crystals

The main aim of this study is to test the applicability of Model Hopping to the growth of molecular crystals. Unlike in single atom crystals, where the ordering is determined by the translational degrees of freedom, the orientations of the molecules play a key role in the ordering process.
In order to limit computational time and thus allow for a thorough analysis of the MH algorithm parameters, simple models are used. All models comprise non-interacting particles in an artificial energy landscape that is adapted to mimic the key challenges of the growth of a molecular crystal: The filling of defects in a crystalline layer, the continuous growth of a step edge and the nucleation of several molecules as the starting point of a new layer.

Step 1: Filling of a Defect Repairing a single defect in a crystalline layer is dominated by the particle position and orientation of molecules plays a minor role. Hence, this process is mimicked by a single isotropic particle in a slightly rugged substrate potential. To account for the site of the defect, a potential well is added at the center. The composite substrate potential is given by

$$U_{substrate} = Z(z)\, U_{xy}(x, y) \tag{4.5.7}$$

The potential in z-direction is a LJ-like potential inducing adsorption of the free atom on the surface:

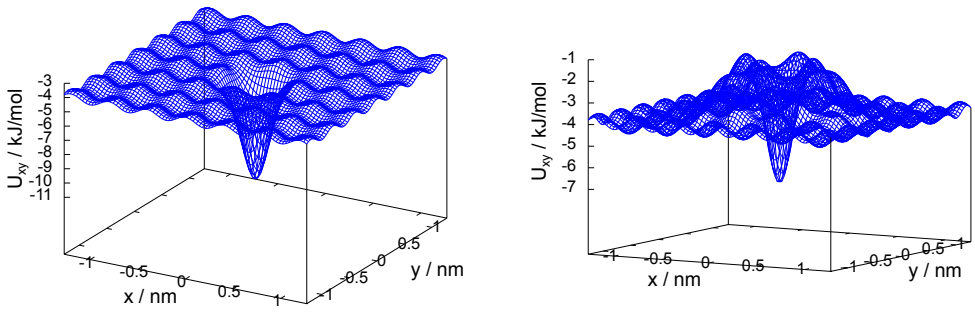
$$Z(z) = 4.0\,\epsilon \left(\left(\frac{\sigma_z}{z} \right)^{12} - \left(\frac{\sigma_z}{z} \right)^{6} \right) \tag{4.5.8}$$

Here, ϵ was set to $\epsilon = 1.0\,\mathrm{kJ/mol}$. In order to reproduce a rugged energy surface, $U_{xy}(x, y)$ was set to

$$U_{x,y}(x, y) = c + \epsilon_{x,y}(\cos(\alpha x) + \cos(\alpha y))\,\mathrm{kJ/mol} + U_{pit} \tag{4.5.9}$$

where U_{pit} emulates the defect site:

$$U_{pit} = \epsilon_{pit}\, \exp\left(-\frac{x^2 + y^2}{2\sigma_{xy}^2} \right)\,\mathrm{kJ/mol} \tag{4.5.10}$$



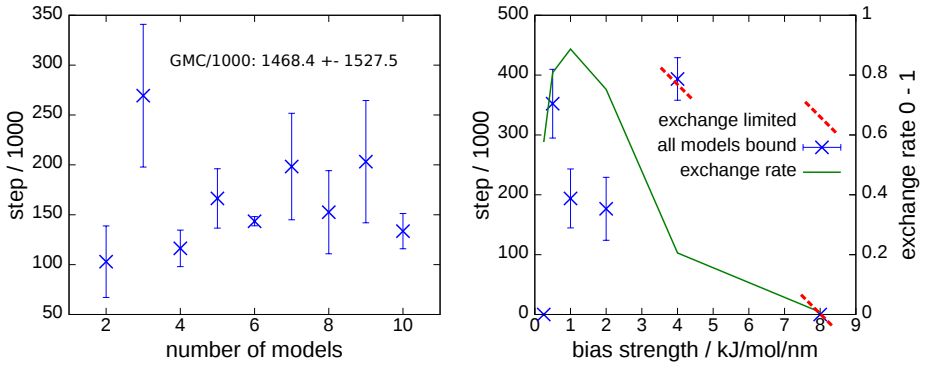
(a) (b)

Figure 4.5.1.: Model Hamiltonian for the simulation of the growth of molecular crystals. (a) A periodic potential in x-y-direction with an additional binding site in the middle mimics the binding of a molecule in a single defect site on a rugged surface. (b) The binding site surrounded by an energy barrier is used to model activated processes such as the continuous growth of an existing step edge. The parameters were adapted to reproduce the energies of a single particle on a physical LJ-substrate using standard carbon LJ parameters.

We used $c = 4.0\,\mathrm{kJ/mol}$, $\epsilon_{x,y} = 0.5\,\mathrm{kJ/mol}$ and $\epsilon_{pit} = 7.5\,\mathrm{kJ/mol}$ in order to reproduce the energy landscape of a realistic Lennard-Jones surface with standard carbon force field parameters, where the binding energy in the defect position was measured to be $-11\,\mathrm{kJ/mol}$ and the energies on the rugged surface varied between $-4.13\,\mathrm{kJ/mol}$ and $-2.81\,\mathrm{kJ/mol}$. This artificial potential is illustrated in Fig. 4.5.1a.
A simple bias potential of

$$E_B = c_0\, r, \quad r = \Big((x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2\Big)^{1/2} \tag{4.5.11}$$

was applied and the substrate potential of Eq. 4.5.7 was used as physical potential E_A.
Deploying this model, the dependence of the convergence speed on the number of models n_M and the bias strength c_0 was analysed. Obviously, computational effort scales linearly with n_M, and the number of models should be reduced as far as possible. While PT simulations are known to require multiple intermediate models, this assumption is questionable for MH, as explained previously. The model mimicking the single defect in the crystalline surface via the model Hamiltonian given by Eqs. 4.5.7 – 4.5.10 were simulated using $n_M \in [2, 10]$ models. Four independent runs per value of n_M were performed at $T = 35\,\mathrm{K}$. Ten independent GMC runs were performed for comparison and the number of steps to repair the defect was averaged. Convergence in MH was reached once the single atom was at the site of the defect in all models in MH.
The results are displayed in Fig. 4.5.2 and promote the following conclusions: First, there is no significant dependence of the convergence on the number of models n_M. Consequently,



(a) (b)

Figure 4.5.2.: Dependence of convergence speed on MH parameters n_M and c_0: a) Number of MC steps necessary to converge the MH runs in dependence of the number of parallel models. The bias strength was set to $c_0 = 1.0\,\mathrm{kJ/mol/nm}$ for all runs and results were averaged over independent runs. Ten independent Generic MC runs were performed for comparison (label inset). No significant dependence of the convergence on n_M is visible. b) Using $n_M = 5$ the convergence was averaged over 5 independent runs for each value of c_0. For $c_0 = 4.0\,\mathrm{kJ/mol/nm}$ some of the five independent runs and for $c_0 = 8.0\,\mathrm{kJ/mol/nm}$ all of the five independent runs did not converge at all as the exchange probability is decreasing with increasing bias strength. For further analysis the acceptance rate of the exchanges between the physical model and the model with the weakest bias was also plotted as a function of the bias strength. As expected, the exchange acceptance decreases if c_0 exceeds a certain value (in this case, $c_0 = 1.0\,\mathrm{kJ/mol/nm}$). The simulation with $c_0 = 0.25\,\mathrm{kJ/mol/nm}$ did not converge after 500×10^3 MC steps. For the system studied, values up to $c_0 = 2.0\,\mathrm{kcal/mol/nm}$ are reasonable.

unlike in PT, $n_M = 2$ is the optimal setup for efficient sampling. This supports the hypothesis that exchange probabilities are solely determined by the choice of the energy models and not by the number of models, as discussed in section 4.5.2. Second, the average number of steps needed for convergence in GMC is about ten times as large as the number of steps necessary in a MH simulation (see inset in Fig. 4.5.2a). As the computational effort in MH is $n_M \times n_{MC}$ with n_{MC} being the number of MC steps, only $n_M \leq 8$ is computationally more efficient than GMC. Using $n_M = 2$ induces a speedup of approximately 7. Keeping in mind that GMC relies on complete random diffusion for short range interaction (which is also indicated by the large standard-deviation of the convergence speed of GMC), the speedup can be expected to increase with size and complexity of the system at hand.
The number of steps necessary to converge the MH simulation with $n_M = 5$ models in dependence of the bias potential are displayed in Fig. 4.5.2b. The averaged exchange rate is included. Exchange is radically impeded for bias strengths of $c_0 = 4.0\,\mathrm{kJ/mol/nm}$ and $c_0 = 8.0\,\mathrm{kJ/mol/nm}$ and not all systems were fully converged. While the weakest bias ($c_0 = 0.25\,\mathrm{kJ/mol/nm}$) shows significant exchange rates, it is too weak to bias the system efficiently and convergence is not reached at all. However, simulations using bias strengths of $c_0 = 0.5\,\mathrm{kJ/mol/nm}$, $c_0 = 1.0\,\mathrm{kJ/mol/nm}$ and $c_0 = 2.0\,\mathrm{kJ/mol/nm}$ show exchange rates significantly larger than zero and good convergence.

Step 2: activated processes For the second challenge in the deposition of molecular crystals the potential was modified to mimic activated processes. In activated processes the system has to overcome an energy barrier in order to reach a thermodynamically stable state. Examples for activated processes include the binding of small molecules to proteins or, as explained above, the continuing growth of the step edge of a molecular crystal. In the latter, a deposited molecule has to undergo a reorientation into an energetically unfavorable state, before attaching itself to molecules in a crystalline phase. The potential given by Eqs. 4.5.7 - 4.5.10 was modified by adding a potential barrier around the defect site:

$$U_{pit} = -7.5\,\exp\left(-\frac{x^2+y^2}{2\sigma_{xy}^2}\right)\,\mathrm{kJ/mol} + c_b\,\exp\left(-\frac{x^2+y^2}{2(3\,\sigma_{xy})^2}\right) \tag{4.5.12}$$

The potential is illustrated in Fig. 4.5.1b. MH and GMC runs with 100×10^6 MC steps were performed using this potential with different values for c_b, leading to energy barriers of up to $7.63\,\mathrm{kJ/mol}$ (which is more than $25 \times k_\mathrm{B}\,T$ at $T = 35\,\mathrm{K}$). In MH, two parallel models with a bias strength $c_0 = 2.0\,\mathrm{kJ/mol/nm}$ were used and the number of MC steps that are needed for the particle to reach the binding site in both models was averaged over three independent runs for the MH simulations and five independent runs for GMC. The averaged results are displayed in Tab. 4.5.1.
While the performance of MH and GMC is comparable for the lowest barrier, $c_b = 1.0\,\mathrm{kJ/mol}$ (1.76×10^6 energy evaluations with MH vs. 2.4×10^6 energy evaluations with GMC), MH outperforms GMC for $c_b = 2.0\,\mathrm{kJ/mol}$ with a speedup of 1.86. For higher energy barriers, the defect site was not reached in GMC, as the probability to cross the energy barrier drops exponentially with the barrier height ΔE_b. MH on the other hand is capable of reaching the

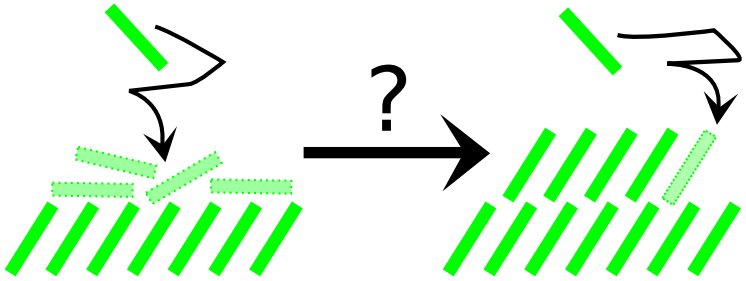


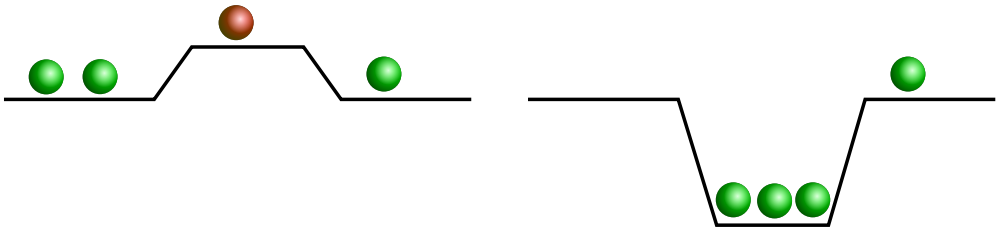
Figure 4.5.3.: Schematic illustration of the deposition of crystals consisting of planar molecules. Only when a certain number of molecules is deposited on a smooth layer, the new layer can form in a crystalline orientation. Up to this critical number, the flat orientation is energetically favorable.

defect site nevertheless, as the bias potential is pushing the system across the energy barrier in the biased model and exchange is accepted.

Step 3: Nucleation events A possible approach to the simulation of molecular crystal is layer-by-layer growth using single molecule deposition [1]. A major challenge in this protocol is the nucleation of crystalline seeds consisting of several molecules: single molecules that are newly deposited on top of a perfect crystalline layer are often arranged in a configuration that is different from the crystalline orientation and position (surface reconstruction). Only when the number of molecules exceeds a certain value, the crystalline state becomes energetically favorable and nucleation takes place. To achieve nucleation, single molecules need to first leave their stable single-molecule-configuration, which requires energy, and subsequently arrange in the crystalline position and orientation to reach the global optimum. The time scale for such correlated activated processes often exceeds the limits of MD, while GMC fails to construct the necessary correlated moves with a significant probability. This is illustrated in 4.5.3 for the case of planar moleculesFneg. For a limited number of molecules, the flat orientation is favorable. For nucleation to take place, they have to be brought in upright orientation, which costs energy, before the system gains energy upon nucleation into the global optimal state.

Table 4.5.1.: Convergence of activated processes using MH, modeled by a binding site surrounded by an energy barrier. GMC and MH runs were performed for different values of c_b (see Eq. 4.5.12) resulting in different barrier heights.

c_b / kJ/mol	barrier ΔE_b / kJ/mol	convergence MH / 10^6 steps	convergence GMC / 10^6 steps
1.0	0.52	0.88 ± 0.44	2.4 ± 2.2
2.0	1.11	0.97 ± 0.34	3.6 ± 3.0
4.0	2.37	2.1 ± 1.5	> 100
8.0	4.97	5.4 ± 0.5	> 100
12.0	7.63	> 10	> 100



(a) (b)

Figure 4.5.4.: Schematic illustration of a nucleation event: a) before the critical number of nucleation partners is in a state similar to the nucleation state (stand-up configuration of pentacene, etc) this state is energetically not favorable. b) As soon as the number of particles in the for single particles energetically unfavorable nucleation state exceeds a certain value, nucleation can take place, leading to a very stable configuration. As soon as there is one micro-crystal, other particles can easily attach to it.

To emulate this behaviour, multiple non-interacting particles were simulated and the surface potential of Eqs. 4.5.7-4.5.10 was changed as follows: All particles in the system are biased towards a certain region. The occupation of this region is energetically unfavorable, as long as the number of particles in the region is below a critical value. This corresponds to the unfavorable reorientation of single molecules (e.g. molecules being forced into the upright position). As soon as the number of particles in this "nucleation zone" exceeds a certain value – being the equivalent of enough molecules in the crystalline configuration – nucleation takes place and the occupation of the nucleation zone becomes energetically favorable. This corresponds to the nucleated state being stable, when the number of molecules exceeds a critical value. This is illustrated in Fig. 4.5.4.
To achieve this behavior, the potential mimicking the defect site, U_{pit}, is modified to

$$U_{pit} = \begin{cases} -c_{hill} / \left(e^{(r-\mu)/\sigma} + 1 \right), & n_{nucl} < n_{crit} \\ +c_{nucl} / \left(e^{(r-\mu)/\sigma} + 1 \right), & n_{nucl} \geq n_{crit} \end{cases} \tag{4.5.13}$$

Here, $c_{hill} > 0$ and $c_{nucl} > 0$ represent the energetical cost for reorientation of single molecules and the gain per molecule upon nucleation, respectively (negative sign will be compensated by $U(z)$-LJ-like potential). μ is the size of the nucleation zone and r the distance from the middle of the nucleation site. In this model, the distance to the nucleation zone emulates the orientation of a molecule in a realistic molecular system. n_{nucl} is the number of particles already inside the nucleation zone and n_{crit} the number of particles necessary for nucleation to take place. σ is a measure for the slope of the potential at the edge of the nucleation zone. For $c_{hill} = 2.0\,\mathrm{kJ/mol}$ and $c_{nucl} = 20.0\,\mathrm{kJ/mol}$, MH simulations using 5×10^6 steps were performed using the bias potential given in Eq. 4.5.11 with $c_0 = 2.0\,\mathrm{kJ/mol/nm}$. GMC simulations for comparison were omitted as the simulation of activated processes suggested that an energy barrier of $c_{hill} = 2.0\,\mathrm{kJ/mol}$ cannot be crossed without bias potential. The

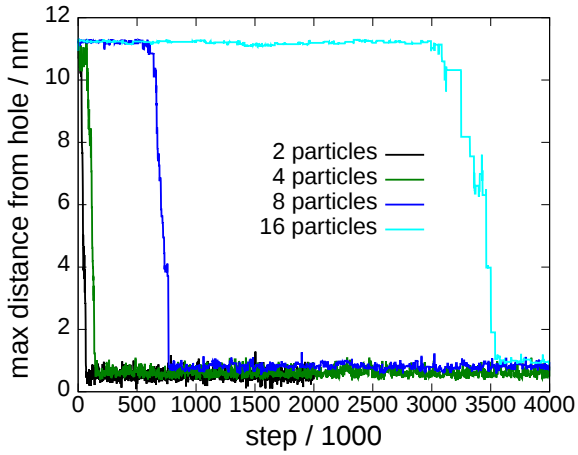


Figure 4.5.5.: Simulation of a nucleation event for different numbers of particles. For 2, 4, 8 and 16 particles nucleation was observed in less than 5×10^6 MC steps. The time needed for complete convergence scales approximately linearly with the number of particles in the system squared, n_p^2.

critical number of particles needed for nucleation was set to n_p =2, 4, 8 and 16. Exchange was attempted every $100 \times n_p$ steps. Convergence was reached once nucleation was observed in all models. The convergence behaviour is displayed in Fig. 4.5.5. For all systems, nucleation was observed in less then 5×10^6 MC steps. As this simple model system contains all relevant features of a nucleation event, this indicates that nucleation events can in principle be modeled using the Model Hopping approach.
Subsequently, in order to investigate the behaviour of the MH algorithm in dependence of barrier heights, the simulation of this nucleation event was attempted using Model Hopping for different combinations of values for c_{hill} and c_{nucl}. The strength of the bias was kept at 2.0 kJ/mol/nm and the simulation was performed with 8 (non-interacting) particles. Parameter combinations of c_{hill} = 2.5 kJ/mol, 4.0 kJ/mol and 8.0 kJ/mol and c_{nucl}/c_{hill} = 1.0, 2.0, 4.0, 8.0 and 16.0 for value of c_{hill} were applied. Regarding the simulation result, there are three possibilities: nucleation is successful in the biased model and exchange between the biased and the unbiased model leads to nucleation also in the unbiased model (✓/✓); the energy barrier is crossed in the biased model but exchange is not realized (✓/✗); nucleation occurs in neither the biased nor the unbiased run (✗/✗). The results for the MH runs are displayed in Tab. 4.5.2. GMC results are included for comparison. While nucleation is not observed at all in GMC (as was expected due to $c_{hill} > k_\mathrm{B}\,T$), nucleation occurred in all biased simulations independent of c_{hill}. However, exchange between the models leads to an increase in the bias potential of the biased model, if the configuration that was simulated in the unbiased energy function is farther from the desired state than the configuration simulated with the biased energy function. Only if this is overcompensated by the energy gain upon nucleation, which occurs due to the fact that the physical energy model is scaled by 0.5 in the biased energy model, exchange is accepted and the application of Model Hopping successful. This was the case for ratios of c_{nucl} and c_{hill} larger than 2.0.

Table 4.5.2.: Nucleation results for different energy barriers c_{hill} and ratios of c_{nucl}/c_{hill}. The MH simulations were performed using two models and 8×10^6 MC steps. For comparison, GMC simulations were performed for every value of the energy barrier using 7×10^7 MC steps. Results are divided in three categories: Nucleation was successful in all models (✓/✓), the energy barrier could be overcome but exchange did not occur (✓/✗) and nucleation was not observed at all (✗/✗).

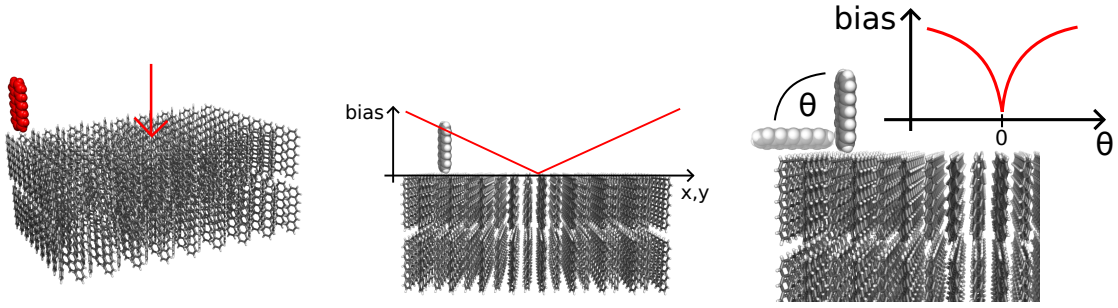
c_{nucl}/c_{hill} \ c_{hill}	2.5 kJ/mol	4.0 kJ/mol	8.0 kJ/mol
1.0	✓/✗	✓/✗	✓/✗
2.0	✓/✗	✓/✗	✓/✗
4.0	✓/✓	✓/✓	✓/✓
8.0	✓/✓	✓/✓	✓/✓
16.0	✓/✓	✓/✓	✓/✓
GMC	✗/✗	✗/✗	✗/✗

4.5.4. Transfer to Realistic Systems: Pentacene

The results above strongly suggest that it is, in principle, possible to apply MH for the fast deposition of molecular crystals, where the orientation of the molecules is essential to the ordering of the structure. To verify the applicability to real systems, the first challenge discussed above, namely the filling of a defect site in the smooth layer of the molecular crystal, will be analysed using a molecular crystal of pentacene.
In addition to the position dependent bias potential introduced above, a second bias is applied affecting the orientation of the pentacene molecules:

$$U_{orientation} = -c_z \left| \cos\left(\frac{p_z}{2.0} \left(\phi - \phi_0\right) \right) \right| \tag{4.5.14}$$

A periodicity of $p_z = 2$ is used and the equilibrium angle is set to $\phi_z = 5.74°$, as extracted from a perfect crystalline pentacene structure.
The bias potentials are illustrated in Fig. 4.5.6. In order to increase diffusion, basin hopping by simulated annealing (SA) was applied using cycles using 50 SA cycles with 4×10^3 MC steps each, annealing from $T = 3500\,\mathrm{K}$ to $T = 500\,\mathrm{K}$. Exchange of the configurations between the unbiased and the biased energy model was attempted after each SA cycle. After some rough pre-tuning of the bias strength, MH was applied to the system using $c_0 \in [0.5, 1.0]\,\mathrm{kcal/mol/Å}$ and an orientation strength of $c_z = 30.0\,\mathrm{kcal/mol}$.
Exchange was accepted in 33 out of 50 attempts for $c_0 = 0.5\,\mathrm{kcal/mol/Å}$ and in 32 out of 50 attempts for $c_0 = 1.0\,\mathrm{kcal/mol/Å}$. The distance of the molecule to the hole in both energy models after each exchange attempt is displayed in Fig. 4.5.7. Especially the simulation using $c_0 = 0.5\,\mathrm{kcal/mol/Å}$ illustrates quite well how the MH algorithm works: In the beginning, no exchange is performed. Model 1 (containing the biased energy function) converges quite fast to the defect position. Once the defect position is reached in the biased model, exchange is accepted almost every attempt and the defect is repaired after less than 80×10^3 MC steps



(a) (b) (c)

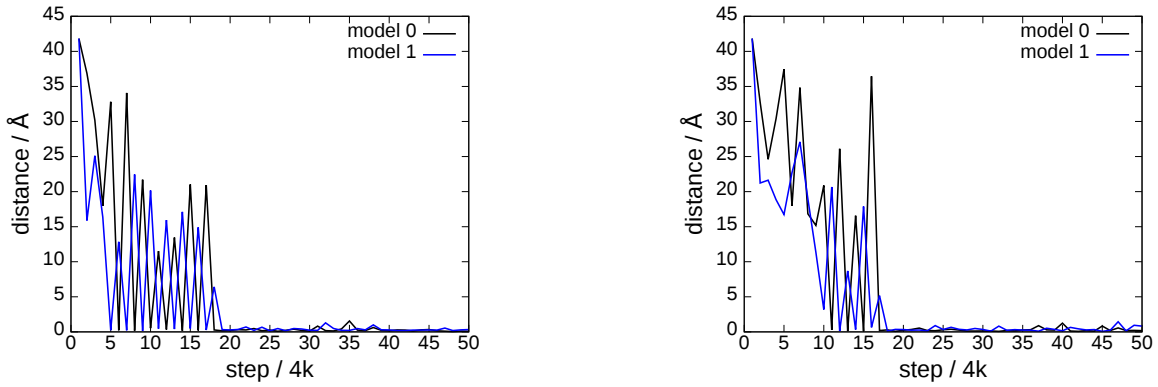
Figure 4.5.6.: Fixing a defect in a smooth pentacene surface: (a) in order to repair the defect, the molecule has to be put into an upright position and placed at the defect position. A position (b) and orientation (c) bias potential are used to achieve the correct position and orientation of the pentacene molecule in the biased model in the MH run.

in both models.
For comparison, three independend GMC runs were performed with the same number of SA cycles as in the MH simulations. The number of steps per SA cycles in GMC was increased to 10×10^3. As diffusion is pretty high at the temperatures in the simulation, the starting point of one SA cycle is de facto irrelevant. Hence, the three simulations are comparable to one large GMC run consisting of 150 MC cycles.
The defect was not repaired in any of the simulations, suggesting a speedup factor of at least $\frac{150}{19}\,\frac{10}{2\cdot 4} = 9.87$ of MH compared to GMC for this specific problem - in fact, it is not clear that the defect will be repaired in GMC in any feasible simulation time at all.

4.6. Summary and Outlook

Towards Unit Acceptance Rates

Molecular simulations are presently limited in system size and time-scale by the computational effort required to overcome long autocorrelation times. In this chapter the AROMoCa algorithm as a generic approach to generate complex collective and correlated moves with high acceptance rates in Monte Carlo simulations was investigated. By performing changes to the system based on an estimator for the change in energy induced by a potential move, higher acceptance rates for collective and even correlated moves can be achieved.
Two versions of the AROMoCa approach were presented. In the first version, the forces were used to identify degrees of freedom far from equilibrium and to construct collective, uncorrelated moves with high acceptance rates. AROMoCa using the first order energy approximation converged important physical observables, such as order parameters, faster than Metropolis MC methods for all of the systems studied in this chapter. Moreover, the performance of AROMoCa was comparable to molecular dynamics simulations for the



(a) $c_0 = 0.5\,\mathrm{kcal/mol/Å}$ (b) $c_0 = 1.0\,\mathrm{kcal/mol/Å}$

Figure 4.5.7.: Distance of the single molecule to the defect position in the substrate before each exchange attempt. Once the molecule is in the defect position in both models, the simulation is converged. Model 0 started in the unbiased force field and Model 1 in the force field including the bias. Exchange happens frequently once the desired configuration is reached in the biased system.

crystallization of a three dimensional Lennard-Jones liquid. Protocols similar to the first order AROMoCa, which moves a single particle, has been previously investigated, amongst others for the equilibration of water [41, 200]. The results in this chapter however imply that the change of a single coordinate per MC step is often insufficient to significantly accelerate the simulation convergence. Rather, a balance between the number of particles moved in a single MC step and the displacement of each degree of freedom is essential to obtain an optimized algorithm. The computation of gradients incurs roughly a factor of three in computational effort compared to the energy evaluation alone and efficient computationally optimized implementations are available in many MD programs. In the first order AROMoCa algorithm the total computational effort is therefore still dominated by the evaluation of gradients and energy. Accordingly, the cost induced by the move construction was below 10 % in the simulations reported above. Although similar ideas have already been explored in fbMC (force biased Monte Carlo) or uniform-acceptance force-bias Monte Carlo [43, 44], previously established methods lack one key feature: the identification of the degrees of freedom with values far from the local equilibrium, which is the central element of the first order AROMoCa method, is the essential new ingredient that generates acceptance rates close to unity for collective moves, while exactly preserving detailed balance at all times.
In the second part of this chapter the second order AROMoCa approach was applied successfully to macro-molecular systems. Our results indicate that detailed balance is preserved exactly while achieving large acceptance rates. A measure for the sampling efficiency of the macro-molecular system was introduced and the performance of AROMoCa compared to the gold standard MD. We found that in small systems, thermalization happens basically instantaneously in AROMoCa, while this process is limited by the femtosecond time step in MD. Depending on system size and temperature, AROMoCa outperformed MD by several

orders of magnitude. While the calculation of the energy estimator applied in the second version of AROMoCa exceeds the calculation of atomic forces and thus induces additional computational costs, the second order energy derivatives therein lead to a coupling of single degrees of freedom. This is the key feature to the construction of correlated moves that are necessary for the efficient simulation of many molecular systems, but cannot be realized in previously presented accelerated MC based protocols such as fbMC. In these methods, like in the first order AROMoCa algorithm, multiple uncorrelated moves are combined into a collective, yet still uncorrelated change to the system. A major issue to address in future studies is the system size dependence of the AROMoCa acceptance rates. In fact, the mapping of cartesian coordinates onto internal degrees of freedom and thus the estimator for the energy change is only accurate for small changes. Furthermore, the absolute error in this mapping increases with the distance between cartesian coordinate and internal degree of freedom, i.e. the size of the system. An improvement of the mapping accuracy will therefore increase the acceptance rate and hopefully remove the system size dependence of the overall performance.
In summary we found that AROMoCa is a generic MC based algorithm that can accelerate the simulation of many systems. The large speedup of the second order AROMoCa in comparison to MD simulations for small systems makes it a promising tool for the study of polymeric films, which are of high interest in organic electronics presented in the next chapter. Further, AROMoCa can be combined with or embedded in other methods such as simulated annealing (SA) [107], Multiple Try Monte Carlo (MTM) [197, 198], parallel tempering (PT) [108] or Model Hopping methods [111] that presently are based on generic Metropolis Monte Carlo or MD methods. This opens many possibilities for improvement of computational efficiency for a variety of applications.

Model Exchange in Molecular Simulations

In addition to the derivation of AROMoCa, the applicability of the Model Hopping approach [111] to the growth of molecular crystals was analysed. Therefore, a model system mimicking the key challenges in this process was constructed. The study of this model system suggested that, unlike in PT, the application of only two models, a biased and an unbiased one, is sufficient for the effective application of MH. Further, the choice of the bias potential and its strength is crucial for the exchange probabilities. In all three applications, namely the filling of a single defect in a smooth crystalline layer, activated processes and the model for nucleation events, convergence was reached where generic MC protocols showed limited performance or failed to produce the desired result completely. Notably, activated processes with activation barriers of $\Delta E > k_\mathrm{B} T$ were overcome by the MH approach. Moreover, the collective activation of multiple degrees of freedom that is necessary for nucleation events was achieved. Here, the global energetic minimum occupies only a tiny fraction of the configuration space. As a result of the large physical time scales of these processes, long simulation times are required in MD, while the probability to collectively activate the relevant degrees of freedom at the same time in GMC is vanishingly small due to large autocorrelation times. Successful application of MH to a realistic molecular crystal using the example of pentacene concluded this study.

In summary, the MH approach is a promising tool for the simulation systems where convergence is impeded by correlated effects or configuration space characteristics such as energy barriers or small volume fractions of global minima. MH can be applied to induce a drift towards relevant parts of the configurational space in MD and, in addition, to reduce autocorrelation times in MC based protocols.

5. Summary and Outlook

Inorganic semiconductors are well established and frequently used in a variety of present-day technological applications, such as lighting or electronic and photovoltaic devices. As a complementary technology with promising prospects, organic semiconductors have been intensely investigated for applications in energy conversion and electronics for several decades. On account of this, organic electronic devices have already reached the market for some applications, such as displays in mobile devices and TV-sets. Further, materials on the basis of organic small molecules or polymers have been successfully employed in large area OLED and OPV technology, promising inexpensive, large scale production of efficient and durable devices.

Nevertheless, there remain many challenges in the realization of inexpensive, market-ready thin-film devices. In principle, thin layers of amorphous organic compounds can be generated via solution processing, instead of small scale, troublesome vacuum vapor deposition. Besides jet-printing of customized electronic devices, this allows large-scale roll-to-roll processing of organic materials and thereby the production of large-area multi-layer OLED or OPV devices in a time- and cost-efficient manner. The power conversion efficiency (PCE) of solution processed devices, however, is significantly lower and their degradation much faster than in devices generated using vacuum vapor deposition.

The reasons of observed limitations in OLED and OPV devices are to date not fully understood. For instance, there is an active debate on the fundamental mechanisms of exciton dissociation at interfaces in OPV devices. Similarly it is generally acknowledged that the low mobility of charge carriers and excitons in OLED materials prevent large-area applications and it remains a mystery, how charge carriers are transported in guest-host systems, where the guest molecules (typically emitters) trap charge carriers in concentrations way below the percolation limit. In summary, despite decades of research, many fundamental functional principles in organic devices are not fully understood, in part owing to the complexity of the systems.

While experimental analysis and characterization of organic materials is essential for the development and enhancement of organic electronic devices, computer-based analysis of organic material properties and *in silico* device development may contribute significantly to the overall progress of organic electronics: First, device level simulations could give insight on microscopic processes that are responsible for limited PCE and fast degradation. Second, the identification of suitable materials by the presently practised trial-and-error approach, involving experimental synthesis, small scale production and characterization of organic electronic devices, is both costly and time-consuming. Efficient pre-screening of organic compounds via modeling of device characteristics on the basis of single molecule properties may speed up the search for novel, efficient and durable materials.

However, the amorphous arrangement in many organic materials renders *ab initio* calculation of material properties challenging, as the amorphous phase induces charge carrier localization and disorder in the electronic structure of each individual constituent. While DFT or Hartree-Fock methods can be applied to the calculation of molecular electronic structures in the gas phase, the treatment of organic materials on the nanometer-scale using these methods is presently not feasible. In fact, although computational resources become more powerful, there is no single method available that covers all necessary length- and time scales for the simulation of organic electronic devices. Therefore, a promising approach is the combination of multiple methods designed for the calculation of effects and processes on specific time and length-scales.

In this work I presented a multiscale modeling approach that allows the simulation of organic electronic materials including morphology and electrical properties on the electronic scale, aiming at the *in silico* design of organic electronic devices: In this approach the geometry of organic compounds is optimized using quantum-mechanical calculation methods in a first step. Subsequently, molecular modeling methods, i.e. molecular dynamics or Monte-Carlo simulations, are applied to generate atomistic morphologies on the nanometer scale using force-fields parametrized from QM. Rates for charge transfer between localized orbitals can then be calculated using Marcus theory of hopping transport, and Kinetic Monte Carlo simulations or analytical models yield macroscopic observables, such as the charge carrier mobility. Details on this multiscale protocol were given in chapter 2.

In chapter 3, this approach was applied to a variety of materials: First, I studied the electronic properties of thin layers of small organic molecules to demonstrate the power of the multiscale modeling approach. Using molecule specific force-fields, I generated atomistic morphologies for nine different materials with experimental mobilities ranging over several orders of magnitude. Microscopic hopping parameters were extracted from the electronic structure and the charge carrier mobilities were calculated in good agreement with experimental values. This indicates that macroscopic bulk-properties can be calculated from first principle using the multiscale modeling approach, which is an important step towards *in silico* material design.

While amorphous layers of organic compounds are widely investigated and realized in experiment, the encapsulation of organic materials in solid state junctions remains challenging. Due to limited layer stability, the materials are damaged when electrode material is deposited on top and sophisticated methods have to be applied to realize electronic contacts. However, these approaches either include specialized electrodes such as Hg-drops or require additional intermediate layers, increasing complexity or costs and thus hampering the development of thin and flexible large-area devices. Therefore, in a second investigation, I studied a system where layer stability is increased by coordinating Fe^{II} redox centers to terpyridine units, resulting in rigid, extended one-dimensional molecular wires (MCMWs). The geometry of the MCMW-units was optimized and force-field parameters were extracted on quantum-mechanical level. Monte Carlo simulations yielded a densely packed arrangement of a set of wires that could be extended into thin periodic layers. Further structural analysis shows that high stability is induced by strong electrostatic interaction between the components of the material. QM calculations yielded high electron mobilities in agreement with experimental

observations, showing that transport can proceed off-wire.
Subsequently, I studied the electronic properties of three-dimensional porous metal-organic frameworks (MOFs). MOFs have been intensely studied over the past decade and the combination of organic compounds with metal centers opens the prospect of thin, light-weight devices with intriguing properties. To date, over 20000 different MOFs have been realized and employed successfully in a variety of applications, such as CO_2 capture, hydrogen storage, sensing and photovoltaics. Despite continuous efforts, only insulating MOFs have been reported to date, impeding their application in electronic devices. Recently, loading of HKUST-1, a prototype MOF, with TCNQ and F_4TCNQ, however, was reported to increase the MOF conductivity by several orders of magnitude, but an explanation for this radical change in the electronic properties of HKUST-1 had yet to be established. Therefore, I investigated the electronic structure of HKUST-1 using different methods. DFT calculations on the periodic HKUST-1 yielded the band-structure along the high symmetry lines. A band-gap beyond the optical range and weak dispersion for bands close to the Fermi-level were found to be the fundamental reasons for limited charge transport in HKUST-1. Excitation energies extracted from refined DFT cluster calculations verified the value calculated for the band-gap, in agreement with experimental data. In the last study presented in chapter 3 I analysed the impact that small molecules loaded into the pores of HKUST-1 have on its electronic structure in order to explain the increased conductivity of HKUST-1 loaded with TCNQ and F_4TCNQ. I first studied the loaded state by optimizing the geometry of TCNQ and F_4TCNQ in HKUST-1 and found a configuration of host and guest molecules that explains the formation of additional charge channels through HKUST-1 via MOF-guest bonds. Electron mobility was calculated for first-order and second-order charge transfer processes using Marcus theory and Kinetic Monte Carlo simulations. The results indicate that electron transport through the combined MOF-guest-system is dominated by second-order effects for which an increase in the mobility of several orders of magnitude was observed in agreement with experimental data.
In the last part of this thesis, I investigated a method to accelerate molecular simulations using molecular dynamics or Monte Carlo methods. These molecular modeling techniques are frequently used in a wide range of scientific problems concerned with structure formation on the scale between 1 nm and 1 μm, such as the folding of proteins, phase transition of glassy films and prediction of amorphous organic structures. Therefore, they are also an important component of the multiscale modeling approach. Despite continuing efforts to develop better methods and increased computational power, standard methods where structure formation is relevant suffer from intrinsic limitations, such as the femto-second integration time-step (MD) or low acceptance rates and long autocorrelation times (MC). In chapter 4 I developed a novel approach to generate complex collective and correlated moves with high acceptance rates in Monte Carlo simulations. In this Acceptance Rate Optimized Monte Carlo (AROMoCa) method, correlated MC moves are constructed on the basis of an estimator for the change in energy that is induced by potential moves, thus increasing system dynamics by achieving high acceptance rates for large step sizes.
In a first version of the method, energy gradients were used to focus system changes on regions of the phase space far from equilibrium. Collective moves with high acceptance

rates were achieved by biasing the move construction with an energy estimator based on the atomic forces, while preserving detailed balance exactly. This approach was applied to several systems of increasing complexity. In a direct comparison of AROMoCa and MD simulations of the crystal transition of a three-dimensional Lennard-Jones liquid, the well established MD protocol and the novel MC method were competitive for the first time.
In a second version of the algorithm, I included the second-order Taylor expansion in the energy estimator to allow larger system changes and induce correlated moves affecting all degrees of freedom in a single simulation step. This algorithm was applied to a model system mimicking macro-molecules in order to test its applicability to proteins and polymers and to allow comparison to MD simulations. Comparison of coordinate distributions in the simulations of MD and AROMoCa demonstrated a speedup of AROMoCa over MD. However, results indicate that MD outperforms AROMoCa for larger systems at high temperature. This indicates that the advantage of AROMoCa can be fully exploited in setups where system dynamic is hampered, e.g. near phase transitions.
In many applications of molecular modeling methods, the thermodynamically stable configuration covers only a small fraction of the multidimensional configuration space. Additionally, the region of the phase space where the potential is minimal may be surrounded by an energy barrier, rendering the progress towards global equilibrium an activated process. The simulation of such correlated activated processes, e.g. molecular nucleation events, is infeasible with standard MD or MC based protocols due to millisecond-time-scales and long autocorrelation times, respectively. Concluding chapter 4, I therefore investigated the Model Hopping (MH) approach regarding its applicability to the growth of molecular crystals. In the MH protocol, information on the system of interest can be used to bias replicas of the system towards the global minimum in the potential energy surface. Exchange between replicas and the unbiased system is restricted by an acceptance criterion and allows enhanced sampling while preserving detailed balance. Three critical aspects in the growth of molecular crystals – the filling of a defect in a perfectly smooth layer, the continuous growth of a step-edge and the nucleation on top of a smooth layer – were investigated using model systems. For all three systems I found that the MH simulations outperformed the established Metropolis MC method, indicating that MH can help to increase sampling efficiency of correlated or activated processes.
The work presented here is thus one step towards the development of multi-scale modeling methods for device level simulations and opens the prospect of efficient characterization and prediction of organic material properties, and the *in silico* design of OE devices. The methods presented can be applied for pre-screening of materials, limiting laborious experimental trial-and-error device characterisation and thus facilitating the development of novel organic materials.

Thanks to intensive research over the past decades, organic materials have been used successfully in a variety of interesting applications including market-ready devices, such as mobile displays. Nevertheless, there are many open questions regarding the properties of organic electronic devices, e.g. the structural formation and the behaviour of charge carriers near organic-organic interfaces. Most present-day OLED devices comprise several layers and it is unclear how molecules arrange near these interfaces and, further, how the molecular packing

and the resulting molecular electronic structure in these regions influence the device properties. Furthermore, the formation of excitons by electron-hole recombination in the emission layer of OLED devices, the diffusion of excitons and the radiative decay need to be analyzed for a full explanation of OLED device characteristics. Similar questions arise in the structural formation of OPV devices. By phase separation, a binary mixture of e.g. a polymer and a small molecule form domains of bulk-hereojunctions. Size and arrangement of these domains directly determine the electrical device properties. Neither structural domain formation nor electronic properties such as exciton generation and diffusion within the domains, exciton separation at the interfaces between the domains or the transport of electrons and holes to the electrodes are to date fully understood.
In addition, both OLED and OPV technology comprise organic-inorganic interfaces that pose a twofold challenge. First, the deposition of electrode material can damage the organic material, and novel electrode materials and processing techniques are required for inexpensive large-scale production. Second, the electronic processes at these organic-inorganic interfaces are to date not fully understood but play a crucial role in the device efficiency. Furthermore, in both OPV and OLED devices, the fast degradation of organic materials produced by solution processing is one of the major draw-backs in the realization of inexpensive large-area devices. Therefore, a fundamental understanding of the dependence of molecular packing on the production method, the ageing of material due to charge transport and the sensitivity of charge transport properties to defects is a major goal in the establishment of organic materials in electronic applications.
The results of chapter 3 demonstrate that multiscale simulation methods are able to compute macroscopic properties of organic semiconductors, indicating that multiscale modeling is a promising approach toward device calculations from first principle. In regard of the limitations described above, the extension of the multiscale modeling approach toward the theoretical description of complete devices is a prominent goal in current research. The charge transport calculations, as presented in these studies, however, were limited to the simulation of either holes or electrons in the bulk material. To capture all physical effects, device level simulations require the description of additional processes in the bulk material as well as at organic-organic and organic-inorganic interfaces as mentioned above. A full description should therefore include the quantum-mechanical simulation of exciton generation via photon absorption or thermal activation, excition diffusion through bulk material, exciton separation at interfaces and radiative exciton annihilation.
In order to simulate exciton processes, charge transfer at interfaces and electronic properties on the device level, atomistic morphologies on the micrometer scale need to be generated to capture effects, such as the molecular packing at rough organic-organic interfaces in OLEDs or the formation of heterojunctions in OPV. The most straight-forward approach is the development of accelerated MD and MC methods. Advanced methods and increased computational power may help to overcome fundamental bottle-necks in the simulation of structural formation on the micrometer scale. The developed AROMoCa algorithm and the MH method aim in this direction. However, all studies to date investigated model systems, and while these models were selected to mimic relevant features of realistic systems, the performance of AROMoCa and MH when applied to realistic systems such as proteins, polymers

or small flexible organic molecules has yet to be tested.
An alternative approach to the atomistic simulation of microscale structures is the application of coarse grained models. In these models, several atoms or molecules are combined into a single unit, thus limiting the number of degrees of freedom. This allows the simulation of large systems up to the device scale, but adequate parametrization of the interaction between coarse units and the re-mapping onto atomic resolution that is needed for electronic transport calculations remains ambitious.

Despite various achievements over the past decade, the development of efficient, inexpensive applications of organic electronics still poses a challenge to theory and experiment. In this context, the establishment of device level simulations is a crucial goal in the progress of OLED and OPV technology, as these simulations promise to deliver insight on fundamental limitations of current devices and may support the development of suitable materials as an alternative to practised experimental trial-and-error approaches.

A. Programs and Tools

In addition to programs and packages mentioned in the respective results sections, the following tools were used throughout this thesis:

A.1. Data Analysis

Data analysis was performed using python with the numpy and scipy libraries.

A.2. Graphics and Illustrations

- PyMOL, Version 1.4.1, was used for rendering molecular systems.
- Scientific data was plotted using gnuplot.
- Inkscape was used for post-processing graphs, rendered pictures and the construction of illustrations
- Povray was applied for the construction of rendered objects in the introduction

B. Abbreviations, Figures and Tables

Frequently Used Abbreviations

LED Light emitting diode

PCE Power conversion efficiency

OLED Organic light emitting diode

OFET Organic field effect transistor

OPV Organic photovoltaic

MOF Metal-organic framework

QM Quantum mechanics/mechanical

MD Molecular dynamics

MC Monte-Carlo

GMC generic Monte-Carlo (Metropolis-MC)

AROMoCa Acceptance Rate Optimized Monte Carlo

MH Model Hopping

DFT Density functional theory

KMC Kinetic Monte-Carlo

MM Molecular modeling

ESP Electrostatic potential

NDDO Neglect of Diatomic Differential Overlap

SA Simulated Annealing

PT Parallel Tempering

HOMO Highest occupied molecular orbital

LUMO Lowest unoccupied molecular orbital

SX super-exchange

MCMW Metal centered molecular wire

LJ Lennard-Jones

RMSD Root-mean-square-deviation

List of Figures

List of Tables

C. Calculation of the Derivatives of Cartesian Coordinates with Respect to Internal Coordinates

C.1. Indexing

In order to apply AROMoCa with the second order estimator to bonded macromolecules, it is necessary to express changes in potentials defined in cartesian space in internal coordinates (bonds, angles, dihedrals). Therefore, the derivatives of cartesian coordinates with respect to internal coordinates need to be calculated for the application of Eqs. 4.2.59 and 4.2.60. To allow a straight forward implementation, all indices (i.e. indices for atoms, bonds, angles and dihedrals) start at 0. Accordingly, the bond with index k is the bond between the atoms k and $k+1$, the angle with index l is defined by the vectors from atoms l to $l+1$, $\boldsymbol{v}_l$, and from atoms $l+1$ to $l+2$, $\boldsymbol{v}_{l+1}$, and the dihedral angle with index m is defining the rotation around the bond between the atoms with indices $m+1$ and $m+2$. This is illustrated in Fig. 4.2.2. Further, the rotation axis of a rotation around this angle is the vector normal to $\boldsymbol{v}_l$ and $\boldsymbol{v}_{l+1}$ at the atom with index $l+1$, denoted $\hat{n}_{l+1}$ by the corresponding rotation matrix $R(\hat{n}_{l+1}, \alpha_l)$. The rotation matrices for dihedral angle m is $D(\hat{v}_{m+1}, \phi_m)$. For more details on indexing and definitions, see main text, section 4.2.4.

C.2. First order Derivatives

C.2.1. First Order Bond Derivatives

The position of atom i, $\boldsymbol{x}_i$, can be written as:

$$\boldsymbol{x}_i = \boldsymbol{v}_{0,k-1} + \boldsymbol{v}_k + \boldsymbol{v}_{k+1,i} = \boldsymbol{v}_{0,k-1} + r_k\,\hat{v}_k + \boldsymbol{v}_{k+1,i} \tag{C.2.1}$$

where $\boldsymbol{v}_k = r_k\,\hat{v}_k$ is the bond vector from atoms k to $k+1$ with bond length r_k. As the vectors preceding and following $\boldsymbol{v}_k$ are independent of r_k, the first derivative of $\boldsymbol{x}_i$ in r_k is

$$\frac{\partial}{\partial r_k}\boldsymbol{x}_i = \hat{v}_k \qquad \forall \quad k < i \tag{C.2.2}$$

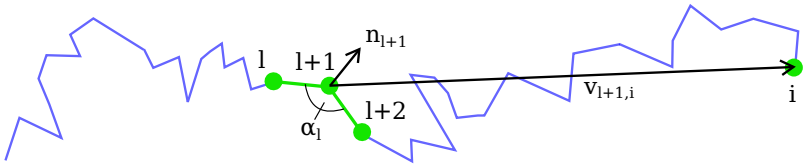


Figure C.2.1.: For the calculation of the first order angle derivatives, the change of the vector from atom $l+1$ to atom i upon a change of the angle α_l is analyzed.

C.2.2. First Order Angle Derivatives

An illustration for the calculation of first order angle derivatives is presented in Fig. C.2.1. Using the left hand rules for cross products as learned in high school[1] and the illustration in Fig. C.2.1, a rotation of $\Delta\alpha_l$ around the vector $\hat{v}_l \times \hat{v}_{l+1}$ would decrease angle α_l. Thus, the normal vector for the rotations of the angles α_l at atoms $l+1$ is chosen as

$$\hat{n}_{l+1} = -\hat{v}_l \times \hat{v}_{l+1} \tag{C.2.3}$$

The derivative of position $\boldsymbol{x}_i$ with respect to angle α_l is calculated by splitting the vector into two parts:

$$\boldsymbol{x}_i = \boldsymbol{v}_{0,l+1} + \boldsymbol{v}_{l+1,i} \tag{C.2.4}$$

As $\boldsymbol{v}_{0,l+1}$ is independent of the change of angle l, only the second contribution is considered. Rotation of $\boldsymbol{v}_{l+1,i}$ around rotation axis $\hat{n}_{l+1}$ by an infinitesimal angle $\Delta\alpha_l$ results in a modified vector $\boldsymbol{v}'_{l+1,i}$:

$$\boldsymbol{v}'_{l+1,i} = R(\hat{n}_{l+1}, \Delta\alpha_l)\, \boldsymbol{v}_{l+1,i} \tag{C.2.5}$$

$$\approx (1 + \frac{\partial}{\partial\alpha_l} R(\hat{n}_{l+1}, \alpha_l)|_{\alpha_l=0}\, \Delta\alpha_l)\, \boldsymbol{v}_{l+1,i} \tag{C.2.6}$$

The change of the atom position $\boldsymbol{x}_i$, $\Delta\boldsymbol{x}_i$ then equals the change of vector $\boldsymbol{v}_{l+1,i}$:

$$\Delta\boldsymbol{x}_i = \boldsymbol{v}'_{l+1,i} - \boldsymbol{v}_{l+1,i} \approx \frac{\partial}{\partial\alpha_l} R(\hat{n}_{l+1}, \alpha_l)|_{\alpha_l=0}\, \Delta\alpha_l\, \boldsymbol{v}_{l+1,i} \tag{C.2.7}$$

resulting in the first angle derivative:

$$\frac{\partial}{\partial\alpha_l}\boldsymbol{x}_i \approx \frac{\Delta\boldsymbol{x}_i}{\Delta\alpha_l} = \frac{\partial}{\partial\alpha_l} R(\hat{n}_{l+1}, \alpha_l)|_{\alpha_l=0}\, \boldsymbol{v}_{l+1,i} \qquad \forall \quad l < i-1 \tag{C.2.8}$$

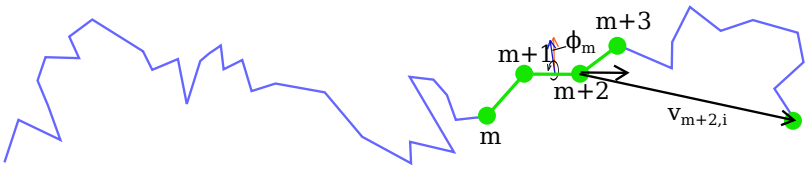


Figure C.2.2.: For the calculation of the first order dihedral derivatives, the change of the vector from atom $m+1$ to atom i upon a change of the dihedral angle ϕ_m is analyzed.

C.2.3. First Dihedral Derivative

Accordingly, the derivative of $\boldsymbol{x}_i$ in dihedral m is calculated by measuring the change of vector $\boldsymbol{v}_{m+2,i}$ from the atom position at the end of the bond that is rotated, atom $m+2$, to the atom i, upon rotation of dihedral m by $\Delta\phi_m$ around the rotation axis of the corresponding bond, $\hat{v}_{m+1}$:

$$\begin{aligned} \boldsymbol{v}'_{m+2,i} &= D(\hat{v}_{m+1}, \Delta\phi_m)\, \boldsymbol{v}_{m+2,i} && \text{(C.2.9)} \\ &\approx \left(1 + \frac{\partial}{\partial\phi_m} D(\hat{v}_{m+1}, \phi_m)|_{\phi_m=0} \Delta\phi_m\right) \boldsymbol{v}_{m+2,i} && \text{(C.2.10)} \end{aligned}$$

This results in:

$$\frac{\partial}{\partial\phi_m}\boldsymbol{x}_i \approx \frac{\partial}{\partial\phi_m} D(\hat{v}_{m+1}, \phi_m)|_{\phi_m=0}\, \boldsymbol{v}_{m+2,i} \qquad \forall \quad m < i-2 \qquad \text{(C.2.11)}$$

Notably, the derivative of cartesian coordinates with respect to bonds is exact, whereas the angle- and dihedral derivatives are only an approximation.

C.3. Second Order Derivatives

C.3.1. Second Order Derivative Bond Bond

As the unit vector of each bond is independent of all other bond unit vectors, this is fairly easy:

$$\frac{\partial^2}{\partial r_k^2}\boldsymbol{x}_i = \frac{\partial}{\partial r_k}\frac{\partial}{\partial r_k}\boldsymbol{x}_i = \frac{\partial}{\partial r_k}\hat{v}_k = 0 \qquad \forall \quad k \qquad \text{(C.3.1)}$$

[1] (thumb × index finger) is parallel to middle finger (when pointing in 90° away from thumb and middle finger); rotation around an axis aligned to your left hand thumb rotates in direction of all other fingers while doing the "thums up" sign).

C.3.2. Second Derivative Angle Angle – Same Index

From the considerations above, we get:

$$\frac{\partial^2}{\partial \alpha_l^2} \boldsymbol{v}_{l+1,i} = \frac{\partial^2}{\partial \alpha_l^2} R(\hat{n}_{l+1}, \alpha_l)|_{\alpha_l=0} \boldsymbol{v}_{l+1,i} \tag{C.3.2}$$

C.3.3. Second Derivative Dihedral Dihedral – Same Index

... and the same for dihedrals:

$$\frac{\partial^2}{\partial \phi_m^2} \boldsymbol{v}_{m+2,i} = \frac{\partial^2}{\partial \phi_m^2} D(\hat{v}_{m+2}, \phi_m)|_{\phi_m=0} \boldsymbol{v}_{m+2,i} \tag{C.3.3}$$

C.3.4. Second Derivative Angle Bond

If the bond is preceding the angle ($k <= l$), the two degrees of freedom are independent and

$$\frac{\partial^2}{\partial r_k \partial \alpha_l} \boldsymbol{x}_i = 0 \tag{C.3.4}$$

as a change in the bond length does not influence the rotation of $\boldsymbol{v}_{l+1,i}$ around the normal vector $\hat{n}_{l+1}$.
If, however, the bond is after the angle in the construction of the chain, $k > l$, the second derivative is non-zero, as the change of r_k along the rotated vector $\boldsymbol{v}'_k$ induces a different shift in the position of atom i:

$$\begin{aligned}
\boldsymbol{v}'_{l+1,i} =& \boldsymbol{v}'_{l+1,k} + \boldsymbol{v}''_k + \boldsymbol{v}'_{k+1,i} && \text{(C.3.5)}\\
=& R(\hat{n}_{l+1}, \Delta\alpha_l)\, \boldsymbol{v}_{l+1,k} + R(\hat{n}_{l+1}, \Delta\alpha_l)\, (\boldsymbol{v}'_k) + R(\hat{n}_{l+1}, \Delta\alpha_l)\, \boldsymbol{v}_{k+1,i} && \text{(C.3.6)}\\
=& R(\hat{n}_{l+1}, \Delta\alpha_l)\, \boldsymbol{v}_{l+1,k} + R(\hat{n}_{l+1}, \Delta\alpha_l)\, (\boldsymbol{v}_k + \Delta r_k \hat{v}_k) + R(\hat{n}_{l+1}, \Delta\alpha_l)\, \boldsymbol{v}_{k+1,i} && \text{(C.3.7)}\\
=& R(\hat{n}_{l+1}, \Delta\alpha_l)\, (\boldsymbol{v}_{l+1,k} + \boldsymbol{v}_k + \boldsymbol{v}_{k+1,i}) && \text{(C.3.8)}\\
&+ R(\hat{n}_{l+1}, \Delta\alpha_l)\, (\Delta r_k \hat{v}_k) && \text{(C.3.9)}\\
=& R(\hat{n}_{l+1}, \Delta\alpha_l)\, (\boldsymbol{v}_{l+1,k} + \boldsymbol{v}_k + \boldsymbol{v}_{k+1,i}) && \text{(C.3.10)}\\
&+ (1 + \frac{\partial}{\partial \alpha_l} R(\hat{n}_{l+1}, \alpha_l)|_{\alpha_l=0} \Delta\alpha_l)(\Delta r_k \hat{v}_k) && \text{(C.3.11)}\\
=& \mathcal{O}(\Delta\eta^1) + \frac{\partial}{\partial \alpha_l} R(\hat{n}_{l+1}, \alpha_l)|_{\alpha_l=0}\, \Delta\alpha_l\, \Delta r_k \hat{v}_k && \text{(C.3.12)}
\end{aligned}$$

The second derivative in bonds and angles is the expression containing two $\Delta\eta_m$:

$$\frac{\partial^2}{\partial \alpha_l \partial r_k} = \frac{\partial}{\partial \alpha_l} R(\hat{n}_{l+1}, \alpha_l)|_{\alpha_l=0}\, \hat{v}_k \qquad \forall \quad l < k < i \tag{C.3.13}$$

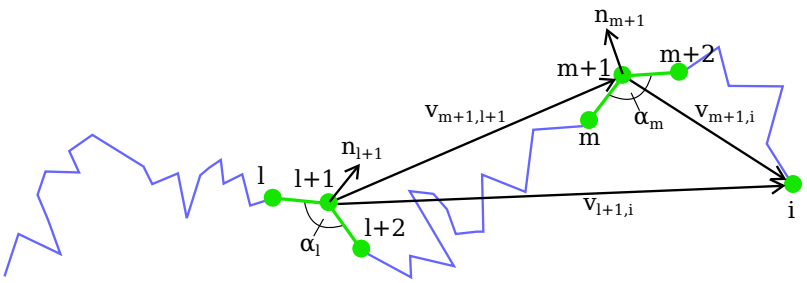


Figure C.3.1.: For the calculation of second order angle derivatives, first the change of the vector from atom $m+1$ to atom i, $\boldsymbol{v}_{m+1,i}$ upon change of angle α_m and subsequently, the change of vector $\boldsymbol{v}_{l+1,i}$ from atom $l+1$ to atom i upon a change of the angle α_l is analyzed.

C.3.5. Second Derivative Bond Dihedral

Without going further into detail: This is the same as the angle-bond case. I have it written on some paper but right now I am too lazy to tex it. We get:

$$\frac{\partial^2}{\partial\phi_m \partial r_k} = \frac{\partial}{\partial\phi_m} D(\hat{v}_{m+1}, \phi_m)|_{\phi_m=0}\, \hat{v}_k \qquad \forall \quad l+1 < k < i \tag{C.3.14}$$

C.3.6. Second Derivative Angle Angle – Different Indices

Let l and m be the indices of two angles and let w.l.o.g. be $l < m$. Then there are two possible approaches:

1. first, the second angle along the chain, m, is changed and then the first angle l
2. the first angle along the chain, l, is changed first, followed by the change of angle m

While it is possible to calculate the derivatives using both ways, both are expected to reproduce the same result in the infinitesimal limit, as derivatives commute:

$$\frac{\partial}{\partial x}\frac{\partial}{\partial y} f(x,y) \equiv \frac{\partial}{\partial y}\frac{\partial}{\partial x} f(x,y) \tag{C.3.15}$$

We expect the first approach to be slightly less intensive, as the normal vectors of the rotation remain fixed; following approach 2, the normal vector $\hat{n}_{m+1}$ will change upon rotation of $\boldsymbol{v}_{l+1,m+1}$ around the first normal vector $\hat{n}_{l+1}$, leading to a extensive calculation of the dependence of $\hat{n}_{m+1}$ on α_l. Therefore, we will focus on the first approach, changing angle m first and subsequently modifying angle l. A supporting illustration is given in Fig. C.3.1
Change of angle m modifies the vector from atom $m+1$ to atom i, $\boldsymbol{v}_{m+1,i} \rightarrow \boldsymbol{v}'_{m+1,i}$, therefore inducing a modification of the vector from atom $l+1$ to atom i, $\boldsymbol{v}_{l+1,i} \rightarrow \boldsymbol{v}'_{l+1,i} =$

$\boldsymbol{v}_{l+1,m+1} + \boldsymbol{v}'_{m+1,i}$. An additional change of angle l results then in:

$$\begin{aligned}
\boldsymbol{v}''_{l+1,i} =& R(\hat{n}_{l+1}, \Delta\alpha_l)(\boldsymbol{v}_{l+1,m+1} + \boldsymbol{v}'_{m+1,i}) && \text{(C.3.16)}\\
=& R(\hat{n}_{l+1}, \Delta\alpha_l)(\boldsymbol{v}_{l+1,m+1} + R(\hat{n}_{m+1}, \Delta\alpha_m)\boldsymbol{v}_{m+1,i}) && \text{(C.3.17)}\\
\approx& \left(1 + \frac{\partial}{\partial\alpha_l} R(\hat{n}_{l+1}, \alpha_l)|_{\alpha_l=0}\Delta\alpha_l\right) \times && \text{(C.3.18)}\\
& \left(\boldsymbol{v}_{l+1,m+1} + \left(1 + \frac{\partial}{\partial\alpha_m} R(\hat{n}_{m+1}, \alpha_m)|_{\alpha_m=0}\Delta\alpha_m\right)\boldsymbol{v}_{m+1,i}\right) && \text{(C.3.19)}\\
=& \boldsymbol{v}_{l+1,m+1} + \boldsymbol{v}_{m+1,i} && \text{(C.3.20)}\\
& + \mathcal{O}(\Delta\eta^1) && \text{(C.3.21)}\\
& + \frac{\partial}{\partial\alpha_l} R(\hat{n}_{l+1}, \alpha_l)|_{\alpha_l=0} \frac{\partial}{\partial\alpha_m} R(\hat{n}_{m+1}, \alpha_m)|_{\alpha_m=0}\, \boldsymbol{v}_{m+1,i}\, \Delta\alpha_l\, \Delta\alpha_m && \text{(C.3.22)}
\end{aligned}$$

and the second derivative can be extracted:

$$\frac{\partial^2}{\partial\alpha_l\, \partial\alpha_m} \boldsymbol{x}_i^{(1)} = \frac{\partial}{\partial\alpha_l} R(\hat{n}_{l+1}, \alpha_l)|_{\alpha_l=0} \frac{\partial}{\partial\alpha_m} R(\hat{n}_{m+1}, \alpha_m)|_{\alpha_m=0}\, \boldsymbol{v}_{m+1,i} \qquad \text{(C.3.23)}$$

C.3.7. Remaining Derivatives

The derivatives for angle-dihedral and dihedral-dihedral are calculated accordingly and are given in the main text in Tab. 4.2.1.

Bibliography

[1] C. Groves, Organic light-emitting diodes bright design, Nature Materials 12 (7) (2013) 597–598.

[2] L. S. Hung, C. H. Chen, Recent progress of molecular organic electroluminescent materials and devices, Materials Science & Engineering R-Reports 39 (5-6) (2002) 143–222.

[3] G. Li, R. Zhu, Y. Yang, Polymer solar cells, Nature Photonics 6 (3) (2012) 153–161.

[4] B. Kippelen, J.-L. Bredas, Organic photovoltaics, Energy & Environmental Science 2 (3) (2009) 251–261.

[5] A. W. Hains, Z. Liang, M. A. Woodhouse, B. A. Gregg, Molecular semiconductors in organic photovoltaic cells, Chemical Reviews 110 (11) (2010) 6689–6735.

[6] C. D. Dimitrakopoulos, P. R. L. Malenfant, Organic thin film transistors for large area electronics, Advanced Materials 14 (2) (2002) 99–117.

[7] C. N. Hoth, P. Schilinsky, S. A. Choulis, S. Balasubramanian, C. J. Brabec, Solution-processed organic photovoltaics, in: Applications of Organic and Printed Electronics, Springer, 2013, pp. 27–56.

[8] J. Hast, M. Tuomikoski, R. Suhonen, K.-L. Väisänen, M. Välimäki, T. Maaninen, P. Apilo, A. Alastalo, A. Maanineny, 18.1: Invited paper: Roll-to-roll manufacturing of printed oleds, in: SID Symposium Digest of Technical Papers, Wiley Online Library, 2013, pp. 192–195.

[9] C. W. Tang, S. A. VanSlyke, Organic electroluminescent diodes, Applied physics letters 51 (12) (1987) 913–915.

[10] M. A. Green, K. Emery, Y. Hishikawa, W. Warta, E. D. Dunlop, Solar cell efficiency tables (version 49), Progress in photovoltaics: research and applications 20 (1) (2015) 12–20.

[11] S. Shrestha, Photovoltaics literature survey (no. 104), Progress in Photovoltaics: Research and Applications 21 (6) (2013) 1429–1431.

[12] N. T. Kalyani, S. Dhoble, Organic light emitting diodes: energy saving lighting technology—a review, Renewable and Sustainable Energy Reviews 16 (5) (2012) 2696–2723.

[13] I. G. Lezama, A. F. Morpurgo, Progress in organic single-crystal field-effect transistors, Mrs Bulletin 38 (1) (2013) 51–56.

[14] M. Mesta, M. Carvelli, R. J. de Vries, H. van Eersel, J. J. M. van der Holst, M. Schober, M. Furno, B. Luessem, K. Leo, P. Loebl, R. Coehoorn, P. A. Bobbert, Molecular-scale simulation of electroluminescence in a multilayer white organic light-emitting diode, Nature Materials 12 (7) (2013) 652–658.

[15] D. Sun, S. Ma, Y. Ke, D. J. Collins, H.-C. Zhou, An interweaving mof with high hydrogen uptake, Journal of the American Chemical Society 128 (12) (2006) 3896–3897.

[16] J. An, N. L. Rosi, Tuning mof co2 adsorption properties via cation exchange, Journal of the American Chemical Society 132 (16) (2010) 5578–5579.

[17] J. Liu, P. K. Thallapally, B. P. McGrail, D. R. Brown, J. Liu, Progress in adsorption-based co2 capture by metal-organic frameworks, Chem. Soc. Rev. 41 (2012) 2308–2322.

[18] O. K. Farha, A. M. Shultz, A. A. Sarjeant, S. T. Nguyen, J. T. Hupp, Active-site-accessible, porphyrinic metal−organic framework materials, Journal of the American Chemical Society 133 (15) (2011) 5652–5655.

[19] A. M. Shultz, O. K. Farha, J. T. Hupp, S. T. Nguyen, A catalytically active, permanently microporous mof with metalloporphyrin struts, Journal of the American Chemical Society 131 (12) (2009) 4204–4205.

[20] L. E. Kreno, K. Leong, O. K. Farha, M. Allendorf, R. P. V. Duyne, J. T. Hupp, Metal–organic framework materials as chemical sensors, Chemical Reviews 112 (2) (2012) 1105–1125.

[21] D. Y. Lee, D. V. Shinde, S. J. Yoon, K. N. Cho, W. Lee, N. K. Shrestha, S.-H. Han, Cu-based metal–organic frameworks for photovoltaic application, The Journal of Physical Chemistry C 118 (30) (2014) 16328–16334.

[22] R. J. Kuppler, D. J. Timmons, Q.-R. Fang, J.-R. Li, T. A. Makal, M. D. Young, D. Yuan, D. Zhao, W. Zhuang, H.-C. Zhou, Potential applications of metal-organic frameworks, Coordination Chemistry Reviews 253 (23–24) (2009) 3042 – 3066.

[23] N. Stock, S. Biswas, Synthesis of metal-organic frameworks (mofs): Routes to various mof topologies, morphologies, and composites, Chemical Reviews 112 (2) (2012) 933–969.

[24] B. Liu, O. Shekhah, H. K. Arslan, J. Liu, C. Wöll, R. A. Fischer, Enantiopure metal–organic framework thin films: oriented surmof growth and enantioselective adsorption, Angewandte Chemie International Edition 51 (3) (2012) 807–810.

[25] A. A. Talin, A. Centrone, A. C. Ford, M. E. Foster, V. Stavila, P. Haney, R. A. Kinney, V. Szalai, F. El Gabaly, H. P. Yoon, F. Léonard, M. D. Allendorf, Tunable electrical conductivity in metal-organic framework thin-film devices, Science 343 (6166) (2014) 66–69.

[26] P. Friederich, F. Symalla, V. Meded, T. Neumann, W. Wenzel, Ab initio treatment of disorder effects in amorphous organic materials: Toward parameter free materials simulation, Journal of Chemical Theory and Computation 10 (9) (2014) 3720–3725.

[27] V. Rodin, F. Symalla, V. Meded, P. Friederich, D. Danilov, A. Poschlad, G. Nelles, F. von Wrochem, W. Wenzel, Generalized effective-medium model for the carrier mobility in amorphous organic semiconductors, Physical Review B 91 (15) (2015) 155203.

[28] J. J. Kwiatkowski, J. Nelson, H. Li, J. L. Bredas, W. Wenzel, C. Lennartz, Simulating charge transport in tris(8-hydroxyquinoline) aluminium (alq3), Physical Chemistry Chemical Physics 10 (14) (2008) 1852–1858.

[29] P. Kordt, J. J. van der Holst, M. Al Helwi, W. Kowalsky, F. May, A. Badinski, C. Lennartz, D. Andrienko, Modeling of organic light emitting diodes: From molecular to device properties, Advanced Functional Materials 25 (13) (2015) 1955–1971.

[30] T. Mönch, T. S. Sherkar, L. J. A. Koster, P. Friederich, M. Riede, P. Formanek, C. Koerner, K. Vandewal, W. Wenzel, K. Leo, Experimental and theoretical study of phase separation in znpc: C 60 blends, Organic Electronics 27 (2015) 183–191.

[31] S. Mohr, L. E. Ratcliff, L. Genovese, D. Caliste, P. Boulanger, S. Goedecker, T. Deutsch, Accurate and efficient linear scaling dft calculations with universal applicability, Physical Chemistry Chemical Physics 17 (47) (2015) 31360–31370.

[32] A. K. Rappé, C. J. Casewit, K. Colwell, W. Goddard Iii, W. Skiff, Uff, a full periodic table force field for molecular mechanics and molecular dynamics simulations, Journal of the American chemical society 114 (25) (1992) 10024–10035.

[33] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids, Journal of the American Chemical Society 118 (45) (1996) 11225–11236.

[34] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general amber force field, Journal of computational chemistry 25 (9) (2004) 1157–1174.

[35] D. Dubbeldam, S. Calero, T. Vlugt, R. Krishna, T. L. Maesen, E. Beerdsen, B. Smit, Force field parametrization through fitting on inflection points in isotherms, Physical review letters 93 (8) (2004) 088302.

[36] V. Ruehle, A. Lukyanov, F. May, M. Schrader, T. Vehoff, J. Kirkpatrick, B. Baumeier, D. Andrienko, Microscopic simulations of charge transport in disordered organic semiconductors, Journal of Chemical Theory and Computation 7 (10) (2011) 3335–3345.

[37] D. E. Shaw et al., Anton, a special-purpose machine for molecular dynamics simulation, ACM SIGARCH Computer Architecture News 35 (2) (2007) 1–12.

[38] D. E. Shaw et al., Millisecond-scale molecular dynamics simulations on anton, in: High Performance Computing Networking, Storage and Analysis, Proceedings of the Conference on, 2009, pp. 1–11.

[39] W. K. Hastings, Monte carlo sampling methods using markov chains and their applications, Biometrika 57 (1) (1970) 97–109.

[40] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller, Equation of state calculations by fast computing machines, The Journal of Chemical Physics 21 (6) (1953) 1087–1092.

[41] M. Rao, C. Pangali, B. J. Berne, On the force bias monte carlo simulation of water: methodology, optimization and comparison with molecular dynamics, Molecular Physics 37 (6) (1979) 1773–1798.

[42] G. Dereli, Stillinger-weber type potentials in monte carlo simulation of amorphous silicon, Molecular Simulation 8 (6) (1992) 351–360.

[43] M. Timonova, J. Groenewegen, B. J. Thijsse, Modeling diffusion and phase transitions by a uniform-acceptance force-bias monte carlo method, Phys. Rev. B 81 (14) (2010) 144107.

[44] E. C. Neyts, B. J. Thijsse, M. J. Mees, K. M. Bal, G. Pourtois, Establishing uniform acceptance in force biased monte carlo simulations, Journal of Chemical Theory and Computation 8 (6) (2012) 1865–1869.

[45] U. Hansmann, Free energy landscape and folding mechanism of a β-hairpin in explicit water: A replica exchange molecular dynamics study, Chem. Phys. Lett 281 (1997) 140–150.

[46] B. Lüssem, M. Riede, K. Leo, Doping of organic semiconductors, physica status solidi (a) 210 (1) (2013) 9–43.

[47] S. Scholz, D. Kondakov, B. Lüssem, K. Leo, Degradation mechanisms and reactions in organic light-emitting devices, Chemical Reviews 115 (16) (2015) 8449–8503.

[48] V. Coropceanu, J. Cornil, D. A. da Silva Filho, Y. Olivier, R. Silbey, J.-L. Brédas, Charge transport in organic semiconductors, Chemical reviews 107 (4) (2007) 926–952.

[49] D. L. Cheung, A. Troisi, Modelling charge transport in organic semiconductors: from quantum dynamics to soft matter, Physical Chemistry Chemical Physics 10 (39) (2008) 5941–5952.

[50] O. V. Mikhnenko, P. W. M. Blom, T.-Q. Nguyen, Exciton diffusion in organic semiconductors, Energy Environ. Sci. 8 (2015) 1867–1888.

[51] A. Mishra, P. Bäuerle, Small molecule organic semiconductors on the move: Promises for future solar energy technology, Angewandte Chemie International Edition 51 (9) (2012) 2020–2067.

[52] I. Kondov, Multiscale Modelling Methods for Applications in Materials Science: CECAM Tutorial, 16-20 September 2013, Forschungszentrum Jülich, Lecture Notes, Vol. 19, Forschungszentrum Jülich, 2013.

[53] R. M. Nieminen, From atomistic simulation towards multiscale modelling of materials, Journal of Physics: Condensed Matter 14 (11) (2002) 2859.

[54] J. Nelson, J. J. Kwiatkowski, J. Kirkpatrick, J. M. Frost, Modeling charge transport in organic photovoltaic materials, Accounts of Chemical Research 42 (11) (2009) 1768–1778.

[55] A. Troisi, D. L. Cheung, D. Andrienko, Charge transport in semiconductors with multiscale conformational dynamics, Physical review letters 102 (11) (2009) 116602.

[56] P. Friederich, V. Meded, F. Symalla, M. Elstner, W. Wenzel, Qm/qm approach to model energy disorder in amorphous organic semiconductors, Journal of chemical theory and computation 11 (2) (2015) 560–567.

[57] T. Neumann, D. Danilov, C. Lennartz, W. Wenzel, Modeling disordered morphologies in organic semiconductors, Journal of Computational Chemistry 34 (31) (2013) 2716–2725.

[58] C. Lee, R. Waterland, K. Sohlberg, Prediction of charge mobility in amorphous organic materials through the application of hopping theory, Journal of Chemical Theory and Computation 7 (8) (2011) 2556–2567.

[59] P. Hohenberg, W. Kohn, Inhomogeneous electron gas, Phys. Rev. 136 (1964) B864–B871.

[60] F. Bloch, The electron theory of ferromagnetism and electrical conductivity, Z. Physik 57 (1929) 545.

[61] S. Vosko, L. Wilk, M. Nusair, Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis, Canadian Journal of physics 58 (8) (1980) 1200–1211.

[62] J. P. Perdew, A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems, Physical Review B 23 (10) (1981) 5048.

[63] L. A. Cole, J. Perdew, Calculated electron affinities of the elements, Physical Review A 25 (3) (1982) 1265.

[64] J. P. Perdew, Y. Wang, Accurate and simple analytic representation of the electron-gas correlation energy, Physical Review B 45 (23) (1992) 13244.

[65] U. von Barth, L. Hedin, A local exchange-correlation potential for the spin polarized case. i, Journal of Physics C: Solid State Physics 5 (13) (1972) 1629.

[66] O. Gunnarsson, B. Lundqvist, Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism, Physical Review B 13 (10) (1976) 4274.

[67] J. P. Perdew, Y. Wang, Pair-distribution function and its coupling-constant average for the spin-polarized electron gas, Physical Review B 46 (20) (1992) 12947.

[68] J. P. Perdew, K. Burke, Y. Wang, Generalized gradient approximation for the exchange-correlation hole of a many-electron system, Physical Review B 54 (23) (1996) 16533.

[69] A. D. Becke, Densityfunctional thermochemistry. iii. the role of exact exchange, The Journal of Chemical Physics 98 (7) (1993) 5648–5652.

[70] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, Electronic structure calculations on workstation computers: The program system turbomole, Chemical Physics Letters 162 (3) (1989) 165–169.

[71] G. Kresse, J. Furthmuller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, Computational Materials Science 6 (1) (1996) 15–50.

[72] G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Physical Review B 54 (16) (1996) 11169.

[73] C. C. J. Roothaan, New developments in molecular orbital theory, Reviews of modern physics 23 (2) (1951) 69.

[74] J. Pople, R. Nesbet, Self-consistent orbitals for radicals, The Journal of Chemical Physics 22 (3) (1954) 571–572.

[75] J. A. Pople, D. P. Santry, G. A. Segal, Approximate self-consistent molecular orbital theory. i. invariant procedures, The Journal of Chemical Physics 43 (10) (1965) S129–S135.

[76] M. J. Dewar, W. Thiel, Ground states of molecules. 38. the mndo method. approximations and parameters, Journal of the American Chemical Society 99 (15) (1977) 4899–4907.

[77] J. J. Stewart, Mopac: a semiempirical molecular orbital program, Journal of computer-aided molecular design 4 (1) (1990) 1–103.

[78] J. J. Stewart, Optimization of parameters for semiempirical methods vi: more modifications to the nddo approximations and re-optimization of parameters, Journal of molecular modeling 19 (1) (2013) 1–32.

[79] S. Sastry, P. G. Debenedetti, F. H. Stillinger, Signatures of distinct dynamical regimes in the energy landscape of a glass-forming liquid, Nature 393 (6685) (1998) 554–557.

[80] Y. Sugita, Y. Okamoto, Replica-exchange molecular dynamics method for protein folding, Chemical Physics Letters 314 (1–2) (1999) 141–151.

[81] Y. Sugita, A. Kitao, Y. Okamoto, Multidimensional replica-exchange method for free-energy calculations, The Journal of Chemical Physics 113 (15) (2000) 6042–6051.

[82] A. Schug, T. Herges, W. Wenzel, Reproducible protein folding with the stochastic tunneling method, Phys Rev Lett 91 (15) (2003) 158102.

[83] A. Lukyanov, C. Lennartz, D. Andrienko, Amorphous films of tris(8-hydroxyquinolinato)aluminium: Force-field, morphology, and charge transport, Physica Status Solidi a-Applications and Materials Science 206 (12) (2009) 2737–2742.

[84] K. Lindorff-Larsen, S. Piana, R. O. Dror, D. E. Shaw, How fast-folding proteins fold, Science 334 (6055) (2011) 517–520.

[85] K. A. Dill, J. L. MacCallum, The protein-folding problem, 50 years on, Science 338 (6110) (2012) 1042–1046.

[86] T. Strunk, M. Wolf, M. Brieg, K. Klenin, A. Biewer, F. Tristram, M. Ernst, P. J. Kleine, N. Heilmann, I. Kondov, W. Wenzel, Simona 1.0: An efficient and versatile framework for stochastic simulations of molecular and nanoscale systems, Journal of Computational Chemistry 33 (32) (2012) 2602–2613.

[87] T. Neumann, D. Danilov, C. Lennartz, W. Wenzel, Modeling disordered morphologies in organic semiconductors, Journal of Computational Chemistry 34 (31) (2013) 2716–2725.

[88] H. J. Berendsen, D. van der Spoel, R. van Drunen, Gromacs: a message-passing parallel molecular dynamics implementation, Computer Physics Communications 91 (1) (1995) 43–56.

[89] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, et al., Gromacs 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, Bioinformatics 29 (7) (2013) 845–854.

[90] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, Journal of computational physics 117 (1) (1995) 1–19.

[91] S. Plimpton, B. Hendrickson, A new parallel method for molecular dynamics simulation of macromolecular systems, Journal of Computational Chemistry 17 (3) (1996) 326–337.

[92] P. Eastman, M. S. Friedrichs, J. D. Chodera, R. J. Radmer, C. M. Bruns, J. P. Ku, K. A. Beauchamp, T. J. Lane, L.-P. Wang, D. Shukla, T. Tye, M. Houston, T. Stich, C. Klein, M. R. Shirts, V. S. Pande, Openmm 4: A reusable, extensible, hardware independent library for high performance molecular simulation, Journal of Chemical Theory and Computation 9 (1) (2013) 461–469.

[93] D. J. Wales, J. P. Doye, Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms, The Journal of Physical Chemistry A 101 (28) (1997) 5111–5116.

[94] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general amber force field, Journal of computational chemistry 25 (9) (2004) 1157–1174.

[95] J. Huang, A. D. MacKerell, Charmm36 all-atom additive protein force field: Validation based on comparison to nmr data, Journal of computational chemistry 34 (25) (2013) 2135–2145.

[96] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, Development and testing of the OPLS All-Atom force field on conformational energetics and properties of organic liquids, J. Am. Chem. Soc. 118 (45) (1996) 11225–11236.

[97] B. H. Besler, K. M. Merz, P. A. Kollman, Atomic charges derived from semiempirical methods, Journal of Computational Chemistry 11 (4) (1990) 431–439.

[98] U. C. Singh, P. A. Kollman, An approach to computing electrostatic charges for molecules, Journal of Computational Chemistry 5 (2) (1984) 129–145.

[99] J. Gao, M. A. Thompson, et al., Combined quantum mechanical and molecular mechanical methods, American Chemical Society, 1998.

[100] H. Köuppel, W. Domcke, L. Cederbaum, Multimode molecular dynamics beyond the born-oppenheimer approximation, Advances in Chemical Physics, Volume 57 (1984) 59–246.

[101] K. Laasonen, A. Pasquarello, R. Car, C. Lee, D. Vanderbilt, Car-parrinello molecular dynamics with vanderbilt ultrasoft pseudopotentials, Physical Review B 47 (16) (1993) 10142.

[102] J.-P. Ryckaert, G. Ciccotti, H. J. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, Journal of Computational Physics 23 (3) (1977) 327–341.

[103] H. C. Andersen, Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations, Journal of Computational Physics 52 (1) (1983) 24–34.

[104] D. J. Evans, B. L. Holian, The nose–hoover thermostat, The Journal of chemical physics 83 (8) (1985) 4069–4074.

[105] A. Lemak, N. Balabaev, On the berendsen thermostat, Molecular Simulation 13 (3) (1994) 177–187.

[106] P. Glassermann, Monte carlo methods in financial engineering, Springer Science.

[107] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, Optimization by simulated annealing, Science 220 (4598) (1983) 671–680.

[108] R. H. Swendsen, J.-S. Wang, Replica monte carlo simulation of spin-glasses, Phys. Rev. Lett. 57 (21) (1986) 2607–2609.

[109] K. Hukushima, K. Nemoto, Exchange monte carlo method and application to spin glass simulations, Journal of the Physical Society of Japan 65 (6) (1996) 1604–1608.

[110] T. Whitfield, L. Bu, J. Straub, Generalized parallel sampling, Physica A: Statistical Mechanics and its Applications 305 (1) (2002) 157–171.

[111] W. Kwak, U. H. Hansmann, Efficient sampling of protein structures by model hopping, Physical review letters 95 (13) (2005) 138102.

[112] C. J. Geyer, Practical markov chain monte carlo, Statistical Science (1992) 473–483.

[113] D. J. Earl, M. W. Deem, Parallel tempering: Theory, applications, and new perspectives, Physical Chemistry Chemical Physics 7 (23) (2005) 3910–3916.

[114] R. A. Marcus, Electron transfer reactions in chemistry. theory and experiment, Reviews of Modern Physics 65 (3) (1993) 599.

[115] V. Arkhipov, E. Emelianova, P. Heremans, H. Bässler, Analytic model of carrier mobility in doped disordered organic semiconductors, Physical Review B 72 (23) (2005) 235202.

[116] R. A. Marcus, N. Sutin, Electron transfers in chemistry and biology, Biochimica et Biophysica Acta (BBA)-Reviews on Bioenergetics 811 (3) (1985) 265–322.

[117] H.-S. Ren, M.-J. Ming, J.-Y. Ma, X.-Y. Li, Theoretical calculation of reorganization energy for electron self-exchange reaction by constrained density functional theory and constrained equilibrium thermodynamics, The Journal of Physical Chemistry A 117 (33) (2013) 8017–8025.

[118] S. F. Nelsen, S. C. Blackstock, Y. Kim, Estimation of inner shell marcus terms for amino nitrogen compounds by molecular orbital calculations, Journal of the American Chemical Society 109 (3) (1987) 677–682.

[119] J. Jortner, M. Bixon, A. A. Voityuk, N. Rösch, Superexchange mediated charge hopping in dna, The Journal of Physical Chemistry A 106 (33) (2002) 7599–7606.

[120] B. Baumeier, O. Stenzel, C. Poelking, D. Andrienko, V. Schmidt, Stochastic modeling of molecular charge transport networks, Physical Review B 86 (18) (2012) 184202.

[121] E. Hillard, A. Vessières, L. Thouin, G. Jaouen, C. Amatore, Ferrocene-mediated proton-coupled electron transfer in a series of ferrocifen-type breast-cancer drug candidates, Angewandte Chemie International Edition 45 (2) (2006) 285–290.

[122] M. Kasper, K. Sattler, K. Siegmann, U. Matter, H. Siegmann, The influence of fuel additives on the formation of carbon during combustion, Journal of Aerosol Science 30 (2) (1999) 217–225.

[123] D. Conroy, A. Moisala, S. Cardoso, A. Windle, J. Davidson, Carbon nanotube reactor: Ferrocene decomposition, iron particle growth, nanotube aggregation and scale-up, Chemical Engineering Science 65 (10) (2010) 2965–2977.

[124] A. J. Bergren, K. D. Harris, F. Deng, R. L. McCreery, Molecular electronics using diazonium-derived adlayers on carbon with cu top contacts: critical analysis of metal oxides and filaments, Journal of Physics: Condensed Matter 20 (37) (2008) 374117.

[125] Z. Karipidou, B. Branchi, M. Sarpasan, N. Knorr, V. Rodin, P. Friederich, T. Neumann, V. Meded, S. Rosselli, G. Nelles, W. Wenzel, M. A. Rampi, F. von Wrochem, Ultrarobust thin-film devices from self-assembled metal–terpyridine oligomers, Advanced Materials (2016) 1521–4095.

[126] H.-C. Zhou, J. R. Long, O. M. Yaghi, Introduction to metal-organic frameworks, Chemical Reviews 112 (2) (2012) 673–674.

[127] J. Liu, W. Zhou, J. Liu, I. Howard, G. Kilibarda, S. Schlabach, D. Coupry, M. Addicoat, S. Yoneda, Y. Tsutsui, T. Sakurai, S. Seki, Z. Wang, P. Lindemann, E. Redel, T. Heine, C. Woell, Photoinduced charge-carrier generation in epitaxial mof thin films: High efficiency as a result of an indirect electronic band gap?, Angewandte Chemie-International Edition 54 (25) (2015) 7441–7445.

[128] D. Y. Lee, C. Y. Shin, S. J. Yoon, H. Y. Lee, W. Lee, N. K. Shrestha, J. K. Lee, S.-H. Han, Enhanced photovoltaic performance of cu-based metal-organic frameworks sensitized solar cell by addition of carbon nanotubes, Scientific reports 4 (3930) (2014) 1–5.

[129] V. Stavila, A. A. Talin, M. D. Allendorf, Mof-based electronic and optoelectronic devices, Chemical Society Reviews 43 (16) (2014) 5994–6010.

[130] H. Q. Pham, M. Toan, P.-T. Nguyen-Nguyen, Y. Kawazoe, H. Mizuseki, N.-M. Duc, Engineering of band gap in metal organic frameworks by functionalizing organic linker: A systematic density functional theory investigation, Journal of Physical Chemistry C 118 (9) (2014) 4567–4577.

[131] A. Kuc, A. Enyashin, G. Seifert, Metal-organic frameworks: Structural, energetic, electronic, and mechanical properties, Journal of Physical Chemistry B 111 (28) (2007) 8179–8186.

[132] K. T. Butler, C. H. Hendon, A. Walsh, Electronic chemical potentials of porous metal-organic frameworks, Journal of the American Chemical Society 136 (7) (2014) 2703–2706.

[133] L.-M. Yang, P. Ravindran, P. Vajeeston, M. Tilset, Properties of irmof-14 and its analogues m-irmof-14 (m= cd, alkaline earth metals): electronic structure, structural stability, chemical bonding, and optical properties, Physical Chemistry Chemical Physics 14 (14) (2012) 4713–4723.

[134] K. Trepte, S. Schwalbe, G. Seifert, Electronic and magnetic properties of dut-8 (ni), Physical Chemistry Chemical Physics 17 (26) (2015) 17122–17129.

[135] L.-M. Yang, Quantum chemistry investigation of rigid a-irmof-m0 series (a = zinc, cadmium, and alkaline-earth metals) on crystal structure, electronic structure, formation energy, chemical bonding, and optical properties, Microporous and Mesoporous Materials 183 (2014) 218–233.

[136] D. E. Vanpoucke, J. W. Jaeken, S. De Baerdemacker, K. Lejaeghere, V. Van Speybroeck, Quasi-1d physics in metal-organic frameworks: Mil-47 (v) from first principles, Beilstein journal of nanotechnology 5 (1) (2014) 1738–1748.

[137] S. Schiefer, M. Huth, A. Dobrinevski, B. Nickel, Determination of the crystal structure of substrate-induced pentacene polymorphs in fiber structured thin films, Journal of the American Chemical Society 129 (34) (2007) 10316–10317.

[138] R. Campbell, J. M. Robertson, J. Trotter, The crystal structure of hexacene, and a revision of the crystallographic data for tetracene, Acta crystallographica 15 (3) (1962) 289–290.

[139] C. Caleman, P. J. van Maaren, M. Hong, J. S. Hub, L. T. Costa, D. van der Spoel, Force field benchmark of organic liquids: density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant, Journal of chemical theory and computation 8 (1) (2011) 61–74.

[140] G. G. Malliaras, Y. Shen, D. H. Dunlap, H. Murata, Z. H. Kafafi, Nondispersive electron transport in alq3, Tech. rep., DTIC Document (2001).

[141] P.-O. Löwdin, On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals, The Journal of Chemical Physics 18 (3) (1950) 365–375.

[142] V. Stehr, J. Pfister, R. Fink, B. Engels, C. Deibel, First-principles calculations of anisotropic charge-carrier mobilities in organic semiconductor crystals, Physical Review B 83 (15) (2011) 155208.

[143] S. F. Nelsen, M. N. Weaver, Y. Luo, J. R. Pladziewicz, L. K. Ausman, T. L. Jentzsch, J. J. O'Konek, Estimation of electronic coupling for intermolecular electron transfer

from cross-reaction data, The Journal of Physical Chemistry A 110 (41) (2006) 11665–11676.

[144] A. D. Becke, A new mixing of hartree–fock and local density-functional theories, The Journal of Chemical Physics 98 (2) (1993) 1372–1377.

[145] A. Schäfer, C. Huber, R. Ahlrichs, Fully optimized contracted gaussian basis sets of triple zeta valence quality for atoms li to kr, The Journal of Chemical Physics 100 (8) (1994) 5829–5835.

[146] A. Schäfer, H. Horn, R. Ahlrichs, Fully optimized contracted gaussian basis sets for atoms li to kr, The Journal of Chemical Physics 97 (4) (1992) 2571–2577.

[147] R. Coehoorn, W. Pasveer, P. Bobbert, M. Michels, Charge-carrier concentration dependence of the hopping mobility in organic materials with gaussian disorder, Physical Review B 72 (15) (2005) 155206.

[148] Y. Roichman, Y. Preezant, N. Tessler, Analysis and modeling of organic devices, physica status solidi (a) 201 (6) (2004) 1246–1262.

[149] H. Fong, S. So, Hole transporting properties of tris (8-hydroxyquinoline) aluminum (alq3), Journal of applied physics 100 (9) (2006) 094502.

[150] S. Naka, H. Okada, H. Onnagawa, Y. Yamaguchi, T. Tsutsui, Carrier transport properties of organic materials for el device operation, Synthetic metals 111 (2000) 331–333.

[151] R. Kepler, P. Beeson, S. Jacobs, R. Anderson, M. Sinclair, V. Valencia, P. Cahill, Electron and hole mobility in tris (8-hydroxyquinolinolato-n1, o8) aluminum, Applied physics letters 66 (26) (1995) 3618–3620.

[152] J. Kalinowski, N. Camaioni, P. Di Marco, V. Fattori, A. Martelli, Kinetics of charge carrier recombination in organic light-emitting diodes, Applied physics letters 72 (5) (1998) 513–515.

[153] A. G. Mückl, S. Berleb, W. Brütting, M. Schwoerer, Transient electroluminescence measurements on organic heterolayer light emitting diodes, Synthetic metals 111 (2000) 91–94.

[154] K. Okumoto, K. Wayaku, T. Noda, H. Kageyama, Y. Shirota, Amorphous molecular materials: charge transport in the glassy state of n, n′-di (biphenylyl)-n, n′-diphenyl-[1, 1′-biphenyl]-4, 4′-diamines, Synthetic metals 111 (2000) 473–476.

[155] P. Borsenberger, J. Shi, Hole transport in a vapor deposited phenylenediamine molecular glass, physica status solidi (b) 191 (2) (1995) 461–469.

[156] T. Mori, E. Sugimura, T. Mizutani, Estimate of hole mobilities of some organic photoconducting materials using the time-of-flight method, Journal of Physics D: Applied Physics 26 (3) (1993) 452.

[157] S. Tse, K. Kwok, S. So, Electron transport in naphthylamine-based organic compounds, Applied physics letters 89 (26) (2006) 262102.

[158] M. Stolka, J. Yanus, D. Pai, Hole transport in solid solutions of a diamine in polycarbonate, The Journal of Physical Chemistry 88 (20) (1984) 4707–4714.

[159] R. De Boer, T. Klapwijk, A. Morpurgo, Field-effect transistors on tetracene single crystals, Applied physics letters 83 (21) (2003) 4345–4347.

[160] H. Klauk, M. Halik, U. Zschieschang, G. Schmid, W. Radlik, W. Weber, High-mobility polymer gate dielectric pentacene thin film transistors, Journal of Applied Physics 92 (9) (2002) 5259–5263.

[161] M. Ratner, A. Aviram, Molecular rectifiers, Chem. Phys. Lett 29 (1974) 277–283.

[162] R. L. Carroll, C. B. Gorman, The genesis of molecular electronics, Angewandte Chemie International Edition 41 (23) (2002) 4378–4400.

[163] H. B. Akkerman, B. de Boer, Electrical conduction through single molecules and self-assembled monolayers, Journal of Physics: Condensed Matter 20 (1) (2007) 013001.

[164] T.-W. Kim, G. Wang, H. Lee, T. Lee, Statistical analysis of electronic properties of alkanethiols in metal–molecule–metal junctions, Nanotechnology 18 (31) (2007) 315204.

[165] H. Haick, J. Ghabboun, D. Cahen, Pd versus au as evaporated metal contacts to molecules, Applied Physics Letters 86 (4) (2005) 042113.

[166] H. Haick, M. Ambrico, J. Ghabboun, T. Ligonzo, D. Cahen, Contacting organic molecules by metal evaporation, Physical Chemistry Chemical Physics 6 (19) (2004) 4538–4541.

[167] A. Vilan, D. Cahen, Soft contact deposition onto molecularly modified gaas. thin metal film flotation: principles and electrical effects, Advanced Functional Materials 12 (11-12) (2002) 795–807.

[168] K. T. Shimizu, J. D. Fabbri, J. J. Jelincic, N. A. Melosh, Soft deposition of large-area metal contacts for molecular electronics, Advanced Materials 18 (12) (2006) 1499–1504.

[169] H. B. Akkerman, P. W. Blom, D. M. De Leeuw, B. De Boer, Towards molecular electronics with large-area molecular junctions, Nature 441 (7089) (2006) 69–72.

[170] T. Li, J. R. Hauptmann, Z. Wei, S. Petersen, N. Bovet, T. Vosch, J. Nygård, W. Hu, Y. Liu, T. Bjørnholm, et al., Solution-processed ultrathin chemically derived graphene films as soft top contacts for solid-state molecular electronic junctions, Advanced Materials 24 (10) (2012) 1333–1339.

[171] M. J. Preiner, N. A. Melosh, Creating large area molecular electronic junctions using atomic layer deposition, Applied Physics Letters 92 (21) (2008) 213301.

[172] A. J. Bergren, K. D. Harris, F. Deng, R. L. McCreery, Molecular electronics using diazonium-derived adlayers on carbon with cu top contacts: critical analysis of metal oxides and filaments, Journal of Physics: Condensed Matter 20 (37) (2008) 374117.

[173] A. D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, Physical review A 38 (6) (1988) 3098.

[174] J. P. Perdew, Density-functional approximation for the correlation energy of the inhomogeneous electron gas, Physical Review B 33 (12) (1986) 8822.

[175] J. P. Perdew, Erratum: Density-functional approximation for the correlation energy of the inhomogeneous electron gas, Physical Review B 34 (10) (1986) 7406.

[176] F. Weigend, R. Ahlrichs, Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: design and assessment of accuracy, Physical Chemistry Chemical Physics 7 (18) (2005) 3297–3305.

[177] J. C. Scott, G. G. Malliaras, Charge injection and recombination at the metal–organic interface, Chemical Physics Letters 299 (2) (1999) 115–119.

[178] S. S.-Y. Chui, S. M.-F. Lo, J. P. Charmant, A. G. Orpen, I. D. Williams, A chemically functionalizable nanoporous material [cu3 (tma) 2 (h2o) 3] n, Science 283 (5405) (1999) 1148–1150.

[179] K. Schlichte, T. Kratzke, S. Kaskel, Improved synthesis, thermal stability and catalytic properties of the metal-organic framework compound cu3(btc)(2), Microporous and Mesoporous Materials 73 (1-2) (2004) 81–88.

[180] L. Alaerts, E. Seguin, H. Poelman, F. Thibault-Starzyk, P. A. Jacobs, D. E. De Vos, Probing the lewis acidity and catalytic activity of the metal-organic framework cu-3(btc)(2) (btc = benzene-1,3,5-tricarboxylate), Chemistry-a European Journal 12 (28) (2006) 7353–7363.

[181] L. Heinke, Z. Gu, C. Wöll, The surface barrier phenomenon at the loading of metal-organic frameworks, Nature communications 5 (4562) (2014) 1–6.

[182] Z.-G. Gu, L. Heinke, C. Wöll, T. Neumann, W. Wenzel, Q. Li, K. Fink, O. D. Gordan, D. R. T. Zahn, Experimental and theoretical investigations of the electronic band structure of metal-organic frameworks of hkust-1 type, Applied Physics Letters 107 (18).

[183] P. E. Blöchl, Projector augmented-wave method, Physical Review B 50 (24) (1994) 17953.

[184] G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Physical Review B 59 (3) (1999) 1758.

[185] E. Runge, E. K. Gross, Density-functional theory for time-dependent systems, Physical Review Letters 52 (12) (1984) 997.

[186] H. J. Monkhorst, J. D. Pack, Special points for brillouin-zone integrations, Physical Review B 13 (12) (1976) 5188.

[187] M. Kühn, F. Weigend, Phosphorescence lifetimes of organic light-emitting diodes from two-component time-dependent density functional theory, The Journal of chemical physics 141 (22) (2014) 224302.

[188] O. Laporte, W. F. Meggers, Some rules of spectral structure, JOSA 11 (5) (1925) 459–460.

[189] C. H. Hendon, A. Walsh, Chemical principles underpinning the performance of the metal–organic framework hkust-1, Chemical Science 6 (7) (2015) 3674–3683.

[190] H. L. Skriver, N. Rosengaard, Surface energy and work function of elemental metals, Physical Review B 46 (11) (1992) 7157.

[191] K. J. Erickson, F. Léonard, V. Stavila, M. E. Foster, C. D. Spataru, R. E. Jones, B. M. Foley, P. E. Hopkins, M. D. Allendorf, A. A. Talin, Thin film thermoelectric metal–organic framework with high seebeck coefficient and low thermal conductivity, Advanced Materials 27 (22) (2015) 3453–3459.

[192] A. R. Leach, Molecular Modelling: Principles and Applications, 2nd Edition, Pearson Education Limited, Harlow, 2001.

[193] M. Tuckerman, B. J. Berne, G. J. Martyna, Reversible multiple time scale molecular dynamics, The Journal of Chemical Physics 97 (3) (1992) 1990–2001.

[194] J. A. Izaguirre, S. Reich, R. D. Skeel, Longer time steps for molecular dynamics, The Journal of Chemical Physics 110 (20) (1999) 9853–9864.

[195] J.-S. Wang, R. H. Swendsen, Monte carlo renormalization-group study of ising spin glasses, Physical Review B 37 (13) (1988) 7745–7750.

[196] J.-S. Wang, R. H. Swendsen, Cluster monte carlo algorithms, Physica A: Statistical Mechanics and its Applications 167 (3) (1990) 565–579.

[197] J. S. Liu, F. Liang, W. H. Wong, The multiple-try method and local optimization in metropolis sampling, Journal of the American Statistical Association 95 (449) (2000) 121–134.

[198] R. Craiu, C. Lemieux, Acceleration of the multiple-try metropolis algorithm using antithetic and stratified sampling, Statistics and Computing 17 (2) (2007) 109–120.

[199] D. A. Kofke, Getting the most from molecular simulation, Molecular Physics 102 (4) (2004) 405–420.

[200] C. Pangali, M. Rao, B. J. Berne, On a novel monte carlo scheme for simulating water and aqueous solutions, Chemical Physics Letters 55 (3) (1978) 413–417.

[201] S. Whitelam, P. L. Geissler, Avoiding unphysical kinetic traps in monte carlo simulations of strongly attractive particles, The Journal of Chemical Physics 127 (15) (2007) 1–20.

[202] S. Whitelam, P. L. Geissler, Erratum: Avoiding unphysical kinetic traps in monte carlo simulations of strongly attractive particles [j. chem. phys.127, 154101 (2007)], The Journal of Chemical Physics 128 (21) (2008) 1–3.

[203] L. Velazquez, J. C. Castro-Palacio, Improving the efficiency of monte carlo simulations of systems that undergo temperature-driven phase transitions, Physical Review E 88 (1) (2013) 013311.

[204] M. J. Mees, G. Pourtois, E. C. Neyts, B. J. Thijsse, A. Stesmans, Uniform-acceptance force-bias monte carlo method with time scale to study solid-state diffusion, Physical Review B 85 (13) (2012) 134301.

[205] J. Liu, E. Luijten, Rejection-free geometric cluster algorithm for complex fluids, Physical Review Letters 92 (3) (2004) 035504.

[206] W. Krauth, Quantum monte carlo calculations for a large number of bosons in a harmonic trap, Physical Review Letters 77 (18) (1996) 3695–3699.

[207] M. P. Allen, D. J. Tildesley, Computer simulation of liquids, Oxford University Press, 1989.

[208] N. Heilmann, J. Setzler, M. Brieg, T. Strunk, M. Wolf, C. Seith, W. Wenzel, Thermodynamic characterization of protein folding equilibriums at the all atom level, Biophysical journal 104 (2) (2013) 369a–370a.

[209] T. Neumann, D. Danilov, W. Wenzel, Multiparticle moves in acceptance rate optimized monte carlo, Journal of computational chemistry 36 (30) (2015) 2236–2245.

[210] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, M. Parrinello, Plumed: A portable plugin for free-energy calculations with molecular dynamics, Computer Physics Communications 180 (10) (2009) 1961–1972.

[211] L. Alessandro, L. G. Francesco, Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science, Reports on Progress in Physics 71 (12) (2008) 126601.

[212] V. Limongelli, M. Bonomi, M. Parrinello, Funnel metadynamics as accurate binding free-energy method, Proceedings of the National Academy of Sciences 6358-6363 (16) (2013) 110.

[213] P. J. Steinhardt, D. R. Nelson, M. Ronchetti, Bond-orientational order in liquids and glasses, Physical Review B 28 (2) (1983) 784–805.

Danksagung

Am Ende meiner Promotion gilt es, all denjenigen zu danken, die mich während der dreieinhalb Jahre inhaltlich, ideologisch, finanziell, materiell oder mental unterstützt haben.
Dabei möchte ich zunächst der Landesgraduiertenförderung des Landes Baden-Württemberg danken, durch deren Finanzierung ich die in dieser Arbeit präsentierten Studien anfertigen konnte. An zweiter Stelle gilt mein aufrichtiger Dank Prof. Dr. Wolfgang Wenzel, der in vielerlei Hinsicht zum Gelingen dieser Arbeit beigetragen hat: zum einen durch die Bereitstellung der Projekte und Entwicklung der Forschungsvorhaben, zum anderen durch intensive Betreuung, bereitwillige Hilfestellungen und ergiebige Diskussionen bis in die philosophischen Tiefen des detaillierten Gleichgewichts, sowie durch die frühe Einbindung in eine Vielzahl an Kollaborationen. Desweiteren danke ich Prof. Dr. Gerd Schön herzlich für die Übernahme des Korreferats dieser Arbeit.
Weiterhin sei an dieser Stelle der International Graduate School des Bio-Interfaces Programs für die zahlreichen kurzweiligen und bereichernden Kurse und Vorträge, denen ich beiwohnen durfte, gedankt. In diesem Rahmen danke ich weiterhin den Mitgliedern meines Thesis-Advisory-Committees (TAC), Prof. Dr. Christof Wöll und Prof. Dr. Stefan Bräse, für ihren Input während der TAC meetings. Bei meinen in- und externen Projektpartnern möchte ich mich für die interessanten Projekte bedanken, die es mir ermöglichten, Teile der Arbeit frühzeitig zu publizieren: den Mitgliedern der AG Wöll (IFG, KIT), PD Dr. Karin Fink und Qiang Li (INT, KIT), sowie den Mitgliedern des MSL (Sony Stuttgart).
Selbstredend ist eine solche Promotion nicht im Alleingang möglich und die Unterstützung durch eine kompetente Gruppe unbezahlbar. Daher danke ich herzlich allen Mitgliedern der AG Wenzel: Denis für die Betreuung in den ersten Jahren und für zahlreiche Hilfestellungen während der gesamten Arbeit; Velimir für das 10-Uhr-ICE-Bier, schöne Wetteinsätze, musikalische Exkurse und die immer volle Packung Kaugummis auf seinem Schreibtisch, sowie nicht zuletzt für die Hilfestellungen und Korrekturen beim Anfertigen der Dissertationsschrift; Timo für die unerschütterliche Geduld bei der Hilfe mit sämtlichen großen und kleinen Software-Problemen; Franz und Pascal für so manche QM-Erleuterung für ignorante Molecular-Mechanics-Kollegen und, ebenso wie Florian, für das Korrekturlesen meiner Dissertation. Den übrigen (auch ehemaligen) Kollegen möchte ich für eine schöne Athmosphäre, die Bake-Offs und insgesamt viereinhalb schöne Jahre danken. Ein kleines Dankeschön geht an die Kaffeepausen Crew, Nadine, Geertje und Simon, für die erheiternden Minuten in der Cafeteria.
Zu guter Letzt danke ich meiner gesamten Familie für die Unterstützung jeglicher Art während Schulzeit, Studium und Promotion, sowie meinem privaten Fahrradmechaniker Moritz Sander für die vielen hands-on-Kurse im Schrauben, durch die es mir gelang, den alten Renn-Esel während der insgesamt viereinhalb Jahre größtenteils pannenfrei über schätzungsweise 20000 km Arbeitsweg zu pedalieren.

Publications

[1] T. Neumann, D. Danilov, C. Lennartz, W. Wenzel, Modeling disordered morphologies in organic semiconductors, Journal of Computational Chemistry 34 (31) (2013) 2716–2725.

[2] W. Guo, J. Liu, P. G. Weidler, J. Liu, T. Neumann, D. Danilov, W. Wenzel, C. Feldmann, C. Wöll, Loading of ionic compounds into metal–organic frameworks: a joint theoretical and experimental study for the case of la 3+, Physical Chemistry Chemical Physics 16 (33) (2014) 17918–17923.

[3] P. Friederich, F. Symalla, V. Meded, T. Neumann, W. Wenzel, Ab initio treatment of disorder effects in amorphous organic materials: Toward parameter free materials simulation, Journal of Chemical Theory and Computation 10 (9) (2014) 3720–3725.

[4] L. E. Ratcliff, L. Grisanti, L. Genovese, T. Deutsch, T. Neumann, D. Danilov, W. Wenzel, D. Beljonne, J. Cornil, Toward fast and accurate evaluation of charge on-site energies and transfer integrals in supramolecular architectures using linear constrained density functional theory (cdft)-based methods, Journal of Chemical Theory and Computation 11 (5) (2015) 2077–2086.

[5] T. Neumann, D. Danilov, W. Wenzel, Multiparticle moves in acceptance rate optimized monte carlo, Journal of computational chemistry 36 (30) (2015) 2236–2245.

[6] Z.-G. Gu, L. Heinke, C. Wöll, T. Neumann, W. Wenzel, Q. Li, K. Fink, O. D. Gordan, D. R. T. Zahn, Experimental and theoretical investigations of the electronic band structure of metal-organic frameworks of hkust-1 type, Applied Physics Letters 107 (18) (2015) 1–6.

[7] Z.-G. Gu, H. Fu, T. Neumann, Z.-X. Xu, W.-Q. Fu, W. Wenzel, L. Zhang, J. Zhang, C. Wöll, Chiral porous metacrystals: Employing liquid-phase epitaxy to assemble enantiopure metal–organic nanoclusters into molecular framework pores, ACS Nano 10 (1) (2016) 977–983.

[8] W. Guo, C. Zhi, C. Yang, T. Neumann, C. Kübel, W. Wenzel, A. Welle, W. Pfleging, O. Shekhah, C. Wöll, E. Redel, Bi2o3 nanoparticles encapsulated in surface mounted metal-organic frameworks thin films, Nanoscale 8 (2016) 6468–6472.

[9] Z. Karipidou, B. Branchi, M. Sarpasan, N. Knorr, V. Rodin, P. Friederich, T. Neumann, V. Meded, S. Rosselli, G. Nelles, W. Wenzel, M. A. Rampi, F. von Wrochem, Ultrarobust thin-film devices from self-assembled metal–terpyridine oligomers, Advanced Materials (2016) 1521–4095.