CHARGE TRANSPORT IN DISORDERED MATERIALS

by

Adam Gerald Gagorik

B.S. University of Pittsburgh, 2008

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Adam Gerald Gagorik

It was defended on

December 3, 2013

and approved by

Geoffrey R. Hutchison, Professor, Department of Chemistry

Lillian T. Chong, Professor, Department of Chemistry

Daniel S. Lambrecht, Professor, Department of Chemistry

Tomasz Kowalewski, Professor, Department of Chemistry, Carnegie Mellon University

Dissertation Director: Geoffrey R. Hutchison, Professor, Department of Chemistry

CHARGE TRANSPORT IN DISORDERED MATERIALS

Adam Gerald Gagorik, PhD

University of Pittsburgh, 2013

This thesis is focused on on using Monte Carlo simulation to extract device relevant properties, such as the current voltage behavior of transistors and the efficiency of photovoltaics, from the hopping transport of molecules. Specifically, simulation is used to study organic field-effect transistors (OFETs) and organic photo-voltaics (OPVs). For OFETs, the current was found to decrease with increasing concentration of traps and barriers in the system. As the barrier/trap concentration approaches 100 %, the current recovers as carrier begin to travel through the manifold of connected trap states. Coulomb interactions between like charges are found to play a role in removing carriers from trap states. The equilibrium current in OFETs was found to be independent of charge injection method, however, the finite size of devices leads to an oscillatory current. Fourier transforms of the electrical current show peaks that vary non-linearly with device length, while being independent of device width. This has implications for the mobility of carriers in finite sized devices. Lastly, the presence of defects and high barriers ($> 0.4\,\mathrm{eV}$) was found to produce negative differential resistance in the saturation region of OFET curves, unlike traps. While defects and barriers prohibit carriers from reaching the drain at high voltages, the repulsive interaction between like charged carriers pushes charges around the defects.

For OPVs, the effects of device morphology and charge delocalization were studied. Fill factors increased with domain size in monolayer isotropic morphologies, but decreased for band morphologies. In single-phase systems without Coulomb interactions, astonishingly high fill factors ($\approx 70\,\%$) were found. In multilayer OPVs,a complex interplay of domain size, connectivity, tortuosity, interface trapping, and delocalization determined efficiency.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to take a moment to reflect on the help of others throughout my graduate career. First, I want to thank my family members. My mother Nancy has provided a place for me to live during my graduate studies, allowing me to become debt free. My father Gerald has provided emotional support by meeting me for dinner every Friday for the last 5 years. My brothers Ian and Ethan have provided me knowledge on computers and graphics, and have shown interest in my work.

I would also like to thank my friends and colleagues at Pitt. Thank you for helping with my project and being my friends. Tamika Madison and Eric Kratz, thank you for all the coffee. I haven't slept in years. To my group members, thank you for all the support you have given me.

I want to thank all the Professors for providing assistance, mentoring, serving on my committees, and writing letters. In particular, I want to thank the late Peter Siska for solidifying my love of chemistry. Without his assistance, I would not be where I am today. I would also like to thank David Earl for teaching me about Monte Carlo simulation. I want to thank David Pratt for the best teaching experience any teaching assistant could ask for. Thank you Akila Gothandaraman for assisting me with programming and GPU calculations.

I am also thankful to the University of Pittsburgh, Carnegie Mellon University, and Bayer Materials Science for financial support. I am grateful for the supercomputing resources provided by the University of Pittsburgh Center for Simulation and Modeling.

Last, but not least, I am indebted to Geoffrey Hutchison, my advisor. Thank you for letting me join your group and work on my project in my own way. It can not be expressed enough how positive a work environment you provide. Not many graduate students get as lucky in their advisor choice as I have.

1.0 INTRODUCTION

The mechanism of charge transport in disordered materials is highly debated and fundamentally different than charge transport in inorganic materials. For inorganic semiconductors, carriers are highly delocalized. This leads to the familiar picture of energy bands for conduction. In organic materials, conformational disorder and the coupling of electronic motions to vibrations leads to the picture of localized carriers. The charge transport is then described as hopping between these sites described by theories such as Marcus-Hush[1] electron transport and Miller-Abrahams[2] conduction. Questions remain, however. How does the hopping picture relate to the behavior of macroscopic devices? Why can we make devices based on organic semiconductors that perform as well as they currently do? How do we make better devices? To understand these questions, the relationship between charge transport and device behavior must be studied. The problem is not an easy one; the transport mechanism is still not very well understood. Nevertheless, *it is the goal of this thesis to use the molecular scale picture of hopping charges to predict device relevant properties such as efficiency and current-voltage behavior.*

This thesis aims to address the following questions. How does the presence of barriers, traps, and defects affect the current extracted from OFETs? Does the way we describe charge injection matter for the equilibrium current? Can the mechanism of negative differential resistance be described with the simulation? For OPVs, what are the effects of domain size? How does charge delocalization affect charge extraction and exciton separation? Finally, how significant of a role do the Coulomb interactions between charges play in charge transport?

A major goal of this work was to use simple models to describe charge transport in OFET and OPV devices. The inclusion of physically relevant processes, such as the interaction of charge carriers via explicit Coulomb interactions, leads to a simple, straight forward

interpretation of device behavior. This physical intuition is often absent from methods based on continuous differential equations. In the spirit of physical chemistry and modeling in general, the Monte Carlo approach allows us to include only the most significant physical processes and build up complexity only as needed.

1.1 REVIEW

Inexpensive, flexible, and synthetically variable organic materials are an alternative to inorganic materials in organic field-effect transistors (OFETs), organic light-emitting diodes (OLEDs), and organic photo-voltaics (OPVs).[3–12] Many have realized that organic materials promise to allow tailoring of device properties, lower the environmental impact, and lower the cost of mass producing devices. The devices are inexpensive to produce because most materials are easily solution processed, and can be assembled using preexisting manufacturing infrastructure, such as ink-jet or roll printing. Such devices can be recyclable, synthetically variable, and possess other unique properties, such as flexibility.

However, the low efficiency of devices compared to their inorganic counterparts still limit their use. This limitation can be overcome with a better and more fundamental framework to describe charge dynamics in these materials. For OFETs, drawing on such understanding, the parameters that most affect device operation, such as the *on-off* current ratio, the turn-on voltage, or the effects of defects and charge mobility on current, can be optimized. For OPVs, a more fundamental framework will allow one to probe the many processes that affect OPV function, such as morphology, exciton injection, and exciton recombination.

In this section, we review some concepts needed to motivate and understand this thesis. We will talk about charge hopping, charge mobility, the process of charge injection, the effects of morphology, and other simulation models used to describe charge transport.

1.1.1 Organic Materials

Common organic materials for OPVs are P3HT and PCBM, shown in Figure 1.1A and 1.1B. However, many other candidates exist.[3,11] PCBM is electron conducting, and generally called an acceptor or n-type material. The molecular volume of PCBM is around $1\,\mathrm{nm}^3$, the same size as the sites used in the Monte Carlo model. P3HT is hole conducting, and called a donor or p-type material. Pthalocyanine, shown in Figure 1.1C, is another hole conductor, also around $1\,\mathrm{nm}^2$ in size. Pthalocyanines with varying metal centers, such as Cu, Ni, or Zn, can serve as traps and barriers for OFETs compared to metal-less pthalocyanines.

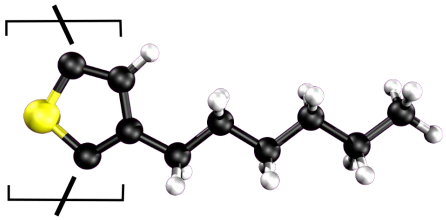
1.1.2 Charge Hopping

Charge transport in organic materials is described as hopping between localized states. We can view it as excess electrons hopping between the lowest unoccupied molecular orbitals (LUMOs) of adjacent molecules, or missing electrons (holes) hopping between the highest occupied molecular orbitals (HOMOs) of adjacent molecules. The energy bands normally present in inorganic materials localize because of the poor electronic overlap between adjacent molecules and disorder present in the system. Charge carrier can then occupy the localized states, which are essentially just orbitals. Transport is then described as hopping between these states. Two expressions can be used. The first, is due to Miller and Abrahams, and is shown in Equation 1.1, which is the transport rate, k_{ij}, for a charge hopping between two molecules.[2]

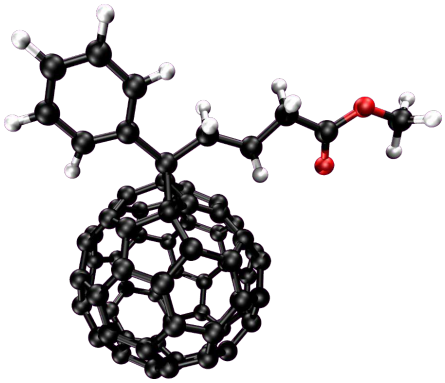
$$k_{ij} = \nu_0 e^{-2\gamma_{ij} r_{ij}} \min\left[1, e^{-\beta \Delta E}\right] \tag{1.1}$$

In Equation 1.1, the ν_0 is a hopping attempt rate, γ_{ij} is a parameter representing the coupling between molecules separated by a distance r_{ij}. ΔE is the energy change associated with a charge hopping between molecule i and molecule j, and β^{-1} is kT, where T is the temperature and k Boltzmann's constant. The equation tells us that it is exponentially less likely for charges to hop between molecules farther apart. Likewise, the hopping is exponentially less likely the more positive the energy change. If the energy change is negative, then the transport rate only depends on the distance between the molecules. The Miller-Abrahams formalism allows for the possibility of transport at low temperatures, for example,

(A) P3HT



(B) PCBM



(A) metal-PC

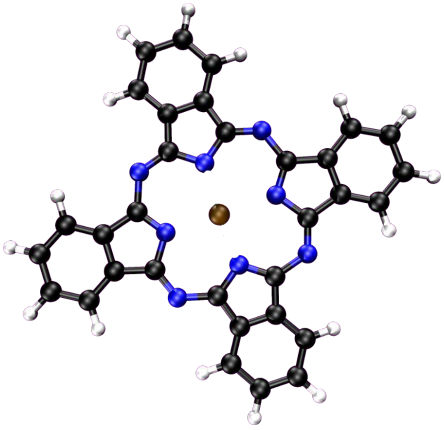


Figure 1.1: Common organic semiconductors.

through tunneling.

The second formalism is the Marcus-Hush electron transport theory. According to Marcus-Hush, the transport rate, k_{ij} is given by Equation 1.2.

$$k_{ij} = \frac{2\pi < H_{ab} >^2}{\hbar} \frac{1}{\sqrt{4\pi\lambda kT}} e^{-\frac{(\Delta G+\lambda)^2}{4\lambda kT}} \tag{1.2}$$

In Equation 1.2, $\hbar$ is Planck's constant, $\langle H_{AB} \rangle$ is the electronic coupling between states, λ is the reorganization energy, and ΔG is the free energy change. Both formalisms are well adapted for Monte Carlo techniques. Each involves a rate in terms of a pre-factor and a Boltzmann like exponential. The Miller-Abrahams is even in the form of a Metropolis acceptance criterion.[13] The probability of a process and the rate of the process are closely tied. For example, consider Equations 1.3 and 1.4.

$$R = \frac{S}{T} \tag{1.3}$$

$$P = \frac{S}{A} \tag{1.4}$$

P is the probability of a process occuring, R is the rate of the process, A is the number of attempts of the process, and S is the number of successes of the process. The probability of the process can be directly related to the rate, as shown in Equation 1.5.

$$P = \frac{RT}{A} \tag{1.5}$$

1.1.3 Mobility

Mobility is the "constant" of proportionality between the drift velocity of carriers and the electric field, as shown in Equation 1.6.

$$v = \mu E \tag{1.6}$$

In Equation 1.6, v is the drift velocity of carriers, μ is the mobility, and E is the electric field. μ has units of $(\mathrm{cm\ s^{-1}})/(\mathrm{V\ cm^{-1}}) = \mathrm{cm^2\ V^{-1}\ s^{-1}}$. The mobility of organic materials is low. For example, in the best of cases, consider a mobility of $10^{-1}\,\mathrm{cm^2\ V^{-1}\ s^{-1}} = 10\,\mathrm{nm^2\ V^{-1}\ ps^{-1}}$. In an OFET with a length of 1000 nm, when a 100 V bias is applied between the source and drain, the carrier drift velocity is only $1\,\mathrm{nm\ ps^{-1}} = 1000\,\mathrm{m\ s^{-1}}$. Compare this to a mobility of $1500\,\mathrm{cm^2\ V^{-1}\ s^{-1}}$, which is close to the mobility of electrons in silicon. For the same field, the electrons would be traveling with a drift velocity of $1.5 \times 10^4\,\mathrm{nm\ ps^{-1}} = 1.5 \times 10^7\,\mathrm{m\ s^{-1}}$, 5 % of the speed of light.

The mobility need not be a constant. For example, the Poole-Frenkel effect predicts the mobility to scale exponentially with the square root of the electric field[14] at high electric field. Monte Carlo disorder models can capture this effect.[15]

$$\mu \propto e^{\sqrt{E}} \tag{1.7}$$

We can calculate the mobility from the source-drain current, I_{SD} in an OFET simulation.[16] The source drain current is given by Equation 1.8.

$$I_{SD} = \frac{\mu \epsilon W}{LD}\left[(V_{SG} - V_0)V_{SD} - \frac{V_{SD}^2}{2}\right] \tag{1.8}$$

In Equation 1.8, ϵ is the dielectric constant, W is the width of the device, L is the length of the device, V_{SG} is the source gate voltage, and V_0 is the threshold voltage. In the saturation region of the current voltage curve, $V_{SD} > V_{SG}$. In this regime, Equation 1.8 can be reduced to Equation 1.9.

$$\sqrt{I_{SD}} = \sqrt{\frac{\mu \epsilon W}{2LD}}(V_{SG} - V_0) \tag{1.9}$$

The slope of a plot of $\sqrt{I_{DS}}$ vs. V_{SG} can be used to calculate the mobility. V_{SG} is related to the carrier concentration by Equation 1.10.

$$V_{SG} = \frac{eND}{\epsilon A} \tag{1.10}$$

1.1.4 Charge Injection

The performance of a device depends upon factors intrinsic to the semiconductor, such as the resistance of the semiconductor - a property ultimately tied to the mobility of the carriers in the material. However, outside factors also alter the device performance. For example, a contact resistance is present when electrodes are connected to the device, because an energetic barrier to charge injection forms. Lowering the contact resistance is often accomplished by matching the work function with the frontier molecular orbitals of the organic semiconductor, using self assembled monolayers between the contacts,[17–19] or doping.[20,21]

Device geometry also plays a role in the injection barrier. The contract resistance in bottom contact devices is large compared to top contact devices due to poor charge electrode surface area and non-optimal morphology growth. Unfortunately, traditional lithographic techniques are tailored towards bottom-contact manufacturing.[19] Recently, however, Caironi et al. have made top gate bottom contact n-type OFETs with contact resistances in the range of 10 to 20 $k\Omega$.[22]

For systems with a small channel size, optimization of contact resistance is crucial. The overall resistance becomes dominated by the contract resistance for short channels because the intrinsic resistance is proportional to channel length. Additionally, contact resistance optimization is especially needed for n-type devices, where the Fermi energy of the noble metal electrode is often mismatched with the LUMO energy of the semiconductor. For example, Cheng et al. have made downscaled, low-voltage, n-channel OFETs based on core-cyanated perylenediimide. Although they measured a low contact resistance with Au electrodes, Ag electrodes (which have a lower work function) were needed to counteract severe contact effects in the short channel devices.[23]

1.1.5 OPV Morphology

The first OPV, based on a bilayer junction, was described in 1986, by Tang,[24] with an efficiency of 1 %, and was fabricated from copper phthalocyanine and a perylene tetracarboxylic derivative. Such devices consisting of separate p-type and n-type layers have been extensively studied, reaching efficiencies of up to 5 %.[25,26] Initial studies on the photoexcitation

in a mixture of poly-(p-phenylene vinylene) (PPV) and fullerene C_{60} by Sariciftci et al.,[27] revealed that excitons in organic materials have short lifetimes and strong binding energy, necessitating a p-n interface to facilitate charge separation.[28] A structure suitable for charge harvesting was proposed to be a bulk heterojunction (BHJ), in which p-type (donor) polymer and n-type (acceptor) fullerene are dispersed in a disordered two-phase structure, on a length scale comparable to the exciton diffusion length. Blends of regioregular poly(3-hexyl)thiophene (rr-P3HT) as the donor with fullerene derivatives (such as [6,6]-phenyl-C_{61}-butyric acid methyl ester (PCBM)) as an electron acceptor became popular due to their high efficiency.[9,17,29–38] The increased performance of P3HT/PCBM solar cells was largely attributed to their ability to form distinct phase interfaces, forming a bulk heterojunction, without macroscopic phase separation. The size and structure of a bi-continuous, two-phase bulk heterojunction is critical to exciton harvesting and charge carrier movement within OPVs, determining their overall power conversion efficiency.

The performance of OPVs is highly sensitive to the device morphology, which describes the physical connectivity and sizing of electron donors and electron acceptors within the active layer. Device morphology not only depends upon the chemical and physical behavior of the components, but also on the processes used to deposit them. Therefore, the understanding of how charge transport in OPVs is affected by morphology is of the utmost importance.

1.1.6 Charge Delocalization

In organic materials, charge transport is typically described by localized carrier hopping through a manifold of energy states. The charge carriers are localized to molecular sites through disorder of the morphology and electron-phonon coupling. In contrast, highly delocalized carriers move through the energy bands of an inorganic material.

However, delocalization can still play a role in organic materials. Charge delocalization is well known from electronic spectroscopy (i.e. π-stacked dimer excitations). It has been suggested as a mechanism for efficient long range charge separation in organic semiconductors.[39,40] For example, well ordered, semi-crystalline domains likely lead to charge

delocalization and improved device performance in organic photovoltaics (OPVs). Devices with well ordered morphologies have been reported with very large fill factors ($76 - 80\,\%$) and power conversion efficiencies of up to $8.7\,\%$.[41] However, the complete picture is not clear. While charge transport is enhanced by crystalline domains, exciton transport may be inhibited in well ordered systems.[42] This additional level of complexity is due to the dominance of exciton diffusion occurring by an inter-chain mechanism. In fact, exciton delocalization in regioregular P3HT is estimated to be as low as 1-2 nm, and therefore may not play a major role in charge separation.[43] Studying the interplay of delocalization with electrostatics may help elucidate these effects.

The combination of electrostatics and delocalization can play a major role in the charge separation process. For example, along with hot charge transfer (CT) states, calculations by Tamura and Burghardt show that charge delocalization can lead to ultra fast (< 100 fs) charge separation via the lowering of Coulomb barriers.[44] While delocalization can be treated very accurately in quantum calculations,[43,45,46] the effect is generally ignored in meso-scale simulations of charge transport. Here, highly accurate electrostatics are achieved by largely ignoring the effects of delocalization and treating carriers as point charges.

1.1.7 Other Charge Transport Models

Other models have been used to describe charge dynamics in organic semiconducting systems, usually based upon the gaussian disorder model (GDM) of Bässler, which assumes charge hopping through transport states that have a Gaussian distribution for energy.[47] When incorporating the spatial correlation of sites,[15] or the presence of traps,[48–50] these models give good agreement with experiment. Results of similar accuracy to the GDM can also be obtained from continuum models, such as drift-diffusion equations.[51] Our model differs from the GDM in that site energies (and thus the density of states) are not assumed Gaussian, and instead are explicitly assigned, as will be explained later. The model aims to compare theory and experiment while deriving needed values from first principle calculations.[45]

The use of explicit electrostatics is critical in the model. Explicit Coulomb interactions have been avoided by others due to their long range nature, making them computationally

demanding in systems of realistic sizes, or requiring special treatment when using periodic boundary conditions.[52] Some authors have included explicit Coulomb calculations into their models, but the simulation methods are unlike those used in this work. For example, when solving diffusion and continuity equations, Poisson's equation has been used to handle carrier interactions.[53] Sometimes, an approximate approach, such as a mean field or homogeneous medium, is taken.[52] Event-based Monte Carlo simulations by Watkins et al. have included Coulomb calculations, but used a small cutoff (9 nm), far below the expected interaction distance (∼15-30 nm) in low dielectric materials, even when using periodic boundary conditions.[54] More recent event-based Monte Carlo simulations, however, by Lyons et al. have included explicit Coulomb interactions.[55] Additionally, correlations between carriers, representing interactions, have been added to Gaussian Disorder Models,[56] but obscure the detail of underlying physical mechanisms.

1.2 PROJECT DESCRIPTION

In this section we give a brief summary of all the work performed and discussed in this thesis. The work has focused on development of a Monte Carlo simulation framework for charge transport, and application of that framework to OFETs and OPVs. Though simulation results were compared to and complimented by various experimental projects, *all experimental work was not performed by the author of this thesis.* Finally, we note that the chapters adaptations of previously published works.[57–62]

1.2.1 OFET Simulation

OFETs simulation involved studying charge injection and the presence of traps, barriers, and defects in OFETs. For charge injection, two main methods were probed, named injection and seeding. Seeding involves randomly placing charges during initiation, while injection involves injecting carriers from the source electrode. It was found that the method of charge injection does not affect the equilibrium current. However, both methods produce an oscil-

lating current. The oscillating current, as well as carrier lifetime and path length, occurs at frequencies that vary non-linearly with device length. These simulations were performed at various device geometries. Additionally, the incorporation of multiple layers and a gate potential into the simulation model for OFETs.

The effect of trap concentration on current in OFETs was examined. An asymmetric curve was found, which has been verified by experimental studies. The falloff and recovery of current with trap percentage becomes more extreme with increasing trap potential. Current voltage curves for OFETs with traps, barriers, and defects were calculated. It was found that defects and barriers of a high enough energy will lead to negative differential resistance.

1.2.2 OPV Simulation

To simulate OPVs, the simulation framework had to be reworked and extended, as discussed below. OPVs simulations were first performed on monolayer systems. A series of morphologies were studied. Isotropic two phase morphologies of varying domain size were simulated with and without the presence of scattered acceptor sites (pepper) in the donor phase. Additionally, ideal band morphologies of increasing domain size, and single phase systems were studied. It was found that device performance increases with increasing domain size in isotropic systems, but decreases in band systems. The presence of pepper in the donor phase was found to decrease the efficiency of isotropic systems. Similar to the monolayer systems, isotropic two-phase morphologies were studied with increasing domain size. Ideal systems based on the minimally curved surfaces called gyroid, p-surface, and d-surface, as well as 3D band morphologies were also simulated. Lastly, the effects of charge delocalization of device performance was checked. For all morphologies, increased charge delocalization leads to efficient separation of charge carriers.

1.2.3 Coding

A simulation framework was developed to perform all simulations. Though an extension of previous work from another student and post doc, the code has been altered in significant ways to accomplish the work in this thesis. Major changes included extension of the

systems from monolayers (2D) to multilayers (3D), the inclusion of a gate potential for transistors, the output of carrier lifetime and path length, the simulation of holes and electrons simultaneously, the implementation of checkpoint files, the extension of the source and drain systems to include multiple faces of the device, the incorporation of recombination and exciton injection, and the incorporation of charge delocalization into the Coulomb calculations. Additionally, major refactoring of the code to be legible and commented. To aid in the interpretation of simulations, a real time visualization of carrier movements was implemented using OpenGL. Finally, CPU algorithms for Coulomb interactions were made more efficient, and a GPU implementation was made.

1.3 OVERVIEW

First, the detailed description of the Monte Carlo model will be given, with a description of approximations and algorithms. The novel approach to Coulomb interaction calculations via GPUs will be described in this section. Next, the comparison of injection methods for OFETs is presented. In the following chapter, there is a discussion of the effects of traps, barriers, and defects in OFETs. In the final chapters, the extension of the Monte Carlo model to OPVs is presented. The effects of domain size, morphology idealization, and charge delocalization are discussed in the context of monolayer and multilayer OPVs. The appendix includes data, such as IV curves, to supplement the main chapters. Also in the appendix, previous work from unrelated projects in a previous group is included for completeness. This work was used for the comprehensive examination portion of the graduate program before the group dissolved. The project was implementation of novel Monte Carlo move for atomistic simulations of polymers in explicit solvent. The move, called solvent shift monte carlo (SSMC) aims to increase the sampling rate of inner torsional angles of the polymer. Additionally, molecular dynamics (MD) simulations of the polymer poly(n-isopropylacrylamide) (PNIPAAM) were performed. In these simulations, a temperature dependent collapse of the polymer was observed.

2.0 THEORETICAL MODEL OF CHARGE TRANSPORT

2.1 INTRODUCTION

A framework was developed to study charge transport in OFETs and OPVs using Monte Carlo simulations. Guided by first principles and experiment, the model is based on charge hopping.[16,45,60] The model attempts to connect molecular details, such as energy levels, to macroscopic device behaviors, such as current-voltage curves. Due to the complexity of the problem, coarse graining is needed.

In this chapter, a detailed description of the simulation model for transistors and solar cells is presented. The model is a Monte Carlo model. This means it consists of a parameterized energy function, a set of rules for proposing moves, and a way to calculate the probability of performing moves. Basics moves include hopping between sites, injection from electrodes, injection via illumination, recombination between holes/electrons, and extraction of charges at the electrodes. The energy function consists of the interaction with the electrodes (potential), the interaction between charges (Coulomb potential), and the assignment of orbital energies (traps, barriers, and defects). All of these aspects will be discussed. In addition, special focus is given to the calculation of the Coulomb potential between charges on graphics processing units (GPUs).

2.2 SITES

The model consists of a rectangular lattice of sites in between two electrodes. The rectangular lattice of sites can be a single layer or multiple layers. Each lattice site represents a

single, small, organic semiconducting molecule on the order of $1\,\mathrm{nm}^3$. The size of the site corresponds to the size of a phenyl-C_{61}-butyric acid methyl ester (PCBM) or phthalocyanine (PC) molecule (see Figure 1.1). OFET simulations use a single set of sites (grid), while OPV simulations use two sets of sites (grids).

An energy value is associated with each site. In OPVs, one grid represents the HOMO, while the other grid represents the LUMO. In OFETs, the grid represents either the HOMO or the LUMO energy, depending upon the majority carrier. The HOMO sites are used for hole dynamics, while the LUMO sites are used for electron dynamics.

When an excess electron occupies a site, this electron is sitting in the LUMO, and the site is an anion. Likewise, when an excess hole occupies a site, there is an electron missing from the HOMO, and the site is a cation. Whenever a hole and electron occupy the same site simultaneously, the site is considered to be an exciton. For an exciton, the molecule is in an excited state. There is both an electron missing from the HOMO and an excess electron sitting in the LUMO. This may arise, for example, when the molecule absorbs light, promoting an electron from the HOMO to the LUMO. Additionally, an exciton may form when an electron and hole hop to the same site. The model does not allow for more than one electron or hole to occupy a site at one time. Specifically, only the following situations are possible: empty state, anionic state, cationic state, and first excitonic state. Currently, there is no way for us to consider the spin of carriers, so there is no way to distinguish between triplet or singlet excitons.

Extensive use is made of a unique mapping between sites and their x, y, z positions. The position of an agent in the grid can be thought of as a 3-tuple of integers (x_i, y_i, z_i). This 3-tuple can be hashed into a single number called the site-id, s_i. This mapping reduces the amount of information that needs to be passed to the GPU during Coulomb calculation. Also, the mapping allows for easy look-up of pre-computed values of distance on the central processing unit (CPU). The following equations hold for site-ids, where all quantities are

integers, and integer division applies. The dimensions of the grid are L_x, L_y, and L_z.

$$0 \leq x_i < L_x \qquad x_i = s_i \% L_x \tag{2.1}$$

$$0 \leq y_i < L_y \qquad y_i = s_i / L_x - (s_i/(L_x L_y)) L_y \tag{2.2}$$

$$0 \leq z_i < L_z \qquad z_i = s_i/(L_x L_y) \tag{2.3}$$

$$0 \leq s_i < L_x L_y L_z \qquad s_i = L_x(y_i + z_i L_y) + x_i \tag{2.4}$$

2.3 SIMULATION

A typical simulation can be described as follows. Carriers arrive on the lattice by being injected. For OFETs, carriers can be randomly seeded at initiation. Random seeding represents the gate electrode injecting charges at the semiconductor-dielectric interface. Carriers can also be injected from the source electrode at a constant probability, or a probability that depends upon energetics such as the image potential, the Coulomb energy, and the injection barrier. The choice of a constant probability for the injection and removal of charge carriers represents an assumption that the energetic barriers at the electrode-semiconductor interfaces are constant and small. For OPVs, carriers are injected as excitons (electron and hole on the same molecule) onto a randomly selected site. Carrier injection by the electrodes has not been simulated in OPVs.

This means that the model does not simulate the dark current. We have assumed an ideal situation, where hole conducting and electron conducting sites absorb light at the same rate - i.e. excitons have an equal probability to form anywhere in the entire system. In reality, this situation is probably not true. Excitons are more likely to form away from the electrodes, and one of the materials is more likely to form excitons than the other. Additionally, when considering three dimensional systems, absorption rate should decrease linearly deeper into the material.

Once on the lattice, the movement of carriers is assumed to occur via thermally activated hopping, based on a bimolecular charge transfer rate that can be described using Marcus-Hush theory.[1] Hopping occurs because electronic states are localized on molecular sites in organic semiconductors, a result of the weak electronic interactions and disorder present in the system.[63,64] The criterion is described in section 2.4. As carriers reach the drain electrodes, the number of carriers leaving the system is monitored. This information is used to calculate a current.

2.4 METROPOLIS CRITERION

The Metropolis criterion for the probability of accepting a Monte Carlo move, with a coupling constant, A, takes the following form:

$$P_{accept}^{1\rightarrow 2} = min\left[A, Ae^{-\frac{\Delta E_{12}}{kT}}\right] \tag{2.5}$$

In traditional Metropolis criterion, A takes the value of 1. We have used a value of 1/3. What this means is that, regardless of the energy change being positive or negative, moves are always rejected at least 2/3 of the time. When the energy change is positive, the moves are rejected with probability greater than 2/3 Essentially, the coupling constant is increasing the likelihood that charge carriers remain still – because the molecular sites are not well "coupled" to one another.

Detailed balance is still maintained when including the coupling constant. Consider, for example, the case when the energy change when going from state $1 \rightarrow 2$ is positive. In this case the acceptance probabilities are:

$$P_{accept}^{1\rightarrow 2} = Ae^{\frac{-\Delta E_{12}}{kT}} \quad P_{accept}^{2\rightarrow 1} = A \tag{2.6}$$

When simulating in the canonical ensemble, the probability to be in state 1 or state 2 is given by the Boltzmann distribution:

$$P_1 \propto e^{\frac{-E_1}{kT}} \quad P_2 \propto e^{\frac{-E_2}{kT}} \tag{2.7}$$

When the probability of proposing a move from state $1 \rightarrow 2$ is equal to the probability of proposing a move from state $2 \rightarrow 1$, the detailed balance condition is still satisfied:

$$P_{accept}^{1\rightarrow 2} P_1 = P_{accept}^{2\rightarrow 1} P_2 \rightarrow \left(A e^{\frac{-\Delta E_{12}}{kT}} \right) \left(e^{\frac{-E_1}{kT}} \right) = A \left(e^{\frac{-E_2}{kT}} \right) \tag{2.8}$$

The coupling constant is a parameter of the model whose value changes the magnitude of currents calculated. Currents in the range of nA are calculated when a value of 1/3 is used. The physical motivation for the coupling constant comes from the Miller-Abrahms theory of phonon-assisted electron hopping, or the Marcus-Hush charge transfer theory, which is more reliable in organic materials. For example, the Marcus-Hush charge transfer rate is given by:

$$k_{et} = \frac{2\pi \langle H_{ab} \rangle^2}{\hbar} \frac{1}{\sqrt{4\pi \lambda k_b T}} e^{\frac{-\left(\lambda + \Delta G^0\right)^2}{4\lambda k_b T}} \tag{2.9}$$

And so, the coupling constant encodes the electronic overlap between adjacent molecules H_{ab}, as well as the reorganization energy λ. These quantities are material specific and can be calculated from quantum electronic structure calculations. For this work, we assumed an ideal situation where the donor and acceptor phases have the same coupling constant, and that two-site hops have an exponentially smaller coupling constant, reflecting smaller orbital overlap and a higher reorganization barrier.

2.5 TIMESCALE

In the strictest sense, Monte Carlo methods should not provide information about dynamics. This is routed in the fact that the potential V, through the Boltzmann factor, $e^{-\beta \Delta V}$, gives information about the probability distribution of particles in space, not time. On the other hand, the kinetic energy, T, gives information about the distribution of particles in time (i.e. their velocity distribution). In the derivation of the Monte Carlo acceptance probability in the canonical ensemble, the configurational partition function is used, which has had the kinetic information integrated away.

However, with the following argument, we can recover dynamical information from a Monte Carlo simulation. If Monte Carlo moves are kept physical or "small", a series of Monte Carlo configurations can be interpreted as a trajectory. That is, a timescale ($\sim$1 ps) can be associate with a Monte Carlo move that is "small" enough to be a physical transition between states of the system. This has the practical implication of moves being restricted to single-site to two-site hops. At no point are moves which transport carriers over large distances allowed to occur.

This assumption is not hard to make. Consider a full MD trajectory. Such a trajectory, has a starting state and an ending state. Multiple trajectories could connect these two states, each with their own well defined timescale. An average over these trajectories can be used to compute an average timescale that can be associated with the transition between the two states. The MD trajectory is itself a valid Monte Carlo move. As the MD trajectory becomes shorter, and the states more physically connected or correlated, the spread in the timescale becomes smaller. In this way, one could view a MD trajectory as a series of very small Monte Carlo moves. We use this argument to interpret the configurations produced by the Monte Carlo as a trajectory.

2.6 ENERGY

The energy change, $\Delta E = q\Delta V$, used in the Metropolis criterion has multiple contributions, and is represented in terms of potentials, as shown in Equation 2.10.

$$V_{sim} = V_{field} + V_{Coulomb} + V_{trap} \tag{2.10}$$

V_{field} represents a electric field across the device and has intrinsic and applied components. We assume that the field varies linearly between the left and right electrodes.

$$V_{field}(x_j, y_j, z_j) = \frac{V_R - V_L}{L} x_j \tag{2.11}$$

V_L is the potential of the left electrode, V_R is the potential of the right electrode, L is the length between the electrodes, and x_j is the position of the charge carrier j. The field is assumed not to vary in the y direction (width of the device).

$$V_{Coulomb}(x_j, y_j, z_j) = \sum_{|r_{ij}|>0}^{N} \frac{q_i}{4\pi\epsilon_0\epsilon|r_{ij}|} \tag{2.12}$$

$V_{Coulomb}$ represents Coulomb forces arising from the interaction of all N carriers. The Coulomb interaction is subjected to a cutoff of 50 nm and damped by a dielectric constant ϵ of 3.5, which is typical of organic semiconductors. Special care is needed for the Coulomb interaction when a hole and electron occupy the same molecule. $V_{Coulomb}$ evaluates to zero in this case, implying that there is no energetic force to form an exciton in the model. Therefore we introduce $V_{exciton}$ to represent the exciton binding energy. $V_{exciton}$ simply replaces the case in the Coulomb sum where $|r_{ij}| = 0$ and $i \neq j$. A value of 0.5 eV is used for electrons, and a value of -0.5 eV is used for holes. This is just enough energy to overcome the Coulomb interaction between a hole and electron separated by 1 nm. The net effect of this process is an exciton binding energy of $0.5\,\text{eV} - 0.411\,\text{eV} = 0.089\,\text{eV}$. This means the dissociation probability is $P = 1.0\,\%$ in the absence of all other potentials. Interfaces, the electrode potential, and other carriers make the dissociation increase.

When moving to delocalized charges, the exciton binding energy does not remain constant with the delocalization parameter, σ (see chapter 6). To overcome this, the interaction of a hole and an electron on the same site was set to $V_{SI} + V_{Exciton}$, where V_{SI} is the self interaction. The self interaction is the interaction energy between two charges separated by 1 nm. This formalism forces the exciton binding energy to be $V_{exciton}$ regardless of the value of σ. $V_{exciton}$ was set 0.089 eV to keep comparison with point charges consistent. The probability to dissociate excitons into free carriers is a parameter that deserves further investigation.

V_{trap} represents the shift in HOMO or LUMO energy when considering sites as different molecules, and encodes the morphology into the energy landscape. V_{trap} depends on position in space and if one is considering the HOMO or LUMO. For example, when a site is in the hole conducting phase, V_{trap} is some $\epsilon < 0$ for the HOMO. This encourages holes to travel in this phase.

Special meaning is assigned to the electrode potentials in solar cell simulations. To calculate an IV curve for a solar cell, the voltage applied across the cell *in the simulation* was varied over some range, for example, -2.0 and $2.0\,\mathrm{V}$. It was then assumed that the voltage in the simulation (V_{sim}) was really made up of an intrinsic voltage (V_{int}) and an extrinsic voltage ($\mathrm{V}_{applied}$), as shown in equation 2.13.

$$V_{sim} = V_{int} + V_{applied} \tag{2.13}$$

V_{int} is the result of the difference between the work functions of the two electrodes. We consistently used a value of $\mathrm{V}_{int} = 1.5\,\mathrm{V}$ for each system. To get an IV curve, we plotted the calculated current vs. $V_{applied}$. The calculated IV curves are s-shaped, and exhibit saturation regions when charge transport is very efficient.

2.7 CONCENTRATION

While we impose a maximum carrier concentration as a parameter (mainly for algorithmic reasons, the maximum is almost never reached, unless a morphology has very poor conductivity pathways. Instead, the actual carrier concentration reached (at equilibrium) is dictated by various other factors, such as the injection rate of excitons into the system or the rate of removal of carriers (at the electrodes or through recombination). Additionally, the spatial distribution of carriers at equilibrium is not uniform, but rather anisotropic. On the other the hand, the max concentration in an OFET is dictated by the gate electrode. In OFET simulation, an equilibrium concentration is reached as the source injects charges at a constant rate.

2.8 SOURCES AND DRAINS

To calculate a current, a dynamical interpretation of the Monte Carlo moves was made. Strictly speaking, Monte Carlo samples the configuration space only (see section 2.5). In

other words, it doesn't make sense to associate a trajectory with the configuration space, especially when the sampled configurations are disconnected from each other physically. For example, you can't arrange the generated configurations in a movie where particles continuously move from state to state. However, we feel a dynamical interpretation is valid when the proposed moves are physical and can be associated with a timescale. By physical moves, we mean that the Monte Carlo moves are between adjacent sites, or next nearest neighbors. An unphysical move would be moving carriers over much larger distances. By associated timescale, we mean the calculable $\sim$1 ps electron/hole transfer rate when carriers hop between molecules on the order of $\sim 1^3$ nm^3 in size.

Calculating current in an OFET simulation is straight forward. There is only one drain, on the right side of the system. We define positive current to be electrons flowing to the right. Therefore the flux of carriers through the right face is used to calculate a current. The current is simply the number of charges that flow through the drain divided by the total time.

An equilibration time is accounted for. To account for equilibration, one must subtract out the statistics from the equilibration period. For example, consider a system where the drain has a success rate of 100 at 100 000 steps into the simulation and a success rate of 500 at 500 000 steps into the simulation, the end of the simulation. To account for 100 000 steps of equilibration, the final success rate will be $500 - 100 = 400$.

The situation is slightly more complicated when multiple carrier types and drains are present. To extract a current from the simulation, the following method was used. We kept track of the carrier type and cell face (electrode) a carrier exited. There are four possibilities: holes leave through the right electrode (N_{HR}), holes leave through the left electrode (N_{HL}), electrons leave through the right electrode (N_{ER}), electrons leave through the left electrode (N_{EL}). The convention remains that electrons leaving the right electrode as positive current. Therefore, the current (I) through the device is calculated as:

$$I = (N_{HL} + N_{ER} - N_{HR} - N_{EL})\frac{e}{\tau} \tag{2.14}$$

τ is the simulation time, e is the elementary charge, and the different N are carrier counts. Additionally, the current density (J), is calculated as I/A, where A is the cell surface area.

2.9 EXCITON INJECTION

Exciton injection, was implemented in a probabilistic fashion using an injection rate. The injection rate quantifies the probability that a molecule (simulation site) will absorb a photon. The energy from the photon excites an electron from the HOMO orbital into the LUMO orbital. A rough estimate for an injection rate was calculated by examining the AM1.5 solar spectrum. By integrating the irradiance ($\mathrm{W\,m^{-2}\,nm^{-1}}$) over wavelength to obtain the incident power ($\mathrm{W\,m^{-2}}$), using the average energy of an photon (around 700 nm, 1.77 eV), and the area of a solar cell device ($256^2\,\mathrm{nm^2}$), a rate of $10^{-3}\,\mathrm{ps^{-1}}$ was used. Not every single photon incident is absorbed, so this is definitely an over estimation. However, we did observe that current in a single-phase morphology scales linearly with injection rate. Since the injection rate depends upon the area of the solar cell, the rate was scaled linearly with surface area for systems larger or smaller than $256^2\,\mathrm{nm^2}$.

We have not tested how the characteristic parameters like fill factor change with exciton injection rate. However, we have tested individual current-voltage points, and found that the current calculated at a given voltage point increases linearly. Also, if the injection rate is abnormally high, a large amount of carriers in the system can lead to a situation where charge carriers form Coulombically-bound clusters. Basically the system can freeze. Although we believe that this situation is unrealistic, we note that the formation of these clusters can be mitigated by increasing the recombination rate. Finally, we note that in the two dimensional case, the photon flux was constant. In a three dimensional system, an enhanced photon flux could increase layers which receive low light.

2.10 EXCITON RECOMBINATION

We implemented recombination by first identifying all excitons in the simulation and removing some with a small probability called the recombination rate. The recombination rate quantifies the probability that an exciton (electron-hole pair on one site) will recombine. In this process, light or heat is emitted as an excess electron in the LUMO transitions into

the vacant orbital represented by the hole (lack of an electron) in the HOMO orbital. Note that we are not talking about the event where a hole and electron on separate molecules combine to form an exciton, but rather that a molecule with a hole and electron already present decays into the neutral molecule. The process of recombination was implemented by identifying excitons at the end of each Monte Carlo step, and removing excitons if some random number between 0 and 1 was less than the recombination rate. Simulations were ran with a recombination rate of $10^{-5}\,\mathrm{ps}^{-1}$. This is equivalent to saying that one would expect a recombination event at least once every $\sim 10^5\,\mathrm{ps}$, if there is at least one exciton present in the system.

2.11 OPENCL

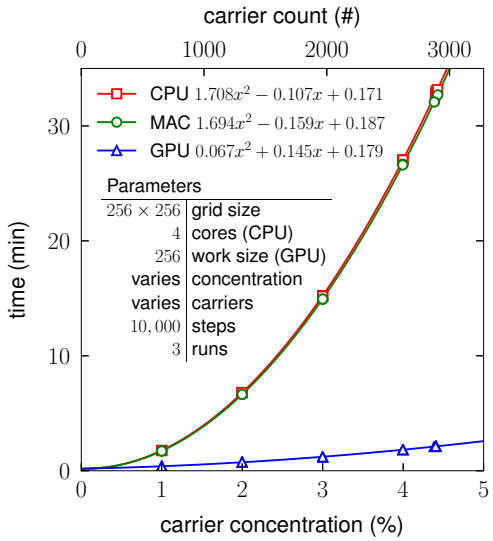
In this section, the use of the open compute language (OpenCL) for Coulomb calculations is discussed. OpenCL is a programming framework for heterogeneous computing developed by the Khronos Group. With OpenCL, one can harness the parallel processing power of both the GPU and the CPU in a single computer program. OpenCL is an open source, platform independent alternative to proprietary languages such as Nvidia's compute unified device architecture (CUDA). The speedup using a GPU with OpenCL for Coulomb calculations is shown in Figure 2.1.

The calculation of the Coulomb interactions between a set of point charges must be performed repeatedly and accurately during the simulation. The problem can be stated as follows. Given a set of charges, $\mathbf{Q} = \{q_i \mid 0 \leq i < N\}$ and their positions $\mathbf{R}_i = \{\mathbf{r}_i \mid 0 \leq i < N\}$ find the potentials $\mathbf{V}_j = \{v_j \mid 0 \leq j < M\}$ at the points $\mathbf{R}_j = \{\mathbf{r}_j \mid 0 \leq j < M\}$ using Equation 2.15, where ϵ_0 is the permittivity of free space and ϵ is the dielectric constant.

$$V_j = \sum_{i=0}^{N} \frac{q_i}{4\pi\epsilon\epsilon_0 |\mathbf{r}_j - \mathbf{r}_i|} \tag{2.15}$$

A given component of the sum in Equation 2.15 is called the kernel.

(A)



(B)

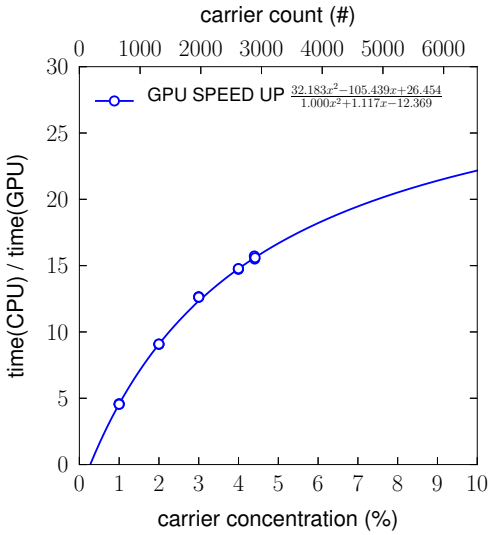
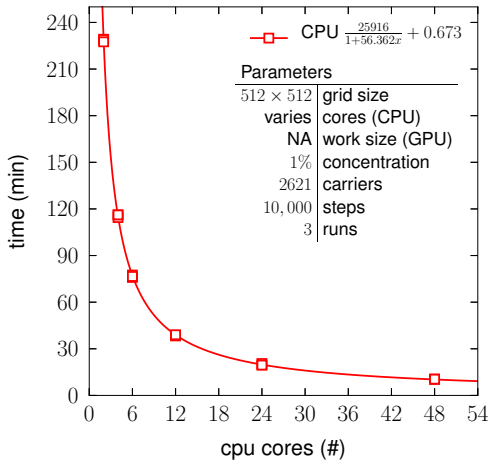


Figure 2.1: Simulation timing (GPU/CPU) vs carrier concentration.

(A)



(B)

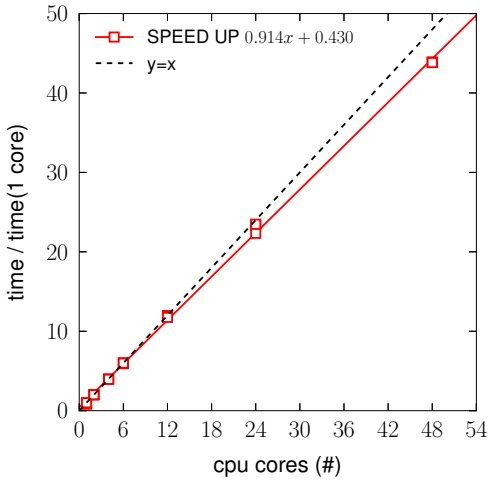


Figure 2.2: Simulation timing (CPU) vs cores.

Unfortunately, brute force calculation of Equation 2.15 would scale as $M \times N$, where N is the number of charges and M is the number of locations to calculate the potential at. This is due to a double loop over M and N, as shown below.

```
for j in range(M):                      # loop over M potential points
    vj = 0.0                            # zero potential at point j
    for i in range(N):                  # loop over N charges
        rij = (xj - xi) * (xj - xi) +   # calculate r_ij^2
              (yj - yi) * (yj - yi) +
              (zj - zi) * (zj - zi)
        if rij > 0:                     # make sure potential is finite
            rij = sqrt(rij)             # square root (slow)
            vj += qi / rij              # division (slow)
    vj = vj / (4 * pi * e * e0)         # fix units
```

Typically, M will be $2N$, encompassing the current and trial positions of charge carriers at a given simulation step. Therefore, the simulation will scale as kN^2, where k is some prefactor. If the number of charges doubles, the simulation will take four times as long. Additionally, the presence of a square root, division, and "if" statement inside the inner loop slow the calculation down greatly. Sometimes, the calculation of the potential is desired at every single lattice point in the grid. Such a calculation is extremely slow due to the large number of lattice points. For example, consider a system of 1000 charges. A normal Coulomb calculation, at the current and trial locations would involve $2 \times 1000 \times 1000 = 2 \times 10^6$ evaluations of the kernel in Equation 2.15. On the other hand, if the simulation occurs on a grid of $256 \times 256 \times 32 = 2\,097\,152$ sites, the number of kernel evaluations to calculate the potential everywhere is $2\,097\,152\,000$, nearly 1048 times as many calculations.

There are many tricks to speeding up Coulomb calculations. While some are algorithmic in nature, others make physical assumptions that may sacrifice accuracy. One of the most basic methods is the use of a neighbor list. A neighbor list is a record of particle pairs that fall within a certain cutoff distance of one another. In this case, only the pairs need be considered in the summation. The neighbor list assumes that the state of the system does not change significantly between simulations steps. This makes the reconstruction of the neighbor list, an N^2 process in its own right, less frequent than the Coulomb calculation. The neighbor list also assumes that the Coulomb interactions are negligible outside a certain cutoff distance. This assumption tends to be a poor one for potentials that are long ranged.

The Coulomb potential is of the form r^n, where $n = -1$. Potentials of this form are long ranged if the integral over a sphere, $4\pi \int_0^\infty r^2 r^n$, diverges. Only potentials with $n < -2$ will converge. Neighbor lists are bad for Monte Carlo simulations, where the state of the system can change rapidly with "non-physical" transitions.

In the lattice simulation, exact calculation of the Coulomb interactions are desired. To combat the intense cost, parallel computation can be used. First off, the calculation of the sum in Equation 2.15 can be performed for different j independently by different threads. Threads need only have read access to the charge locations and write access to a memory location of a specific j. This form of parallel computation is straightforward and easily achieved with the open multi-processing (OpenMP) library. The CPU implementation of the Coulomb loop uses this form. The speedup depends upon the number of cores the CPU can dedicate, anywhere from 2 to 48.

```
for j in range(M):          # loop over M potential points
    spawnThread()
...
def thead(j):
    for i in range(N):      # CPU threads perform inner loop
        ...
```

The speedup using the parallel CPU implementation is shown in Figure 2.2. Although the speedup is near linear in the number of cores, the CPU algorithm is quadratic in the number of charges. A much faster algorithm can be accomplished with GPUs.

The heart of the GPU Coulomb calculation is performed by the kernel. The kernel can be viewed as the instructions followed by a single worker (thread). To understand the kernel, a little background on the architecture of a GPU is needed. As shown in Figure 2.3, a GPU has various components. The GPU is divided into workers (threads). These workers have their own private memory, that is small and fast to access, but restricted to the worker. The workers are then grouped together into work groups (blocks). A work group has its own local memory. This memory is also very fast to access. However, although all workers in the work group share the memory, any workers outside of the work group cannot access the memory. There are many work groups, all with their own workers and local memory. Every work on the GPU has access to the global memory of the GPU. However, accessing this memory is costly. If all the workers try to access the same region of the global memory at the same

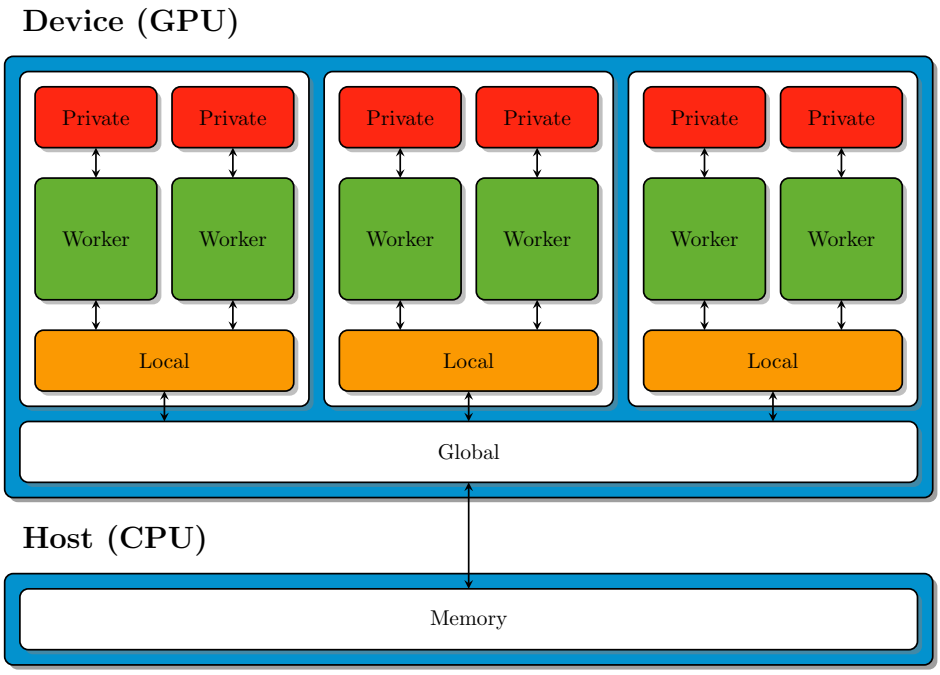


Figure 2.3: OpenCL host and device memory layout.

time, the algorithm will become very slow.

With the layout in mind, we can now describe the kernel. The kernel takes eight arguments. Three are arrays, prefixed by the compiler directive `__global`, which means the array is located in the global memory of the GPU. Although all workers can read and write to the global memory, repeated access is extremely inefficient. The signature of the kernel and description of the arguments is shown below.

```
__kernel void coulomb(
    __global double *o,  // array for output, size = number of grid sites
    __global int    *s,  // site-ids
    __global int    *q,  // charges
    int              n,  // number of electrons, holes, and defects
    int             c2,  // square of cutoff
    int          xsize,  // size of grid in x-direction (length)
    int          ysize,  // size of grid in y-direction (width)
    double prefactor     // electrostatic prefactor
) {
// note: site-ids (s) and charges (q) contain:
//    electron ids/charges (current) size=E
```

```
//    hole     ids/charges (current) size=H
//    defect   ids/charges           size=D
//    electron ids/charges (future)  size=E
//    hole     ids/charges (future)  size=H
```

The kernel works by assigning a charge (in the present or the future) to each work group. The work group id, `get_group_id(0)`, is used to index the site ids and charge arrays. This variable is the same for every worker in the work group. The dimensions of the grid are used to convert the site id into the position of the charge.

```
    // each worker of work group loads the same charge and site,
    // using the "work group id"
    int qi = q[get_group_id(0)];
    int si = w[get_group_id(0)];

    // extract position from site using the grid dimensions
    int zi = (si) / (xsize * ysize);
    int yi = (si) / (xsize) - (zi * ysize);
    int xi = (si) % (xsize);
```

The total number of work groups, W, needed is shown in Equation 2.16, where E is the number of electrons, H is the number of holes, and D is the number of defects. The array "s" of the kernel must at least be this size. The factors of 2 represent that calculations must be performed using the current and future sites of carriers. Although useless information, the kernel also calculates the potential felt at defect sites to simplify the algorithm. If the defects are uncharged, they are excluded from the calculation.

$$W = 2E + 2H + D \tag{2.16}$$

The real work begins as workers in the work group load the site ids and charges of interacting charges into local memory. The workers must be made aware of the local memory, which must be large enough to store all information (charges, sites, potentials) for a chunk. Each worker in a work group keeps a running sum of the calculations it has performed. This is what the local memory `vlocal` is for. The work must set the initial value to 0 using its unique local id, `get_local_id(0)`. After this, the workers wait for the other workers to catch up.

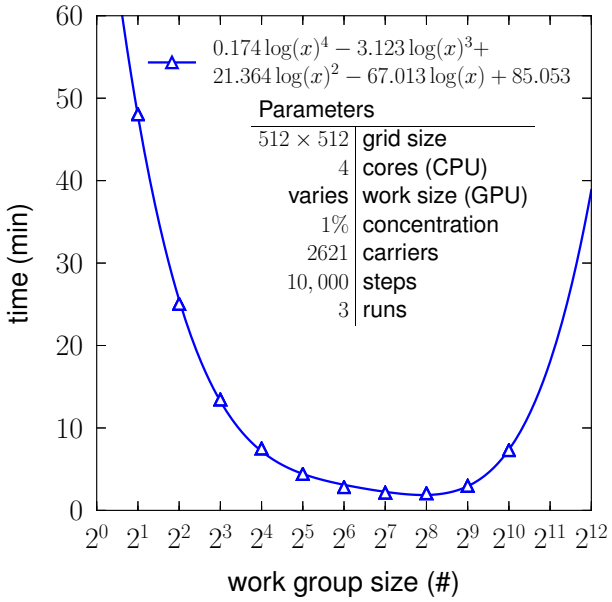


Figure 2.4: Simulation timing (GPU) vs work size

```
// allocate local memory for this work group
__local int    slocal[1024]; // sites      : large enough...
__local int    qlocal[1024]; // charge     : optimal size = 256
__local double vlocal[1024]; // potentials : 256 < 1024

// each worker sets local potential to zero
vlocal[get_local_id(0)] = 0;

// wait for the other workers to get to this point
barrier(CLK_LOCAL_MEM_FENCE);
```

Each worker in the work group is responsible for loading a different portion from the global memory. Due to the limited size of the work group, this process is done in chunks until all interactions have been accounted for. The optimum number of workers in the work group is 256, as shown in Figure 2.4.

The number of chunks needed is the total number of charges, n divided by the size of the work group, `get_local_size(0)`. It is unlikely that these numbers will divide evenly, so an extra chuck is added.

```
// calculate how many pieces we can divide the total list of
// charges/positions into
int num_loads = n / (get_local_size(0)) + 1;
```

The main loop is over the chunks. For each chunk, a worker calculates its unique load id. This load id indexes into the global site and charge arrays. The worker then uses its local work group id, `get_local_id(0)`, to save the information from global memory to the local memory. The workers then wait for the other workers to finish loading the chunk information into the local memory. A special site id of -1 is used to indicate a worker that should not calculate an interaction during this chunk. This happens when the load id is larger than the number of charges. It can happen when the work group size and number of charges are not multiples of one another.

```
// start loading chunks of charges to calculate on
for (int load_number = 0; load_number < num_loads; load_number++)
{
    // each worker loads a different charge
    int load_id = get_local_size(0) * load_number + get_local_id(0);
    if (load_id < n) //number of charges
    {
        slocal[get_local_id(0)] = s[load_id];
        qlocal[get_local_id(0)] = q[load_id];
    }
    else
    {
        slocal[get_local_id(0)] = -1;
        qlocal[get_local_id(0)] = -1;
    }
    barrier(CLK_LOCAL_MEM_FENCE);
```

Once all information is loaded into local memory, each worker calculates a contribution to the Coulomb interaction. The loaded site id is mapped to the x, y, and z positions of the interacting charge. If the distance is greater than zero and less than the cutoff, the contribution is added to the local potential array. Workers must wait for other workers in the work group to finish. This is because the next chunk can not be loaded while other workers are calculating Coulomb interactions.

```
    // each worker performs a different q/r calculation
    int sj = slocal[get_local_id(0)];
    if (sj >= 0)
    {
```

```
        int   zj = (sj) / (xsize * ysize);
        int   yj = (sj) / (xsize) - (zj * ysize);
        int   xj = (sj) % (xsize);
        double r = (xi - xj) * (xi - xj) +
                   (yi - yj) * (yi - yj) +
                   (zi - zj) * (zi - zj);

        // compute the interaction
        if (r > 0 && r < c2)
        {
            vlocal[get_local_id(0)] += qlocal[get_local_id(0)] * rsqrt(r);
        }

        }
        barrier(CLK_LOCAL_MEM_FENCE);
    }
```

Once the loop over chunks has been finished, the potentials in the local potential array, `vlocal` must be summed. There is no easy way to do this sum in parallel. Therefore, a single worker (worker 0 of the work group) is responsible for summing the potentials. Finally, the result is saved to the global output array.

```
    barrier(CLK_LOCAL_MEM_FENCE);
    if (get_local_id(0) == 0)
    {
        double v = 0;
        for (int l = 0; l <  get_local_size(0); l++)
        {
            v = v + vlocal[l];
        }
        //o[si] = v;
        o[get_group_id(0)] = prefactor * v;
    }
}
```

The full kernel is given below, without interruption.

```
__kernel void coulomb(
    __global double *o,  // array for output, size = number of grid sites
    __global int    *s,  // site-ids
    __global int    *q,  // charges
    int              n,  // number of electrons, holes, and defects
    int             c2,  // square of cutoff
    int          xsize,  // size of grid in x-direction (length)
    int          ysize,  // size of grid in y-direction (width)
    double prefactor     // electrostatic prefactor
) {
```

```
    int qi = q[get_group_id(0)];
    int si = w[get_group_id(0)];

    int zi = (si) / (xsize * ysize);
    int yi = (si) / (xsize) - (zi * ysize);
    int xi = (si) % (xsize);

    __local int    slocal[1024];
    __local int    qlocal[1024];
    __local double vlocal[1024];

    vlocal[get_local_id(0)] = 0;
    barrier(CLK_LOCAL_MEM_FENCE);

    int num_loads = n / (get_local_size(0)) + 1;

    for (int load_number = 0; load_number < num_loads; load_number++)
    {
        int load_id = get_local_size(0) * load_number + get_local_id(0);

        if (load_id < n)
        {
            slocal[get_local_id(0)] = s[load_id];
            qlocal[get_local_id(0)] = q[load_id];
        }
        else
        {
            slocal[get_local_id(0)] = -1;
            qlocal[get_local_id(0)] = -1;
        }
        barrier(CLK_LOCAL_MEM_FENCE);

        int sj = slocal[get_local_id(0)];
        if (sj >= 0)
        {
        int   zj = (sj) / (xsize * ysize);
        int   yj = (sj) / (xsize) - (zj * ysize);
        int   xj = (sj) % (xsize);
        double r = (xi - xj) * (xi - xj) +
                   (yi - yj) * (yi - yj) +
                   (zi - zj) * (zi - zj);

        if (r > 0 && r < c2)
        {
            vlocal[get_local_id(0)] += qlocal[get_local_id(0)] * rsqrt(r);
        }

        }
```

```
        barrier(CLK_LOCAL_MEM_FENCE);
    }

    barrier(CLK_LOCAL_MEM_FENCE);
    if (get_local_id(0) == 0)
    {
        double v = 0;
        for (int l = 0; l <  get_local_size(0); l++)
        {
            v = v + vlocal[l];
        }
        //o[si] = v;
        o[get_group_id(0)] = prefactor * v;
    }
}
```

The Coulomb kernel described above is for calculating the potential *at the* charge positions, in the future and past, due to other charges. This is all that is needed for the Monte Carlo simulation. However, a related problem, useful for analysis and comparison to experiments, such as kelvin probe atomic force microscopy, is calculating the Coulomb potential at every point in the grid. The first problem is slow on the CPU. The second is second problem is painfully slow, if not impractical. However, a GPU implementation is very practical and fast.

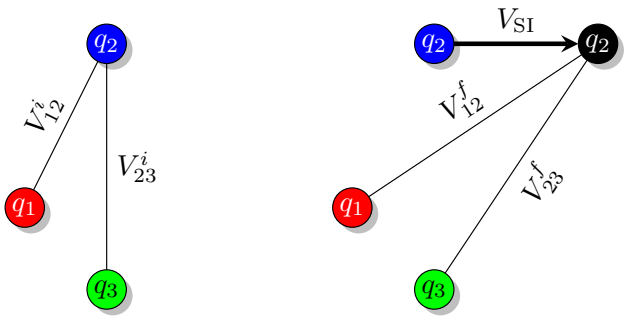
Below is the GPU implementation for calculating the Coulomb potential everywhere in the grid. The algorithm is very similar. However, it uses the multi-dimensional nature of work groups to simplify the indexing. The underlying algorithm is very similar to the previous one. With a 3D kernel, the simulation grid is mapped directly onto the GPU architecture. This time, each work group represents a single site in the simulation. There are as many work groups as there are sites. Contrast this to the above kernel, where each work group represented a location to calculate an interaction at. This puts a limit on the size of the grid that the GPU can handle. There is a maximum number of work items that a kernel can handle, specific to the GPU.

```
__kernel void coulomb(
    __global double *o,  // array for output, size = number of grid sites
    __global int    *s,  // site-ids
    __global int    *q,  // charges
    int              n,  // number of electrons, holes, and defects
    int             c2,  // square of cutoff
    double    prefactor  // electrostatic prefactor
```

```
) {
    // map 3D local work item indices to 1D index j
    int j = get_local_id(0) +
            get_local_id(1) * get_local_size(0) +
            get_local_id(2) * get_local_size(0) * get_local_size(1);

    // map 3D work group indices to 1D index k
    int k = get_group_id(0) +
            get_group_id(1) * get_num_groups(0) +
            get_group_id(2) * get_num_groups(0) * get_num_groups(1);

    // let 'this work item' know about the local memory for the
    // work group it belongs to
    // ... to be accessed using the index 'j'
    __local int    slocal[64];
    __local int    qlocal[64];
    __local double vlocal[64];

    // have 'this work item' set its own initial potential to zero
    vlocal[j] = 0;
    barrier(CLK_LOCAL_MEM_FENCE);

    // loop over global memory in chunks of size local_volume = (Sx * Sy * Sz)
    // There are n charges, so you need to do at least n / local_volume loads,
    // maybe 1 more if the numbers don't divide nice.
    int local_volume = get_local_size(0) * get_local_size(1) * get_local_size(2);
    int num_loads = n / (local_volume) + 1;

    for (int load_number = 0; load_number < num_loads; load_number++)
    {
        // 'this work item' is responsible for loading information for a
        //  single charge from the arrays s and q
        int load_id = local_volume * load_number + j;
        if (load_id < n) //number of charges
        {
            slocal[j] = s[load_id];
            qlocal[j] = q[load_id];
        }
        else
        {
            // don't load from s and q if our load_id was greater than the
            // number of charges in s and q - it can happened when
            // the numbers don't divide nicely.
            slocal[j] = -1;
            qlocal[j] = -1;
        }
        barrier(CLK_LOCAL_MEM_FENCE);
```

```
        // 'this work item now calculates a q / r for the local id it
        // loaded ... only if it loaded one (s >= 0)
        int s = slocal[j];
        if (s >= 0)
        {
            // calculate the x,y,z for the charge 'this work item' loaded
            // from the 1D index s
            int   z = (s) / (get_num_groups(0) * get_num_groups(1));
            int   y = (s) / (get_num_groups(0)) - (z * get_num_groups(1));
            int   x = (s) % (get_num_groups(0));
            // calculate the distance between x,y,z and 'this work group' -
            // remember each point in the 3D space we are calculating
            // the coulomb potential in got assigned to a work group;
            // The assignment was done in such a way so that the work group
            // ids corresponded to the position of the point the work group
            // is assigned to.
            double r = (get_group_id(0) - x) * (get_group_id(0) - x) +
                       (get_group_id(1) - y) * (get_group_id(1) - y) +
                       (get_group_id(2) - z) * (get_group_id(2) - z);
            // Check for cutoff and make sure r != 0
            // which happens when a charge is present at the work groups
            // position
            if (r > 0 && r < c2)
            {
                vlocal[j] = vlocal[j] + qlocal[j] * rsqrt(r);
            }
        }
        barrier(CLK_LOCAL_MEM_FENCE);
    }
    barrier(CLK_LOCAL_MEM_FENCE);

    // If 'this work item' is work item 0 in the work group, tally up the
    // results from the other work groups and write them to the
    // output vector o.
    if (j == 0)
    {
        double v = 0;
        for (int l = 0; l < local_volume; l++)
        {
            v = v + vlocal[l];
        }
        o[k] = prefactor * v;
    }
}
```

Both coulomb kernels were easily modified to include delocalization of charges. The error function is used to represent delocalized charges.



$$\Delta V = \left(V^f_{12} + V^f_{23} + V_{\rm SI}\right) - \left(V^i_{12} + V^i_{23}\right) - V_{\rm SI}$$

Figure 2.5: Self interaction.

```
vlocal[get_local_id(0)] += qlocal[get_local_id(0)] * erf(erffactor * r)/r;
```

When calculating the Coulomb interactions over future and present positions of charges, the future position of a charge can interact with itself in the past. We call this interaction the self interaction. On this CPU, one can easily check to see if the interacting charge is the same as the acting charge. However, this information is not present on the GPU. Therefore, after the GPU calculation results are sent back to the CPU, the self interaction is subtracted off of the answer. The same process was adopted for pure CPU calculations, to save the evaluation of an "if" statement in the inner Coulomb loop. A simple example of the self interaction is shown in Figure 2.5.

3.0 CHARGE INJECTION DYNAMICS : CHARACTERISTICS OF DEVICE TURN-ON

The text in this chapter has been adapted from Gagorik and Hutchison, *Simulating Charge Injection and Dynamics in Microscale Organic Field-Effect Transistors.*[59]

3.1 INTRODUCTION

Monte Carlo simulations were used to investigate the carrier dynamics in realistic, finite-sized, small-molecule, organic field-effect transistors (OFETs) within the first few nanoseconds of device turn-on, as well as the system equilibrates. The results show that the device current exhibits large magnitude oscillations (64 ± 27 nA) during device turn-on if the initial configuration assumed no carriers in the device (i.e., carriers only arrive through injection from the source electrode). After equilibration (125 ns), the current continues to oscillate, however, at lower magnitude (64 ± 2 nA), even if the initial configuration assumed randomly placed charges. Fourier Transforms of device current as a function of simulation time show that these oscillations occur at well-defined device-geometry dependent frequencies, independent of initial configuration of the system. Examination of the carrier lifetimes and path lengths, which were found to vary non-linearly with device length, are used to argue that the oscillations are the result of the charge injection procedure, which assumed a constant probability event. The results suggest that carriers travel in waves in realistically finite-sized devices and that carrier lifetime and path length vary non-linearly by device geometry. Alternating current studies of OFETs may be useful in confirming these findings.

In this section, the Monte Carlo simulation framework to examine the dynamics of charge

carriers during device turn-on in a single-molecule (Cu-phthalocyanine), realistically sized (128 by 128 to 1024 by 1024 nm^2) organic field-effect (OFET) transistor. We wish to discuss the effects of charge injection and transport, and predict behavior not discussed elsewhere in the simulation or experimental literature. In particular, we examine the case of constant probability injection (90 %) limited by carrier concentration as a function of simulation time, device geometry, and energetics. A constant probability of injection ignores the details of the organic semiconductor metal interface, and this view is tested against more detailed models.

We find that the rate at which carriers traverse the organic semiconductor oscillates with time. This oscillation leads to a turn-on source-drain voltage at low simulation times. The frequencies present in the oscillation are shown to decrease non-linearly with source-drain length, an effect ultimately linked to a nonlinear increase in carrier lifetime and path length. We find that the ultimate source of the oscillations is the restriction of free carrier concentration in the device. In real devices, free carrier concentration is fixed by using a gate electrode. The result is discussion relevant for experimentalists interested in time-dependent behavior of OFETs. For example, while the quality of OFET devices has been judged using carrier mobility, evaluated mostly from drain current in the steady-state, the understanding of transient behavior of OFET is needed for potential high-frequency applications.[65]

3.2 COMPUTATIONAL METHODS

First, the model used to simulate charge transport is described. Second, all simulation parameters are outlined for the systems used to investigate charge injection in the model.

3.2.1 Description of Model

Our group has developed a coarse-grained Monte Carlo simulation model to study charge dynamics in organic semiconductors.[16,60] The model aims to compare theory and experiment while deriving needed values from first principle calculations.[45] Unlike our previous work, the present study focuses on injection and dynamics in a pure OFET material with no defects

or traps.

Geometrically, the model consists of a source electrode, a rectangular lattice of sites, a drain electrode, and charge carriers. The arrangement is representative of a thin-film transistor. The rectangular lattice of sites can be a single layer or multiple layers. The individual sites correspond to single organic semiconductor molecules, such as phthalocyanine. The device geometry is illustrated in Figure 3.1A. The model assumes a single carrier device, that is a p-type or n-type semiconductor. In this work we study hole transport, however, the results are symmetric with respect to the charge used for carriers. Disorder in the system is explicitly included in the energy landscape (discussed below). A single layer is normally sufficient, because injected charge tend to be confined to a region very near the dielectric interface.[66]

A gate electrode is not explicitly included in this work, however, a fixed concentration of carriers is maintained in the device. The desired target concentration is an average value. The carrier concentration can be related to the voltage applied between the source electrode and gate (v_{sg}) in real FET devices.[67] Higher v_{sg} allows more charge carriers to be introduced into the semiconductor film, which results in a higher source-drain current i_{ds}. For example, a target concentration of 1 % corresponds to 2621 carriers in a $1024 \times 256\,\mathrm{nm}^2$ Si/SiO_2 device, with a v_{sg} of 15.7 V and 300 nm thick oxide layer.[16]

Although the target concentration is an average value, the resulting distribution of charges tends to be anisotropic – even without the explicit inclusion of trap sites. See Figure 3.9A for a picture of the electrostatic potential, which is the root cause of the anisotropy, to see how carriers distribute themselves.

The inclusion of a gate can be accomplished in the model by simulating the device with multiple layers and a potential gradient along the z-direction. The effect of the potential gradient forces carriers to travel in layers close to the gate electrode, essentially reproducing most results of monolayer simulations. We are focused on simulating monolayer systems, and we note that real devices can be monolayers.

A typical simulation can be described as follows. Carriers arrive on the lattice by being injected from the source electrode. Once on the lattice, the movement of carriers is generally assumed to occur via thermally activated hopping, based on a bimolecular charge

transfer rate that can be described using Marcus-Hush theory.[1] Hopping occurs because electronic states are localized on molecular sites in organic semiconductors, a result of the weak electronic interactions and disorder present in the system.[63,64]

The hopping is accomplished by proposing Monte Carlo moves for carriers. The difference in energy a carrier undergoes when moving from its current to proposed location is used in a Metropolis criterion to simulate transport. While moves to lower energy are accepted one-third of the time, moves to higher energy states are accepted one-third of a Boltzmann factor involving the energy change. The Monte Carlo moves are restricted to be single site hops, and a dynamic interpretation of the states produced is accomplished by associating a timescale with the moves, regardless of if they are accepted or rejected. For example, carriers will "spend more time" on states of low energy, like trap states, by nature of a small probability to move to higher energy states. The timescale chosen for the moves is $\sim$1 ps, and is consistent with quantum chemistry calculations and fast, efficient charge transfer, however, the charge transfer rate only affects the magnitude of electrical currents calculated by our code and not the trends observed.

The energy of carrier j, of charge q, at position (x, y), on a two dimensional lattice, shown in equation 3.1, and diagrammed in Figure 3.1B, has multiple contributions.

$$E(x, y, q) = q\left[\left(\frac{V_D - V_S}{L_x}\right)x + \sum_{i \neq j}^{N} \frac{q_i}{4\pi\epsilon_0\epsilon\sqrt{(x - x_i)^2 + (y - y_i)^2}} + V_{disorder}(x, y)\right] \quad (3.1)$$

First, the potential difference between the source, V_S, and the drain, V_D, electrodes creates an electric field that is assumed to be constant along the length of the device, L_x. This field is assumed zero along the y direction (the width of the device). Secondly, Coulomb contributions arising from the interaction of all N carriers are explicitly added. The Coulomb interaction is subjected to a cutoff of 50 nm and damped by a dielectric constant ϵ of 3.5,[68] which is typical of the phthalocyanine semiconductors studied experimentally by our group. Changing the dielectric constant in our simulation was found to have little effect. Traps and barriers are modeled as perturbations in the energy landscape, $V_{disorder}(x, y)$, and have been considered elsewhere.[16,60] While most models assume that trap/barrier energies yield a Gaussian distribution,[48–51,69] we have frozen this assumption out. Instead, for a given site, $V_{disorder}(x, y)$ is chosen explicitly to be 0.0 eV or some constant, typically on the order of $\pm$0.1

to $\pm 3.0\,\mathrm{eV}$. Note that a positive value indicates a trap for electrons, while being a barrier for holes, since it is framed in terms of a potential. Sites that can not be traveled to, termed defects, can also be included, and one may view them as sites with $E(x, y, q)$ some energy large enough that the Metropolis always rejects transfer to such a site. Defects, traps, and barriers were not considered in this work, however, some disorder in the system can still be argued to be present, for the Coulomb interactions add both spatial and energetic disorder to the energy levels. See Figure 3.9A for a contour plot of the Coulomb energy in a device.

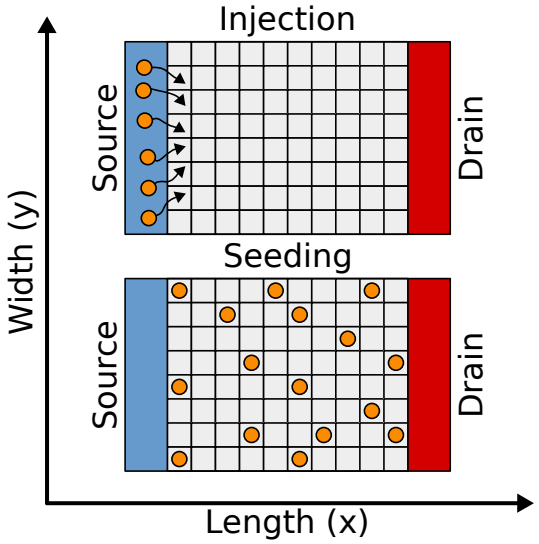
Device current is calculated as carriers arrive at the drain electrode. We count the number of carriers collected at the drain per iteration and then use the transfer rate to calculate a current. An example current-voltage curve can be seen in Figure 3.2. The current calculated is the conduction current. There may also be a displacement current. When no charges are allowed to enter from the source or drain, if a time varying voltage is present, then a displacement current should be measurable. There is no time varying component from the applied external potential. However, just before charges reach the drain, there is a time varying local internal potential from the moving point charges. As charges approach the drain, the change in induced charge on the drain generates the displacement current. While we are not considering the induced charge on the drain electrode, which would simply pull the carriers slightly towards the drain, the Coulomb interactions between the carriers are capturing the local internal potential.

3.2.2 Systems Studied

Carrier dynamics were examined for device geometries consisting of a single monolayer of well-defined small molecule sites. A single layer is one site or one molecule thick in the model. The various device lengths and widths chosen all fell within the range of 128 to 1024 sites. Each site represents a single organic semiconducting molecule, and corresponds roughly to $1\,\mathrm{nm}^2$, so devices range up to $\sim 1\,\mu\mathrm{m}^2$ in size.

For a given system, either the current-voltage curve was studied or a single current-voltage point. When studying a single point, carrier lifetime and carrier path length were recorded for every carrier. The carrier lifetime is the total number of simulation steps,

(A) path length



(B) lifetime

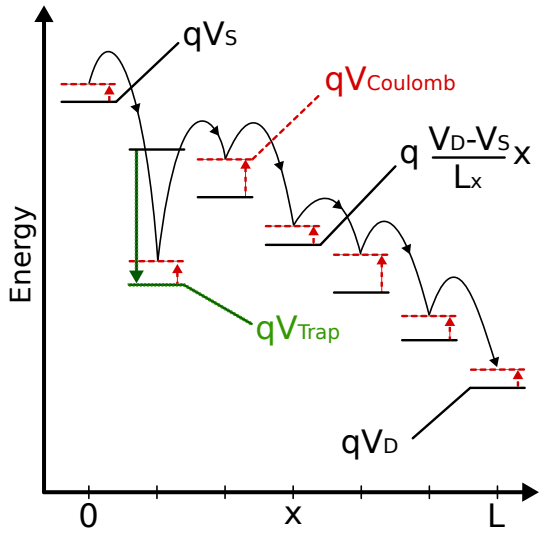


Figure 3.1: Injection and energy model.

or Monte Carlo attempts, for a given carrier's existence. Carrier path length is the total distance a carrier travels when moving from source to drain. Note that the path length is not a displacement or the mean free path and that the lifetime is not the time between "collisions". Indeed, due to electrostatic repulsion between carriers, actual collisions are rare events.

Two additional parameters that were also examined during simulations were the potential calculation procedure, and the initial condition. For the potential calculation procedure, Coulomb interactions between carriers were either included or left out. For the initial condition, two methods were employed. The first method is referred to as the seeding procedure, while the second is called the injection procedure. The methods are illustrated in Figure 3.1A. During the seeding procedure, carriers were first randomly placed on the lattice until the desired carrier concentration was obtained. Then, the system was allowed to equilibrate for a set number of steps, after which statistics were collected. For all steps after the random seeding, charges were injected from the source if the system was below the desired carrier

concentration. During the injection procedure, no random seeding occurred. The lattice started off empty, and charges were injected from the source if the system was below the desired carrier concentration. No equilibration period was employed so that equilibration behavior at short time could be examined.

3.3 RESULTS

First we consider how source-drain current varies with source-drain voltage, and find a turn-on source-drain voltage. Second, the instantaneous current is closely examined at short and long simulation times to determine the origin of the turn-on. Third, Fourier transforms of the instantaneous current are taken. Finally, in a further effort to explain the turn-on voltage, carrier lifetime and path length are investigated. The effects of Coulomb interactions are examined at each step, as well as the variance with lattice geometry where appropriate.

3.3.1 Current-Voltage Curves

Figure 3.2 shows the current-voltage curves for a 1024 by 512 nm system. Coulomb interactions were included for this system and the injection procedure was used. The system ran for 300 ns, and curves are shown at different simulation times.

The red curve (top) shows a typical current-voltage curve with a linear region below a source-drain voltage of 40 V, and a saturation region above 40 V. The curve, corresponding to simulation times after 200 ns, represents a well-equilibrated system. The curve lacks a turn-on source-drain voltage, where source drain current is not observed until a sufficient voltage is applied.

The green and blue curves show different current-voltage behavior, found under 20 ns, early in the simulation. The system is not well equilibrated. During this period, there is no current when the source-drain voltage is below 10 to 20 V. This region is termed the turn-on source-drain voltage. The turn-on voltage is not to be confused with the *threshold* voltage of a gate electrode. The simulation suggests it may be kinetic in nature, only appearing at

short times in non-equilibrated systems. It arises because of field-dependent mobility akin to the Poole-Frenkel effect (carriers do not have enough thermal energy to hop all the way to the drain at low electric field, which pulls them along, or in the absence of other carriers, which pushes them apart). Turn-on voltages are expected in the output characteristics of OFETs, as seen, for example, in the top-gate ambipolar light emitting OFETs demonstrated by Zaumseil et al., where the turn-on is gate-voltage dependent.[70]

The inclusion of Coulomb interactions alters the current-voltage curves. With Coulomb interactions, the magnitude of the current is non-zero at zero source-drain voltage. This occurs partly because carriers are not allowed to re-enter the source electrode once on the grid. A slight build up of charges near the source electrode creates repulsion that drives a small current. The effect is not present when Coulomb interactions are not included. Without Coulomb interactions, traps, or defects, at zero source-drain voltage all molecular sites have equal energy. In this unrealistic equal energy scenario, one might expect carriers to randomly walk to the drain given sufficient simulation time. The time is so long ($> 300\,\mathrm{ns}$), however, that no appreciable current develops.

If the seeding procedure is used (the starting configuration has a random distribution of carriers), even at early time, current-voltage curves reflect the prototypical OFET curve in red. The turn-on effect is not present, unless traps are added to the system.

In a real device, saturation is commonly explained by a pinch-off effect. The channel through which current flows is "pinched off" near the drain electrode as v_{ds} rises. A source-drain current continues to flow, however, because carriers still make it to the drain by traveling through a resistive region outside the pinched off channel. Although further increases in v_{ds} creates a larger pull on carriers towards the drain, the width and thus overall resistance of the resistive region increases. These effects balance, and lead to the saturation region.

In our simulation, saturation and a "pinched off" region also form. The maximum current is limited by the number of carriers in the device, which is inevitably tied to the injection rate and target concentration of the carriers. Both of these parameters are related to the gate electrode potential, which is constant, and only alters the magnitude of the current seen in the saturation region and the value of v_{ds} at which saturation arises. While the injection rate and

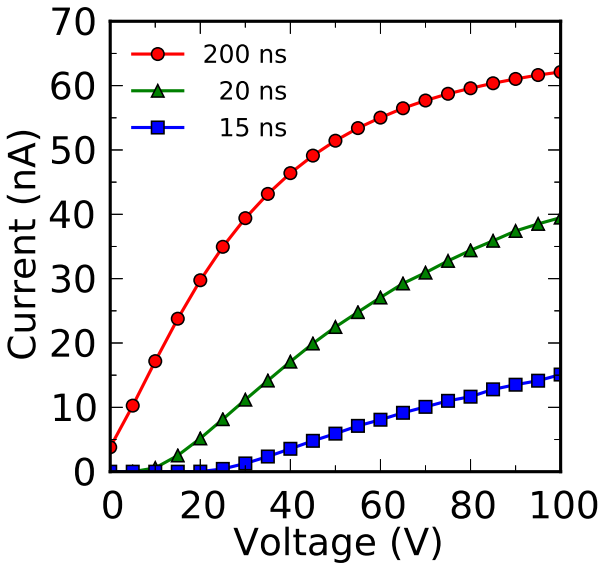


Figure 3.2: IV curve during turn-on.

average carrier concentration are channel approximations in our model, they are sufficient to yield pinch-off, channel formation, and current saturation behaviors because we include explicitly carrier electrostatic interactions. We tried more detailed injection procedures and they showed little differences from the simple approximation used in the bulk of this work, suggesting our results are independent of the injection procedure.

We also note that the maximum current is related to how efficiently carriers can travel across the device. The maximum rate at which carriers can travel from source to drain happens when the carriers travel the shortest path - a straight line perpendicular to the electrode surface. Carriers are likely to travel this path as v_{ds} increases, because the energy change associated with moving along the channel becomes increasingly large compared to other interactions, such as the Coulomb force, that may knock the carrier off this path. With a constant carrier concentration and carriers traveling at a maximum rate, the current must saturate.

3.3.2 Instantaneous Current

Typically, the current, as seen in Figure 3.2, is calculated as the average across the total simulation time. This rate, in units of carriers accepted per step, is converted to the average current by using the transfer rate (∼1 ps) and the charge of the carriers (1 e). The process can be repeated for smaller sampling intervals to calculate an instantaneous current which is a function of the simulation time. This instantaneous current oscillates, but approaches the average current as the system equilibrates. Note that we are talking about the magnitude of the current oscillating with time and not its sign.

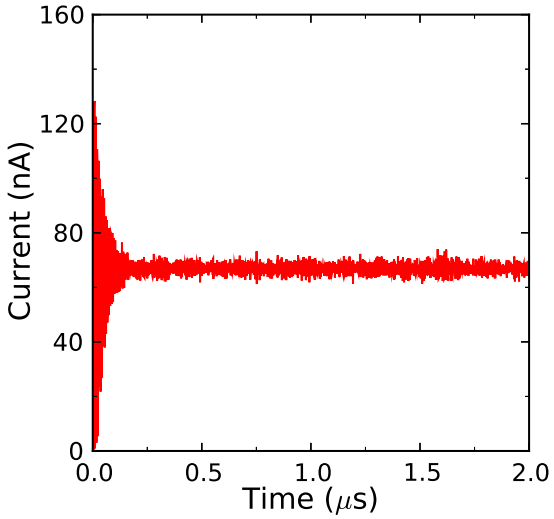
An oscillating instantaneous current can be seen in Figure 3.3A, where the injection procedure is used for a 512 by 512 nm system with Coulomb interactions. Note that the plot represents the value of the current at a single source-drain voltage point, 50 V, on a given current-voltage curve. At simulation times before 125 ns, the current is wildly oscillating at 64 ± 27 nA. Before 4 ns, the instantaneous is zero. At times after 125 ns, the system becomes equilibrated. While the average current remains 64 nA and still oscillates, the oscillations are much smaller at ± 2 nA.

As Figure 3.3B shows, when the seeding procedure is used, corresponding to an already equilibrated device, the current still oscillates with a standard deviation of 2 nA. However, the wild oscillations are not present before 125 ns and the overall magnitude resembles the injection procedure value of 64 nA after 125 ns.

An oscillating instantaneous current occurs with or without Coulomb interactions. However, the oscillations are smaller when Coulomb interactions are included. For example, after 125 ns, oscillations without Coulomb interactions were ± 5 nA as compared to ± 2 nA with Coulomb interactions.

As channel length increases, the current magnitude does not change because the carrier concentration is kept constant between the systems, and the surface area of the drain does not change with device length. However, the large oscillation region at short simulation time takes longer to equilibrate. As channel width increases, the current magnitude increases, as expected because the drain surface area increases with device width.

(A) injection proceedure



(B) seeding proceedure

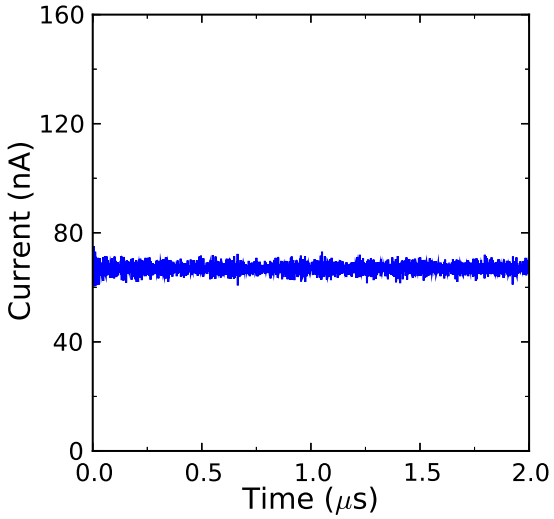
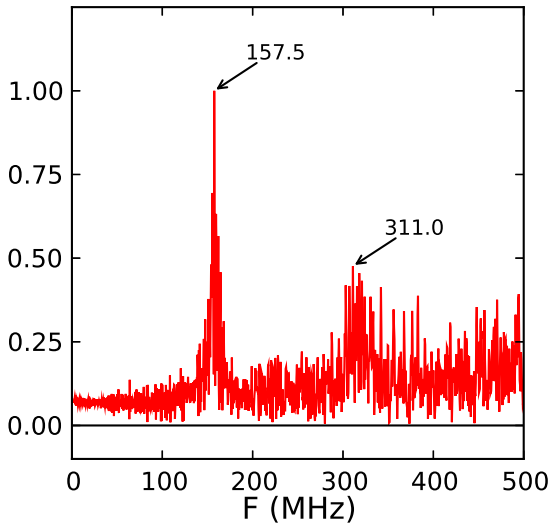


Figure 3.3: I(t) injection vs seeding.

3.3.3 Fourier Transformed Current

To better examine the oscillatory behavior, a Fourier transform of the instantaneous current was taken for a range of device geometries. Figure 3.4A and Figure 3.4B show the Fourier transform power spectra for a 512 by 512 nm system with Coulomb interactions. The sampling interval was 1 ns and the sampling period was 2 µs. In the frequency domain, this gave a resolution of 500 kHz, a bandwidth of 1 GHz, and Nyquist frequency of 500 MHz. Before performing the Fourier transform, a linear fit of the signal was subtracted and a Hamming window applied. A given power spectrum shows relative intensity, where all peak intensities are scaled by the maximum intensity. Figure 3.4A shows a Fourier transform power spectrum of a simulation using the injection procedure. Peaks are seen at 157.5 and at 311.0 MHz. Figure 3.4B shows a Fourier transform power spectrum of a simulation using the seeding procedure. Peaks are seen at 155.5, 315.0 and 479.0 MHz. The peak location is only minutely affected by equilibration procedure. The higher frequency peaks are roughly integer multiples or overtones of the lowest frequency or fundamental peak (at ∼155-157 MHz).

(A) injection proceedure



(B) seeding proceedure

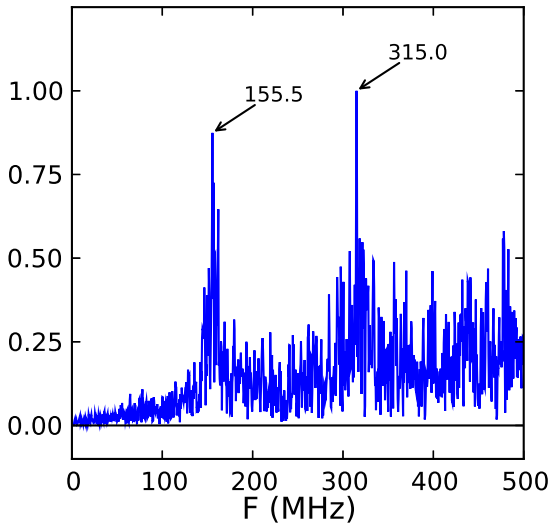


Figure 3.4: FFT of I(t) injection vs seeding.

Although Figure 3.4B is an exception, the fundamental almost always has the highest intensity. The exception may be an artifact of the bin size or frequency domain resolution, which is determined by the sampling interval in the time domain.

When Coulomb interactions are not included, a given power spectrum becomes sharper because, as mentioned when discussing the instantaneous current above, there are larger amplitude oscillations when not including Coulomb interactions. As device length increases, the fundamental peak and all other peaks are found to shift towards lower frequencies. The shift occurs in a non-linear fashion. As device width increases, the fundamental peak and all other peaks do not shift. We note that decoherence of the Fourier peaks arises when moving from monolayers to multilayers.

Figure 3.5 shows Fourier transforms of instantaneous current for geometries with heights and widths in the range 128 nm to 384 nm. Notice that the peak locations shift to lower frequency with increasing device length, but do not change with device height.

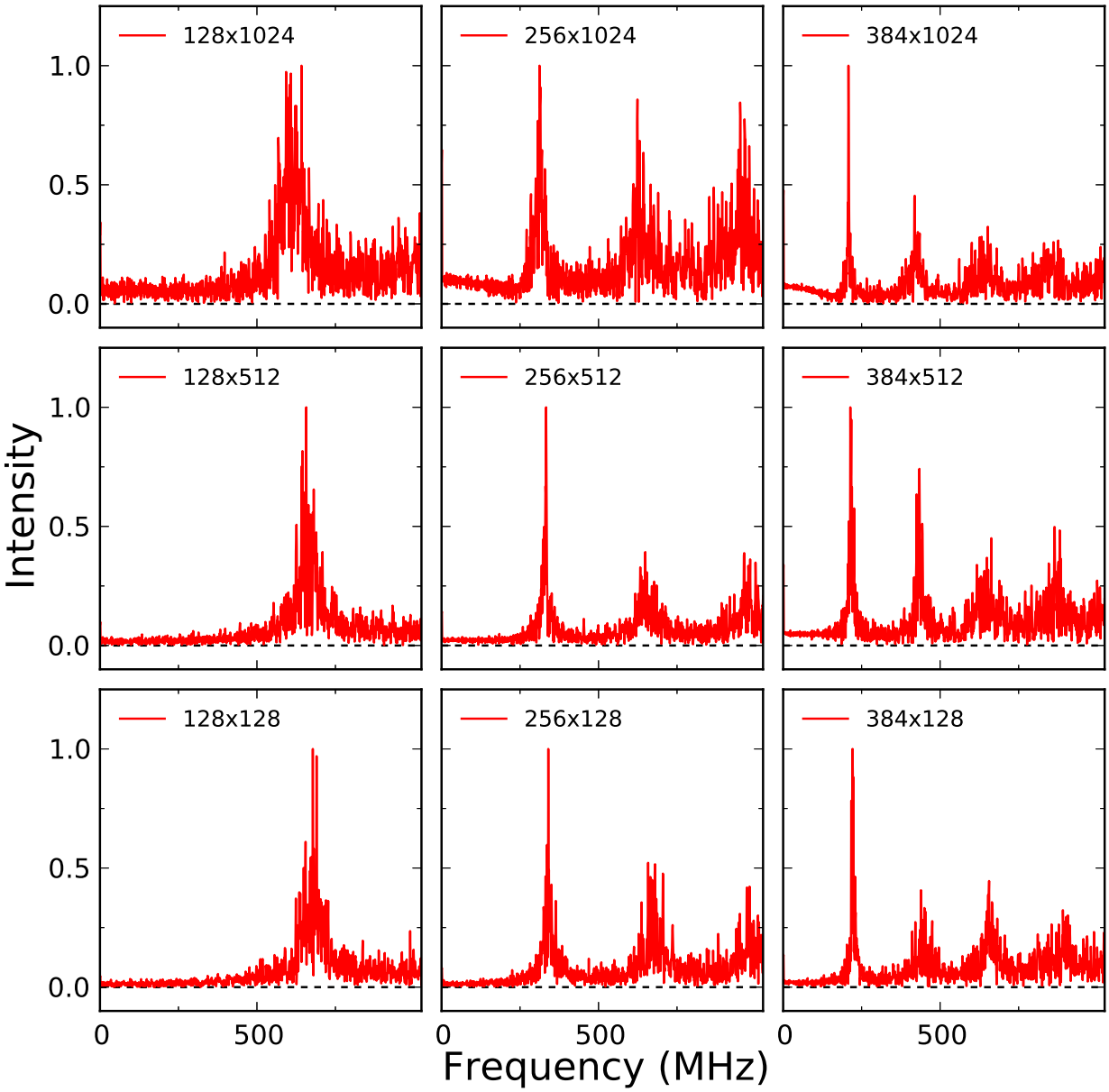


Figure 3.5: FFT of I(t) many geometries.

3.3.4 Carrier Lifetime

Based on our findings from the Fourier transforms, which exhibit clear frequency peaks, the carrier lifetime was monitored for systems with and without Coulomb interactions for same device geometries. For every carrier in a simulation, a counter is created when the carrier is injected from the source. Every time this carrier attempted a move, successful or not, the counter was increased. When a given carrier reached the drain electrode, the carrier reported the total number of counts. Each count contributes 1 ps to the carrier's lifetime. Histograms of the lifetimes were performed, and the shapes were nearly Gaussian.

Average lifetime remained relatively constant with device width for a given device length. This suggests that the path taken by carriers is not influenced by device width. In contrast, as Figure 3.6B shows, the average lifetime increased with device length for a given device width. The quadratic fit performs better than the linear fit (see R^2 values in Figure 3.6B), suggesting that the lifetime varies non-linearly with device length. This behavior was seen for all device widths examined. When Coulomb interactions are not included, the histograms remained nearly Gaussian and the averages increased by approximately 1.0 %. The shape and location of the histograms did not depend upon the seeding or injection procedure.

3.3.5 Carrier Path Length

Much like carrier lifetime, the carrier path length was monitored for systems with and without Coulomb interactions for the same device geometries. The path length is defined here as the total distance traveled by a carrier as it journeys from source to drain. In a similar spirit to carrier lifetime, each carrier starts a counter when it comes into existence. The counter is only incremented if a carrier successfully performs a move. When a given carrier reaches the drain electrode, the carrier reports the total number of counts. Each count contributes 1 nm, corresponding to the lattice spacing, to a carrier's path length. If carriers traveled a perfectly straight path from source to drain, then this path length would simply be the device length.

The behavior of path length with device geometry was similar to carrier lifetime, but not identical. For a given device length, path length remained constant with device width,

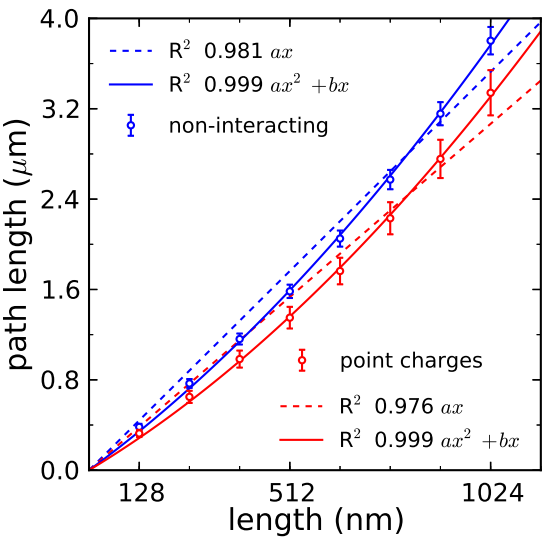
suggesting carriers do not sample the entire width of the channel. However, there was a "narrow channel effect" for devices with widths less than or equal to 256 nm when Coulomb interactions were included. For these narrow channels, the path length is roughly 50 nm less and the lifetime 150 ps less than the constant width values, suggesting carriers sample ∼256 nm of device width on average. As Figure 3.6A shows, for a given device width, path length increased with device length. Again the R^2 values suggest that the path length increased non-linearly. In contrast to lifetime, when not including Coulomb interactions, the magnitude of path length became noticeably larger. For example, the path length is approximately 15 % larger without Coulomb interactions for a 512 by 512 nm geometry. This suggests, not surprisingly, that carriers sample more of the width when Coulomb interactions are present, due to electrostatic repulsion. Additionally, the histograms are Gaussian without including Coulomb interactions but have an asymmetric tail towards lower path lengths with Coulomb interactions. The shape and location of the histograms do not depend upon the seeding or injection procedure.

Figure 3.7B shows the average lifetime as a function of device width for a constant device length of 384 nm including Coulomb interactions. The lifetime remains constant with device width except below 384 nm, where the lifetime becomes noticeably smaller. The same behavior is seen for the path length, shown in Figure 3.7A. The standard deviation of the data, shown by the error bars or seen visually in the histograms shown in Figure 3.8, also does not change significantly with device width.

3.3.6 Energy

Figure 3.8 shows histograms of path length and lifetime for a 256 × 256 nm device with and without Coulomb interactions. Regardless of Coulomb interactions, the lifetime takes a Gaussian shape and similar values. For the case shown, the maximum is only 0.95 % smaller when including Coulomb interactions. The path length maximum, however, is 13.0 % smaller when including Coulomb interactions. Additionally, the shape is non-Gaussian, with a tail towards low path length when including Coulomb interactions. The result is independent of the seeding or injection procedure. Changing the device length changes the location of the

(A) path length



(B) lifetime

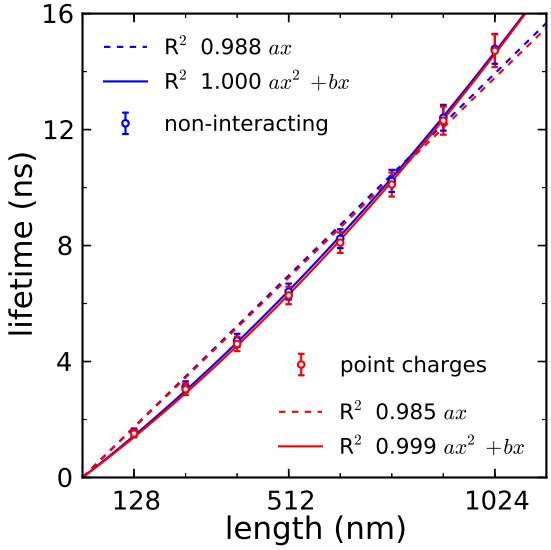
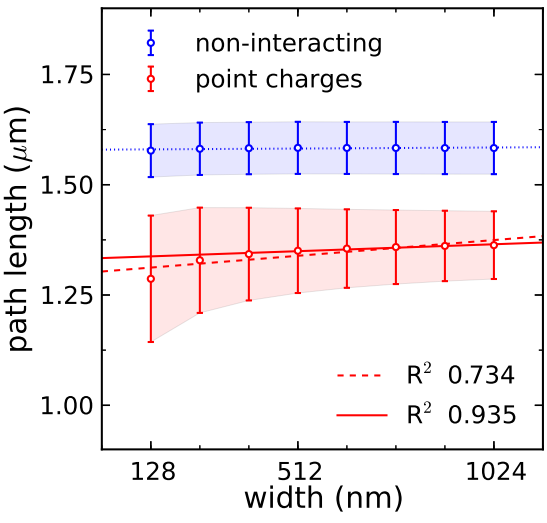


Figure 3.6: Lifetime and path length vs device length.

(A) path length



(B) lifetime

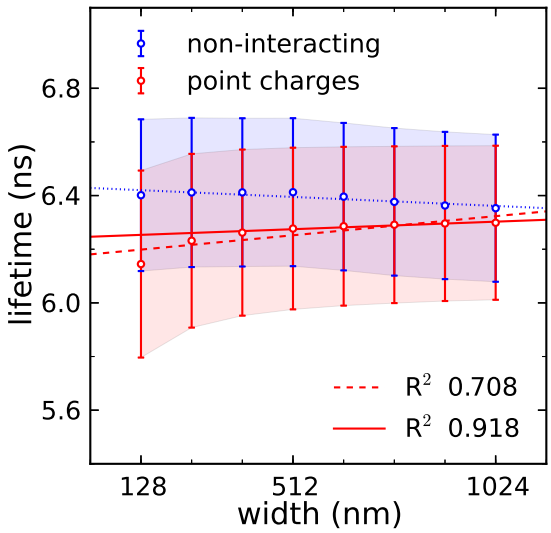
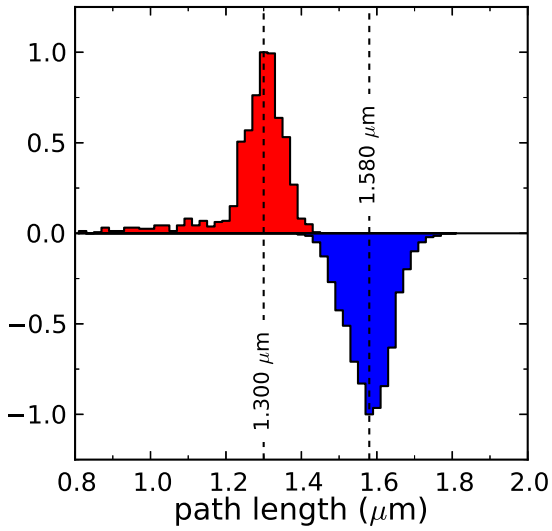


Figure 3.7: Lifetime and path length vs device width.

(A) path length



(B) lifetime

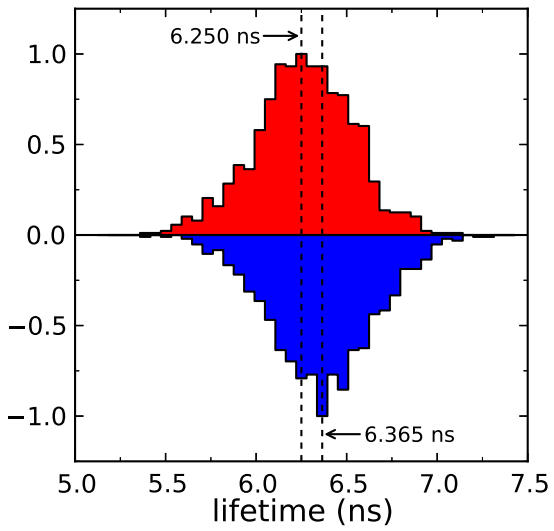


Figure 3.8: Lifetime and path length histograms.

maximum.

Figure 3.9A shows the Coulomb potential at every location in a 256×256 nm. Figure 3.9B shows a linescan of this potential taken at $y = 128$ nm. While the system has reached the desired concentration and an equilibrated state, the carriers are not distributed isotropically. There is a depletion region near the drain electrode (right). Also, the majority of carriers are traveling paths in the central region of the device because they are being repelled by other carriers lining the surface.

3.4 DISCUSSION

The behavior of the source-drain device current in this model, namely its oscillatory nature, is a function of charge injection at the source electrode. As part of the simulation model, the number of carriers on the lattice is maintained by only allowing the source to inject charges when the number of charges is below some desired concentration. Injection is then

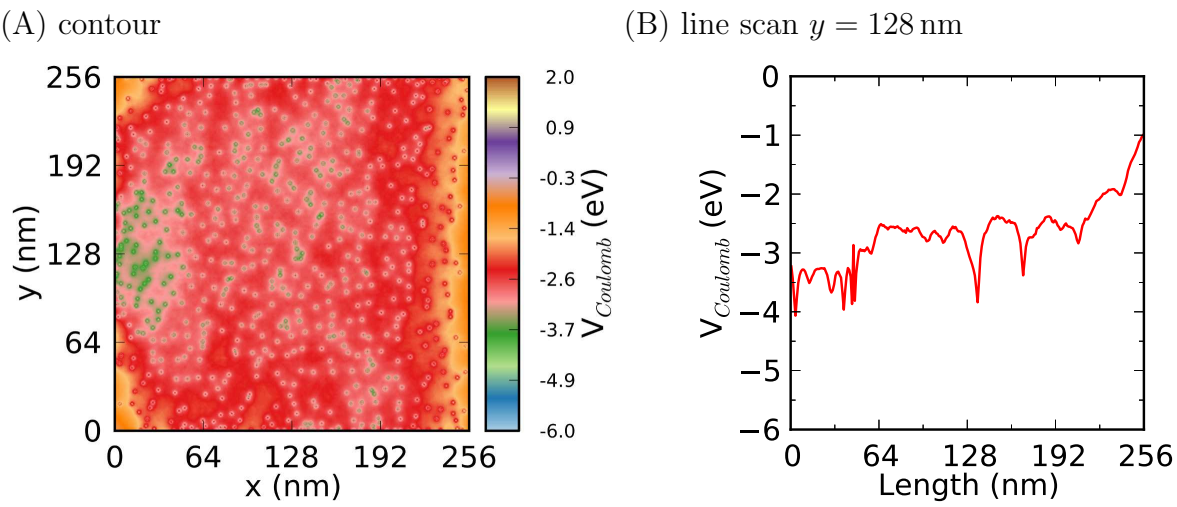


Figure 3.9: Energy contour of thin film.

allowed to occur with a certain constant probability. The procedure reflects the role of the gate electrode and a charge injection barrier in the system.

The gate electrode is related to how charges are injected in the system. In a real system, the gate opens up the conduction pathway by freeing charges in the system, a process captured by the random seeding procedure. We note that the current oscillations are present with or without this procedure. Otherwise, the gate lowers the molecular energy levels relative to the source so that carriers can be injected and a current flow, a process captured by the injection procedure. Equivalently, as the gate turns on, charges discharge. The freed up sites allow for more carriers to flow in from the source. This effect, and others, such as an injection barrier, are captured by the probabilistic injection procedure, whether constant or calculated on the fly.

The understanding of charge injection barrier is important, because the development of high efficiency organic (opto-)electronic devices is primarily based on charge injection at the interface between metallic electrodes and organic semiconductors.[71] The energetics of this metal/organic interface play a key role in charge injection,[72,73] and there are many models to describe the mechanism.[74–76] Models differ because the details of the process depend on

the system, however there are several common themes. For an up to date review, see Natali and Caironi.[77]

Many parameters can, and do modulate charge injection (electric field, image potential, barrier height, temperature, mobility, mechanism, etc). The mechanism determines which parameters are significant. A carrier may tunnel from delocalized states in the metal to localized states in the organic semiconductor at low temperature (low field), or surmount a barrier by hopping at high temperature (high field). The contact between metal and semiconductor may not be perfect, and an energy barrier will be present due to many factors.

For example, there is an energy difference between the semiconductor HOMO/LUMO and metal work function, and a surface dipole may form. Additionally, there may be significant rearrangement of the organic molecules after charge transfer, or the organic molecules may have reacted with the metal surface. Whatever the mechanism or origin of the energy barrier may be, we are assuming that carriers are accomplishing injection in an average probabilistic fashion. The environments a carrier sees upon injection are averaged out both over different locations near the source, and as time progresses. Simply put, the steady-state rate of injection is roughly constant. This assumption is a coarse graining of the many factors. There is one exception, namely the limitation of carrier concentration, that alters this picture and produces the subtle effects on the simulation that are the subject of this paper.

We have also considered and tested other injection procedures in these simulations. For example, an energetic barrier to injection based on the Coulomb interaction of the injected charge with other charges in the system, as well as their image charges was used. The calculated barrier was used in the Boltzmann factor to determine an injection probability. However, this procedure did not show any significant differences from a constant probability model. This is because, after equilibration, the distribution of carriers near the source does not change significantly. Therefore, the calculated injection probability is essentially constant. The exact value of the constant probability of injection only influences the exact magnitude of the current. It does not alter the shape of current-voltage curves or the oscillatory behavior of the instantaneous current discussed above. The oscillatory behavior, instead, arises from restrictions placed on the carrier concentration, as will now be explained.

When the lattice is completely empty, as is the case at short times during the device turn-on, the source injects charges at almost every time step. This in turn creates a "wave" of charges that build up near the source electrode, and begins to travel towards the drain electrode. The source stops injecting once it reaches the desired concentration, ending the wave of carriers. The carrier wave spreads out as it travels to the drain. The rate of travel is influenced by the energy landscape. As the magnitude of the source-drain potential increases, the carriers reach the drain more quickly. When Coulomb interactions are present, the carriers repel each other and the wave spreads out more rapidly.

Before the carrier wave reaches the drain, there is essentially no current. This is seen at short times and low voltage in Figure 3.2. Once the wave reaches the drain, a large spike in current occurs. This can occur at short times only if the source-drain voltage is large enough or Coulomb interactions included. This is why a turn-on source-drain voltage is seen. Experimental verification would require capturing and examining the current-voltage curve within the first 15 ns.

Only as carriers reach the drain, can the source begin to inject charges again. A new wave of carriers is released into the lattice, and as this process repeats, the waves become more and more diffuse. Eventually, the system reaches a more uniform equilibrated steady state which is similar to that created by an initial seeding of random carriers. Regardless of using the seeding procedure or injection procedure, systems evolve to this equilibrated spread out state.

Even when equilibrated, however, the pulsing nature of the source still causes the oscillatory current and peaks in the Fourier transforms. For example, the fundamental for the 512 by 512 nm system occurs, regardless of equilibration procedure, near 158 MHz. This means that every $\sim$6 ps, carriers reach the drain in the equilibrated system. The oscillatory nature remains because, in the equilibrated system, the number of charges is very close to the desired concentration. When the system is at the desired concentration the source will not inject charges. Likewise, when the system falls below the desired concentration a wave of charges is injected by the source. While this is a consequence of requiring constant carrier concentration in our simulations, the use of more detailed injection models shows little effect.

The non-linear decrease in fundamental frequency in the Fourier transforms and the

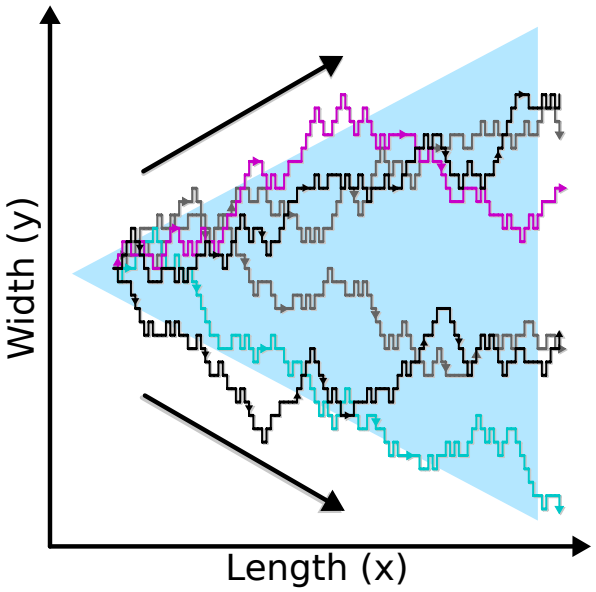


Figure 3.10: Carrier trajectory spreading with device length.

non-linear increase in carrier lifetime with device length are connected. As the device length increases, a given carrier has more of a chance to stray from a straight path, as illustrated in Figure 3.10. Thus, the path it travels from the source to the drain becomes longer, and therefore the time it takes to get from the source to the drain longer. As the lifetime increases, the time between when the source can and cannot inject charges increases. This increase correlates to the decrease in the fundamental Fourier frequency with device length. In a similar manner, the lifetime and path length do not change with device width because the average path taken remains similar with a constant device length. Consequently, the fundamentals do not change with device width either.

3.5 CONCLUSIONS

We explored charge transport in thin-film organic field effect transistors during and after device turn-on using Monte Carlo simulation, paying close attention to the effect of device dimensions and Coulomb interactions. We used two initial configurations for charge carriers, an empty lattice and a randomly seeded lattice. In both cases, charges were injected with a constant probability throughout the simulation and a desired carrier concentration was maintained, and we found that a more sophisticated injection model was not needed.

The results for an empty lattice, independent of the presence of Coulomb interactions, showed a turn-on source-drain voltage in the current-voltage curve that was, however, transient, and disappeared after the system equilibrated, resembling a typical current-voltage curve for a randomly seeded system. On closer examination, we found that the turn-on for an empty lattice was linked to an oscillating instantaneous current during device turn-on. More surprisingly, the oscillating current was present in the well-equilibrated lattice, though lower in magnitude, even for a randomly seeded lattice, which showed no turn-on voltage.

The Fourier transform of current as a function of simulation time, independent of initial configuration of charge carriers (empty lattice or randomly seeded) and the presence of Coulomb interactions, showed well-defined fundamental peaks and overtones that were a function of device dimensions. While independent of device width, peak frequencies and their overtones shift, non-linearly, towards lower frequencies as device length increased. Additionally, carrier lifetime and path length were found to increase non-linearly with device length and remain constant with device width. We proposed that the nonlinear behavior of the carrier lifetime, related to the frequencies observed in the Fourier transform, stems from nonlinear increase of carrier path length with device length.

Finally, we proposed that the oscillating behavior of the current was a result of carrier injection from the source, which could not inject charges when the system was at the free charge carrier concentration limit. As a result of this injection process, carriers were found to travel in waves of charge density even in the well equilibrated device.

Observing non-linear behavior in experimental devices would be a worthwhile exercise. That is, how do carrier lifetime and path length change with device dimensions, and what

would this mean for concepts like carrier mobility? Likewise, do charge carriers travel in waves in real devices? If our proposed mechanism is correct, then they may, for real devices do have a restricted free carrier concentration, fixed by a gate electrode, as well as injection probability. To answer these questions, however, device properties like source-drain current would need to be resolved to timescales on the order of nanoseconds. This is not outside the realm of techniques like complex impedance spectroscopy, used to measure AC conductivity, and provide information on processes with diverse timescales. For example, dielectric spectroscopy measurements have been made of doped polythiophene, poly(phenylene vinylene),[78] or polypyrrole,[79] which have measurements in the sub-THz range ($\tau \approx \omega^{-1}$, THz $\sim$ps).

4.0 TRAPS, BARRIERS, AND DEFECTS : CONCENTRATION AND NEGATIVE DIFFERENTIAL RESISTANCE

The text in this chapter has been adapted from Chen et al., *Charge Transport in Imperfect Organic Field Effect Transistors: Phthalocyanine Mixtures as Charge Trapping Models*[62] and Madison et al., *Charge Transport in Imperfect Organic Field Effect Transistors: Effects of Charge Traps*.[60] These works were coauthored.

4.1 INTRODUCTION

In this chapter, we discuss simulations of OFETs when defects, barriers, or traps are present in the organic semiconductor. Barriers and traps are molecules with higher or lower HOMO or LUMO levels than the native semiconducting molecule. When a carrier hops to a trap site, it has significantly lower probability to hop out of the site. This manifests itself as a lowering of the current observed in OFET current voltage curves. Likewise, there is a lower probability for carriers to hop onto barriers sites. Defect may be thought of as barrier sites with an HOMO-HOMO or LUMO-LUMO offset so high, that the probability to hop to the site is essentially zero.

Figure 4.1 shows the energy of an electron in the field produced by a source and drain electrode. Electrons are attracted to regions of positive potential. This means that a "trap" for electrons is higher in potential. Electrons carry a negative charge. Therefore, higher regions of potential turn into lower regions of energy. A similar picture can be drawn for holes. However, holes are attracted to regions of negative potential. They still jump "down" in energy, However, it is the custom to always draw energy levels in terms of the energy of

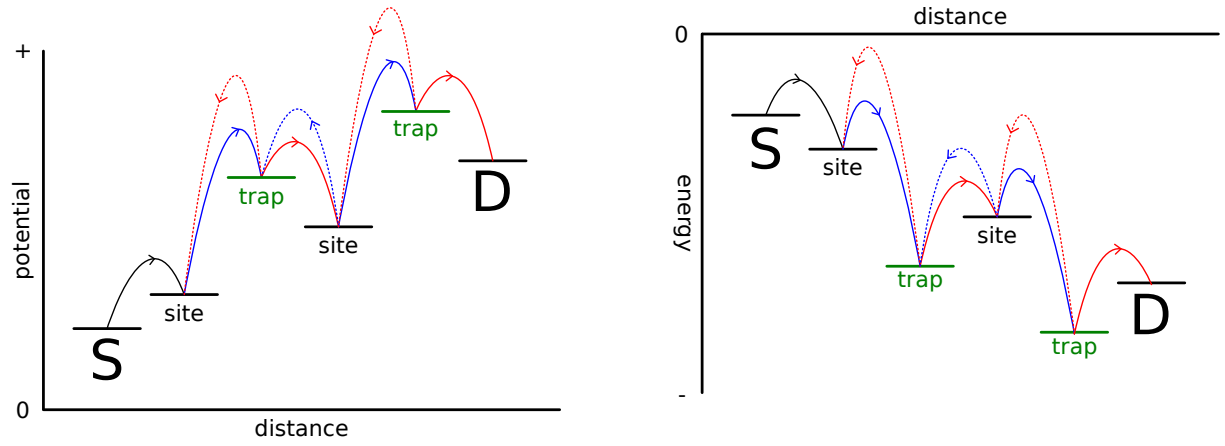


Figure 4.1: Potential and energy with traps for electrons.

the electron. Therefore, one will often encounter the idea that holes "jump" up in energy. What is meant is that the holes jump up the *electron* energy landscape. In reality, as holes move towards more negative potentials, they release energy. The behavior is summarized in Table 4.1.

Why is it necessary to differentiate between barriers and defects? As will be shown in this chapter, there is a sharp cutoff in energy where barriers start behaving like defects. On a technical note, barriers and defects are treated differently in the simulation. Defects are sites which carriers are not allowed to travel to. Barriers, like traps, are sites with shifted energies. In this sense, very high defects act as *neutral* defects.

What about very deep traps? An example might be interspersed PCBM in P3HT (dis-

		$\Delta V > 0$	$\Delta V < 0$
hole	$q > 0$	barrier	trap
electron	$q < 0$	trap	barrier

Table 4.1: Potential for holes vs electrons.

cussed in chapter 5 where it is called "pepper"). One might term such sites, "defective" as well. If electrons transport to "pepper", then will get stuck there, making the site in some sense defective. Once the electron is trapped, other carriers can not travel to the site. In this way, deep traps act as *charged* defects.

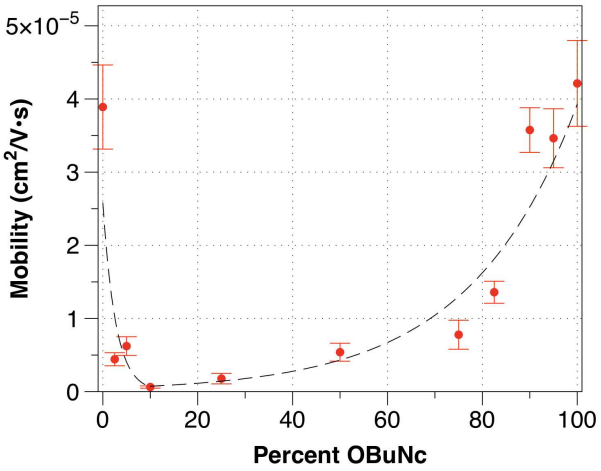
Defects, barriers, and traps all affect the magnitude of current and shape of an OFET curve. As we will see, increasing the concentration of traps and barriers of modest energy shift ($< 0.05 - 0.15\,\mathrm{eV}$) leads to a decrease in the observe current of an OFET. While the minimum is observed around 40 %, the current recovers as the concentration of traps or barriers passes 60 %.

In the saturation region of an OFET curve, the story changes. The large electric field ($e * 100\,\mathrm{V}/1000\,\mathrm{nm} * 1\,\mathrm{nm} = 0.1\,\mathrm{eV}$) provides a driving force to remove carriers from traps as deep as or a little deeper than 0.1 eV. This is not the case for barriers and defects. There is a high probability for electrons to move towards the drain electrode, which is at a more positive potential relative to the source electrode. However, while this potential can drive a carrier onto modest barrier sites, it can not push a carrier through a defect or very large barrier site ($> 0.4\,\mathrm{eV}$). Instead, carriers will actually get stuck behind the defects. As this chapter will show, this leads to negative differential resistance.

4.2 TRAP AND BARRIER CONCENTRATION

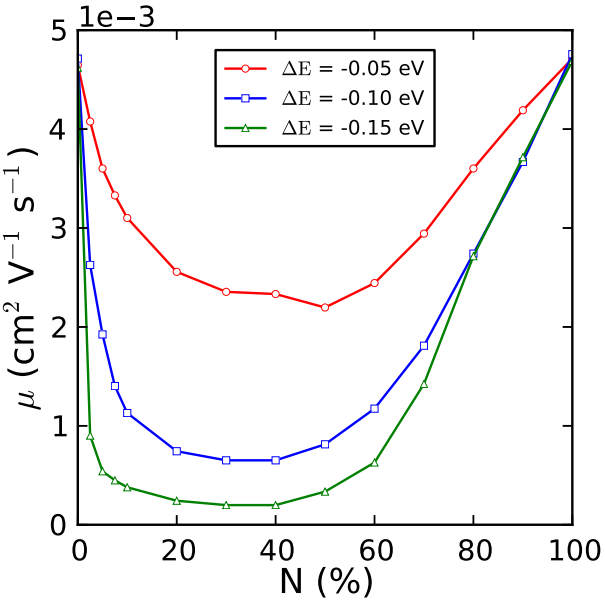
The behavior of current in OFETs as a function of traps concentration at modest trap energies ($0.0\,\mathrm{eV} < \Delta V \leq 0.3\,\mathrm{eV}$ for electrons, $-0.3\,\mathrm{eV} \leq \Delta V < 0.0\,\mathrm{eV}$ for holes) is anti-symmetric. That is, the current does not reach a minimum at 50 % traps, as would be predicted by percolation theory.

Figure 4.2 shows the current as a function of trap percentage for an experimental system obtained by others.[62] The system consisted of Ni(II)-octabutoxy-phthalocyanine (NiOBuPc) and octabutoxy-naphthalocyanine (OBuNc) as a barrier model. The current decreases rapidly with only a small introduction of OBuNc. This suggests that OBuNc is acting as a high barrier or defect, or other defects are in the system. The current begins to recover



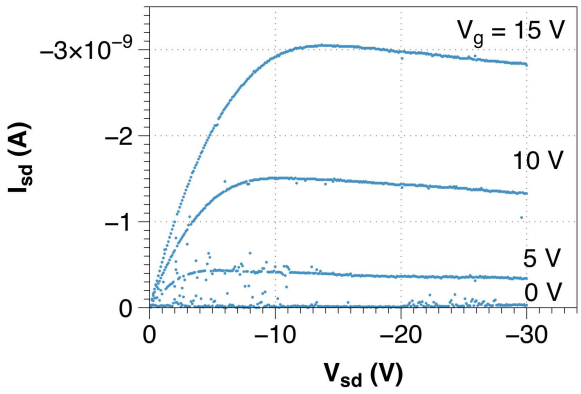
From the work of Chen et al..[62]

Figure 4.2: Experimental trap concentration scan.



Adapted from Madison et al..[60]

Figure 4.3: Mobility vs trap concentration.



From the work of Chen et al..[62]

Figure 4.4: Experimental IV curve with barriers.

as early as 20 %, although significant recover does not occur until 80 %.

Rather, the minimum is reached more rapidly ($\leq$ 20 % traps) for energies as low as $|\Delta V| \leq 0.15\,\text{eV}$. Likewise, recovery of the current happens much earlier ($\sim$60 %) than expected based on the initial falloff. Barriers of modest energy ($0.0\,\text{eV} < \Delta V \leq 0.3\,\text{eV}$ for holes, $-0.3\,\text{eV} \leq \Delta V < 0.0\,\text{eV}$ for electrons), produce a current vs. concentration curve that is a mirror image of the trap system.

4.3 NEGATIVE DIFFERENTIAL RESISTANCE

Simulations of thin film OFETs with dimensions of $1024 \times 256 \times 3\,\text{nm}^3$ with 1 % charge carriers (7864 max carriers) were performed. Current was calculated for 0.5 µs and results averaged over three simulation replicas. Only hole transport was considered. Coulomb interactions were treated using point charges, delocalized Gaussian charges, and non-interacting charges. For Gaussian charges, the standard deviation chosen was 1, 1.25 and 1.5 nm. IV curves were calculated for systems with a mixture of 75 % semiconductor and 25 % shallow traps ($-0.05\,\text{eV}$), 25 % high barriers (0.50 eV), and 25 % defects. V_{DS} was varied between

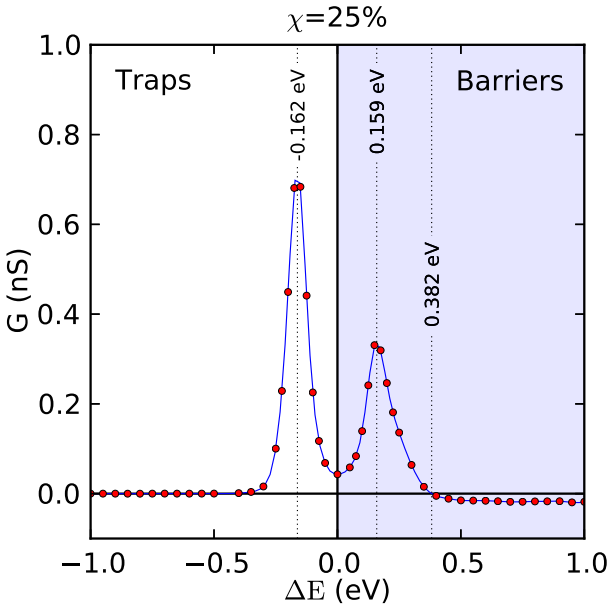


Figure 4.5: Scan of saturation slope vs trap energy.

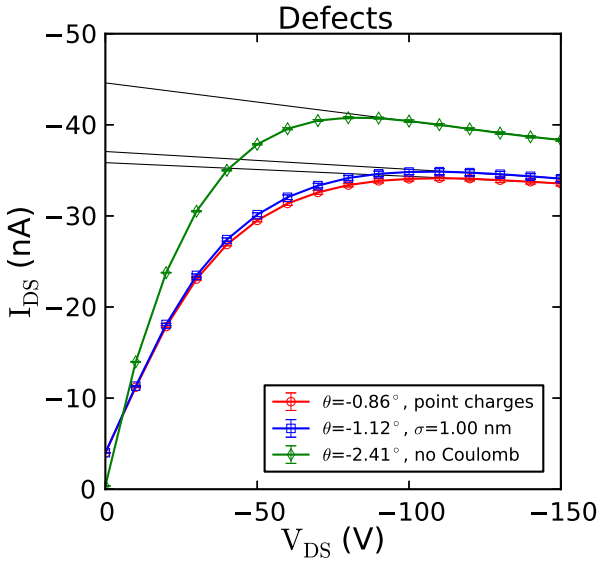
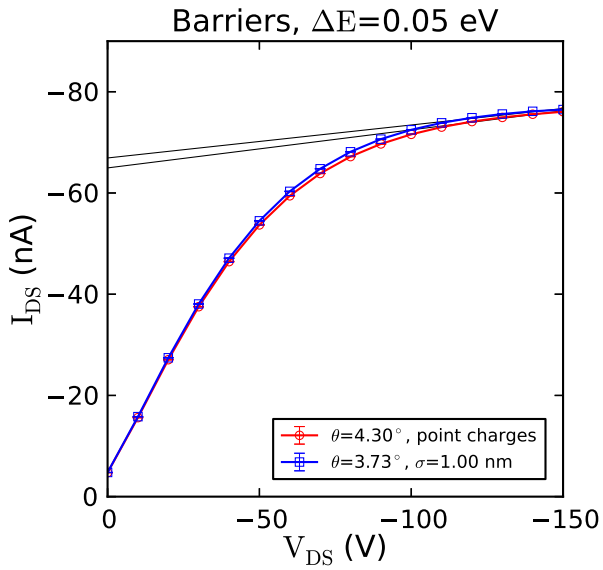


Figure 4.6: IV curve with defects.

(A) shallow barriers



(B) high barriers

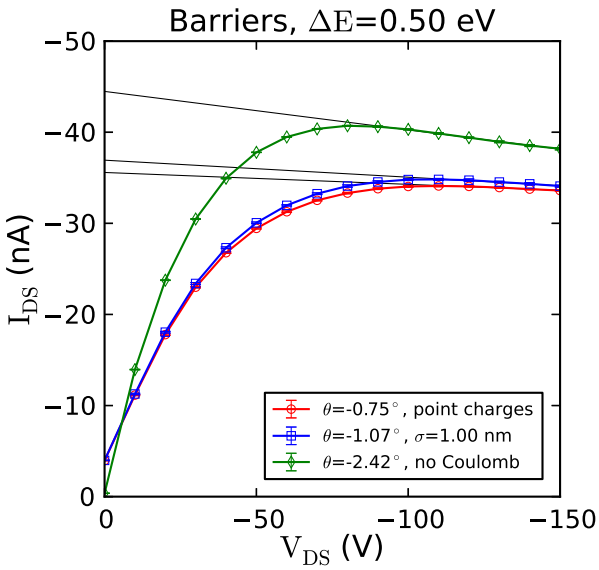


Figure 4.7: IV curve with barriers.

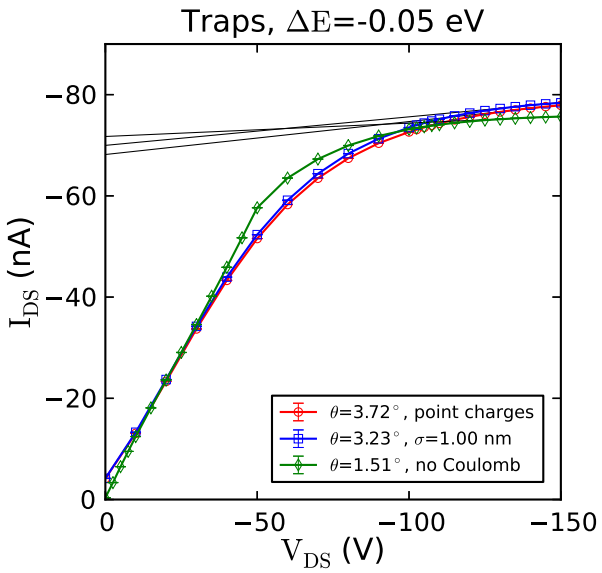


Figure 4.8: IV curve with shallow traps.

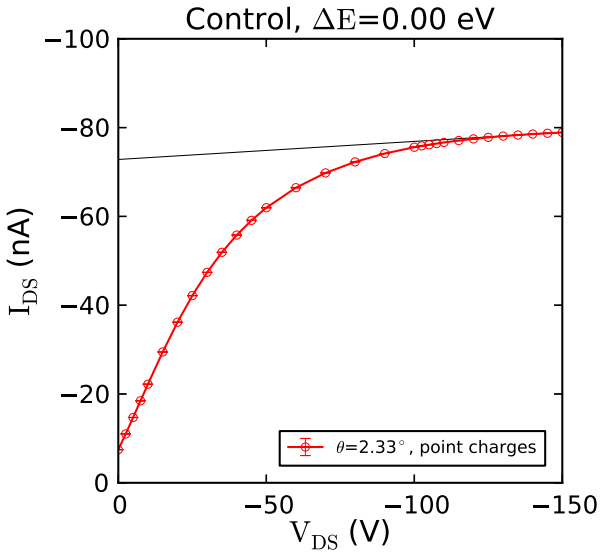


Figure 4.9: IV curve with no traps, barriers, or defects.

0 and −150 V. Additionally, the slope of the IV curve in the saturation region (differential conductance) was measured as a function of trap/barrier energy (−1.0 to 1.0 eV) by calculating a few IV points in the saturation region (120, 140 and 150 V). Figure 4.5 shows the differential conductance in the saturation region as a function of trap or barrier energy for a system with 25 % traps or barriers. Differential conductance indicates how the conductance is changing as the voltage increases and is simply the slope of the IV curve. When the trap energy is zero (pure system), the differential conductance is small and positive (0.04 nS). Ideal saturation would mean that the conductance does not change with voltage. This system is very near, but not completely saturated. Interestingly, as the trap or barrier energy increases, the differential conductance reaches a maximum at ±0.16 eV. The peak for traps (0.68 nS) is nearly twice the height as the peak for barriers (0.32 nS). A larger differential conductance indicates a perturbed onset of ideal saturation. Therefore, this suggests that traps are a larger hindrance to charge transport than barriers, likely because it is thermodynamically favorable to fall into a trap. The peak location is near the energy associated with the potential difference between the source and drain electrodes at satu-

ration for an adjacent site hop ($e(150\,\mathrm{V})/(1024\,\mathrm{nm})(1\,\mathrm{nm}) \approx 0.15\,\mathrm{eV}$). It is this energetic driving force, along with the Columbic interactions between carriers, which drives carriers out of traps or into/past barriers. The presence of a peak indicates that traps/barriers are (at first) an increasing hindrance to ideal saturation. As traps deepen beyond $-0.16\,\mathrm{eV}$, the differential conductance begins to fall and remains zero for traps deeper than $-0.4\,\mathrm{eV}$. Although the differential conductance calculated is zero beyond $-0.4\,\mathrm{eV}$, a saturated IV curve does not exist. When traps are this deep, there is no current, because nearly all carriers become trapped. There are many more trap sites than carriers. The simulation assumptions, such as constant carrier concentration, may be invalid in this region. Barriers behave differently as they become larger than 0.16 eV. The differential conductance falls off, and becomes negative beyond 0.4 eV. At this energy, the barriers become insurmountable hills or defects. Figure 4.7B shows IV curves for systems with 25 % high barriers (0.50 eV). Curves are shown for Coulomb interactions treated as point charges, delocalized Gaussian charges (σ = 1.0, 1.25 and 1.50 nm), and non-interacting charges. All curves exhibit negative differential conductance in the saturation region (-0.013, -0.019, -0.021, -0.025 and $-0.042\,\mathrm{nS}$). As charge interactions become weaker (higher sigma), the negative differential conductance becomes more negative and thus the system more resistive. A mechanism for defect induced negative differential conductance was proposed by Hanwell.[16] Non-interacting carriers become blocked/immobilized by the presence of defects, leading to a negative differential conductance. However, when Coulomb interactions are present, other carriers are funneled around the defects by the Coulomb interaction with blocked carriers. This leads to a smaller differential conductance and faster onset of ideal saturation. Figure 4.6 shows IV curves for systems with 25 % defects. Again, curves are shown for the varying treatments of Coulomb interactions. Curves show a negative differential conductance in the saturation region (-0.015, -0.020 and $-0.042\,\mathrm{nS}$) of similar magnitude to the corresponding curves for high barriers (-0.013, -0.021 and $-0.042\,\mathrm{nS}$). This indicates that the barriers are behaving like defects. Figure 4.8 shows IV curves for systems with 25 % shallow traps ($-0.05\,\mathrm{eV}$). Curves are shown for Coulomb interactions treated as point charges, delocalized Gaussian charges with $\sigma = 1.00\,\mathrm{nm}$, and non-interacting charges. All three exhibit positive differential conductance (0.065, 0.057 and 0.026 nS), in agreement with Figure 4.5. Non-interacting

charges show the smallest differential conductance in the saturation region and reach near ideal saturation at lower voltages ($\sim -100\,\mathrm{V}$) than interacting charges ($> 150\,\mathrm{V}$). However, non-interacting charges show lower current (-75 nA) in the saturation region compared to interacting charges ($-80\,\mathrm{nA}$). This is a result of Coulomb interactions removing carriers from trap sites.

4.4 CONCLUSION

The behavior of current in OFETs as a function of traps concentration at modest trap energies ($0.0\,\mathrm{eV} < \Delta V \leq 0.3\,\mathrm{eV}$ for electrons, $-0.3\,\mathrm{eV} \leq \Delta V < 0.0\,\mathrm{eV}$ for holes) is anti-symmetric. That is, the current does not reach a minimum at 50 % traps, as would be predicted by percolation theory. Rather, the minimum is reached more rapidly ($\leq 20\,\%$ traps) for energies as low as $|\Delta V| \leq 0.15\,\mathrm{eV}$. Likewise, recovery of the current happens much earlier ($\sim$60 % traps/barriers) than expected based on the initial falloff. Barriers of modest energy ($0.0\,\mathrm{eV} < \Delta V \leq 0.3\,\mathrm{eV}$ for holes, $-0.3\,\mathrm{eV} \leq \Delta V < 0.0\,\mathrm{eV}$ for electrons), produce a current vs. concentration curve that is a mirror image of the trap system.

Beyond a critical value ($|\Delta V| > 0.38\,\mathrm{eV}$ for the systems studied), the mirror symmetry between trap and barrier current vs. concentration curves is broken. At this point, traps are significantly deep that carriers can not escape. For example, without Coulomb interactions, to have a escape probability of 1.0 % from a trap that is 0.3 eV deep, an electric field of $0.391\,\mathrm{eV\ C^{-1}\ nm^{-1}}$ is needed. That is 391 V over 1000 nm; the device would decompose! All carriers in the OFET will be trapped and no current will flow. The Coulomb interactions between carriers will drastically lower the required voltage to free the carriers. For example, the repulsive energy from *another* like-charged carrier just 1 nm away is 0.411 eV. Ignoring the mean field energy from all other carriers, this means an extra 0.206 eV of energy is released when the carrier escapes the trap. This reduces the electric field needed (for a 1.0 % probability) to $0.185\,\mathrm{eV\ C^{-1}\ nm^{-1}}$. Although 185 V over 1000 nm is very high, this calculation only includes the effect of a single extra like-charged charge. The addition of more charges will lower the voltage to a reasonable level. In this way, Coulomb interactions

are significant for the de-trapping process in OFETs.

This is not the case for barriers, however. For a barrier of 0.3 eV height, the above probabilities and voltage also apply. However, the process is transport to the barrier site, rather than escape from a trap site. This difference is crucial in the behavior of the OFET. While deep traps will reduce the current to zero, making a dead device, current can still flow in the presence of high barriers, when the barrier concentration is $< 25\,\%$.

There will still be an IV curve with high barriers. However, negative differential resistance will emerge in the saturation region. The mechanism is as follows. Carriers, which can not easily transport to the barrier molecule, will be blocked from moving towards the drain electrode. This is because, in the saturation region, the probability to move towards the source is very low compared to other options. The most likely case is for the carrier to remain in place. For example, consider a system with no barrier, and an electric field of $0.1\,\mathrm{V\,nm^{-1}}$. Ignoring Coulomb interactions, the probabilities are $P_{right}, P_{up}, P_{down} = 8.325\,\%$, $P_{left} = 0.174\,\%$, $P_{stay} = 74.85\,\%$. If there is a high barrier in front of the carrier, then $P_{right} \approx 0\,\%$. This raises the probability of the carrier staying in place to $P_{stay} = 83.18\,\%$. Coulomb interactions will reduce the probability for the carrier to remain in place in both cases. However, the probability for the carrier to stay put will remain higher when a barrier is present, because $P_{right} \approx 0\,\%$.

When the voltage is lower (before the saturation region), the probability to move backwards is increased. This has the effect of reducing the probability of the carrier remaining in place. For example, at $0.01\,\mathrm{V\,nm^{-1}}$, the probability to move backwards increases to $5.654\,\%$. The net effect is that carriers can move around the barrier and navigate to the drain electrode (increasing the measured current).

However, in the saturation region, increasing the voltage increasingly makes the probability for carriers to remain in place larger. The carrier become "stuck" behind the barriers. Since they can not make it to the drain electrode, the current decreases with increasing voltage. This is the so called negative differential resistance. The same phenomenon will arise when sites are defective (i.e. carriers can not transport to the site).

Finally, Coulomb interactions can lead to a lessening of the negative differential resistance. That is, the slope of the IV curve in the saturation region will have a smaller

magnitude with Coulomb interactions. The probability of carriers that are "stuck" behind defects to stay put remains high in this case. Actually, the probability is higher than the case without Coulomb interactions. This is because the carrier must move against the field of other carriers to escape. However, the probability of other carriers to move near the blocked carrier is significantly reduced. For example, the probability goes from 8.325 % (without Coulomb interactions) to 0.139 % with Coulomb interactions. The carrier is much more likely to move around the blocked carrier. This increases the chance that this carrier will make it to the drain electrode, increasing the current, and reducing the negative differential resistance.

5.0 MONOLAYER OPVS : DOMAIN SIZE EFFECTS IN IDEAL AND ISOTROPIC SYSTEMS

The text in this chapter has been adapted from Gagorik et al., *Monte Carlo Simulations of Charge Transport in 2D Organic Photovoltaics*.[57]

5.1 INTRODUCTION

The effect of morphology on charge transport in organic photovoltaics is assessed using Monte Carlo. In isotopic two-phase morphologies, increasing the domain size from 6.3 nm to 18.3 nm improves the fill factor by 11.6 %, a result of decreased tortuosity and relaxation of Coulombic barriers. Additionally, when small aggregates of electron acceptors are interdispersed into the electron donor phase, charged defects form in the system, reducing fill factors by 23.3 % on average, compared to systems without aggregates. In contrast, systems with idealized connectivity show a 3.31 % decrease in fill factor when domain size was increased from 4 nm to 64 nm. We attribute this to a decreased rate of exciton separation at donor-acceptor interfaces. Finally, we notice that the presence of Coulomb interactions increases device performance as devices become smaller. The results suggest that for commonly found isotropic morphologies, the Coulomb interactions between charge carriers dominates exciton separation effects.

In this section, we have applied the Monte Carlo hopping code to charge transport in a two-dimensional model OPV, by calculating current-voltage curves (IV curves). We assess idealized charge transport in different morphologies by extracting quantities from the IV curves such as open circuit voltage ($V_{oc}; I(V_{oc}) = 0$), short circuit current ($I_{sc} = I(0)$),

theoretical power ($P_{th} = I_{sc} \times V_{oc}$), voltage at maximum power ($V_{mp}; d(IV)/dV(V_{mp}) = 0$), current at maximum power ($I_{mp} = I(V_{mp})$), maximum power ($P_{mp} = I_{mp} \times V_{mp}$), and fill factor ($FF = P_{mp}/P_{th}$). The fill factor is an important device parameter for optimizing performance. Solar cells ultimately deliver power to a load, which has a resistance. When the device is connected in series to the load, the current through the load has to equal the current through the solar cell. We can find how much power the solar cell delivers by intersecting the load line (Ohm's Law for the load) with the IV curve of the solar cell. The maximum power in a real device is quantified by the fill factor. In the IV Curves in Figures 5.2 and 5.3, the fill factor is geometrically the ratio of the areas of the rectangles, but more importantly is the percent of the maximum theoretical power obtainable from the actual device.

The focus was on the effect of morphology alone on device performance; as such, many assumed device properties were idealized and are unattainable in real materials. For example, we have assumed that exciton injection, exciton recombination, and electrode collection probabilities are constant, independent of material. In real devices, phases will differ in their ability to absorb light, have different dielectric constants, and have different electrode work function to frontier molecular orbital energy matching. Additionally, we have assumed that charge transport in n-type and p-type materials is equal and efficient, and that v_{oc} is maximized. In this work, we have not yet tested the effect of exciton injection rate. Instead, we have assumed the averaged AM1.5 solar spectrum[80] used for device verification.

5.2 SYSTEMS

Three classes of morphologies were examined: isotropic two-phase, band two-phase, and single-phase. In the two-phase morphologies (isotropic and band), the highest occupied molecular orbital (HOMO) of the electron donor (hole conducting) sites was chosen to be 0.5 eV higher in energy relative to the HOMO of electron acceptor (electron conducting) sites. Likewise, the lowest unoccupied molecular orbital (LUMO) of the electron acceptor (electron conducting) sites was chosen to be 0.5 eV lower in energy relative to the electron

donor (hole conducting) sites.

The isotropic two-phase morphologies simulated are shown in Figure 5.1A. Morphologies were simulated with the goal of achieving a randomly separated system two-phase with a readily definable domain size $\langle L_P \rangle$. Most simulated two-phase morphologies require costly numerical calculations of differential equations, such as the Cahn-Hilliard[55,81] approach. Other methods, such as Ising lattice relaxation, can introduce high-frequency noise and difficulty in targeting a specific domain size.[54] To avoid these issues while maintaining the desired morphological characteristics, a macroscopic, algorithmic approach was developed to produce realistic morphologies in a cheap, efficient and repeatable way.

The simulation of two-phase morphologies started with $257^2\,\mathrm{nm}^2$ "canvas" arrays of uniformly random numbers between 0 and 1. These canvasses were then convoluted (locally averaged) with isotropic 2D Gaussian kernels $e^{-\frac{i^2}{2\sigma_l^2}-\frac{j^2}{2\sigma_l^2}}$ to produce the set of four correlated arrays with $\sigma_l = 3, 5, 7, 9$. These real-valued sets were then converted into binary maps by a thresholding rule, with zeros corresponding to D (donor) and ones to A (acceptor) phases. The average domain size in such binary maps was set by the value σ_l. To produce morphologies with equal volume fractions of both phases ($\frac{\phi_1}{\phi_2} = 0.5$), thresholding was performed at the 50 % level of the cumulated probability of the normalized distribution of intensities. In $\frac{\phi_1}{\phi_2} = 0.5$ morphologies, the domain size $\langle L_P \rangle$is twice the size of σ_l. A pictorial illustration of this process can be found in the supporting information.

Furthermore, to introduce simulated small-size aggregates, some pixels of one phase were selected at random to be changed to the opposite phase, producing a "peppered" morphology. This morphology simulates small islets of PCBM trapped in a network of P3HT, postulated by others to be present in real devices.

5.3 RESULTS

Figure 5.2 shows IV curves for isotropic two-phase morphologies, and Figure 5.3 shows IV curves when 2.5 % pepper is added. $\langle L_P \rangle$for these morphologies were $\approx$ 6, 10, 14 and 18 nm. These systems are named with the shorthand LP6, LP10, LP14, LP18. With small-sized

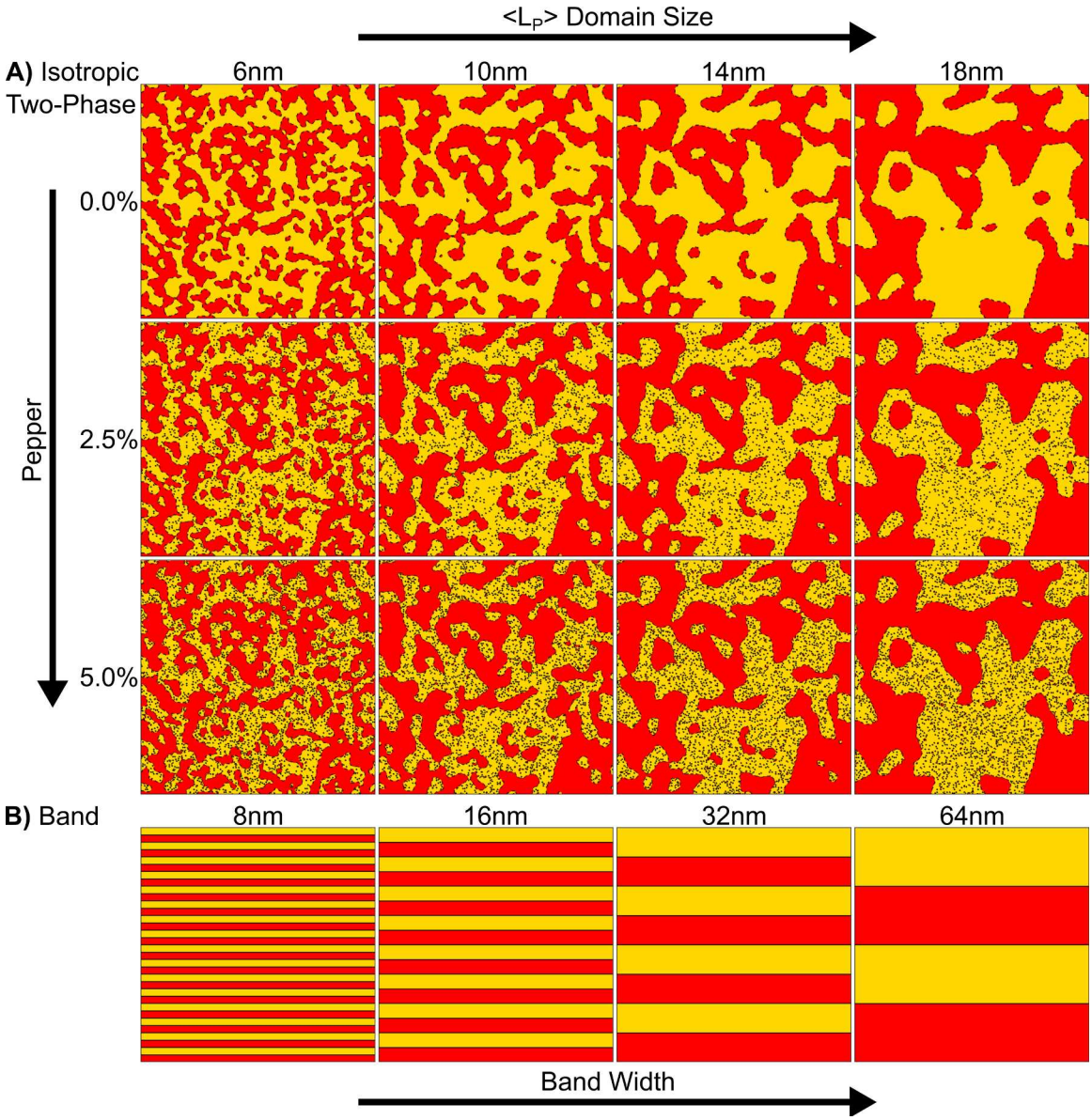


Figure 5.1: 2D morphologies.

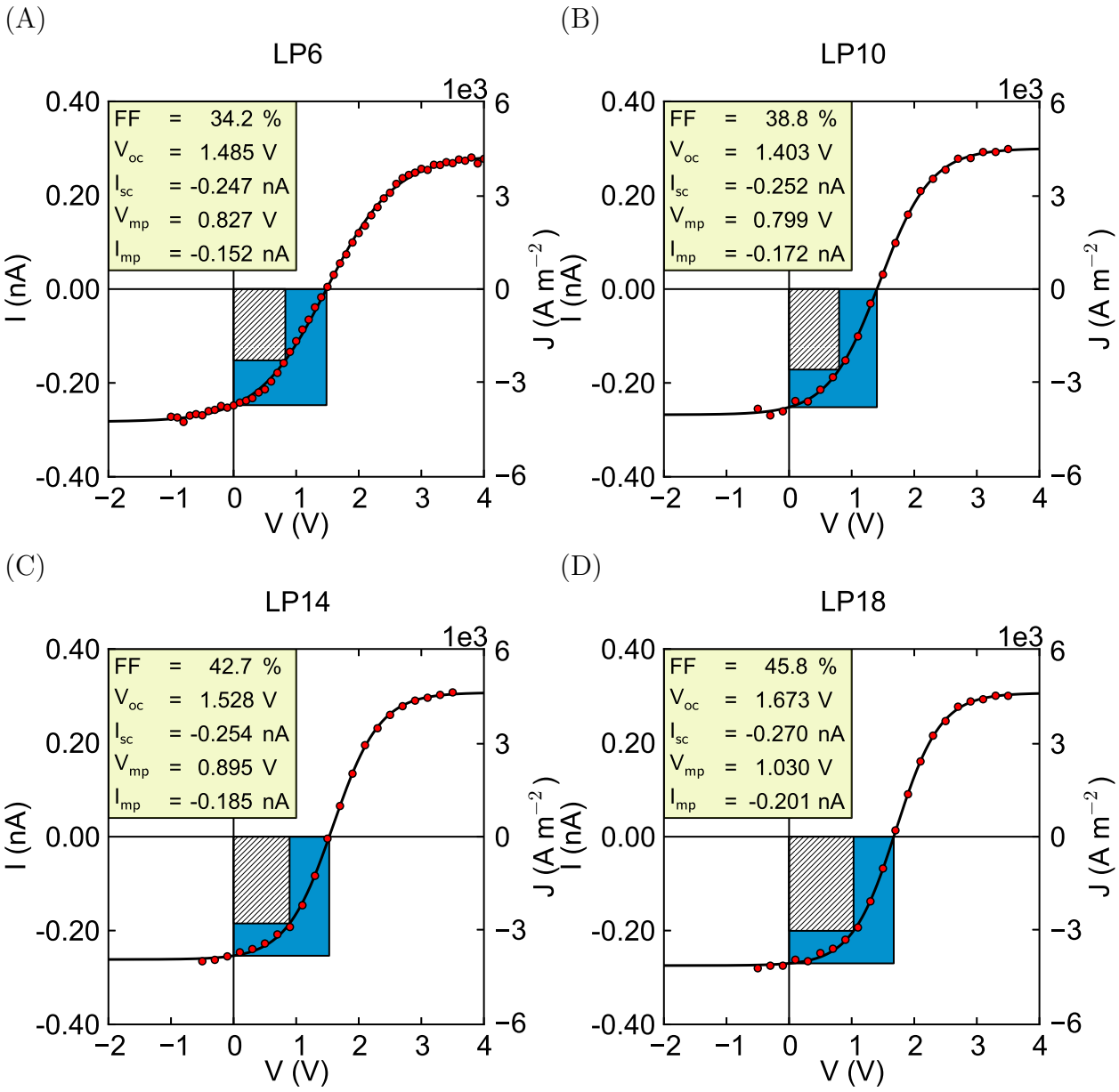


Figure 5.2: IV curves for isotropic 2D morphologies.

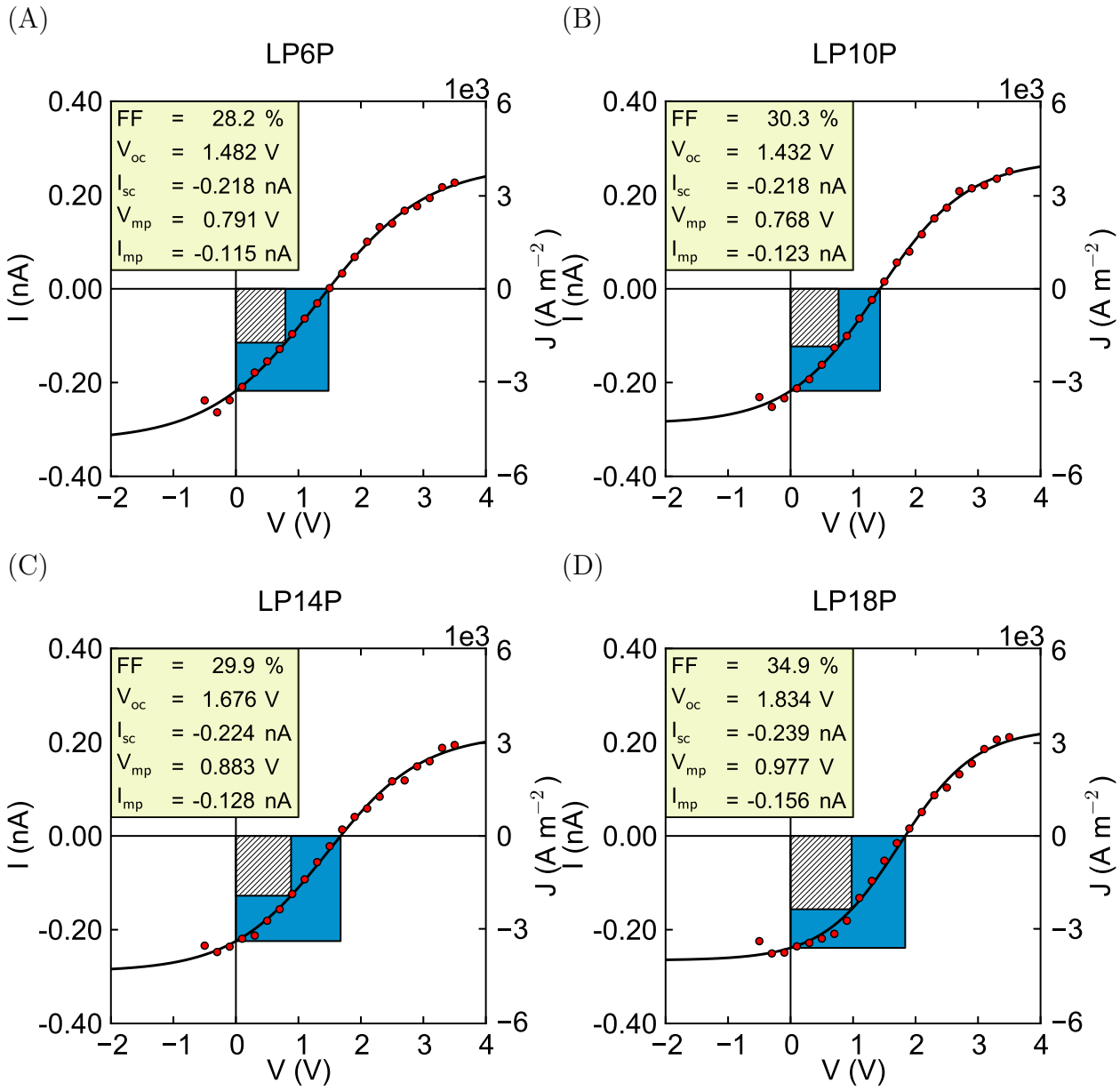


Figure 5.3: IV curves for isotropic 2D morphologies with pepper.

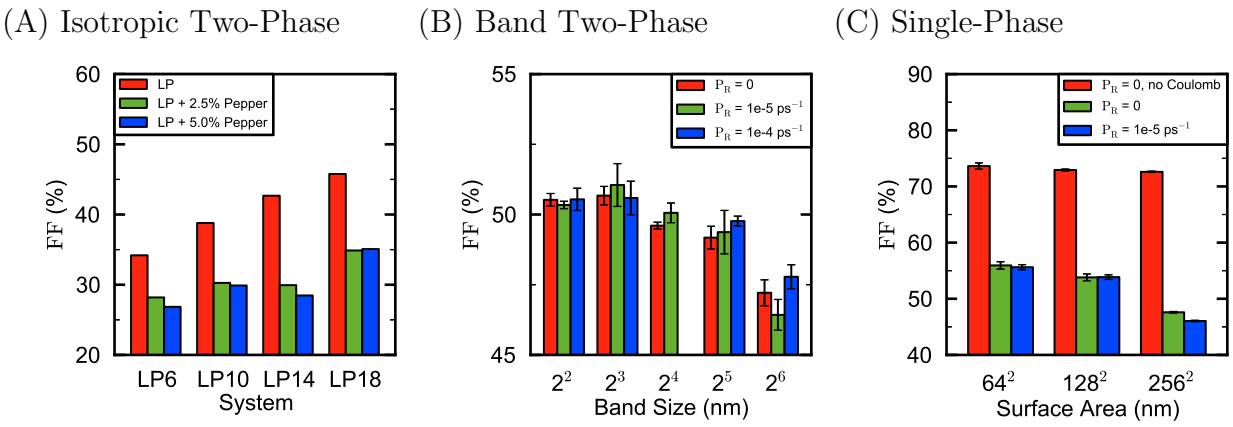


Figure 5.4: Fill factors for 2D morphologies.

acceptor aggregates added, (colloquially referred to as "pepper"), a P is appended to the name. Without pepper, the calculated fill factors were $FF \approx 34$, 38, 42 and 45 %. With pepper, the calculated fill factors were $FF \approx 28$, 30, 30 and 34 %, about a 20 % decrease. Additional increases in small-sized acceptor aggregates reduces the magnitude of current, but does not have a large effect on fill factor. Figure 5.4A shows calculated fill factors in order of increasing $\langle L_P \rangle$. Other extracted IV curve parameters can be found in the supporting information.

Simulated IV Curves indicate that larger $\langle L_P \rangle$, in both the peppered and non-peppered isotropic two-phase morphologies, yields a larger fill factor. Excitons must diffuse to interfaces in order to separate into free carriers, and the exciton diffusion length is on the order of ∼10 nm. Logic would dictate that interfaces should be within one to two exciton diffusion lengths, so it it counter-intuitive that morphologies with larger domain sizes performed better than those with smaller domain size. Despite the fact that recombination was not allowed to occur in these systems, the trend observed would not change because, as will be shown for the band and single-phase systems, a recombination rate $P_R \leq 10 \times 10^{-5}\,\mathrm{ps}^{-1}$ does not have a large effect on calculated fill factor in this model. This idea was tested for the two-phase isotropic system with $\langle L_P \rangle \approx 18$ nm, where the fill factor changed from 45.8 % to 45.2 % when using a recombination rate of $10 \times 10^{-5}\,\mathrm{ps}^{-1}$.

The fill factor increases with larger domain size because more carriers are able to diffuse to the electrodes at low bias voltages. In Figure 5.2, we can identify the increase in fill factor as increase of the area of the gray rectangle (maximum power) relative to the area of the blue rectangle (theoretical power).

After excitons dissociate at an interface, charge carriers still have to take a tortuous rout to their respective electrodes. Although smaller domain size features may facilitate exciton dissociation, narrower pathways through each phase in such morphologies pose problems for current extraction. First, there may not be a complete path to an electrode for a carrier to take in tortuous networks. Second, narrow pathways through small domain-size morphologies inevitably force carriers to remain within the Coulomb influence of other charges, complicating the energy landscape. Carriers will not move easily in narrow pathways already filled with other carriers of the same charge, and Coulomb traps readily form at interfaces between opposite charges. Consider, for example, that in a dielectric environment of $\epsilon = 2$-4, the thermal energy kT (0.026 eV at 300 K), offsets the Coulomb interaction between two charges at a distance of 14-28 nm. Therefore, carriers traveling in phase domains with widths ≤ 30 nm are strongly affected by the Coulomb force. These problems are mitigated as charge carrier pathways widen, allowing more carriers to reach the electrodes. Others have noticed that some degree of phase separation is desired, particularly in bulk heterojunction solar cells.[38,82] Additionally, it has been found that decreased miscibility of the donor-acceptor phases leads to improved charge collection efficiency.[83]

In the peppered systems, the situation becomes worse, although a similar trend in fill factor with $\langle L_P \rangle$is observed. The small acceptor aggregates act as electron traps in the hole conducting phase. Trapped electrons then serve as charged defects for holes,[16] reducing the total current. For example, i_{mp} is on average 26 % smaller in peppered simulations. As a result, fill factors are on average 23.3 % smaller with pepper. We note that the size of the pepper is in correspondence with the size of a single PCBM molecule, on the order of $\sim 1^3$ nm^3. In a previous work, we studied the effects of homogeneous vs. heterogeneous traps in OFETs.[60] That work leads us to believe that incorporating larger pepper would either lead to trapping islands or increased connectivity with very large pepper. Additionally, we would expect percolation issues when the phases deviate largely from a 40:60 ratio. We have

briefly probed the effect of the size of pepper, by simulating the LP6 and LP18 systems with 4-site sized ($2^2\,\mathrm{nm}^2$) pepper at the same concentration as 1-site sized pepper. We found that transport was still inhibited in the LP6 system, which showed little change in fill factor (28 %). In contrast, we found that larger sized pepper in the LP18 system lead to an increased fill factor (45 % vs. 34 % for small pepper), likely because trapped negative charge is averaged over the larger defect. More work is needed to elucidate the effect of the larger pepper.

For band morphologies, alternating stripes (bands) of electron conducting and hole conducting material run from electrode to electrode. The width of the bands was varied (4, 8, 16, 32 and 64 nm), corresponding to the $\langle L_P \rangle$in the isotopic two-phase systems. In this way, banded systems remove the complication of tortuous pathways to the electrodes, while simultaneously testing varying conduction pathway widths. Band morphologies are shown in Figure 5.1. The effect of exciton recombination rate was tested by changing the recombination rate from $10 \times 10^{-5}\,\mathrm{ps}^{-1}$ to $0\,\mathrm{ps}^{-1}$ for each system. The fill factors for band systems with recombination were $FF \approx 50.5$, 50.7, 49.6, 49.2 and 47.2 %. Without recombination, the fill factors were $FF \approx 50.3$, 51.0, 50.1, 50.2 and 46.8 %, respectively. The fill factors show little difference with recombination. Figure 5.4B shows calculated fill factors in order of increasing $\langle L_{Ps} \rangle$. Other extracted IV curve parameters can be found in the supporting information.

In contrast to isotopic two-phase systems, we find that the device performance does not increase with increasing domain size, but instead reaches a maximum when channel width is 8 nm. At a channel width of 4 nm, the fill factor is only 1.4 % smaller than the best performing device. There is slightly increased likelihood to encounter Coulomb barriers and traps in the 4 nm case. Otherwise, as the channel width increases, the fill factor decreases. This occurs because excitons have to traverse too large a distance to reach an interface. The results in the band morphologies suggest that tortuous networks and connectivity to electrodes are playing a larger role in the isotropic two-phase systems than the size of conduction pathways within each phase.

In the single-phase morphology, all hopping sites have the same HOMO energy assigned, and likewise for the LUMO. In these systems, the only source of energetic disorder arises from

Coulomb interactions. In a real system, other sources of disorder, such as configurational variation of molecular geometries or thermal fluctuations, would perturb the molecular orbital energies. Additionally, with no phase interfaces, separation of excitons into free charges is unlikely in this system because the energetic cost to overcome the exciton binding energy is not easily met.

The surface area of the device was varied in single-phase systems ($64^2\,\mathrm{nm}^2$ – $512^2\,\mathrm{nm}^2$). Varying the surface area has an indirect effect on the exciton injection rate, because larger areas will absorb more photons per step. The target injection rate for a $256^2\,\mathrm{nm}^2$ system was $1 \times 10^{-3}\,\mathrm{ps}^{-1}$, which is an overestimate compared to the AM1.5 solar spectrum. This rate was then scaled linearly with surface area for systems larger or smaller than $256^2\,\mathrm{nm}^2$. For example, the $512^2\,\mathrm{nm}^2$ system had a $(1 \times 10^{-3}\,\mathrm{ps}^{-1})(512^2\,\mathrm{nm}^2)/(256^2\,\mathrm{nm}^2) = 4 \times 10^{-3}\,\mathrm{ps}^{-1}$ injection rate. Second, the effect of exciton recombination rate was tested by changing the recombination rate from 10×10^{-5} to $0\,\mathrm{ps}^{-1}$. Lastly, the effect of turning off Coulomb interactions was tested. Figure 5.4C shows calculated fill factors for single-phase morphologies vs grid surface area. The IV curves and extracted IV curve parameters can be found in the appendix A.

The results for the single-phase systems again showed little difference when considering recombination rate, and this difference is even less pronounced for surface areas $\leq 256^2\,\mathrm{nm}^2$. It was expected that the current *density* would not depend upon surface area, however, when Coulomb interactions are included, we found that fill factors are larger in smaller devices. The fill factor increases because i_{sc} grows, while i_{mp} shrinks with surface area. v_{mp} and v_{oc} remain essentially constant. We have noticed that removing the Coulomb interactions between carriers removes the scaling of fill factor with surface area in the single-phase systems. As carriers become confined in smaller systems, a repulsive Coulomb environment may be leading to the observed changes in fill factor with surface area. When the recombination rate was set to zero, there were minor changes in extracted values. For example, $|\Delta FF| \leq 1.8\,\%$, and $\Delta FF = 0.01 \pm 1.0\,\%$. The largest change occurs for the large channel width of 64 nm, where recombination leads to a reduction in i_{mp} and fill factor. It is also interesting to note that recombination may be helping in the smaller channels, as a slight reduction in the number of carriers makes the energetic landscape easier to traverse.

5.4 CONCLUSION

In conclusion, a Monte Carlo hopping model was employed to simulate charge transport in three classes of morphologies. Increasing the domain size, $\langle L_P \rangle$, in isotropic two-phase morphologies increased device performance. In contrast, increasing domain size in two-phase band systems decreased device performance. We attributed increasing device performance in isotropic two-phase systems to decreased tortuosity of charge pathways and increased connectivity to the electrode surfaces. We attributed decreasing device performance with increasing domain size in band systems to a lack of interfaces for exciton dissociation. When electron accepting aggregates are dispersed randomly throughout the electron donor, we observed severe decreases in device performance in the isotropic two-phase systems. We ascribed this to the formation of negatively charged defects in the hole conducting phase.

Finally, we observed that Coulomb interactions are the origin of increased efficiency in the small devices due to a field dependent mobility. For all systems, Coulomb interactions were found to play a crucial role in device behavior, and suggest that Coulomb interactions are adequately accounted for in simulation models.

6.0 MULTILAYER OPVS: CHARGE DELOCALIZATION AND MORPHOLOGY

The text in this chapter has been adapted from Gagorik et al., *Effects of Delocalized Charges on Organic Photovoltaics: Nanoscale 3D Monte Carlo Simulations*.[61]

6.1 INTRODUCTION

Monte Carlo simulations of charge transport in organic solar cells were performed for ideal and isotropic bulk heterojunction morphologies while altering the delocalization length of charge carriers. Previous device simulations have either treated carriers as point charges or with a highly delocalized mean-field treatment. Our new model of charge delocalization leads to weakening of Coulomb interactions and improved device performance at moderate delocalization lengths, relative to point charges and more realistic predicted current and fill factors. We find that charge delocalization leads to significantly increased acceptance probability when escaping interface traps. In isotopic two-phase morphologies, increasing the domain sizes leads to decreases in device efficiencies. We previously showed that tortuous pathways in systems with small domain sizes can decrease device performance in thin film systems. However, the diminishing effects of Coulomb interactions with delocalization, the increased number of pathways in bulk systems, and efficient separations of excitons by small domains reverses this effect in 3D. We emphasize that delocalization, which has largely been ignored in the past, is an important parameter to consider and optimize when choosing materials for organic solar cells.

The understanding of charge transport in disordered organic semiconductors is far from

complete, and mechanisms are very different from inorganic materials. In these materials, charge transport is typically described by localized carrier hopping through a manifold of energy states. The charge carriers are localized to molecular sites through disorder of the morphology and electron-phonon coupling. In contrast, highly delocalized carriers move through the energy bands of an inorganic material.

However, delocalization can still play a role in organic materials. Charge delocalization is well known from electronic spectroscopy (i.e. π-stacked dimer excitations). It has been suggested as a mechanism for efficient long range charge separation in organic semiconductors.[39,40] For example, well ordered, semi-crystalline domains likely lead to charge delocalization and improved device performance in organic photovoltaics (OPVs). Devices with well ordered morphologies have been reported with very large fill factors ($76 - 80\,\%$) and power conversion efficiencies of up to $8.7\,\%$.[41] However, the complete picture is not clear. While charge transport is enhanced by crystalline domains, exciton transport may be inhibited in well ordered systems.[42] This additional level of complexity is due to the dominance of exciton diffusion occurring by an inter-chain mechanism. In fact, exciton delocalization in regioregular P3HT is estimated to be as low as 1-2 nm, and therefore may not play a major role in charge separation.[43] Studying the interplay of delocalization with electrostatics may help elucidate these effects.

The combination of electrostatics and delocalization can play a major role in the charge separation process. For example, along with hot charge transfer (CT) states, calculations by Tamura and Burghardt show that charge delocalization can lead to ultra fast (< 100 fs) charge separation via the lowering of Coulomb barriers.[44] While delocalization can be treated very accurately in quantum calculations,[43–45] the effect is generally ignored in meso-scale simulations of charge transport. Here, highly accurate electrostatics are achieved by largely ignoring the effects of delocalization and treating carriers as point charges.

To address these questions, we have incorporated delocalization into our existing Monte Carlo model[16,57,59,60] of charge transport in OPVs. We have used the Monte Carlo model to examine a series of four isotropic two-phase morphologies with increasing domain size, three ideal morphologies based on minimum curvature surfaces, and morphologies based on bands (see Figure 6.4). The effects of delocalization have been examined by varying the

delocalization length of the carriers, from 0.75 to 1.50 nm. We have also simulated point charges and non-interacting charges.

We have used the Monte Carlo simulation to predict current-voltage curves (IV Curves). From the IV curves we have computed the efficiency of each system by calculating fill factors. The fill factor (FF) is the ratio of the maximum power that can be extracted from the cell to the theoretical power, as shown in Equation 6.1:

$$FF = \frac{v_{mp} \times i_{mp}}{v_{oc} \times i_{sc}} \times 100\,\% \tag{6.1}$$

In Equation 6.1, i_{mp} is the current at maximum power, v_{mp} is the voltage at maximum power, i_{sc} is the short circuit current, and v_{oc} is the open circuit voltage. The fill factor ranges between 0 and 100 %, and can be viewed as the ratio of the areas of two rectangles - a smaller one formed by i_{mp} and v_{mp}, and a large one formed by i_{sc} and v_{oc}. The closer the fill factor is to 100 %, the more square (ideal) the IV curve, and the more efficient the performance of the solar cell. In a real device, the solar cell is connected to an external load, which has a resistance, R. Only when R is the ratio v_{mp}/i_{mp} can the maximum power be extracted from the device. The short circuit current i_{sc} corresponds to $R = 0$, while the open circuit voltage corresponds to $R = \infty$.

6.2 MODEL

The model consists of a lattice of sites representing molecules ~1 nm^3 in size, approximately the same size as an isolated fullerene. Each site is considered a neutral molecule. In this work, we have used a lattice of $192 \times 192 \times 32$ nm^3. Charge carriers, such as holes and electrons, can occupy the sites. Holes are to be viewed as an electron missing from the highest occupied molecular orbital (HOMO) of a molecule. Electrons are to be viewed as an excess electron in the lowest unoccupied molecular orbital (LUMO). When an electron and hole occupy the same site, they are an exciton. Excitons can recombine to form unoccupied sites.

The simulation proceeds by injecting excitons randomly in the system. We have assumed that the system is ideal in that all sites have the same probability to form an exciton, regardless of thickness. In the future, we will relax these assumptions. For now, we have used an injection probability estimated from the AM1.5 solar spectrum, and is proportional to the surface area of the device. For a $256 \times 256\,\mathrm{nm}^2$ area, we have estimated an injection probability of $10^{-3}\,\mathrm{ps}^{-1}$. This injection probability is scaled by a factor of $192^2/256^2 = 0.5625$ for an area of $192 \times 192\,\mathrm{nm}^3$. Recombination is also allowed to occur at a fixed probability for excitons in the system. Unlike injections, a fixed recombination probability does not correspond to a fixed rate because excitons are only considered when holes and electrons occupy the same site. A constant probability is a simplification; recombination in organic solar cells can happen by many mechanisms with different timescales.[84] For example, for a bimolecular recombination mechanism, the Langevin expression predicts recombination rate to depend upon the hole/electron mobilities, the electron/hole concentrations, and the intrinsic carrier concentration.[84] We have used a fixed recombination probability of $10^{-4}\,\mathrm{ps}^{-1}$, inspired by kinetic Monte Carlo simulations.[85–87]

Charges hop between adjacent sites using the Metropolis criterion. The probability to hop between sites is $H_{AB}e^{-\beta\Delta E}$ if $\Delta E \leq 0$. If $\Delta E > 0$, the probability to hop between sites is H_{AB}. H_{AB} is a coupling constant representing the approximate electronic overlap between molecular wavefunctions. We have used $H_{AB} = 1/3$, independent of hopping direction. H_{AB} is made exponentially smaller, $H_{AB} = 1/27$, for a two-site hop. Carriers are considered individually for trial moves.

Sites are assigned energies representing the HOMO and LUMO energies of the molecules, as shown in Figure 6.1. There are three main contributions: the donor/acceptor offset, the overall device potential between electrodes (including differences between metal workfunctions), and the Coulomb interaction between charges.

The sites are classified as donor or acceptor, according to the system morphology chosen, as discussed below. The donor/acceptor LUMO and HOMO are offset by a constant energy $\Delta E_{LL} = \Delta E_{HH} = 0.5\,\mathrm{eV}$. This value is slightly larger than the Coulomb interaction between two opposite charges spaced 1 nm apart (0.411 eV). It provides a driving force to separate carriers at the interface.

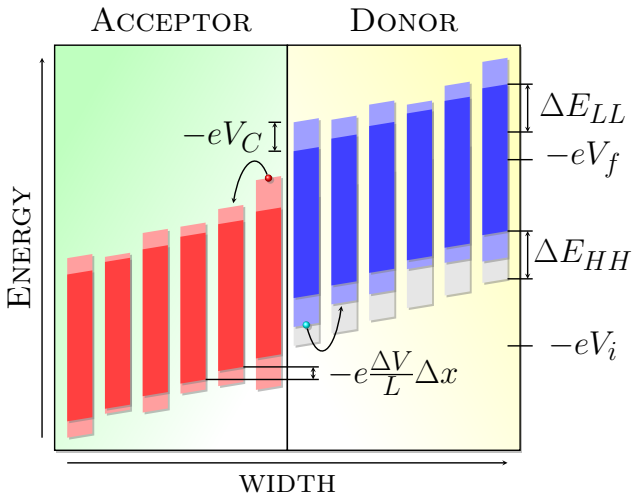


Figure 6.1: Energy model for solar cells.

Next, a linear electrostatic potential is applied between the electrodes, using the potential difference $\Delta V = V_f - V_i$ and width of the device, L. This potential includes the external and intrinsic bias. To construct an IV curve, this potential is varied from -2.0 to $2.0\,\mathrm{V}$.

Disorder is added to the system using the Coulomb interaction, as shown in Equation 6.2. In Equation 6.2, N is the number of charges (carriers or charged defects), q is the charge, r_{ij} is the distance between charges i and j, and $\epsilon = 3.5$ is the dielectric constant. For every Monte Carlo move, the Coulomb interaction must be calculated for the initial and final state - around $2 \times N \times N_{steps}$ times. Even with a cutoff of 50 nm and lookup table for inter-cell distances, the brute force evaluation of Equation 6.2 is the most time consuming step of the simulation, because it scales as N^2. Therefore, a parallel GPU implementation of the Coulomb sum was used to make the calculation feasible. With the GPU code, carrier numbers of 10 000 or more can be handled.

$$V_C(\mathbf{r}_j) = \sum_{i=0}^{N} \frac{q}{4\pi\epsilon\epsilon_0 |\mathbf{r}_{ij}|} \tag{6.2}$$

Equation 6.2 describes the interaction between point charges. To compute the potential due to a charge distribution, on uses Poisson's Equation, $\nabla^2 V = \rho/(\epsilon\epsilon_0)$, where ρ is the charge density. When using a charge density described by spherically symmetric Gaussians

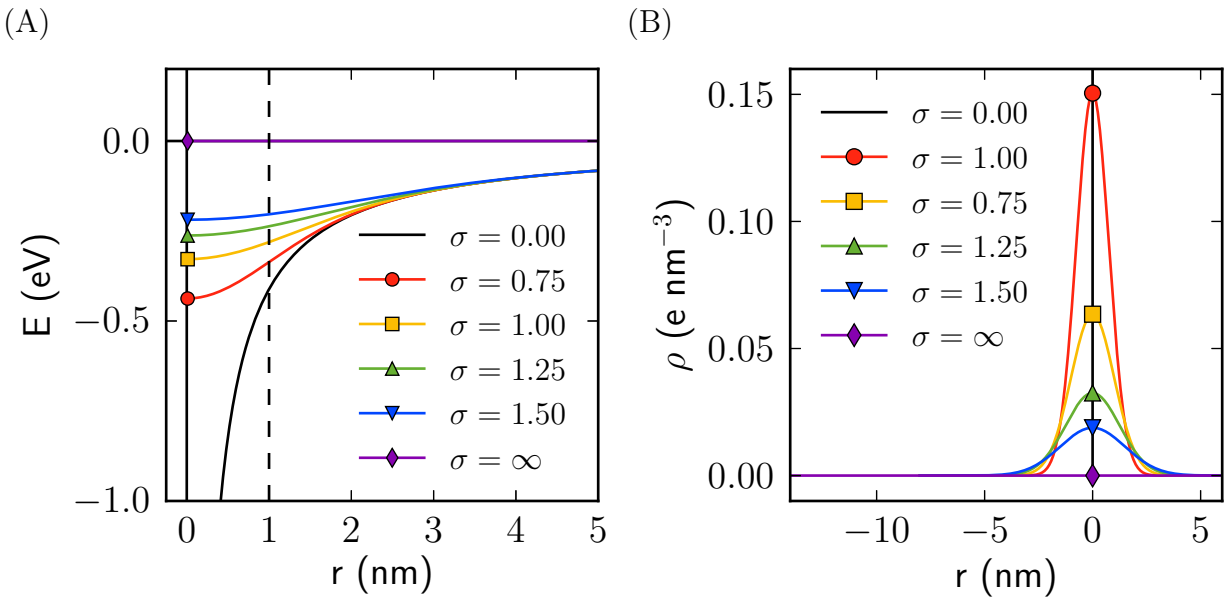


Figure 6.2: Coulomb potential with delocalized charges.

(Equation 6.3), an analytical solution can be found (Equation 6.4). The result is is simply the Coulomb law multiplied by the error function, with a parameter σ describing the delocalization length.

$$\rho(\mathbf{r}_j) = \sum_{i=0}^{N} \frac{q}{\sigma^3(2\pi)^{3/2}} e^{-\frac{|\mathbf{r}_{ij}|^2}{2\sigma^2}} \tag{6.3}$$

$$V_C^{\text{erf}}(|\mathbf{r}_j|) = \sum_{i=0}^{N} \frac{q}{4\pi\epsilon\epsilon_0|\mathbf{r}_{ij}|} \operatorname{erf}\left(\frac{|\mathbf{r}_{ij}|}{\sigma\sqrt{2}}\right) \tag{6.4}$$

The error function only modifies the Coulomb interaction at short range. For example, when $\sigma = 1.0\,\text{nm}$, $0.95 \leq V_C^{\text{erf}}/V_C \leq 1$ at $r \geq 2.0\,\text{nm}$. However, at short distances, the Coulomb potential is significantly diminished. For example, when $\sigma = 1.0\,\text{nm}$, $0.0 \leq V_C^{\text{erf}}/V_C \leq 0.68$ at $r \leq 1.0\,\text{nm}$. This effect can be seen in Figure 6.2A, where the potential at each value of σ converges to the Coulomb potential at large r (i.e. $\geq 3.0\,\text{nm}$). The charge density as a function of r is shown in Figure 6.2B.

Figure 6.3 shows energy and acceptance probability as a function of sigma. The system is a hole at the origin ($r = 0\,\text{nm}$), and an electron at 1 to 3 nm away. The energy change

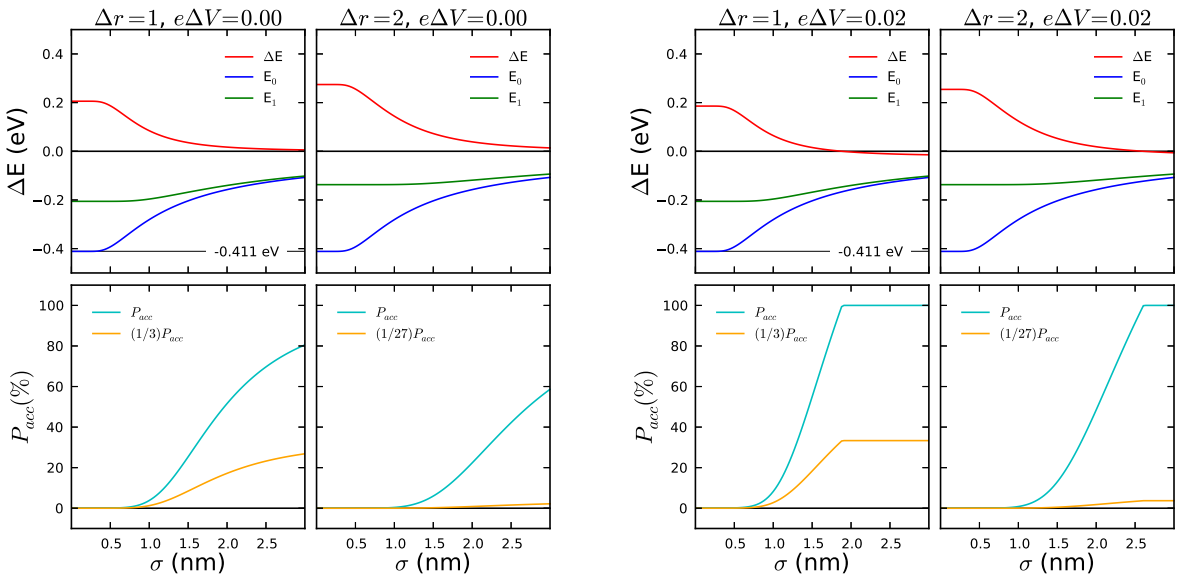


Figure 6.3: Probability vs sigma

for moving the electron from 1 to 2 nm, or from 1 to 3 nm, is shown in zero field and in $0.02\,\mathrm{V\ nm^{-1}} = 2\,\mathrm{V}/200\,\mathrm{nm}$ field. Table 6.1 lists numerical values associated with the graphs in Figure 6.3. The acceptance probability increases significantly with sigma.

6.3 SYSTEMS

Morphologies were constructed with a domain of $192 \times 192 \times 32\,\mathrm{nm}^3$, or 1 179 648 sites. Periodic boundary conditions were not used, to properly include finite size effects in real materials. All morphologies consisted of a donor and an acceptor phase whose frontier molecular orbital were offset by 0.5 eV, as described above. In addition to more realistic morphologies, idealized morphologies were also considered.

The ideal morphologies studied are shown in Figure 6.4. The first three ideal morphologies were constructed from equations of triply periodic minimal surfaces.[88] Triply periodic minimal surfaces are free of self intersections and contain perfect connectivity to the elec-

σ(nm)	r_{hole}(nm)	r_{elec0}(nm)	r_{elec1}(nm)	Δr(nm)	E_0(eV)	E_1(eV)	ΔE(eV)	$\beta\Delta E$	P_{acc}(%)	$H_{ab}P_{acc}$(%)
0.00	0	1	2	1	-0.411	-0.206	0.186	7.184	0.076	0.025
0.25	0	1	2	1	-0.411	-0.206	0.186	7.183	0.076	0.025
0.50	0	1	2	1	-0.393	-0.206	0.167	6.460	0.156	0.052
0.75	0	1	2	1	-0.336	-0.204	0.112	4.341	1.302	0.434
1.00	0	1	2	1	-0.281	-0.196	0.065	2.496	8.243	2.748
1.25	0	1	2	1	-0.237	-0.183	0.034	1.313	26.913	8.971
1.50	0	1	2	1	-0.204	-0.168	0.015	0.599	54.958	18.319
0.00	0	1	3	2	-0.411	-0.137	0.254	9.836	0.005	0.000
0.25	0	1	3	2	-0.411	-0.137	0.254	9.835	0.005	0.000
0.50	0	1	3	2	-0.393	-0.137	0.236	9.112	0.011	0.000
0.75	0	1	3	2	-0.336	-0.137	0.179	6.933	0.097	0.004
1.00	0	1	3	2	-0.281	-0.137	0.124	4.800	0.823	0.030
1.25	0	1	3	2	-0.237	-0.135	0.082	3.180	4.159	0.154
1.50	0	1	3	2	-0.204	-0.131	0.053	2.041	12.993	0.481

Table 6.1: Interaction energies vs sigma.

trodes. The minimal surfaces chosen, which are of the form $f(x, y, z) = 0$, were the gyroid (Eq. 6.5, Figure 6.4A), p-surface (Eq. 6.6, Figure 6.4B), and d-surface (Eq. 6.7, Figure 6.4C).

$$\cos(k_x x)\sin(k_y y) + \cos(k_y y)\sin(k_z z) + \cos(k_z z)\sin(k_x x) = 0 \tag{6.5}$$

$$\cos(k_x x) + \cos(k_y y) + \cos(k_z z) = 0 \tag{6.6}$$

$$\begin{aligned}\sin(k_x x)\sin(k_y y)\sin(k_z z) + \sin(k_x x)\cos(k_y y)\cos(k_z z) + \\ \cos(k_x x)\sin(k_y y)\cos(k_z z) + \cos(k_x x)\cos(k_y y)\sin(k_z z) = 0\end{aligned} \tag{6.7}$$

In Equations 6.5, 6.6, and 6.7, k is an angular wavenumber, defining the periodicity of the surface. We choose a k that corresponds to a wavelength $\lambda = 8\,\text{nm}$, as shown in Equation 6.8, where L is a dimension of the grid, and n is the number of periods along the dimension. Only isotropic systems ($k = k_x = k_y = k_z$) were studied. In isotropic systems, the minimum possible distance between surface boundaries is $\frac{\lambda}{2} = 4\,\text{nm}$.

$$\lambda = \frac{L}{n} = \frac{192\,\text{nm}}{8} = \frac{32\,\text{nm}}{4} = 8\,\text{nm} \tag{6.8}$$

$$k = \frac{2\pi}{\lambda} = \frac{2\pi}{8\,\text{nm}} = \frac{\pi}{4\,\text{nm}} \tag{6.9}$$

In practice, sites were assigned by evaluating $f(x, y, z)$ on the domain and thresholding. Donor sites were placed where $f(x, y, z) >= 0$, while acceptor sites were placed where

$f(x, y, z) < 0$. This assignment makes the zero-level contour-isosurface of $f(x, y, z)$ the boundary between donor and acceptor sites.

The last ideal morphology studied was the "band" morphology. The band morphology is a checkerboard pattern of alternating donor and acceptor sites in the yz-plane, extended along x-direction. Two band morphologies were used. The band4 morphology had a checker size of $4 \times 4\,\mathrm{nm}^2$. The band8 morphology had a checker size of $8 \times 8\,\mathrm{nm}^2$. Figure 6.4D shows a $32 \times 32 \times 32\,\mathrm{nm}^3$ slice of the band4 morphology.

More realistic isotropic morphologies, akin to typical organic bulk heterojunctions, are also shown in Figure 6.4. Isotropic morphologies were constructed using an algorithm discussed previously,[57] extended to three dimensions. In the algorithm, a grid of random noise is constructed in the domain. Then, this noise is convoluted with a three-dimensional Gaussian of desired width, σ. Finally, thresholding is applied to make a roughly 50/50 mixture of isotropic two-phased morphology. Four isotropic systems were studied, named iso3, iso4, iso6, and iso9. The value of σ used for the Gaussian kernels were 3, 4, 6 and 9 nm. These σ lead to a domain size of 5.1, 8.1, 9.57 and 13.97 nm and inter-facial area of 196.04, 123.31, 104.39 and $71.52\,\mathrm{m}^2\,\mathrm{cm}^{-3}$. The iso3 morphology domain size is comparable to the ideal morphologies minimum distance between phases of 4 nm, excluding the band8 morphology. The band8 morphology is more in line with the iso4 system.

6.4 DISCUSSION

Figure 6.5A shows computed fill factors for ideal morphologies. Aside from the band8 morphology, each ideal system has the same domain size of 4 nm. As the delocalization is increased from point charges ($\sigma = 0.00\,\mathrm{nm}$) to delocalized charges ($\sigma = \infty$), the average FF increases from 35 % to 58 %. The FF continues to increase, less dramatically, as the delocalization is increased to $\sigma = 1.50\,\mathrm{nm}$. Non interacting charges, which can be viewed as infinitely delocalized carriers in this model, yield devices with the best FF (70 %). While there is no clear trend in morphology with delocalized charges, the trend with point charges is gyroid < d-surface < p-surface < band4.

Figure 6.5B shows computed fill factors for isotropic two phase morphologies as a function of sigma. The iso4 and iso6 systems have very similar domain size (8.1 and 9.57 nm). Within error, their FF are very similar. As the delocalization of carriers increases, the device performance increases. The most dramatic increase occurs between point charges and $\sigma = 0.75\,\text{nm}$, where the fill factors increase from $< 40\,\%$ to $> 50\,\%$. However, the differences of FF for $\sigma = 0.75 - 1.50\,\text{nm}$ is much less pronounced. Only when non-interacting charges ($\sigma = \infty$) are used, does the FF increase significantly.

In a previous study, on simulations of 2D monolayer devices, we found that decreasing the domain size leads to poorer charge extraction in thin films.[57] We argued that while smaller domain sizes lead to efficient exciton separation, small tortuous pathways can hinder charge extraction. The close proximity of charges and tortuous morphology to the electrodes lead to a disordered energy landscape with a limited number of viable pathways to the electrodes. The trend appears to reverse in larger bulk 3D systems. That is, smaller domains show improved device performance at each value of the delocalization parameter σ. However, we note that the effects are relatively small and within error bars. More replicas need to be simulated and averaged to elucidate the effect. The interplay between tortuous pathways and efficient exciton separation, the separation wins. We propose that more pathways have opened up in the 3D environment than were present in the 2D thin film simulations.

Figure 6.6 shows snapshots of the Monte Carlo simulation for the iso3 morphology at v_{oc}. At v_{oc}, the potential between the electrodes balances the internal intrinsic bias. Therefore, there is very little incentive for excitons to separate and free carriers to travel to the electrodes. Recombination is the most likely fate of the carriers in this regime. However, the effects of charge delocalization are easily visualized at v_{oc}. When $\sigma = 0.75\,\text{nm}$ (6.6A, 6.6C), holes and electrons remain bound in interface traps. However, as the delocalization is increased to $\sigma = 1.50\,\text{nm}$ (6.6B, 6.6D), the number of bound interface traps decreases significantly. The energetics of escaping the interface trap are more favorable when the carriers are delocalized.

Consider, for example, an electron hole pair trapped at an interface. The exciton has separated, putting the electron in the LUMO of the acceptor, and the hole in the HOMO of the donor. The Coulomb attraction between the hole and electron ($\epsilon = 3.5$) separated by

1 nm is 0.411 eV, nearly 15.9 times larger than kT at 300 K (0.0259 eV). A potential difference of 1.0 V between electrodes separated by 256 nm only releases 0.003 90 eV of energy during the 1.0 nm hop out of this state. Given that the Coulomb attraction is only halved at 2.0 nm, this means the hole and electron will "separate" only a minuscule fraction of the time ($e^{-\beta\Delta E} = 0.0407\,\%$). Even if the carriers reach this state, it is extremely likely the charges fall back into the charge trap state, and possibly recombine. There must be other mechanisms to explain the efficient charge separation in organic materials.

The delocalization of charges can weaken the Coulomb interaction between trapped carriers. The interaction of spherically symmetric Gaussian charge distributions can be modeled using the standard Coulomb law and the error function (Eq. 6.4). Using this model, the probability to separate carriers, of only modest delocalization lengths (0.5-1.5 nm) can increase dramatically. For example, a delocalization length of 1.0 nm dampens the Coulomb interaction by $\mathrm{erf}(1.0/\sqrt{2}) = 68\,\%$ at an electron hole separation of 1.0 nm. The probability to escape the charge trap state increases by over a factor of 100 ($\Delta E = 0.0806\,\mathrm{eV}, P = 4.42\,\%$).

6.5 CONCLUSION

In summary, we have predicted that the charge delocalization leads to increased charge extraction in OPVs. We propose that the weakening of energetic barriers, particularly in Coulomb interface traps leads to the efficient separation of charges. Lastly, in isotropic two-phase morphologies, decreased domain size leads to increased device performance. Unlike thin films, where the Coulomb force dominates the energy landscape of separated charges, more pathways open up in 3D systems. Small domains lead to efficient exciton separation, while the increased number of pathways in 3D alleviate interface trap formation between separated carriers.

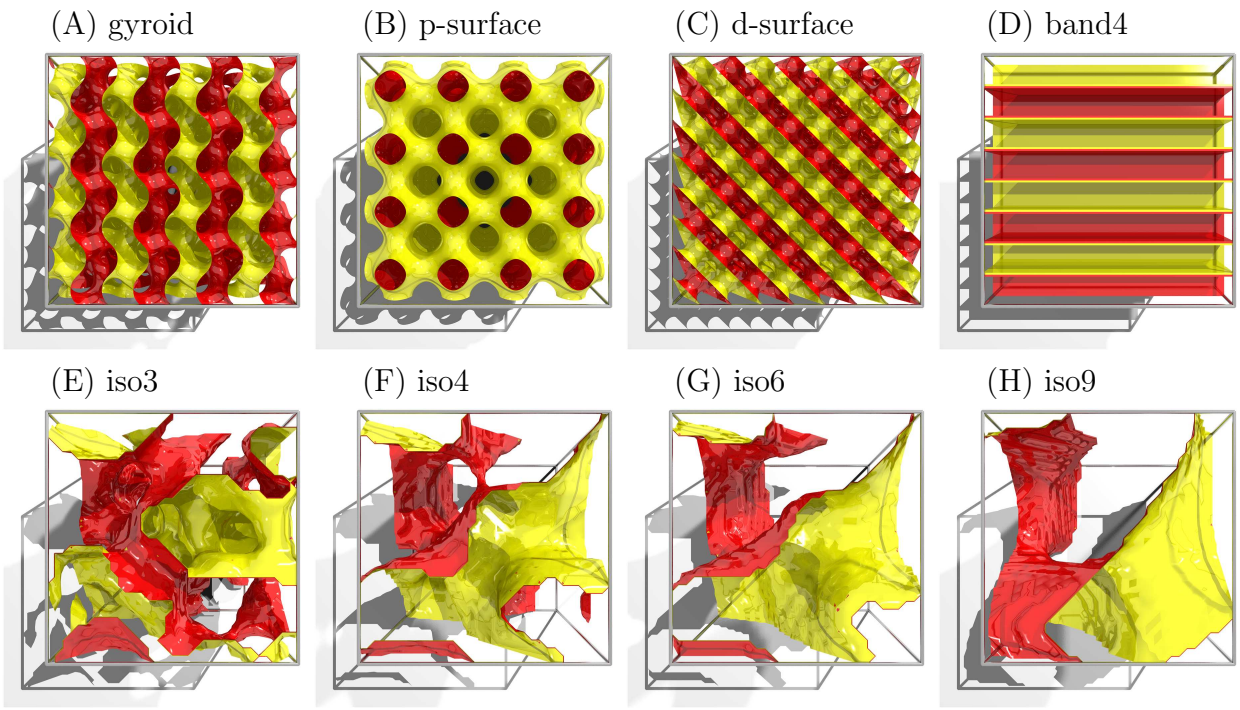


Figure 6.4: 32x32x32 slices of 3D surfaces.

(A) (B)

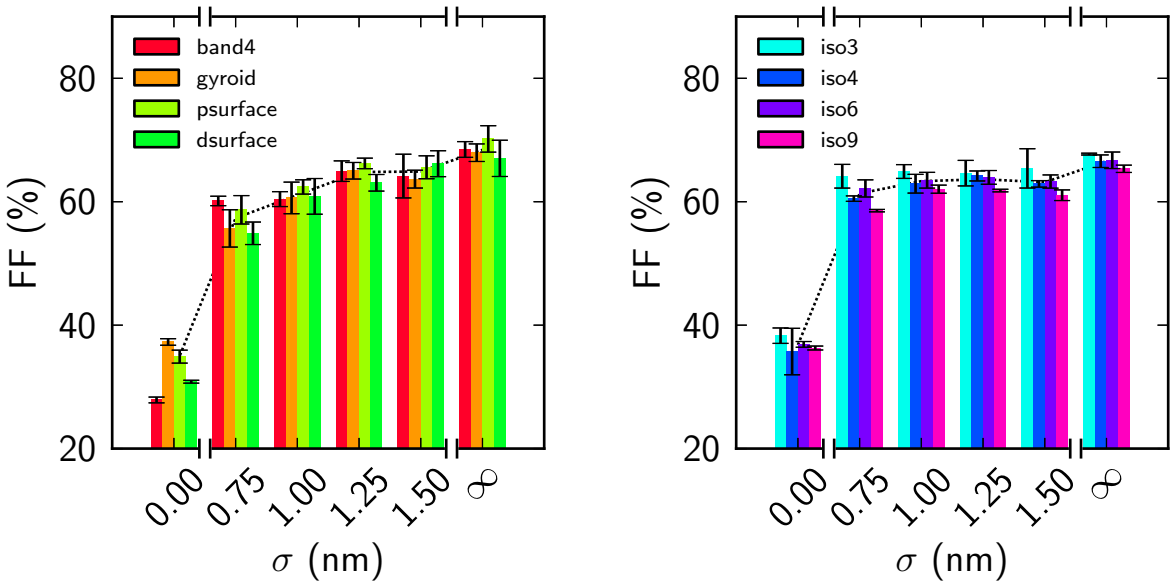


Figure 6.5: Fill factors for 3D simulations.

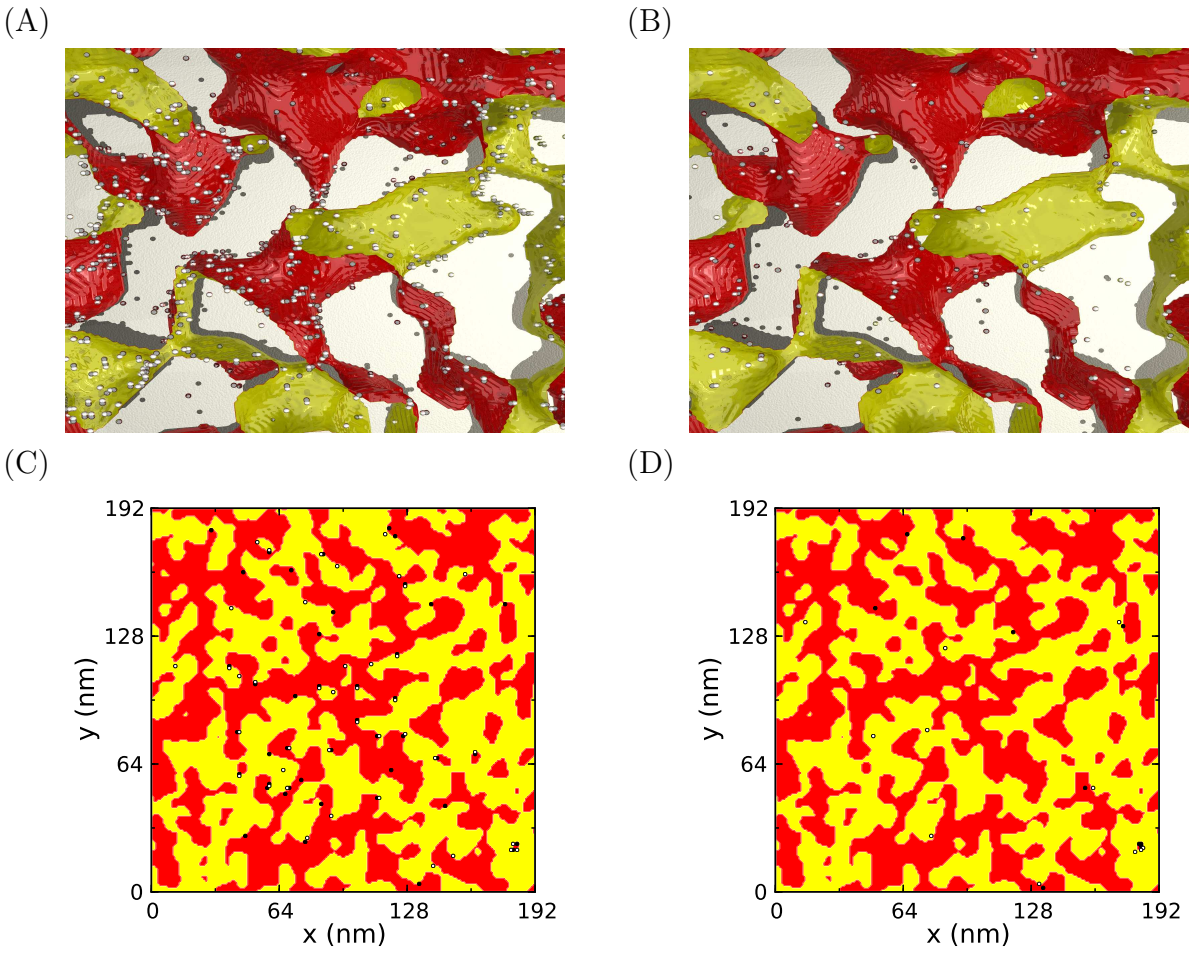


Figure 6.6: Snapshot of trajectory in isotropic morphology.

7.0 CONCLUSION

In this chapter, the findings of this thesis are summarized.

7.1 SUMMARY

7.1.1 Injection

The physical picture of charge injection in OFETs matters when attempting to understand charge transport in these devices. On the one hand, the "seeding" procedure represents the introduction of carriers at random locations into the semiconductor. In this case, the potential difference between the gate and source electrodes lead to the oxidation of the semiconductor, introducing carriers into the conduction channel. On the other hand, the "injection" procedure represents the effect the gate electrode has on the energy of frontier orbitals relative to the Fermi energy of the electrodes. As the orbital energy is brought in line with the electrode Fermi level, charge injection from the electrodes becomes an energetically favorable process. While both procedures are aspects of the same phenomenon, it is important to examine the effects of these processes in finite sized devices.

I explored charge transport in thin-film organic field effect transistors during and after device turn-on using Monte Carlo simulation, paying close attention to the effect of device dimensions and Coulomb interactions. I used two initial configurations for charge carriers, an empty lattice and a randomly seeded lattice. In both cases, charges were injected with a constant probability throughout the simulation and a desired carrier concentration was maintained, and I found that a more sophisticated injection model was not needed.

The results for an empty lattice, independent of the presence of Coulomb interactions, showed a turn-on source-drain voltage in the current-voltage curve that was, however, transient, and disappeared after the system equilibrated, resembling a typical current-voltage curve for a randomly seeded system. On closer examination, I found that the turn-on for an empty lattice was linked to an oscillating instantaneous current during device turn-on. More surprisingly, the oscillating current was present in the well-equilibrated lattice, though lower in magnitude, even for a randomly seeded lattice, which showed no turn-on voltage.

The Fourier transform of current as a function of simulation time, independent of initial configuration of charge carriers (empty lattice or randomly seeded) and the presence of Coulomb interactions, showed well-defined fundamental peaks and overtones that were a function of device dimensions. While independent of device width, peak frequencies and their overtones shift, non-linearly, towards lower frequencies as device length increased. Additionally, carrier lifetime and path length were found to increase non-linearly with device length and remain constant with device width. I proposed that the nonlinear behavior of the carrier lifetime, related to the frequencies observed in the Fourier transform, stems from nonlinear increase of carrier path length with device length.

Finally, I proposed that the oscillating behavior of the current was a result of carrier injection from the source, which could not inject charges when the system was at the free charge carrier concentration limit. As a result of this injection process, carriers were found to travel in waves of charge density even in the well equilibrated device. The fundamental oscillation frequencies decrease non-linearly with device length. Therefore, this effect is a phenomenon of finite sized devices.

Observing non-linear behavior in experimental devices would be a worthwhile exercise. That is, how do carrier lifetime and path length change with device dimensions, and what would this mean for concepts like carrier mobility? Likewise, do charge carriers travel in waves in real devices? If our proposed mechanism is correct, then they may, for real devices do have a restricted free carrier concentration, fixed by a gate electrode, as well as injection probability. To answer these questions, however, device properties like source-drain current would need to be resolved to timescales on the order of nanoseconds. This is not outside the realm of techniques like complex impedance spectroscopy, used to measure AC conductivity,

and provide information on processes with diverse timescales. For example, dielectric spectroscopy measurements have been made of doped polythiophene, poly(phenylene vinylene),[78] or polypyrrole,[79] which have measurements in the sub-THz range ($\tau \approx \omega^{-1}$, THz $\sim$ps).

7.1.2 Trap/Barrier Concentration

The behavior of current in OFETs as a function of trap concentration at modest trap energies ($0.0\,\text{eV} < \Delta V \leq 0.3\,\text{eV}$ for electrons, $-0.3\,\text{eV} \leq \Delta V < 0.0\,\text{eV}$ for holes) is antisymmetric. That is, the current does not reach a minimum at 50 % traps, as would be predicted by simple percolation theory. Rather, the minimum is reached more rapidly (≤ 20 % traps) for energies as low as $|\Delta V| \leq 0.15\,\text{eV}$. Likewise, recovery of the current happens much earlier ($\sim$60 % traps/barriers) than expected based on the initial falloff. Barriers of modest energy ($0.0\,\text{eV} < \Delta V \leq 0.3\,\text{eV}$ for holes, $-0.3\,\text{eV} \leq \Delta V < 0.0\,\text{eV}$ for electrons), produce a current vs. concentration curve that is a mirror image of the trap system.

7.1.3 Negative Differential Resistance

Beyond a critical value ($|\Delta V| > 0.38\,\text{eV}$ for the systems studied), the mirror symmetry between trap and barrier current vs. concentration curves is broken. At this point, traps are significantly deep that carriers cannot escape. However, beyond this energy, barriers begin to behave like defects. The probability for carriers to sit behind defects increases as the voltage is increased. This leads to a decrease of current in the saturation region. Coulomb interactions with the "stuck" carriers push other carriers around the barriers. This leads to a reduction in the magnitude of the negative differential resistance.

7.1.4 Monolayers

I extended the Monte Carlo model to solar cells simulation. The extended model was employed to simulate charge transport in three classes of morphologies. Increasing the domain size, $\langle L_P \rangle$, in isotropic two-phase morphologies increased device performance. In contrast, increasing domain size in two-phase band systems decreased device performance. I

attributed increasing device performance in isotropic two-phase systems to decreased tortuosity of charge pathways and increased connectivity to the electrode surfaces. I attributed decreasing device performance with increasing domain size in band systems to a lack of interfaces for exciton dissociation. When electron accepting aggregates are dispersed randomly throughout the electron donor, I observed severe decreases in device performance in the isotropic two-phase systems. I ascribed this to the formation of negatively charged defects in the hole conducting phase.

Larger domains leading to more efficient exciton separation is in opposition to the common held belief that small domains to mitigate recombination losses. While this is in some sense true, as shown by the monolayer band systems, a more complex mechanism is at play. This is because there are still hindrances to charge transport after exciton separation. Escaping charge transfer states is an energetically expensive process. While the LUMO-LUMO offset provides energy to push the exciton into the charge transfer state near an interface, more energy is still needed to escape the charge transfer state. Other process must be at play, such as kinetically hot electrons or weakening of Coulomb barriers through charge delocalization.

Finally, I observed that Coulomb interactions are the origin of increased efficiency in the small devices due to a field dependent mobility. This was a hint that charge delocalization can lead to more efficient devices. This is because non-interacting charges can be viewed as infinitely delocalized charges. For all systems, Coulomb interactions were found to play a crucial role in device behavior, and suggest that Coulomb interactions are adequately accounted for in simulation models.

7.1.5 Multilayers

Multiple layer simulations are expensive. The equilibration time and number of charge carriers increase significantly. Calculations were feasible only by extensive use of GPUs for the Coulomb interactions. Such novel algorithms are a breakthrough for OPV simulation.

Isotropic morphologies of increasing domain size showed less correlation with domain size than monolayer systems. The increased number of charge pathways in three dimensions

reduces the effects of tortuosity observed in monolayers. In monolayer systems, the fill factor increased with domain size, because the tortuosity decreased. Even the smallest of domain sizes in multilayer systems shows efficient charge extraction.

Device efficiency for ideal morphologies was not significantly different than isotopic systems. This suggests that morphology, at least at the level studied, is not playing as much of a role in charge extraction. While ideal morphologies contain perfect connectivity to the electrodes, multilayer isotopic systems still contain appreciable connectivity. For this reason, it is not surprising that efficiencies between the two were similar.

Lastly, I have predicted that the charge delocalization leads to increased charge extraction in OPVs. In some sense, charge delocalization is a finer aspect of system morphology. Alignment of aromatic rings, for example, can lead to charge delocalization and increased mobility in preferential directions. I propose that the weakening of energetic barriers, particularly in Coulomb interface traps leads to the efficient separation of charges. Other mechanisms, absent from the model, such as hot carriers, could lead to interface trap escape. This work has made it apparent that exciton separation is an easier step when compared to breaking apart charge transfer states.

In summary, in isotropic two-phase morphologies, decreased domain size leads to increased device performance. Unlike thin films, where the Coulomb force dominates the energy landscape of separated charges, more pathways open up in 3D systems. Small domains lead to efficient exciton separation, while the increased number of pathways in 3D alleviate interface trap formation between separated carriers.

7.1.6 Conclusion

This thesis has focused on on using Monte Carlo simulation to extract device relevant properties, such as the current voltage behavior of transistors and of efficiency photovoltaics, from the molecular picture of hopping. The physical picture presented by the model, which is often absent from other models, has been invaluable to interpreting OFET measurements. The approximation of the Gaussian disorder of energy levels (traps) has been relaxed. This has provided a model with a highly tunable energy level distribution, that can be compared

easily with experiments. From the model, the mechanism by which current is mediated by the concentration traps and barriers at different regions of the current voltage curve is now understood and verified. Likewise, a mechanism for negative differential resistance, which is observed in experiments,[62] has been proposed in terms of the presence of barriers and defects.

Extending the hopping model to OPVs is a new approach, markedly different from kinetic Monte Carlo and continuum models. Predicting device efficiency from morphology is a hard problem, involving effects from multiple length and time scales. Connecting the molecular picture of electron transfer to the macroscopic world of functional devices is necessary for fill factor prediction. The model has begun to challenge long help beliefs about the transport of charges in OPVs. The common argument of small domain size leading to efficient exciton separation in OPVs is only part of a more complex picture involving the interplay between domain size, connectivity to electrodes, tortuosity of pathways, Coulomb bound interface traps, and charge delocalization.

APPENDIX A

OPV MONOLAYER DATA

This appendix contains data that accompanies chapter 5.

A.1 BAND

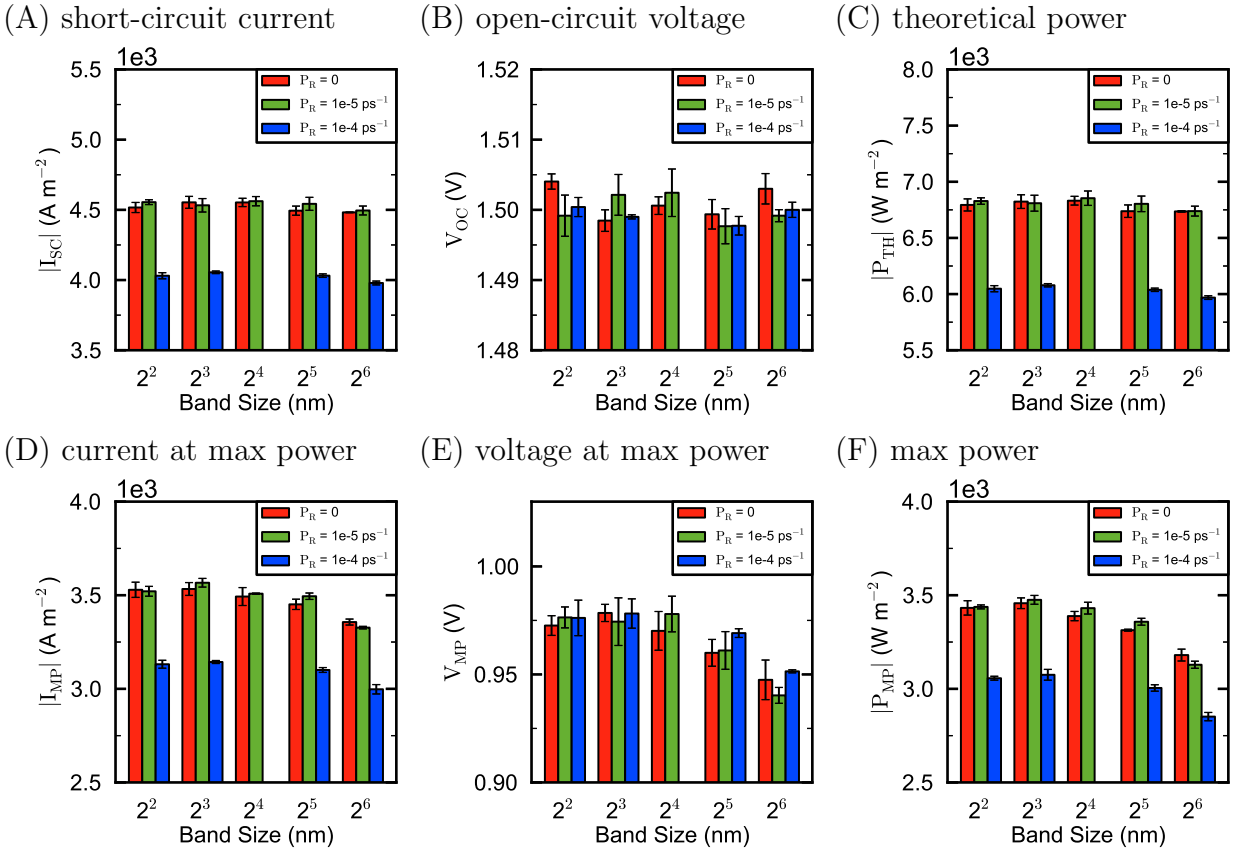


Figure A1: IV curve parameters for 2D band morphologies.

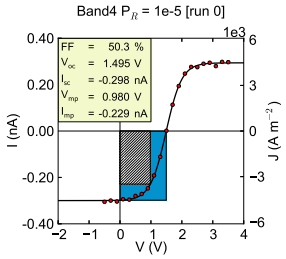
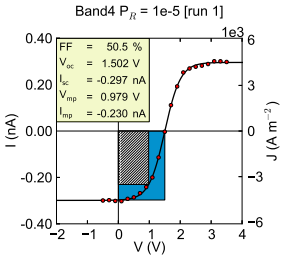
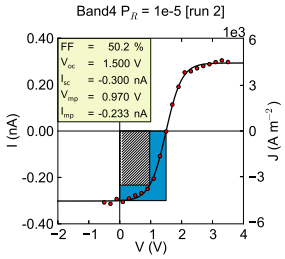
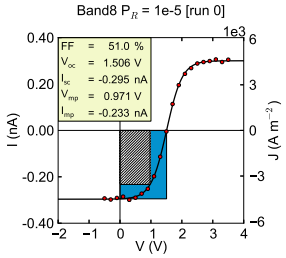
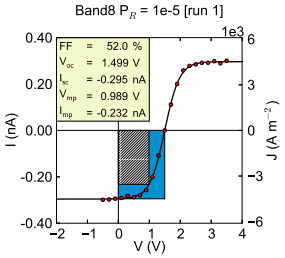
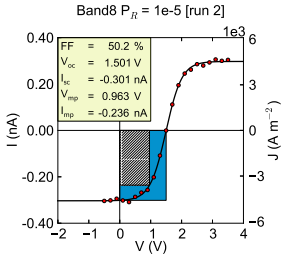
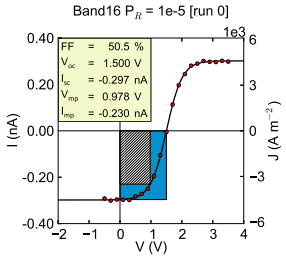
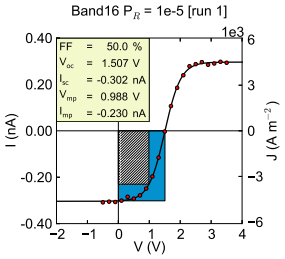
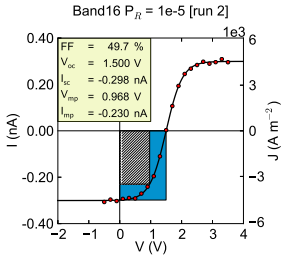
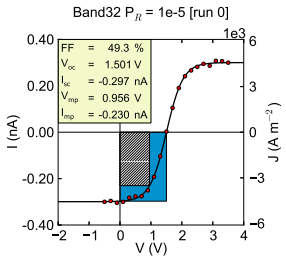
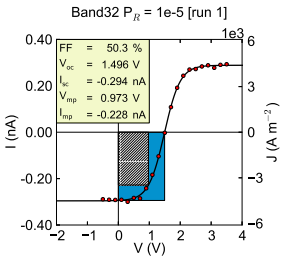
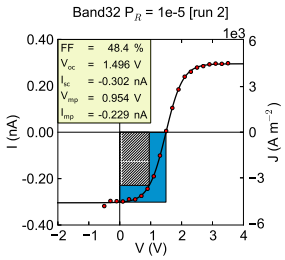
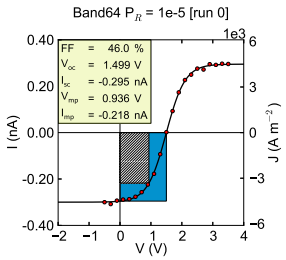
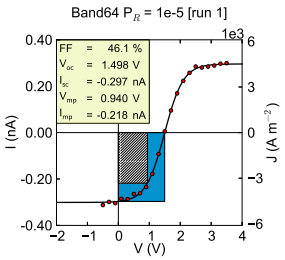
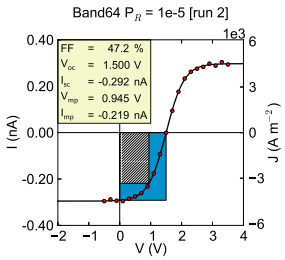
$P_R = 10^{-5}$	**run 0**	**run 1**	**run 2**
4 nm			
8 nm			
16 nm			
32 nm			
64 nm			

Figure A2: IV curves - band with recombination

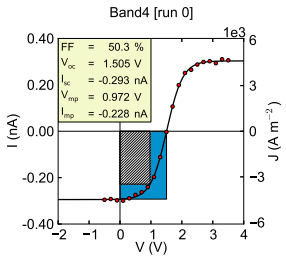
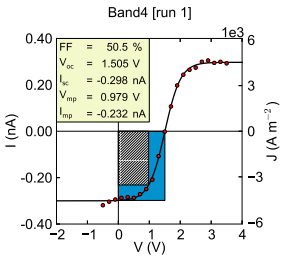
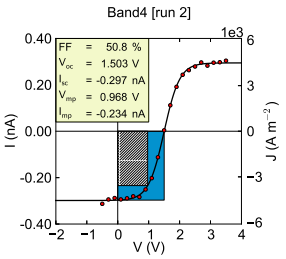
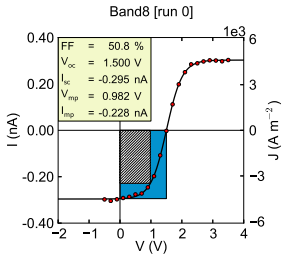
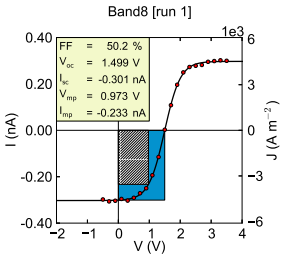
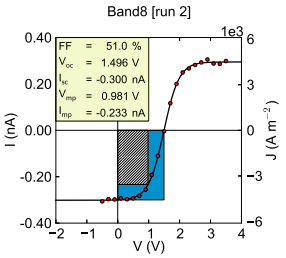
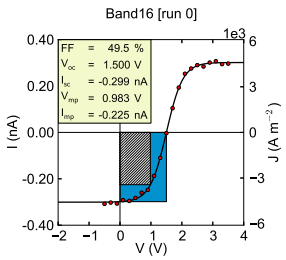
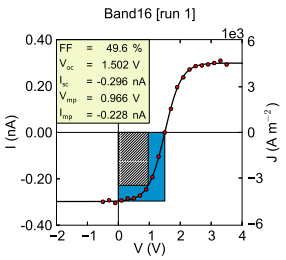
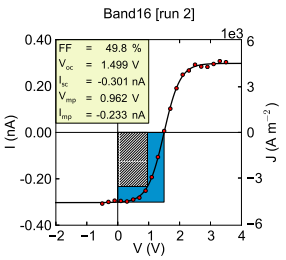
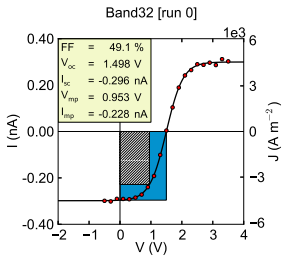
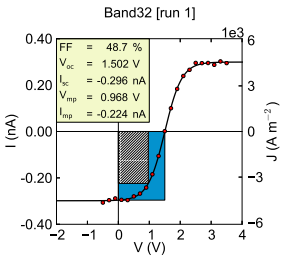
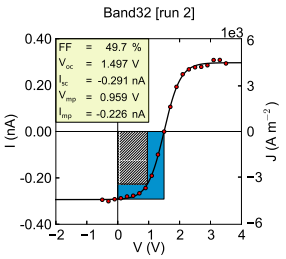
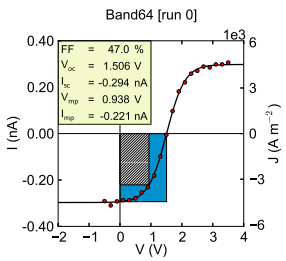
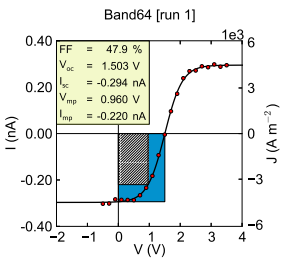
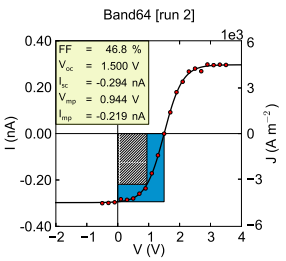
$P_R = 0$	**run 0**	**run 1**	**run 2**
4 nm			
8 nm			
16 nm			
32 nm			
64 nm			

Figure A3: IV curves - band without recombination

A.2 MONO

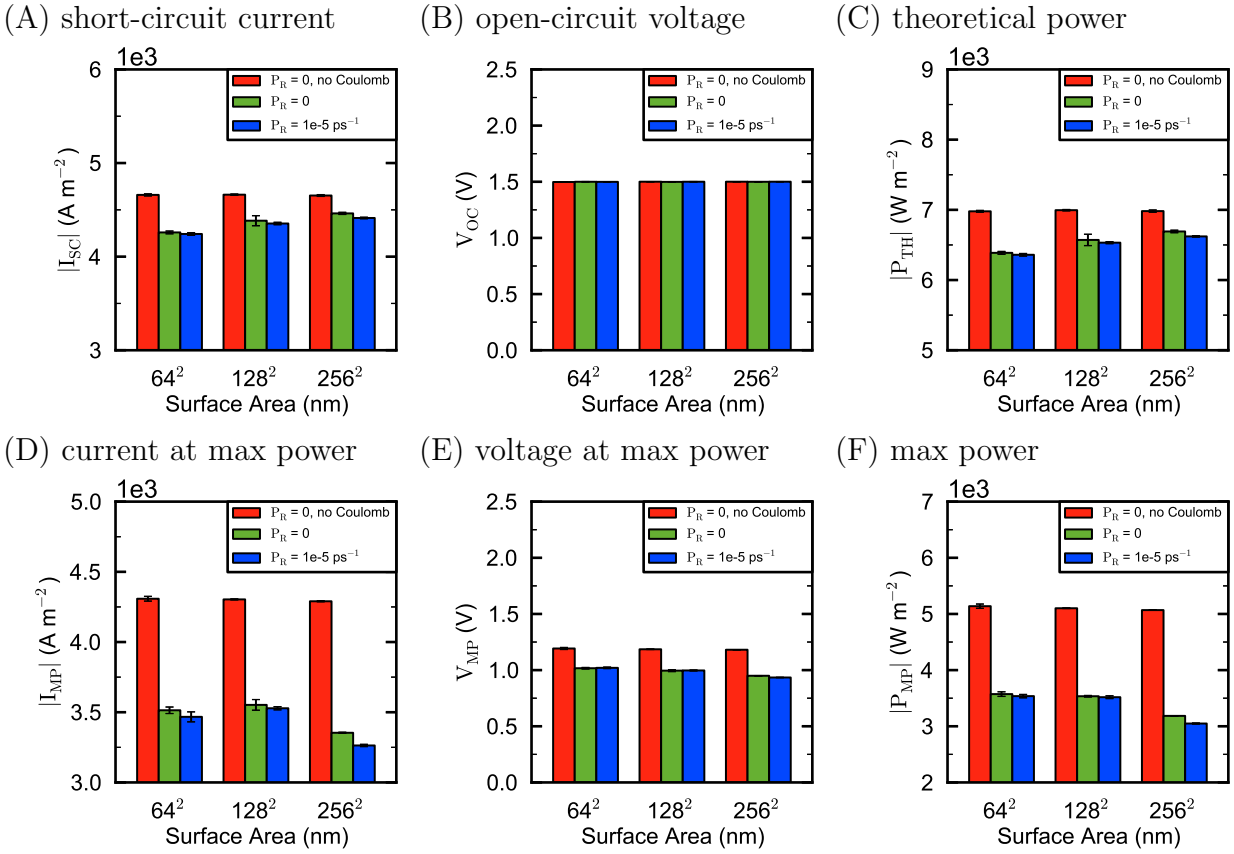


Figure A4: IV curve parameters for 2D mono morphologies.

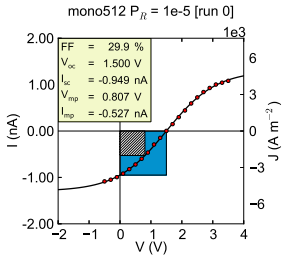
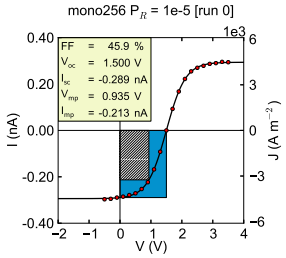
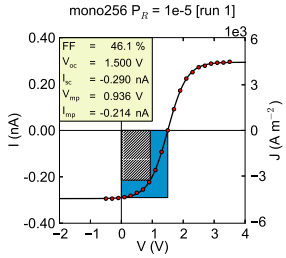
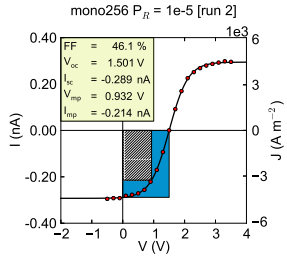
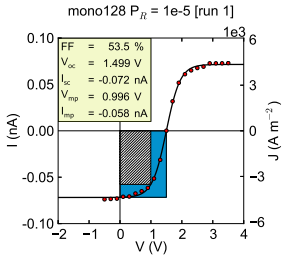
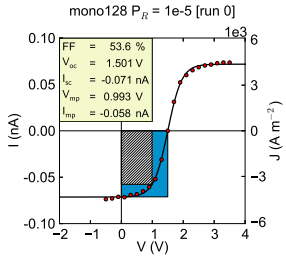
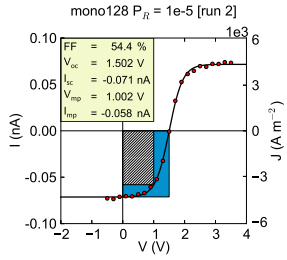
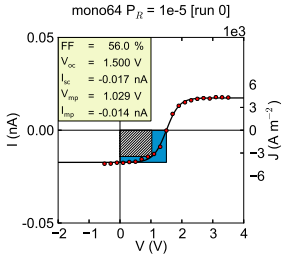
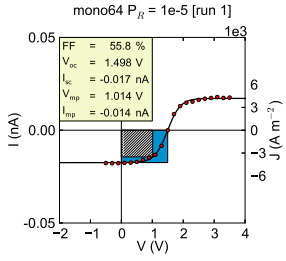
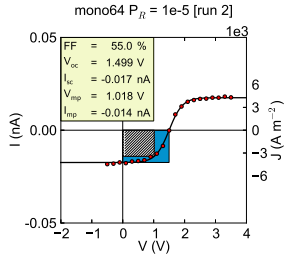
$P_R = 10^{-5}$	**run 0**	**run 1**	**run 2**
$512^2\,\mathrm{nm}^2$			
$256^2\,\mathrm{nm}^2$			
$128^2\,\mathrm{nm}^2$			
$64^2\,\mathrm{nm}^2$			

Figure A5: IV curves - mono with recombination

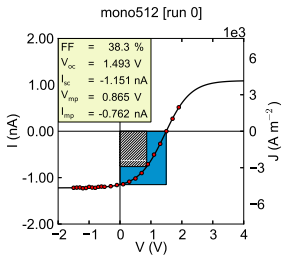
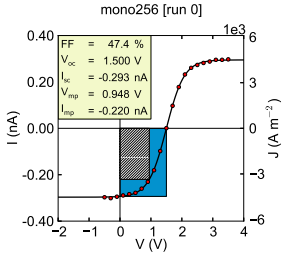
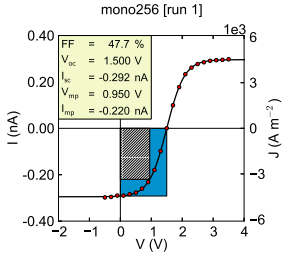
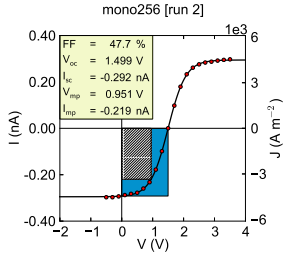
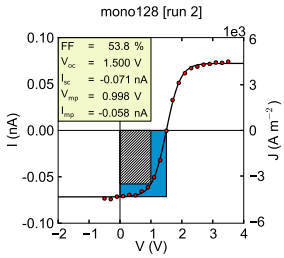
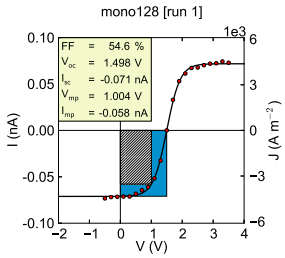
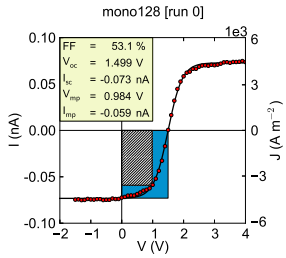
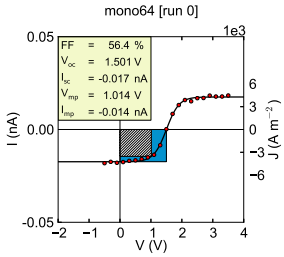
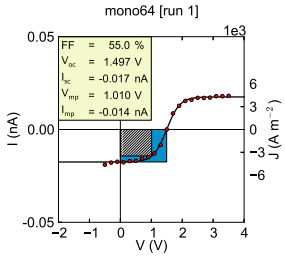
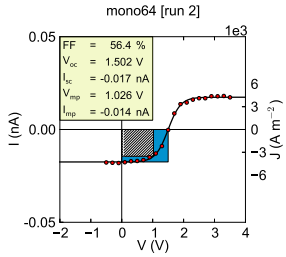
$P_R = 0$	**run 0**	**run 1**	**run 2**
$512^2\,\mathrm{nm}^2$			
$256^2\,\mathrm{nm}^2$			
$128^2\,\mathrm{nm}^2$			
$64^2\,\mathrm{nm}^2$			

Figure A6: IV curves - mono without recombination

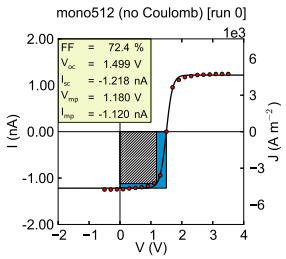
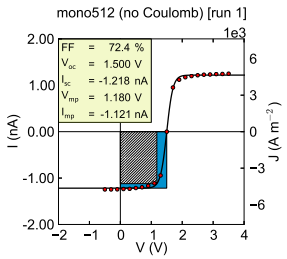
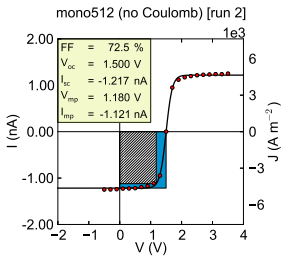
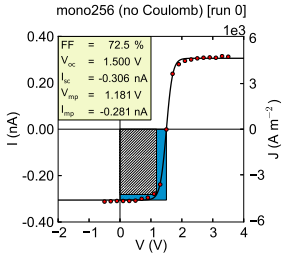
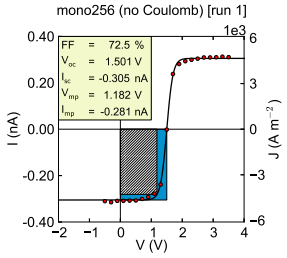
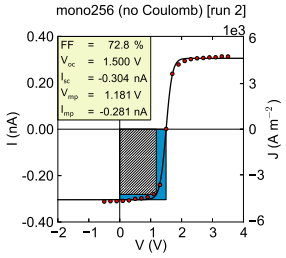
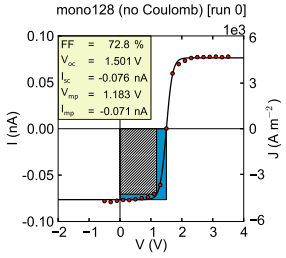
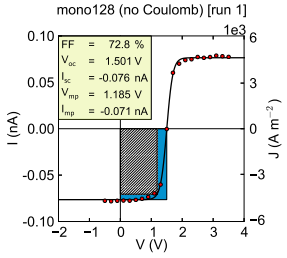
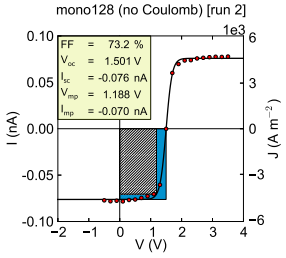
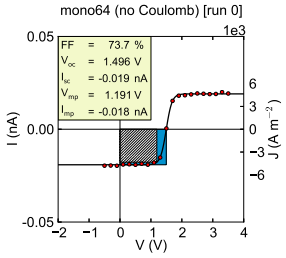
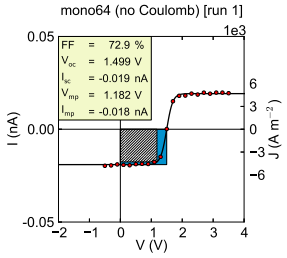
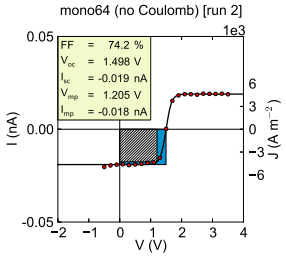
$P_R = 0$	**run 0**	**run 1**	**run 2**
$512^2\,\mathrm{nm}^2$			
$256^2\,\mathrm{nm}^2$			
$128^2\,\mathrm{nm}^2$			
$64^2\,\mathrm{nm}^2$			

Figure A7: IV curves - mono without Coulomb

A.3 ISO

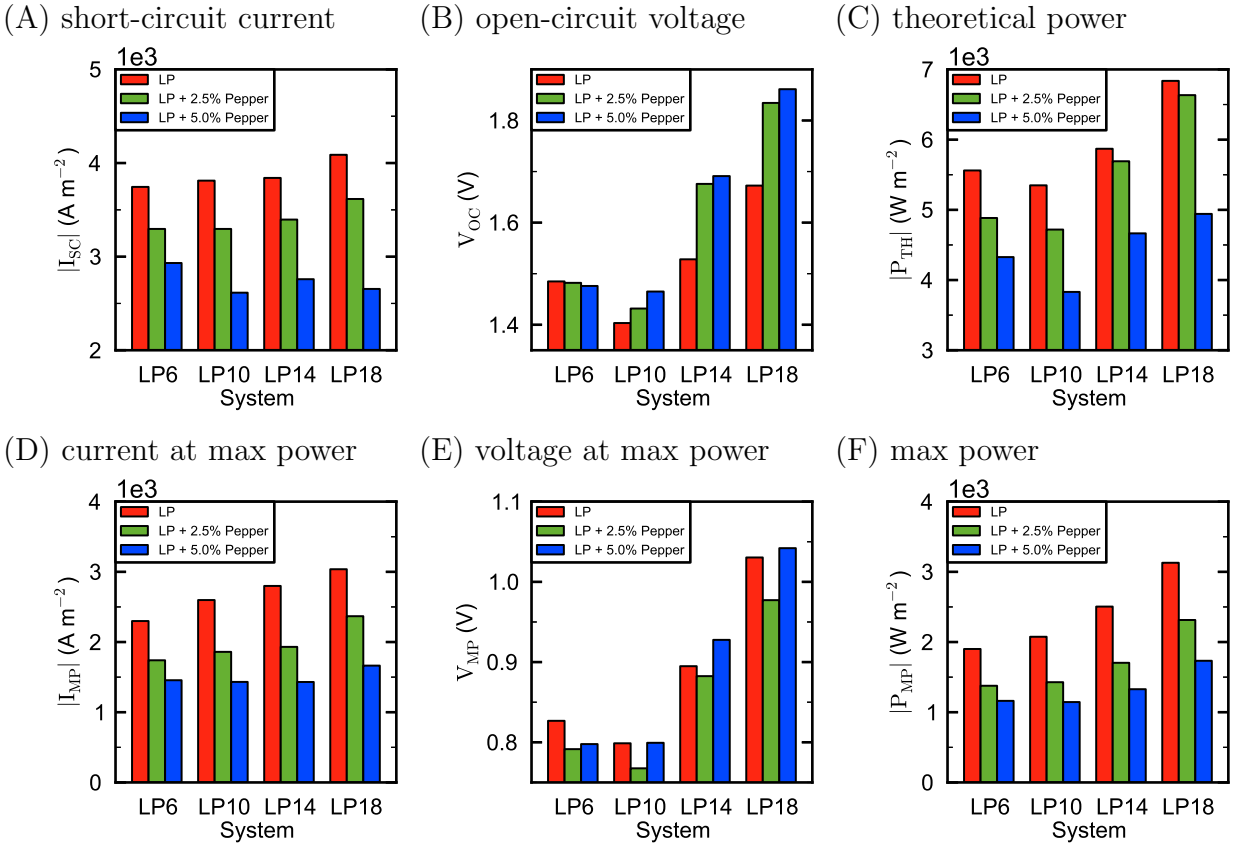


Figure A8: IV curve parameters for isotropic 2D morphologies.

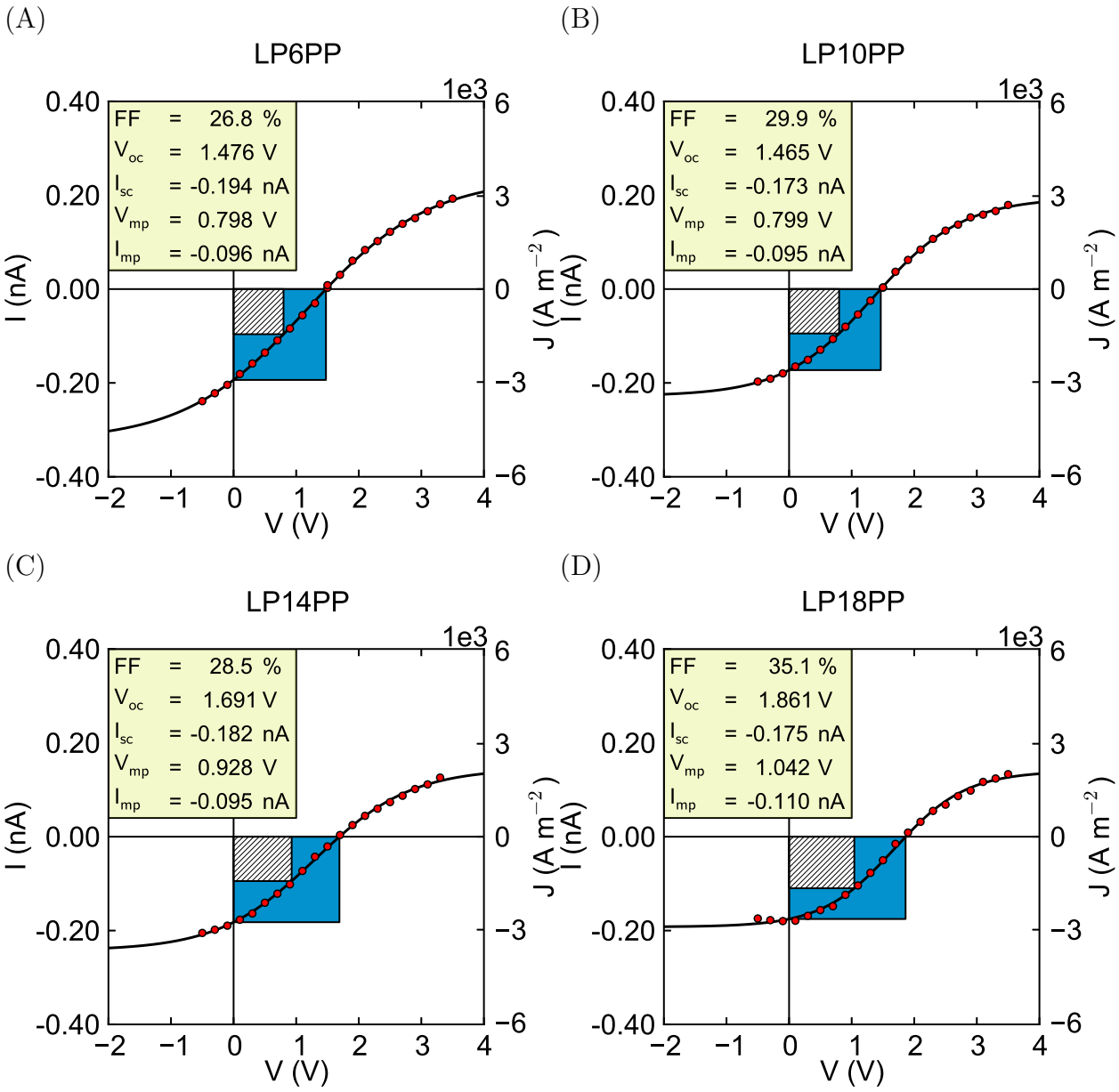


Figure A9: IV curves for isotropic 2D morphologies with pepper 10.

APPENDIX B

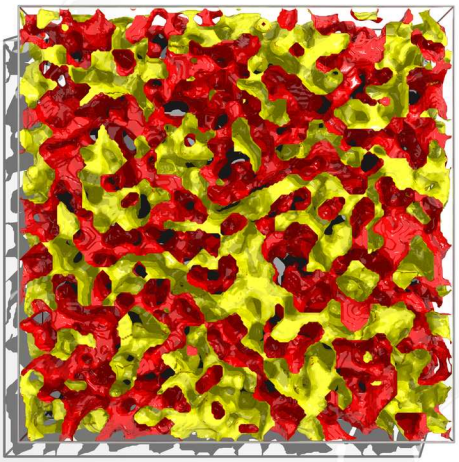
OPV MULTILAYER DATA

This appendix contains data that accompanies chapter 6. All IV curves, surface morphologies, and IV curve parameters are included.

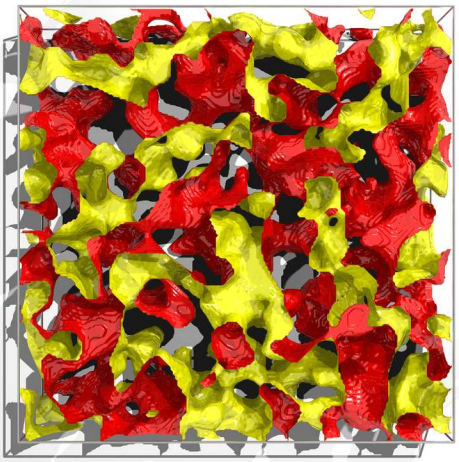
B.1 ISOTROPIC MORPHOLOGIES

B.1.1 Surfaces

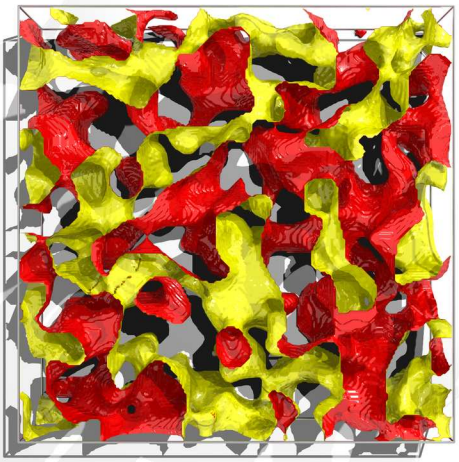
(A) iso3



(B) iso4



(C) iso6



(D) iso9

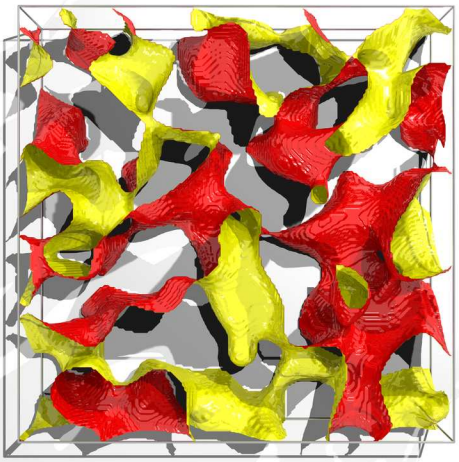


Figure B1: Isotropic 3D surfaces.

B.1.2 Parameters

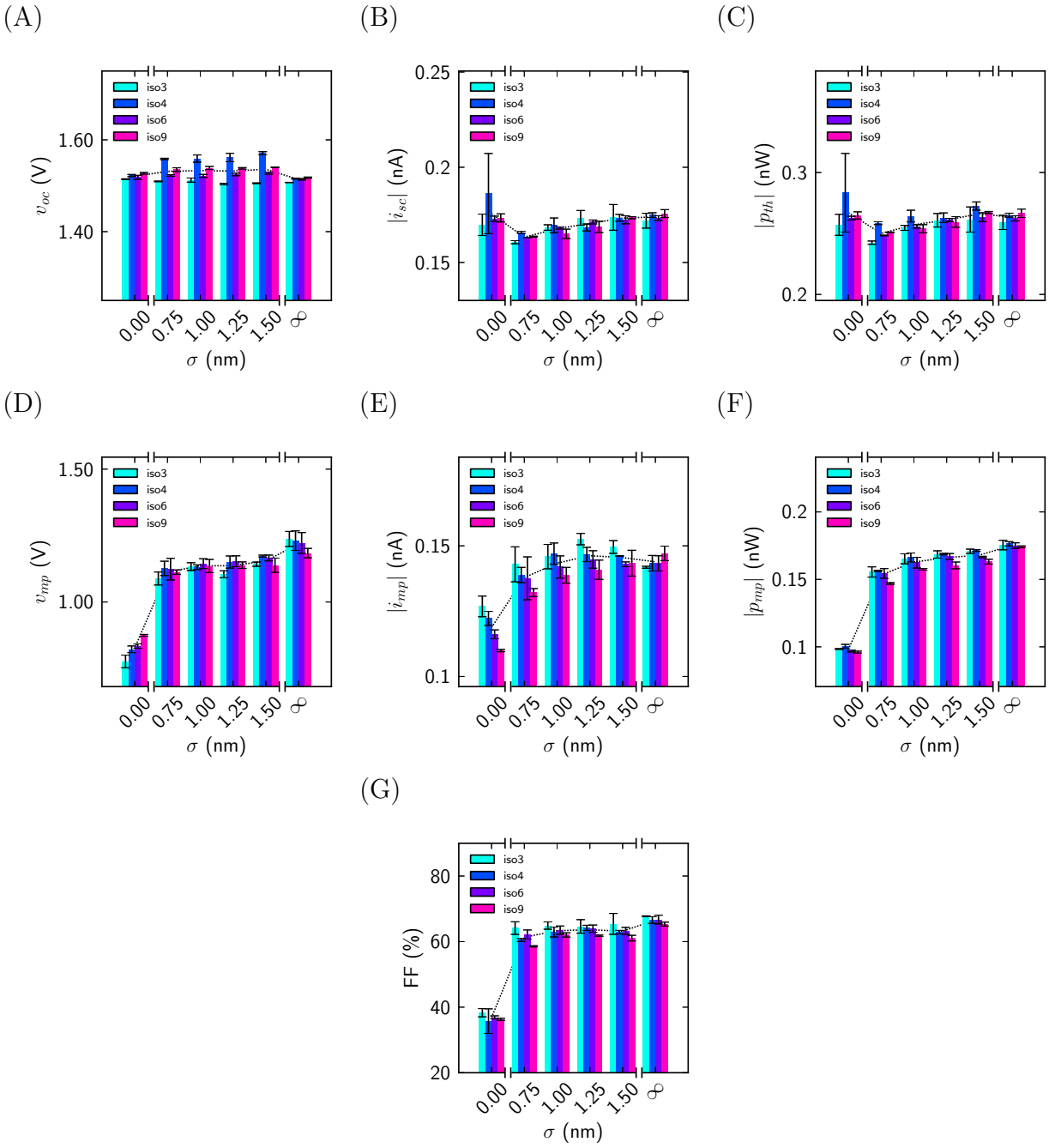


Figure B2: Summary of isotropic morphologies.

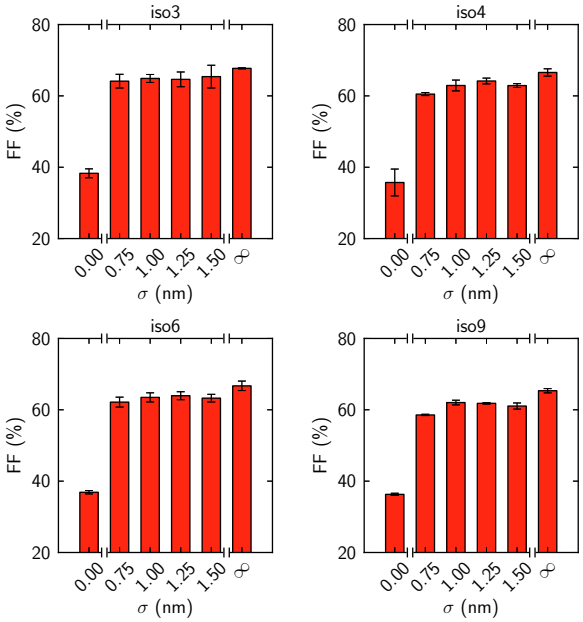


Figure B3: Iso 3D - FF

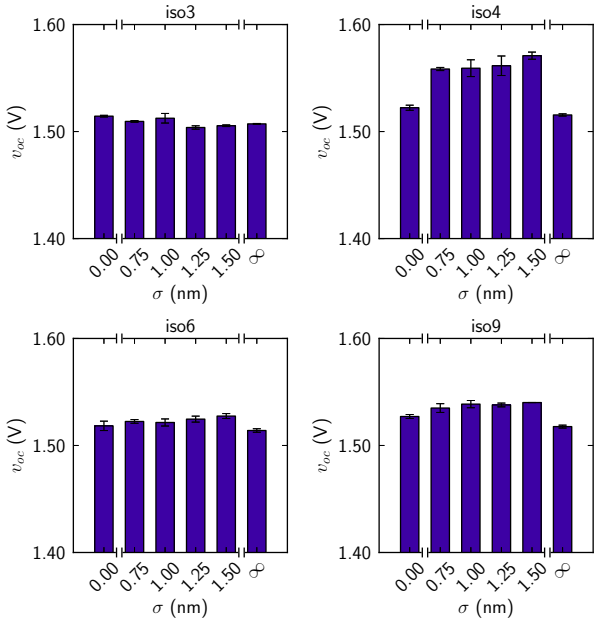


Figure B4: Iso 3D - Voc

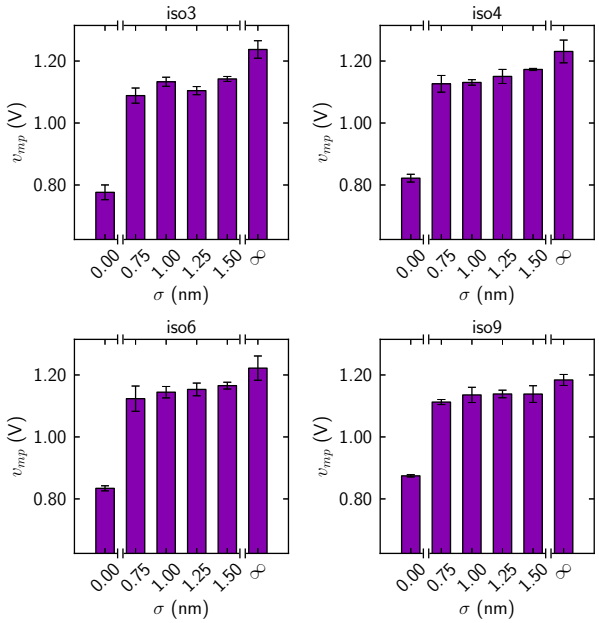


Figure B5: Iso 3D - Vmp

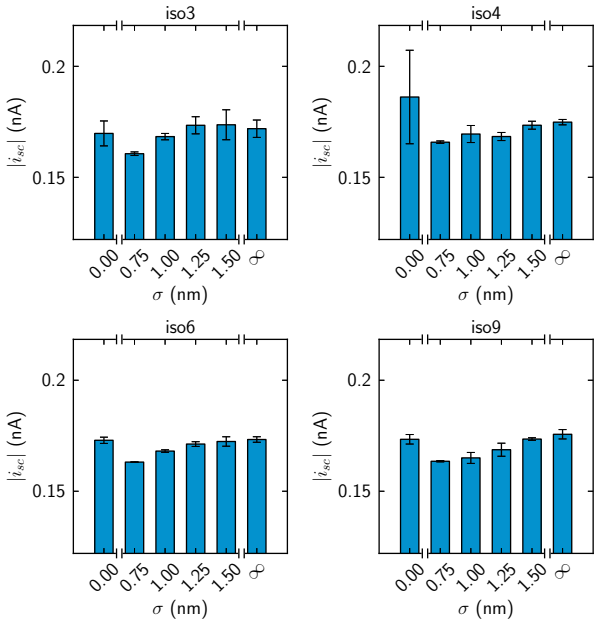


Figure B6: Iso 3D - Isc

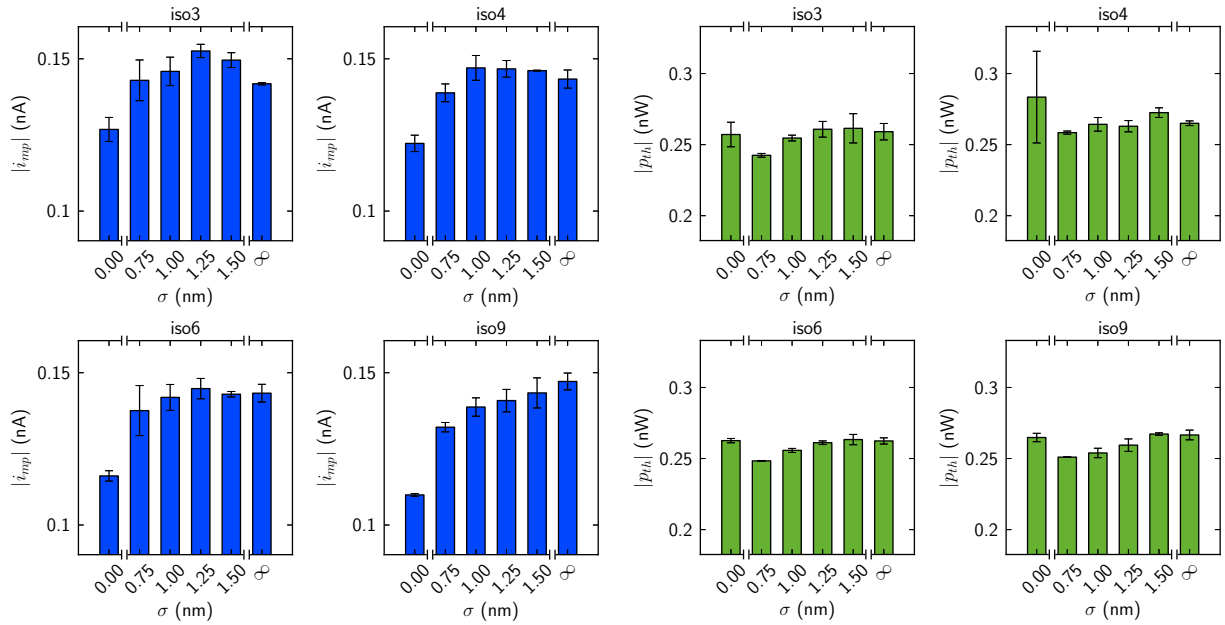


Figure B7: Iso 3D - Vmp

Figure B8: Iso 3D - Pth

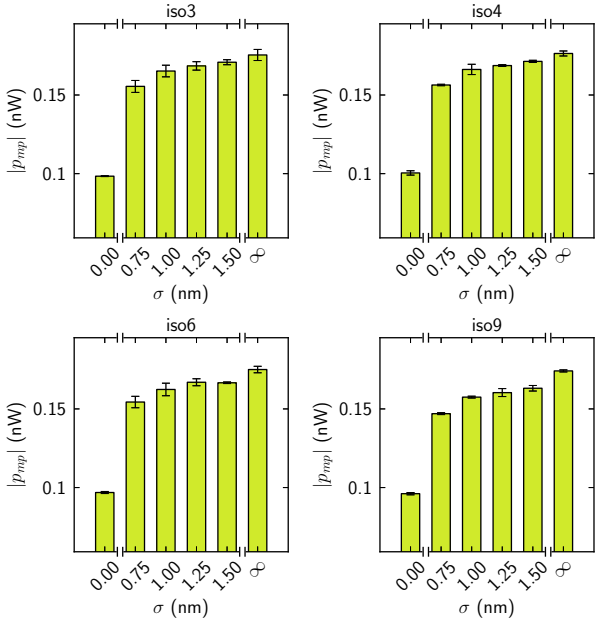


Figure B9: Iso 3D - Pmp

B.1.3 IV Curves

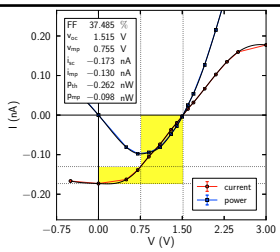
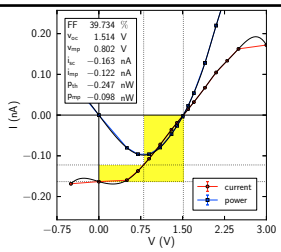
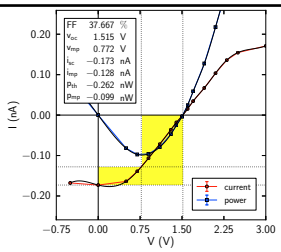
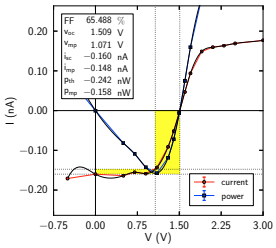
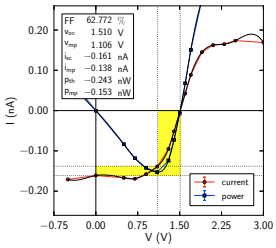
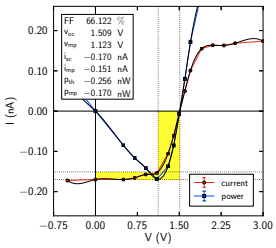
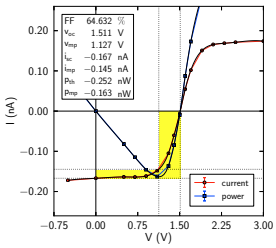
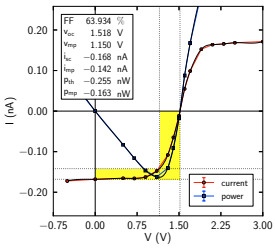
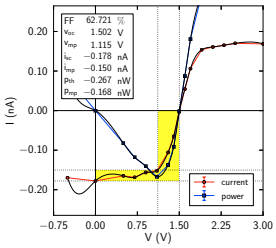
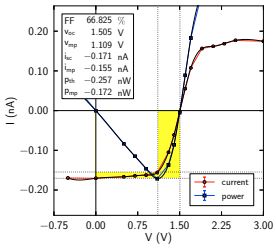
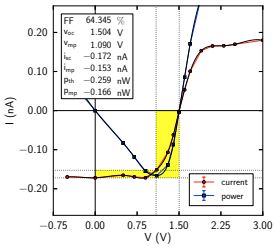
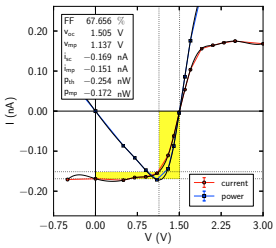
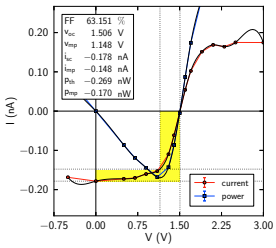
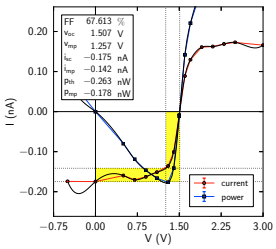
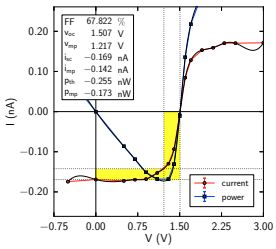
iso3	run 0	run 1	run 2
$\sigma =$ 0.00 nm			
$\sigma =$ 0.75 nm			
$\sigma =$ 1.00 nm			
$\sigma =$ 1.25 nm			
$\sigma =$ 1.50 nm			
$\sigma = \infty$			

Figure B10: IV curves - iso3

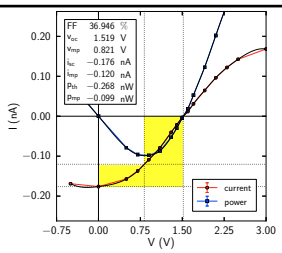
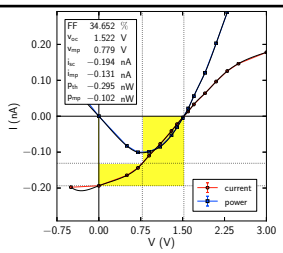
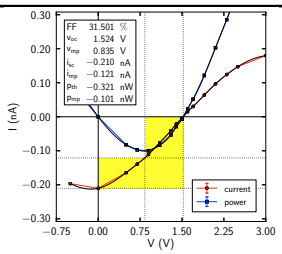
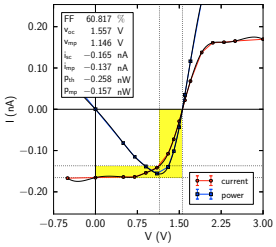
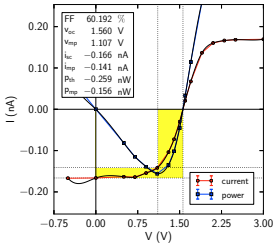
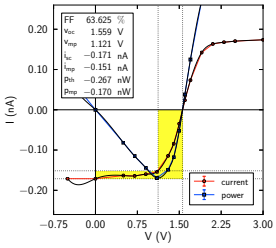
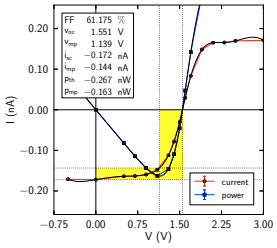
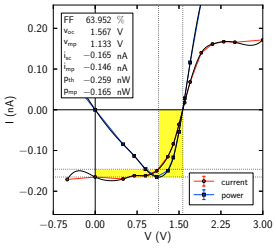
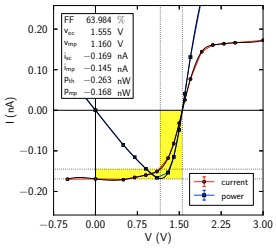
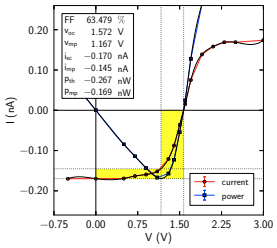
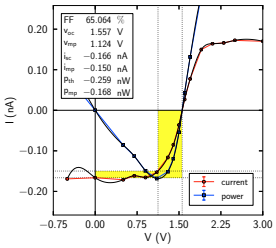
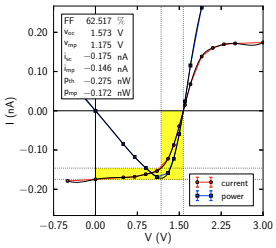
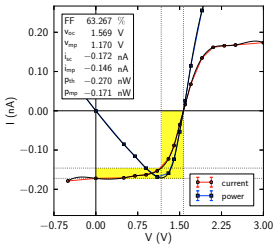
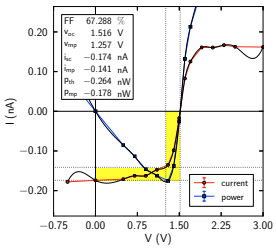
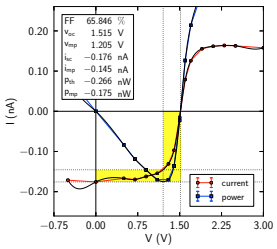
iso4	run 0	run 1	run 2
$\sigma =$ 0.00 nm			
$\sigma =$ 0.75 nm			
$\sigma =$ 1.00 nm			
$\sigma =$ 1.25 nm			
$\sigma =$ 1.50 nm			
$\sigma = \infty$			

Figure B11: IV curves - iso4

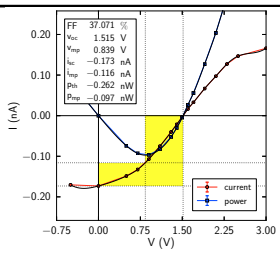
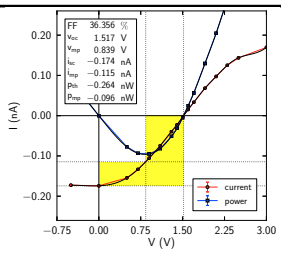
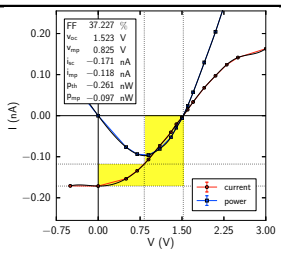
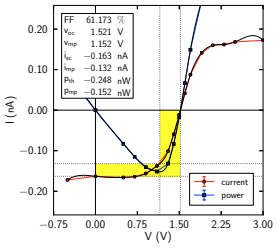
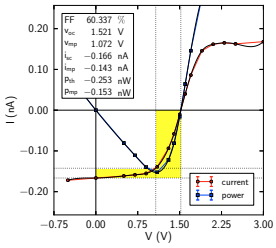
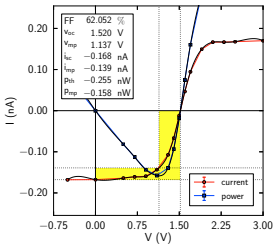
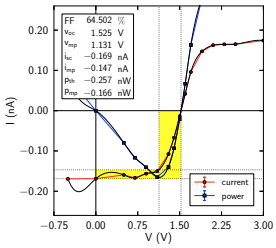
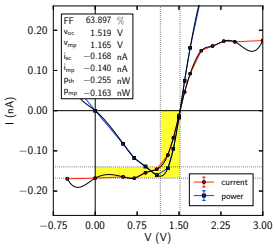
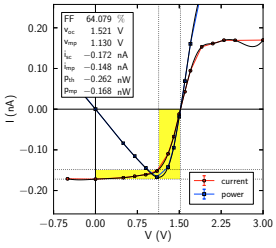
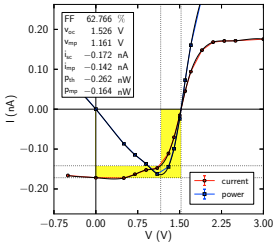
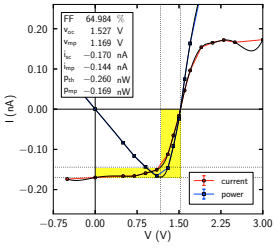
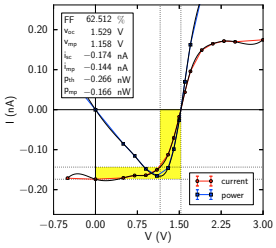
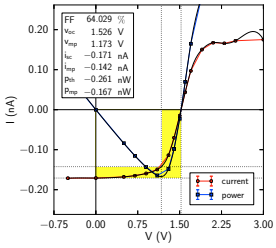
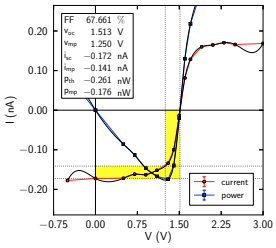
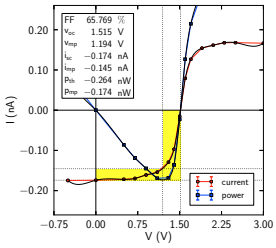
iso6	run 0	run 1	run 2
$\sigma =$ 0.00 nm			
$\sigma =$ 0.75 nm			
$\sigma =$ 1.00 nm			
$\sigma =$ 1.25 nm			
$\sigma =$ 1.50 nm			
$\sigma = \infty$			

Figure B12: IV curves - iso6

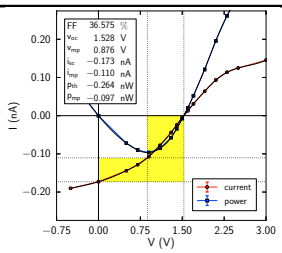
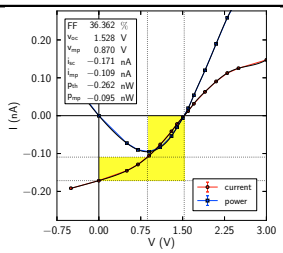
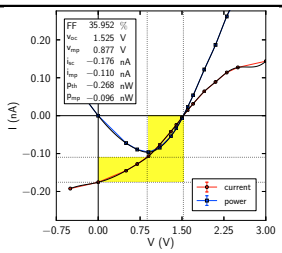
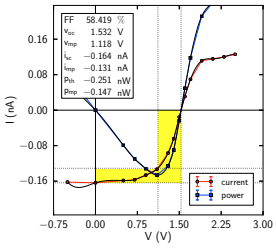
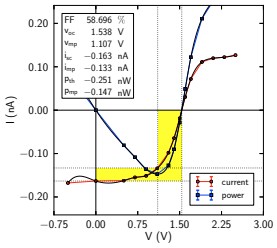
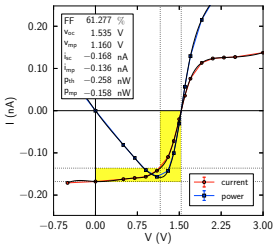
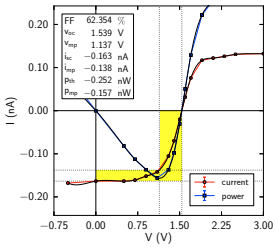
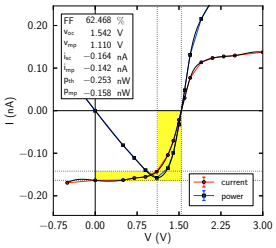
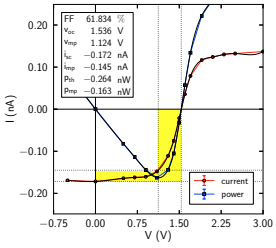
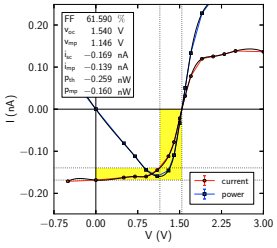
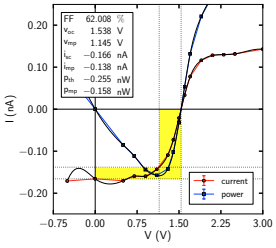
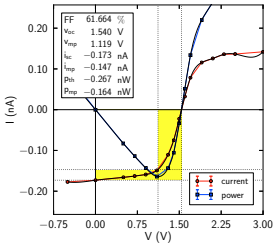
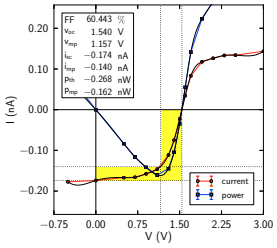
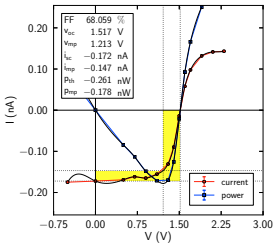
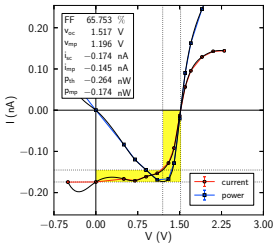
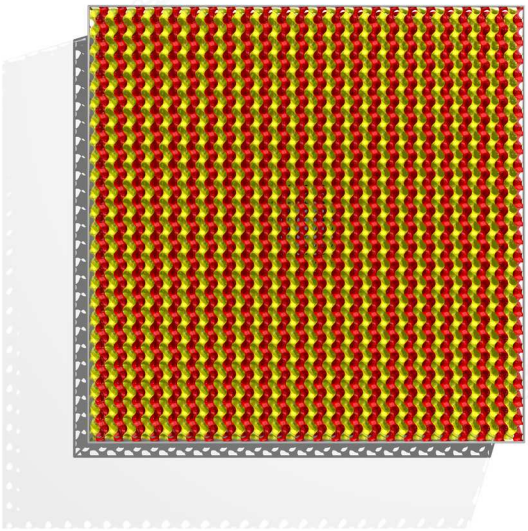
iso9	run 0	run 1	run 2
$\sigma =$ 0.00 nm			
$\sigma =$ 0.75 nm			
$\sigma =$ 1.00 nm			
$\sigma =$ 1.25 nm			
$\sigma =$ 1.50 nm			
$\sigma = \infty$			

Figure B13: IV curves - iso9

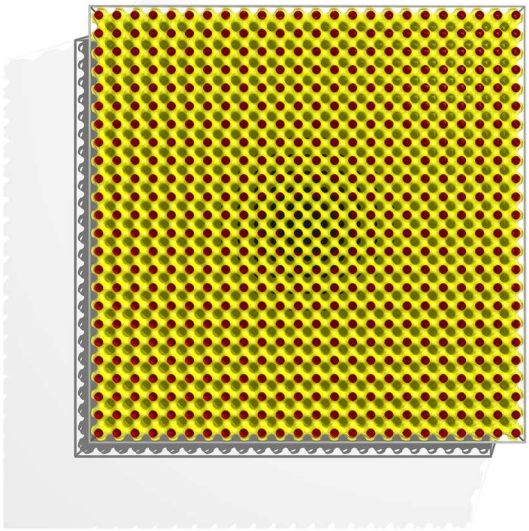
B.2 IDEAL MORPHOLOGIES

B.2.1 Surfaces

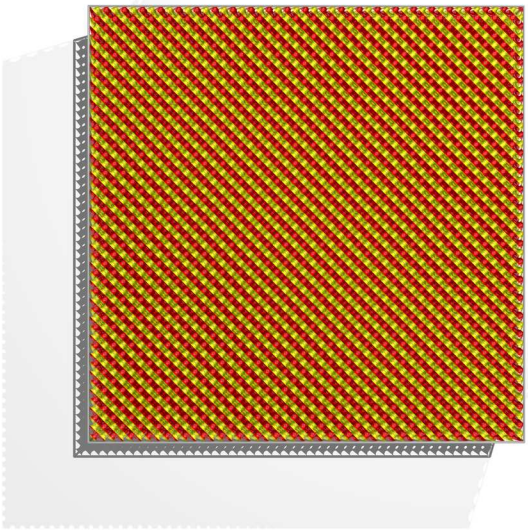
(A) gyroid



(B) p-surface



(C) d-surface



(D) band4

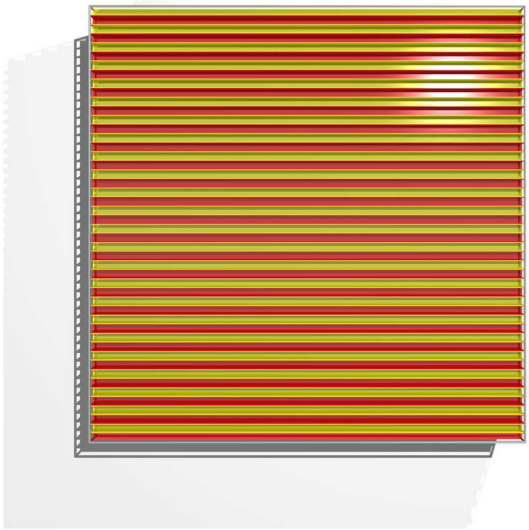


Figure B14: Ideal 3D surfaces.

B.2.2 Parameters

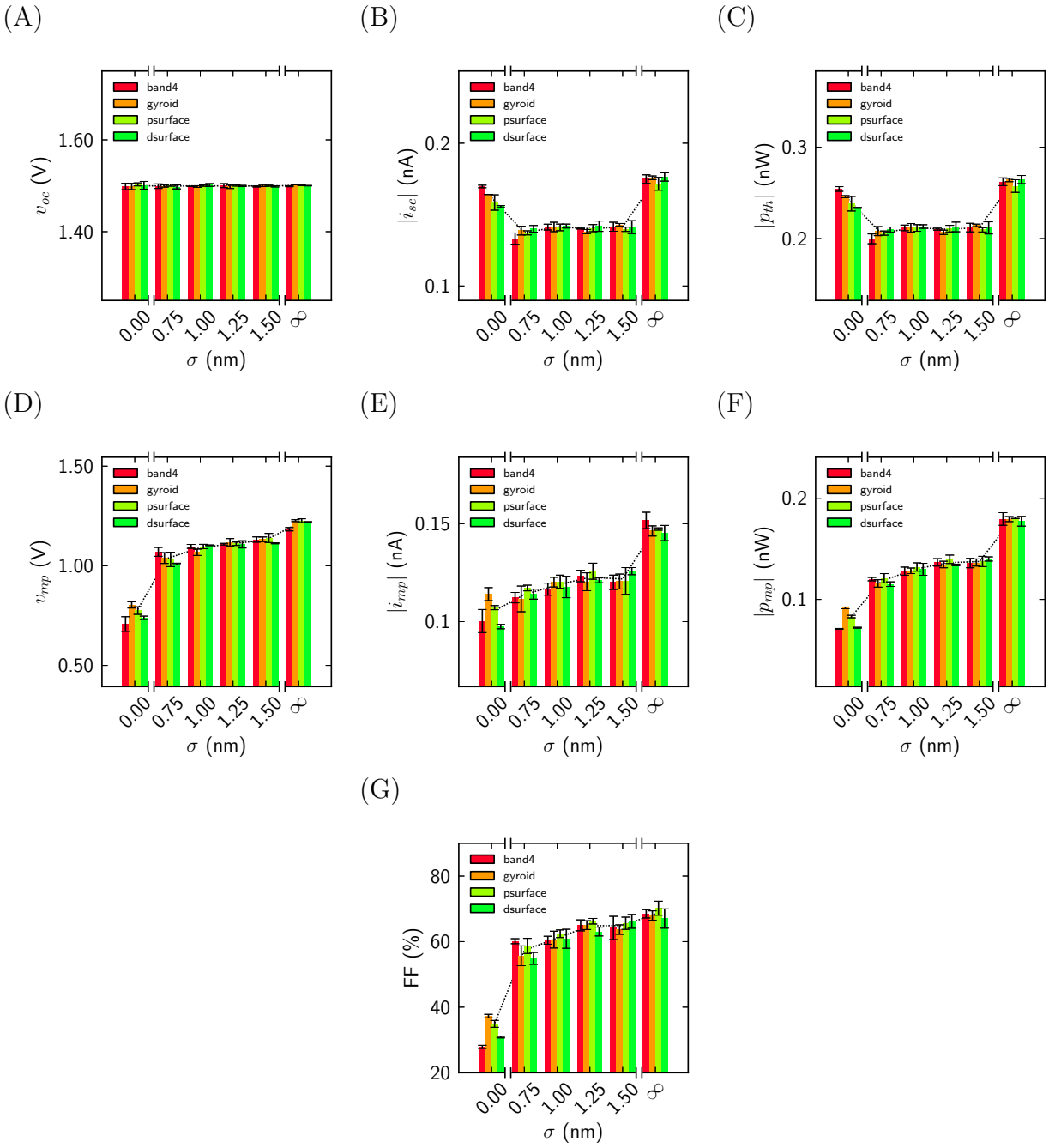


Figure B15: Summary of ideal morphologies.

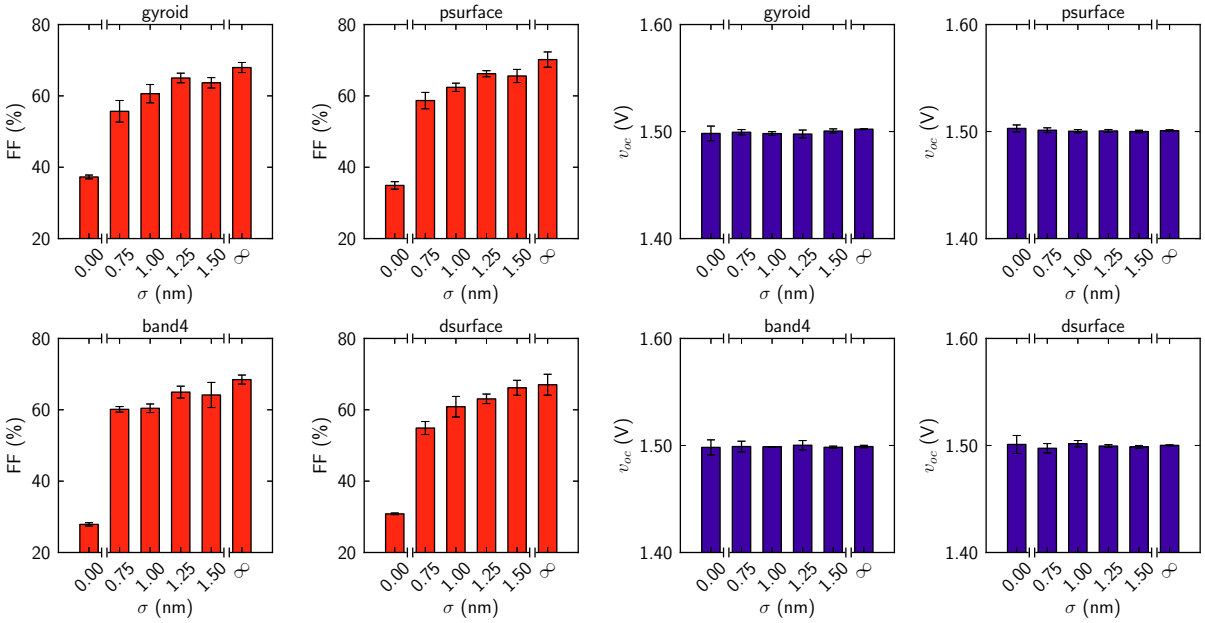


Figure B16: Ideal 3D - FF

Figure B17: Ideal 3D - Voc

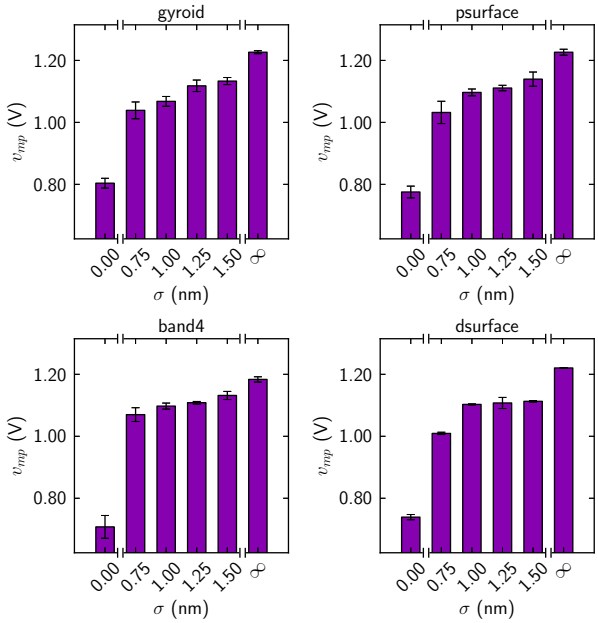


Figure B18: Ideal 3D - Vmp

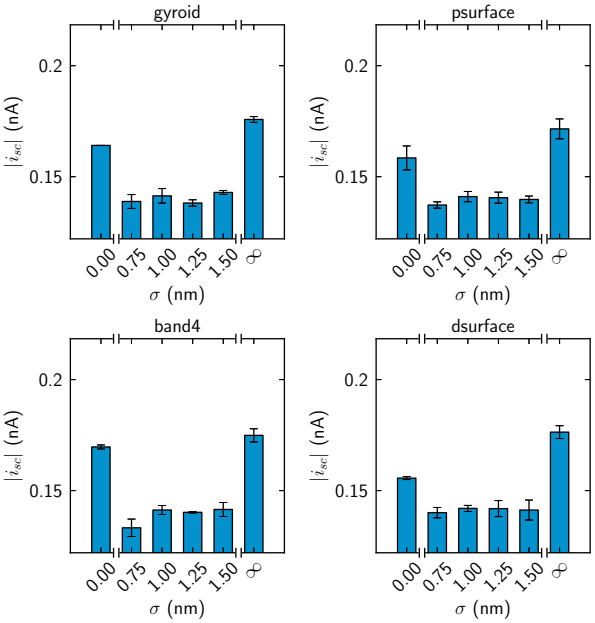


Figure B19: Ideal 3D - Isc

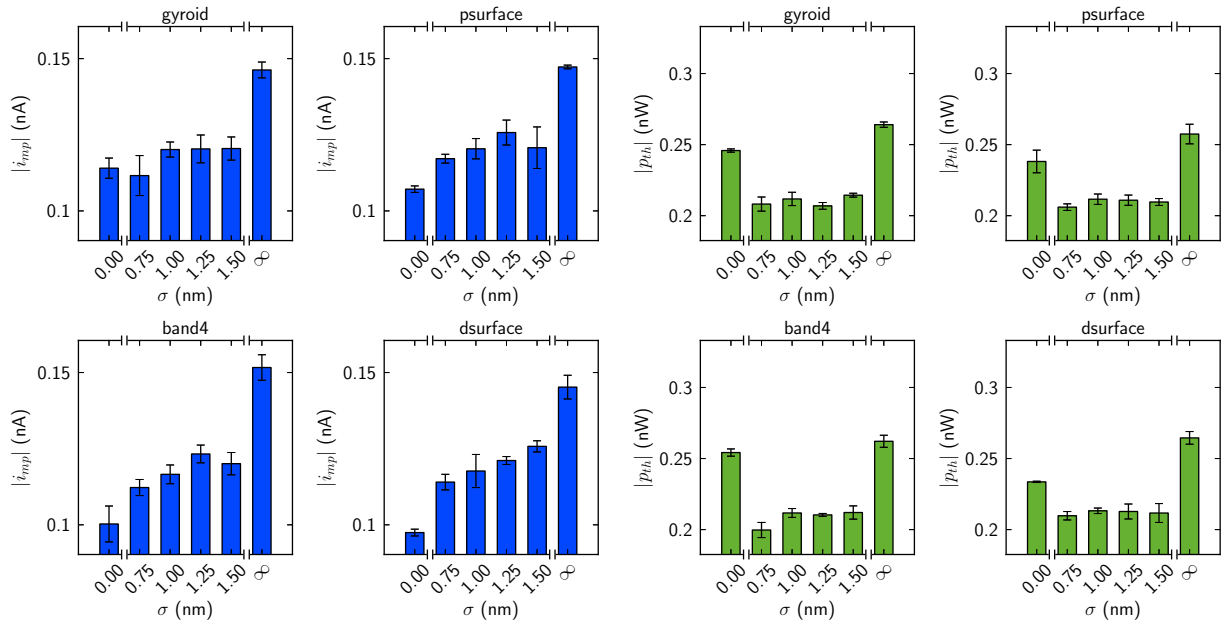


Figure B20: Ideal 3D - Vmp

Figure B21: Ideal 3D - Pth

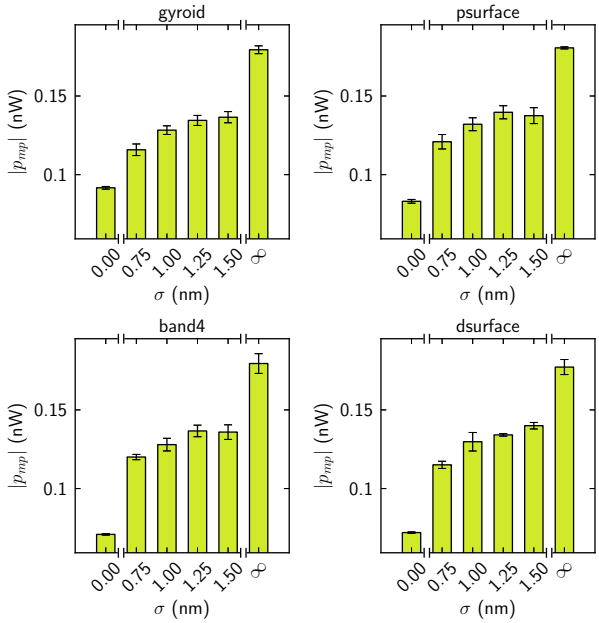


Figure B22: Ideal 3D - Pmp

B.2.3 IV Curves

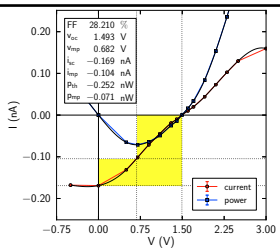
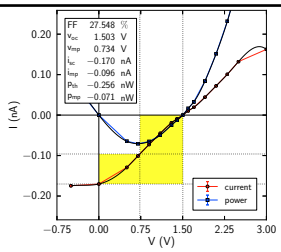
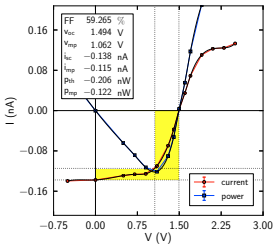
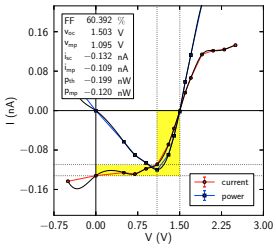
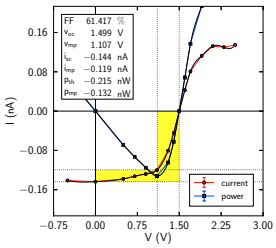
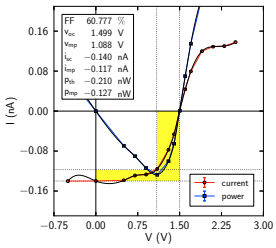
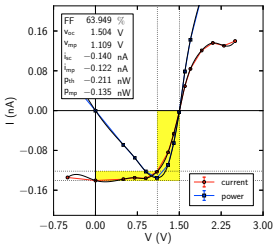
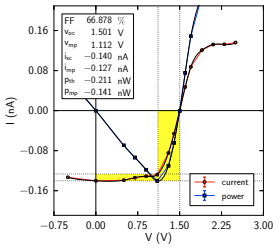
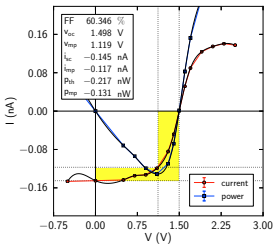
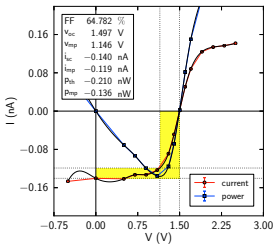
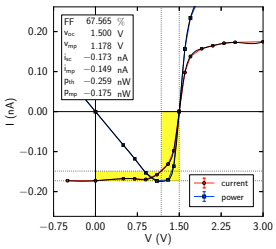
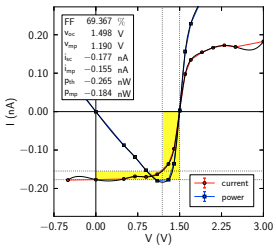
band4	run 0	run 1	run 2
$\sigma =$ 0.00 nm			
$\sigma =$ 0.75 nm			
$\sigma =$ 1.00 nm			
$\sigma =$ 1.25 nm			
$\sigma =$ 1.50 nm			
$\sigma = \infty$			

Figure B23: IV curves - band4

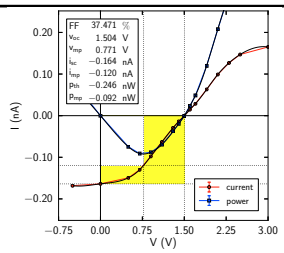
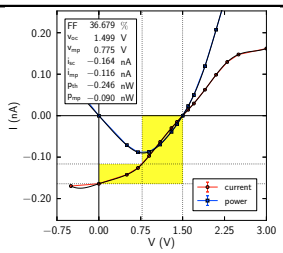
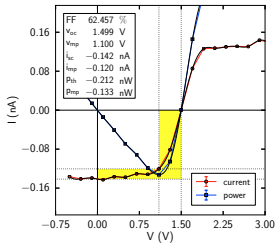
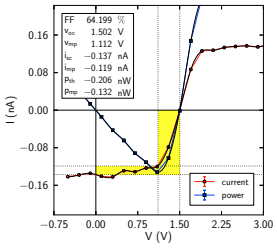
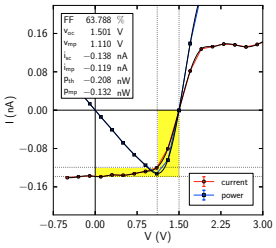
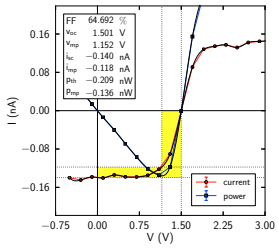
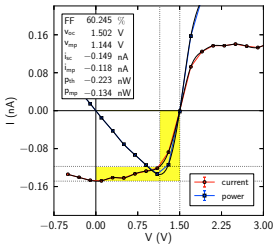
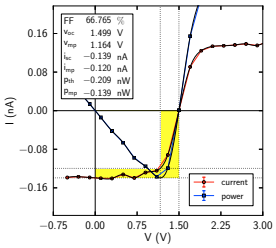
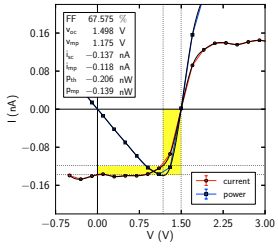
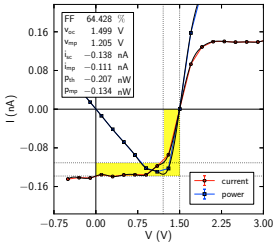
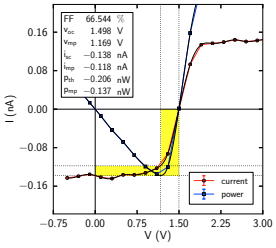
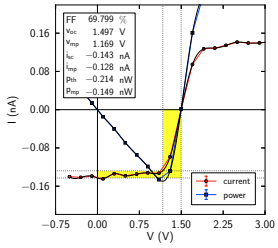
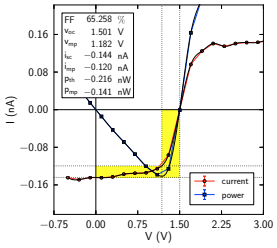
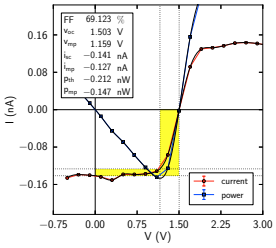
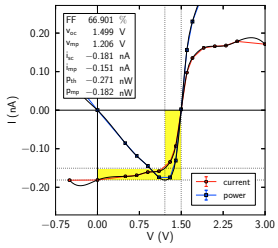
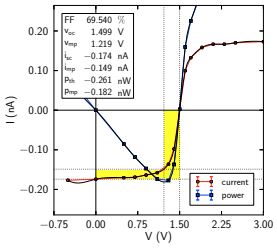
band8	run 0	run 1	run 2
$\sigma =$ 0.00 nm			
$\sigma =$ 0.75 nm			
$\sigma =$ 1.00 nm			
$\sigma =$ 1.25 nm			
$\sigma =$ 1.50 nm			
$\sigma = \infty$			

Figure B24: IV curves - band8

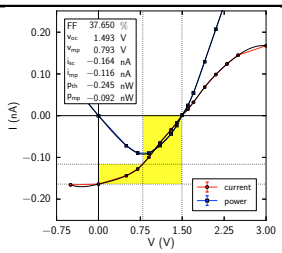
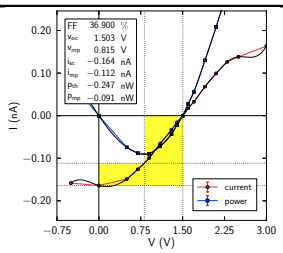
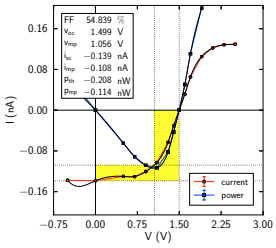
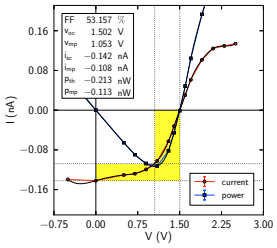
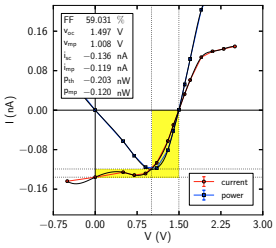
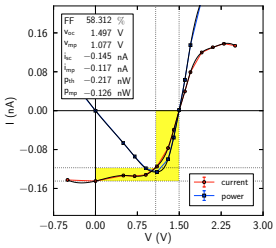
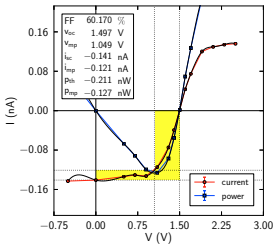
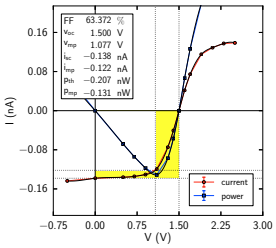
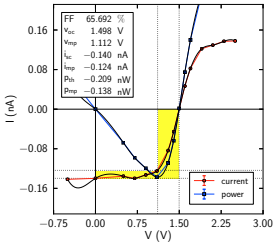
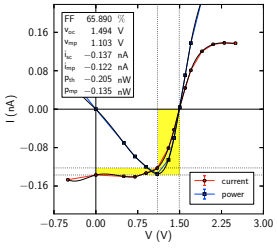
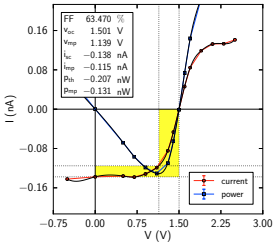
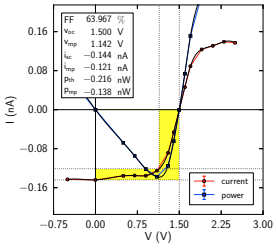
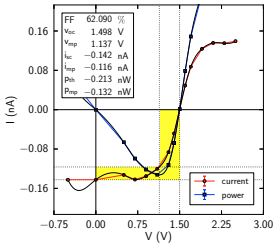
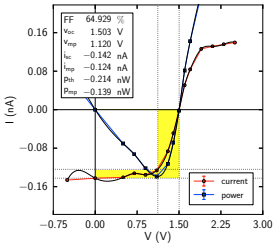
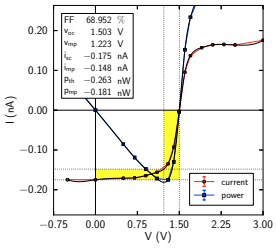
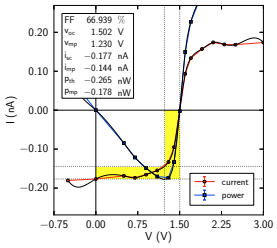
gyroid	run 0	run 1	run 2
$\sigma =$ 0.00 nm			
$\sigma =$ 0.75 nm			
$\sigma =$ 1.00 nm			
$\sigma =$ 1.25 nm			
$\sigma =$ 1.50 nm			
$\sigma = \infty$			

Figure B25: IV curves - gyroid

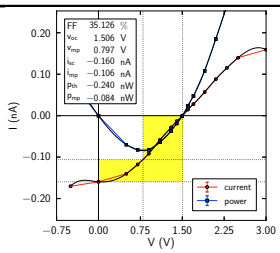
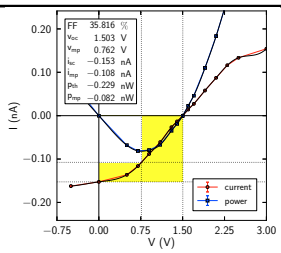
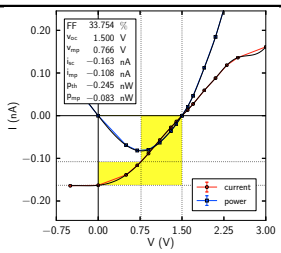
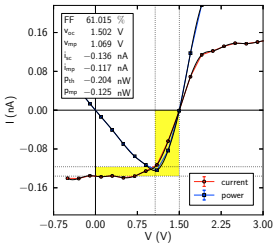
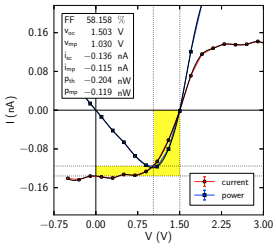
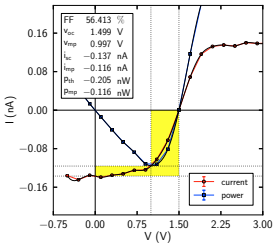
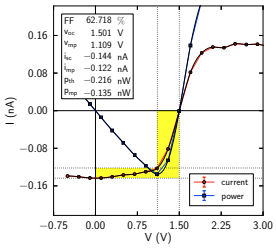
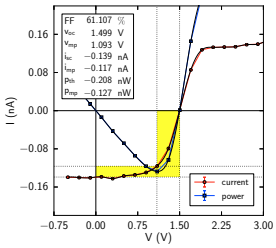
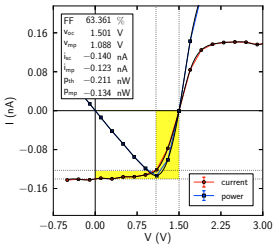
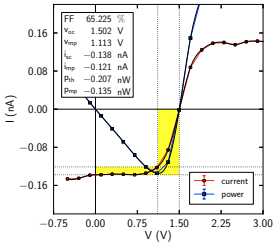
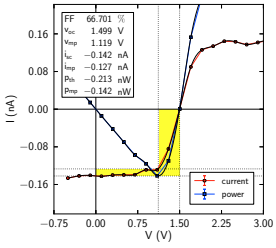
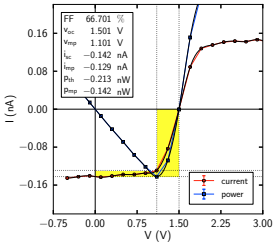
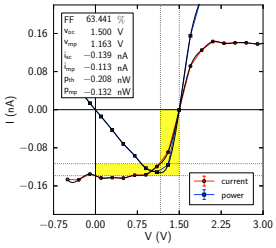
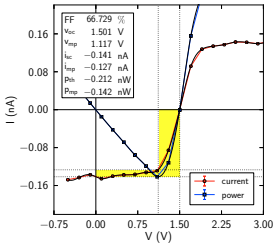
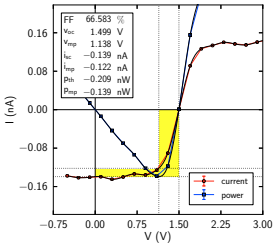
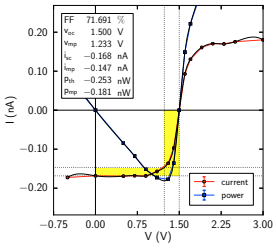
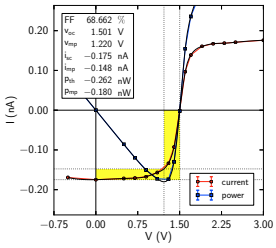
p-surface	run 0	run 1	run 2
$\sigma =$ 0.00 nm			
$\sigma =$ 0.75 nm			
$\sigma =$ 1.00 nm			
$\sigma =$ 1.25 nm			
$\sigma =$ 1.50 nm			
$\sigma = \infty$			

Figure B26: IV curves - p-surface

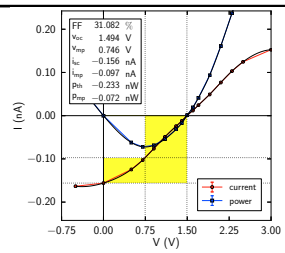
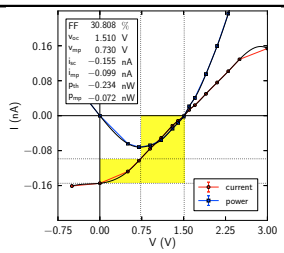
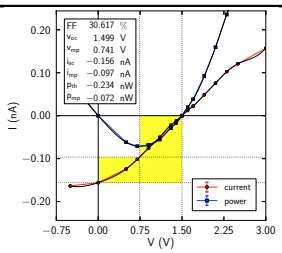
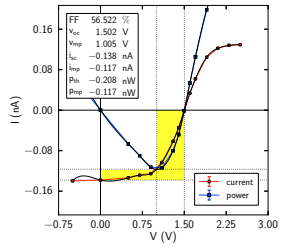
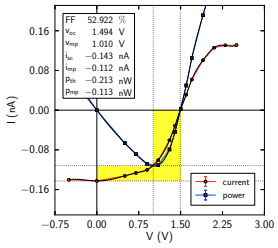
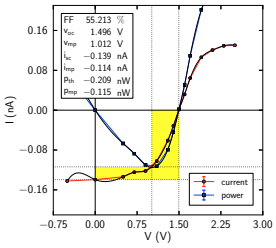
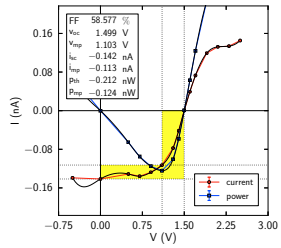
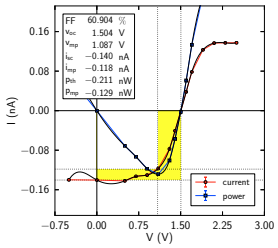
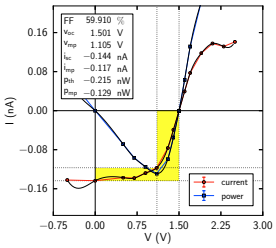
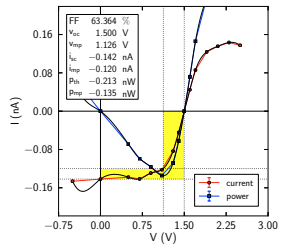
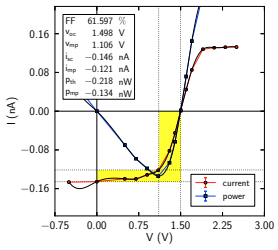
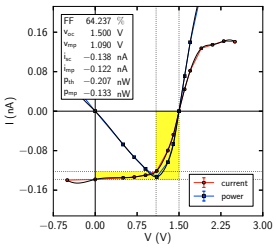
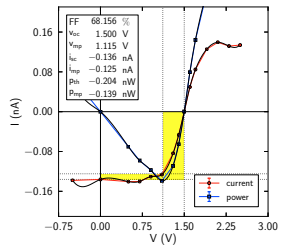
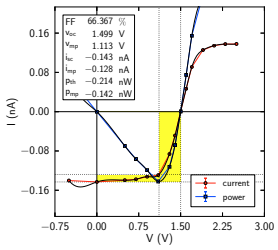
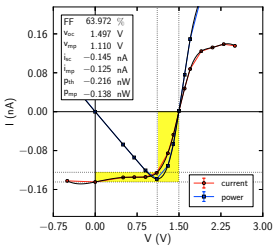
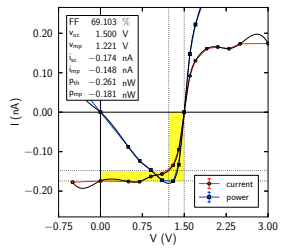
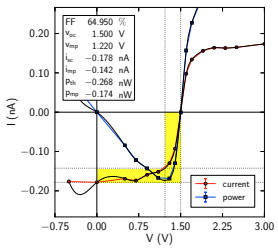
d-surface	run 0	run 1	run 2
$\sigma =$ 0.00 nm			
$\sigma =$ 0.75 nm			
$\sigma =$ 1.00 nm			
$\sigma =$ 1.25 nm			
$\sigma =$ 1.50 nm			
$\sigma = \infty$			

Figure B27: IV curves - d-surface

APPENDIX C

SOLVENT SHIFT MONTE CARLO

C.1 INTRODUCTION

In this section, the goals of this project are first presented. Then, an introduction to the problem is given.

C.1.1 Goals

The aim of this research was to extend the SSMC method,[89] proposed by Hixson et al. to the atomistic simulation of polymers in explicit solvents. The main hypothesis was that the SSMC method would significantly increase the phase space sampling of atomistic systems compared to existing methods. The project work was performed by first implementing a Monte Carlo simulation code for atomistic systems. The SSMC method was then examined for simple polymer systems in explicit solvent.

C.1.2 Problem

The monte carlo method is a powerful method with a lot of promise for the statistical mechanical simulation of polymers in chemistry. As an alternative to the molecular dynamics technique, Monte carlo offers the theoretical potential to explore phase space very efficiently. In practice, however, MD is often the method of choice for the simulation of chemical systems in explicit solvents.

The success of a Monte carlo technique intimately relies on the ability to produce reasonable random input. In the context of chemistry, one desires to create physically reasonable configurations of a system in a random way. The task is difficult and impractical when systems are dense. The short comings of the Monte carlo method in dense systems make the MD method more desirable.

The problem faced in this work was to apply a relatively new Monte carlo move, or way of generating random input, to dense systems.

C.2 BACKGROUND

In this section, background on the Monte carlo method is presented. The core principle of importance sampling is explained by first discussing a brute force averaging example. The Metropolis method to importance sampling is then introduced, and put into context by discussing Monte carlo moves in chemical simulations. Finally, the issues faced by the Monte carlo method in dense systems are explained.

C.2.1 Monte Carlo

A very well established and commonly used method for computer simulation is called Monte carlo. Though applicable in many different ways in many different subjects, all forms of Monte carlo generally use random numbers to solve a problem of interest. In the context of the chemistry, Monte carlo (MC) can be used to sample the configurational space of a system of particles, among other things. In particular, one can use the Monte carlo method to evaluate the thermal averages defined by statistical mechanics. Thermodynamic averages offer a solid connection between experimental observables and theory, making their efficient and accurate evaluation important. Consider the standard form for the average value of a quantity A, as prescribed by statistical mechanics.

$$< A >= \frac{\int d\mathbf{q}^N d\mathbf{p}^N A(\mathbf{q}^N, \mathbf{p}^N) P(\mathbf{q}^N, \mathbf{p}^N)}{\int d\mathbf{q}^N d\mathbf{p}^N P(\mathbf{q}^N, \mathbf{p}^N)} \tag{C.1}$$

In equation C.1, the average of A is simply the sum of the possible values of A weighted by the likelihood of A. The possible values of A depend upon the domain over which A can be defined. In statistical mechanics, A is any quantity that can be calculated by quantum mechanics or approximated by classical mechanics. Statistical mechanics gives the likelihood of A, in the canonical ensemble, to be the Boltzmann distribution. In a classical approximation, A takes on values which are defined by all possible positions and momenta of the particles in a system. This very large domain, of dimension 6^n, is often called phase space. For all but the most special cases, the thermodynamic average is a multidimensional integral that cannot be evaluated with traditional methods. The dimensionality is simply too large.

C.2.2 Brute Force

To understand how Monte carlo works, first consider the steps of a brute force monte carlo scheme. The steps can be summarized as follows.

1. domain
2. random input
3. evaluate function
4. average (over distribution)

In this scheme, one chooses random points in some domain, evaluates some function on that domain, and averages the result over some known probability distribution. The classic example of this process is determining the average depth of the Nile river. For example, the depth of water may be measured at random points on the surface of the earth. The depth of the Nile is then determined by averaging the depths subject to a probability distribution. The probability distribution describes how likely it is for a point on the earth to be in the Nile river.

The brute force scheme suffers from some obvious drawbacks. First, one must choose a great deal of points, the number of which can never be known, to obtain accurate results. That is, points not in the Nile River do not contribute to the average. Additionally, we are less likely to pick points the contribute significantly to the average as the domain increases

in size. Second, one must know the probability distribution fully. The brute force method illustrates the principles behind a monte carlo average, but is not used in practice.

C.2.3 Importance Sampling

Instead of evaluating integrals using the brute force scheme, the method of importance sampling scheme can be used. Importance sampling can be viewed as a variation on the brute force method. The steps are as follows.

1. domain
2. random input $\propto$ distribution
3. evaluate function
4. average (equally)

There are two main differences. First, the random input is chosen to be proportional to the probability distribution. Second, the average is not performed over the probability distribution. All points are considered to contribute equally to the average. In the context of the Nile river example, random points on the earth are more likely to be chosen in the region of containing the Nile River. One can average the resulting depths equally because the depths appear with their correct weight.

C.2.4 Metropolis Method

To apply the method of importance sampling, one must devise a scheme to produce random input proportional to a probability distribution. Metropolis et al.[13] were able to devise such a scheme by appealing to concepts in probability theory. The method provides a way to decide if freely chosen random input is reasonable (proportional to the distribution). It works by making a given state of a system depend only upon the previous state.

To understand how Metropolis produced this acceptance criterion, one first imagines a system with a set of possible states. First, there is a probability ρ_n that the system is in a given state n, called the current state. Second, there is a probability $\pi_{n\to m}$ that the system transitions from the current state n to another state m. This transition probability depends

only upon the current state. The probability to be in a new state is simply the probability to be in the current state times the likelihood to transition to the new state, $\rho_m = \pi_{n\to m}\rho_n$. By continually multiplying by the likelihood to transition between states, the probability to be in a new state approaches a well defined limiting distribution ρ^*. That is, successive applications of $\pi_{n\to m}$ to ρ^* leave the system at equilibrium. By satisfying this condition on $\pi_{n\to m}$, called balance, Metropolis was able to form the acceptance criterion. With the acceptance criterion, and the assumption that sampling is ergodic, the importance sampling method is easily employed.

Here is how it works. One can be more strict on the balance condition. At equilibrium, one can say that the net flux between two states is zero. The system is said to satisfy detailed balance.

$$\alpha_{m\to n}p_{m\to n}\rho_m = \alpha_{m\leftarrow n}p_{m\leftarrow n}\rho_n \tag{C.2}$$

Equation C.2 is a law dictating how states of the system change. The states of the system, or their location in the domain, are n and m. The probability that a system be in any particular state is ρ_m or ρ_n. The probability that a system *attempt* a transition from one state to another is given by $p_{m\to n}$ or $p_{m\leftarrow n}$. The probability that a system *accept* a transition from one state to another is given by $\alpha_{m\to n}$ or $\alpha_{m\leftarrow n}$. Note that the probability to transition between states, $\pi_{n\to m}$ or $\pi_{n\leftarrow m}$, is simply the product of attempting and accepting a transition.

To take advantage of equation C.2, Metropolis assumed that the probability to propose transitions was symmetric.

$$\alpha_{m\to n} = \alpha_{m\leftarrow n} \tag{C.3}$$

In statistical mechanics, the probability to exist in some state is simply the Boltzmann distribution for the canonical ensemble.

$$\rho_n = \frac{e^{-\beta U_n}}{Z\left(N,V,T\right)} \tag{C.4}$$

In equation C.4, U_n is the potential energy of state n and $Z\left(N,V,T\right)$ is the configurational partition function at constant volume, temperature, and particle number. With

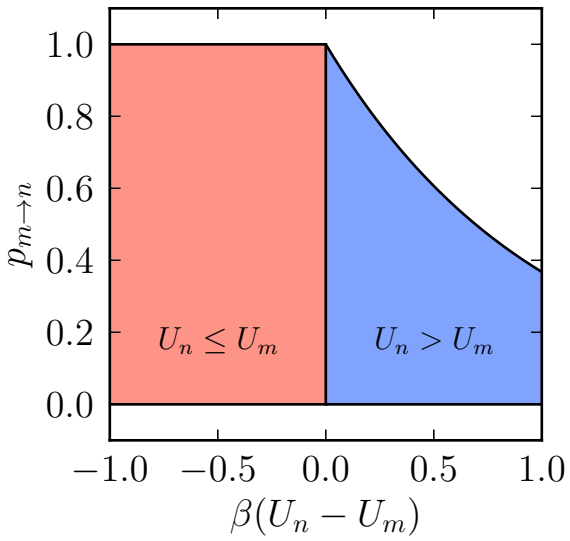


Figure C1: Metropolis method acceptance criterion.

these assumptions, detailed balance can be satisfied by making the probability of accepting a transition have the following form.

$$p_{m\to n} = min\left\{1, e^{-\beta[U_n - U_m]}\right\} \tag{C.5}$$

In equation C.5, one takes the probability to transition from state m to state n as the minimum of 1 or the Boltzmann factor describing the potential energy difference between the two states. The criterion is visualized graphically in figure C1. Whenever the energy of the new state lower that the energy of the current state, the new state is accepted with unit probability. However, whenever the energy of the new state is higher than the energy of the current state, the new state is accepted is probability that decays exponentially with the difference in energy between the states.

The Metropolis criterion allows random input to be generated according to the Boltzmann distribution, an essential step to importance sampling and the application of the monte carlo method to chemical systems.

C.2.5 Chemical Context

In chemical calculations, one can easily apply the importance sampling method discussed above. In this situation, one evaluates the thermodynamic integral C.1 for the chemical system. The domain, or limits of integration, are all possible configurations or locations of particles for the system. Of course, not all configurations contribute to the integral because the Boltzmann distribution says some configurations are not likely. And so, in the spirit of importance sampling, we approximate the integral by performing averages over configurations that appear with their proper weight. We generate these configurations by producing random configurations of the system and subjecting them to the Metropolis criterion described earlier.

C.2.6 Moves

A very powerful aspect of the Metropolis scheme is the absolute freedom one has in choosing random configurations. We are free to perform nonphysical transitions, allowing the possibility to rapidly sample the domain. There are several standard moves used to explore conformational degrees of freedom in molecular systems. The simplest possible move is a simple translation of a randomly chosen particle. However, such moves are only acceptable for all but the smallest displacements in systems where the atom is in a molecule. This is because of somewhat large energetic changes from multiply disrupted bond lengths, angles, and torsions.

The more common approach is to choose internal degrees of freedom for possible configurational changes. During these moves, all other degrees of freedom are kept frozen. Examples include bond length changes, bond angle changes, and torsion angle changes. Additionally, translations and rotations of entire molecules are often applied.

While some of these small scale moves work well, additional large scale moves can be much more advantageous. The phase space can be sampled extremely rapidly with the use of clever moves, especially when used in combination. Some examples include parallel tempering, configurational-bias monte carlo, hybrid Monte carlo, pivot moves, and geometric cluster algorithms. Torsional moves, in particular, greatly change the configuration of a

polymer.

C.2.7 Dense Systems

As powerful as monte carlo can be for statistical mechanical simulations, there is a reason MD is often preferred for the simulation of polymers in explicit solvents. It is unfortunate that, almost all moves that significantly change the configuration of the polymer are rejected in explicit solvent. This means that sampling is probably not ergodic, and accurate averages can not be obtained.

For example, consider performing a Monte Carlo simulation of a polymer in explicit solvent in the canonical ensemble. Trial moves that change the configuration of the polymer could easily generate overlapping particles. Overlapping particles are a major problem in Monte carlo. Recall that the Metropolis criterion leads to the probability of accepting a trial move depending upon the energy difference between the old state configuration and the new proposed state. The energy of these states is calculated using a parameterized force field which undoubtedly must contain some inter-particle pair potential. Computing the energy of a configuration that contains overlapping particles would clearly cause a pair potential, such as the Lennard-Jones potential, to become large. Thus, the energy of such a configuration would be large compared to the configuration it was generated from and such a trial move would be rejected.

C.3 METHODS

In this section, the Solvent Shift Monte Carlo move and its implementation are described. A brief overview of the simulation program is given.

C.3.1 Solvent Shift

The solvent shift monte carlo move aims to alleviate the problem of overlapping particles by combining an internal coordinate move for a solute with a cluster rotational move for any

overlapping solvent particles. In this section, the steps of the SSMC move are outlined.

The solvent shift monte carlo move involves small and large internal torsional moves of a solute polymer followed by an iterative rotational cluster move of surrounding solvent to remove particle overlap. The steps might be summarized as follows.

1. choose random dihedral
2. choose random angle
3. choose rotation method
4. find atoms that move
5. generate rotation matrix
6. perform dihedral rotation
7. identify overlapping particles
8. rotate overlapping solute
9. iteratively identify and rotate solute
10. propose or terminate

One must choose a random solute dihedral angle to change. While a dihedral angle can be defined between any four points in space, for a molecule it is common to define dihedral angles between four atoms connected along a line. If the atoms 1-2-3-4 are connected along a line, then a dihedral angle is defined between them. This type of dihedrals is called a proper dihedral angle, and turn out to be the most convenient types of dihedral to perform Monte carlo moves with. A list of all proper dihedral angles for a system is stored, and a randomly generated integer in the index space of this list is generated in order to choose the dihedral angle. Though any proper dihedral angle may be chosen this way, it is important to note that proper dihedral angles near the core of a large polymer are generally more tricky to accept in dense systems.

One must choose a random angle. A random floating point number is first generated between negative and positive one half. This number is then multiplied by a preset desired maximum angular displacement. The end product is a number between negative and positive the maximum angular displacement. This method is necessary to ensure that the probability of proposal remains symmetric in the Metropolis criterion. Often, the preset

desired maximum angular displacement is allowed to optimize during equilibration periods of the simulation. Another option for angle choosing exists if one desires angles to be within a certain range. A pre-chosen angular range and average value is a first specified. A random floating point number is generated between negative and positive one half. This number is then multiplied by the pre-chosen desired range. Another random number between zero and one is then generated. If this number is larger than one half, the desired average angle is added. Otherwise, the desired average angle is subtracted.

One must choose how the molecule is to rotate. For this work, one side of a dihedral is kept frozen while the other side rotates. Consider a proper dihedral defined between atoms 1-2-3-4. One may choose atoms 1-2-3 to remain fixed while atom 4 (and all atoms connected to atom 4) rotate. Conversely, one may choose atoms 2-3-4 to remain fixed while atom 1 (and all atoms connected to atom 1) rotate. Another option for rotation might be to fix atoms 1-2 and rotate atoms 1 and 4 (and their connections) by half angles. This option is an unnecessary complication for the SSMC move. The side of the dihedral to rotate is randomly chosen by generating a random number between zero and one. If this number is greater than one half, then a rotation axis is defined between atoms 2-3, while atom 4 and its connections are targeted for rotation. Otherwise, a rotation axis is defined between atoms 3-2, while atom 1 and its connections are targeted for rotation.

One must find all atoms that rotate based on the axis chosen and the side of the dihedral that moves. A list of all connections in the system is stored. An algorithm is devised to use this list to generate all atoms that should move from these connections. A recursive function is used with care for this algorithm.

One must generate a rotation matrix based on the axis chosen and the angle chosen. Various methods exist for producing rotation matrices. The method of quaternions was used to produce the rotation matrix in this work. Quaternions are simply four dimensional vectors of unit length, i.e. points on the surface of a four dimensional hyper-sphere. The space of rotations in three dimensions can be shown to be equivalent to the surface of this hyper-sphere. A quaternion can be generated easily from an axis and angle using a simple formula. The usefulness of quaternions lies in the simple arithmetic needed to generate the rotation matrix (no expensive sin or cos functions need to be evaluated). The inverse of the

rotation matrix, also needed, is simply its transpose because rotation matrices are unitary.

With the axis, angle, moved atoms, and rotation matrix, a rotation move on the solute is performed. One must choose an atom to use as a reference for rotation calculations. An atom in the proper dihedral axis is chosen as a reference point. For example, atom 3 in dihedral angle 1-2-3-4 is chosen when atoms 2-3 define the axis and atom 1 remains fixed. This step is necessary because rotation matrices can not be trivially applied to atoms restricted to move under periodic boundary conditions. Correct application of rotation matrix is ensured by first calculating the minimum distance vectors between the reference atom and atoms targeted to move. This may be visualized as switching to a minimum image frame of reference about the reference atom. The rotation matrix is applied to the set of calculated minimum image vectors. After rotation, the rotated vectors become the new moved atom positions by adding the reference atom vector and fixing periodic boundaries. Everything described up until this step is essentially equivalent to a simple torsional move.

After the solute rotation, particles that overlap are identified. This is done by looping over the neighboring atoms that moved and calculating atom-atom distances. If the distance is less than some tolerance, the molecule an overlapping atom belongs to is targeted. Should the molecule identified as overlapping is the solute molecule itself, the move is rejected immediately.

The identified overlapping molecules are simply rotated in the reverse direction (using the transpose rotation matrix) about the dihedral axis. The process is repeated. Molecules are continually identified until no overlapping particles are identified. Should this condition never be met, the move is rejected after some set number of iterations. During this process, intimate tracking of molecules that move is kept.

The configuration reached is then subjected to the Metropolis criterion kept earlier. Energy change is calculated for particles that moved only. No bond angles or bond lengths change during a dihedral move. No non-bonded terms change between particles that do not move. This includes non-bonded forces between atoms that moved and neighboring fixed atoms, improper dihedrals that changed, and proper dihedrals that changed. List of proper and improper dihedrals that may change during a dihedral move are pre-computed at the beginning of a simulation.

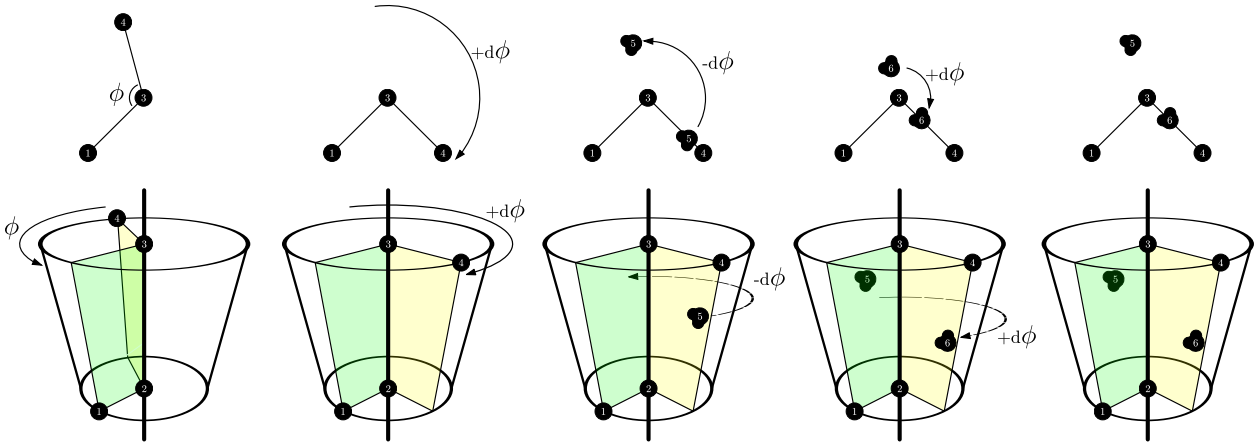


Figure C2: SSMC move steps.

C.3.2 Programming

This section contains a brief discussion of the simulation program written to investigate the SSMC move. The code was written using the C++ programming language (1998 standard) using an object oriented approach.

The code uses a simulation run object to perform a Monte carlo simulation. The simulation run object orchestrates the behavior of move objects and an energy calculator object. The simulation run object implements the main Monte carlo loop, handles input, and reports output. Input information included a list of atoms, molecules, molecular terms (bonds, angles, dihedrals, non-bonded information, connect information), an initial configuration, and program options. This information is rigorously checked for consistency and stored in many objects types. The object types include points, matrices, atoms, molecules, bonds, angles, proper dihedrals, improper dihedrals, simulation box, topology, and simulation run. Output information includes a configuration trajectory as well as various simulation parameters and statistics.

The move objects perform standard Monte carlo moves. There is a base move class through which various specific moves are derived. The base class knows how to apply the

Metropolis criterion, implements the accept and reject behavior of a move, and suggests a virtual function to perform a given move. Derived moves implemented were molecule translation, molecule rotation, bond displacement, angle bending, small and large proper dihedral moves, and of course the SSMC move.

The energy calculator object knows how to calculate total energy and energy changes by communicating with a topology object. The topology object contains a list of atoms, molecules, bonds, angles, dihedrals, and non-bonded information. The topology object knows how to report potential contributions from its various terms. The energy calculator object knows how to sum up these contributions to obtain total energies and energy changes. Additionally, the energy calculator object performs non bonded calculations by communicating with a cell linked list object. The cell linked list object informs the energy calculator about how atoms are arranged into neighboring cells in the simulation. Finally, a simulation cell object allows the correct application of periodic boundary conditions.

C.4 RESULTS

C.4.1 Implementation Test Systems

To implement the move a simple (small) test system was used to ensure that the move was working as intended. The test system was a alkane molecule and a single water molecule. The AMBER force field, TIP4P water model, and OPLS charges were used. The alkane was in a extended conformation. The water molecule was placed so that it would be overlapping with the alkane if a central proper dihedral angle of the alkane were changed by 180 degrees. The Lennard-Jones contribution to the potential energy as a central dihedral is rotated for this system is shown in figure C3A.

It is clear from the figure that the energy spikes when the dihedral angle is changed by by 180 degrees. Any torsional moves near this amount are rejected due to large energy changes.

For this system, SSMC torsional moves were performed with 100 percent frequency. Only torsional changes of 180 degrees were attempted and only the central dihedral was selected

for this test. The move showed 100 percent acceptance. The simulation trajectory indicated that, as the torsional angle of the alkane changed by 180 degrees, the water molecule move rotated about the dihedral axis by negative 180 degrees, as expected. When torsional moves of a random magnitude are chosen, the acceptance drops slightly. This is because small angle changes that lead to overlap than can not effectively be removed.

This test did not indicate how the move would begin to behave in denser systems. So, two more water molecules were added to the simple alkane system. The water molecules were placed such that, should the alkane dihedral change by 90, 180, or 270 degrees there would be overlap. The Lennard-Jones contribution to the potential energy as a central dihedral is rotated for this system is shown in figure C3B. It is clear from the figure that it would be very difficult to accept a torsional move in this system. When torsional moves of only 90, 180, or 270 degrees were attempted, the systems showed 100 percent acceptance. When torsional moves of a random magnitude are chosen, the acceptance drops slightly.

Finally, the number of waters leading to possible overlap was increased to seven. The water molecules were placed such that, should the alkane dihedral change by any multiple of $\frac{2\pi}{7}$ there would be overlap. Unfortunately, the both random and fixed angle attempts showed very poor acceptance rates in such a system. The arrangement of water molecules chosen was most likely unfavorable.

C.4.2 Simple Alkane System

This section describes results for a simple polymer in a well solvated water box. A 10 carbon alkane was simulated using the AMBER force field in TIP4P water with OPLS charges. The simulation was in a roughly $3\,\mathrm{nm}^3$ cell with around 1000 water molecules. A water box was first prepared and equilibrated using molecular dynamics at constant pressure. The alkane was inserted into the box and equilibrated to ensure the proper box size and density of water was reached.

Taking the initial configuration of the system to be the well equilibrated polymer + water box, various Monte carlo simulation were performed at constant volume. Equilibration was also performed during Monte carlo to allow move parameters to optimize. In addition

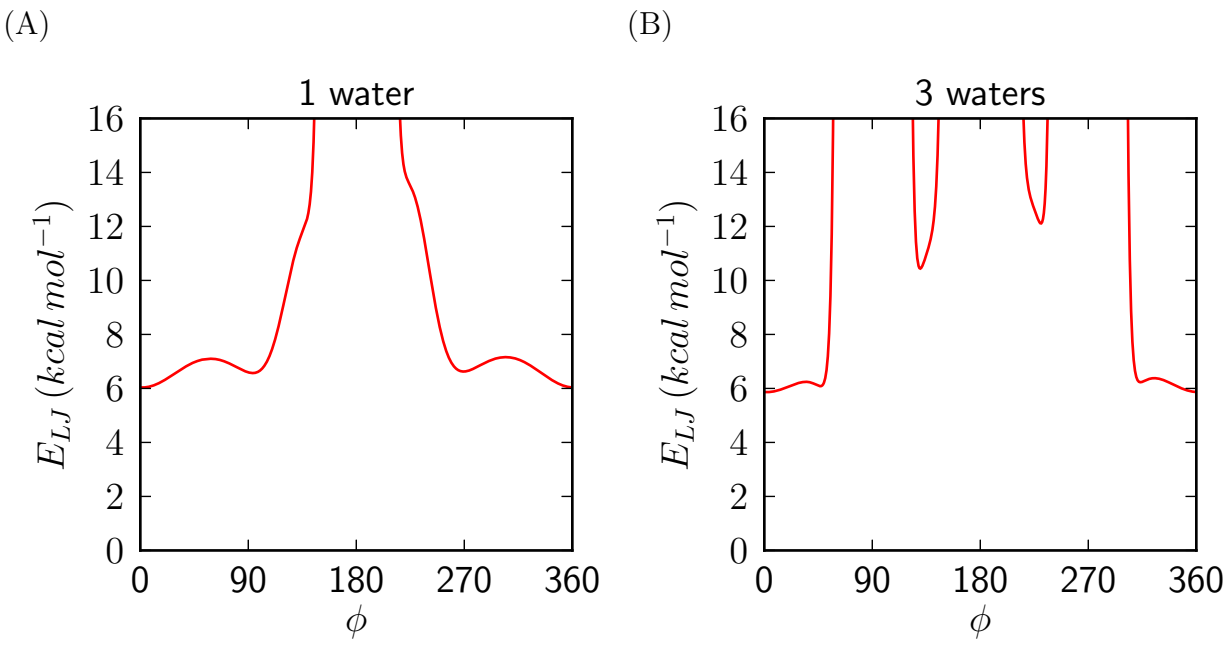


Figure C3: LJ potential during SSMC move.

to SSMC moves, translation, rotation, bond, angle, and standard torsional moves were attempted with equal frequencies. Different distance overlap criteria and torsional angle choosing methods were attempted for the SSMC move. Small moves were attempted by choosing random torsional angle changes such that $0 \leq |\theta| \leq 60$ or $0 \leq |\theta| \leq 120$. Large moves were attempted by choosing random torsional angle changes such that $120 - \delta \leq |\theta| \leq 120 + \delta$ or $60 - \delta \leq |\theta| \leq 60 + \delta$. The acceptance rate, configurational trajectories, and proposed energy changes were examined to diagnose the SSMC move.

Distance overlap criteria was examined between 1.0 and 5.0 Å. While values between 1.0 and 2.0 Å did not seem to differ, values of 3.0 Å and above would cause the move not to converge on a proposed configuration. With high distance overlap criteria, it is too difficult to reach a no overlap state during iterative rotation of solvent molecules. Values below 1.0 Å did not effectively target water molecules for movement.

For large moves, where proposed angle changes are forced to be around a desired value, SSMC were not accepted. The SSMC acceptance rate for small moves, where proposed

angle changes may be in a range that includes zero, showed an acceptance rate of roughly 5.0 percent. However, the acceptance rate for simple torsional moves when angles were chosen with the same method was also roughly 5.0 percent.

Analysis of configurational trajectories showed that, when a torsional angle did change, the angle change was not significantly large. This suggests that only small torsional angle moves were being accepted. If a torsional angle move is small no water molecules would be moved during a SSMC proposal. Sometimes a large change would be accepted, even in the large move cases. However, these large changes never occurred for central dihedral angles of the alkane. Only small groups near ends would rotate, something not difficult to accomplish with standard torsional moves. None of the accepted moves seemed to involve rotation of solute molecules.

The energy changes of proposed moves were collected. Figure C4 shows proposed energy changes histograms in three cases. Shown on the left are energy changes for moves proposed with $0 \leq |\theta| \leq 120$. Some proposed configurations have energy changes close to zero and less than zero. These configurations are most likely the small angle and end group rotations discussed earlier. The majority of energy changes occur at energies much to high to be accepted.

Shown in the middle are energy changes for moves proposed with $60-10 \leq |\theta| \leq 60+10$. The scale on this graph is extended to show that no moves are proposed with energies negative or near zero. Almost all changes are clustered at very high energies. It is possible that effective rotation of solvent can not be achieved at this angle.

Shown on the right are energy changes for moves proposed with $120-10 \leq |\theta| \leq 120+10$. There are fewer negative energy counts compared to the $0 \leq |\theta| \leq 120$ case. The peak near zero most likely corresponds to end group rotation changes.

The energy change histograms indicate that the move is suffering in dense systems. The energy changes are too high to remedy by parallel tempering methods. If we desired a 1.00 percent acceptance probability, then the ratio of a proposed move energy change to kT should be around 4.61. This means energy changes above 4.61 kT are not very likely to be accepted. With kT at 300 K being 0.592 kcal/mol, energy changes above 2.73 kcal/mol are not very likely to be accepted. Even at 1000 K, a move will not be accepted with energy

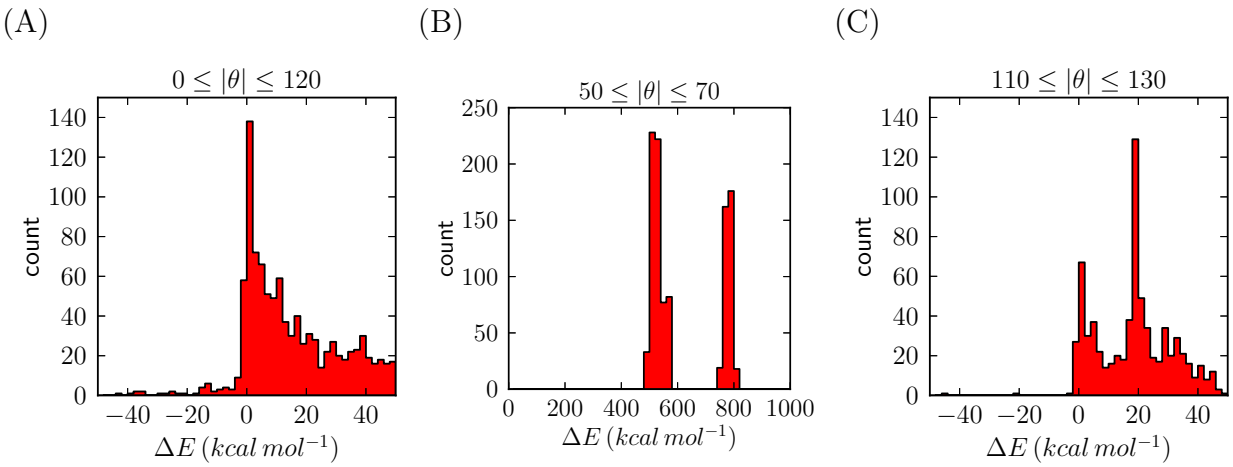


Figure C4: SSMC move energy changes in realistic density.

changes above 9.09 kcal/mol.

C.4.3 Betapeptide Various Densities

This section describes results for a betapeptide polymer in a solvated water box at various densities. The idea was to see if the SSMC move could be accepted at various densities. The betapeptide was βVal-βAla-βLeu-βAla-βVal-βAla-βLeu. The AMBER force field, TIP4p water model, and derived charges were used. The simulation was in a roughly 6.5 nm^3 cell with around 9000 water molecules. A water box was first prepared and equilibrated using molecular dynamics at constant pressure. The betapeptide was inserted into the box and equilibrated to ensure the proper box size and density of water was reached. After equilibration, molecular dynamics was run at constant volume for a short time. Nine initial configurations were taken from the constant volume molecular dynamics trajectory.

Taking the initial configurations of the system, various Monte carlo simulation were performed at constant volume. Solvent molecules were removed randomly to produce initial configurations at lower densities. Systems were produced with 2000, 4000, and 6000 solvent molecules. These densities are far from the realistic 9000 solvent molecules. Each

configuration underwent 2 million Monte carlo moves.

In addition to SSMC moves, translation, rotation, bond, angle, and standard torsional moves were attempted with equal frequencies. SSMC moves chose angles such that $0 \leq |\theta| \leq 120$. For standard torsional moves, angles were chosen such that $0 \leq |\theta| \leq 30$. In this way, any large torsional change observed is due to the SSMC move. The distance overlap criterion used was 1.0 Å. The acceptance rate, configurational trajectories, and proposed energy changes were examined to diagnose the SSMC move.

The acceptance rates in the presence of 2000, 4000, and 6000 solvent molecules were all close to 9.0 percent. Observation of the configurational trajectories showed that some of these acceptances were non central dihedral angles or small torsional changes. However, there were noticeable large changes in central dihedral angles. If or not solvent molecules moved during these transitions was not kept track of. It is possible that the SSMC move worked in the reduced density systems. However, it is also possible that no solvent molecules moved during the SSMC moves. Simulations with no SSMC moves and large torsional moves should be performed.

C.5 CONCLUSIONS

C.5.1 Struggles

The SSMC has low to no acceptance rate in the systems studied. This is because proposed configurations have large energy changes compared to previous configurations. Large energy changes cause the moves to be rejected. Why were the proposed configurations higher in energy?

Consider what the SSMC is saying about energy of a system. SSMC says there is a distance overlap criterion that quantifies if or not molecules are overlapping. Should the distance between two molecules be too low, SSMC decides the system must be somewhere too high on the potential energy surface. The solution to finding a lower energy configuration is to march a specific amount along a generalized coordinate. The generalized coordinate

is the angle the water molecule makes about a dihedral axis. The specified amount is the randomly chosen angle change for the move. There is no special rule saying that we should reach a local minimum by performing this transformation. It is almost like saying walking a random direction by some random amount will always bring a system to a local minimum. The entire idea of rotating molecules around a dihedral axis to approach a low energy configuration is wrong. This is especially true as the potential energy surface becomes less coarse grained. Systems at lower densities accept moves because the potential energy surface becomes less spiky.

The distance overlap criterion is probably not correctly quantifying overlap. Overlapping means something along the lines of being high on the potential energy surface. The orientation of solvent molecules is going to play role in the energy. The potential around a water molecule is far from spherically symmetric. How can distance alone correctly quantify overlap?

The idea of a single distance criterion for all atoms in the stem is completely wrong. Different atoms are different sizes and have different strengths of interaction. A hydrogen atom can get closer to other atoms than an oxygen atom can. The Lennard-Jones sigma and epsilon are different for every single possible combination of atoms in a system. Therefore the average distances between atoms will be different depending upon the atoms involved.

The distance criterion required to get convergence was most likely too small to begin with. Consider the distance between waters in a realistically dense system. It might be somewhere on the order of 3.0 Å according to the radial distribution function for water. Choosing an overlap criterion this high, however, causes the SSMC move not to converge on a non-overlapping condition. Above a certain value, there comes a point where spheres drawn around atoms intersect no matter their orientations about the dihedral axis.

It is safe to conclude that the SSMC does not work for systems with complicated potential energy surfaces. The SSMC produces unrealistic configurations of the system. This is a result of some bad assumptions made about the potential energy surface. The SSMC move needs a lot of work in order to be feasible.

C.5.2 Future Work

The SSMC move needs a more realistic way of generating lower energy configurations from configurations with overlap. While rotation of solvent molecules about the dihedral axis does not produce configurations of low energy, the transformation might be a good enough start. The fact is we know that after a dihedral angle change of the polymer, there is a huge gaping hole where part of the polymer used to be. Rotation about the dihedral axis can put the overlapping particles in this hole. While the resulting configuration is almost never of lower energy, it can be made lower energy by employing minimization techniques. There are numerous minimization techniques, and they can be combined. To speed things up, not all the particles in the system would have to move during the minimization. In the end, a minimization would be a small price to pay to generate such a large change in configuration of the stem. Another possibility would be to avoid the rotation. Instead, simply identify the overlapping particles and grow them in the hole region with random orientations. Place them with probability based on the Boltzmann factor. This could generate low energy configurations.

The SSMC move needs a more realistic way of identifying overlapping particles. It is difficult to correctly identify if particles overlap based on a single distance. One could compare distance calculations between atoms to some fraction of the Lennard-Jones sigma for the pair. This comparison might give a better indication of overlap. The idea of using an angular criterion to identify overlap is somewhat impractical. Quaternions could be used to quantify the orientation of molecules, however, it is unclear what a good criterion for proper orientation would be.

The SSMC move needs to focus on internal angles. When proposing small or large angle changes on dihedrals near the ends of molecules the SSMC move is not accomplishing anything an ordinary torsional move could not already perform. One can restrict the dihedral angles chosen to be core angles of the polymer during SSMC moves.

The SSMC move needs to focus on large angle changes. Small torsional angle changes do not contribute to large conformational changes of the system. Standard torsional moves can handle the small torsional changes. SSMC does not need to be proposing moves with

small angle changes. Focus should be put on optimizing and making the move work with large angle changes.

Other move types could benefit from SSMC if it was working. SSMC does not necessarily need to be combined with torsional moves, although torsional moves are ideal for changing conformations greatly. Other moves, such as angle changes, could benefit from the removal of overlapping particles in proposed configurations.

SSMC could be combined with more advanced techniques. Once the SSMC generates new configurations with a non-negligible acceptance rate, the possibilities are endless. The move can be combined with any other valid Monte carlo move to make a valid Monte carlo scheme.

APPENDIX D

PNIPAAM

D.1 INTRODUCTION

In this section, an introduction to the polymer PNIPAAM is given. An attempt is made to discuss some previous experimental and theoretical information on PNIPAAM and derived copolymers. Reasons why PNIPAAM is worthy of study and the goals of this theoretical work are presented.

D.1.1 Goals

The goal of this project was to theoretically probe the conformations of the polymer poly(n-isopropylacrylamide) (PNIPAAM). In particular, the role of water in the temperature dependent collapse of the polymer was to be investigated. These goals were to be met using the molecular dynamics method. The project was motivated by experimental studies that estimated the number of waters withing amide and carbonyl groups on PNIPAAM, and suggested a mechanism of collapse.[90] Previous MD results on the simulation of n-isopropylacrylamide (NIPAAM) polymers by other parties[91] were to be confirmed and extended upon.

An extension of the project was to investigate copolymers of PNIPAAM made with the comonomers acrylic acid (AAC) and hydroxyethyl methacrylate poly-(trimethylene carbonate) (HEMAPTMC). In particular, how the behavior of collapse changes and the effect of

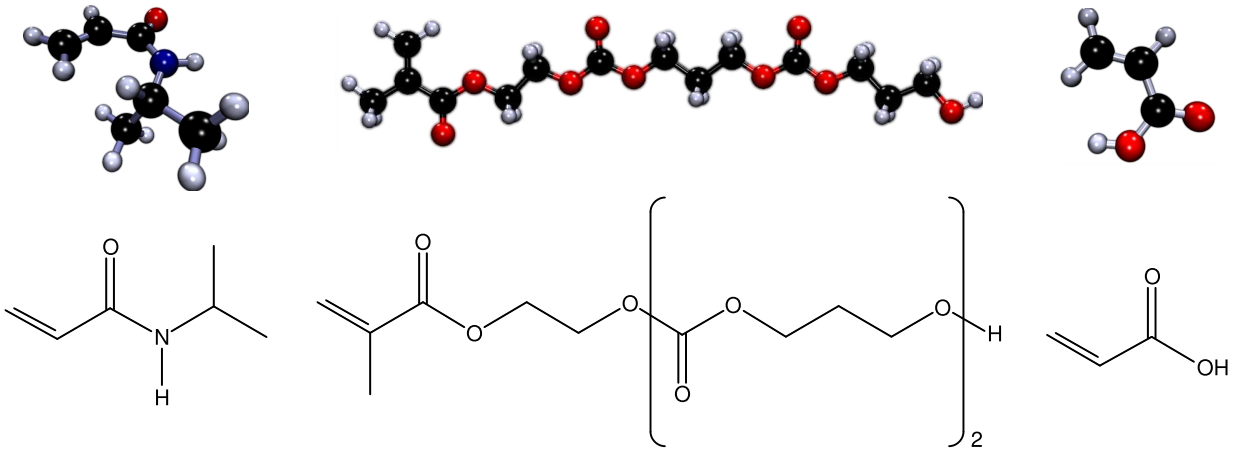


Figure D1: NIPAAM, HEMAPTMC, and AAC.

pH on col-polymer systems was to be examined. These goals were also to be met using MD. The extension was motivated by a biological application of the copolymer system and related copolymer hydrogels to mydocardial infraction repair.[92]

D.1.2 Structure

When talking about the structure of a polymer, it is convenient to divide the discussion between chemical structure and geometrical structure or conformation. The chemical structure is the information about the bits and pieces that make up the polymer and how they are molded together. The conformation is how the molded pieces are arranged in space at any given moment in time. While the chemical structure plays an important role in determining the geometrical structure, the chemical structure is not the only factor. Environmental factors also play a huge role. The environment includes information about the state of the system such as temperature or pressure. Other chemical species may interact with the polymer and different regions of the polymer may even interact with each other. These intermolecular and intramolecular interactions affect the overall geometrical conformation.

Additionally, it must be pointed out that even at equilibrium a polymer does not exist as a single geometrical conformation. There is and must be an ensemble of conformations that give rise to the observed average conformation.

First, the chemical structure of PNIPAAM is discussed. PNIPAAM is a polymer made from the monomer NIPAAM. Chemically, NIPAAM monomers contain an isopropyl group, an amide group, and a vinyl group. The amide group consists of a carbonyl bound to a secondary amine. The polymerization of NIPAAM occurs at the vinyl group. Upon polymerization, the vinyl carbon bound to amide can become chiral. This allows the polymer to exhibit tacticity. The relative steriochemistry of monomers can be an ordered pattern or random. If the pattern involves alternating steriochemistry, the polymer is called syndiotactic. If the pattern involves repeating steriochemistry, the polymer is called isotactic. If the pattern is random, then the polymer is atactic. Polymers of PNIPAAM should be atactic in most situations.

Second, factors that could influence the geometrical structure of PNIPAAM are discussed. Special emphasis is put on the role of water in the environment. The chemical structure of PNIPAAM allows for self interaction and intimate interaction with water. Specifically, the amine hydrogen along with the carbonyl oxygen allow NIPAAM monomers to hydrogen bond to one another. Additionally, the amide brings hydrophilic character to PNIPAAM in aqueous solution. The carbonyl carbon can hydrogen bond to water hydrogens and the amine hydrogen can hydrogen bond to water oxygen. Multiple hydrogen bonds can occur to the carbonyl. For example, one or two hydrogen bonds may form on the carbonyl. The structure of the polymer as well as the presence of other water molecules will determine how many hydrogen bonds are allowed sterically. The hydrogen bonds themselves may serve as nucleating sites for the formation of highly structured water clusters. The formation of water clusters can provide an extra enthalpic diving force for formation of hydrogen bonds.

The isopropyl and vinyl groups bring hydrophobic character to PNIPAAM. That is, hydrogen bonds are not expected to form near isopropyl groups or vinyl groups. Waters forced to remain near an isopropyl group pay an entropic penalty and have no enthalpic stabilization from hydrogen bonding. Overall this is an unfavorable situation. In the end, the interplay between enthalpy and entropy govern the conformation of PNIPAAM, the

amount of hydrogen bonding, the formation of water clusters, and structure of the solvent.

D.1.3 Phase Behavior

When aqueous solutions of PNIPAAM are made, the system exhibits phase behavior. There is a phase where the solution is cloudy with low transmittance and a phase where the solution shows high transmittance. The abrupt phase transition, which is independent over a range of concentrations and polymer molecular weights, occurs around 32 °C. The reason why a phase transition occurs is known to be the result of expansion and collapse of PNIPAAM polymers with changes in temperatures. At low temperatures PNIPAAM exists in a solvated extended state. At high temperatures PNIPAAM exists in a collapsed insoluble state.

The system comprised of PNIPAAM and water exhibits what is known as a lower critical solution temperature (LCST). The LCST is simply a critical point on the water-polymer-mixture phase diagram. Below this temperature, phase boundaries cease to exist. That means the compounds are miscible in all proportions. This is the point where the solution has high transmittance. Any temperature above the LCST will induce phase separation or precipitation. This is the point where the solution is cloudy with low transmittance. The value of the LCST, and therefore the phase transition, is also known to depend upon other thermodynamic variables such as pH or salt concentration. Typically, lowering the pH or increasing the ionic strength of the solution will decrease the LCST. Chemical substitution of monomers to form PNIPAAM copolymers will also alter the LCST.

The LCST behavior exists in other forms of PNIPAAM. As polymers often do, PNIPAAM can form a polymer network. This happens when the polymer chains physically aggregate in a nonlinear fashion, creating a web of entangled polymer chains. PNIPAAM actually forms a covalent polymer network; chains are bound to each other covalently. The network may have holes and pockets. Water or other substances can fill the holes. When a PNIPAAM polymer network is expanded into it full volume by a fluid such as water, it is by definition a gel. If the fluid or swelling agent is water, PNIPAAM is said to form a hydrogel. A PNIPAAM hydrogel will also exhibit a phase behavior near the LCST. The hydrogel will change dimensions macroscopically, shrinking with increases in temperature. Volume phase

transition is another name for the phase transition in this situation.

Sometimes, PNIPAAM is classified as a thermally responsive polymer because of the phase behavior. As was mentioned above, PNIPAAM basically becomes less soluble in water as the temperature is increased. Often compounds become more soluble when temperature of the system is increased, making this behavior counterintuitive. Thermo-responsive polymers, and more generally polymers which respond to environmental variables have many practical applications to be discussed below.

D.1.4 Applications

PNIPAAM can be used in drug delivery applications. For example, Dufresne et al. synthesized several PNIPAAM copolymers containing hydrophobic moieties and pH-sensitive units.[93] They showed the copolymers to be a potential safe and efficient alternative for the solubilization of hydrophobic drugs in drug delivery applications. In their work, the copolymers would self-assemble in aqueous solution into micellar structures whose stability depended upon pH. The root of this behavior was a result of the phase behavior of PNIPAAM. Hydrophobic drug molecules were then enclosed in the hydrophobic region of the micelle. Since the micelle was only stable at certain pH, it would destabilize (releasing the drug) when the pH changed (in the interior of a tumor cell). Particularly interesting, these PNIPAAM copolymers, already known to exhibit low cytotoxicity, where shown to be more effective in delivering the photosensitizer aluminum chloride phthalocyanine used in photo-radiation treatment of cancer in an implanted mouse mammary tumor model.

PNIPAAM can be used in sensor applications. For example, Zhang et al. produced glucose sensitive PNIPAAM copolymer microgel.[94] In their work, they functionalized a PNIPAAM copolymer microgel with the common glucose sensitive molecule phenylboronic acid (PBA). The resulting system exhibited the volume phase transition at a reduced temperature dependent on PBA content. The microgel particles, however, swelled in the presence of glucose, an effect that depended upon pH, ionic strength, and PBA content. The microgel showed significant response to glucose even under physiological conditions, suggesting a sensor application for diabetes.

D.2 BACKGROUND

In this section various concepts in MD that need to be understood to simulate PNIPAAM and copolymers are discussed briefly.

D.2.1 Integration

In MD one needs to integrate the system forward in time according to the laws of classical mechanics. This must be done numerically, and so requires an integration algorithm. Traditional methods such as Euler and Runge Kutta are not used very often. Instead, a common choice for integration is the Verlet algorithm as shown in equations D.1, D.2, and D.3.

$$\mathbf{r}_i(t+h) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t-h) + h^2\mathbf{a}_i(t) \tag{D.1}$$

$$\mathbf{v}_i(t) = \frac{\mathbf{r}_i(t+h) - \mathbf{r}_i(t-h)}{2h} \tag{D.2}$$

$$\mathbf{a}_i(t) = \frac{\mathbf{f}_i(t)}{m_i} \tag{D.3}$$

$\mathbf{r}_i(t)$, $\mathbf{v}_i(t)$, $\mathbf{a}_i(t)$, and $\mathbf{f}_i(t)$ are the position, velocity, acceleration, and force vector of or on particle i at time t. h is an algorithmic parameter called the time step. The choice of h can greatly affect the stability of the algorithm and the accuracy of the results. The Verlet algorithm uses the current position, some previous position, and the force on a particle to calculate the position at a later time. Velocity is not needed to perform the integration, making equation D.2 optional. Unfortunately, the Verlet algorithm suffers from precision loss and is not self starting.[95]

Another algorithmic choice is a variation of the Verlet algorithm called the leap-frog algorithm as shown in equations D.4 and D.5.

$$\mathbf{r}_i(t+h) = \mathbf{r}_i(t) + h\mathbf{v}_i(t+\frac{1}{2}h) \tag{D.4}$$

$$\mathbf{v}_i(t+\frac{1}{2}h) = \mathbf{v}_i(t-\frac{1}{2}h) + h\mathbf{a}_i(t) \tag{D.5}$$

While the leap-frog algorithm explicitly includes the velocity and avoids the calculations of large number differences, the velocities obtained are not synchronized with the positions obtained. In the leap-frog algorithm, position, velocity, and acceleration for at least four different times are involved in the acceleration. Since the velocities are out of sync, an extra step is needed to calculate velocities which are in sync with the positions.[95]

D.2.2 Force Field

As discussed in section D.2.1, MD is a classical dynamics problem that requires the integration of Newton's equations of motion. Newton's second tells us the time rate change of momentum of a particle is intimately tied to the force on that particle. One needs to know the forces present to perform the integration. The net force on an entire simulation cell should be zero and not time varying when energy is conserved. However, inter-particle forces are present. A common practice is to use an empirically or theoretically derived force field to describe the forces in the system. Often the force field consists of a set of parameterized potential functions and their parameters. Force functions are then the negative gradient of the the potential functions.

Depending upon how interactions are defined, a force field is sometimes termed as all-atom, coarse-grained, or hybrid. All-atom force fields explicitly represent all atoms in the system and define potentials including all atoms explicitly. Coarse-grained force fields do not necessarily represent all atoms in the system as single particles. Hybrid force fields may represent some atoms explicitly and others as coarse-grained. Solvent molecules in a system form a special class of molecules. It is common for solvent molecules to have a very specific force field or be modeled without the use of explicit particles. The entire set of potential functions can be viewed to form a multi-dimensional potential energy surface A coarse-grained surface can be understood as a smoothed out version of the atomistic surface. Dynamics occur on a faster timescale with less resolution on the coarse-grained surface.

Potentials can be divided into bonded interactions and non-bonded interactions. Bonded interactions exist for molecules in the system and describe how the potential energy of a molecule changes as the molecules conformation changes. It is most convenient to describe

bonded interactions in terms of generalized internal coordinates such as bond distances, bond angles, and dihedral angles.

Non-bonded interactions encompass how individual atoms in the system respond to the presence of all other atoms in the system. Atoms need not participate in chemical bonding or be in the same molecule to experience the non-bonded forces. Non-bonded interactions usually depend upon the distance between one, two, or more atoms. Mostly, but not always, the dependence on position is finite Laurent polynomial. Should the non-bonded interactions in this case depend upon some power of distance no greater than negative three, then the interactions are termed short ranged. Likewise, for powers greater than negative three, the interactions are termed long ranged. A few potentials common to the AMBER force field,[96] an all-atom force field, are discussed below. First, bonded interactions are discussed.

Probably the three most useful internal generalized coordinates to a chemist are atom-atom bond distances, atom-atom-atom bond angles, and atom-atom-atom-atom bond proper dihedral angles. Emphasis is put on the fact that only distances, angles, and dihedrals where atoms are properly bonded are being considered here. For each of these simple internal coordinates, atomistic potentials can be defined. For example, a potential can be defined for each bond in a system.

$$V(|\mathbf{r}_{ij}|) = k_{ij}\left(|\mathbf{r}_{ij}| - r_{ij0}\right)^2 \tag{D.6}$$

In equation D.6, $|\mathbf{r}_{ij}|$ is the distance between particles i and j, r_{ij0} is the equilibrium bond length for atoms i and j, and k_{ij} is the force constant describing the curvature of the potential or magnitude of the energy involved in the interaction. This is only one possible way to represent a bond interaction potential. Other potentials might more accurately represent the bond, such as the Morse potential. However, most of the time, MD does not allow for the possibility of chemical changes like bond breaking. The bonding potential shown is not without flaws. Should an atom move some very large distance with respect to it's bonded partner a great deal of instability to a calculation can be introduced into the simulation. Likewise, poor initial configurations of systems could cause similar problems. A similar harmonic potential can be defined for angles in the system, by replacing the equilibrium

length with an equilibrium angle and position with the intermolecular angle.

$$V(\theta_{ijk}) = k_{ijk}(\theta_{ijk} - \theta_{ijk0})^2 \tag{D.7}$$

Equation D.7 is similar to equation D.6. The equation is now defined for the angle θ_{ijk} between atoms i, j, and k, as well as their respective force constant k_{ijk} and equilibrium angle θ_{ijk}.

The last most useful internal coordinate defining a potential, the dihedral angle, needs special consideration. A dihedral angle is the angle formed between two *defined* planes formed by four atoms. One can form many different planes and combinations of two planes from four atoms. When four atoms are bonded end on end, two planes can be uniquely defined with the first three atoms and the last three atoms. The angle between these two planes in termed the proper dihedral angle when the angle domain is between 0 and 2π. A potential is often defined for all proper dihedrals in a system.

$$V(\phi_{ijkl}) = k_{ijkl} \cos\left(n_{ijkl}\phi_{ijkl} - \phi_{ijkl0}\right) \tag{D.8}$$

In equation D.8, ϕ_{ijkl} is the dihedral angle for atoms i, j, k, and l. The magnitude of the interaction is determined by the force constant k_{ijkl}, while the multiplicity n_{ijkl} determines the period or number of maxima and minima seen. The location of the minima and maxima is controlled by the phase angle ϕ_{ijkl0}. For any single dihedral angle in a system, a sum of terms like equation D.8 can be used to define the overall dihedral potential. The sum is not unlike a Fourier series.

When four atoms are bonded in a T-shaped pattern, two planes can be formed by each half of the *T*. The angle between these two planes is termed the improper dihedral angle. A potential is often defined for angles of this type *only* when the atoms confined to a plane at equilibrium geometry. The potential is not defined for every possible improper dihedral, rather only for certain ones such as around the planar carbonyl carbon. The same potential form used for proper dihedrals can be used for improper dihedrals, although harmonic functions are occasionally used also.

The non-bonded interactions are discussed next. To simplify the calculation, only pair-

wise non-bonded interactions are often considered. For short range interactions, the six-twelve pair potential called the Lennard-Jones potential is used.

$$V\left(|\mathbf{r}_{ij}|\right) = 4\epsilon\left[\left(\frac{\sigma_{ij}}{|\mathbf{r}_{ij}|}\right)^{12} - \left(\frac{\sigma_{ij}}{|\mathbf{r}_{ij}|}\right)^{6}\right] \tag{D.9}$$

In equation D.9, $|\mathbf{r}_{ij}|$ is the distance between atoms i and j, ϵ is a depth determining the strength of attraction, and σ is a distance identifying the onset of repulsion. The Lennard-Jones potential encompasses the many possible intramolecular short range interactions into a single form. Dispersive attraction and hydrogen bonding forces, among other things, are encompassed by the attractive inverse-six term. Pauli exclusion is included by the repulsive inverse-twelve term.

For long range interactions, the familiar Coulomb potential for point charges is often used.

$$V\left(|\mathbf{r}_{ij}|\right) = \frac{1}{4\pi\epsilon_0}\frac{q_i q_j}{|\mathbf{r}_{ij}|} \tag{D.10}$$

In equation D.10, $|\mathbf{r}_{ij}|$ is distance between atoms i and j, q_i and q_j are the charges on atoms i and j, and ϵ_0 is the permittivity of free space. The use of the Coulomb potential is somewhat of an approximation. Coulombs law is the correct solution of Poisson's Equation for point charges in electrostatics, however, atoms are not point charges. The more realistically point like charges, the electrons and nucleons, are not explicitly represented in a MD simulation. As a consequence atomic charges are derived and assigned to reside at atomic positions. This is done by picking a set of charges for atomic centers that reproduce a quantum mechanically derived charge distribution. The process does not really capture the idea that the electronic distribution of charges changes as atoms change their environment. Atoms are polarizable; the idea of an atomic charge on each atom *constant* throughout a simulation is a vast over simplification.

One needs more than just the forms of potentials to calculate. Be it the Lennard-Jones epsilon and sigma, or the harmonic bond equilibrium length, a suitable set of parameters need to be provided. The parameters can be extracted from experiment, theory, or a combination. Parameters such as the equilibrium bond length or force constant can be obtained

by fitting the a potential function to quantum mechanically derived equivalent. A set of pre-chosen molecules is used to do the fitting. The Lennard-Jones parameters can be derived by optimization of MC liquid simulations to reproduce experimental densities. Parameters for specific molecules, like water, other solvents, or amino acids, are often optimized to reproduce experimental phase points or enthalpies.

D.2.3 Periodic Boundary Conditions

To perform atomistic MD on a system of particles of macroscopic size is unrealistic. One can only hope to simulate a system on the order of a few hundred thousand particles in any reasonable amount of time. To better produce the properties of a bulk material while simultaneously keeping the number of particles explicitly represented in a simulation to a minimum, special techniques are required. A common possibility is to apply periodic boundary conditions (PBC).

Consider a simulation cell containing N particles. The idea of PBC is to *imagine* the simulation cell to be tessellated in all directions in space. All the particles inside the simulation cell are also replicated into all mirror image cells. The concept is not unlike the idea of the unit cell used to represent the infinite periodic lattice of a solid. Only particles inside the central unit cell are kept track of. As a MD simulation proceeds, particles inside the unit cell move around according to the equations of motion describing the system. PBC ensure that, should a particle pass through some face of the central unit cell, the particle will re-enter the face opposite the face it passed through. In this way, PBC allow one to replicate the behavior of an infinite system. The assumption is, however, that the infinite system is truly an isotropic replication of the central unit cell.

In reality, all particles in all images interact with another via non-bonded forces described in section D.2.2. This concept presents a challenge conceptually and computationally. How to handle it appropriately is discussed below in sections D.2.4 and D.2.5. A common practice to remove the infinite nature of the problem is to apply a the minimum image convention concept. The minimum image convention ensures that the cutoff to a long range interaction is no larger than half the shortest primitive lattice vector describing the simulation

cell. Doing so makes sure that particles only interact with their nearest neighbors, reducing the sum over all images to a sum over nearest images. Special care is needed to correctly identify nearest neighboring particles and compute minimum inter-particle distances.

How PBC are implemented in a simulation depends largely upon the translation symmetry of the simulation cell. If there is a geometry that tessellates in space then it can be used as a unit cell. Only regular polygons are used, greatly restricting the possible choices. Some examples include a simple cube or a truncated octahedron. Clever choice of the simulation cell can be used to reduce the number of particles used in the simulation, thus speeding up the simulation. Any simulation cell can be represented most generally by three lattice vectors which form a triclinic cell.

D.2.4 Short Range Interactions

The calculation of non-bonded forces can be a computational bottleneck. Often a computational scientist relies on parallelization of simulation codes and tricky algorithms to alleviate this bottleneck. The nature of the non-bonded calculation, some algorithmic tricks, and the idea of a neighbor list are discussed here.

The calculation of short range non-bonded forces is a complicated problem which requires various approximations to be tractable. In brute force, the problem has computational complexity N^2, where N is the number of particles. This means that the time it takes to perform the calculation scales quadratically.

One might consider the complexity of the problem arising from the application of PBC, among other things. For example, consider the calculation of a short range pair-wise interaction between N particles in a simulation. Each particle should interact with each other particle. If using PBC, each particle should also interact with all other periodic images of particles. With these considerations in mind, the problem is the infinite sum. The sum converges when the non-bonded forces are short range.

The application of an appropriate cutoff, particularly one adhering to the minimum image convention, allows the problem to become commutable. The spatial cutoff limits the number of interactions that need to be calculated, essentially truncating the sum and introducing

some error.

To evaluate the sum, a double loop over all particles would be performed and the nearest neighbor interaction calculated if two particles lie within the cutoff. The double loop becomes slower as nested if statements are added. It is hard to avoid at least four nested if statements inside the inner loop. Three are required for computing the correct minimum image distance in three dimensions. A fourth is required to compare the distance to the cutoff. Tricks such as comparing square inter-particle distances to square cutoffs are often applied to avoid necessary calculation of square roots. Additionally, intelligent choice of loop variables are chosen to take advantage of Newton's third law or avoid double counting of potential interactions. Still, the computational complexity is order N^2 because of the double loop over all particles. The task becomes extremely time consuming on the computer as the system size increases.

Often parallelization of the problem helps to speed up the calculation. However, not all problems are suited for parallel computing. The structure of the problem, the cost of communication, and computational overhead created by parallelization all need to be considered. The calculation of short range non-bonded forces in combination with the structure of a MD simulation lend themselves well to parallelization. Methods such as decomposing the domain of a simulation across multiple computers are often employed to greatly speed up the computation.

Another method to accelerate the calculation is the application of a neighbor list. The neighbor list is used to speed up the calculation of non-bonded forces by reducing the computational complexity of short range non-bonded forces. The idea of a neighbor list is to reduce the size of the inner loop. The inner loop is changed from a loop over all particles to a loop over neighboring particles. There are many variants of the neighbor list method, however in the simplest version, lists of the neighbors are pre-calculated and stored for each particle. The neighbors of a particle are those particles who lie within the cutoff. The neighbor list can reduce the computational complexity to order N. While the construction of the neighbor list is in itself a time consuming task of order N^2, the construction need not be performed every step if the neighbors are not changing very fast.

D.2.5 Long Range Interactions

Not unlike the short ranged non-bonded forces, the long range non-bonded forces are a computational bottleneck in a simulation. Although often suited for parallelization, long range non-bonded forces are not suited for the simple application of a cut off in an infinite system. The problem is rooted in the fact that if a potential varies as r^n with n greater than negative 3, the lost accuracy due to truncation is infinite.[95] Probably the most important example of long ranged non-boned potential is Coulomb's law.

A cut off distance can not be accurately used with a potential like Coulomb's Law in an infinite system. This means that minimum image convention principle is not applicable to the sum of Coulomb terms in a simulation. With no cut off to apply, a neighbor list can not be used for the long range forces (actually it can, but in a very different way). The sum is truly infinite over all periodic images of the simulation cell. It can be shown that the sum is only conditionally convergent, approaching a well defined value on if performed in a certain manner. Parallel computation can not remedy the problem of accuracy.

There are numerous ways to deal with the divergent behavior of long-range forces. The method used and discussed here stems from the Ewald sum calculation of coulomb energies in the periodic lattices of solids. A MC simulation cell is very like a periodic lattice.

The backbone of the Ewald sum method is to separate the sum of coulomb terms into a short range contribution that can be *safely truncated* and a long range contribution that can be *accurately estimated*. Essentially, the conditionally convergent coulomb sum over all images of a periodic lattice is broken down into two rapidly converging sums. How the portioning is performed is left up to the specific implementation of the method. The first sum is made to be a short ranged sum that converges rapidly in real space. This means a cut off, the minimum image convention, and the neighbor list can be used to accelerate this short range contribution. The second sum is essentially the long range tailing correction. The partitioning is done in such a way as to ensure the long range tailing correction is a slowly varying function. This mean the long range tailing correction can be represented in Fourier space by a finite number of wave vectors. In the end, the long range tailing correction converges rapidly in reciprocal space, allowing it to be truncated in reciprocal space.

The specific implementation of the Ewald method used here is called the particle mesh ewald (PME) method.[97] In the PME method, the coulomb sum is partitioned in the traditional sense by screening atomic charges with Gaussian charge distributions. The screened charge distribution is the short ranged contribution of the sum and is performed in real space with a cut off and the neighbor list. The canceling Gaussian screened distribution along with the tailing corrections form the reciprocal space sum. Although analytic formula exist for the evaluation of the reciprocal sum, PME, like other mesh methods, chooses to evaluate the reciprocal sum on a mesh of grid points. This evaluation is accomplished by interpolating the atomic charges onto the grid using interpolating polynomials. The choice of the Lagrange interpolating polynomials is one characteristic that distinguishes PME from some other mesh methods .[98] The main advantage of evaluating the reciprocal sum on a grid is that, instead of performing the sum, one can solve Poisson's equation. The numerical solution to Poisson's equation is very fast when using parallel forms of the fast Fourier transform, which only works on finite difference grids. The PME method has computational complexity $Nlog(N)$.

D.2.6 Other Ensembles

It is often said that the natural ensemble for Newton's equations of motion is the micro-canonical ensemble. That is, systems evolving according to dynamics prescribed by Newton's laws are those with a constant number of particles, a constant volume, and a constant energy. Gibbs postulated that time averages (over the dynamics of a system) are equivalent to ensemble averages. The problem is that ensemble averages are, in general, not equivalent in different ensembles. How does one deal with performing time averages for other ensembles?

Statistical mechanics tells us the meaning of a system at constant temperature. To be at constant temperature, a system is in contact with a heat bath. The heat bath and system are at a constant temperature. Individually, the bath or the system are not at constant energy. They exchange energy with each other to reach the same temperature. Together, both form an overall system that is at constant energy. Newton's mechanics correctly describes the dynamics of overall system, which is at constant energy. Newton's mechanics does not

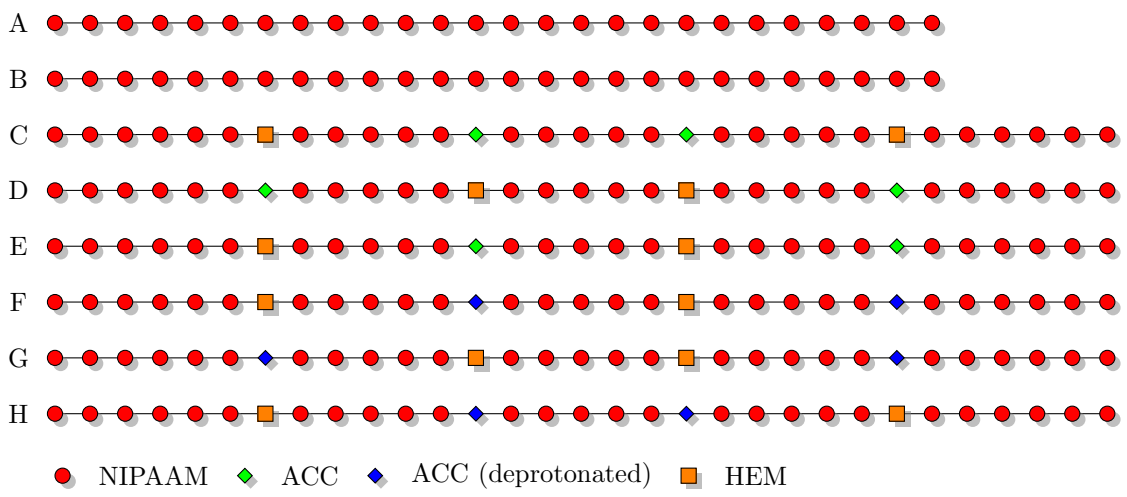


Figure D2: Topology of the NIPAAM polymers studied.

correctly describe the dynamics of the individual system, which is at constant temperature. If one can find a dynamics which does describe the individual system at constant temperature, then one can perform the correct ensemble averages.

The Nose-Hoover method produces a deterministic dynamics that can be used to simulate a system at constant temperature .[99,100] It is based on a reformulation of the Lagrangian equations of motion of the system.[95] The method introduces artificial variables into the Lagrangian that represent a heat bath. The dynamics observed represent the correct canonical distribution. A similar approach, the Parrinello-Rahman method, is used to introduce a volume bath or barostat.[101] This allows a dynamics to be produced at constant pressure.

D.3 METHODS

The simulations performed are described in this section. First, the systems chosen for study are presented. Second, simulation parameters used are discussed. Third, algorithmic parameters common to all simulations are given. Finally, the general simulation procedure is outlined.

D.3.1 Systems

A total of eight systems were examined. The systems are named A through G and shown in figure D2. System A was a pure PNIPAAM polymer with an atactic arrangement of monomers. The system contained 26 monomers and 8815 water molecules. System B was also a pure PNIPAAM polymer with 26 monomers. However, system B had a syndiotactic arrangement of monomers. The system contained 26 monomers and 8790 water molecules. Systems C through H were copolymers of PNIPAAM. Hydroxyethyl methacrylate poly-(trimethylene carbonate) (HEMAPTMC) and neutral acrylic acid (AAC) were the comonomers used. A medical study by Fujimoto et al.[92] examined these comonomers. Many possibilities existed for choosing how many comonomers to use and their positions in the polymer. Systems with four comonomers were chosen, with three possible arrangements of comonomers along the polymer backbone. In each case, two AAC comonomers and two HEMAPTMC comonomers were placed along the polymer chain. There were 5 PNIPAAM monomers placed in between each comonomer and 6 placed to cap the ends. This made a total of 31 monomers. In system C, a HEMAPTMC comonomer was placed near each end of the polymer chain and two AAC comonomers near the middle of the chain. In system D, an AAC monomer was placed near each end of the polymer chain and two HEMAPTMC comonomers near the middle of the chain. In system E, the two HEMAPTMC and two AAC comonomers were alternated along the chain. The relative arrangements of the comonomers are indicated in D2. These systems contained 15 314, 14 826, and 15 484 water molecules respectively. Systems F, G, and H were identical to systems C, D, and E. However, the AAC monomers were deprotonated. These systems were chosen in a crude attempt to examine the effect a low pH environment. Sodium counter ions were added to keep the systems neutral. These systems contained 15 766, 14 170, and 16 071 water molecules respectively.

D.3.2 Simulation Parameters

The AMBER force field (ff94) of Cornell et al.[96] along with the ff99sb modifications of Hornak et al.[102] were used to model system interactions. The force field variant was chosen for its extended parameterization beyond amino acids and nucleic acids. Additionally, the

all atom, explicit hydrogen representation of molecules was desired. See section D.2.2.

The TIP4P-EW explicit water model of Horn et al.[103] was used to model a water solvent. The model was used for its explicit representation of water molecules and an optimized performance with Ewald electrostatics methods. The number of water molecules used with each system was dictated by the size of the simulation cell used for each system. The volume of the simulation cell along with a rough density of 1.0 g/mL was used estimate the number of solvent molecules. Proper equilibration fixed the density.

The size of the simulation cell was determined by the length of a fully extended polymer and the cutoff for short range interactions. While satisfying the minimum image convention, the simulation cell size had to be so that a fully extended polymer had little or no chance of interacting with itself. Octahedral boundary conditions were used for pure PNIPAAM systems, while more general triclinic boundary conditions were used for copolymer systems.

The AMBER force fields do no provide atomic charges for a great deal of molecules. Atomic charges for the NIPAAM monomer were taken from a previous PNIPAAM MD study.[91] Atomic charges were derived for HEMAPTMC and AAC monomers according to the restrained electrostatic potential (RESP) method.[104] The RESP method can be summarized in a few steps.

1. perform a quantum mechanical geometry optimization of a molecule or residue.
2. perform a quantum mechanical single point calculation of the electrostatic potential
3. fit the derived electrostatic potential to point charges associated with each atom

During the geometry optimization, density functional theory (DFT) with the Becke 3-parameter Lee-Yang-Parr (B3LYP) functional was used. A $6-31G^*$ basis was used during geometry optimization and during electrostatic potential calculation. This basis set is recommended for the procedure.[104] The Gaussian[105] simulation software was used for these calculations. The RESP fitting procedure was done with the aid of the Antechamber[106] package.

D.3.3 Algorithmic Parameters

The MD simulations were performed using the GROMACS simulation software.[107–110] Many algorithmic parameters were based on recommendations found in the GROMACS manual.[111]

The leap frog integration algorithm was chosen with a time step of 2 fs. See section D.2.1.

For short range interactions, a spatial cutoff of 9.0 Å was used. A simple neighbor list was used with update frequency of 10 steps to accelerate short range calculations. The neighbor list was constructed using the same 9.0 Å cutoff. See section D.2.4.

For long range interactions, the particle mesh ewald (PME) method was used. See section D.2.5. The long range interaction was partitioned into a rapidly varying short range contribution and a slowly decaying long range contribution using the parameter $\alpha = 10^{-5}$. The rapidly varying contribution was evaluated in real space with a spatial cutoff of 9.0 Å. The slowly decaying contribution was evaluated in reciprocal space using the discrete fast Fourier transform. Initially, the charge density was assigned to a isotropic lattice of grid points spaced by 12.0 Å. The grid spacing was allowed to optimize for parallelization of the Fourier transform according to GROMACS protocol. Atomic charges were interpolated onto grid points using fourth order Lagrange interpolating polynomials. The relative permittivity of the system was set to 1 and metallic boundary conditions were assumed.

For simulations in the canonical ensemble, the Nose-Hoover temperature algorithm was employed. See section D.2.6. Temperature was different for various simulations. See section D.3.2 or figure D3. The period of temperature fluctuations was set to 5 ps.

For simulation in an isothermal isobaric ensemble, the Parrinello-Rahman pressure algorithm was also employed. See section D.2.6. All simulations were performed at 1.0 bar. The simulation cell and particle positions were scaled in an isotropic manner. The period of pressure fluctuations was set to 5 ps. The compressibility of the system was set to 4.5×10^{-5} 1/bar.

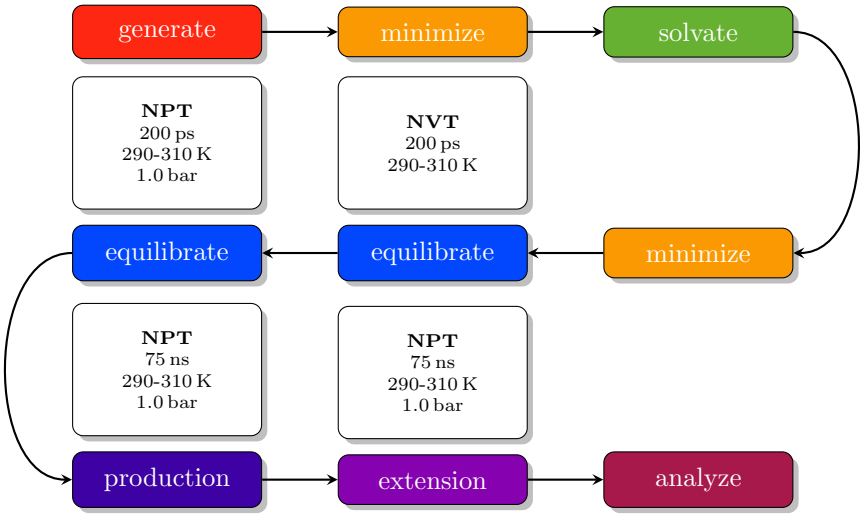


Figure D3: Molecular dynamics procedure.

D.3.4 Simulation Protocol

The general simulation procedure was a multi-step process and is summarized in figure D3. First, system configurations and topologies needed to be generated. The process to generate initial configurations can be summarized as follows.

1. derive atomic charges for possible monomers
2. place the desired number and type of monomers end on end in a simple cubic coordinate system.
3. find all atom types, bonds, angles, proper dihedrals, and improper dihedrals in the system.
4. assign the correct force field terms.
5. minimize the geometry with a steepest descents algorithm.
6. place the resulting extended configuration in an appropriately sized octahedral or triclinic coordinate system.
7. add enough water molecules to make the density approximately 1.0 g/mL.

The method used to derive atomic charges was in described in section D.3.2. To place monomers end on end, atomic positions for a monomer were first generated in a chemical software program. By carefully keeping tack of connection information, atactic and syndiotactic configurations could be generated as monomers were joined together. Systems A and B were placed in an octahedral boundary system, while systems C through H were placed in a triclinic boundary system. Roughly, the octahedral and triclinic systems had initial volumes around 900 and 1800 nm[3]. Box sizes were chosen to satisfy the minimum image convention and remove the possibility of the polymer interacting with itself. This corresponds to roughly 30 000 and 60 000 atoms per system. Steepest descents energy minimization needed to be performed after generating the configuration and adding solvent molecules to remove any bad energy contacts.

After the initial configurations were generated, the systems needed to be equilibrated at different temperatures and pressures. The temperatures used were 290, 300, 310, and 320 K. These temperatures are above, below, and around the LCST of PNIPAAM. All systems were simulated at a pressure of 1.0 bar. The equilibration process included 2 stages. First, a 200 ps equilibration in the canonical ensemble was performed at a desired temperature. Initial velocities were generated according to the Maxwell distribution for this step. It was checked that the system was stable and had a temperature fluctuating about the desired value. Second, a second 200 ps equilibration was performed in an isothermal-isobaric ensemble at a desired pressure and temperature. Velocities were not regenerated. It was checked that the system had reached a stable density and box size.

After equilibration, a production run was performed in an isothermal-isobaric ensemble for 75 ns at a desired temperature. Finally, production runs for systems A and B were extended for an additional 75 ns.

The 200 ps stages, run at 2 fs per step, amounted to 100 000 steps each. The 75 ns runs, also at 2 fs per step, amounted to 37 500 000 steps each. During that stage, statistics were collected every 25 000 steps. This gave rise to 1500 or 3000 configurations to perform time averages with.

D.4 RESULTS

After simulations were complete, various methods of analysis was performed and is discussed below in turn.

D.4.1 RMSD

In order to compare two conformations of a molecule, a molecular fitting algorithm can be used. A molecular fitting algorithm requires a numerical measure of the *difference* between two structures when they are positioned in space.[95] The root mean square deviation (RMSD) between two molecular conformations is one such numerical measure. When the RMSD is minimized, by changing the relative orientation between two structures, the RMSD will indicate how structures deviate from one another. The RMSD is defined to be the square root of the average distance squared between corresponding atoms in two different conformations.

$$\mathrm{RMSD}_{12} = \sqrt{\frac{\sum_{i=1}^{N} |\mathbf{r}_{i1} - \mathbf{r}_{i2}|^2}{N}} \tag{D.11}$$

In equation D.11, $\mathrm{RMSD}_{\mathbf{12}}$ is the root mean square deviation between conformations **1** and **2**, N is the number of atoms in the polymer, $\mathbf{r}_{i\mathbf{1}}$ is the position of atom i in conformation **1**, and $\mathbf{r}_{i\mathbf{2}}$ is the position of atom i in conformation **2**.

To observe how the conformation of systems change throughout the simulation, the RMSD was calculated with respect to the initial structure. Therefore, large values of RMSD indicate that a conformation is different from the initial conformation. Note that the RMSD between adjacent time frames offers much less information. That two time frames have similar RMSD does not necessarily indicate two structures have similar structures. The RMSD calculated is only to be interpreted as closeness to the initial conformation.

The goal of calculating the RMSD is to provide quantification to the possibly that PNIPAAM and copolymers are in a collapsed state. Since the initial structure of systems were relatively extended, significant deviation of the RMSD from zero at later time frames opens the possibility that the structure *may* be collapsed. It is important to emphasize may, for we can only know the polymer is some state dissimilar from the initial conformation. There

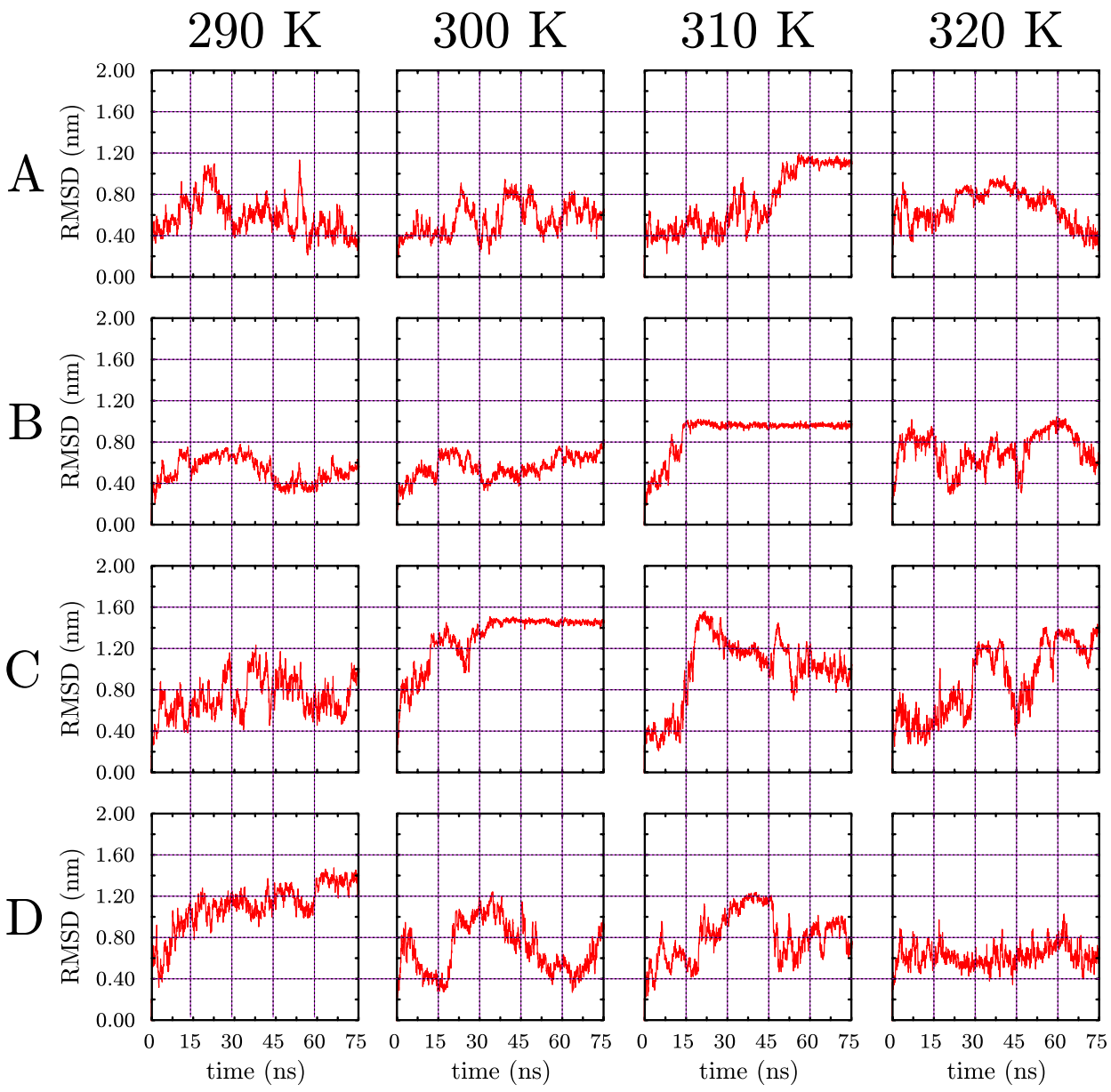


Figure D4: RMSD vs. time for systems A, B, C, and D.

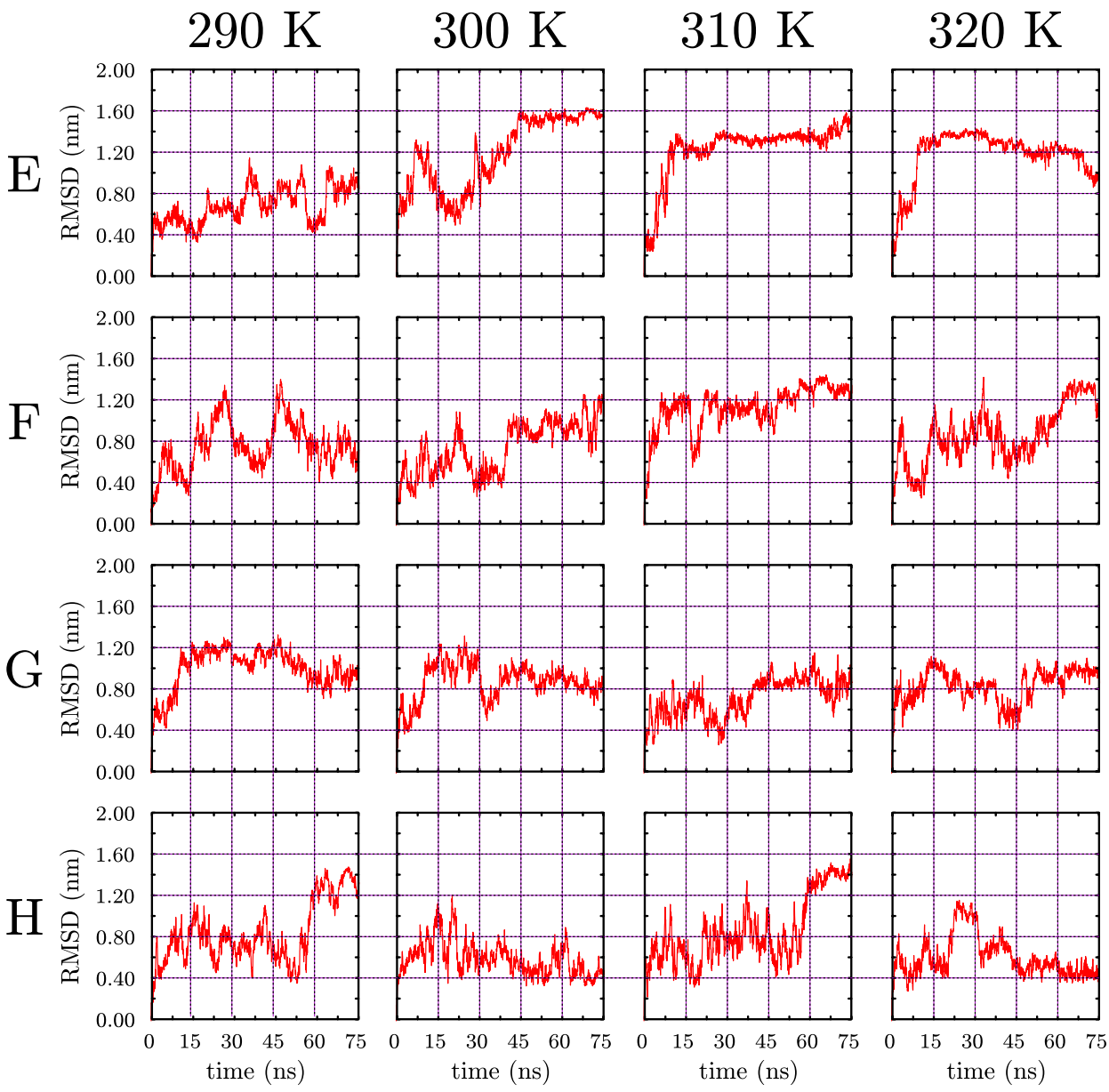


Figure D5: RMSD vs. time for systems E, F, G, and H.

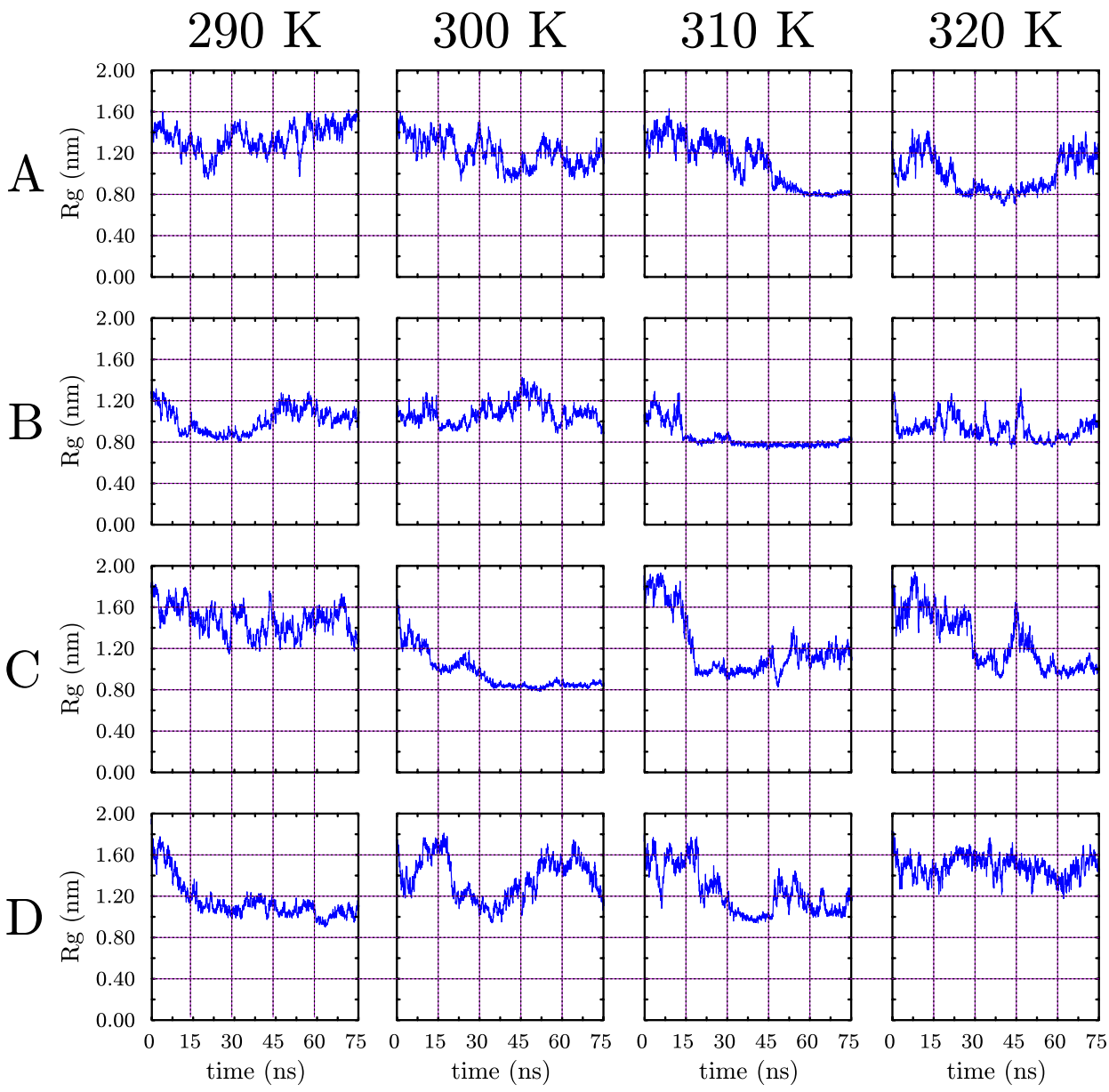


Figure D6: Radius of gyration vs. time for systems A, B, C, and D.

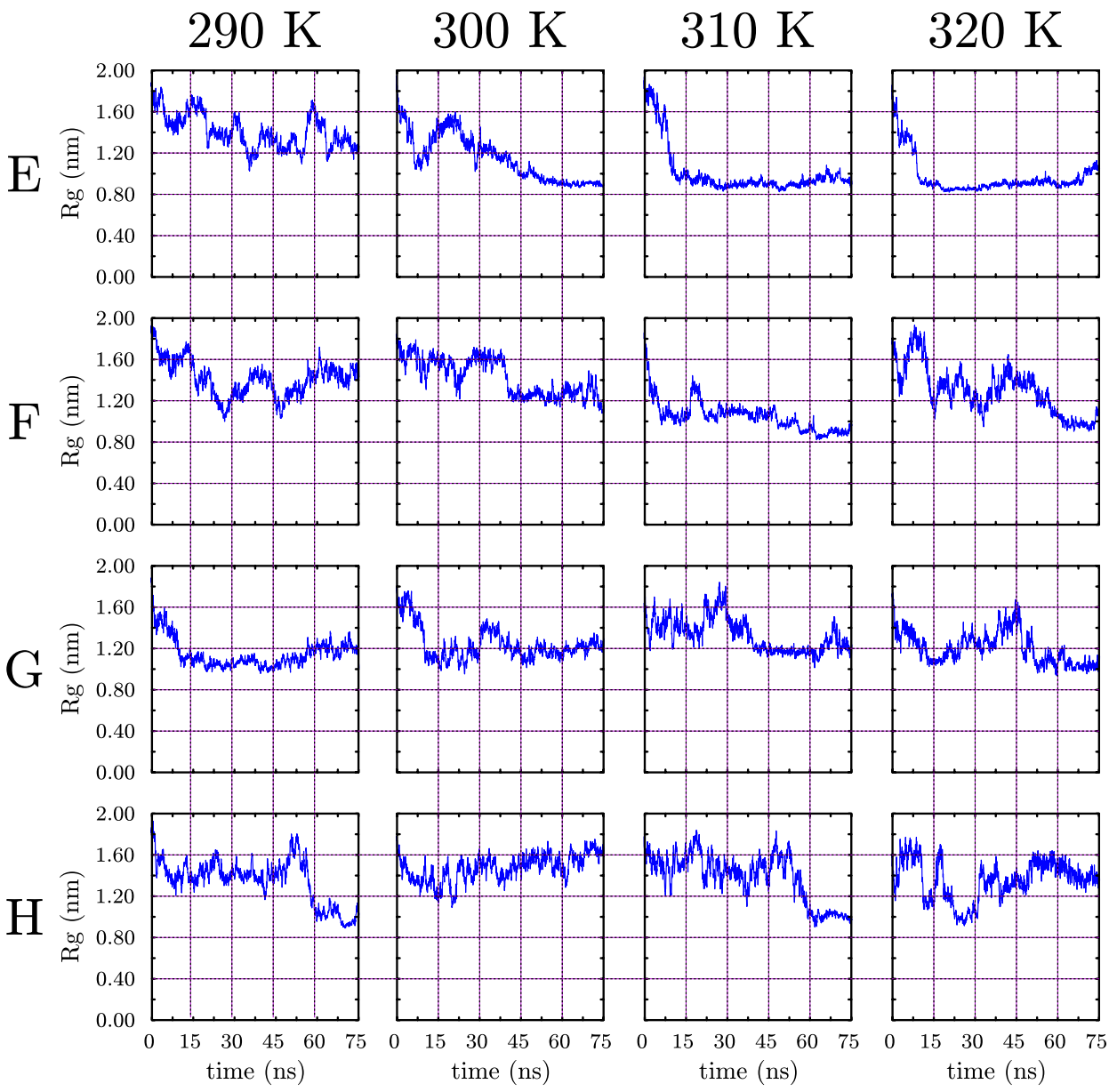


Figure D7: Radius of gyration vs. time for systems E, F, G, and H.

is no guarantee that state is collapsed.

When calculating RMSD, only the backbone atoms were considered. Backbone atoms are the alkene carbons of the monomers. Figure D4 shows the RMSD verse time for system A, B, C, and D. The RMSD is shown for 0 to 75 ns at 290 K, 300 K, 310 K, and 320 K. Figure D5 shows the same for systems E, F, G, and H.

In general, the RMSD for every system is a quantity that constantly changes as the simulation proceeds. The polymer is moving around; the conformation is constantly changing. That being said, regions where the RMSD of a system becomes large with low fluctuations appear at various temperatures. For example, system A takes on an average value of 1.108 ± 0.022 nm from 65 to 75 ns at 310 K. Compare this to average value of 0.421 ± 0.010 nm, 0.616 ± 0.069 nm, and 0.422 ± 0.074 nm during the same time period at the other temperatures. Clearly, the the RMSD at 310 K in this region is much larger than the other temperatures. Similar behavior can be found for the other systems. The behavior of the copolymer systems, however, seems to be more disordered.

The RMSD varied throughout the simulation for each system. This only means that the conformation has deviated from the initial extended state. It does not necessarily mean that systems are extended or collapsed. The pure systems showed large deviations near the known LCST of PNIPAAM. It is probable for the pure polymer to be found in some set of states that are not extended near 310 K. The copolymer systems occasionally showed large deviations, but at no real consistent temperature. In any case, it is best to think of the polymer existing in many states at a given temperature. The changing RMSD shows that the state is always changing. The polymer fluctuates between these states.

D.4.2 Radius of Gyration

The radius of gyration (R_g) is a concept in many disciplines used to quantify the overall size of an object. In polymer science, R_g is the mean squared distance of atoms from the center of mass of a polymer.[112]

$$R_g = \sqrt{\frac{\sum_{i=1}^{N} |\mathbf{r}_i - \mathbf{R}_{cm}|^2}{N}} \tag{D.12}$$

In equation D.12, R_g is the radius of gyration, N is the number of atoms in the polymer,

$\mathbf{r}_i$ is the position of atom i, and $\mathbf{R}_{cm}$ is the center of mass of the polymer.

To observe how the average size of the polymer throughout simulations, the R_g was calculated for each system. The R_g provides a direct indication of the size of the polymers at given times in the simulation. When combined with the RMSD calculated in section D.4.1, one can argue if or not the polymer is in a collapsed state. In contrast to the RMSD, it makes sense to directly compare the R_g between any time frame.

When calculating R_g, only the backbone atoms were considered. Backbone atoms are the alkene carbons of the monomers. Figure D6 shows the R_g verse time for systems A, B, C, and D. The R_g is shown for 0 to 75 ns at 290 K, 300 K, 310 K, and 320 K. Figure D7 shows the same for systems E, F, G, and H.

Examining system A, one can see that from 65 to 75 ns, the average R_g is 0.806 ± 0.015 nm at 310 K. Compared to other temperatures, whose averages are 1.484 ± 0.070, 1.122 ± 0.074 and 1.164 ± 0.091 nm, the R_g at 310 K is much smaller. Recall that, during the same time period, the RMSD of section D.4.1 increased greatly at 310 K. System A is considered collapsed in this region. Similar arguments could be made for other regions of different systems.

Note also that system A is also noticeably smaller from 30 to 60 ns at 320 K. However, the RMSD was not very large during this time period. This is because system A at 320 K had an initial conformation that was probably not as extended as has been said previously.

Overall, the R_g for each system is a quantity that is always changing. The polymer is moving around and its average size generally changes. However, not unlike the RMSD results, the R_g appears to take on small and large values of various time periods. More importantly, the moment R_g decreases often matches up exactly with the moment RMSD increases in section D.4.1. For this reason, the polymer is argued to be collapsed when the R_g is relatively small and extended when the R_g is relatively large. A good indication of the meaning of relatively large is the value of R_g near the start of a simulation. The starting configurations are already known to be extended. A good indication of the meaning of relatively small is to assign a cutoff value for R_g, possibly around 1.1 nm. In the end, the R_g and RMSD provide a way to tell if the polymer is extended or collapsed.

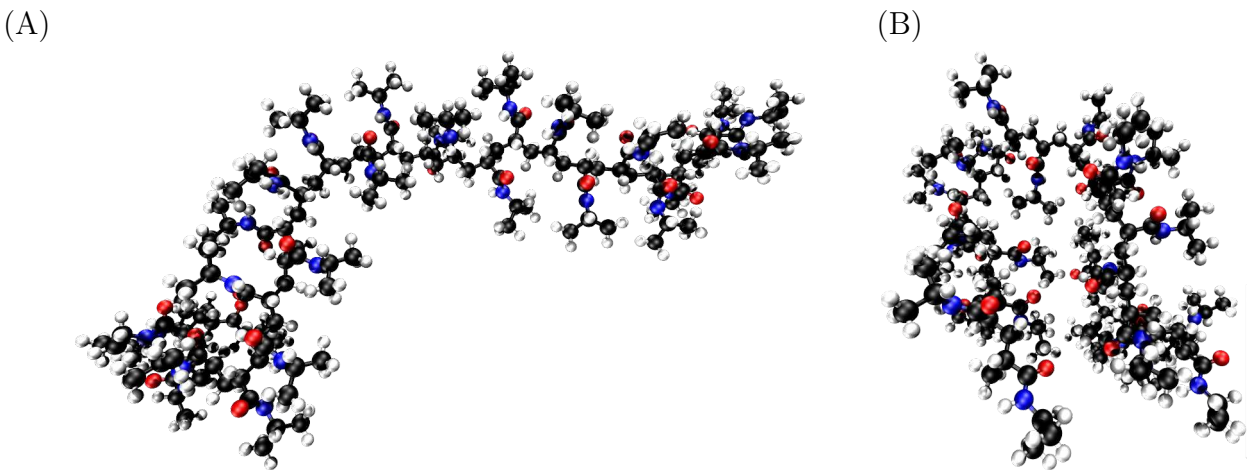


Figure D8: Extended and collapsed NIPAAM configurations.

D.4.3 Radial Distribution Function

The radial distribution function (RDF), pair distribution function, or pair correlation function, gives the probability of finding an atom a distance r from another atom compared to the ideal gas distribution.[95] In a simulation, the RDF is simply the ratio between the average number density $\rho(r)$ at a distance r from any given atom and the density at a distance r from an atom in an ideal gas at the same overall density. Any deviation of the RDF from unity reflects correlations between the particles due to the intermolecular interactions.[113] The radial distribution need not be defined between all atoms in a system. For example, to examine hydrogen bonding structure in a water model, the RDF could be defined between the water oxygen atoms and water hydrogen atoms.

To calculate various RDFs, the algorithm described by Allen and Tildesley[114] was used. To handle the correct calculation of minimum particle distance in triclinic cells, the method of W. Smith was applied.[115] The RDF was investigated for possible hydrogen bonds and for possible hydrophobic interactions. The goal of calculating RDF was to investigate the role of water in the collapsed and extended state of PNIPAAM and copolymers.

An example set of RDFs are shown in figure D9. From left to right, the RDF was

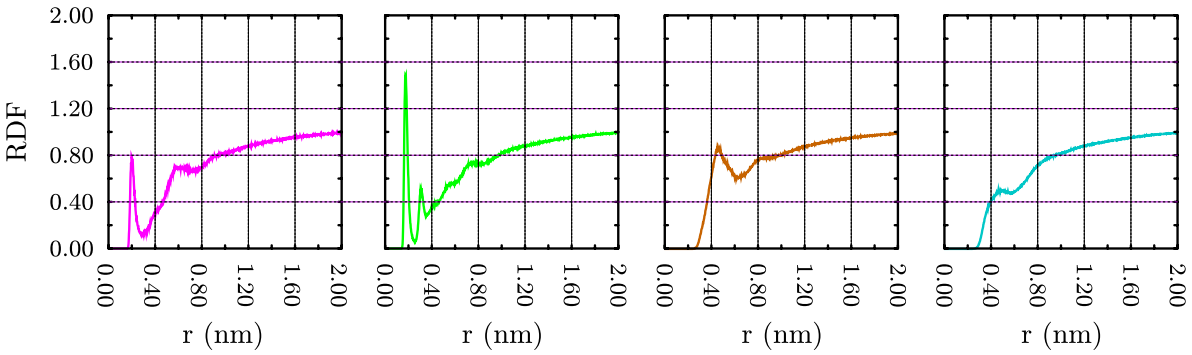


Figure D9: Radial distribution functions for system A.

calculated for various functional groups PNIPAAM and the copolymer. Defined first was an RDF between all amide hydrogens on NIPAAM monomers and all water oxygens. A peak appears around 3.0 Å. It is more probable to find a water at this distance near an amide than in an ideal gas of the same average density. The distance most likely corresponds to a hydrogen bond.

Second from the left in figure D9 is an RDF defined between all carbonyl oxygens on NIPAAM monomers and all water hydrogens. Two peaks appear near 2.0 and 3.0 Å. It seems probable that two water molecules are hydrogen bonded to carbonyl at these two distances.

Third from the left and fourth from the left in figure D9 are RDFs defined for hydrophobic interactions. The first is an RDF defined between all isopropyl groups on NIPAAM monomers and all water oxygens. The second is an RDF defined between all alkene carbons on NIPAAM monomers (the backbone atoms) and all water oxygens. There is a peak in the first graph, however, it is almost 5.0 Å away. There are no peaks in the second graph. It is not as likely to find water molecules near the isopropyl or backbone atoms of the polymer. This makes sense, because there should be no hydrogen bonding here.

A NIPAAM monomer making on average 3 hydrogen bonds, in the extended state at least, is in agreement with spectroscopic results.[90] However, according to the same paper, it is to be expected that the NIPAAM amide remained unchanged while the NIPAAM carbonyl

sheds a single hydrogen bond in the collapsed state. This would bring the total hydrogen bond count down to 2 in the collapsed state. The RDF did not seem to indicate this.

The RDFs were calculated for all the systems at all the temperatures. They were calculated for the entire trajectory, and for portions of the trajectory deemed collapsed or extended based on values of R_g. Unfortunately, little or no change was observed for any given RDF. To confirm, various programs and routines were used. This was surprising because it is thought that water plays an important role in the collapse of PNIPAAM and copolymers. However, one possible explanation is that the RDF is in some sense rotationally averaged out. Just because a hydrogen bound to one electronegative atom is within some distance of another electronegative atom does not necessarily mean a hydrogen bond is being formed. It should be that some sort of angular geometric criterion be met. In particular, one might expect the angle at donor electronegative atom to be somewhat linear.

D.4.4 Water Counting

The number of solvent molecules around a solute within a cut off distance gives an indication of solvent structure around the solute. When examined as a function of simulation time, one can observe if corresponding changes in solvent structure accompany changes in solute structure. To further probe the role of solvent molecules in conformational changes of PNIPAAM and copolymers, the number of waters around each polymer at a given time was counted. First, the smallest distance between a water molecule and a polymer was identified.

$$d_{\mathbf{WP}} = min(|\mathbf{r}_{i\mathbf{W}} - \mathbf{r}_{j\mathbf{P}}|) \tag{D.13}$$

In equation D.13, $d_{\mathbf{WP}}$ is the *smallest* distance between a water molecule $\mathbf{W}$ and polymer $\mathbf{P}$. $\mathbf{r}_{i\mathbf{W}}$ is the position of atom i in water molecule $\mathbf{W}$. $\mathbf{r}_{j\mathbf{P}}$ is the position of atom j in the polymer $\mathbf{P}$. If the distance was less than some cut off, then water was counted as being around the polymer.

$$\mathbf{W} = \begin{cases} counted & \text{if} d_{\mathbf{WP}} \leq c \\ not\ counted & \text{if} d_{\mathbf{WP}} > c \end{cases} \tag{D.14}$$

In equation D.14, $\mathbf{W}$ is a water molecule in question, and c is a desired cut off. The

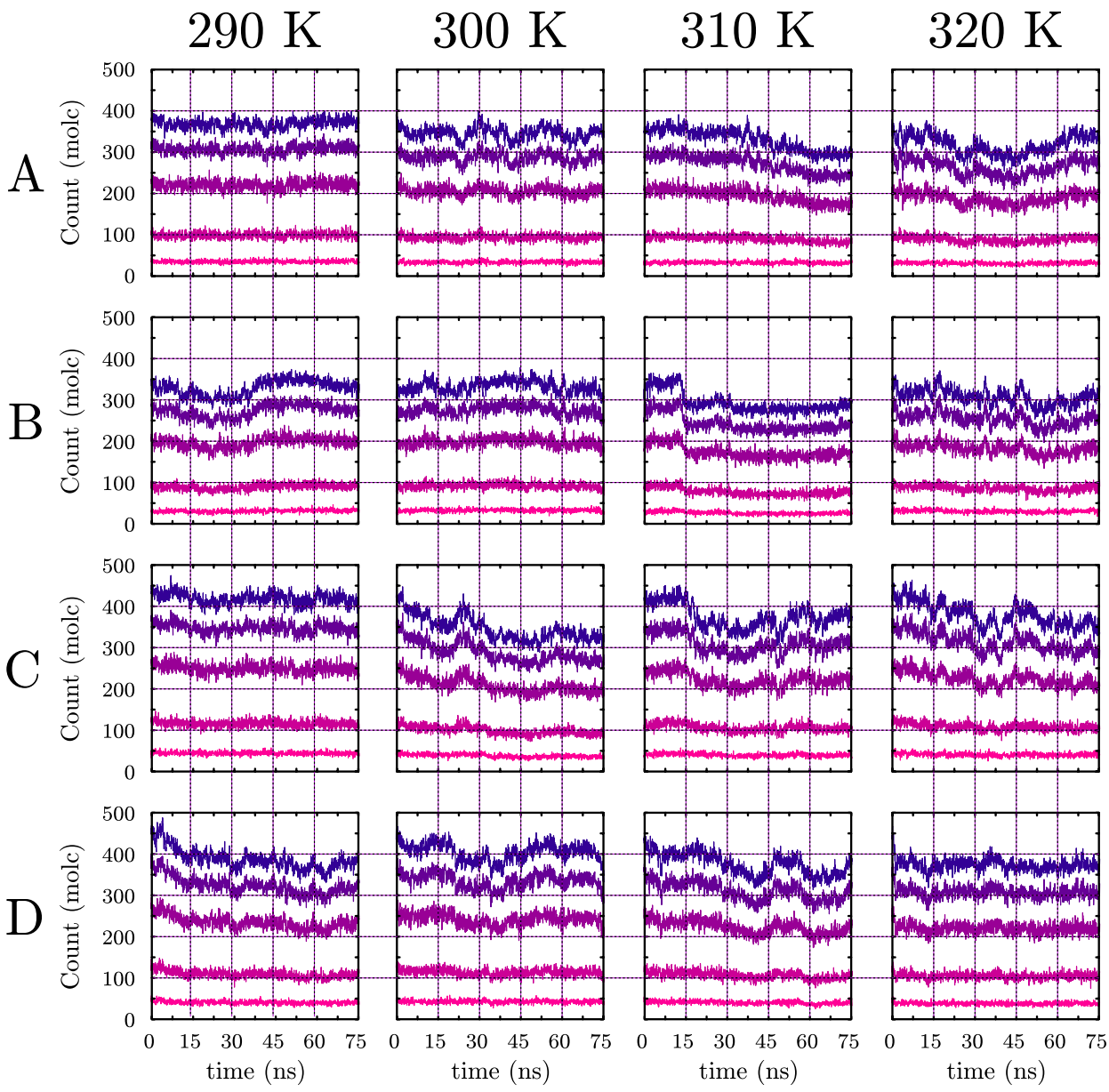


Figure D10: Water count vs. time for systems A, B, C, and D.

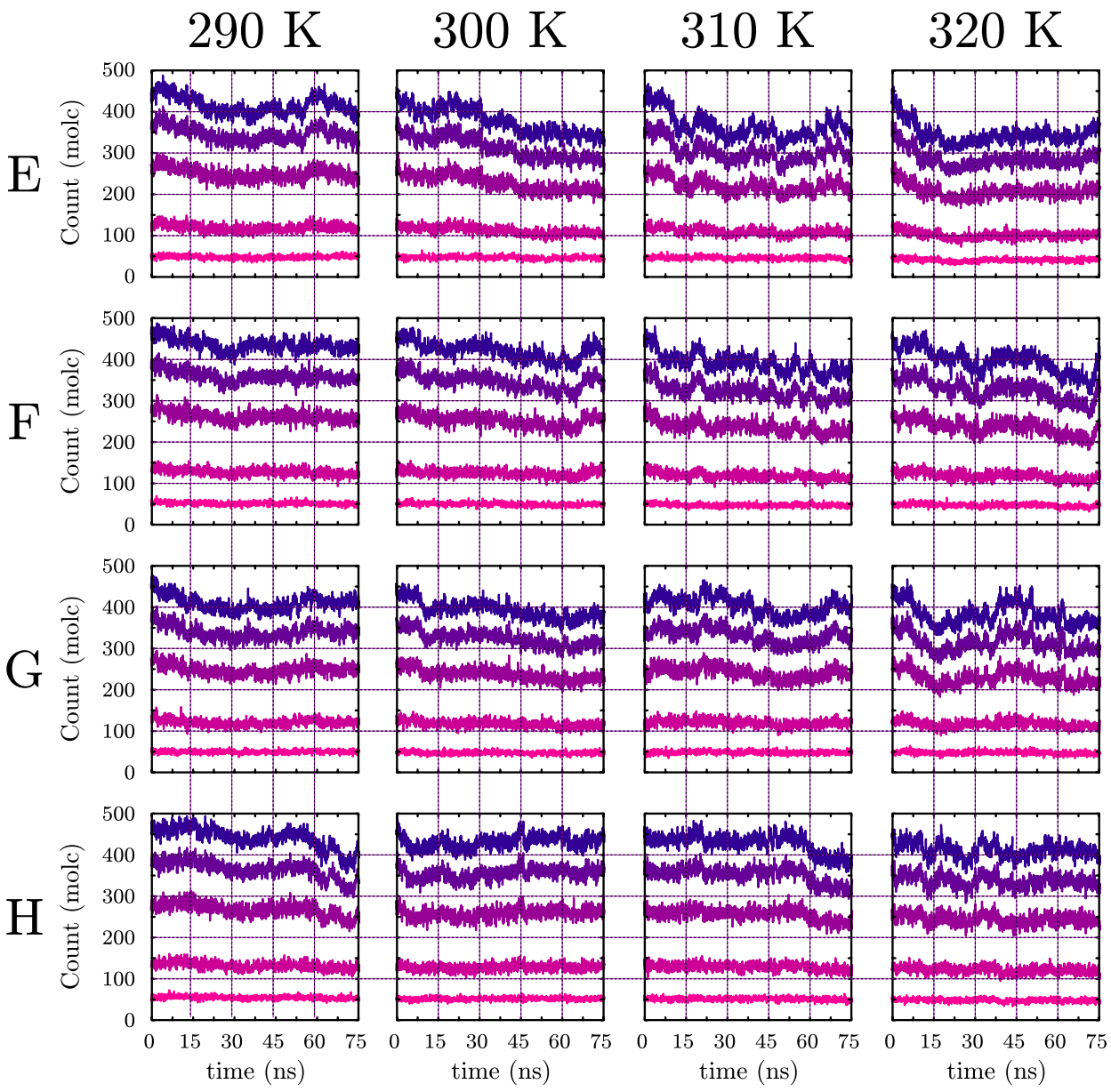


Figure D11: Water count vs. time for systems E, F, G, and H.

operation was performed for each water molecule in the system and monitored as a function of simulation time. Distances were calculated at each trajectory frame using the minimum image distance method of W. Smith.[115] All atoms in a polymer and all atoms in a given water molecule were included in the calculation. Cut off distances of 0.40 nm, 0.35 nm, 0.30 nm, 0.25 nm, and 0.20 nm. These distances are above and below common hydrogen bonding distances.

Figure D10 shows waters counted around polymers in systems A, B, C, and D. The water count is shown for 0 to 75 ns at 290 K, 300 K, 310 K, and 320 K. Figure D11 shows that same for systems E, F, G, and H.

Over all systems, one can observe that the number of waters within 0.20 nm is relatively constant. For systems A and B, the number of waters within this distance is roughly 35. For systems C through H, the number of waters within this distance changes between 40 and 50. Recall systems C through H were the larger copolymer systems. More waters can fit around larger polymers. The amide and carbonyl hydrogen bonding RDFs of section D.4.3 contained peaks at 2.0 nm and were also observed not to change very much throughout a simulation. It is possible that the waters counted within 0.20 nm correspond to carbonyl and amide hydrogen bonds made to the polymers. However, the numbers are slightly larger than the total number of amide and carbonyl groups in the polymer.

In a similar manner, waters counted withing 0.25 nm remain relatively constant. For systems A and B, the number of waters within this distance is on average between 90 and 100. For systems C through H, the number is higher around 115 to 125. These numbers are slightly more than twice but slightly less than three times the waters seen within 0.20 nm. It seems likely that the waters counted within 0.25 nm encompass all hydrogen bonds made to carbonyl and amide. This would include the second peak present in carbonyl hydrogen bonding RDFs, which corresponds to a saturated carbonyl.

For distances of 0.30 nm, 0.35 nm, and 0.40 nm, one starts to see the number of waters around the polymer begins to fluctuate. It can be seen that the number of waters generally decreases when the polymer R_g decreases. At first it might seem odd that the number of waters counted around PNIPAAM should decrease as PNIPAAM becomes smaller. Surely more waters can *fit* around a polymer when the polymer is smaller. However, consider that

the surface area of a polymer decreases with the overall size of the polymer. Since the surface area of the polymer is smaller when it is collapsed, there are less polymer atoms near waters and the count must decrease.

D.5 CONCLUSIONS

In this section the results of the study of PNIPAAM and copolymers are summarized. Possible future work is suggested.

D.5.1 Polymer Collapse

RMSD and R_g showed some general properties. Observing large RMSD and small R_g for a polymer allowed it to be classified as collapsed or extended. At no point did any system remain completely extended or collapsed throughout the entire simulation. The polymers were always fluctuating between conformations at all temperatures. However, clearly distinct regions of low R_g and high RMSD appear during specific time periods. It was often the case that the times when a low R_g appeared corresponded to a time when RMSD was large. When systems tended to enter a collapsed state for a significant time period, the amount of fluctuations in both RMSD and R_g greatly decreased. The result suggests systems in collapsed states are in deep wells on the potential energy surface.

Clearly collapsed states were observed for both pure NIPAAM polymers, systems A and B. These states are identified by a large RMSD and low R_g with small fluctuations. Below 310 K, system A is mainly in an extended states. Above 310 K, system A spends significant time in a collapsed state. System B appears to spend some time collapsed below 310 K, is defiantly collapsed at 310 K, and is more extended above 310 K. System B shows a collapsed behavior much earlier in the simulation compared to system A. It may be easier for a syndiotactic system to find a collapsed conformation where side chains do not interfere sterically. The behavior of systems A and B agree with the experimentally known LCST of pure PNIPAAM.

System C is mainly in a collapsed state at 300 K and above. System D does not spend significant time in collapsed states at any given temperature. If it is in a collapsed state, the overall size of the polymer is still larger than systems C or E. This result can be explained by the fact that system C contains large HEMAPTMC near the core of the polymer. Steric interactions may be preventing system C from obtaining a collapsed state for a long period of time. System E shows similar behavior to system C. The results suggest the LCST of protonated copolymer systems is slightly lower than pure PNIPAAM systems.

Charged copolymer systems existed in extended states most of the simulation. A possible exception is at 310 K, where system F is collapsed. Even when in collapsed states, systems F, G, and H show generally larger R_g. This result is most likely due to the inability to force charged AAC groups near each other for long in collapsed states. System G may also suffer from having large HEMAPTMC near the core of the polymer. The results suggest the LCST of deprotonated copolymer systems is slightly higher or out of range of the temperatures examined.

To explain the behavior of systems, consider the free energy change in going from an extended to collapsed state. An extended state is enthalpically stabilized by polymer-water contacts. In the collapsed state, these contacts are broken. A contribution may be recovered by favorable polymer-polymer contacts in the collapsed state. However, the overall enthalpic change to move from an extended state to a collapsed state is positive. That is, it takes more energy to break polymer-water hydrogen bonds and destroy solvent structure than is recovered by polymer-polymer contacts. However, as solvent structure is destroyed and polymer-water contacts broken, the overall entropy of the water will increase. The increase in entropy of the surrounding solvent is larger than the decrease in entropy of a collapsed polymer. The entropic change to move from an extended state to a collapsed state must be negative. In the end, it is the combination of enthalpy and entropy that determine if the polymer remains extended or collapsed. At low temperatures, the enthalpy of stabilization wins. The polymer remains extended because the free energy of going from a extended to collapsed state is positive at low temperatures. As higher temperatures, the free energy becomes negative because the entropy change in going from an extended to collapsed state is positive. The entropy term beats the enthalpy term at higher temperatures. Deprotonated

copolymer systems show a higher LCST because the enthalpy is more positive. It takes more energy to bring protonated groups close together, and more energy to strip water molecules away from highly charged groups. Therefore, a higher temperature is needed to make the entropy term win in free energy change of collapse. The explanation does not seem to explain why protonated systems show a lower temperature for collapse. It could be that the entropy change is even more positive due to the large HEMAPTMC groups moving around, even in the collapsed state.

D.5.2 Role of Water

To investigate how the solvent changed when NIPAAM collapsed the counting procedure and RDF were calculated. The counting procedure showed two trends. Waters within 0.25 nm of the polymer were not changing significantly as the polymer collapsed. These waters are most likely involved in hydrogen bonding with the polymer. The result would suggest that hydrogen bonding to the polymer is not changing significantly. The result is supported by the lack in change of hydrogen bonding RDF. Waters from 0.30 nm to 0.45 nm decreased in number as systems became collapsed. This result is because the surface area of exposed polymer decreases as it collapses. It also suggest that it could be possible some polymer-water contacts are being broken. By conservation of the number of water molecules, it must be that the number of water molecules is increasing even farther away from the polymer during collapse. This result agrees with the idea that the entropy of water molecules increases during collapse.

Amide to water hydrogen bonding RDF and water counting suggest a single hydrogen bond is made to NIPAAM in both the collapsed and extended state. Carbonyl to water hydrogen bonding RDF and water counting suggest two hydrogen bonds are made to NIPAAM in both the collapsed and extended state. The result contradicts experiment, which shows the number of carbonyl water hydrogen bonds should decrease by one in the collapsed state. It is questionable how well hydrogen bonding can be represented and identified in the simulation.

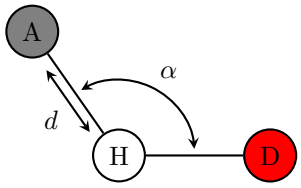


Figure D12: Geometric criterion for hydrogen bonds.

D.5.3 Future Work

More works needs to be done to better probe the conformational changes of NIPAAM and copolymers. The resolution with which one could observe hydrogen bonding in a simulation could be increased. One could use a force field that explicitly includes potentials for hydrogen bonding. The AMBER force field relies on non-bonded forces to include hydrogen bonding effects. All hydrogen bonding phenomena are absorbed into the Lennard-Jones 6-12 potential and Coulomb interactions. The idea is somewhat flawed. Hydrogen bonding, like many types of bonding, is geometric in nature. In particular, one might expect the donor-hydrogen-acceptor angle to be close to linear. How are spherically symmetric potentials, with no angular terms, supposed to capture the geometric nature of hydrogen bonding? There are better potentials that include angular effects and could be used to improve the results of the simulation.

Along the same lines, one could come up with better methods to identify hydrogen bonds in the simulation. For example, a geometric criterion could be used along with the water counting method. Rather than just count the number of waters within some distance of the polymers, count the number of water hydrogens within a distance of a hydrogen bonding donor atom. Ensure that the hydrogen is only counted if the donor-hydrogen-acceptor angle is some value. Consider the diagram in figure D12. It may also be possible to use an energy criterion to identify hydrogen bonds.

More variations on systems need to be examined. For example, systems more representative of experimental systems. While experimental systems may be solutions of polymers, the MD simulations are only of a single polymer in solution. Simulating a multi-polymer system should not be out of the realm of possibilities. Coarse graining is a possibility if

computational complexity becomes too large. A polymer network with water, or PNIPAAM hydrogel, would be a very informative system to study. Likewise, more work needs to be done in order to better understand the role of the HEMAPTMC and AAC copolymer systems. System variation would be a good next step for copolymer systems.

D.5.4 Summary

In summary, one can observe collapsed an extended states of PNIPAAM and copolymers near the known LCST. The temperature of collapse is less well defined in the simulation compared to experimental values. However, the temperature of collapse seems to be at the experimental LCST for pure system, slightly lower for copolymer systems, and slightly higher for deprotonated copolymer systems. While hydrogen bonding does not seem to change close to the polymers during collapse, the number of waters away from the polymer increases as the polymer collapses.

APPENDIX E

ACRONYMS

RMSD root mean square deviation. 177–179, 186, 191

R_g radius of gyration. 178, 179, 182, 185–187, 191

AAC acrylic acid. 152, 168, 169, 187, 190, 191

B3LYP Becke 3-parameter Lee-Yang-Parr. 169, 191

CPU central processing unit. 12, 14, 23, 26, 33, 36, 191

CUDA compute unified device architecture. 23, 191

DFT density functional theory. 169, 191

GDM gaussian disorder model. 9, 191

GPU graphics processing unit. 12–14, 23, 26, 27, 33, 36, 100, 191

HEMAPTMC hydroxyethyl methacrylate poly-(trimethylene carbonate). 152, 168, 169, 187, 188, 190, 191

HOMO highest occupied molecular orbital. 3, 14, 19, 22, 23, 55, 60, 191

LCST lower critical solution temperature. 155, 172, 178, 186–188, 190, 191

LUMO lowest unoccupied molecular orbital. 3, 7, 14, 19, 22, 55, 60, 100, 191

MC monte carlo. 131–134, 136–139, 142, 144, 147, 148, 151, 162, 165, 191

MD molecular dynamics. 12, 18, 131, 132, 138, 152, 153, 157–159, 161, 162, 164, 169, 170, 189, 191

NiOBuPc Ni(II)-octabutoxy-phthalocyanine. 62, 191

NIPAAM n-isopropylacrylamide. 152, 154, 169, 181, 186, 188, 189, 191

OBuNc octabutoxy-naphthalocyanine. 62, 191

OFET organic field-effect transistor. iv, 1–3, 6, 7, 10–15, 20, 21, 37, 38, 44, 60, 62, 64, 69, 70, 79, 97, 99, 101, 191

OLED organic light-emitting diode. 2, 191

OpenCL open compute language. 23, 191

OpenMP open multi-processing. 26, 191

OPV organic photo-voltaic. iv, 1–3, 7, 8, 10–15, 72, 100, 102, 191

P3HT poly(3-hexylthiophene). 61, 191

PBA phenylboronic acid. 156, 191

PBC periodic boundary conditions. 162, 163, 191

PC phthalocyanine. 14, 191

PCBM phenyl-C_{61}-butyric acid methyl ester. 14, 61, 191

PME particle mesh ewald. 166, 170, 191

PNIPAAM poly(n-isopropylacrylamide). 12, 152, 154–157, 168, 169, 172, 177, 178, 181, 182, 185–187, 190, 191

PPPM particle-particle particle-mesh. 191

RDF radial distribution function. 179, 181, 182, 185, 188, 191

RESP restrained electrostatic potential. 169, 191

SSMC solvent shift monte carlo. 12, 131, 138–140, 142, 143, 145–151, 191

Bibliography

[1] Marcus, R. A. *J. Chem. Phys.* **1956**, *24*, 979.

[2] Miller, A.; Abrahams, E. *Phys. Rev.* **1960**, *120*, 745–755.

[3] Sirringhaus, H. *Proc. IEEE* **2009**, *97*, 1570–1579.

[4] Singh, T. B.; Sariciftci, N. S. *Ann. Rev. Mater. Res.* **2006**, *36*, 199–230.

[5] Meerheim, R.; Lussem, B.; Leo, K. *Proc. IEEE* **2009**, *97*, 1606–1626.

[6] Rand, B. P.; Genoe, J.; Heremans, P.; Poortmans, J. *Prog. Photovoltaics Res. Appl.* **2007**, *15*, 659–676.

[7] Katz, H. E.; Huang, J. *Ann. Rev. Mater. Res.* **2009**, *39*, 71–92.

[8] Facchetti, A. *Chem. Mater.* **2011**, *23*, 733–758.

[9] Günes, S.; Neugebauer, H.; Sariciftci, N. S. *Chem. Rev.* **2007**, *107*, 1324–1338.

[10] Kola, S.; Sinha, J.; Katz, H. E. *J. Polym. Sci., Part B: Polym. Phys.* **2012**, *50*, 1090–1120.

[11] Mori, T. *J. Phys. Condens. Matter* **2008**, *20*, 184010.

[12] Servaites, J. D.; Ratner, M. A.; Marks, T. J. *Energy Environ. Sci.* **2011**, *4*, 4410.

[13] Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.

[14] Frenkel, J. *Phys. Rev.* **1938**, *54*, 647–648.

[15] Gartstein, Y. N.; Conwell, E. M. *Phys. Rev. B* **1995**, *51*, 6947–6952.

[16] Hanwell, M. D.; Madison, T. A.; Hutchison, G. R. *J. Phys. Chem. C* **2010**, *114*, 20417–20423.

[17] Cheng, X.; Noh, Y.-Y.; Wang, J.; Tello, M.; Frisch, J.; Blum, R.-P.; Vollmer, A.; Rabe, J. P.; Koch, N.; Sirringhaus, H. *Adv. Funct. Mater.* **2009**, *19*, 2407–2415.

[18] Noh, Y.-Y.; Cheng, X.; Tello, M.; Lee, M.-J.; Sirringhaus, H. *Semicond. Sci. Technol.* **2011**, *26*, 034003.

[19] Youn, J.; Dholakia, G. R.; Huang, H.; Hennek, J. W.; Facchetti, A.; Marks, T. J. *Adv. Funct. Mater.* **2012**, *22*, 1856–1869.

[20] Gwinner, M. C.; Pietro, R. D.; Vaynzof, Y.; Greenberg, K. J.; Ho, P. K. H.; Friend, R. H.; Sirringhaus, H. *Adv. Funct. Mater.* **2011**, *21*, 1432–1441.

[21] Gwinner, M. C.; Jakubka, F.; Gannott, F.; Sirringhaus, H.; Zaumseil, J. *ACS Nano* **2012**, *6*, 539–548.

[22] Caironi, M.; Newman, C.; Moore, J. R.; Natali, D.; Yan, H.; Facchetti, A.; Sirringhaus, H. *Appl. Phys. Lett.* **2010**, *96*, 183303.

[23] Cheng, X.; Caironi, M.; Noh, Y.-Y.; Newman, C.; Wang, J.; Lee, M. J.; Banger, K.; Pietro, R. D.; Facchetti, A.; Sirringhaus, H. *Org. Electron.* **2012**, *13*, 320–328.

[24] Tang, C. W. *Appl. Phys. Lett.* **1986**, *48*, 183.

[25] Peumans, P.; Yakimov, A.; Forrest, S. R. *J. Appl. Phys.* **2003**, *93*, 3693.

[26] Walzer, K.; Maennig, B.; Pfeiffer, M.; Leo, K. *Chem. Rev.* **2007**, *107*, 1233–1271.

[27] Sariciftci, N. S.; Smilowitz, L.; Heeger, A. J.; Wudl, F. *Science* **1992**, *258*, 1474–1476.

[28] Gao, J.; Heeger, A.; Hummelen, J.; Wudl, F.; Yu, G. *Science* **1995**, *270*, 1789–1791.

[29] Benanti, T. L.; Venkataraman, D. *Photosynth. Res.* **2006**, *87*, 73–81.

[30] Blom, P. W. M.; Mihailetchi, V. D.; Koster, L. J. A.; Markov, D. E. *Adv. Mater.* **2007**, *19*, 1551–1566.

[31] Mayer, A. C.; Scully, S. R.; Hardin, B. E.; Rowell, M. W.; McGehee, M. D. *Mater. Today* **2007**, *10*, 28–33.

[32] Thompson, B. C.; Fréchet, J. M. J. *Angew. Chem. Int. Ed.* **2008**, *47*, 58–77.

[33] Vanlaeke, P.; Swinnen, A.; Haeldermans, I.; Vanhoyland, G.; Aernouts, T.; Cheyns, D.; Deibel, C.; D'Haen, J.; Heremans, P.; Poortmans, J.; Manca, J. *Sol. Energy Mater. Sol. Cells* **2006**, *90*, 2150–2158.

[34] Vanlaeke, P.; Vanhoyland, G.; Aernouts, T.; Cheyns, D.; Deibel, C.; Manca, J.; Heremans, P.; Poortmans, J. *Thin Solid Films* **2006**, *511–512*, 358–361.

[35] Coakley, K. M.; McGehee, M. D. *Chem. Mater.* **2004**, *16*, 4533–4542.

[36] Dennler, G.; Scharber, M. C.; Brabec, C. J. *Adv. Mater.* **2009**, *21*, 1323–1338.

[37] Kippelen, B.; Brédas, J.-L. *Energy Environ. Sci.* **2009**, *2*, 251.

[38] Peet, J.; Heeger, A. J.; Bazan, G. C. *Acc. Chem. Res.* **2009**, *42*, 1700–1708.

[39] Bakulin, A. A.; Rao, A.; Pavelyev, V. G.; van Loosdrecht, P. H. M.; Pshenichnikov, M. S.; Niedzialek, D.; Cornil, J.; Beljonne, D.; Friend, R. H. *Science* **2012**, *335*, 1340–1344.

[40] Deibel, C.; Strobel, T.; Dyakonov, V. *Adv. Mater.* **2010**, *22*, 4097–4111.

[41] Guo, X.; Zhou, N.; Lou, S. J.; Smith, J.; Tice, D. B.; Hennek, J. W.; Ortiz, R. P.; Navarrete, J. T. L.; Li, S.; Strzalka, J.; Chen, L. X.; Chang, R. P. H.; Facchetti, A.; Marks, T. J. *Nat. Photon.* **2013**, *7*, 825–833.

[42] Healy, A. T.; Boudouris, B. W.; Frisbie, C. D.; Hillmyer, M. A.; Blank, D. A. *J. Phys. Chem. Lett.* **2013**, *4*, 3445–3449.

[43] Heiber, M. C.; Dhinojwala, A. *J. Phys. Chem. C* **2013**, *117*, 21627–21634.

[44] Tamura, H.; Burghardt, I. *J. Am. Chem. Soc.* **2013**, *135*, 16364–16367.

[45] Madison, T. A.; Hutchison, G. R. *J. Phys. Chem. C* **2011**, *115*, 17558–17563.

[46] Tamura, H.; Burghardt, I. *J. Phys. Chem. C* **2013**, *117*, 15020–15025.

[47] Bässler, H. *Phys. Status Solidi B* **1993**, *175*, 15–56.

[48] Wolf, U.; Bässler, H.; Borsenberger, P.; Gruenbaum, W. *Chem. Phys.* **1997**, *222*, 259–267.

[49] Fishchuk, I.; Kadashchuk, A.; Bässler, H.; Weiss, D. *Phys. Rev. B* **2002**, *66*, 205208.

[50] Fishchuk, I.; Kadashchuk, A.; Vakhnin, A.; Korosko, Y.; Bässler, H.; Souharce, B.; Scherf, U. *Phys. Rev. B* **2006**, *73*, 115210.

[51] van der Holst, J. J. M.; van Oost, F. W. A.; Coehoorn, R.; Bobbert, P. A. *Phys. Rev. B* **2011**, *83*, 085206.

[52] Williams, J.; Walker, A. B. *Nanotechnology* **2008**, *19*, 424011.

[53] Kodali, H. K.; Ganapathysubramanian, B. *Modell. Simul. Mater. Sci. Eng.* **2012**, *20*, 035015.

[54] Watkins, P. K.; Walker, A. B.; Verschoor, G. L. B. *Nano Lett.* **2005**, *5*, 1814–1818.

[55] Lyons, B. P.; Clarke, N.; Groves, C. *Energy Environ. Sci.* **2012**, *5*, 7657.

[56] Novikov, S. V.; Dunlap, D. H.; Kenkre, V. M.; Parris, P. E.; Vannikov, A. V. *Phys. Rev. Lett.* **1998**, *81*, 4472–4475.

[57] Gagorik, A. G.; Mohin, J. W.; Kowalewski, T.; Hutchison, G. R. *J. Phys. Chem. Lett.* **2013**, *4*, 36–42.

[58] Hoffmann, P. B.; Gagorik, A. G.; Chen, X.; Hutchison, G. R. *J. Phys. Chem. C* **2013**, *117*, 18367–18374.

[59] Gagorik, A. G.; Hutchison, G. R. *J. Phys. Chem. C* **2012**, *116*, 21232–21239.

[60] Madison, T. A.; Gagorik, A. G.; Hutchison, G. R. *J. Phys. Chem. C* **2012**, *116*, 11852–11858.

[61] Gagorik, A. G.; Mohin, J. W.; Kowalewski, T.; Hutchison, G. R. *J. Phys. Chem. C* **2013**, In Progress.

[62] Chen, X.; Gagorik, A. G.; Quan, X.; Gillespie, J.; Hutchison, G. R. *Nat. Mater.* **2013**, In Progress.

[63] Coropceanu, V.; Cornil, J.; da Silva Filho, D. A.; Olivier, Y.; Silbey, R.; Brédas, J.-L. *Chem. Rev.* **2007**, *107*, 926–952.

[64] Tessler, N.; Preezant, Y.; Rappaport, N.; Roichman, Y. *Adv. Mater.* **2009**, *21*, 2741–2761.

[65] Weis, M.; Lin, J.; Taguchi, D.; Manaka, T.; Iwamoto, M. *J. Phys. Chem. C* **2009**, *113*, 18459–18461.

[66] Katz, H. E.; Bao, Z. *J. Phys. Chem. B* **2000**, *104*, 671–678.

[67] Demeyu, L.; Stafström, S.; Bekele, M. *Phys. Rev. B* **2007**, *76*, 155202.

[68] Shi, N.; Ramprasad, R. *Appl. Phys. Lett.* **2006**, *89*, 102904.

[69] Mohan, S. R.; Singh, M. P.; Joshi, M. P. *J. Phys. Chem. C* **2012**, *116*, 2555–2562.

[70] Zaumseil, J.; Donley, C. L.; Kim, J.; Friend, R. H.; Sirringhaus, H. *Adv. Mater.* **2006**, *18*, 2708–2712.

[71] Mitrasinovic, P. M. *Curr. Org. Chem.* **2010**, *14*, 198–211.

[72] Cahen, D.; Kahn, A.; Umbach, E. *Mater. Today* **2005**, *8*, 32–41.

[73] Heimel, G.; Romaner, L.; Zojer, E.; Bredas, J. *Acc. Chem. Res.* **2008**, *41*, 721–729.

[74] Braun, D. *J. Polym. Sci. Pol. Phys.* **2003**, *41*, 2622–2629.

[75] Scott, J. C.; Malliaras, G. G. *Chem. Phys. Lett.* **1999**, *299*, 115–119.

[76] Scott, J. C. *J. Vac. Sci. Technol. A* **2003**, *21*, 521–531.

[77] Natali, D.; Caironi, M. *Adv. Mater.* **2012**, *24*, 1357–1387.

[78] Romijn, I. G.; Pasveer, W. F.; Martens, H. C. F.; Brom, H. B.; Michels, M. A. J. *Synth. Met.* **2001**, *119*, 439–440.

[79] Romijn, I.; Hupkes, H.; Martens, H.; Brom, H.; Mukherjee, A.; Menon, R. *Phys. Rev. Lett.* **2003**, *90*, 176602.

[80] NREL. Reference Solar Spectral Irradiance: Air Mass 1.5. `http://rredc.nrel.gov/solar/spectra/am1.5/` (accessed Jun, 2012).

[81] Wodo, O.; Ganapathysubramanian, B. *J. Comput. Phys.* **2011**, *230*, 6037–6060.

[82] Groves, C.; Reid, O. G.; Ginger, D. S. *Acc. Chem. Res.* **2010**, *43*, 612–620.

[83] Treat, N. D.; Varotto, A.; Takacs, C. J.; Batara, N.; Al-Hashimi, M.; Heeney, M. J.; Heeger, A. J.; Wudl, F.; Hawker, C. J.; Chabinyc, M. L. *J. Am. Chem. Soc.* **2012**, *134*, 15869–15879.

[84] Proctor, C. M.; Kuik, M.; Nguyen, T.-Q. *Prog. Polym. Sci.* **2013**, *38*, 1940–1960.

[85] Casalegno, M.; Raos, G.; Po, R. *J. Chem. Phys.* **2010**, *132*, 094705.

[86] Wei, F.; Liu, L.; Liu, L.; Li, G. *IEEE J. Photovoltaics* **2013**, *3*, 300–309.

[87] Li, G.; Liu, L.; Wei, F.; Xia, S.; Qian, X. *IEEE J. Photovoltaics* **2012**, *2*, 320–340.

[88] Schoen, A. H. *Infinite periodic minimal surfaces without self-intersections*; 1970.

[89] Hixson, C. A.; Benigni, J. P.; Earl, D. J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 6335.

[90] Ahmed, Z.; Gooding, E. A.; Pimenov, K. V.; Wang, L.; Asher, S. A. *J. Phys. Chem. B* **2009**, *113*, 4248–4256.

[91] Gangemi, F.; Longhi, G.; Abbate, S.; Lebon, F.; Cordone, R.; Ghilardi, G. P.; Fornili, S. L. *J. Phys. Chem. B* **2008**, *112*, 11896–11906.

[92] Fujimoto, K. L.; Ma, Z.; Nelson, D. M.; Hashizume, R.; Guan, J.; Tobita, K.; Wagner, W. R. *Biomaterials* **2009**, *30*, 4357–4368.

[93] Dufresne, M.-H.; Garrec, D.; Sant, V.; Leroux, J.-C.; Ranger, M. *Int. J. Pharm.* **2004**, *277*, 81–90.

[94] Zhang, Y.; Guan, Y.; Zhou, S. *Biomacromolecules* **2006**, *7*, 3196–3201.

[95] Leach, A. R. *Molecular Modeling Principles and Aplications*; Addison Wesley Longman Limited, 1996.

[96] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

[97] Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.

[98] Deserno, M.; Holm, C. *J. Chem. Phys.* **1998**, *109*, 7678.

[99] Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511.

[100] Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.

[101] Parrinello, M. *J. Appl. Phys.* **1981**, *52*, 7182.

[102] Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712–725.

[103] Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665.

[104] Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

[105] Frisch, M. J. et al. Program.

[106] Wang, J. Program.

[107] Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43–56.

[108] Lindahl, E.; Hess, B.; Van Der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.

[109] van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

[110] Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

[111] van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. Program.

[112] Ferrario, M.; Ciccotti, G.; Binder, K. *Computer Simulations in Condensed Matter: From Materials to Chemical Biology*; Springer, 2006.

[113] Frenkel, D.; Smit, B. *Understanding Molecular Simulation From Algorithms to Applications*; Elsevier, 2002.

[114] Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Oxford University Press, 1987.

[115] Smith, W. The minimum image convention in non-cubic MD cells. 1989.