Modeling and Numerics for Two Partial Differential Equation Systems Arising From Nanoscale Physics

Daniel Brinkman

Darwin College

DAMTP, Center for Mathematical Sciences

UNIVERSITY OF CAMBRIDGE

A thesis submitted for the degree of

Doctor of Philosophy

December 2012

I dedicate this thesis to my Fiancée, whose love and support were always appreciated, even from 4000 miles away.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

Daniel Brinkman

Acknowledgements

This work would not have been possible without the help and support of my colleagues, friends, and family.

First and foremost I would like to thank my supervisors Professor Peter Markowich and Professor Klemens Fellner. Professor Markowich introduced me to the topics contained within my thesis and helped me to find both coauthors and funding through his KAUST grant. Professor Klemens Fellner was instrumental in directing my day-to-day research beginning from my first week in Cambridge. Even after he accepted a position at the University of Graz, he continued to help me find direction both in my overall career development and the research which makes up Chapter 2 of this thesis.

I am very grateful to the post-doctoral researchers in the Applied Partial Differential Equations group at the University of Cambridge for many helpful discussions, especially Marie-Therese Wolfram and Clemens Heitzinger. Marie's assistance was vital for my understanding of the finite-element method used in Chapter 2 of this thesis (especially with regards to the use of NetGen and NGSolve). Clemens provided excellent and immediate hands-on direction in the project associated with Chapter 3 of the thesis.

I would also like to thank the other PhD students of Peter Markowich at the University of Cambridge: Daniel Marahrens, Jan Pietschmann, and Alex Lorz. I would especially like to thank my officemate Daniel for many interesting discussions, mathematical and otherwise. I would also like to thank Professor Peter Olver for kindling my love of PDEs as an instructor and research advisor during my undergraduate years.

Outside of the world of mathematics, I would like to thank my flatmate Matt for helping me maintain perspective (or at least extending my awareness to PhDs in sciences other than math and physics) and for many evenings of videogames and deep conversations. I would also like to thank my friends from the USA who took the time to cross the Atlantic and visit me: Nic, Tyler, Ben, Amanda, Katrina, Jon, and Kory.

None of this would have been possible without the support of my family. I would like to thank my parents, Tom and Diane, for their support and encouragement to leave my home country and my comfort-zone. I would also like to thank my sister, Beth, for helping me to feel connected to home through various online conversations.

Last but certainly not least I would like to thank my fiancée Denise for all of the love and support she has given me over the last three and a half years. Thank you for encouraging me to follow my dreams and pursue my studies far far away from home.

Abstract

The mathematical field of analysis of partial differential equations has undergone rapid changes in recent years. Consistent improvement of mathematical computation allows more and more questions to be addressed in the form of numerical simulations. At the same time, novel materials arising from advances in physics and material sciences are creating new problems which must be addressed. This thesis focuses on two such materials, organic semiconductors and graphene.

In Chapter 2 we derive a generalized model for organic photovoltaic devices – solar cells based on organic semiconductors. After selecting an appropriate self-consistent model, we derive asymptotic estimates for the operation of the device in several regimes. We use these estimates to partially justify several assumptions used in the formulation of simplified models which have already been discussed in the literature. Furthermore, we simulate such devices using a 2D hybrid discontinuous Galerkin finite element scheme and compare the numerical results to our asymptotics. We then discuss the potential applicability of the simulations to real-devices by identifying which regimes correctly model physical device behavior and discussing the limitations of the model choice. Finally, several perspectives are given on proving existence and uniqueness of the model.

In Chapter 3 we derive a convergent second-order finite difference numerical scheme for simulation of the 2D Dirac equation. We demonstrate the convergence of the numerical scheme with several examples for which explicit solutions are known and consider how errors appear in the simulations. We furthermore extend the Dirac system with Poisson's equation (in 3D) for the electrical potential and investigate the application to electrons in graphene. In particular, we show that the numerical scheme captures several interesting physical effects which have been predicted to appear in graphene.

Contents

List of Figures

Nomenclature

General

χ_I The characteristic function of I

$\frac{\partial}{\partial t}$ Partial derivative with respect to time

$\nabla \cdot$ Divergence operator

∇ Gradient operator

ν Unit normal vector

Ω Domain of the problem

Γ The boundary ($\partial \Omega$)

C^1 The space of continuously differentiable functions

i $\sqrt{-1}$

L^2 The Lebesgue space of order 2

Chapter 2 (Photovoltaics)

n Electron concentration

J_n Electron current

p Hole concentration

J_p Hole current

J Total electric current

U	Work function potential
V	The electric potential
E	The electric field, $-\nabla V$
X	Exciton (or polaron pair) concentration
α	Iterative damping parameter
ϵ_r	Relative electric permittivity
Γ_D	The part of Γ with mostly Dirichlet conditions
Γ_N	The part of Γ with mostly Neumann conditions
λ	Debye length
μ_n	Electron mobility
μ_p	Hole mobility
μ_X	Exciton mobility
∇_ν	Normal derivative (with respect to I)
c_r	Bimolecular recombination rate
c_r'	Rate of excition formation due to bimolecular recombination
d	Interface width
E_{SC}	The typical electric field at short circuit
G	Exciton generation rate
h	Dipole charge separation length
I	Interface region of the device (in the domain sense)
I^C	The complement of I with respect to Ω; the bulk of the device
k_d	Exciton dissociation rate
k_r	Exciton recombination rate

V_{diff} Potential difference

Chapter 3 (Dirac)

ψ Wave vector

Δx The step size in the x-direction

ϵ Electric permittivity

Φ Electric potential

A Magnetic vector potential

c Wave speed (not equal to speed of light for graphene)

m Electron mass

Chapter 1

Introduction

The role of mathematics in modern science cannot be overstated. Every discipline takes advantage of mathematical models and statistical tests, with numerical simulations becoming ever more common. As the science begins to demand more and more mathematical sophistication, classical mathematical techniques are applied to more and more interesting problems. New mathematical tools can be applied to garner new insights to old problems, and to answer new questions which require a high level of sophistication.

With this backdrop, partial differential equations have become increasingly valuable. The broad array of problems which can be modeled using variations of classical PDEs is staggering. Simple variants of the prototypical elliptic, parabolic, and hyperbolic PDE systems can be used to model phenomena ranging from electronic devices to crowd dynamics to the structure of our universe. Adding to the complexity are the rich applications which result from coupling multiple PDEs into a system. Such systems can consist of several similar equations with complicated coupling mechanics, or of completely different types of equations.

This thesis focuses on two systems of the second type, which couples the elliptic Poisson's equation for the electrical potential to evolution equations governing the motion of the particles, which then in turn govern the potential. The lack of a specific time-scale in the elliptic equation requires special consideration in order to apply appropriate numerical schemes. The most prevalent of such techniques are iterative. Solving Poisson's equation can be done using well-understood elliptic techniques at each time step, allowing us to insert a "known" potential into the evolution equations in order to solve systems which can be essentially arbitrarily non-linear in the potential.

Despite this underlying structural similarity and the common application to novel nanoscale electronic devices, the problems addressed in this thesis have very little in common. In Chapter 2 we deal with a coupled system of parabolic second-order drift-diffusion-convection equations modeling the behavior of organic photovoltaic devices using finite element methods for numerical simulation. In Chapter 3 we deal with a coupled system of hyperbolic first-order advection equations modeling the behavior of a sheet of graphene using finite difference methods. In this introduction, each problem will be introduced from both a physical and mathematical perspective, including an overview of the current literature and a discussion of the role of this work within the larger scheme of the problems.

1.1 Organic Photovoltaic Devices

Solar energy has been touted as an energy generating alternative to fossil fuels for ages. First patented for electricity generation near the end of the 19th century [94], solar power had already been used (for instance as a method for heating water) for many years. Modern commercial solar cells are all largely based on semiconductor technologies, the first of which was patented in 1946 [66]. The patent was submitted in 1941, thus predating all forms of nuclear technology. However, difficulties in nanoscale fabrication lead to solar energy lagging behind nuclear power for many years. Recent concerns about the safety and environmental impact of nuclear waste have lead to exceptional renewed interest in solar power.

Other forms of renewable energy sources also exist. Hydroelectric power and wind power, in particular, have received considerable interest. However, the environmental impact of such technologies are also questioned. Whether disrupting the mating and migration of fish species or killing endangered birds, the impact of harnessing the transportation media of animals is fairly controversial. In contrast, solar cells can be installed on artificial structures which are already in place, with significantly reduced environmental impact.

The ubiquitous semiconductor solar cell does have several disadvantages, however. High efficiency cells often rely on rare elements which only result from environmentally damaging mining processes, and some reports indicate that the total carbon cost of producing, transporting, and installing the devices outweighs the carbon savings from the clean energy. Furthermore, since classical semiconductor devices tend to be etched

onto a rigid substrate, they are by nature inflexible and somewhat unwieldy.

New research into organic photovoltaic (OPV) devices is intended to rectify both of these difficulties. Consisting of polymers which can be readily synthesized in the bulk from common materials, these devices can be loosely called plastics. Often flexible and significantly cheaper than traditional devices, OPVs have the potential to revolutionize the solar industry. Unfortunately, these devices currently have solar efficiencies which are significantly below those of modern inorganic semiconductor devices (i.e. at present 8.3% compared to 27.6% for "one-sun" Gallium Arsenide and 42.3% for concentrator cells [37]). However, significant research from a variety of disciplines has focused on methods for improving these efficiencies.

Many mathematical models have been proposed to study the behavior of bilayer OPV cells as simple implementable organic devices (for instance [5, 13, 15, 21, 50]). A shared mathematical basis for these models comes from extensive literature on inorganic semiconductor models. We begin by presenting the classical mathematical system for inorganic semiconductor devices, then present the necessary extensions to apply the model to OPVs.

1.1.1 Classical Semiconductor Theory

The PDE Model System

In classical semiconductor theory there are two charge carriers. Free electrons (n) are the negative charge carriers, but are only considered at sufficiently high energies (the conduction band) enabling them to move throughout the device. Lower energy levels are nearly completely filled and thus usually display no dynamics. However, if an electron is missing from one of these lower levels (the valence band), it creates the dynamics of a positively charged particle called a hole (p) – see Figure 1.1. In a real device, the equilibrium state can be modified chemically. A region with extra electrons is called n-type and a region with extra holes is called p-type. Because the carriers are electrically charged, they move in accordance to the local electric potential (V). The electric potential exerts a force on charged particles proportional to its negative gradient ($E = -\nabla V$) which is called the electric field.

For inorganic devices, the standard macroscopic models couple two drift-diffusion-reaction equations for the behavior of the free charge carrier densities (i.e. electrons and holes) in a system with Poisson's equation governing the self-consistent electrical
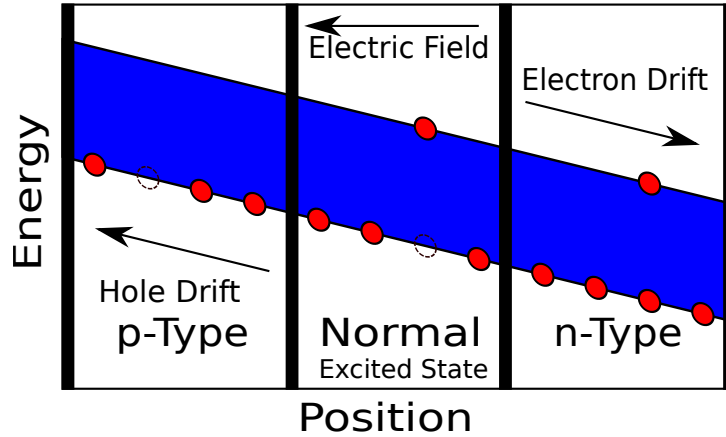
Figure 1.1: Schematic illustration of electron and hole transport for the conduction band (top line) and valence band (lower line) of a classical semiconductor with an applied field tilting the bands. Electrons are red circles, holes are black outlines. The blue area is a forbidden energy region called the band gap. Note that this simplified schematic ignores the effects of the self-consistent potential, which would cause the bands to "bend" in the n-type and p-type materials.

potential [74, 81]. The usual drift-diffusion system for classical semiconductors (also called van-Roosbroeck's system [89]) is as follows:

$$\nabla \cdot (\epsilon \nabla V) = q(n - p - C), \tag{1.1a}$$

$$\frac{\partial n}{\partial t} = \nabla \cdot (D_n \nabla n - \mu_n n \nabla V) - R, \tag{1.1b}$$

$$\frac{\partial p}{\partial t} = \nabla \cdot (D_p \nabla p + \mu_p p \nabla V) - R; \tag{1.1c}$$

where the variables and parameters will be explained in the following paragraphs.

It is possible for a free electron to fill a hole, resulting in the "loss" of both an electron and a hole. This process, called recombination, is taken to occur at rate R, which depends on the concentrations n and p. The reverse process is also possible, but requires an input of energy equal to the band gap (the difference between the valence band and the conduction band – the blue region in Figure 1.1). When this energy is provided by light, we call the process photogeneration, which gives an additional negative contribution to R.

Within the semiconductor, the influence of the electric potential on the charge carriers is modeled by a drift coefficient proportional to $\mu \nabla V$. The electrical current in the device

12

is given by the sum of the electron current, J_n, and the hole current, J_p, given by:

$$J_n = (D_n \nabla n - \mu_n n \nabla V) \qquad J_p = -(D_p \nabla p + \mu_p p \nabla V).$$

The mobility μ and the diffusion constant D account for how the material properties govern the motion of the charges. Even for general semiconductors several models exist for the dependence of μ on other variables in the system, especially the electric field [31]. For most applications, D is taken to be proportional to μ (the Einstein relation) [74].

The interesting behavior of semiconductors arises because the electric potential is not specified, but is strongly influenced by the charged carriers via Poisson's equation. Each of the charge carriers has the same fundamental charge q which contributes to the electric potential modified by the electrical permittivity ϵ. The electrical permittivity is a material property, so in a classical semiconductor with only one material (often silicon), it is usually a constant.

Semiconductor devices (such as transistors and diodes) form the basic components of most modern technology and are usually created by creating spatial inhomogeneity via doping – adding stationary charges (ions) to the system. This enters the system as a doping concentration C which is spatially dependent and can be positive or negative (depending on the nature of the doping). When C is very negative, it tends to repel the negative free electrons and creates a p-type region which consists predominantly of free holes; the corresponding electron-rich region is called an n-type region (as shown in Figure 1.1). Modifying how boundary conditions are applied and the geometry of the n- and p-regions allows for incredibly varied systems. For further information on the mathematical system, please refer to [55].

P-N Diodes and Solar Cells

The device which most resembles the OPV case is the P-N diode, which consists of exactly one p-region and one n-region, with a single (not necessarily linear or planar) interface between them (see, for instance Figures 1.2 and 2.1). An extensive mathematical treatment of such devices is given in [74, Chapter 4]. Such a device is often used to limit current to a single specific direction. An applied potential in the forward direction (corresponding to $V_{\text{app}} > 0$ in Chapter 2), results in a current roughly following Ohm's law. However, an applied potential in the reverse direction (corresponding to $V_{\text{app}} < 0$) gives only a small leakage current. When such a device has charges generated within the

device (say, from a negative contribution $-G$ to R resulting from particles generated by light), this small leakage current is supplemented by the photogenerated current.

Because the device is designed with two intrinsically different energy levels (usually as a result of the choice of contact metals), it is possible for this current to provide power to the system. The power of a circuit is given by the product of the current and the applied potential. For a device with an internal (or built-in) potential V_{int}, the power generated by the device is given by the formula

$$P = -JV_{\mathrm{app}} = J(V_{\mathrm{int}} + V_{\mathrm{diff}}).$$

Here V_{diff} is the potential difference in the device, which corresponds to the total potential difference between the two metallic contacts. Since the direction of the photogenerated current is usually determined by the potential difference V_{diff} and NOT the applied potential V_{app}, there exists a regime of positive applied potentials where the current is negative and the generated power is positive. As explained in the previous paragraph, for a typical P-N diode J has the same sign as V_{app} and thus power is only consumed by the device, not produced. Note that the power consumption regime can correspond to using the device as a light-emitting diode (LED). In the case where R corresponds (at least partially) to light emission, the diode will produce light when a positive field is applied.

Unfortunately, the notation for these quantities is not uniform throughout the literature. In particular, for the rest of this thesis, we generally treat V_{diff} as the dependent variable, as it represents the overall boundary data for the system. However, in many other references (including [74]) V_{app} is treated as the dependent variable. We choose to look at the potential difference mainly because it simplifies the notation for the 1D asymptotics discussed in Section 2.5.1. The other possible source of confusion is the sign of the potential quantities. Because the convention for a 1D P-N diode model is for the P material to be on the left and for the potential difference to be the difference of the P material and the N material, the potential difference is negative when the potential on the right is higher than the potential on the left. To reduce the possible confusion, for this work we will take zero boundary conditions on the right boundary and take the left hand boundary condition to be V_{diff}. For more details, see Figure 2.1 and the discussion at the beginning of Section 2.5.1.

Mathematical Analysis

Much is known about classical semiconductor systems, including existence and uniqueness results for the steady state. The steady state is of particular interest for DC devices which generally operate on time scales much longer than the duration of the transient (i.e. the switching-on) dynamics. In particular, [55] uses an iterative fixed point method to prove the existence of solutions on a domain $\Omega \subset \mathbb{R}^2$ of the steady state system for a general class of devices with moderate assumptions on the form of $C(x)$, $\mu(x, \nabla V)$, and $R(n, p)$. The boundary of the device $\Gamma = \partial\Omega$ is split into two disjoint parts, with homogeneous Neumann boundary conditions on one part (Γ_N) and Dirichlet boundary conditions on the remainder (Γ_D). The same reference also demonstrate uniqueness for small applied voltages. For general doping parameters, it is possible to create a device with non-unique solutions, but this is not expected to occur for bilayer devices (such as the P-N diodes of the previous section or the organic bilayer devices considered below).

Existence and uniqueness results are also known for the parabolic time-dependent system 1.1. One of the most powerful tools for such calculations is the entropy of the system. Under similar assumptions to those above, one can show that the entropy functional:

$$\int_\Omega n \log \frac{n}{n_\infty} - (n - n_\infty) + p \log \frac{p}{p_\infty} - (p - p_\infty) + \frac{\epsilon}{2}(\nabla V - \nabla V_\infty)^2 \qquad (1.2)$$

is positive and monotonously decreasing in t. The steady-state variables n_∞, p_∞, and V_∞ are the solution to the time-independent steady-state problem. Using standard regularity results for elliptic and parabolic PDEs, it is thus possible to prove that a unique solution exists for domains $\Omega \subset \mathbb{R}^{2,3}$ [32]. With a few additional assumptions on the structure of the coefficients and artificial trapping potentials (instead of more general boundary conditions), it is also possible to prove exponential decay of the solution to the steady state solution [23].

Existence and uniqueness of the system with more general coefficients (particularly μ and R) is also possible, but the proof becomes much more technical because a simple gradient-flow structure is not evident. However, it is still possible to prove global existence and uniqueness as well as convergence to a unique equilibrium through fixed point methods [97].

Numerical simulation of semiconductor devices has been an extremely active field for decades, with extensive work done at Bell Labs as transistor devices became vitally

important for modern technology. A 1964 article by Gummel [39] described the most ubiquitous method for resolving the difficulty with simultaneously simulating the elliptic Poisson's equation with the parabolic reaction-diffusion systems for n and p by iterating their solutions (now called Gummel iteration or the Gummel map). For an extensive consideration of the numerical simulation of such systems see [8].

1.1.2 Physics of Organic Devices

The basic mathematical structures of organic photovoltaic devices are nearly identical to the system for silicon semiconductors. However, the physical processes are significantly more complicated, in large part because organic materials are not generally good semi-conductors. More specifically, the band gap in organic polymers is significantly larger than that of typical inorganic semiconductors. Because most light cannot excite an electron from the valence band to the conduction band, it is incapable of directly creating electrons and holes [41] (as adding the term $-G$ to the recombination rate R in System 1.1 would imply). All is not lost, however, because the light is still capable of creating an excited electron energy state which remains bound to the hole it left behind. This excited state acts like a neutral particle and is called an exciton [44]. The first organic bilayer solar cell to take advantage of this alternate light generation mechanism was produced in 1985 by Tang [86].

Excitons and Polaron Pairs

The dynamics of these excitons have been studied in various materials for many years. Perhaps the earliest comprehensive analysis of their behavior is given in a 1934 paper by Lars Onsager [67] (unrelated to the work for which he was awarded the 1968 Nobel prize in chemistry). In particular, he posited the existence of charge-transfer-states (excitons) – bound pairs of oppositely charged particles which could either recombine or dissociate with specific rates. His work focused on explaining why the resistance of electrolytic fluids did not follow the usual Ohm's law. However, his method was extended by paired 1983 papers [7, 36] for the specific case of organic materials. In this context, the recombination refers to the the electron falling back into the valence band, emitting energy in the form or light or heat, whereas dissociation refers to the (productive) creation of free charge carriers. For a detailed examination of dissociation rate modeling, see Section 2.2.1.

16

Because of the large energy gap in silicon, an exciton will almost always recombine unless an external force acts to separate the pair. Thus in a region with homogeneous properties, excitons will usually not create free charges. Luckily for the field of organic photovoltaics, photogenerated excitons are not stationary, but move within the device. In a typical device they might diffuse $10 - 50$nm before recombining [58]. This allows us to design a device which contains regions with strong external forces, for instance the interface between two polymers with properties analogous to the p-type and n-type silicon discussed in Section 1.1.1. At the intersection of two such materials, the exciton settles in a configuration for which the bound electron and hole each reside in a material which is energetically favorable (see Figure 1.2). In such a state, the particles become much more stable, and eventually dissociate (as opposed to recombining) with very high probability [3]. Such a configuration is often called a polaron pair. In solid state physics, a polaron consists of a quasi-particle coupled to a localized electromagnetic wave. Because our polaron pair has a preferred direction, it creates a dipole moment near the interface. These dipoles can change the intrinsic energy levels of the device and lead to complicated electronic structure [90].
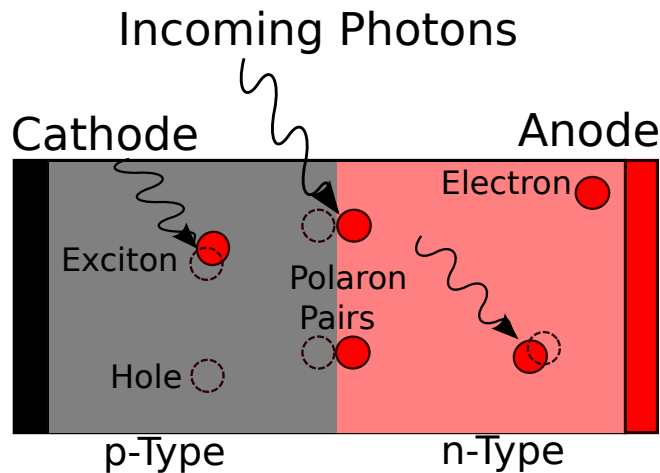
Figure 1.2: Schematic illustration of an organic photovoltaic bilayer, showing excitons and polaron pairs. Electrons are shown in red, holes in black. The drawing is not to scale, but the charge separation distance truly is longer for polaron pairs than for excitons. Compare to Figure 1.1 where electrons and holes exist only in n and p-type materials respectively.

Device Geometry

Given that the majority of exciton dissociation occurs at the interface between the n-type and p-type regions of the device, it is natural to consider ways in which to increase the length of the interface. In particular, if excitons generally diffuse at least 10nm [58], the optimal device would have an interface within 10nm of any given point. Because it is difficult to create coherent polymer layers thinner than 50nm, an alternative to the traditional flat bilayer device is necessary.

One possibility is to use two polymers which tend to phase-separate as a common solvent evaporates. Two groups demonstrated such bulk heterojunction (BHJ) devices independently in 1995 [40, 99]. These devices have the benefit of increasing the number of excitons which reach an interface and thus dissociate. This number is often measured by the internal quantum efficiency, the percentage of absorbed photons which yield escaping charges. The problem facing BHJs is the possible creation of "islands" of one type of polymer (see Figure 1.3). A charge created in an energetically favorable such region can become trapped and will not productively reach the contacts of the device. Carefully optimizing the feature size of the BHJ can mitigate this result, with recent devices exhibiting an internal quantum efficiency of nearly 100% [69].

Another strategy is to construct interdigitated polymers layer-by-layer so that "fingers" of each polymer protrude into the bulk of the opposite material [57]. This maintains the charge-transport advantages of a bilayer, since each generated charge still has a clear path to the contacts of the device – see Figure 1.3. More recent advantages in nano-fabrication have allowed for bulk processes which create small features such as nano-columns in a regular pattern across the device [42]. However, it is still extremely difficult to design such interfaces in practice, especially in the bulk. Nevertheless, such interfaces are extremely common in modeling because they represent a periodic structure governed by simple parameters (i.e. the frequency and amplitude of the interface oscillations). See, for instance [13, 95].

Coefficient Modifications

Although the addition of excitons is the most obvious modification to the classical semiconductor system, it is by far from the only such change. The electric-field dependent mobilities included in classical semiconductor systems become even more complicated for organic semiconductors. The Poole-Frenkel mobility discussed above [31] can still

Interdigitated Bilayer

Bulk Heterojunction
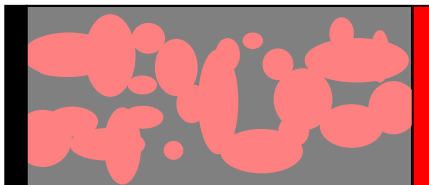
Figure 1.3: Schematic illustration of an interdigitated organic photovoltaic bilayer and a bulk heterojunction device (with medium-sized features). The red material is a n-type material and the black material is a p-type material. Note that for the BHJ some material might be isolated from the corresponding electrode. Compare to the structure Figure 1.2.

be used, but the justification is less clear because in general organic systems are not crystalline. A similar $e^{\sqrt{|E|}}$ form can still be derived [49], but experimental effects indicate that the regime for which these forms are applicable is not particularly broad. For example, temperature and charge concentration effects can become quite important for small electric fields [70]. Such effects can be even more important if self-heating is considered, adding another Poisson's equation for the temperature in the device [28]. We do not consider these effects in our model, especially since the exact forms for such modifications are not well-developed. See Section 2.2 for comprehensive details on the coefficient choices we make in our model.

Another potential issue with the modeling is the behavior of the boundary conditions of the device. In the first organic mono-layer devices, the charge injection from the metal contacts into the polymer device played a very important role [86]. However, bilayer devices show less impact from the metal-organic interfaces and instead depend heavily on the organic-organic interface [41]. For the majority of our simulations, we use the nonlinear organic-metallic contact boundary conditions given by [80]. These have shown great use in a variety of other papers on OPV modeling, including the majority of the

drift-diffusion modeling papers cited in this thesis (including [5, 13, 21, 47, 91]), but we show in Section 2.4 that they have little impact on the simulated behavior of the device.

1.1.3 Mathematical Modeling of the Organic System

The long diffusion ranges of the excitons above suggest supplementing System (1.1) with a reaction-diffusion equation for the excitons:

$$\frac{\partial X}{\partial t} = \nabla \cdot (D_p \nabla X) + cR - (k_d + k_r)X, \tag{1.3}$$

where X is the concentration of excitons in the device, k_d is the rate at which excitons dissociate and k_r is the rate at which excitons recombine. The recombination rate R for free electrons and holes (which usually depends on n and p) is the same R which appears in the continuity equations for n and p. The coefficient c represents the portion of recombining free electrons and holes that create excitons; the remaining portion represents radiative losses, generating either heat or re-emitted light. The term $k_d X$ must be added to the right-hand side of each of the continuity equations to represent the generation of free charges via exciton dissociation. Polaron pairs can be treated with similar dynamics (up to the meaning of the diffusion term if they are trapped on an interface).

Drift-Diffusion Models

Drift-diffusion equations have been used with different degrees of generality in a variety of articles on mathematical modeling of OPV devices. In fact, every drift-diffusion model for organic photovoltaic devices in the literature can be classified based on its treatment of the excitons and polarons in the device. Because the main contribution of the excitons to the dynamics of the device is only through the generation of electrons n and holes p, the primary importance of the model is the behavior of the term $k_d X$ throughout the device. There are several different alternatives treated in the literature. It is particularly important to distinguish between whether the model considers excitons, polaron pairs, both individually, or both combined. In particular, if the model focuses on a BHJ, then the interface regions are indistinguishable from the non-interface regions and the two types of charge pairs are indistinguishable.

Because the dissociation of polaron pairs is expected to be the main contribution to free charge carriers in an organic device, it is sensible to consider a model where the only charge generation comes from polaron pairs. The most influential article with this

type of modeling is the 2002 paper by Barker, Ramsdale, and Greenham [5]. The main downside of their analysis is that it applies only to a homogeneous bilayer device in one spatial dimension. Although this does allow for very general forms for the impact of the polaron pairs on the system (including what seems to be the only consistent dipole treatment given in the literature), it also ignores several interesting details. In particular, the formula for the polaron pair concentration becomes an ODE, and in the steady state the concentration can be explicitly written as a function of the other variables. They do extend the model for transient photocurrents in [46], but a 2D extension is still required in order to consider phenomena related to polaron pair orientation.

The other possibility for a bilayer device is to simulate only the excitons in the device. In this case, the dissociation rate can be made dependent on the background properties of the material being investigated. The most influential work of this type is described in two 2006 papers by Buxton and Clark [12, 13]. In [13], they describe their 2D finite difference numerical scheme and apply it to sinusoidal interfaces with different characteristic amplitudes and frequencies. In [12], they first simulate a realistic 2D polymer morphology and then use their model to simulate the characteristics of a typical disordered polymer system (the only such drift-diffusion modeling for such complicated interfaces). They then compare the resulting current-voltage characteristics arising from a realistic system and the ideal interdigitated system above. The key to the modeling is that they include a contribution from energy differences in the polymers as part of the electric field in the device (what we call U in Chapter 2). This allows for the use of an electric-field dependent dissociation rate which is small for the average field expected in the bulk of the device. For more details about this modeling, see Section 2.2.1.

Modeling only one type of bound electron-hole pairs is especially reasonable when simulating BHJ devices. Because it can be more or less impossible to distinguish whether an individual point consists of the n-type polymer or the p-type polymer, it becomes reasonable to simply consider averaged characteristics of the system. In particular, physical observations on sample polymer blends can be used to determine the mobilities and dissociation rates throughout the device. In this regime, excitons and polaron pairs are indistinguishable and only one type of particle needs to be simulated. For instance, the polaron pair modeling above can be used with a bulk concentration of polaron pairs which are assumed to be stationary (since they will always rest on some type of local interface), both in the steady state [21, 50] or considering transient phenomenon [47].

Alternatively, once can calculate the average creation and dissociation of polaron pairs as a function of the optical spectrum of the incoming light and use this simulation to calculate a spatially dependent generation term G for free carriers [51].

The most general perspective is to consider separate dynamics for both excitons and polaron pairs. To this end, Section 2.1.1 below justifies the creation of a system for which Equation 1.3 simulates the behavior of both types of bound particles depending on the location within the device. Taking 1nm as a typical charge separation distance for an exciton, we assume that bound states within 1nm of the interface are polaron pairs and that bound states in the bulk are excitons. By smoothing the interface over a thin region, we allow polaron pairs to diffuse along the interface (as opposed to the work of [95], which again treats the polaron pair concentrations as an ODE). We consider the model in 2D, generalizing the effects of [5], including a dipole effect generated by dipoles on the interface. In particular, we demonstrate that our dipole model exactly replicates [5] in 1D.

Asymptotic Results

Another common tactic in the literature is to begin with the full drift-diffusion models presented above and then study simplified asymptotic systems. Of particular note are early works focused on organic light emitting diodes [19]. These devices are essentially equivalent to the organic bilayers presented above (see Figure 1.2), except that current is forced into the device (as discussed for P-N diodes above). Because there is no photogeneration term, the model becomes much simpler. In particular, since electrons and holes are only injected at the appropriate contact, the minority carriers (i.e. electrons in the hole-transporting material and holes in the electron-transporting layer) are negligible. If we further assume that the applied field dominates the device so strongly that diffusion is negligible, then the system (1.1) is explicitly solvable [60].

The assumption that the minority carrier concentrations are zero is extremely prevalent in the organic photovoltaic literature. It simplifies many of the calculations, in particular since the right-hand-side of Poisson's equation (1.1a) becomes the density of the majority carrier. As we show in Section 2.5.2, we can thus write the whole system (1.1) as decoupled PDEs which only interact on the interface of the device. The zero-current (open circuit) case was given in the literature by [15], while the explicit results in Section 2.5.2 for nonzero currents were developed independently in [73]. They further-

more extend the modeling to include boundary conditions at a thin interface between the n-type and p-type regions and then solve the resulting equations in several different current regimes.

A major goal of our asymptotic calculations in Section 2.5 is to justify the zero minority carrier limit for general organic photovoltaic devices. We give numerical evidence supporting this qualitative conclusion in Section 2.6. However, we also show in Section 2.8 that a physically relevant strict limit of this kind cannot be obtained from a bulk drift-diffusion model of the type we have chosen. Motivated by the large-field approximations in [60], we develop an asymptotic scheme (including minority carriers) based on a strong electric field in the device (see Section 2.5.1). This gives reasonable results for large potential differences, but begins to fail as the potential difference approaches zero.

Of particular interest is the response of the current in the device, J, to the potential difference V_{diff} (see Figure 2.3). Of specific importance is the short circuit current J_{SS}, the open circuit voltage V_{OC} (the point where the current is zero), and the fill factor. The fill factor represents the ratio of the optimal power output to the ideal power output of the device [22]. The ideal power output is $J_{\text{SS}} * (V_{\text{int}} + V_{\text{OC}})$, these being the largest current and the largest applied potential in the energy-generating regime. We see in Section 2.6.3 that our asymptotic result matches the numerical results (discussed below) well for negative fields and approaching V_{OC}. To better simulate the intermediate region, we show in Section 2.7 that exactly integrating the PDE system yields simple expressions which can determine important quantities for device operation. In particular, we can then use Gummel iteration to approximate the electric potential and charge carriers in the device and we obtain excellent current-voltage characteristic results in Section 2.7.3. However, we also demonstrate in Section 2.8 that our model gives quantitatively unrealistic results at V_{OC}, even though it demonstrates the correct qualitative behavior.

Finite Element Numerical Simulation

Following the example of the majority of the literature, we choose to simulate the device in the steady state regime. In Section 2.3 we implement the model equations using a hybrid discontinuous Galerkin [16, 17] finite element discretization. These elements allow for discontinuities or near-discontinuities along edge boundaries, providing an ideal situation for the interface between the n-type and p-type regions within the device. We solve the system by iterating solution of Poisson's equation with solution of the continuity

23

equations (Gummel iteration [39]). Although potentially slower than Newton-methods used elsewhere [20], Gummel iteration is known to be extremely robust. We create the mesh using NetGen [78] and the related NGSolve wrapper. The matrix equations are solved using the PARDISO solver [76, 77]. The same combination has been used extensively for other reaction-diffusion problems [52, 96]. For details on the numerical implementation, see Appendix 2.B.

In Section 2.9 we discuss the use of our system for complicated interfaces such as those discussed in [12] and [95]. The use of finite elements allows for complicated interfaces to be implemented directly. However, choosing to have a thin interface region around the interface introduces numerous difficulties with implementation in the current form, especially with regards to the dipole term and interfaces which are not continuously differentiable (such as the interdigitated bilayer in Figure 1.3).

Existence and Uniqueness Results

Based on the extensive literature on classical semiconductor devices, there have been several results on the existence and uniqueness of drift-diffusion models for organic semiconductor devices. However, because of the variety of models presented in the literature, few general results are given. Several particular difficulties exist in the organic modeling case, most of which are mathematically intractable. In particular, the convection and diffusion coefficients become quite complicated. The mobilities μ and the permittivity ϵ are usually bounded away from zero for physical models, so uniform ellipticity is usually not problematic. However, the two materials in the system usually have different values for the coefficients, so for a bilayer with a sharp interface, discontinuous mobilities must be considered. Although difficult, this can be dealt with by using approximating systems and taking an appropriate weak limit [30]. Likewise, the mobilities given in physics literature are not a-priori bounded from above. Because this is generally unphysical (since very strong electric fields will cause a breakdown of the material), it is usually artificially bounded by a very large saturating mobility (which is never reached in simulations).

Existence is proven for the steady state and transient regimes in [21]. However, they have $\mu_X = 0$ and require that the other coefficients in the exciton equation (1.3) are constant. In particular, this avoids potential feedback difficulties of $k_d(E)$ and allows the exciton equation to be solved explicitly in the steady state. With an explicit formula for the excitons, the system fits within the framework of [74] and the proof of existence

24

is straightforward. A more general work given in [35] allows for a countable number of defect states with a similar structure to the exciton equation. They again use an approximation method to handle the worst nonlinearities, and also use an entropy similar to Equation (1.2). In many respects, their model is close to the one we consider, but once again they do not include field dependent rates or exciton dissociation.

In Section 2.10, we give several perspectives on proving the existence and uniqueness of our system. In particular, we consider how the fixed point methods might be extended to cover our case, examine modifications to the entropy in order to control excitons which can diffuse, and consider how duality methods could be applied to the problem.

Modeling Extensions

There are several possibilities which we do not consider here but which could provide interesting extensions to the current modeling. Some coefficient generalizations have already been discussed above and fit perfectly into the scheme we present below. One such topic is optical modeling. For most devices light is only incident from one direction and is absorbed at different rates in different materials [71]. Since this type of absorption profile is static, it would only require a preprocessing step to calculate the (now spatially dependent) value of G (as in [51, 95]). Likewise, including mobilities which depend on the charge concentrations would require only minor changes to our numerical implementation. Adding consistent temperature effects (as in [28]) would be less straight-forward, requiring another Poisson's equation in the system. However, this too could fit into our framework, likely being implemented with an additional iteration step using the previous values of n, p, and V.

In addition, there are also extensions which would require an entirely different implementation system. For the simple organic bilayer considered in Section 2.6, the interface region is isolated from the rest of the device and simulation in the bulk is relatively well understood. However, the behavior of excitons on the interface is less clear, and could benefit from extended small-scale modeling techniques. One such possibility would be the use of Monte Carlo simulations.

Because the drift-diffusion systems discussed above all rely upon local concentrations being well-defined, they are not necessarily applicable to devices which have features on a very small scale. In particular, BHJ devices can have features on the same scale over which quantum mechanical dissociation of excitons can occur near the interface. For

such simulations, Monte Carlo statistical modeling can be used to understand exciton dissociation [65]. Such simulations can be extended and used for an entire OPV device [56], showing the importance of both the local interface orientation and the size of the features. Further work by the same authors [38] indicates that larger features result in better efficiencies (at least in the applicable range of 4nm-16nm), mirroring the experimental results discussed above for BHJs.

1.2 Graphene

Graphene is an allotrope of carbon, consisting of a single 2D hexagonal lattice. The more common graphite (common in pencils around the world) consists of stacks of millions of layers of graphene. Writing with a graphite pencil leaves behind tens or hundreds of graphene sheets along the paper. Although studied mathematically as an auxiliary to the study of graphite rods over sixty years ago [92], graphene was ignored as a 2D material until quite recently. Long considered thermodynamically unstable, recent advances in fabrication techniques have resulted in a relatively simple recent method for isolating graphene [64], creating significant interest in the material.

1.2.1 Physics of Graphene

In recent years graphene has received attention due to its interesting mechanical, optical, and electronic properties. Although the bulk mechanical properties are difficult to test using the sample sizes available, there is no doubt that graphene exhibits impressive electronic qualities (for a review, see [59]). The electronic properties of graphene are interesting due to both useful applications and because they provide a test-bed for the Dirac equation [82].

The extensive list of interesting properties is one of the main reasons that Andre Geim and Konstantin Novoselov were awarded the 2010 Nobel Prize in physics [1]. Their seminal 2004 article, [61], demonstrated the experimental properties of graphene for the first time. The next year, they published their technique for isolating single sheets of graphene from bulk graphite, [64]. In that same year, they also published an article, [63], discussing the possibility of using graphene to experimentally test phenomena predicted by the Dirac equation (albeit at longer time-scales). This allows for exciting tests of quantum mechanical properties which are not possible (given current technology) for

fundamental particles. From these three articles, research in the area has exploded, with study proceeding in an astounding number of directions.

The electronic properties of graphene have been studied extensively from a variety of vantage points. Largely due to the low resistance of graphene, precise measurements can be made for previously infeasible applications or for challenging environments. In a pair of 2007 papers, two such directions are considered. In [62], the quantum hall effect is demonstrated in graphene at room temperature, a substantial improvement over the liquid-helium temperatures previously required. In [75], detection of single gas molecules by graphene sheets is demonstrated. Practical applications have also abounded, with particular emphasis on high-frequency transistors [54, 98]. Such devices could prove key to future improvement in computing speeds, with the potential to replace the aging complementary metal oxide semiconductor architecture. Other research has indicated that because electron transport is nearly ballistic, it should be possible to observe optical effects in graphene [14]. In particular, we investigate possible applications to beam splitters and Veselago lenses in Section 3.4.3.

The geometry of graphene is also nontrivial. In fact, some effects (especially the Hall effect) in graphene ribbons depend non-trivially on the relationship between the hexagonal structure of graphene and the geometry of the device [59]. Furthermore, graphene is not necessarily limited to flat geometries [68]. Carbon nanotubes can be considered as graphene sheets rolled into a tube (with different properties depending on the structure relationship discussed above). Such results emphasize the need for modeling to determine the stable structures which will be formed in different regimes [6] as well as to predict their electronic and optical properties.

1.2.2 The Dirac Equation

Many interesting studies have focused on the behavior of low-energy electrons in neutral graphene [63, 82]. In this regime, one can show that the tight-binding model for the graphene lattice yields a Hamiltonian identical to that of the 2D Dirac equation (albeit with a velocity $v_f \approx .01c$).

Using the trigonal symmetry of the graphene lattice, it is well known [92] that the tight-binding model yields the Hamiltonian

$$E(k) = \pm\gamma\sqrt{3 + 2\cos{(\sqrt{3}k_y a)} + 4\cos{(\sqrt{3}k_y a/2)}\cos{(3k_x a/2)}},$$

where $k = (k_x, k_y)$ is the wave vector, γ is a hopping frequency and a is the lattice spacing (assuming symmetric energy levels for convenience) [59]. Because each atom has one free electron, the negative branch of the square root is usually completely filled, and the minimum energy states will be the minimums of the positive branch. It is easy to verify that six such minimum momenta K exist, oriented at the corners of the hexagonal Brillouin zone of graphene. For wave vectors $k = K + q$ with $|q| << 1$, we find

$$E(k) = \frac{3at}{2}|q| + O(|q|^2)$$

where $E(K) = 0$ due to the symmetry assumptions above.

This dispersion relation is identical to the dispersion relation of the 2D Dirac equation with the speed of light, c replaced by the velocity $v_f = 3a\gamma/2$. One can therefore derive the Hamiltonian for the system and verify that indeed we obtain the Dirac Hamiltonian [82]. Thus for the low-energy wave vectors sufficiently close to the Dirac points K, we can use the 2D Dirac equation to simulate the behavior of electrons in graphene.

With this motivation, we turn to the mathematical modeling of the Dirac equation:

$$i\hbar\frac{\partial\psi}{\partial t} = (c\bar{\alpha} \cdot \bar{P} + \beta mc^2)\psi$$

for the wave vector ψ, where t is time, c is the characteristic speed (usually the speed of light), $\bar{P}$ is the spatial momentum operator (with three components), and m is the mass. The coefficients α_i and β must obey structural assumptions in order for the equation to correctly model physical behavior. The momentum operator $\bar{P}$ is primarily determined by choosing position-space as a basis. Furthermore, coupling with an electromagnetic field alters both the time derivative and the momentum operator. Choosing the usual minimal coupling [79] we obtain:

$$i\hbar\frac{\partial\psi}{\partial t} = \left(c\bar{\alpha} \cdot (-i\hbar\nabla + \frac{e}{c}\bar{A}) + \beta mc^2 + e\Phi\right)\psi,$$

where e is the fundamental charge, Φ is the electric potential, and A is the magnetic vector potential.

First introduced by Paul Dirac in 1928 [24], the eponymous Dirac equation models the relativistic motion of electrons and other spin-1/2 fermions. It now forms the basis of huge sections of theoretical physics and is a cornerstone of modern study of quantum mechanics. The most novel aspect of the equation is that in 3D the coefficients (α_i and β) are 4×4 matrices, so the equation is actually a system of four partial differential

equations. This was particularly unexpected since only two "types" of electrons (spin up and spin down) are physically observed and the other two equations seem to indicate the existence of electrons with negative energy. Dirac himself proposed one solution to this apparent paradox by speculating that the unobserved states were completely filled with electrons and that we only observe "holes" where such electron states are empty [25]. This provides the nomenclature of electrons and holes used extensively in semiconductor theory (as in Chapter 2 of this thesis). Global existence and uniqueness of results for the electromagnetic system (for instance coupled with Maxwell's equations) is well known [29].

Although some work has been published addressing the Dirac equation with electromagnetic fields from a numerical perspective, it has largely focused on the 3D regime, often using a fluid-dynamics framework to speed computation [83, 84] or specifically in the semi-classical and non-relativistic regimes [45]. Further work has been done specifically in 2D, but it has largely focused on illustrating the phenomena associated with quantum mechanics [87, 88] rather than the simulation of graphene as a 2D material.

For a 2D material, we only have to deal with two systems, because the orthogonality conditions on α and β can be satisfied by 2x2 matrices (for instance by the 2×2 Pauli matrices). After scaling (see Section 3.2), we obtain the following system:

$$\partial_t u^1 = - \left(\partial_x u^2 + i A_x u^2 - i \partial_y u^2 + A_y u^2 \right) - i(1 + V)u^1,$$
$$\partial_t u^2 = - \left(\partial_x u^1 + i A_x u^1 + i \partial_y u^1 - A_y u^1 \right) + i(1 - V)u^2.$$

We show in Section 3.2.2 that a naïve homogenization of the 3D Dirac equation with one isotropic direction results in two coupled copies of the 2D Dirac equation (provided that the electric potential and magnetic fields also obey consistency requirements).

By choosing the Coulomb gauge [79] and further assuming that the magnetic field is transverse and dominated by the applied magnetic field, we show in Section 3.2.4 that we can couple the system to a self-consistent electric potential by introducing the (scaled) Poisson's equation:

$$\Delta V = -\tilde{\lambda}^2 \left(|\tilde{u^1}|^2 + |\tilde{u^2}|^2 \right),$$

where $|u^i|^2$ is the concentration of particles described by the ith wave component of the wave vector.

We create a numerical scheme for solving the equation in several parts in Section 3.3. First, we choose to use the method of Gummel [39] to split the solution of Poisson's

equation and the continuity equations. This has the added advantage that the scheme works identically for the constant-potential case (without Poisson's equation). For the continuity equations, we note that for a given potential, the equations are linear and can be split into two individually solvable parts (using the method of Strang [85] to obtain the combined solution). The term proportional to $\mp i(1 \pm V)$ can be solved exactly as an exponential. For $V(x, t)$, the integral in the exponent is effectively calculated by Riemann summation on the time steps. Because the exponent is complex, both solutions correspond to complex rotation (and therefore conserve mass). We solve the remaining portion using a Crank-Nicolson type implicit scheme [18]. By choosing center-difference spatial derivatives, we can ensure that the scheme conserves the necessary quantities explicitly and prevent artificial dissipation. Similar results for the Schrödinger equation are well-known in the literature [43].

For Poisson's equation, we solve using standard fourth-order compact finite-difference methods [33, 93]. It is important to note that Poisson's equation must still be solved in three dimensions. Even though the charges are constrained to a plane, the electromagnetic energy is not confined. In particular, taking a cubic domain, the boundary conditions on the planes parallel to the graphene sheet can greatly impact the operation of the device. We demonstrate this effect in Section 3.4.5 and provide a connection to the role of gates and back-gates in the design of physical devices.

In Section 3.3.1, we prove that our scheme is consistent and of second order. We first show that the Crank-Nicolson scheme itself (corresponding to the zero mass and zero potential case) is von Neumann stable and thus convergent [53]. We then extend the results to the case of a constant potential and demonstrate that the scheme is still convergent and second order.

In Section 3.4 we show the results of implementing our scheme with a Matlab script. We show that our scheme accurately approximates simplified tests cases where the explicit solution is known and give numerical evidence of the second order convergence. We also demonstrate results in cases where explicit solutions are not known and demonstrate that our scheme can correctly capture interesting effects in graphene, specifically beam splitting and Veselago lensing [14].

Chapter 2

Photovoltaics

This section is the result of collaborative work with Klemens Fellner, Marie-Therese Wolfram, and Peter A. Markowich. Some of the material appeared in our accepted publication [10] and more should appear in a second publication [9] (in prep). Professor Markowich was responsible for the initial direction of the project, and Professor Fellner was instrumental in deciding the final form and motivating the second article. Professor Fellner also provided significant motivation for some of the calculations, although all of the material below was obtained by myself. Dr. Wolfram helped with the implementation of the computer code, accounting for approximately 30% of the C++ code. The simulations were done by myself alone.

2.1   Introduction

The search for cheap, environmentally sustainable energy solutions has led to intense interest and research in the area of organic photovoltaic (OPV) devices. Currently, these devices have solar efficiencies which are significantly below those of modern inorganic semiconductor devices (i.e. at present 8.3% compared to 27.6% for "one-sun" Gallium Arsenide and 42.3% for concentrator cells [37]), currently limiting commercial implementation.

A main difference of OPV devices (in contrast to inorganic PV) is the generation mechanism of free charge carriers, which typically occurs through the generation and dissociation of so called excitons, excited energy states created by incoming light. (We shall discuss excitons and the generation of free charge carriers in more detail below.)

Many mathematical models have been proposed to study the behavior of bilayer OPV cells as simple implementable organic devices: Refs. [5, 13, 15, 21, 50]. These models involve various approximations of the generation of free charges. A mathematical basis for the these models comes from extensive literature on inorganic semiconductor models. For these devices, the standard macroscopic models couple two drift-diffusion-reaction equations for the behavior of the free charge carrier densities (i.e. electrons and holes) in a system with Poisson's equation governing the self-consistent electrical potential [74]. Much is known about such systems, including existence and uniqueness results for the steady state (see e.g. [74, 55] and the references therein). The steady state is of particular interest for photovoltaic devices which are expected to generate power on time scales much longer than the duration of the transient (i.e. the switching-on) dynamics, observed to occur on the microsecond scale [47]. For a brief overview of classical semiconductor theory and standard notation, see Section 1.1.

2.1.1 OPV Model

Drift-diffusion-reaction type models for OPV devices must crucially take into account the specific material properties of inorganic and organic semiconductor materials, which involves in particular electric field dependent mobilities and specific recombination and dissociation rates (see Section 2.2 below). A second crucial amendment has to model the exciton dynamics and couple it appropriately to the original system.

We shall thus consider equations for the four main components, i.e. the concentrations of electrons (n), holes (p), and excitons (X) and the electric potential in the device (V), which must be calculated self-consistently.

Charge Carrier Densities

The evolution of the free charge carrier densities n and p shall be described by typical drift-diffusion-recombination equations and an additional source term due to the excitons:

$$q\frac{\partial n}{\partial t} = \nabla \cdot (qD_n\nabla n - q\mu_n n\nabla(U + V)) - qR_{np} + qk_dX, \qquad (2.1)$$

$$q\frac{\partial p}{\partial t} = \nabla \cdot (qD_p\nabla p + q\mu_p p\nabla(U + V)) - qR_{np} + qk_dX. \qquad (2.2)$$

Here q denotes the positive fundamental charge, D_n, D_p and μ_n, μ_p are the diffusion coefficients and mobilities of n and p, respectively, R_{np} is the recombination rate of

free electrons and free holes, and k_d is the rate of dissociation of excitons into a pair consisting of a free electron and hole. Functional expressions for diffusion coefficients, mobilities and recombination/dissociation rates in organic semiconductor materials will be specified in Section 2.2.

For a bilayer OPV device formed of two different organic semiconductor materials (see Figure 2.1), the potential U (which is not an electric potential and doesn't contribute to the electric field) models an additional convection term that comes from the differences in the energy levels of the two materials. In an organic device, the energy levels involved in charge transfer are the Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular Orbital (LUMO). The HOMO roughly corresponds to the valence band for classical semiconductors, and the LUMO resembles the conduction band (as in Figure 1.1). In general, the HOMO-LUMO gap corresponds to the band-gap.

Although this modeling is based extremely closely on various models in the literature (especially [5] and [13]), several deep problems exist with regards to the behavior of the minority carriers. We discuss these limitations in Section 2.8. In particular, minority carriers in the bulk can carry a nontrivial, nonphysical current - violating the principle of detailed balance in forward bias. This difficulty extends deep into the formulation of the model and since no simple fix is apparent, we carefully investigate the model as given before focusing specifically on its deficiencies.

Excitons

For most organic materials, the HOMO-LUMO gap is too large for a photon to create free electrons and holes. Instead, it creates a bound electron/hole pair, a so-called exciton. In the bulk material, these excitons usually recombine without producing free carriers (with rate k_r, see below). However, at an interface between two materials, a free electron can be generated in the material with a lower energy LUMO while a free hole is generated in the material with a higher energy HOMO. This reduced energy gap allows excitons to dissociate into free holes and electrons more readily at an interface between two materials. This effect is included in the model equations (2.1) and (2.2) by the term U, which represents the energy shift between the two materials. (Note that using the same U in both equations implies that each material has the same bulk HOMO-LUMO gap. This assumption simplifies the asymptotics in Section 2.5, but could probably be eliminated to account for more general materials.)

33

Excitons are typically called polaron pairs (also referred to as charge-transfer states or coulomb bound pairs) when aligned across the interface and are much more stable with respect to recombination. On the other hand, the aligned polaron pairs will either be pushed together or pulled apart by an electric field approximately parallel to the charge separation vector. Accordingly the dissociation rate (k_d) will be heavily field dependent. The combination of these two effects (reduced polaron pair recombination and field-driven dissociation) can lead to very high quantum efficiency (proportion of light creating free charges) for polaron pairs at an interface under an appropriately applied external potential [3]. Note that this is an internal quantum efficiency for polaron pair dissociation, not an external quantum efficiency, which further includes how many of the incoming photons actually create polaron pairs.

Following the above discussion we shall assume that every exciton becomes immediately a polaron pair at the interface (see further explanation below). We therefore consider the following diffusion equation for the evolution of a combined density of excitons and polaron pairs (to be identified by their position), which lacks a convection term since the excitons are electrically neutral:

$$\frac{\partial X}{\partial t} = \nabla \cdot (D_X \nabla X) + cR_{np} + G - k_d X - k_r X. \tag{2.3}$$

The coefficient G is the photo-generation rate for excitons, k_r is the rate of geminate recombination of the excitons, and c represents the proportion of recombining free charges that form excitonic states (as opposed to recombination leading to emission of light, etc). Note that although a dipole will not drift, an applied electric field will generally induce a torque on the dipole. Because the difference in material properties will generally force the dipoles to align to the interface, we assume that the torque generated by the electric field is negligible. In particular, this is justified for a device in which the interface is linear (or planar) and the dominant electric field term is perpendicular to the interface (as in Figure 2.1). See Section 2.9 for some discussion of the difficulties when the interface is no longer linear.

This equation, although apparently quite straightforward, can be quite complicated depending on the expected behavior of the coefficients. Because the device has two materials, the interface between the two materials is expected to be quite thin. It is originally tempting to immediately take the width of this interface to zero, but several issues arise with this approximation. See Section 2.8 for more details. We furthermore discuss in Section 2.6.1 that our chosen model for the polaron-pair diffusion results in
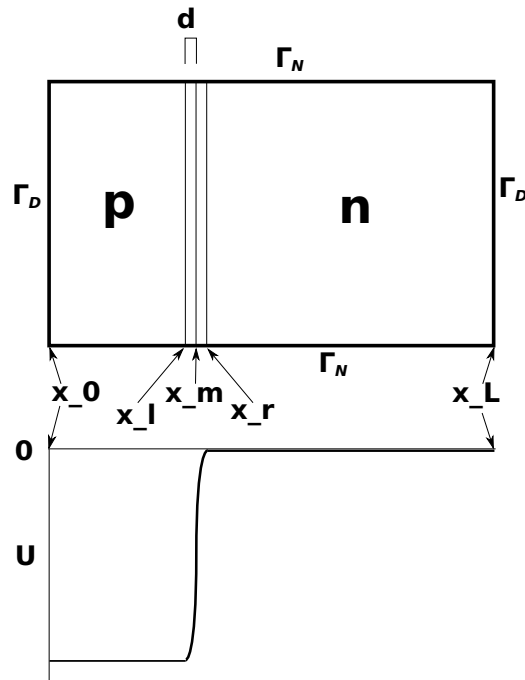
Figure 2.1: Schematic of the 2D device giving labels to the quantities used in the asymptotics and a graphical representation of the function U. For the numerics we take U to be piecewise quadratic connecting the values at either side so that $U \in C^1$. Note that for the asymptotics in Section 2.5 the vertical direction of the first diagram is suppressed and the problem becomes 1D.

non-physical behavior with regards to polaron-pairs diffusing away from the interface (a process generally not physically observed. We discuss one possible model extension for dealing with this difficulty in Section 2.9.

Poisson's Equation

Finally, we have Poisson's equation for the self-consistent electric potential:

$$\epsilon_0 \nabla \cdot (\epsilon_r \nabla V) = q(n - p - h\chi_I \nabla_\nu X), \qquad (2.4)$$

where ϵ_0 denotes the vacuum permittivity and ϵ_r is the relative permittivity of the material. Furthermore, ∇_ν represents the derivative normal to the direction of the material interface I and χ_I denotes the indicator function of the interface region while h is the separation length of the charge pairs of polaron pairs in the interface region. In fact, the extra term proportional to $\nabla_\nu X$ is the field contribution due to the alignment of the polaron pairs at the material interface. The gradient results from taking the continuum limit of a sum of Dirac delta primes assumed to be aligned normal to the interface (corresponding to point dipoles [48]). The exciton distribution X includes both polaron pairs and excitons, with their identity based on their location in the system. Polaron pairs are split over the interface such that the electron and hole (although still bound) each lie in an energetically favorable material. This has the dual effect of aligning a sheet of dipoles (resulting in the aforementioned term in Equation (2.4)) and of making polaron pair states extremely energetically favorable to excitons [3]. We thus assume that any X in an interface region I are polaron pairs and that any $X \in I^C$ (the complement of I) are excitons. In actual devices, organic interfaces are not sharp, and can be blended over a variety of length scales. For our model, we take I to be the region within $d = 1nm$ of the theoretical sharp interface. Note that if we take I to be a straight line parallel to the contacts of the device and assume homogeneity in the parallel direction(s), we exactly recover the model in [5]. No boundary conditions govern the edges of the interface region, but the sharp change in material parameters (as given in Table 2.1) do, however, generate fairly major changes in the charge carrier concentrations.

Domain and Boundary Values

For the remainder of this thesis, we will consider the domain $\Omega \in \mathbb{R}^2$. We take Ω to be rectangular, with Neumann boundary segments Γ_N on the top and bottom edges of the domain and Dirichlet boundary segments on the right and left edges (see Figure 2.1). Unless otherwise noted (as in Sections 2.9 and 2.10) we take the interface I to be linear and parallel to the Dirichlet contacts.

We supplement the system (2.1) to (2.4) with the following boundary conditions:

$$
\begin{cases} V = V_D, & n = n_D, & p = p_D & \text{on} & \Gamma_D, \\ \nabla_\nu V = 0, & \nabla_\nu n = 0, & \nabla_\nu p = 0 & \text{on} & \Gamma_N, \end{cases} \quad \text{and} \quad \nabla_\nu X = 0 \quad \text{on} \quad \Gamma \qquad (2.5)
$$

where the Dirichlet data (V_D, n_D, p_D) are constant on each Dirichlet boundary segment. The Dirichlet boundary conditions for Γ_D in Equation (2.5) correspond to the conducting ends of the device. The difference in the values for the potential V at the two Dirichlet contacts correspond to the potential difference in the device (in Volts) whose offset does not affect the dynamics of the device because only the electric field $E = -\nabla V$ enters the other equations. The usual notation for a photovoltaic device follows the convention scheme of a diode, but in the working case current flows in the direction opposite to that of a usual diode. For this reason, a consistent applied potential is given at the left boundary of the device. We thus take the potential at the right side of the device to be zero, and take the potential at the left side to be the potential difference. Note that this notation has the convenient consequence that the average field in the device has the same sign as the potential difference.

For the boundary conditions of electrons and holes one has to take into account that the HOMO/LUMO levels of the device determine different energies for the electrons and holes implying that the concentrations of n and p at the metallic contacts can be very different. Thus, for majority carriers in their energetically favorable material (p to the left of the interface and n to the right as indicated in Figure 2.1), we take the boundary conditions given by Scott and Malliaras in [80]. For the minority carriers (n to the left and p to the right), homogeneous Dirichlet boundary conditions constitute a very good approximation as the Scott-Malliaras boundary conditions give values approximately seven orders of magnitude smaller than typical concentrations (see Table 2.1 in Section 2.2.2). Note that although the minority carrier boundary conditions are taken to be homogeneous Dirichlet, this does not imply that the minority concentrations are zero (the case considered, for instance by [20, 73]). However, they are usually several orders

37

of magnitude smaller than the majority carrier – see the discussion at the beginning of Section 2.5.2. Further complicating the situation is that even though they are several orders of magnitude smaller, the minority carriers can still play a significant role in the device – see Section 2.8.

On the insulated boundary Γ_N we take homogeneous Neumann conditions for all of the variables because no particles should be able to pass into the insulator and it should be electrically neutral. For the excitons, we take homogeneous Neumann conditions on the whole boundary Γ because the excitons are not be able to transmit from the organic material into the inorganic contacts. Because the exciton equation contains no drift term, the exciton flux is exactly $D_X \nabla X$ and homogeneous Neumann conditions correspond to no-flux boundary conditions.

The first half of this chapter presents the results previously published in [10]. In Section 2.2 we will scale the physical parameters of the equations and detail the specific expressions for the mobilities, dissociation rates and recombination rates in an organic semiconductor material. Next, in Section 2.3 we present the numerical system used for the 2D simulations. In Section 2.5 we analyze the 1D equations in two physically relevant asymptotic regimes – large potential difference and small minority carriers. Moreover, in 2.6 we plot the calculated and asymptotic concentrations in the device at three important points in the device operation – short circuit, optimal power point, and open circuit. We interpret these results in terms of their relation to the asymptotic results and furthermore show the current-voltage characteristic curves generated by both the numerical and asymptotic methods.

We proceed to consider an alternative asymptotic scheme which focuses on implicitly defining the carriers in the device via integral equations. In Section 2.7 we demonstrate that this allows us to generate asymptotic curves which closely match the numerical results presented earlier. We furthermore use this system to demonstrate the physical ramifications of the minority carriers in Section 2.8. We intend to extend these results for a future publication [9]. Finally, we discuss the possibility of extending our numerical simulations to more complicated geometries (Section 2.9) and give several perspectives on proving the existence and uniqueness of the system (Section 2.10).

2.2 Scaling and Coefficient Modeling

In this section, we scale the system presented in Section 2.1.1 and discuss appropriate forms of the various coefficients defined therein. We begin by non-dimensionalizing the system, then proceed by defining the functional forms of the mobility and the dissociation terms, before finally giving physical values for all of our parameters.

We will scale (2.1) - (2.3) and (2.4). The basic scaling introduces the following dimensionless quantities:

$$V = \frac{k_b \mathrm{T}}{q}\tilde{V} = U_\mathrm{T}\tilde{V}, \qquad x = L\tilde{x}, \qquad (n, p, X) = N_r(\tilde{n}, \tilde{p}, \tilde{X}), \qquad \mu = \mu_0\tilde{\mu}, \qquad (2.6)$$

where U_T is the thermal voltage, L is a characteristic length (usually on the order of the device length), N_r is a reference concentration, and μ_0 is a reference mobility. We assume the Einstein's relation $D = U_\mathrm{T}\mu$. Although this might not be justified for some organic materials, it greatly simplifies the equations and as of yet a generally accepted alternative model has not been found. We further introduce the dimensionless quantities:

$$\lambda^2 = \frac{\epsilon_0 U_\mathrm{T}}{q L^2 N_r}, \qquad T = \frac{L^2}{\mu_0 U_\mathrm{T}}, \qquad (2.7)$$

where T denotes a reference time scale corresponding to the diffusion time of a free charge carrier in the bulk of the device. Note that the Debye length λ is typically not a small parameter – see Table 2.1.

For an OPV device under light we follow [5] and choose the Langevin recombination rate $R_{np} = \frac{q(\mu_n + \mu_p)np}{\epsilon_0\epsilon_r}$, which rescales with the reference time T as follows:

$$\frac{L^2}{N_r\mu_0 U_\mathrm{T}}R_{np} = \frac{1}{\lambda^2}\frac{(\tilde{\mu}_n + \tilde{\mu}_p)\tilde{n}\tilde{p}}{\epsilon_r} = c_r\tilde{n}\tilde{p}.$$

Setting $c_r' = cc_r$ this gives the following system:

$$\lambda^2\nabla\cdot(\epsilon_r\nabla\tilde{V}) = \tilde{n} - \tilde{p} - \frac{h\chi_I}{L}\nabla_\nu\tilde{X} \tag{2.8a}$$

$$\frac{\partial\tilde{n}}{\partial\tilde{t}} = \nabla\cdot(\mu_n\nabla\tilde{n} - \mu_n\tilde{n}\nabla(U + \tilde{V})) - c_r\tilde{n}\tilde{p} + k_d T\tilde{X} \tag{2.8b}$$

$$\frac{\partial\tilde{p}}{\partial\tilde{t}} = \nabla\cdot(\mu_p\nabla\tilde{p} + \mu_p\tilde{p}\nabla(U + \tilde{V})) - c_r\tilde{n}\tilde{p} + k_d T\tilde{X} \tag{2.8c}$$

$$\frac{\partial\tilde{X}}{\partial\tilde{t}} = \nabla\cdot(\mu_X\nabla\tilde{X}) + c_r'\tilde{n}\tilde{p} + GT - k_d T\tilde{X} - k_r T\tilde{X}. \tag{2.8d}$$

39

For the charge carrier mobilities μ_n and μ_p we shall apply the Poole-Frenkel model [31] postulating an exponential dependence on the square-root of the electric field $\tilde{E} = -\nabla \tilde{V}$:

$$\mu_n = \mu_n(0)e^{\gamma_n \sqrt{|\tilde{E}|}}, \qquad \mu_p = \mu_p(0)e^{\gamma_p \sqrt{|\tilde{E}|}}. \tag{2.9}$$

2.2.1 Dissociation Term

It is relatively common in the literature [13, 50] to derive exciton and polaron pair dissociation terms based upon the theory of electrolytic fluids outlined in [7, 36]. Several inconsistencies arise in the treatment which become especially important when we attempt to consider the zero-width interface limit. In this case, the polaron pairs will exist only on the interface, and thus they only enter the continuity equations as a distribution. However, since they will change the field locally around the interface, most forms of $k_d(E)$ will be ill-posed functions of distributions (often two delta distributions). Special care needs to be taken for the dissociation rate if an eventual zero-width model is to be successful.

In [13], we have the following integral for the average polaron dissociation rate:

$$k_d(E) = \int_0^\infty \hat{k}_d(E, a) f(a) da = C \int_0^\infty \frac{1}{a} e^{-a^2/a_0^2} da$$

where $\hat{k}_d(E, a)$ a specific dissociation rate depending on the electric field and the charge separation distance a. Here C depends only on E and temperature (and other physical constants, not a). The immediate problem is that this integral diverges for the $a \to 0$ limit. The proposed solution is to define a "physical lower bound" (usually $a_0/2 = .5$ Angstroms) corresponding to the minimal possible exciton separation width. Although this solves the divergence issue, the result is highly dependent on the value of this lower bound (which is not physically well understood).

This integral is cited from [36] for the distribution of what they call the charge-transfer radius, a. This represents the distance between the electron and the hole in the exciton state which lies on an interface between two different polymers. Aside from notational confusion between a, a_0 and r, the more interesting issue is that the integral in this paper is not calculated directly from k_d, but rather from a derived quantity

$$\hat{P}(E, a) = \frac{k_d(E)}{k_d(E) + k_r} \tag{2.10}$$

40

for the "the probability for a hole escape". This is the probability that the decaying particles will dissociate productively (with rate k_d) into a free electron and hole rather than recombining (with rate k_r). Using this notation, the integral over separation distances is given by:

$$P(E) = \int_0^\infty \hat{P}(E, a) f(a) da = C_1 \int_0^\infty \frac{a^2}{C_2 a^3 + 1} e^{-a^2/a_0^2} da$$

for some constants C_1 and C_2. This integral converges.

In [50] they rearrange equation (2.10) and then insert this relationship into the steady state equation for $\tilde{X}$:

$$0 = \nabla \cdot (\mu_X \nabla \tilde{X}) + c_r' \tilde{n} \tilde{p} + GT - (k_d + kr) T \tilde{X}$$
$$= \nabla \cdot (\mu_X \nabla \tilde{X}) + c_r' \tilde{n} \tilde{p} + GT - (1 + \frac{1-P}{P}) k_d T \tilde{X}$$

Since they only treat polaron pairs stuck to the interface (and not bulk excitons), the diffusion term also disappears ($\mu_X = 0$) and they can solve the equation algebraically for $k_d T \tilde{X}$ to obtain

$$k_d T \tilde{X} = P(GT + c_r' \tilde{n} \tilde{p}).$$

One can then insert this relationship into the continuity equations for n and p. This decouples X from the equations for n, and p. Although this simplifies the problem, it is not a feasible direction for a model in higher dimensions in which the polaron pairs themselves are free to move along the interface.

Given the difficulties with this type of integral k_d kernel, we take the polaron pair dissociation rate k_d given by [5] as a function of M, which is a scaled square-root of electric field E:

$$\begin{cases} k_d^-(M) = \frac{2k_d(0)}{M} \left(e^M \left(1 - \frac{1}{M}\right) + \frac{1}{M} \right), \\ k_d^+(M) = \frac{4k_d(0)}{M^2} \left(1 - e^{-M^2/4}\right), \end{cases} \qquad M = \frac{1}{\lambda} \sqrt{\frac{|E|}{L^2 U_T N_r \pi \epsilon_r}},$$

where the $+,-$ superscripts represent positive and negative fields (with respect to the interface normal). Note that this formula is continuous with respect to E, but is not continuously differentiable at $|E| = 0$. As discussed before, we consider all bound pairs within 1nm of the interface to be polaron pairs, and thus this formula only applies in the interface region. For the excitons in the bulk we take a constant dissociation rate since excitons are not generally aligned to the electric field. Note that this means that

41

although k_d is continuous with respect to E, it has spatial discontinuities at the interface edges corresponding to the increased stability of polaron pairs compared to excitons - see Table 2.1. In Section 2.6 we also consider a field-independent dissociation rate and discuss the impact of such a change on the device characteristics.

2.2.2 Device Parameter Values

The physical scaling parameters we take are:

$$V = U_{\mathrm{T}}\tilde{V} = .0258\mathrm{V}\,\tilde{V}$$

$$E = U_{\mathrm{T}}\tilde{E}/L = 2.58 \times 10^5\mathrm{V/m}\,\tilde{E}$$

$$x = L\tilde{x} = 100\mathrm{nm}\,\tilde{x}$$

$$n = N_r\tilde{n} = 10^{20}\mathrm{m}^{-3}\,\tilde{n}.$$

The values for the various constants are taken from [5] (except G which is converted to a spatial density instead of an area density). Table 2.1 collects all the used dimensionless parameters. For the field dependent parameters we also give the values at the short circuit electric-field $E_{\mathrm{SC}} = 3.5 \times 10^6$ (with both signs where applicable), corresponding to the short circuit potential difference of $\mp.5V$. This gives a qualitative understanding of the range of the The expressions k_{out} and k_{in} refer to the rate-values in- and outside of the interface region.

Table 2.1: Values for physical parameters.

$(h/L) = .01$	$(d/L) = .01$	$\epsilon_r = 4$
$\lambda^2 \approx 1.43$	$T = .00386$	$\tilde{E}_{\mathrm{SC}} = 13$
$GT \approx 16990$	$c_r = c_r' \approx .6987$	$U(x_r) - U(x_l) \approx 6$
$k_{d,out}T \approx 1$	$k_{r,in}T \approx 3.86$	$k_{r,out}T \approx 3864$
$k_{d,in}(0)T \approx 386$	$k_{d,in}(+E_{\mathrm{SC}})T \approx 178$	$k_{d,in}(-E_{\mathrm{SC}})T \approx 2763$
$\mu_0 = 10^{-10}$	$(\mu_1/\mu_0) \approx .01$	
$\mu_n(0) = 3$	$\gamma_n = .788$	$\mu_n(E_{\mathrm{SC}}) \approx 53.3$
$\mu_p(0) = 1$	$\gamma_p = .153$	$\mu_p(E_{\mathrm{SC}}) \approx 1.75$
$n(x_L), p(x_0) \approx .04$	$n(x_0), p(x_L) \approx 4 \times 10^{-7}$	

The values of μ_X and $k_{d,out}$ are not easy to calculate physically and no consensus seems to exist in the literature for their values, but we take the given values based on reasonable estimates given in the sources listed above.

2.2.3 Steady State Equations

For usual device operation, a solar cell will be producing current steadily for time-scales on the order of hours. Any transient behavior occurs over such short time-scales (i.e. microseconds [47]) that we neglect them for modeling the long-term behavior and efficiency of the device.

Therefore, for the remainder of the chapter (except Section 2.10 which contains some parabolic regularity results), we will consider only the steady state equations where we drop all the tildes and absorb the time-scaling T into the rates:

$$\lambda^2 \nabla \cdot (\epsilon_r \nabla V) = n - p - \frac{h\chi_I}{L}\nabla_\nu X \qquad (2.11a)$$

$$-\nabla \cdot (\mu_n \nabla n - \mu_n n \nabla(U + V)) = -c_r np + k_d X \qquad (2.11b)$$

$$-\nabla \cdot (\mu_p \nabla p + \mu_p p \nabla(U + V)) = -c_r np + k_d X \qquad (2.11c)$$

$$-\nabla \cdot (\mu_X \nabla X) = c_r' np + G - k_d X - k_r X. \qquad (2.11d)$$

We expect that this system of equations will have a unique steady state solution. If we insist on the physically reasonable requirement that k_d and μ are smooth and bounded from above (corresponding to physical device saturation), then the first three equations fit very nearly into the standard semiconductor framework. See Section 2.10.1 for details.

Concerning the steady state solutions, we emphasize that the light generation term $G > 0$ generally implies a non-zero right-hand-side in Equations (2.11b) and (2.11c) and thus non-constant drift-diffusion fluxes of n and p on the left-hand-side. Thus, we expect a steady flux of electrons and holes away from the interface with nearly negligible bimolecular recombination (due to the work function considerations discussed in the introduction). Even for $G = 0$, we will only recover constant stationary fluxes when the equation

$$\frac{c_r'}{c_r} k_d X - k_d X - k_r X = 0$$

is satisfied. This can hold in particular cases such as $X = 0$ and $c_r = \frac{c_r' k_d}{k_d + k_r}$, but these are not physically realistic given our device parameters. As a consequence, the *model system* (2.11) *is not designed to accurately predict the behavior of the device in the dark* $G = 0$.

2.3 Numerical Scheme

We present a numerical scheme based on a Gummel-type iteration [39], which is well-suited to deal with the nonlinear dependence of the equations for n, p, and X on the electric field ($E = -\nabla V$). At each Gummel-iteration step we calculate first the electric potential V from the current state of n, p, X and then use this new electric potential to update n, p, X. We can write this as follows after the kth step:

$$
-\lambda^2 \nabla \cdot (\epsilon_r \nabla V_{k+1}) = p_k - n_k + \frac{h}{L}\nabla X_k
$$

$$
-\nabla \cdot (\mu_n(E_{k+1})\nabla n_{k+1} - \mu_n(E_{k+1})n_{k+1}\nabla(V_{k+1} + U)) + c_r n_{k+1}\, p_k = k_d(E_{k+1})X_k
$$

$$
-\nabla \cdot (\mu_p(E_{k+1})\nabla p_{k+1} + \mu_p(E_{k+1})p_{k+1}\nabla(V_{k+1} + U)) + c_r n_k\, p_{k+1} = k_d(E_{k+1})X_k
$$

$$
-\frac{\mu_1}{\mu_0}\nabla \cdot (\mu_X \nabla X_{k+1}) + (k_d(E_{k+1}) + k_r)X_{k+1} = c_r' n_k p_k + G.
$$

$$(2.12)$$

Note that we use a semi-implicit discretization for each equation. Specifically, we implicitly discretize the flux and mass terms in (2.12), but use semi-implicit discretization for the recombination term $c_r np$ and an explicit form for the dissociation term $k_d X$. This has the benefit of decoupling the system so that n, p, and X can be effectively updated in parallel since they only require the previously updated values for V. Furthermore, each variable n, p, X is treated completely implicitly within its equation, which greatly reduces numerical instabilities arising from the errors in the previous step.

The iteration continues until a preset L^2 error between the two latest solutions is achieved. In practical numerical simulation, we have found this Gummel iteration to converge well except for high positive values for the applied potential difference (which constitutes the non-working regime for a photovoltaic device). This case, which is characterized by opposite convection terms coming from the work function and the electric potential, leads to large values for n and p near the interface which prevent the Gummel iteration from converging – see Section 2.6.4. To alleviate this difficulty, we add a damping parameter α which interpolates between old and new solutions. More precisely, denoting with V_k', n_k', p_k', X_k' the solutions of the above system (2.12), the damped scheme is:

$$
\begin{aligned}
V_{k+1} &= \alpha V_{k+1}' + (1-\alpha)V_k, & n_{k+1} &= \alpha n_{k+1}' + (1-\alpha)n_k, \\
p_{k+1} &= \alpha p_{k+1}' + (1-\alpha)p_k, & X_{k+1} &= \alpha X_{k+1}' + (1-\alpha)X_k.
\end{aligned}
$$

$$(2.13)$$

In particular, note that since V'_{k+1} and V_k both satisfy linear boundary conditions, so does any convex combination of V'_{k+1} and V_k. This is equally true for the boundary conditions for n,p, and X, which are fixed for a given potential V_{k+1}.

Our numerical experiments in the case of positive applied potentials show that it is possible to optimize the convergence by tuning the damping parameter. In fact it turns out that using a small damping $\alpha \approx .01$ for the first few (e.g. three) iteration steps serves to correct the initial guess and prevents the concentrations from going negative due to badly chosen initial data. Further optimization on the damping, such as a potentially dynamic choice of α, is considered in Section 2.6.4.

Note also that the damped scheme could be modified to represent the parabolic system with a modified Euler method. If the current values (V_k, n_k, p_k, X_k are taken with coefficient 1, then α corresponds to the time step. As discussed above, however, the right-hand side is evaluated semi-implicitly and the equivalence of the iteration scheme to a specific time-dependent model is not immediate.

Space-Discretization

Our choice of a space-discretization has to take into account the two major difficulties inherent in our system. First of all, the material interfaces will cause very strong changes in the densities n, p, and X in the interface region. Secondly, depending on the electric field, the equations can be either convection or diffusion dominated. Furthermore, we want our scheme to easily generalize to complicated 2D interfaces.

The hybrid discontinuous Galerkin (HDG) finite element method is well-suited to handle these challenges. The novel idea for the HDG method is to have separate degrees of freedom on the elements u (an element in the usual sense) and on the facets (boundaries) of the elements u_F. We assume that our domain is decomposed into a mesh $\mathcal{T}$ consisting of triangles with a set of facets $\mathcal{F}$ consisting of the facets (edges) of these triangles. The space we use for each of the equations is $V' = \{(u, u_F) : u \in H^1(\Omega) \cap H^2(\mathcal{T}), u_F \in L^2(\mathcal{F})\}$. Our solution space is $(V, V_f, n, n_f, p, p_f, X, X_f) \in V = V' \times V' \times V' \times V'$ and we denote the approximate solutions to the four equations as u_V, u_n, u_p, u_X with corresponding facet functions $u_{Vf}, u_{nf}, u_{pf}, u_{Xf}$. The facet elements allow for upwind-type calculations for convection-dominated behavior, as well as allowing for large changes of behavior at material boundaries. The functions within the elements are approximated by polynomials of a given order (seven for the results given in Section 2.6). Both in-

creasing the polynomial order and refining the mesh improve the properties of a solution after a given number of iterations at the price of larger computational costs.

A solid introduction to HDG methods is [52], specifically the first section on the general convection-diffusion case. The HDG method can more precisely be called a hybridized symmetric interior penalty discontinuous Galerkin method. The relationship to standard discontinuous Galerkin methods runs quite deeply. The stabilization and symmetrization terms have close analogs in traditional DG methods, see for instance [4, 16, 17].

We rewrite the system (2.12) in the suitable HDG weak form below. For notational convenience, we write the variable without the subscript indicating the equation step, and instead put a bar above the old value of the variable, ie $V_{k+1} \to V$ and $V_k \to \widehat{V}$). Furthermore, we define the variable

$$(n/p)^{up} = \begin{cases} u & \nabla_\nu V > 0 \\ u_F & \nabla_\nu V < 0 \end{cases}$$

so that for an outflow boundary we choose the value given by the bulk variable, but for an inflow boundary we choose the facet value. The additional term of $\mu(E)$ in the convection term has been removed without changing $(n/p)^{up}$ since it is a strictly positive scalar. Note that as before a subscript of ν represents a normal derivative.

We seek $(V, V_f, n, n_f, p, p_f, X, X_f) \in V$ such that:

$$\sum_{T \in \mathcal{T}} \left(\int_T \lambda^2 \epsilon_r \nabla V \cdot \nabla u_V dx - \int_{\partial T} \lambda^2 \epsilon_r \frac{\partial u_V}{\partial n} (V - V_f) ds - \int_{\partial T} \lambda^2 \epsilon_r \frac{\partial V}{\partial n} (u_V - u_{Vf}) ds \right.$$

$$\left. + \int_{\partial T} \frac{A}{h} \lambda^2 \epsilon_r (V - V_f)(u_V - u_{Vf}) ds \right) = \sum_{T \in \mathcal{T}} \int_T (\widehat{p} - \widehat{n} + h/L \nabla_x \widehat{X}) u_V dx + \int_{\Gamma_D} \tau (V_D - V) ds$$

$$\sum_{T \in \mathcal{T}} \left(\int_T \mu_n(E) \nabla n \cdot \nabla u_n dx - \int_{\partial T} \mu_n(E) \frac{\partial u_n}{\partial n} (n - n_f) ds - \int_{\partial T} \mu_n(E) \frac{\partial n}{\partial n} (u_n - u_{nf}) ds \right.$$

$$+ \int_{\partial T} \frac{A}{h} \mu_n(E)(n - n_f)(u_n - u_{nf}) ds - \int_T \mu_n(E) \nabla V n \nabla u_n dx + \int_{\partial T} (\mu_n(E) \nabla V)_n n^{up} u_n ds$$

$$\left. + \int_{\partial T_{out}} (\mu_n \nabla V)_n (n_f - n) u_{nf} ds + \int_T c_r n \widehat{p} u_n dx \right) = \sum_{T \in \mathcal{T}} \int_T k_d(E) \widehat{X} u_n dx + \int_{\Gamma_D} \tau (n_D - n) ds$$

46

$$\sum_{T \in \mathcal{T}} \left(\int_T \mu_p(E) \nabla p \cdot \nabla u_p dx - \int_{\partial T} \mu_p(E) \frac{\partial u_p}{\partial n}(p - p_f)ds - \int_{\partial T} \mu_n(E) \frac{\partial p}{\partial n}(u_p - u_{pf})ds \right.$$

$$+ \int_{\partial T} \frac{A}{h} \mu_p(E)(p - p_f)(u_p - u_{pf})ds + \int_T \mu_p(E) \nabla V p \nabla u_p dx - \int_{\partial T} (\mu_p(E) \nabla V)_\nu n^{up} u_p ds$$

$$\left. - \int_{\partial T_{out}} (\mu_p \nabla V)_\nu (p_f - p) u_{pf} ds + \int_T c_r \widehat{n} p u_p dx \right) = \sum_{T \in \mathcal{T}} \int_T k_d(E) \widehat{X} u_p dx + \int_{\Gamma_D} \tau(p_D - p)ds$$

$$\sum_{T \in \mathcal{T}} \frac{\mu_1}{\mu_0} \left(\int_T \mu_X \nabla V \cdot \nabla u_X dx - \int_{\partial T} \mu_X \frac{\partial u_X}{\partial n}(X - X_f)ds - \int_{\partial T} \mu_X \frac{\partial X}{\partial n}(u_X - u_{Xf})ds \right.$$

$$\left. + \int_{\partial T} \frac{A}{h} \mu_X(X - X_f)(u_X - u_{Xf})ds + \int_T (k_d(E) + k_r) X u_X dx \right) = \sum_{T \in \mathcal{T}} \int_T (c'_r \widehat{n}\widehat{p} + G_x) u_X dx$$

for all test functions $(u_V, u_{Vf}, u_n, u_{nf}, u_p, u_{pf}, u_X, u_{Xf}) \in V$.

It is important to note that quantities of the form $(V_f - V)$ are equal to zero for the true continuous solution, so they all drop out in the analysis (this is the "interior penalty" discussed earlier). These terms are added to the numerics to symmetrize the Laplacian. The integrals multiplied by terms of the form $(V_f - V)(u_V - u_{Vf})$ are also zero and are included with a factor of $A = 10$ to stabilize the system. This term is usually called α in the literature, but we use A here since α refers to our damping parameter.

We take the boundary conditions given in equation (2.5). However, implementing Dirichlet conditions directly will cause issues for the stability of the system. Instead, we define a parameter $\tau = 10^{10} >> 1$. and regularize the boundary condition $u = u_d$ giving $(u + 1/\tau \nabla_\nu u = u_d)$ (on Γ_D). In the weak formulation this gives a penalty term of the form $\int_{\Gamma_D} \tau(u - u_d)ds$. The Neumann conditions are all homogeneous, and thus add no terms to the weak formulation.

All of the numerical examples in Section 2.6 consider the two-dimensional computational domain $\Omega = [0, 1.5] \times [0, 0.2]$ (in the scaled variables from Section 2.2). In order to compare with the one-dimensional asymptotic analysis given in Section 2.5, we take the domain interfaces (see Figure 2.1) to be straight vertical lines located at $x_0 = 0$, $x_l = 0.49$, $x_m = 0.5$, $x_r = 0.51$, and $x_L = 1.5$ respectively. In Sections 2.4 and 2.9, we show example plots for other interfaces which oscillate around these linear boundaries.

The mesh generation was accomplished using the Netgen program created by Schöberl [78]. In particular, the mesh generation automatically preserves the internal interfaces allowing all coefficients to be defined piecewise depending on the properties of the local material. The numerics were done using NGSolve, a solver wrapper for the Netgen pro-

gram. We also included the inverse solver PARDISO to assist with the non-symmetric matrix calculations [76, 77]. See [96] for more details about using the NGSolve program on convection-diffusion equations and Appendix 2.B for details on our implementation.

2.4 Sample Results

Here we present sample results from our numerical scheme. Of primary interest is understanding the qualitative behavior of the device. In the next section we will identify asymptotic regimes with which to understand the device analytically. Because we seek to model a physical device, we seek asymptotics which approximate the results in both a mathematical sense and a physical sense. From a mathematical perspective, the goal is to completely understand the behavior of the model outlined in Section 2.1.1. From an engineering perspective, we would like to understand the operational characteristics of the device. The most important of these factors can be understood by plotting the current though the device versus the potential difference – an IV Curve. (Note that the term "IV curves" uses I instead of J for the current. In general, J is used to prevent confusion with $i = \sqrt{-1}$. For our purposes it also frees I to be used for the interface region.)

2.4.1 Internal Variables

First we consider the results for the device at short circuit. As we will discuss in the following section, this operating voltage is specifically important because it indicates the maximal amount of current that a device can produce. Plots of the concentrations n and p and the two components of the electric field E_x and E_y are given in Figure 2.2. Here the interface is sinusoidal, with the unit normal defined analytically. Details on creating a nonlinear interface with constant width are given in Section 2.9 and Appendix 2.B.

Several important aspects are immediately obvious from the plots. First, all of the significant effects seem to happen orthogonal to the interfaces. Near the metallic contacts we see boundary layers forming, but no effects in the direction parallel to the boundary. Near the organic-organic interface, I, we again see that the primary effects occur perpendicular to the interface. This suggests that it might be possible to understand the dominant behavior of a 2D device with 1D asymptotics. In Section 2.5 we proceed to analyze the 1D system under this assumption.

48
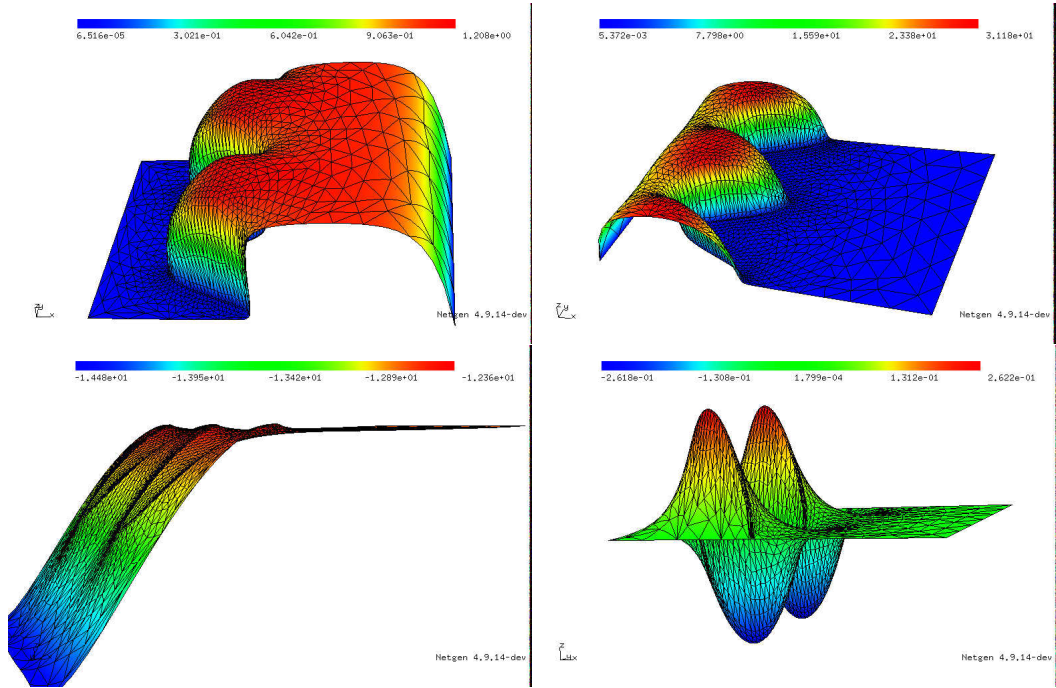
Figure 2.2: Sample simulation results for a 2D device at short circuit with a sinusoidal interface. The first row shows the electron and hole concentrations and the second row shows the electric field in the x and y directions (to the left and right respectively for both rows). Note that the electric field in the x direction dominates (the field in the y direction is two orders of magnitude smaller). In addition, the field stays within 10% of the estimated value $E_{\text{SC}} = 13$ given in Table 2.1.

Further evidence for the 1D model is suggested by the electric field plots in the second row of Figure 2.2. The electric field in the y direction, E_y is approximately two orders of magnitude smaller than E_x. In addition, it is nearly zero away from the interface. Since the convection effects near the interface will be dominated by U, E_y will not be critical to the device function. In addition, we see that $E_x \approx E_{\mathrm{SC}}$ (as defined in Table 2.1). The field varies in the p-region of the device, but remains within 10% of the average field value E_{SC}. In Section 2.5.1 we use this observation to expand the results in terms of deviations from the average field.

Additionally, as expected based on the physics of the device, n is very small in the p-type region to the left of the device and p is very small in the n-type region to the right of the device (refer back to Figure 1.2). This suggests taking the values of the minority carriers to be zero (as considered in Section 2.5.2). We consider this assumption (which is very common in the literature, especially for asymptotic results [5, 15, 73]) more carefully in Section 2.8, showing that it cannot be derived from our chosen model in a physically consistent way.

2.4.2 *IV* Curves

One of the most important characteristics of an OPV device is the effect of changing the applied voltage on the current in the device. We show this relationship by plotting the current as a function of potential difference using multiple simulations with different boundary values. We pick values for V_{diff} and then use these values to calculate the boundary values for n and p according to [80]. Due to the built-in potential of the metallic contacts at the OPV device, we have that zero applied voltage leads to an approximate potential difference over the device of $V_{\mathrm{diff}} = -19.3$ (which we take to be the left edge of the *IV* curves here and below).

Figure 2.3 plots the current for our typical OPV device green) with respect to the potential difference (as well as the power (blue), which is discussed below). This referential *IV* curve compares very closely with a second *IV* curve (red) computed after replacing the stated boundary conditions (from [80]) by homogeneous Dirichlet boundary conditions for each carrier at both Dirichlet boundaries. These conditions (taken for simplicity only) demonstrate that the observable device behavior is relatively independent of boundary conditions.

Both *IV* curves show how the produced current decreases (in magnitude) as the
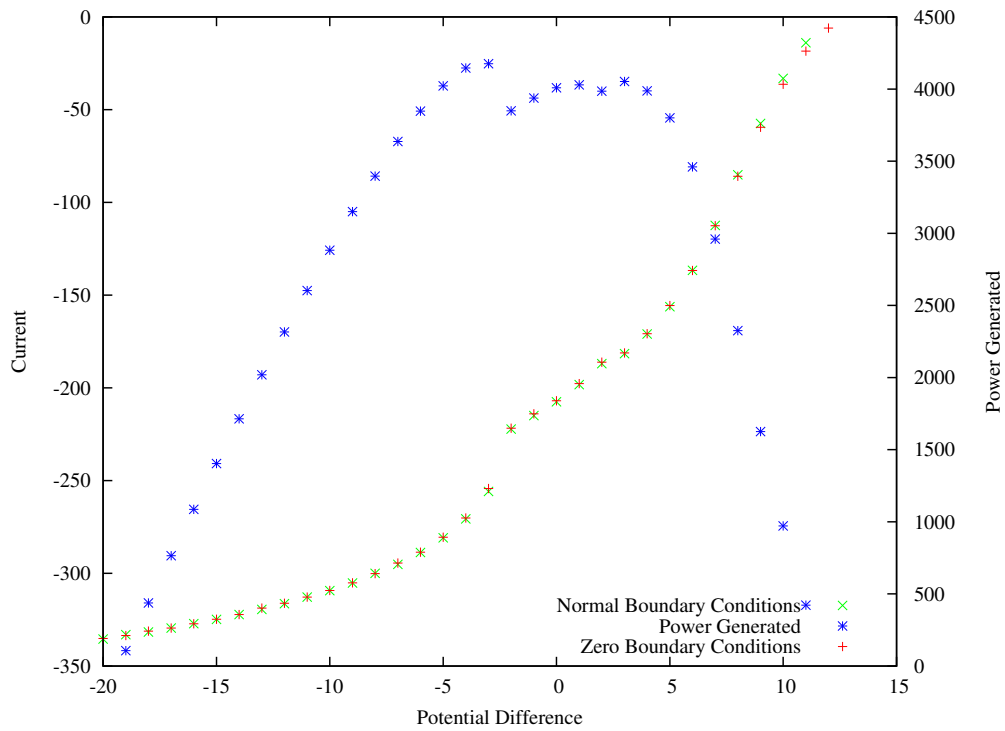
Figure 2.3: A comparison of the IV curves obtained using the boundary conditions outlined above in Table 2.1 above and homogeneous Dirichlet conditions for the charge carriers and a plot of the generated power density for the usual boundary conditions. Note that the power corresponds to the areas of the rectangles with opposite corners at the lower-left and the plotted current at each potential difference. We recall from Table 2.1 that the the scaled internal voltage is taken to be $V_{int} = 19.3$. A different internal voltage would significantly alter the generated power curve.

potential difference increases. In particular we observe a distinctive kink. This is an effect of the dependence of $k_{d,in}(E)$ on the self-consistent electric field E. The kink arises at the potential $V \approx -3$ for which the E-field touches zero from below at the interface. Note that the E-field touching zero happens for a potential V smaller than zero because the holes on the left side of the device (i.e. the majority charges) increase self-consistently the negative applied E-field towards the interface. This argument can be verified by plotting the IV curve obtained from a constant dissociation rate $k_{d,in} = k_{d,in}(E_{\mathrm{SC}})$. A plot comparing this result to the asymptotic results given in Section 2.5.1 is shown in Figure 2.4. Additionally, the effect is significantly reduced by decreasing the exciton mobility μ_X (see Section 2.6.1).
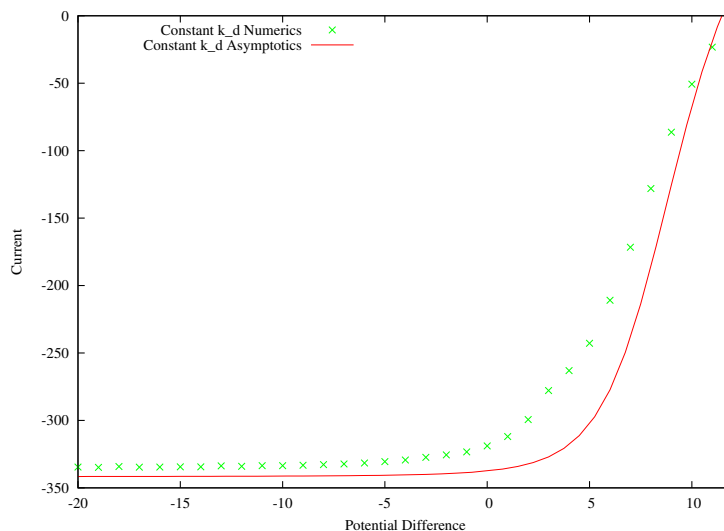
Figure 2.4: A comparison of the IV curves obtained using the numerical and asymptotic schemes with $k_{d,in} = k_{d,in}(E_{\mathrm{SC}}) = 2763$ to be constant. Reprinted below as Figure 2.13.

The power density generated by a device is given by $P = J(V_{\mathrm{int}} + V_{\mathrm{diff}})$. As shown in Figure 2.3, the power (in blue) is zero for two specific characteristic points in the device: $V_{\mathrm{diff}} = -V_{\mathrm{int}}$ (short circuit) and $J = 0$ (open circuit). A third characteristic point, the optimal power point, is defined by the voltage point with the maximal value of P. For potential differences outside of the regime shown in Figure 2.3 the power generated is negative, indicating that the device requires power input to operate and is no longer in the working regime for a photovoltaic device. An upper bound for the maximum power is given by the product of the short circuit current and the open circuit voltage (these

being the maximum current and maximum voltage for the device in the working regime). The *fill factor* denotes the ratio of the maximum obtainable power and this upper bound [22]. Geometrically this is represented by the largest possible rectangle contained under the IV curve and the smallest rectangle containing the IV curve. For the plots given in Figure 2.3 the maximal power point occurs at a potential difference of $V \approx -3$ with a fill factor of 39.5%. This fill factor is in good agreement with other simplified organic devices [37]. In general, fill factors vary heavily depending on the device, and are notably improved for multijunction devices.

2.5   1D Stationary State Asymptotics

In this section we discuss two asymptotic regimes for analyzing System (2.11) in 1D. First, we discuss the large-field approximation. Unlike previous results [60], we do not eliminate the diffusion fluxes, but rather switch to the Slotboom variables (quasi-Fermi-levels) and solve the differential equation system asymptotically in the high-field limit. This gives reasonable results for large potential differences, but begins to fail as the potential difference approaches zero. A more intricate approximation scheme with the same general structure is presented in Section 2.7. Although the latter scheme gives results which more accurately represent the numerical simulations, the large-potential asymptotics still give interesting results (especially for the short-circuit case) and warrant individual consideration.

In Section 2.5.2 we discuss the zero-minority carrier limit which is partially justified in the large-field limit discussed in Section 2.5.1. The zero-current (open circuit) case was given in the literature by [15], while the explicit results here were developed independently in [73]. See Section 2.8 for a discussion of the limitations of the zero-minority carrier limit.

We plot these results along with numerical results in Section 2.6. Several issues become apparent as we carefully analyze the numerical and asymptotic results side-by-side. We discuss possible issues with exciton and polaron pair diffusion in Section 2.6.1. Possible resolution of this issue is further discussed in Section 2.9 with regards to extending the numerical simulations to full 2D problems.

2.5.1 Large Applied Field (Constant μ) Approximation

We begin by considering the steady state equations in 1D:

$$\lambda^2 \epsilon_r V_{xx} = n - p - \frac{h\chi_I}{L} X_x \tag{2.14a}$$

$$-(\mu_n n_x - \mu_n n(U+V)_x)_x = -c_r np + k_d X \tag{2.14b}$$

$$-(\mu_p p_x + \mu_p p(U+V)_x)_x = -c_r np + k_d X \tag{2.14c}$$

$$-\mu_X X_{xx} = c_r' np + G - k_d X - k_r X, \tag{2.14d}$$

together with the boundary conditions

$$n(x_0) = n_0 \qquad p(x_0) = p_0 \qquad V(x_0) = V_{\text{diff}} \qquad X_x(x_0) = 0$$

$$n(x_L) = n_L \qquad p(x_L) = p_L \qquad V(x_L) = 0 \qquad X_x(x_L) = 0$$

where we have assumed that the ϵ_r and μ_X are spatially homogeneous.

In reverse bias, which constitutes most of the operating state of an OPV bilayer device, a negative voltage $V_{\text{diff}} < 0$ is given at the left boundary of the device. Note that with this notation, V_{diff} is the potential difference in the device. For working photovoltaic cells, this potential difference is the combination of the potential applied to the device and a built-in potential from the metallic contacts.

With a constant relative permittivity ϵ_r we introduce the Debye length $\lambda_D^2 := \lambda^2 \varepsilon_r \approx 5$, and rescale the potential according to this potential difference, i.e. $V \to |V_{\text{diff}}|V$ which rewrites Equation (2.14a) as

$$V_{xx} = \varepsilon \left(n - p - \frac{h\chi_I}{L} X_x \right), \qquad \varepsilon := \frac{x_L - x_0}{\lambda_D^2 |V_{\text{diff}}|}, \qquad V(x_0) = -1, \quad V(x_L) = 0. \tag{2.15}$$

Since in a working device a typical short circuit voltage is about -0.5 Volts [5], which is $|V_{\text{diff}}| \approx 19.3$ in our units, we find that physically $\varepsilon \approx 0.01$.

Hence (assuming for the moment that $\varepsilon \frac{h}{L} X_x \ll 1$, where $\frac{h}{L}$ is the fraction of interface width to device width) we expect that the electric potential V is in zeroth ε-order dominated by the potential difference and thus that the electric field is constant in the zeroth order of ε, i.e.

$$V^0(x) = V_{\text{diff}} - E^0(x - x_0), \qquad E^0 := \frac{V_{\text{diff}}}{x_L - x_0} = \pm \frac{1}{\lambda_D^2 \epsilon}. \tag{2.16}$$

We shall see in the following that the approximation (2.16) remains consistent with the assumption $\varepsilon \frac{h}{L} X_x \ll 1$ after being carried over to the equations for the charge carriers and the excitons .

Next, we remark that with the zeroth order approximation E^0, the mobilities μ_n and μ_p are also constant in the zeroth order. For $\epsilon < .01$ (the physical short-circuit value), we calculate:

$$\mu_n = \mu_n(0)e^{\gamma_n \sqrt{|E^0|}} = \mu_n(0)e^{\gamma_n \sqrt{\left|\frac{1}{\lambda_D^2 \varepsilon}\right|}} \gg \frac{1}{100\varepsilon}$$

$$\mu_p = \mu_n(0)e^{\gamma_p \sqrt{|E^0|}} = \mu_p(0)e^{\gamma_p \sqrt{\left|\frac{1}{\lambda_D^2 \varepsilon}\right|}} \gg \frac{1}{\varepsilon}$$

and thus $1/(\mu_n\mu_p) \ll O(\varepsilon^2)$.

Thus, by rescaling

$$\mu_n n \rightarrow n \quad \text{and} \quad \mu_p p \rightarrow p, \tag{2.17}$$

we obtain with $c_r = O(1)$ the following rescaled charge carrier equations

$$-(n_x - n(V+U)_x)_x = -O(\varepsilon^2)\,np + k_d X, \quad n_0 := n(x_0) \ll 1,\ n_L := n(x_L) = O(1), \tag{2.18}$$

$$-(p_x + p(V+U)_x)_x = -O(\varepsilon^2)\,np + k_d X, \quad p_0 =: p(x_0) = O(1),\ p_L := p(x_L) \ll 1. \tag{2.19}$$

(Note that with this scaling, the scales for n and p in Figure 2.2 become identical, with shared maximum values $\mu_n n \approx \mu_p p \approx 60$.)

Moreover, the rescaled equation for the excitons is

$$-\mu_X X_{xx} = O(\varepsilon^2)\,np + G - (k_d + k_r)X, \qquad X_x(0) = X_x(x_L) = 0, \tag{2.20}$$

where $\mu_X \ll 1$ is small.

Letting ε go to zero, we can solve this equation explicitly in terms of hyperbolic sines and cosines. This allows us to check the consistency the approximation (2.16) and the rescaling (2.17) with the underlying assumption $\varepsilon \frac{h}{L} X_x \ll 1$. (See Appendix 2.A.1 for detailed calculations.) With $k_{r,out} \approx 4 \cdot 10^3$ and $k_{d,out} \approx 1$ outside the interface and with $k_{r,in} \approx 1$ and $k_{d,in} \approx 3 \cdot 10^3$ (for typical $E^0 < 0$) inside the interface we have:

$$\|X_x^0\|_{L^\infty([x_0,x_L])} \approx \frac{G}{\sqrt{\mu_X}}\left(\frac{1}{\sqrt{k_{d,in}+k_{r,in}}} - \frac{1}{\sqrt{k_{d,out}+k_{r,out}}}\right) \tag{2.21}$$

and thus $\varepsilon \frac{h}{L}\|X_x^0\|_{L^\infty([x_0,x_L])} \approx O(\varepsilon)$.

Zeroth Order Charge Carriers

Now, we proceed to study the zeroth order solutions n^0 and p^0 of (2.18) and (2.19). Introducing the zeroth order fluxes

$$J_n^0 := n_x^0 - n^0\varphi_x, \qquad J_p^0 := -(p_x^0 + p^0\varphi_x), \qquad \text{with} \quad \varphi = V^0 + U,$$

and neglecting the $O(\varepsilon^2)$ recombination term, the equations (2.18) and (2.19) integrate immediately as

$$J_n^0(x) = J_{n0}^0 - F(x), \qquad J_p^0(x) = J_{p0}^0 + F(x), \tag{2.22}$$

$$F(x) := \int_{x_0}^x k_d(y) X^0(y) dy.$$

Utilizing the Slotboom variables $J_n^0 = e^\varphi (n^0 e^{-\varphi})_x$ and $J_p^0 = -e^{-\varphi}(p^0 e^\varphi)_x$, also called quasi-Fermi levels, we can explicitly solve for n^0 and p^0:

$$n^0(x) = n_0 e^{\varphi(x)-\varphi_0} + J_{n0}^0 \int_{x_0}^x e^{\varphi(x)-\varphi(y)} dy - \int_{x_0}^x F(y) e^{\varphi(x)-\varphi(y)} dy \tag{2.23}$$

$$= n_0 e^{\varphi(x)-\varphi_0} + J_{n0}^0 \Phi_n(x) - \mathcal{F}_n(x),$$

$$p^0(x) = p_0 e^{\varphi_0-\varphi(x)} - J_{p0}^0 \int_{x_0}^x e^{\varphi(y)-\varphi(x)} dy - \int_{x_0}^x F(y) e^{\varphi(y)-\varphi(x)} dy, \tag{2.24}$$

$$= p_0 e^{\varphi_0-\varphi(x)} - J_{p0}^0 \Phi_p(x) - \mathcal{F}_p(x),$$

where $\varphi_0 := \varphi(x_0)$ and we define

$$\Phi_n(x) := \int_{x_0}^x e^{\varphi(x)-\varphi(y)} dy, \qquad\qquad \Phi_p(x) := \int_{x_0}^x e^{\varphi(y)-\varphi(x)} dy,$$

$$\mathcal{F}_n(x) := \int_{x_0}^x F(y) e^{\varphi(x)-\varphi(y)} dy, \qquad\qquad \mathcal{F}_p(x) := \int_{x_0}^x F(y) e^{\varphi(y)-\varphi(x)} dy.$$

Next, we can determine the parameters J_{n0}^0, J_{p0}^0 by evaluating the boundary conditions $n(x_L) = n_L$ and $p(x_L) = p_L$. Upon rearrangement, we have with $\varphi_L := \varphi(x_L)$:

$$J_{n0}^0 = \frac{n_L - n_0 e^{\varphi_L-\varphi_0} + \mathcal{F}_n(x_L)}{\Phi_n(x_L)}, \qquad J_{p0}^0 = \frac{-p_L + p_0 e^{\varphi_0-\varphi_L} - \mathcal{F}_p(x_L)}{\Phi_p(x_L)}, \tag{2.25}$$

and obtain explicit formulas for n^0 and p^0:

$$n^0(x) = n_0 \left(e^{\varphi(x)-\varphi_0} - e^{\varphi_L-\varphi_0} \frac{\Phi_n(x)}{\Phi_n(x_L)} \right) + n_L \frac{\Phi_n(x)}{\Phi_n(x_L)} + \mathcal{F}_n(x_L) \frac{\Phi_n(x)}{\Phi_n(x_L)} - \mathcal{F}_n(x) \tag{2.26}$$

$$p^0(x) = p_0 \left(e^{\varphi_0-\varphi(x)} - e^{\varphi_0-\varphi_L} \frac{\Phi_p(x)}{\Phi_p(x_L)} \right) + p_L \frac{\Phi_p(x)}{\Phi_p(x_L)} + \mathcal{F}_p(x_L) \frac{\Phi_p(x)}{\Phi_p(x_L)} - \mathcal{F}_p(x) \tag{2.27}$$

56

which satisfy both boundary conditions since $\Phi_{n/p}(x_0) = \mathcal{F}_{n/p}(x_0) = 0$.

These equations, although explicit, remain difficult to understand in terms of leading order behavior due to the the competing exponential behavior of their terms. However, using the fact that U is constant outside of the interface, one can explicitly calculate the integrals Φ_n and Φ_p to the zeroth order (see Section 2.7.2 for explicit results).

In the following, we use these formulas to identify the leading contributions to J_{n0}^0 and J_{p0}^0 and thus n^0 and p^0. More precisely, we introduce the parameters

$$\frac{1}{\delta_1} := e^{\varphi(x_l) - \varphi_0}, \qquad \frac{1}{\eta} := e^{\varphi(x_r) - \varphi(x_l)}, \qquad \frac{1}{\delta_2} := e^{\varphi_L - \varphi(x_r)} \qquad (2.28)$$

and observe that $\delta_2 = \delta_1^{-2}$ for the considered device geometry since we have $2(x_l - x_0) = x_L - x_r$ (see Fig 2.1). For the rest of this section we identify $\delta_1 = \delta$. With $\varphi_x = -E^0 + U_x$ and $E^0 < 0$ as in the working device, these parameters are small: $\delta, \eta \approx 10^{-3}$ for $E^0 = E_{\mathrm{SC}}$ with the numerical values given in Table 2.1.

However, using these parameters directly in (2.26) and (2.27) does not directly yield insights into the behavior of n^0 and p^0. Because $\Phi_{n/p}(x)$ can change over many orders of magnitude, the behavior of n^0 and p^0 depends highly on competing exponential terms, none of which can be easily eliminated. One possible way to proceed splits the domain $[x_0, x_L] = [x_0, x_l] \cup (x_l, x_r) \cup [x_r, x_L]$ into three areas: left of the interface, the interface, and right of the interface. For each of these areas, we would obtain formulas for n^0 and p^0 of the form of (2.23) and (2.24) in terms of the values $n^0(x_l), n^0(x_r)$ and $p^0(x_l), p^0(x_r)$ and the boundary terms n_0, n_L and p_0, p_L. Then, imposing continuity of n^0 and p^0 and continuity of the fluxes J_{n0}^0 and J_{p0}^0 at $x = x_l$ and $x = x_r$ would allow us to determine the values $n^0(x_l), n^0(x_r)$ and $p^0(x_l), p^0(x_r)$. This is quite similar to the method which we will discuss below in Section 2.7.

However, a simpler way to proceed considers the zeroth order currents, which are sufficient to understand the short circuit electric current. Using the explicit formulas for (the zero order) Φ_n and Φ_p in terms of δ and η we have from (2.25):

$$J_{n0}^0 = \frac{n_L - \frac{n_0}{\eta \delta^3} + \mathcal{F}_n(x_L)}{\frac{1}{E^0} \frac{1}{\eta \delta^2} \left(1 - \frac{1}{\delta}\right) + \frac{1}{E^0 - U_x(\theta)} \frac{1}{\delta^2} \left(1 - \frac{1}{\eta}\right) + \frac{1}{E^0} \left(1 - \frac{1}{\delta^2}\right)}, \qquad (2.29)$$

$$J_{p0}^0 = \frac{-p_L + p_0 \eta \delta^3 - \mathcal{F}_p(x_L)}{-\frac{1}{E^0} \eta \delta^2 \left(1 - \delta\right) - \frac{1}{E^0 - U_x(\theta)} \delta^2 \left(1 - \eta\right) - \frac{1}{E^0} \left(1 - \delta^2\right)}, \qquad (2.30)$$

where $U_x(\theta)$ denotes a mean value of U_x within the interface.

Because $J_n + J_p = const$, the sum $J_{n0}^0 + J_{p0}^0$ can be used to calculate the total current output of the device (as predicted by the asymptotics). We shall plot and discuss the asymptotic IV curve in Section 2.6. Note that for the given plots therein we take U to be piecewise linear so that $U_x(\theta)$ is explicitly defined.

Short Circuit Current

As mentioned earlier, the sign of E^0 determines if the parameters η, δ are large or small. For the short circuit current, we have $E^0 \approx -13 < 0$ and $\eta, \delta \ll 1$. In order to determine the lowest order terms of these equations, we must also calculate a more precise form of $\mathcal{F}_{n/p}(x_L)$ and thus its order. First we remark that $F(x) = \int_{x_0}^x k_d(y) X^0(y) dy \approx 10^2$ (since $k_{d,in} X^0 \approx \approx 10^4$ on the interface with width 10^{-2} and $k_{d,out} X^0 \approx 10$ outside the interface with device length $x_L - x_0 \approx 1$). Thus, one can verify that the exponentials of $\varphi(x)$ will determine the leading contributions to the integrals $\mathcal{F}_{n/p}(x_L)$. More precisely, since $\varphi(x)$ is strictly increasing in the $E^0 < 0$ regime (recall that $\varphi_x = -E^0 + U_x$) the leading order contributions derive from

$$\mathcal{F}_n(x_L) \approx \int_{x_0}^{x_l} F(y) e^{\varphi_L - \varphi(y)} dy, \qquad \mathcal{F}_p(x_L) \approx \int_{x_r}^{x_L} F(y) e^{\varphi(y) - \varphi_L} dy.$$

Applying the definitions of Equation (2.28), integration by parts and the mean value theorem yields

$$\int_{x_0}^{x_l} F(y) e^{\varphi_L - \varphi(y)} dy = \frac{1}{\eta \delta^2} \int_{x_0}^{x_l} F(y) e^{-E^0(x_l - y)} dy$$

$$= \frac{1}{\eta \delta^2} \left[F(y) \frac{e^{-E^0(x_l - y)}}{E^0} \right]_{x_0}^{x_l} - \frac{1}{\eta \delta^2} \int_{x_0}^{x_l} \frac{F'(y)}{E^0} e^{-E^0(x_l - y)} dy$$

$$= \frac{1}{\eta \delta^2} \frac{F(x_l)}{E^0} - \frac{1}{\eta \delta^2} \frac{F'(\theta_n)}{(E^0)^2} \int_{x_0}^{x_l} e^{-E^0(x_l - y)} dy,$$

for a mean value $\theta_n \in (0, x_l)$. Together with similar calculations for $\mathcal{F}_p(x_L)$ and $\theta_p \in (x_r, x_L)$, we obtain

$$\mathcal{F}_n(x_L) \approx \frac{1}{\eta \delta^2} \frac{F(x_l)}{E^0} - \frac{k_{d,out} X^0(\theta_n)}{(E^0)^2} \frac{1}{\eta \delta^2} \left(1 - \frac{1}{\delta} \right) + O(\eta^{-1} \delta^{-2})$$

$$\mathcal{F}_p(x_L) \approx -\frac{k_{d,out} X^0(\theta_p)}{(E^0)^2} \left(1 - \delta^2 \right) - \frac{F(x_L)}{E^0} + \frac{F(x_r)}{E^0} \delta^2 + O(\delta^2)$$

Inserting only the dominating terms $O(\eta^{-1}\delta^{-3})$ into (2.29) and $O(1)$ into (2.30) yields:

$$J_{n0}^0 \approx E^0 n_0 - \frac{k_{d,out}X^0(\theta_n)}{E^0} \qquad J_{p0}^0 \approx E^0 p_L - \frac{k_{d,out}X^0(\theta_p)}{E^0} - F(x_L)$$

$$J_{tot}^0 \approx E^0(n_0 + p_L) - \frac{k_{d,out}}{E^0}\left(X^0(\theta_n) + X^0(\theta_p)\right) - \int_{x_0}^{x_L} k_d X^0(y)dy. \qquad (2.31)$$

The third term is the easiest to understand, because it is simply the total number of charges generated by the dissociation of excitons in the device. In the working case the majority of the excitons are expected to be produced by the light and thus the final term represents the photogenerated current in the device.

The first term, $E^0(n_0 + p_L)$, gives a current which is proportional to the field, indicating that it represents a resistor. For a usual working device this term will be small ($n_0 + p_L \approx 10^{-6}$ in Section 2.2) since it refers to energetically unfavorable injection of minority carriers into the device. Because the term does not depend on the length of the device, we interpret it as a shunt resistance (parallel to the device). We observe from Equations (2.29) and (2.30) that we have another current term with the form of a shunt resistance: $\eta\delta^3 E^0(n_L + p_0)$. This factor is negligible in the short circuit case, but becomes large as we move to positive fields, where $\delta^3 > 1/\eta$. We investigate the total shunt resistance further in Section 2.6.

Note that in this form Ohm's law becomes $J = E\sigma$ where σ, the conductivity, is the reciprocal of the resistance. Thus a small value of $(n_0 + p_L)$ would correspond to high shunt resistance, which is characteristic of a good device. In particular, a lower shunt resistance will reduce the open circuit voltage as charges can move across the device freely without piling up at the interface. This effect is (barely) visible in Figure 2.3, where the zero boundary condition curve reaches $J = 0$ slightly later than the curve with non-zero boundary conditions. This corresponds to a slightly increased open-circuit voltage – one consequence of increasing the shunt resistance.

The second term of Equation 2.31 represents the interaction of the bulk excitons on the system. For usual device parameters, $k_{d,out} \ll k_{d,in}$ (for the parameters given in Section 2.2, we have $k_{d,out} = 10^{-3}k_{d,in}$). Since $X^0 \approx \frac{G}{k_{d,in}}$ on the interface and $X^0 \approx \frac{G}{k_{r,out}}$ outside the interface, we can neglect the contributions from outside the interface (recall that $\theta_n \in (0, x_l)$ and $\theta_p \in (x_r, x_L)$) compared to the contribution from within the interface (as represented by the integral over the interface region).

Thus the most important contribution to the current for negative fields is:

$$J^0_{\text{approx}} = - \int_{x_l}^{x_r} k_{d,in} X^0(y) dy. \tag{2.32}$$

This approximation works very well in the short circuit case (as examined in Section 2.6, in the discussion preceding Figure 2.7), where it replicates J^0_{tot} within 2% and is actually closer to the numerically calculated J.

Note that J^0_{approx} comes entirely from the J^0_{p0} term. Because we have evaluated J^0_n and J^0_p at the point x_0, we expect that the hole current will be dominant. Using Equation (2.22) we can calculate J^0_{nL} and J^0_{pL}, and note that the only difference from the currents at x_0 is that the $-F(x_L)$ term appears in the electron current. Thus, as expected, in the region where the electrons are favored, the primary contribution to the current comes from the electrons and the total current is conserved. This dominance of the current of the majority carriers (which is expected from the physics) may cease to hold outside of the reverse-bias regime – see Section 2.8.

First Order Terms

Since we have explicit forms for n^0, p^0, and X^0, we can integrate twice to calculate V^1 (with an additional linear term to account for the two boundary values of V). Normally we don't plot V since it is dominated by its boundary terms near short circuit. However, $E^1 = -V^1_x$ can also be calculated explicitly (by a single integration) and we include the first-order field in the plots in Section 2.6. The explicit form is not very illuminating and thus not written here. Note that we need to add a constant value to the field to ensure that its integral gives the correct potential difference in the device (corresponding to the slope of the aforementioned linear term).

In theory we could use the second order form for the potential to calculate the next order n^1 and p^1 solutions. In fact, this is more or less the essence of the iteration scheme outlined in Section 2.3. However, without a constant field, it is no longer possible to explicitly integrate the continuity equations (especially given the sub-exponential forms for the electron and hole mobilities, which become constant in the zeroth order) and we cannot guarantee that the $c_r np$ term remains small. Instead, we will next present another method for asymptotic calculations for the carrier densities. Later, in Section 2.7, we show that integrating the n and p equations implicitly can yield reasonable second order approximations for the current J^0 without requiring explicit calculation of n and

60

p.

2.5.2 Unipolar Approximation

The discussion of the previous section, in particular Equation (2.32), can be summarized by the statement that the electron and hole fluxes J_n^0 and J_p^0 are approximately constant outside the interface but feature a strong variation over the interface with a magnitude of $F(x_r) - F(x_l) = \int_{x_l}^{x_r} k_{d,in} X^0(y) dy$.

As a consequence, using the Equations (2.26) and (2.27), it follows that the zero-order electron and hole densities n^0 and p^0 vary also strongly over the interface. In fact, we obtain for the hole density p^0:

$$p^0(x_l) = \frac{F(x_L) - F(x_l)}{-E^0} + O(\delta) + O(p_L)$$

$$p^0(x_r) = \frac{F(x_L) - F(x_r)}{-E^0 + U_x(\theta_{in})} + \frac{k_{d,in} X^0(\theta_{in})}{(E^0 - U_x(\theta_{in}))^2} + \frac{k_{d,out} X^0(\theta_{out})}{E^0(-E^0 + U_x(\theta_{in}))}$$
$$+ O(\eta, \delta^2) + O(p_L),$$

where $\theta_{in} \in (x_l, x_r)$ is a position within the interface used by the mean value theorem and $\theta_{out} \in (x_r, x_L)$ is a corresponding position outside of the interface. The values of η, δ and p_L are (as mentioned above) small in the working regime. Then, since $U_x(\theta) \gg 1$ for $\theta \in (x_l, x_r)$ and because the dominant part of the integral $F(x)$ comes from the interface region, we see that

$$\frac{p^0(x_l)}{p^0(x_r)} \approx 10^3, \qquad \text{in the working case } E^0 < 0,$$

as a consequence of the work function gap over the interface. A similar result can be obtained for n^0 giving $n^0(x_r) \gg n^0(x_l)$.

Note that in some devices one might expect this ratio to be even larger. However, from Boltzmann's statistics we would expect a ratio of $e^{\Delta U} = e^6$, in reasonable agreement to the results above. In addition, if we narrow the interface while keeping the work function difference ΔU and the generation term given in (2.32) constant, we tend to increase the value of $U_x(\theta_{in})$. Thus in some sense the smallness of $p^0(x_r)$ increases naturally by reducing the size of the interface. See Section 2.8.2 for more details regarding the thin-interface limit.

The previous estimates quantify the observation that a bilayer device is operated such that in one material, $n \ll 1$ (here to the right of the interface) and in the other material

$p \ll 1$ (here to the right of the interface) because of the nature of the HOMO/LUMO levels of the different polymers. We can thus ask for an asymptotic simplification of System (2.14) in the bulk-material away from the interface.

In order to proceed with such an approximation, we observe furthermore that the ratio $\frac{p^0(x_l)}{k_{d,out}X^0} \approx 10 - 20$, i.e. that the boundary value of the majority charger carrier $p^0(x_l)$ given at the left of the interface dominates over the generation of holes from excitons in the bulk material left of the interface. In fact the above factor would even be larger in more realistic exciton models compared to (2.14), which misses a term modeling a electrochemical trapping of excitons (due to stability of polaron pairs on the interface [3]) yielding further decreased the exciton concentrations (see Section 2.9 for a discussion of including this potential). We shall thus entirely neglect here the photogeneration term of charge carriers outside the interface ($k_{d,out}X^0$).

Neglecting the minority charge carriers and excitons outside the interface, we consider the following simplified models for the majority charge carriers:

$$-\lambda_D^2 V_{xx} = \lambda_D^2 E_x = p, \qquad p_x - pE = \frac{-J_{p0}}{\mu_p}, \qquad (2.33)$$

in the p-material at the left of the interface (with the Debye length $\lambda_D^2 = \lambda^2 \epsilon_r$) and

$$-\lambda_D^2 V_{xx} = \lambda_D^2 E_x = -n, \qquad n_x + nE = \frac{J_{n0}}{\mu_n},$$

for the n-material at the right of the interface. Here J_{p0} and J_{n0} are the constants arising from integration.

For the rest of this section we shall focus only on the hole density p. The analysis of the electron equation is analogous. We remark that the Equations (2.33) constitute a generalization of the system considered in Appendix A of [15], where the zero-current case $J_{p0} = 0$ was analyzed. For the present case with non-zero current we assume first that μ_p is constant and exploit then a first-integral of the equation to obtain

$$E_x - \frac{1}{2}E^2 = \frac{-J_{p0}}{\lambda_D^2 \mu_p}x + C_p$$

where C_p is another integration constant. Using the substitution $x = -2u$ and $E = y_u/y$ (see e.g. [15]), we obtain

$$y_{uu} = (\kappa u - 2C_p)y$$

where $\kappa = \frac{-4J_{p0}}{\lambda_D^2 \mu_p}$. With the further substitution $z = \sqrt[3]{\kappa}\left(u - \frac{2C_p}{\kappa}\right)$, this becomes the familiar Airy Equation, $y_{zz} = yz$ and we can write down the formula for y explicitly in

62

terms of Airy Functions. Since $E = y_u/y$, it is more convenient to write the form of $V = -\log(y)$ (derived independently in [73]):

$$V = -\log\left[\alpha Ai\left[\sqrt[3]{\kappa}\left(-\frac{x}{2} - \frac{2C_p}{\kappa}\right)\right] + \beta Bi\left[\sqrt[3]{\kappa}\left(-\frac{x}{2} - \frac{2C_p}{\kappa}\right)\right]\right] \tag{2.34}$$

where we can express α and β in terms of the integration constants and the boundary values. Note that for the n material we take the substitution $x = 2u$ and $-J_p \to J_n$ positive and everything else proceeds identically.

This asymptotic result gives an explicit form for the charge carriers in the bulk in 1D. However, for bilayer systems away from the interface, the system generally approximates a 1D system since the bulk material acts primarily as a charge-transport layer (see the discussion surrounding 2.2). Since we explicitly account for the self-consistent potential, this formulation works much better than our previous asymptotics when we have large carrier concentrations (especially for less negative potential differences). In theory we could combine these equations in the bulk with a numerical calculation of the interface, simplifying the computations. However, in practice, we can change the mesh (or grid) to account for the reduced complexity of the problem in the bulk (see Figure 2.5) and the gain would not necessarily be significant.

The main disadvantage of this formulation is that it relies upon the boundary data of the system. Since it doesn't model the interface, it is thus incapable of predicting the current in the device based on the device parameters, and requires J as a parameter. In fact, we need all of the boundary values p_0, E_0, J_{p0} in order to calculate the field and the hole concentration in the p-region of the device, but generally only p_0 is given. The other values, E_0 and J_{p0}, both require calculating the solution of the whole problem or at least calculating a model of the interface and using asymptotic approximations of the boundary values E_0 and J_{p0}. Alternatively, we can take the values $p(x_l), E(x_l), J_{pl}$ and again calculate the values in the bulk. Either method gives extremely accurate results for the cases in which $J \ll 0$, which we show in Figures 2.7 - 2.9. A treatment of the results coupled with interface conditions on the boundary is given in the recent work [73].

2.6 Discussion of Numerical Examples

In the following we present selected numerical examples of the steady state device behavior of the organic photovoltaic bilayer device plotted in Figure 2.1. In particular, we

shall investigate the steady states calculated from different applied potentials.

Using the scheme outlined in Section 2.3 we use the damped-Gummel iteration given in (2.12) and (2.13) with $\alpha = .01$ for the first three steps (as a sort of preconditioning) and $\alpha = .6$ until convergence (see Section 2.6.4). All of the plots show a mesh consisting of 720 elements, greatly refined near the interface, and use polynomial interpolants of order 7 for the bulk elements – see Figure 2.5. Refining the mesh and increasing the interpolation order both give better accuracy, in particular for the steep concentration changes over the interface (see, i.e. Figures 2.7, 2.9 and 2.10). Note that the geometry of the test cases is taken to be homogeneous in the y-direction and we plot only a sample x-cross section which can be compared with the one-dimensional asymptotics.
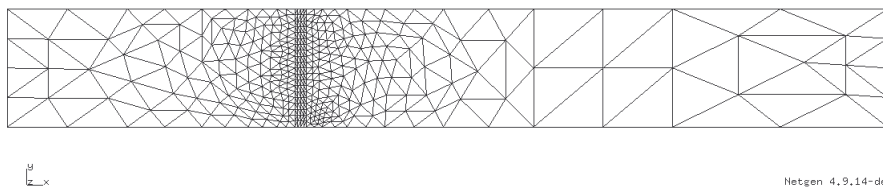
Figure 2.5: Mesh used for all of the plots in Section 2.6

2.6.1   A Note On Exciton Diffusion

It is clear from the previous sections that the behavior of the polaron pairs on the interface is key to the overall current generated in the device. In fact, the approximate current given by (2.32) is simply a weighted integral of X over the interface. Many models for the operation of photovoltaic devices disregard the field dependence of k_d entirely, and doing so with our model still yields qualitatively realistic results (see e.g. Figure 2.4). Since many of the coefficients of the exciton equation are not physically well-understood, this begs the question of whether or not they significantly impact the characteristics of the model. We show in Figure 2.6 that changing the exciton and polaron pair mobility μ_X does in fact have an impact on the qualitative nature of the current-voltage characteristics. This effect is not trivial – we show in Appendix 2.A.1 that L^∞ bounds on X are independent of μ_X. However, given that the equilibrium values for polaron pairs (X inside the interface) are higher than the equilibrium values

64

for excitons, reducing μ_X supplies significantly more polaron pairs near the edges of the interface.
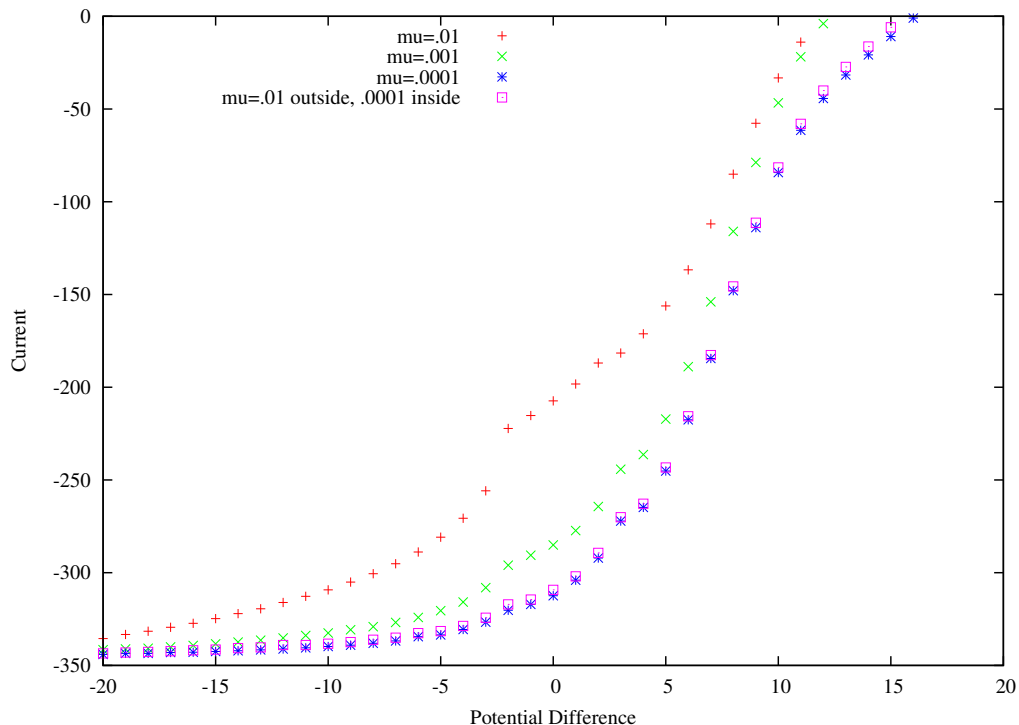
Figure 2.6: Numerical current-voltage characteristics for our given $\mu_X = .01$, $.001$, and $.0001$. Note that the Short Circuit current and open circuit voltage are similar, but that the fill factor is significantly better for the smaller mobility. Also plotted is an IV curve for $\mu_X = .0001$ inside the interface and $\mu_X = .01$ outside the interface. In a physical device, increasing the mobility of the excitons in the bulk should result in many more polaron pairs being produced and thus much more current.

This result, while not necessarily surprising mathematically, does indicate non-physical behavior. Excitons near the interface should be trapped on the interface as polaron pairs. This was modeled by choosing a small μ_X, but clearly the value given in Table 2.1 was not small enough, as decreasing it further results in the changes shown in Figure 2.6. Furthermore, there is also experimental evidence that increasing the bulk exciton mobility in a bilayer should increase the current [58]. However, our numerical results show that reducing the exciton mobility in the bulk by two orders of magnitude does not significantly change the current-voltage characteristic. This suggests that increasing the

65

diffusion in the bulk is not increasing the number of excitons or polaron pairs dissociating. In particular, we note in the plots of the exciton densities in the next section that we do not observe a reduction in the exciton concentration near the interface.

If the excitons were truly moving preferentially towards the interface, we would expect a non-zero flux ∇X towards the interface, probably over a length-scale comparable to the exciton diffusion-length ($\approx .1$ in scaled units). In hindsight, this is not necessarily surprising. For μ_X very small, the concentration of polaron pairs will be very close to the equilibrium value $\frac{G}{k_d+k_r}$ and the behavior of the excitons will have little or no effect on the polaron pair concentration. We discuss a possible solution to this issue in Section 2.9.

Thus the results below thus focus on the effect the excitons and polaron pairs have on the dynamics of the system (especially the current) as opposed to the particulars of the exciton modeling.

2.6.2 Electron/Hole and Exciton Densities at SC, OPP and OV

In this section, we study the details of the charge density and exciton distributions at particular characteristic points of an IV curve: i) Short Circuit (SC), ii) Optimal Power Point (OPP) and iii) Open Circuit Voltage (OV). The short circuit current is the current at zero applied voltage, or a potential difference of $V_{\text{diff}} = -19.3$. The optimal power point is the potential difference for which the power is maximized, which is $V_{\text{diff}} = -3$ in Figure 2.3. The open circuit voltage is the point where $J = 0$, which occurs for $V_{\text{diff}} = 12$.

For each of these points, we will plot the numerical solutions for n (green $\times$, right y-axis) and p (red $+$, left y-axis) in the same chart. Note that we have rescaled n and p by μ_n and μ_p in the asymptotics section, and must appropriately scale them back for the plots. The main effect of this rescaling is that n and p once again have different scales (as in Figure 2.2 in Section 2.4). On the same figure, we also plot the zero-order asymptotics derived in Section 2.5.1 (in blue) and the unipolar asymptotics derived in Section 2.5.2 (in pink). The unipolar results reproduce the numerical solution in the bulk very well when complemented with the numerical data from the simulation of the interface (which is not part of the unipolar approximation). Specifically, we use the boundary data values $p(x_0)$, $E(x_0)$, and J for the holes and $n(x_L)$, $E(x_L)$ and J for the electrons. We observe that in all three cases – SC, OPP and OC – the minority carriers

66

(n to the left of the interface and p to the right of the interface) are strongly dominated by the majority carriers (as discussed above in Section 2.5.2). We also note, however, that small concentrations do not necessarily imply small currents – see Section 2.8.

Next, we plot the numerical solutions of the exciton concentration (in the lower left plot) and the electric field (in the lower right plot). The exciton concentration is compared with the explicit results in Section 2.5.1 and the electric field is compared with the first-order asymptotic approximation (near the end of Section 2.5.1. The unipolar asymptotics do not include exciton behavior to compare with the numerical results. Furthermore, we refrain from plotting the unipolar approximation to the electric field outside the interface (which is discontinuous over the interface since the field is calculated independently on the two sides of the device). However, good agreement similar to the charge carrier densities can be achieved.

We also list the numerically calculated current given at each voltage. The current predicted by the zero-order asymptotics is the sum of J_{nl}^0 and J_{pl}^0 given in Equation (2.29) and (2.30). However, because this predicted current assumes that $E < 0$, it cannot give accurate results in the open circuit case.

Short Circuit

The case with the largest (negative) potential difference V_{diff} is the short circuit case (which has zero applied voltage). This is due to the built-in potential arising from the work functions of the anode and cathode metals. For the device parameters laid out in Section 2.2, this occurs for $V_{\text{diff}} \approx -19.3$.

Since short circuit considers the largest negative potential difference it is thus the case for which the zero-order asymptotics derived in Section 2.5.1 are best suited. This is confirmed by Figure 2.7 which compares the numerical solutions of n, p, X and E with the corresponding asymptotics. Note that we see a dramatic change in n, p, X at the interface, and that either electrons or holes are largely dominant on one side of the device. Moreover, note that the slope of the electric field is significantly larger on the left-side of the device due to the larger concentration of holes on the left-side compared to the smaller concentration of electrons on the right-side (Figure 2.7 plots n and p on different scales).

The numerically calculated current (as listed in Figure 2.7) is $J = -334.2$. The current predicted by Equations (2.29) and (2.30) is $J^0 = -345.15$. Although not in
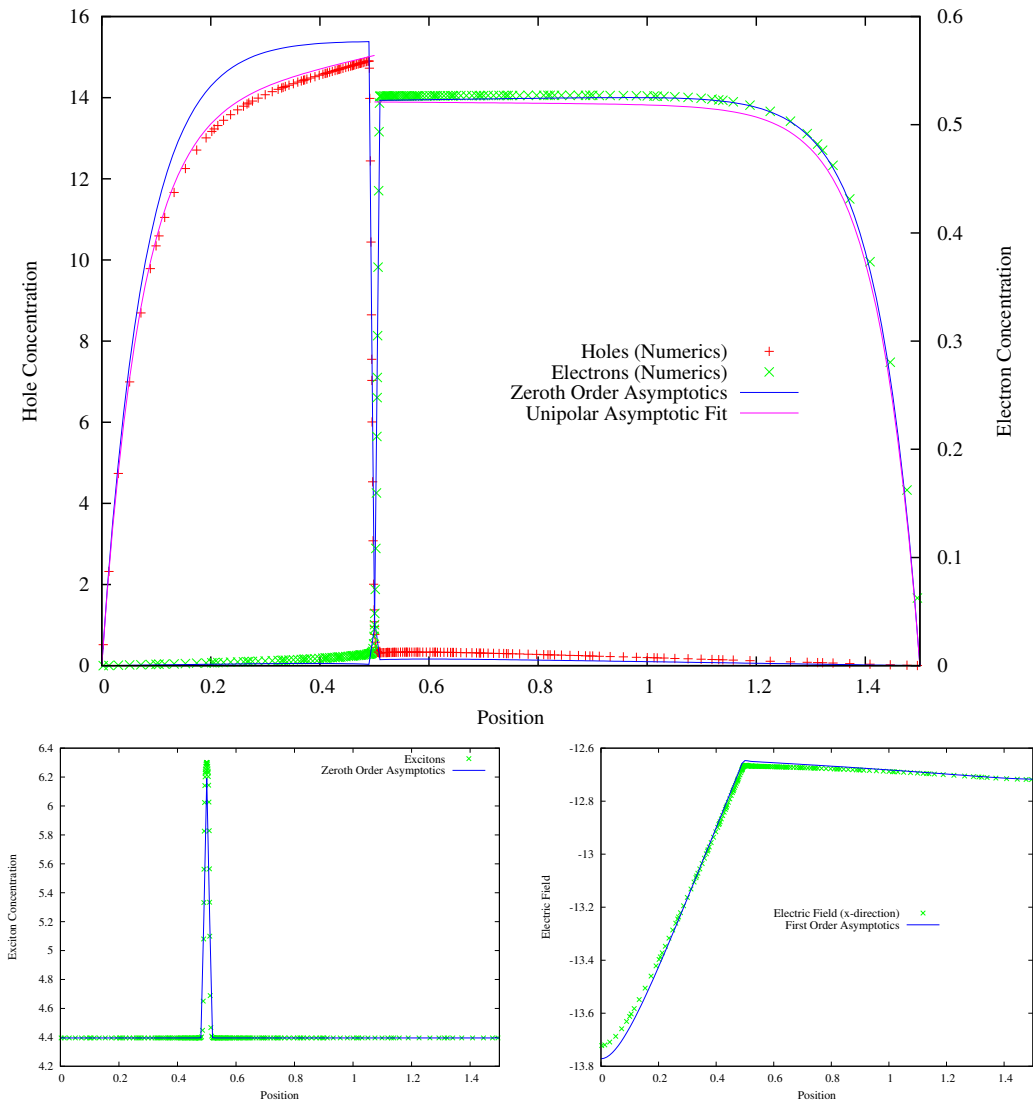
Figure 2.7: Plots of n, p, X and E for the short circuit case $V_{\text{diff}} = -19.3$. Note the different scales for n and p, and that the total variation of E is less than 10% of its average value. $J = -334.2$. The electrons and holes are each dominant in one half of the device, and the asymptotic results provide good agreement to the numerical results in all cases.

exact agreement with the numerically calculated current for the short circuit case, this constitutes a reasonable match for the zeroth order term. Furthermore, the lowest order current (in δ, η) predicted by (2.32) is $J_{\text{approx}} = -339.32$, also in excellent agreement for the short circuit case. However, the current given by the asymptotics does not depend on the potential as dramatically as the simulations suggest, most likely indicating that the $\delta \ll 1$ assumption begins to fail.

Another observed feature of Figure 2.7 are the boundary layers for n to match the boundary data at x_L and for p to match the boundary data at x_0. The values of n and p at their majority side of the interface are predicted by the asymptotics to be $n(x_r) = 0.526$ and $p(x_l) = 15.6$, in good agreement with the numerically calculated values. We emphasize that Equation (2.32) in Section 2.5.1 predicts that the concentrations of n and p at the interface are predominantly a consequence of the leading order flux-changes over the interface $J_{\text{approx}}^0 = -\int_{x_l}^{x_r} k_{d,in} X^0(y) dy$ and are essentially independent of the boundary conditions p_0 and n_L. We can confirm this statement by using the interface values as the boundary conditions (i.e. $n(x_L) = 0.526$ and $p(x_0) = 15.6$). We plot the results in Figure 2.8 without visible difference in the concentrations of the charge carriers at the interface compared to Figure 2.7. In particular note that the current and the maximal values of n and p do not change significantly.

Optimal Power Point

A second key characteristic point for device operation is the optimal power point. This is the point where the power from the device is maximized. As shown in Figure 2.3, this occurs for a voltage difference of $V_{\text{diff}} \approx -3$. This is just below the voltage for which the resulting self-consistent field touches zero from below and the field-dependence of $k_d(E)$ causes a kink in the IV curve. The concentrations n, p and X are plotted along with the E-field in Figure 2.9. In the plot of the E-field we see a slightly more pronounced effect (compared to SC) of the excitons on the electric field (the small bump located at the interface), but again this effect is small compared with the overall changes in the field.

Comparing Figure 2.9 to Figure 2.7 we see that n and p again are primarily located on their specific sides of the device (p to the left, n to the right) with an even larger change in concentrations across the interface. In contrast to Figure 2.7 (which features boundary layers for n and p), the concentrations of n and p in Figure 2.9 appear approximately
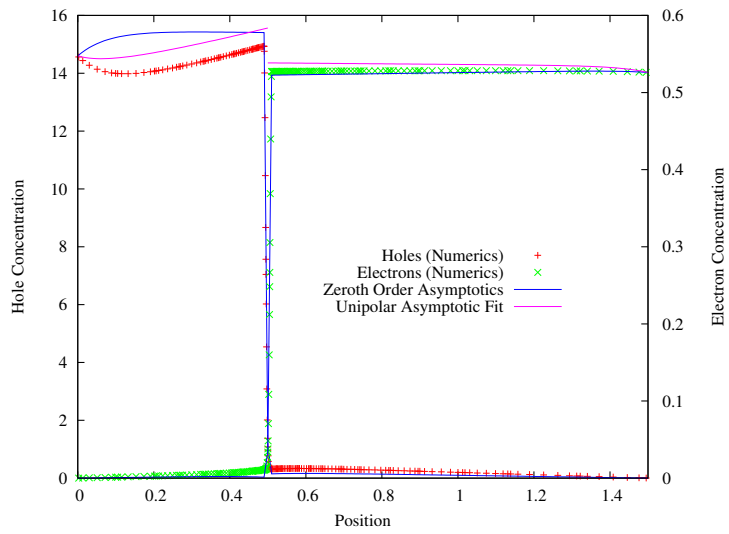
Figure 2.8: Plot of n and p for the short circuit case with artificial boundary conditions, $n_L = 0.526$ and $p_0 = 15.6$. $J = -332.528$, compare these plots to those for the correct boundary conditions in Figure 2.7. The deviation of p from the constant solution results from the self-consistent potential. The effect is less pronounced for the electrons because the concentration is much lower.
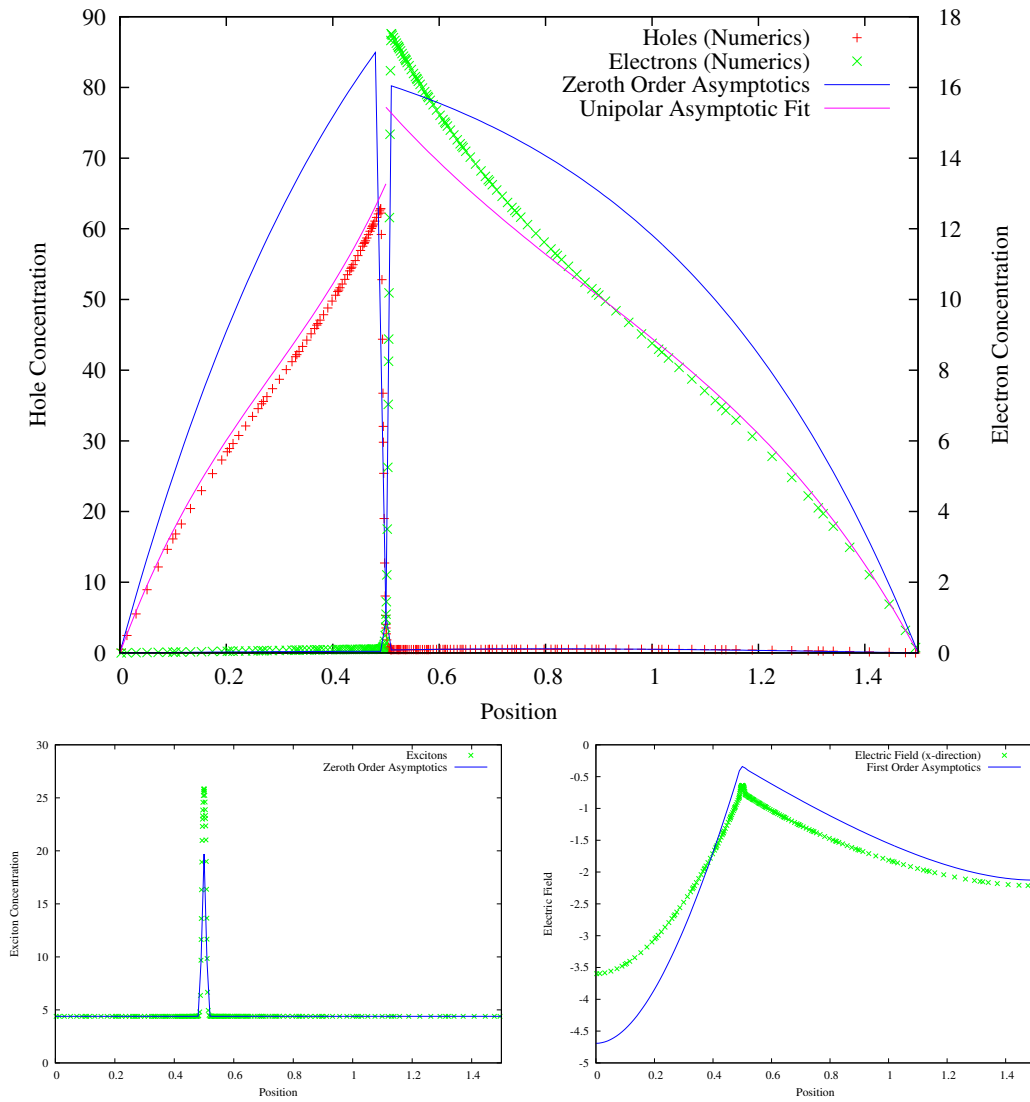
Figure 2.9: Plots of n, p, and E respectively for the optimal power point OPP $V_{\text{diff}} = -3$. The resulting current is $J = -241.585$. The exciton contribution from Poisson's equation is visible as a bump in the field plot. We see that the zeroth order asymptotics begin to diverge from the numerical and unipolar solutions as the potential difference increases.

linear in their majority sides. It is not clear whether this approximate linearity is a characteristic of the optimal power point or coincidental. Our numerical simulations, however, seem to suggest that this is indeed the case and we may conjecture that it is caused by a certain balance between drift and diffusion terms. Moreover, note that the scales for n and p remain different, but are considerably more balanced than in Figure 2.7.

Figure 2.9 also plots the zeroth order asymptotics, which are still reasonable. Moreover, we see excellent fits from the unipolar asymptotic behavior of the quasi-linear behavior of the n and p. The increasing discrepancy of the zeroth order asymptotics suggests the increasing importance of the nonlinearity of the self-consistent E-field. This can be confirmed by using E^0 instead of the numerical values $E(x_0)$ and $E(x_L)$ in the unipolar approximation which is then no longer nearly as accurate.

Open Circuit

The third point of importance for device operation is the open circuit voltage. This is the point at which the current through the device is zero (analogous to an open circuit in electronics). For our simulations this occurs at approximately $V_{\text{diff}} \approx 12$. In physical units this corresponds to a potential difference of 0.31 Volts. Considering the internal voltage of our typical device, this corresponds to an open circuit voltage of 0.81 Volts, larger than expected for a typical OPV (possibly due to the effects outlined in Section 2.8 below).

The plots in Figure 2.10 show a strong alteration of the concentrations n and p compared to the previous Figure 2.7 and 2.9. The hole concentration, p, changes from $p(x_l) \approx 370$ to $p(x_r) \approx 1.4$ over the interface, confirming that our numerical scheme is capable of capturing large changes across the interface (predicted to be $\approx 10^3$ by the zeroth order asymptotics). Note that for this case the maximum values of n and p are of about the same order albeit plotted on different scales (by a factor four).

The behavior of the E-field throughout the device is dominated by the large values of n and p with very little effect of the excitons at the interface. Note that we do not see a very good match of the asymptotic electric field, which is to be expected. In particular, since the zeroth asymptotics ignore the quadratic recombination term ($c_r np$), we are neglecting this loss term for n and p. This is particularly important within the interface, where the concentrations intersect at $n \approx p \approx 18$ (recall that p is strongly decreasing

72
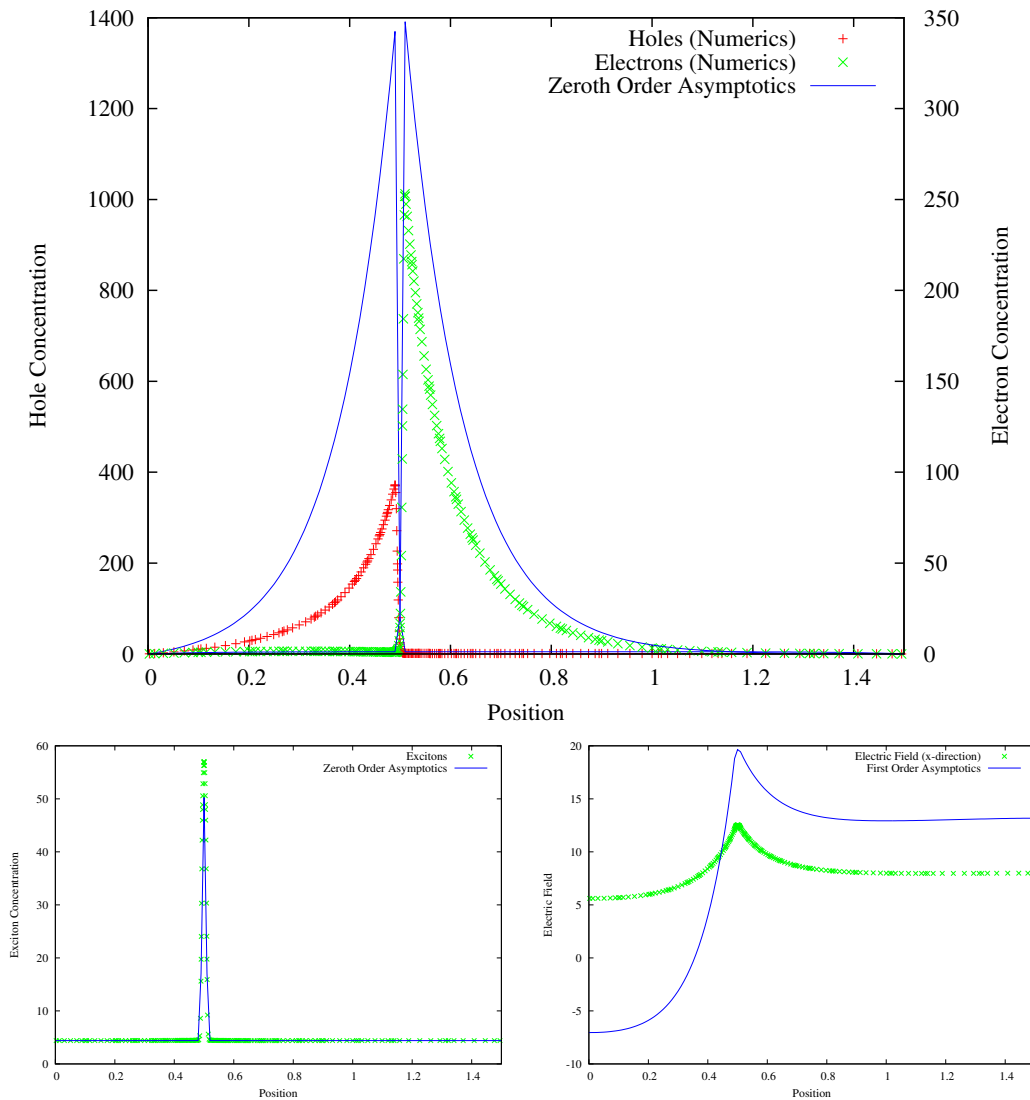
Figure 2.10: Plots of n, p, and E respectively for the (nearly) Open Circuit $V_{\mathrm{diff}} = 12$ case. $J = 0.23$. (The true Open Circuit Voltage is within 0.05 of 12, but further precision is not numerically relevant). Note the different scales for n and p. The qualitative behavior of the zeroth order asymptotics is nearly correct, but they no longer closely match the numerical results. In particular, the estimated value for p is off by at least a factor of 2 nearly everywhere as the majority carrier.

from the left of the interface to the right while n is strongly increasing). This leads n and p to be overestimated. Note, however, that the asymptotics correctly reproduce the qualitative behavior of the carrier concentrations with exponential growth from the boundaries of the device to the interface.

Finally we remark that the presented numerical scheme is capable of continuing further into forward-bias. However, the above model and in particular the parameters are primarily focused on the working-case, and we do not necessarily expect realistic results for the strong forward-bias regime. As the number of polaron pairs on the interface increases (due both to a reduction in k_d and an increase in $c'_r np$), we observe that many of them diffuse away from the interface. This corresponds physically to polaron pairs becoming excitons, a transformation which is not generally physically observed (see Section 2.6.1). We discuss attempts to rectify this problem in Section 2.9.

2.6.3   Shunt Resistance and Asymptotics of the IV Curves

A comparison of the IV curves in Figure 2.3 indicates clearly that the nonlinear boundary conditions proposed in [80] appear to have very little effect on the IV curve. The resulting difference is barely noticeable until we reach open circuit (although it does grow as we pass into the forward bias regime). We can study this discrepancy more clearly by simulating the OPV device in the dark with $G = 0$. The resulting current is plotted in Figure 2.11 and represents the shunt current in the device. Recalling the asymptotic currents Equation (2.29) and (2.30), we can see the increase in the shunt current for $E \gg 0$ (and thus $\delta \gg 1$) proportional to the terms containing n_L and p_0. In the large δ limit we recover a term of the form: $(n_L + p_0)E$. This has the usual form of a resistor following Ohm's law with (scaled) resistivity $1/(n_L + p_0)$.

In Figure 2.11 we show heuristically the equivalence of the shunt current (green $\times$) to the difference between the nonlinear Dirichlet and homogeneous Dirichlet boundary conditions (red $+$). The difference shows exactly the effect of the boundary conditions given by [80] compared to zero boundary conditions. The blue line represents the zeroth order asymptotic fit for the current with $G = 0$. The apparent noise for negative fields is a result of rounding errors in taking the difference of the currents from two different simulations. Choosing a smaller tolerance parameter for the convergence and calculating the current to higher precision (to eliminate errors of magnitude .1 for $J \approx -350$) will reduce the apparent noise. Averaging over slightly different potential differences (i.e.
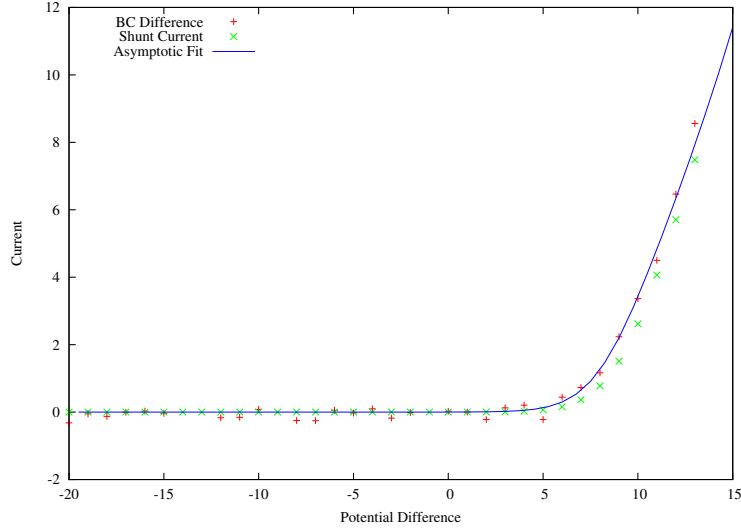
Figure 2.11: An *IV* curve showing the similarity of the shunt current produced by injected carriers with no light generation and the result of subtracting the *IV* curve for zero boundary conditions from the usual *IV* curve. The zeroth-order asymptotic fit is also included. See Figure 2.3 for individual plots of these curves.

-19.99, -20, and -20.01) also reduces the apparent noise.

For a device in the working case the boundary conditions are relatively unimportant, but as we move into the forward bias regime, these boundary conditions become more important. The given shunt current is reminiscent of the exponential growth of the dark-current for a photovoltaic device in forward bias. Although both currents result from taking $G = 0$, our model is not designed to reproduce this dark current, which is normally generated by an equilibrium concentration term in the recombination rate – for instance $c_r(np - n_\infty p_\infty)$. With light generation, this term is negligible compared to the exciton contribution and thus is not included in our modeling assumptions. However, carefully choosing boundary conditions would allow our model to replicate the physically observed diode behavior in the forward bias regime.

Next, Figure 2.12 plots the asymptotic *IV* curve predicted by the sum $J_{n0} + J_{p0}$ of Equations (2.29) and (2.30). Although the asymptotic approximation is questionable for small potential differences (as stated previously) we do observe a qualitatively realistic *IV* curve in Figure 2.12. In particular we observe a similar open circuit voltage and short circuit current as well as a significantly better fill-factor of 87%. The high

fill factor for the asymptotic curves is expected, since the asymptotics neglect the re-combination current, which should be the main cause of power loss for a bilayer device. Indeed, by ignoring the recombination term, the asymptotics describe what is called the photogenerated current, the idealized current generated by the device. We show in Section 2.8 that the common V_{OC} for the numerical and asymptotic results is symptomatic of unphysical behavior near open circuit.
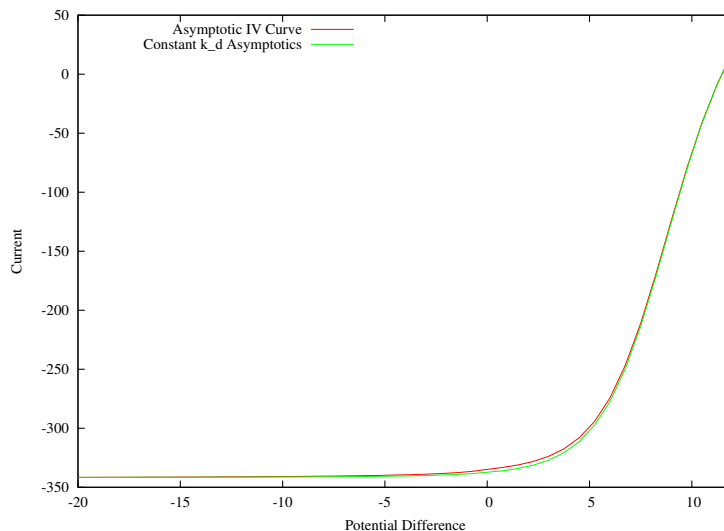
Figure 2.12: The IV Curve predicted by the asymptotics in Section 2.5.1. Note that it has a similar open circuit voltage and short circuit current, but has a better fill factor than the numerical simulations.

A comparable increase in fill-factor is observed in Figure 2.4 (reprinted below as Figure 2.13), where we plot a numerical IV curve assuming a constant dissociation rate $k_{d,in} \gg k_{r,in}$. Moreover, Figure 2.13 plots the corresponding asymptotic IV curve with good agreement. This indicates that a primary loss of efficiency is due to the exciton recombination term, which becomes non-negligible near the interface when approaching the open circuit voltage. This quantifies the physical intuition that device efficiency is increased by forcing the polaron pairs to dissociate more efficiently into free carriers. For the numerical simulation the new fill factor is a much improved 60%, whereas the asymptotics are virtually unchanged from the standard model shown in Figure 2.12.
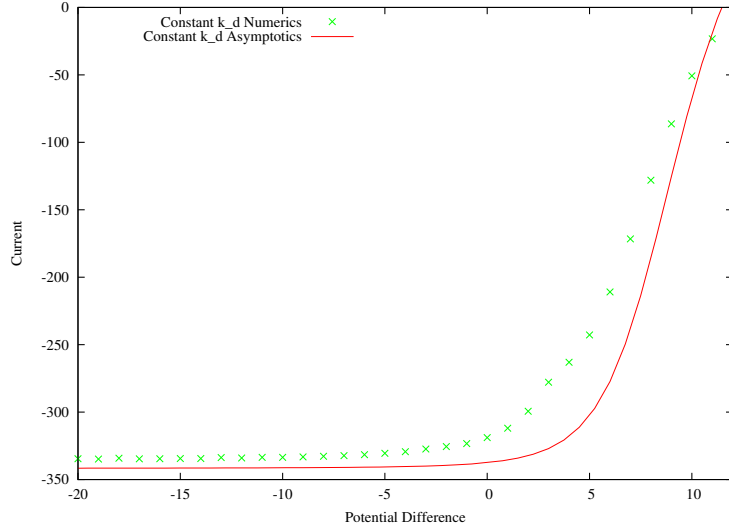
Figure 2.13: A comparison of the IV curves obtained using the numerical and asymptotic schemes with $k_{d,in} = k_{d,in}(E_{\mathrm{SC}}) = 2763$ to be constant.

2.6.4 Damping and Iteration Convergence

Here we present the evidence that our damped iteration scheme indeed converges. First, we show log-plots of the L^2-norm difference between the solutions at consecutive iteration steps. After showing the convergence for the simulations given in Section 2.6, we proceed to discuss the role of the damping parameter in the convergence.

L^2 Errors for the Iteration Scheme

Here we give plots of the L^2 norm of the difference between the subsequent solutions at each step of the numerical iteration scheme. In all of the plots the first three steps have nearly constant error because of the small value of the damping parameter $\alpha = .01$. After the initial stage, we expect exponential convergence.

If our undamped solution u' is always the true solution u^*, then from Equation (2.13), we can calculate the error $\mathcal{E}$ by

$$\mathcal{E}_1 = u_1 - u_0 = \alpha \left(u^* - u_0 \right) \qquad \mathcal{E}_{k+1} = u_{k+1} - u_k = \alpha \left(u^* - u_k \right).$$

We then note that

$$\mathcal{E}_{k+1} = u_{k+1} - u_k$$
$$= \alpha \left(u^* - \left(u_{k-1} + \alpha \left(u^* - u_{k-1} \right) \right) \right)$$
$$= (1 - \alpha)\alpha \left(u^* - u_{k-1} \right)$$

and by induction we have

$$\mathcal{E}_{k+1} = (1 - \alpha)^k \mathcal{E}_1. \tag{2.35}$$

For constant α, we would thus expect the error for a "perfect" damped scheme to decrease exponentially with base $(1 - \alpha)$. Since our u' is only the current best-estimate for the equilibrium solution, we do not expect this relationship to hold. In particular, for the undamped case $\alpha = 1$, this result would be completely meaningless.

However, this equation gives a good approximate decay rate for comparison and we proceed to show the result for $\alpha = .6$ as a cyan line in Figures 2.14 - 2.16. For comparison, we also plot the L^2-norm difference between consecutive steps for V (red), n (green), p (blue), and X (pink).

Short Circuit

We show the error plots for short circuit in Figure 2.14. After approximately three steps, we see exponential decay of the L^2 errors, indicating good convergence to a steady state solution. For short circuit, the errors for p are the limiting factor of the scheme. Note that the concentrations of p are roughly thirty times larger than the concentrations of n (see Figure 2.7) and that this difference probably leads to the absolute L^2 error being roughly 30 times larger for p.

Optimal Power Point

We show the error plots for the optimal power point in Figure 2.15. The early stages are similar to those in Figure 2.14, but the intermediate stage lasts for six points instead of three. In addition, the convergence is no longer better than the damping estimate.

Note that the "noise" in the six intermediate points looks quite similar to the convergence shown for open circuit in Figure 2.16.
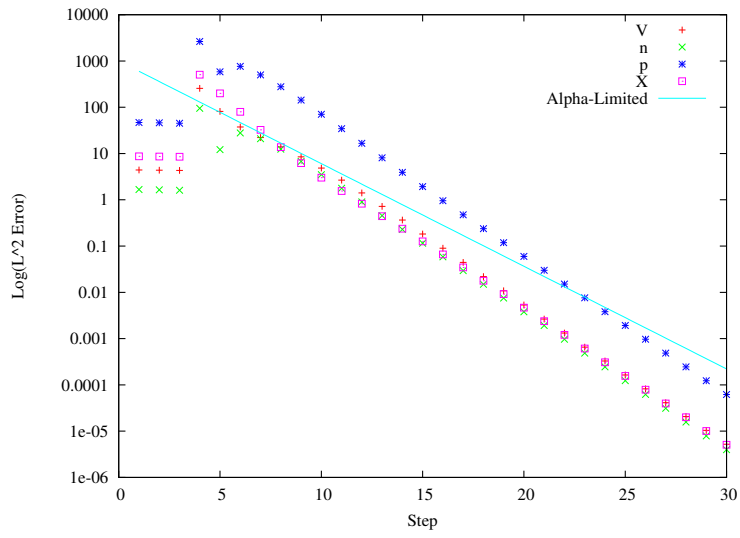
Figure 2.14: Log-plot of the L^2 norm of the difference between subsequent steps for the short circuit simulation shown in Figure 2.7.
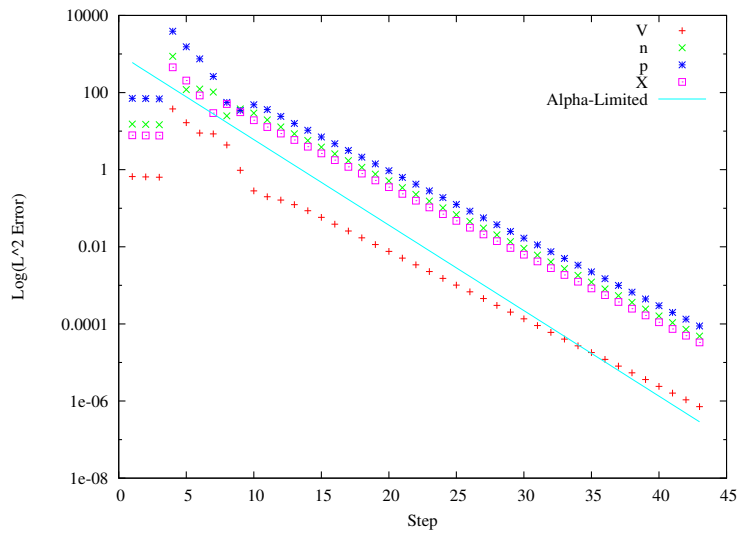
Figure 2.15: Log-plot of the L^2 norm of the difference between subsequent steps for the optimal power point simulation shown in Figure 2.9.

79

Open Circuit

We show the error plots for open circuit in Figure 2.16. The convergence is not as uniform as the short circuit and optimal power point cases (Figures 2.14 and 2.15). However, we still see roughly exponential decay of the L^2 errors, indicating good convergence.
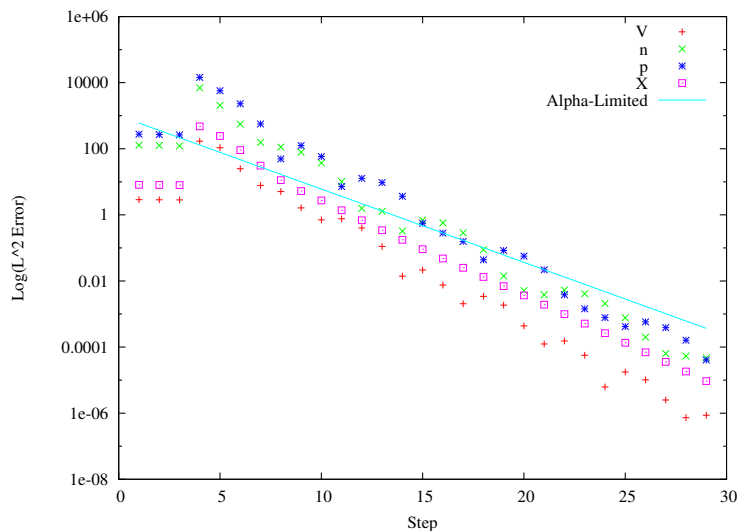
Figure 2.16: Log-plot of the L^2 norm of the difference between subsequent steps for the open circuit simulation shown in Figure 2.10.

Choice of α

The essential reason for the damping is that the iteration scheme will overcompensate for the current error. For example, if n and p are too large, then the $-c_r np$ term will be too large and (all else being equal), n and p will drop below their equilibrium values on the next step. If this effect is too strong, then the simulation can alternate between two states or even diverge. In Figure 2.17 we show two possible scenarios where the iteration scheme does not converge. The first plot shows the iteration switching between two points in solution space (i.e. the L^2 norm of the difference between the steps converges to a constant). The second plot shows the iteration diverging. The errors jump to 10^{80} on the next step before failing completely on the following step.

The damping parameter prevents this overshoot and instead allows the system to converge faster than it might otherwise. However, if the nonlinearities are small (for

80
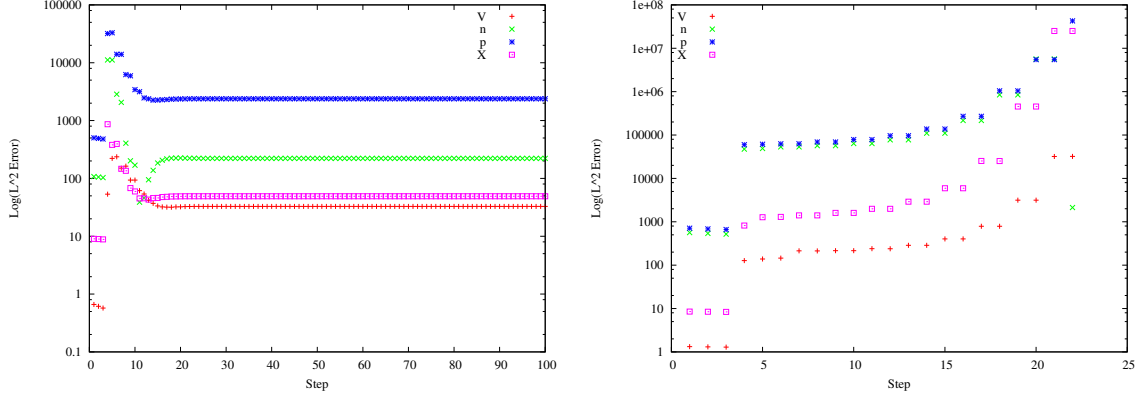
Figure 2.17: Log-plots of the L^2 norm of the difference between subsequent steps for two simulations which fail with no damping ($\alpha = 1$). The first is for the optimal power point ($V_{\text{diff}} = -3$), the second is for $V_{\text{diff}} = 6$.

instance because $c_r n p \ll 0$), then the damping will simply propagate according to Equation (2.35). In particular, as the asymptotics from Section 2.5.1 show, we might expect that damping to be unnecessary at short circuit, but vital as we approach forward bias.

Assuming we have chosen an appropriate mesh and element orders, the computational time for a single iteration is relatively constant. Thus we can evaluate the optimality of α by comparing the number of iterations required for the scheme to reach a given error tolerance. In Figure 2.18 we show the number of iterations required for the simulation to reach the desired error tolerance. Note that for $\alpha = 1$ (cyan), there are missing points. The maximum number of allowed iterations is 100, and reaching this point indicates that the simulation did not converge. For one of these points ($V_{\text{diff}} = 3$), the simulation was still converging (albeit very very slowly).

It is interesting to note that the simulations begin to fail at the optimal power point $V_{\text{diff}} = -3$. We can test the role that the field dependence of k_d has on this behavior by comparing the iteration count for simulations with the usual k_d and with $k_d = k_d(E_{\text{SC}})$. The results of this are shown in Figure 2.19. Note that we see no equivalent bump in the iteration count near the open circuit-voltage but have the same jump for potential differences just larger than zero. However, past the zero field point the constant k_d approximation begins to take significantly longer to converge. This is likely because for $E > 0$, $k_d(E) < k_d(0)$. We test this claim by running the simulation again with an increased value of $k_d(0)$ and note that it always takes more iterations to converge.

81
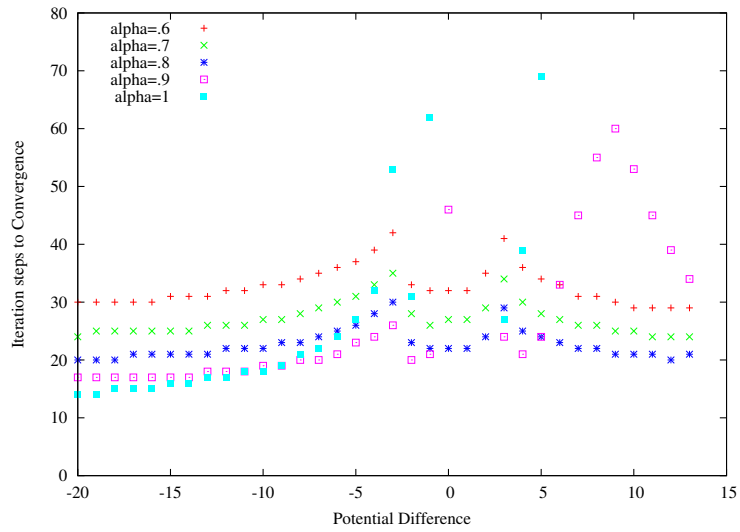
Figure 2.18: Plot of the number of iterations required for convergence for various α.
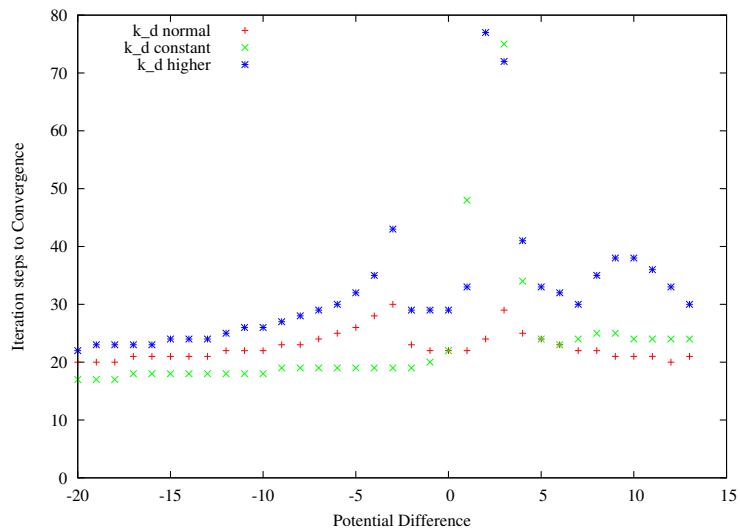
Figure 2.19: Plot of the number of iterations required for convergence for the usual value of k_d, $k_d = k_d(E_{\text{SC}})$ constant, and with $k_d(0) = \frac{k_d(E_{\text{SC}})}{k_d(0)} k_d(E)$ with $\alpha = .8$.

2.7 Improving the Asymptotics

The main objective of this section is to understand the physical characteristics of a photovoltaic device. Section 2.6 demonstrates that our large-field-approximation works well for the reverse-bias case, encompassing the short circuit current point. However, to fully understand the operating characteristics, we also need to understand the optimal power point and the open circuit case. At open circuit, the current is by definition zero, so we only really need to estimate the applied potential for which it occurs. For the optimal power point, we seek to estimate the power output of the device, which requires estimating both the field and the current.

By slightly rewriting the integrals Φ and $\mathcal{F}$ without assuming that the field is constant, we can derive an integral form for the current and the values of n and p. While not explicit, we can use the integrals to extend the asymptotics by manipulation of the exact current before inserting the asymptotic approximations. Furthermore, by considering the order of the integrals, we can also derive a simple approximation for the open circuit voltage which matches our numerical simulations.

This result suggests deep problems with the physical applicability of our model near open circuit, however, especially with regards to the minority carriers. In Section 2.8 we quantify these difficulties and discuss why a bulk drift-diffusion model cannot reasonably approximate the unipolar models discussed in the literature near open circuit.

2.7.1 Implicit Definition of J

We recall the definitions of Φ and $\mathcal{F}$ (in Section 2.5.1), but without rescaling n and p by μ_n and μ_p respectively:

$$\Phi_n(x) := \int_{x_0}^{x} \frac{e^{\varphi(x)-\varphi(y)}}{\mu_n} dy, \qquad\qquad \Phi_p(x) := \int_{x_0}^{x} \frac{e^{\varphi(y)-\varphi(x)}}{\mu_p} dy,$$

$$\mathcal{F}_n(x) := \int_{x_0}^{x} \frac{F(y)e^{\varphi(x)-\varphi(y)}}{\mu_n} dy, \qquad\qquad \mathcal{F}_p(x) := \int_{x_0}^{x} \frac{F(y)e^{\varphi(y)-\varphi(x)}}{\mu_p} dy,$$

where $F(y)$ also still contains the $c_r np$ term:

$$F(x) := \int_{x_0}^{x} \left(k_d(y)X^0(y) - c_r n(y)p(y) \right) dy.$$

Exploiting the symmetry of n and p, we can define integral quantities M^i and M_F^i and manipulate the expressions exactly before examining lowest order asymptotic terms.

The M will depend implicitly on the true values of V, n, p, and X, but the calculations we derive here are exact. We split Φ and $\mathcal{F}$ across the regions $\Omega_1 = [x_0, x_l]$, $\Omega_2 = [x_l, x_r]$ and $\Omega_3 = [x_r, x_L]$ corresponding to the region to the left of the interface, the interface region, and the region to the right of the interface. We have

$$\Phi(x_L) = e^{\pm\varphi(x_L)}(M^1 + M^2 + M^3)$$
$$\mathcal{F}(x_L) = e^{\pm\varphi(x_L)}(M_F^1 + M_F^2 + M_F^3)$$

where

$$M^1 = \int_{x_0}^{x_l} \frac{e^{\mp\varphi(y)}}{\mu(y)}, \qquad M^2 = \int_{x_l}^{x_r} \frac{e^{\mp\varphi(y)}}{\mu(y)}, \qquad M^3 = \int_{x_r}^{x_L} \frac{e^{\mp\varphi(y)}}{\mu(y)},$$

$$M_F^1 = \int_{x_0}^{x_l} \frac{F(y)e^{\mp\varphi(y)}}{\mu(y)}, \qquad M_F^2 = \int_{x_l}^{x_r} \frac{F(y)e^{\mp\varphi(y)}}{\mu(y)}, \qquad M_F^3 = \int_{x_r}^{x_L} \frac{F(y)e^{\mp\varphi(y)}}{\mu(y)}.$$

This notation has two important qualities. First, it allows us to proceed identically for n and p (changing only the sign in the exponential). Second, we have simple expressions for the quantities at x_l and x_r since they correspond with the integral endpoints. We have:

$$\Phi_{n/p}(x_l) = e^{\pm\varphi(x_l)}M^1 \qquad\qquad \Phi(x_r), = e^{\pm\varphi(x_r)}(M^1 + M^2),$$
$$\mathcal{F}_{n/p}(x_l) = e^{\pm\varphi(x_l)}M_F^1, \qquad\qquad \mathcal{F}(x_r) = e^{\pm\varphi(x_r)}(M_F^1 + M_F^2).$$

Inserting these expressions into Equations (2.29) and (2.30) (assuming zero boundary conditions) yields:

$$J_{n0} = \frac{M_F^1 + M_F^2 + M_F^3}{M^1 + M^2 + M^3}, \qquad J_{p0} = -\frac{M_F^1 + M_F^2 + M_F^3}{M^1 + M^2 + M^3}. \tag{2.36}$$

For justification of the zero-boundary-value assumption, see Figure 2.3 and the following discussion.

Since the sign of J_{p0} occurs twice – once in (2.14c) and once in (2.30), we have

$$n, p(x_l) = e^{\pm\varphi(x_l)}\left(\frac{M_F^1 + M_F^2 + M_F^3}{M^1 + M^2 + M^3}M^1 - M_F^1\right)$$
$$= e^{\pm\varphi(x_l)}\frac{M^1(M_F^2 + M_F^3) - (M^2 + M^3)M_F^1}{M^1 + M^2 + M^3}, \tag{2.37}$$

and

$$n, p(x_r) = e^{\pm\varphi(x_r)}\left(\frac{M_F^1 + M_F^2 + M_F^3}{M^1 + M^2 + M^3}(M^1 + M^2) - (M_F^1 + M_F^2)\right)$$
$$= e^{\pm\varphi(x_r)}\frac{(M^1 + M^2)M_F^3 - M^3(M_F^1 + M_F^2)}{M^1 + M^2 + M^3}. \tag{2.38}$$

84

2.7.2 Approximating the M^i

We begin by approximating the order of magnitude of the M^i using the mean value theorem in order to best approximate the value in the case of a large applied field. If the applied field and work function difference are the dominating terms, then the value of $\varphi'(x)$ will take characteristic values in each of the three regions of the device, Ω_1, Ω_2, and Ω_3. Defining E_l, E_m, and E_r as mean values for $-V'$ in these domains and W_m as a mean value for $-U'$ in the interface region, we have:

$$M_n^1 = \frac{e^{-\varphi(x_l)} - e^{-\varphi(x_0)}}{\mu_n(E_l)E_l}, \qquad\qquad M_p^1 = \frac{e^{\varphi(x_0)} - e^{\varphi(x_l)}}{\mu_p(E_l)E_l},$$

$$M_n^2 = \frac{e^{-\varphi(x_r)} - e^{-\varphi(x_l)}}{\mu_n(E_m)(E_m + W_m)}, \qquad\qquad M_p^2 = \frac{e^{\varphi(x_l)} - e^{\varphi(x_r)}}{\mu_p(E_m)(E_m + W_m)},$$

$$M_n^3 = \frac{e^{-\varphi(x_L)} - e^{-\varphi(x_r)}}{\mu_n(E_r)E_r}, \qquad\qquad M_p^3 = \frac{e^{\varphi(x_r)} - e^{\varphi(x_L)}}{\mu_p(E_r)E_r},$$

and

$$M_{F_n}^1 = \frac{F(x_l)e^{-\varphi(x_l)}}{\mu_n(E_l)E_l} - \frac{k_{d,l}X_l - c_r n_l p_l}{\mu_n(E_l)E_l^2}\left(e^{-\varphi(x_l)} - e^{-\varphi(x_0)}\right),$$

$$M_{F_n}^2 = \frac{F(x_r)e^{-\varphi(x_r)}}{\mu_n(E_m)(E_m + W_m)} - \frac{F(x_l)e^{-\varphi(x_l)}}{\mu_n(E_m)(E_m + W_m)}$$
$$- \frac{k_{d,m}X_m - c_r n_m p_m}{\mu_n(E_m)(E_m + W_m)^2}\left(e^{-\varphi(x_r)} - e^{-\varphi(x_l)}\right),$$

$$M_{F_n}^3 = \frac{F(x_L)e^{-\varphi(x_L)}}{\mu_n(E_r)E_r} - \frac{F(x_r)e^{-\varphi(x_r)}}{\mu_n(E_r)E_r} - \frac{k_{d,r}X_r - c_r n_r p_r}{\mu_n(E_r)E_r^2}\left(e^{-\varphi(x_L)} - e^{-\varphi(x_r)}\right),$$

and

$$M_{F_p}^1 = -\frac{F(x_l)e^{\varphi(x_l)}}{\mu_p(E_l)E_l} - \frac{k_{d,l}X_l - c_r n_l p_l}{\mu_p(E_l)E_l^2}\left(e^{\varphi(x_l)} - e^{\varphi(x_0)}\right)$$

$$M_{F_p}^2 = -\frac{F(x_r)e^{\varphi(x_r)}}{\mu_p(E_m)(E_m + W_m)} + \frac{F(x_l)e^{\varphi(x_l)}}{\mu_p(E_m)(E_m + W_m)}$$
$$- \frac{k_{d,m}X_m - c_r n_m p_m}{\mu_p(E_m)(E_m + W_m)^2}\left(e^{\varphi(x_r)} - e^{\varphi(x_l)}\right)$$

$$M_{F_p}^3 = -\frac{F(x_L)e^{\varphi(x_L)}}{\mu_p(E_r)E_r} + \frac{F(x_r)e^{\varphi(x_r)}}{\mu_p(E_r)E_r} - \frac{k_{d,r}X_r - c_r n_r p_r}{\mu_p(E_r)E_r^2}\left(e^{\varphi(x_L)} - e^{\varphi(x_r)}\right).$$

Note that if V and U are piecewise-linear, then the equations above are exact, and in particular we get exact expressions for the Φ and $\mathcal{F}$ defined in Section 2.5.1. There is an apparent discontinuity near $E = 0$, but if φ is constant, then the exponentials are

constant and the limit of the expression approaches the correct value. A sign change of E in Ω will result in the failure of the mean value theorem, but we note that this is not the case for the three important cases (short circuit, optimal power point, and open circuit) presented in Section 2.6.

Using the δ, η defined in Equation (2.28), we can also estimate the order of the M integrals:

$$M_n^1, M_{F_n}^1 = O(\delta_2^{-1}\eta^{-1}\delta_1^{-1}) + O(\delta_2^{-1}\eta^{-1}), \qquad M_p^1, M_{F_p}^1 = O(\delta_2\eta) + O(\delta_2\eta\delta_1),$$
$$M_n^2, M_{F_n}^2 = O(\delta_2^{-1}\eta^{-1}) + O(\delta_2^{-1}), \qquad M_p^2, M_{F_p}^2 = O(\delta_2) + O(\delta_2\eta),$$
$$M_n^3, M_{F_n}^3 = O(\delta_2^{-1}) + O(1), \qquad M_p^3, M_{F_p}^3 = O(1) + O(\delta_2).$$

2.7.3   Improved Asymptotic IV

The key limitation is that the approximation of the previous section ceases to be useful when $E \approx 0$ anywhere in the device, because the mean value expressions depend strongly on the evaluation point. Since the electric field appears in the denominator, small changes around zero will have a large impact. Since these were the explicit forms used in Section 2.6, we conjecture that we should be able to improve our asymptotic IV curves simply by numerically evaluating the M integrals. Although we still ignore the recombination (since we have no estimate on n and p) we can obtain further insights into the current-voltage characteristics of the device. The results are shown in Figure 2.20.

Since the numerically integrated (green) line in Figure 2.20 still uses the linear field approximation, the placement of the kink at V_{diff} is unsurprising, because this is the point where $E = 0$ and thus k_d changes form. In order to calculate the kink at the correct position, we need to better approximate the electric field. However, judging by the difference between the two asymptotic IV curves, we have little reason to trust our previous calculations of n and p as we move towards $V_{\text{diff}} = 0$. It would thus be preferable to calculate values of n and p based on the M^i integrals, but these only give us the values of n and p at the points x_l and x_r. Calculating the right-hand-side of Poisson's equation at any other point will require the evaluation of at least one double integral, and therefore integrating Poisson's equation twice to calculate the second-order potential would result in computationally inefficient quadruple integrals (not to mention then inserting these fields back into the double integrals for the M_F!).
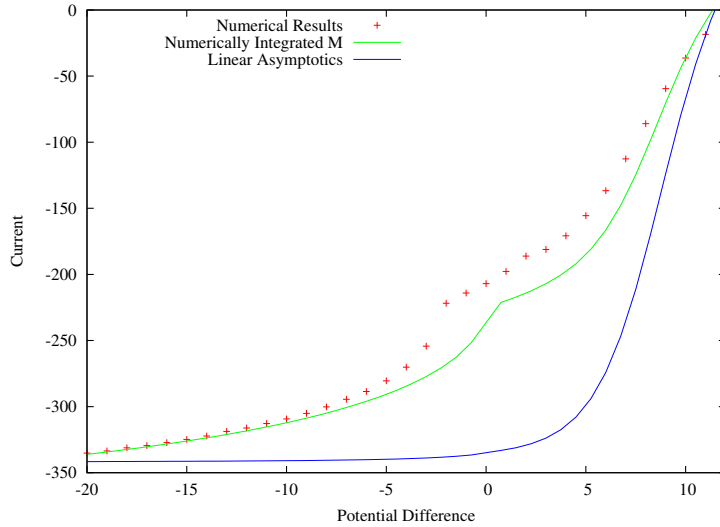
Figure 2.20: Numerical current-voltage characteristics plotted with the linear asymptotics from Section 2.5.1 and numerical integration of the M^i integrals (with no recombination). Note that the numerical integration predicts the kink in the IV curve, but at the wrong position.

Instead, since we are particularly interested in the optimal power point, we observe that n and p are approximately piecewise linear (consider Figure 2.9). Since we can calculate n and p at both x_l and x_r using the M integrals we've already evaluated (care of Equations (2.37) and (2.38)), calculating the linear approximation of n and p and integrating twice can be done analytically. Using the resulting approximation for E to again numerically integrate the M^i results in the asymptotic curve shown in Figure 2.21. We refer to these calculations as "second order" because they use the approximations derived from the original M^i to calculate a new value for the electric potential φ and then use this value to calculate the current.

We see that the linear-approximation for n and p allows us to obtain a much more accurate IV curve in the optimal power point regime, but that we lose accuracy spectacularly as we approach open circuit. In fact, we significantly underestimate the open circuit voltage, even though it was correctly approximated by both of the asymptotic curves presented in Figure 2.20. Heuristically, this seems to be expected since n and p are not linear near V_{OC}. By assuming that n and p were linear, we have thus made significant errors in the short circuit and open circuit regimes. At short circuit, the values
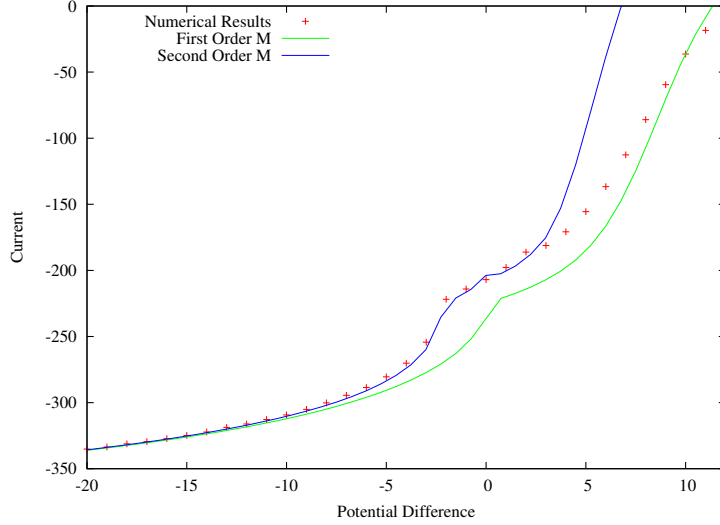
Figure 2.21: Numerical current-voltage characteristics plotted with the first and second-order numerical integration of the M^i integrals. Note that the second-order curve is a better approximation until $V_{\text{diff}} = 0$, but then fails dramatically.

of n and p are nearly constant on each side of the interface (modulo a boundary layer near the contacts). Thus we at most underestimate n and p by a factor of two. However, as we move towards open circuit, n and p decay away from the interface exponentially, and we significantly overestimate them. Thus the field changes more dramatically and we underestimate the open circuit voltage.

We note also that the necessary approximations for these second order M^i integrals are not trivial. If the field is not constant, then neither is $k_{d,in}$, and we can no longer explicitly calculate the equilibrium value of the excitons. Instead of trying to numerically evaluate X using the new field (resulting in undesirable triple integrals), we use the explicit version of X (as given in Appendix 2.A.1) using the local value of $k_{d,in}(E)$ at each point. This undoubtedly creates additional errors, but because the fit in Figure 2.21 is quite good, the error was deemed acceptable. Another option would be to use the equilibrium values of X_{in} and X_{out} (assuming $\mu_X \to 0$), but this leads to a worse fit. This is most likely related to the exciton mobility discussion in Section 2.6.1.

2.8 Obtaining a Unipolar Limit

One of the main goals of considering the full system is to understand the role of the minority carriers. Ideally we would like to find a limit in which our bulk model asymptotically approaches a unipolar model in some limit. The most sensible limits are the large work function limit and the thin-interface limit. We will see that although the minority carrier concentrations become quite small in these limits, the minority carrier current does not. From the drift-diffusion perspective we can take either the concentration or the flux to be zero, but if we take both conditions then the system becomes overdetermined.

Let us consider what happens if we take the original model equation with no minority carriers and no bulk exciton dissociation. Since $k_d X = c_r n p = 0$ in the bulk, we have that

$$F(x) = \begin{cases} 0 & x < x_l \\ \int_{x_l}^{x} k_{d,in} X(y) - c_r n(y) p(y) dy & x_l \leq x \leq x_r \\ F(x_L) & x > x_r \end{cases} . \tag{2.39}$$

This allows us to set $M_F^1 = 0$ and $M_F^3 = F(x_L) M^3$.

In particular, we then have the following expression for the currents at $x = 0$ (in the hole-material):

$$J_{n0} = \frac{M_{F_n}^2 + F(x_L) M_n^3}{M_n^1 + M_n^2 + M_n^3} \qquad J_{p0} = -\frac{M_{F_p}^2 + F(x_L) M_p^3}{M_p^1 + M_p^2 + M_p^3}$$

We use these formulas to plot an asymptotic IV curve and once again compare to the numerical results - see Figure 2.22.

2.8.1 The Small η Limit

Because η must be very small, we can consider the limit $\eta \to 0$, corresponding to an infinitely strong work function difference between the two halves of the device. In particular, this corresponds to a device in which it is physically impossible for electrons to pass into the hole material and for holes to pass into the electron material.
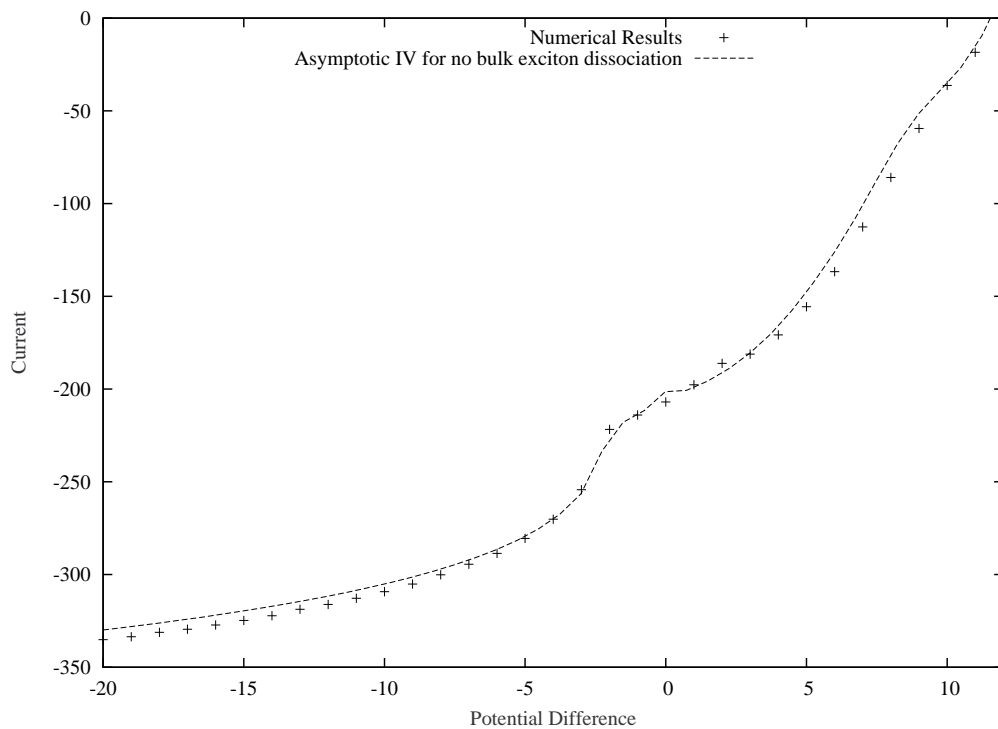
Figure 2.22: The *IV* curve obtained from our original numerical simulations plotted with the (lowest-order) M^i asymptotic *IV* curve for $k_{d,out} = 0$. Note that the fit is very good, suggesting that dynamics in the bulk do not contribute to the *IV* curve. In fact, because the asymptotic calculations do not consider recombination, this suggests that our numerical *IV* curve is relatively independent of recombination – an unphysical result for the bulk model. (Note, however, that this exactly matches some of the unipolar models given in the literature [5].)

This has the following effect on the M orders from the previous section:

$$\eta M_n^1 = O(\delta_2^{-1}\delta_1^{-1}) + O(\delta_2^{-1}), \qquad\qquad M_p^1 = 0,$$
$$\eta M_n^2 = O(\delta_2^{-1}), \qquad\qquad\qquad\qquad M_p^2 = O(\delta_2)$$
$$\eta M_n^3 = 0, \qquad\qquad\qquad\qquad\qquad M_p^3 = O(1) + O(\delta_2).$$

For the M_F (using Equation 2.39), we have:

$$\eta M_{F_n}^1 = 0, \qquad\qquad\qquad\qquad M_{F_p}^1 = 0,$$
$$\eta M_{F_n}^2 = O(\delta_2^{-1}), \qquad\qquad\qquad M_{F_p}^2 = O(\delta_2),$$
$$\eta M_{F_n}^3 = 0, \qquad\qquad\qquad\qquad M_{F_p}^3 = F(x_L)M_p^3.$$

We now have:

$$J_{n0} = \frac{M_{F_n}^2}{M_n^1 + M_n^2} + O(\eta)$$

$$J_{p0} = -\frac{M_{F_p}^2 + F(x_L)M_p^3}{M_p^2 + M_p^3} + O(\eta).$$

At short circuit, we have that $\delta_1, \delta_2 \ll 1$, and thus:

$$J_n^{SS} = 0 + O(\delta_1) + O(\delta_2) + O(\eta)$$
$$J_p^{SS} = -F(x_L) + O(\delta_2) + O(\eta)$$
$$J^{SS} \approx -F(x_L) \qquad\qquad\qquad\qquad\qquad (2.40)$$

which exactly matches our calculation in Equation (2.32) when $k_{d,out} = 0$. In addition, to lowest order we have that $J_n(x_0) = 0$, matching the intuition that there should be no electrons (and thus no electron current) in the p-material.

However, for $E > 0$, we have $\delta_1, \delta_2 > 1$, and in forward bias we expect $\delta_1, \delta_2 \gg 1$. In this case, we obtain

$$J_n^{FB} = O(1) + O(\delta_1^{-1}) + O(\delta_2^{-1})O(\eta) +$$
$$J_p^{FB} = -O(1) + O(\delta_2) + O(\eta).$$

Even with an infinitely strong work function difference, we have an $O(1)$ electron current in the hole material!

Although physically unrealistic, this effect is not necessarily surprising from a reaction-diffusion standpoint. Because of the quasi-positivity of the right-hand side $(k_d X - c_r np)$,

we generally expect that n and p will be positive everywhere. However, for a realistic model of the work function, we expect that the concentrations n and p and their derivatives ∇n and ∇p will be small on the minority side of the interface, even at open circuit. In fact, we see this qualitatively from the numerics in Figure 2.23.
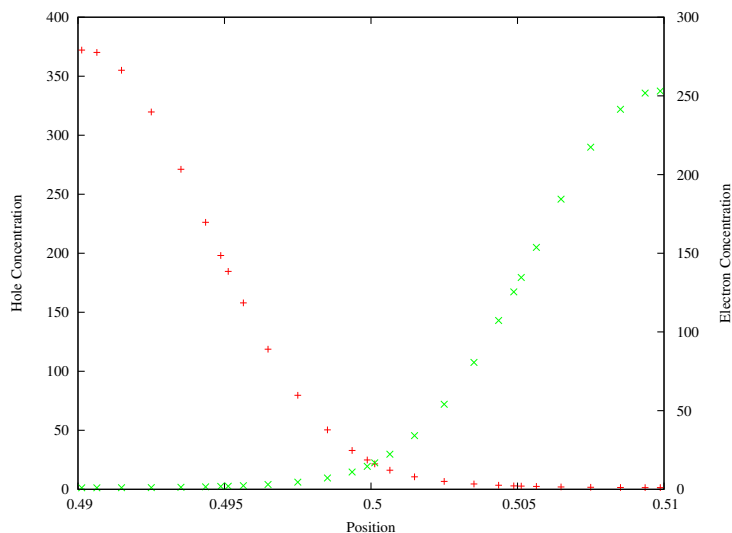
Figure 2.23: Plots of the numerically calculated values of n and p across the interface at open circuit with the usual parameters given in Table 2.1.

This is not the entire story. Although the minority carrier concentrations are exponentially suppressed at the edge of the interface, we know from Section 2.6 (Figures 2.7 and 2.8) that the current in the device is relatively independent of artificial constraints on the concentration values. In Figure 2.24 we plot the numerically calculated concentration for n in the p-type region at open circuit. We see that although the electrons are exponentially suppressed, we still have an electric current of $J_n \approx 100$. This electric current counteracts the hole current of $J_p \approx -100$ resulting in a thermodynamically unrealistic open circuit voltage. We discuss this result in the thin-interface limit regime in the next sections (specifically, see Figure 2.25 and the surrounding discussion).

Although a strict unipolar model will force $n = J_n = 0$ in the p-region, it will also mathematically violate the conservation of current. This can be heuristically described by assuming that the current at the interface entirely recombines (even if this is not included in the mathematical model [5]). However, neither solution suggests compatible mathematical and physical conditions for a unipolar model at open-circuit. Since the

92
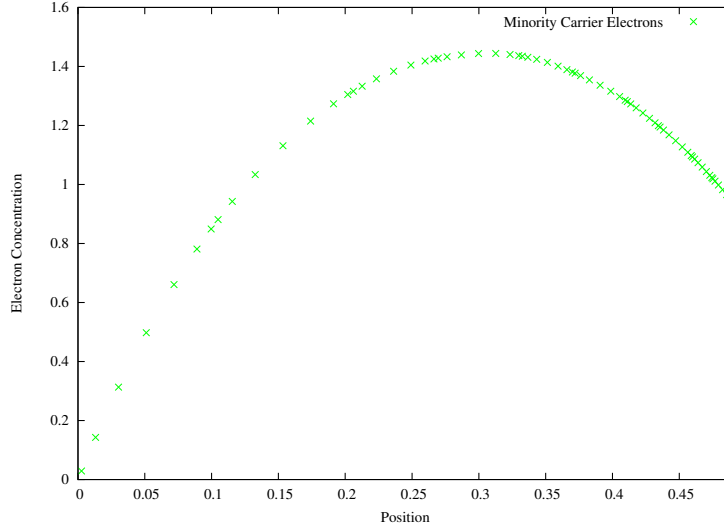
Figure 2.24: Plot of the numerically calculated value of the minority carrier n in the p-type region at open circuit with the usual parameters given in Table 2.1. Note that although n is several orders of magnitude smaller than p, it is not negligible and generates an appreciable current.

large work function limit does not give us a reasonable unipolar model, we proceed to consider the zero-interface limit.

2.8.2 Zero-Interface Limit

An interesting question is what limit our system approaches if we take the width of the interface to zero.If we require that the behavior of the device in the bulk is independent of size of the interface, then we find several necessary conditions.

Since the potential in the device is bounded, we immediately see that $M^2 \to 0$ in the limit. The integrand is bounded and the interval of integration goes to zero. We have to be more careful with the M_F. The forcing integral $F(x)$ is extremely important for the characteristics of the device. For instance, from Equation 2.40 we calculate the short circuit current $J_{tot} \approx F(x_L)$ at strong reverse-bias when $E \gg 1$. Since the short circuit current should not depend on the choice of interface width in the model, we must require that $F(x_L)$ remains constant as $\epsilon \to 0$. Since $F(x_l)$ and $F(x_L) - F(x_r)$ will remain nearly constant based on our bulk independence assumption, we conclude that $F(x)$ will have a jump at the new infinitesimal interface.

93

Since $F(x)$ is the integral of the right-hand-side $(k_d X - c_r np)$, we conclude that either $k_d X$ or $c_r np$ (or both) must weakly converge to a Dirac delta in the limit. This is to be expected for $k_d X$, since polaron pairs will only exist on the interface, and thus if the total number of polaron pairs $\int_I X$ is conserved as the size of the interface goes to zero, we expect that $X \rightharpoonup C\delta$. Note that this separates the polaron pair density from the exciton density, and if exciton dissociation in the bulk is non-negligible, we will retain the usual $k_{d,out} X$ term in the bulk. For the $c_r np$ term, we expect to end up using one-sided limits (from opposite sides) for n and p, corresponding to the dominance of the majority carrier (which we quantify in Equation (2.41)).

Since the zero-interface limit $F(x)$ is still bounded, we also have that $M_F^2 \to 0$ in the limit. With these two approximations, many of our equations become significantly simpler. As required in our assumptions, J_{SS} from the previous section is completely unchanged in the limit, since M^2 and M_F^2 are never dominant at short circuit.

The Jump in n and p

We can also calculate the ratio of $n(x_l)$ and $p(x_l)$ to $n(x_r)$ and $p(x_r)$. Starting from Equations (2.37) and (2.38), we have:

$$\frac{n,p(x_r)}{n,p(x_l)} = e^{\pm(\varphi(x_r)-\varphi(x_l))} \frac{(M^1 + M^2)M_F^3 - M^3(M_F^1 + M_F^2)}{M^1(M_F^2 + M_F^3) - (M^2 + M^3)M_F^1}$$
$$= \eta^{\mp 1} \frac{M^1 M_F^3 - M^3 M_F^1}{M^1 M_F^3 - M^3 M_F^1} = \eta^{\pm 1}. \tag{2.41}$$

Recall that $\eta = e^{\varphi(x_r)-\varphi(x_l)}$, and thus that in the limit, η is the jump in the potential between the left and right side of the interface. Physically, the electric potential should remain continuous, so the jump in φ will be due entirely to the work function gap: $\varphi(x_r) - \varphi(x_l) = \Delta U$. Thus η represents exactly the exponential of the difference in energies and therefore the concentration jump predicted by Boltzmann statistics. We therefore conclude that for the usual organic photovoltaic bilayers (where a large work function gap exists between the materials and exciton dissociation in the bulk of the device is negligible) the unipolar approximation discussed in Section 2.5.2 is justified. (The connection with problems illustrated in the previous section is not immediate, but we demonstrate below that the same issues with the minority carrier current appears.)

The Zero-Interface-Width Current

Continuing our approximations, we note that in the narrow-interface limit, we can also explicitly calculate the open circuit voltage. We begin by assuming that near V_{OC} our mean value approximations are valid because the field is strictly positive. Note that this seems reasonable since the asymptotic IV curves shown previously all match the numerically calculated V_{OC} with reasonable accuracy. In this case, using the results of Section 2.7.2 (with $M^2 = M_F^2 = 0$) we have:

$$\Phi_n(x_L) = \frac{1 - \delta_2^{-1} + \delta_2^{-1}\eta^{-1} - \delta_2^{-1}\eta^{-1}\delta_1^{-1}}{\mu_n(E)E}$$

$$\Phi_p(x_L) = \frac{\delta_1\eta\delta_2 - \eta\delta_2 + \delta_2 - 1}{\mu_p(E)E}$$

$$\mathcal{F}_n(x_L) = \frac{\left(1 - \delta_2^{-1}\right)F(x_L)}{\mu_n(E)E}$$

$$\mathcal{F}_p(x_L) = -\frac{\left(1 - \delta_2\right)F(x_L)}{\mu_p(E)E}.$$

Using Equation (2.36) we have:

$$
\begin{aligned}
J_{tot} &= \frac{\frac{\left(1-\delta_2^{-1}\right)F(x_L)}{\mu_n(E)E}}{\frac{1-\delta_2^{-1}+\delta_2^{-1}\eta^{-1}-\delta_2^{-1}\eta^{-1}\delta_1^{-1}}{\mu_n(E)E}} - \frac{-\frac{(1-\delta_2)F(x_L)}{\mu_p(E)E}}{\frac{\delta_1\eta\delta_2-\eta\delta_2+\delta_2-1}{\mu_p(E)E}} \\
&= \frac{\left(1 - \delta_2^{-1}\right)F(x_L)}{1 - \delta_2^{-1} + \delta_2^{-1}\eta^{-1} - \delta_2^{-1}\eta^{-1}\delta_1^{-1}} + \frac{\left(1 - \delta_2\right)F(x_L)}{\delta_1\eta\delta_2 - \eta\delta_2 + \delta_2 - 1} \\
&= \frac{\left(\delta_1\eta\delta_2 - \delta_1\eta\right)F(x_L)}{\delta_1\eta\delta_2 - \delta_1\eta + \delta_1 - 1} + \frac{\left(1 - \delta_2\right)F(x_L)}{\delta_1\eta\delta_2 - \eta\delta_2 + \delta_2 - 1}
\end{aligned}
\tag{2.42}
$$

It is possible to proceed in generality, but the calculations are simpler and more illuminating for a symmetric device. Thus for notational convenience we show the symmetric case here and relegate the non-symmetric case to Appendix 2.A.2. In a symmetric device, we would expect $\delta_1 = \delta_2 = \delta$. In this case, we have

$$
\begin{aligned}
J_{tot} &= \frac{\left(\delta^2\eta - \delta\eta\right)F(x_L)}{\delta^2\eta - \delta\eta + \delta - 1} + \frac{\left(1 - \delta\right)F(x_L)}{\delta^2\eta - \delta\eta + \delta - 1} \\
&= F(x_L)\frac{\delta\eta(\delta - 1) - (\delta - 1)}{\delta\eta(\delta - 1) + (\delta - 1)} \\
&= F(x_L)\frac{\delta\eta - 1}{\delta\eta + 1}
\end{aligned}
$$

This has several very interesting consequences. In reverse-bias it predicts that up to order $O(\delta\eta)$, the current is simply the net number of generated charges, $F(x_L)$, again

matching the calculations above. If dissociation outside of the interface is negligible, this exactly matches the approximate current given in Equation (2.32). This reinforces the intuition of a device in reverse bias pulling all of the charges out of the device. Even for $\delta \rightarrow 1$, the device continues to extract all of the generated charges if η is small. This case (which exists as a removable pole) corresponds to the zero internal-field case which in Section 2.6 corresponds roughly to the Optimal Power Point.

The Open Circuit Voltage

The next question is how to calculate the open circuit voltage. There are two possibilities, either $F(x_L) = 0$ or $\delta\eta - 1 = 0$. Whichever one of these points occurs first will be the observed open circuit voltage in the model. If $F(x_L) = 0$, then $J_n = J_p = 0$ and we have a consistent physical picture of an open circuit voltage. The other possibility, $\delta = \eta^{-1}$ yields a physically unrealistic case where we have equal and opposite electron and hole currents with no net current.

Let us proceed to calculate the open circuit voltage which corresponds to this unrealistic current. Recalling the definition of δ and η, we have:

$$e^{\varphi(x_r)} = e^{\varphi(x_r)-\varphi(x_l)} \rightarrow \varphi(x_l) = 0.$$

Using the definition of φ, we conclude that $V(x_l) = -U(x_l)$. Because we chose $\varphi(x_L) = 0$, we can rewrite the relationship as:

$$V(x_l) - V(x_L) = -U(x_l) + U(x_r).$$

Finally, if the width of the interface is small with respect to the device length, then our symmetry assumption allows us to write:

$$V(x_0) - V(x_L) = 2\left(U(x_r) - U(x_l)\right)$$

and given the definition of V_{diff}, we calculate the open circuit voltage:

$$V_{\text{OC}} = 2\Delta U.$$

Thus we estimate that the open circuit voltage in the device is approximately twice the work function gap between the materials at the interface. Note that for the value of $\Delta U = 6$ given in Table 2.1, this predicts $V_{\text{OC}} = 12$, which corresponds exactly to the results in Section 2.6. This indicates that the open circuit voltages considered within do

not correspond to physically realistic device operation (as they violate the principle of detailed balance).

In particular, it indicates that the point where $F(x_L) = 0$ (if it exists) requires a larger potential difference than the observed $V_{\text{diff}} = 12$. Our propositions about the behavior of the M integrals in the limit presupposed that the behavior of the device was preserved when taking the thin-interface limit. In particular, we implicitly assumed that $F(x_L)$ would remain constant as we took the interface width to zero. However, observation of Figure 2.23 suggests that this will not be the case. Integrating $F(x_L)$ over the interface gives significantly less recombination than using the values of the carriers at the left and right sides of the interface. Furthermore, results from unipolar models where recombination on the interface is ignored give good qualitative agreement to our model [5].

We speculate that at open circuit the violation of the conservation of current in the unipolar model gives precisely the same majority carrier results as the detailed-balance violation in the two carrier model. Work on establishing this correspondence is ongoing. Unless an extremely precise model of recombination on the interface is given, we must either assume a-priori that the electron and hole fluxes at the interface completely cancel in the form of recombination or accept (large) minority carrier currents.

2.8.3 Possible Routes for Improvement

In the previous two sections, we show that the minority carriers continue to play an important role even if we take an extremely large work function or an exponentially thin interface. In particular, an electron current exists in the hole-material, leading to a physically unrealistic open circuit voltage. In addition, we have shown that increasing the work function or decreasing the interface size will not correct the flaw.

One possible route for improvement is increasing the bimolecular recombination rate c_r inside the interface or decreasing the generation rate G. Increasing c_r by a factor of 100 and decreasing G by a factor of 100 still yields the same problem as before – we do not reach a point where $F(x_L) = 0$, and we obtain an unphysical V_{OC}. Carefully choosing coefficients could potentially produce a physical open circuit voltage, but the coefficients will no longer match physically observed quantities. Abandoning physical parameters to simulate physical behavior is not a particularly attractive option.

We can also consider the general structure of the variables which would allow a

physically realistic open circuit voltage. If we expect minority carrier current to be zero and continuous across the interface, then not only do we need $F(x_L) = 0$, but we also require that the right-hand-side of the charge carrier equations is zero. That is, we require:

$$(\nabla \cdot J_n)(x_l) = k_d(x_l)X(x_l) - c_r n(x_l)p(x_l) = 0$$
$$(\nabla \cdot J_p)(x_r) = k_d(x_r)X(x_r) - c_r n(x_r)p(x_r) = 0$$

where $k_d(x)$ is shorthand for $k_d\big(E(x)\big)$. If these conditions do not hold, then the minority carrier current cannot be zero on both sides of the interface boundary.

Since we have $n(x_l), p(x_r) \ll 1$, we will also require that $k_d X \ll 1$. Our chosen form of $k_d(E)$ does not approach zero for reasonable fields, and thus we will need to enforce that $X(x_l), X(x_r) \ll 1$. However, we model both excitons and polaron pairs with X and this requirement is not possible within the constraints of our model. One option would be to enforce $X(x_l) = X(x_r) = 0$ as boundary conditions for the polaron pairs on the interface. This, however, precludes us from considering excitons diffusing into the interface, because if we specify zero Dirichlet conditions as well as an incoming flux the problem becomes overdetermined.

From a physical perspective, we could simply ignore the minority carrier current by assuming that the fluxes at the interface exactly cancel in the form of recombination, but this is mathematically intractable for a reaction-diffusion interface, resulting in a violation of the conservation of current. In particular, we conclude that considering excitons in the bulk and polaron pairs which are free to diffuse within an interface of non-zero-width does not allow for physically realistic simulation of the open circuit voltage. Although it might be possible to design an alternative model with diffusion excitons and a thin interface region where a work function separates electrons and holes, this would require significant alterations to the model we have discussed here.

Although our model is incapable of describing the open circuit conditions accurately, it still gives interesting insights into the behavior of the device up through the optimal power point. In particular, Figure 2.25 shows that the electron current in the p-material is zero between short circuit and the optimal power point. In particular, since the asymptotic fit does not include recombination but still fits the numerical curves, this confirms the usual belief that bimolecular recombination does not play an important role in a photovoltaic device in the operating regime.

98
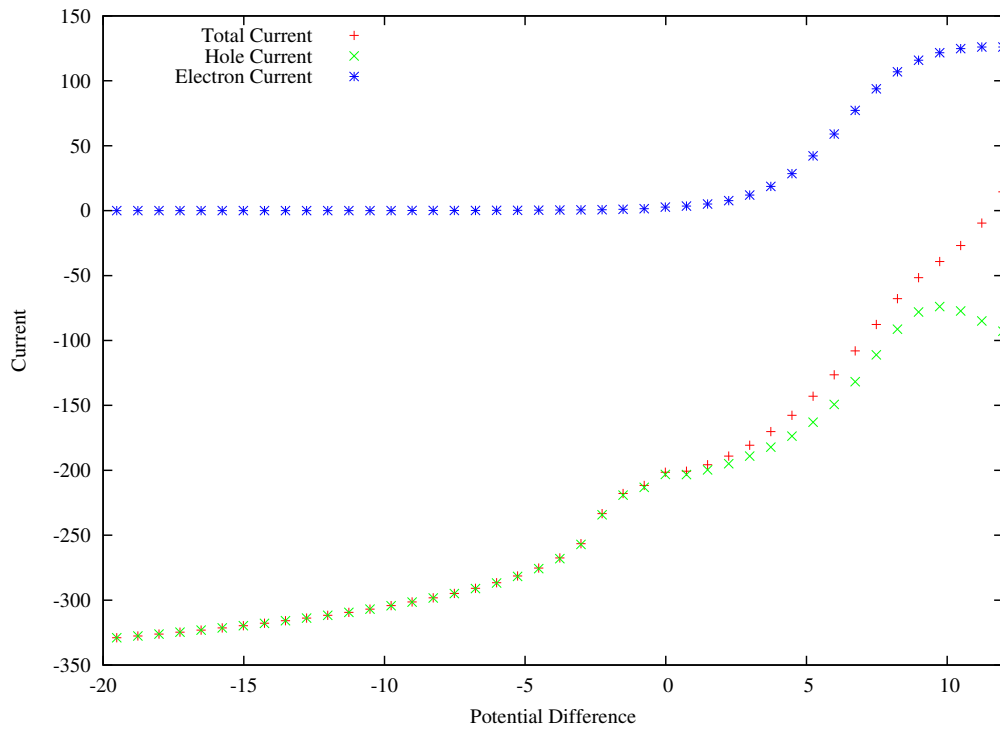
Figure 2.25: Asymptotic IV curve for $k_{d,out} = 0$ with J in blue (as in 2.22), J_n in gold, and J_p in red. Note that the current is taken at the left boundary (in the hole material), so that $J_n \neq 0$ corresponds to an unphysical current.

2.9   Extended 2D Simulations

Although the system presented in Section 2.1.1 is fully 2D, extending the implementation from the 1D analog case to a full 2D case with a nonlinear interface is quite challenging, especially if we want to enforce an interface which always has a width of 2nm. Of additional difficulty is appropriately smoothing the coefficients across the interface in a way which exactly aligns with the interface implemented by the mesh. We start this section by discussing the specific difficulties associated with implementing the work function U. We give a possible route for simplifying these difficulties and then show how the same technique can be used

The most obvious extension is to create piecewise linear interfaces by extending the simple smoothing used in Section 2.6. However, this results in severe problems at the corners of the interface – see Figure 2.26. In particular, since a piecewise quadratic U will have a discontinuous y derivative at the interface corners, numerical instabilities occur. In addition, if the partial derivative U_y is nonzero where the interface intersects with the Neumann boundary, we have unphysical conditions for the flux:

$$J_n|_{\Gamma_N} = \mu_n \nabla_y n - \nu_n n \nabla_y (U + V)) = \mu_n n \nabla_y U$$

where the last equality follows from the homogeneous Neumann conditions on n and V. If U does not also satisfy Neumann boundaries, the electric current through the Neumann boundaries is not zero. This also prevents us from considering simple bilayers where the interface is not perpendicular to the Neumann boundaries.

Because the 2D NetGen interface is only directly capable of creating lines and B-splines, the number of interfaces which can be easily described analytically are quite limited. However, one can create a nonlinear C^1 interface consisting of quarter-circles (the degenerate case for the B-Spline of three points which form an isosceles right-triangle) – see Figure 2.27. However, quarter-circles do not provide a particularly satisfying example since we are unable to change the proportion of the amplitude to the frequency. Instead, we need to consider meshes with geometries that NetGen is not capable of generating automatically.

Dynamically Generating U

One possibility for considering other meshes is to consider a sinusoidal mesh. In this case, we can explicitly sample n points from along the curve and additionally generate n
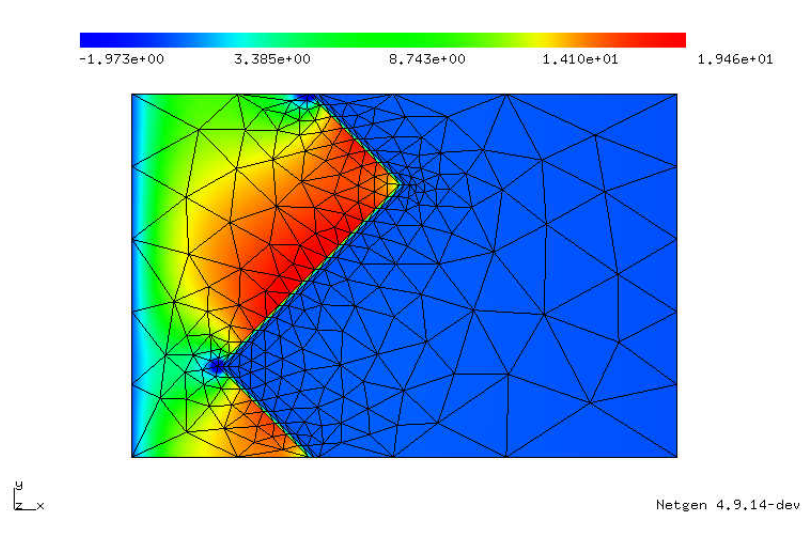
Figure 2.26: Sample numerical results for the hole concentrations for a device with a non-uniform interface at short circuit. Note that there are significant problems (i.e. negative concentrations) near the left-facing corners of the interface. These effects arise from the discontinuous nature of the piecewise quadratic U.

points along the left and right boundaries such that the interface has a constant width at every sample point (see Appendix 2.B). We can then use these points to define the edges of a triangulation and develop a new mesh. Because the equidistant points are difficult to handle explicitly, we then choose to dynamically calculate the value for U.

We can use a diffusion equation

$$\nabla \cdot (D\nabla U) = 0$$

with boundary conditions at the edges of the interface region (I) corresponding to the required maximum and minimum values of U. However, this cannot be done directly because we need U to be C^1 in the whole domain and requiring that $U = \nabla_\eta U = 0$ on both sides of I creates an ill-posed problem. Thus we need to solve the dynamic problem for U on the whole domain. Since we only want a thin transition layer between the two polymers near the interface, we need to change the diffusion coefficient D throughout the device. Once again, changing the coefficient discontinuously will create a discontinuity in the derivative of U, so we need a smooth approximation of a step function near the interface. If this is chosen carefully, we can finally observe dynamically generated work

101
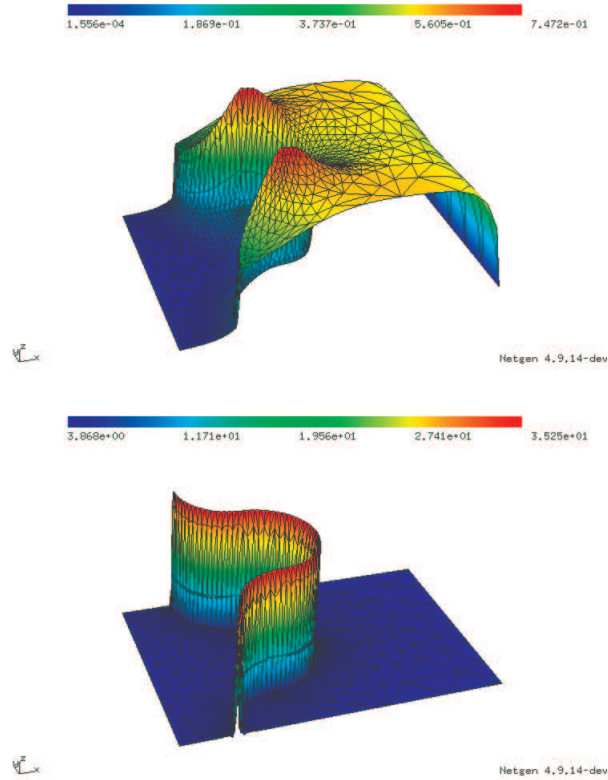
Figure 2.27: Sample numerical results for the electron and exciton/polaron pair concentrations for a device with a non-uniform interface at short circuit. The exciton concentration demonstrates the shape of the interface. The strange peaks in n appear as a result of the given C^1 form of U and the concentrating effects of the electric field in the vertical direction. Additional modeling work is necessary to obtain physically realistic results. Note that away from the interface we see the same nearly-constant behavior with a boundary layer near the metal contact.

functions such as the one shown in Figure 2.28, which uses a Gaussian with extremely small characteristic length in the x-direction centered on the interface. However, if the parameters are not carefully chosen the numerically calculated derivatives are not continuous. The careful tuning necessary to produce a valid U makes large-scale investigation of various interfaces extremely difficult.
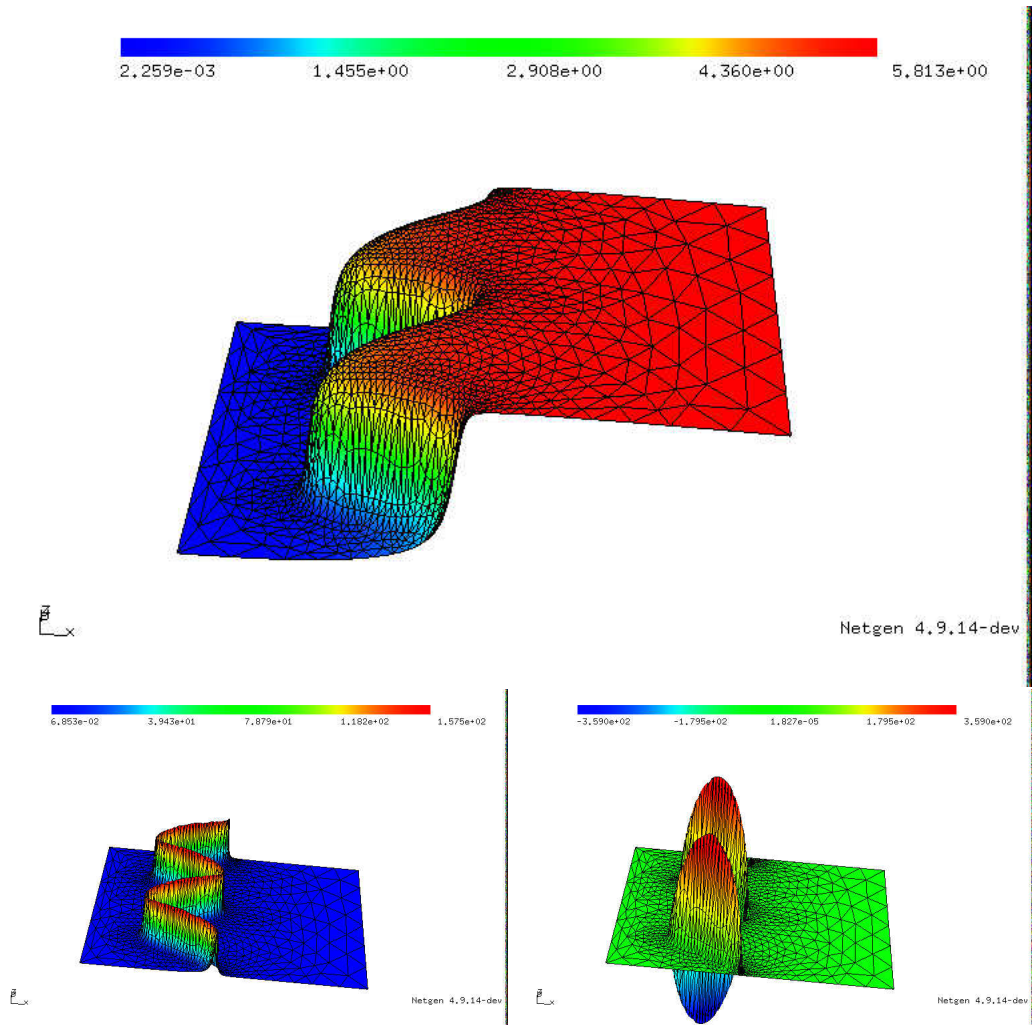
Figure 2.28: Dynamically generated form for the work function U in a device with a sinusoidal interface. The second row depicts the partial derivatives of U with respect to x and y on the left and right respectively.

Exciton Trapping Potential

An interesting benefit is that we can use the method of the previous section to produce an exciton trapping potential. In order to capture the appropriate behavior of the exciton to polaron pair transition (as discussed in Sections 2.6.1 and 2.8), we can add a convection term to our exciton equation:

$$\frac{\partial X}{\partial t} = \nabla \cdot (\mu_X \nabla X - \mu_X \nabla U^X X) + c_r' np + GT - k_d X - k_r TX.$$

where U^X is a confining potential for the polaron pairs. Thus polaron pairs tend to remain on the interface and the physically unrealistic transition from polaron pairs to excitons in suppressed. In practice, it is rather difficult to estimate a physically appropriate form of this function, leaving us free to use a method mirroring the generation of U. In order to solve the polaron pair to exciton transition problem, we must have that $\nabla U^X = O\left(\mu_X^{-1}\right)$ so that the convection contribution is non-negligible.

We can use Poisson's equation with the same diffusion coefficient used for generating the work function and a right-hand side which is negative on the interface. Note that the right-hand side does not need to be continuous, and we can simply take the negative characteristic function of I – the right-hand-side is -1 on the interface region and zero outside. This gives a C^1 function which we can use to trap the excitons on the interface – see Figure 2.29. Unfortunately, these functions do not provide particularly satisfying trapping potentials, because the depth of the well is not equal along the interface. This results because it is difficult to choose diffusion coefficients which match the interface exactly and preserve all symmetries. As such, the regions where the interface is approximately parallel to the contacts have slightly lower averaged diffusion coefficients and thus a more negative trapping potential.

After carefully selecting the necessary coefficients, we can finally obtain results for a sinusoidal interface in Figure 2.30. Note that for each figure, the 1D cross-sections appear to be identical to those given in Figure 2.9. We speculate that averaging over the solution to 1D problems with different interface widths of the device would provide an excellent approximation to the results (this idea was generated with the authors of [10] and [73] during a discussion of current projects). The main issue is that the problems with the work function and exciton-trapping term are not easily solved. In fact, careful selection of parameters is necessary for the numerical scheme to even converge.

It might be possible to handle these problems within the framework of Netgen and

104
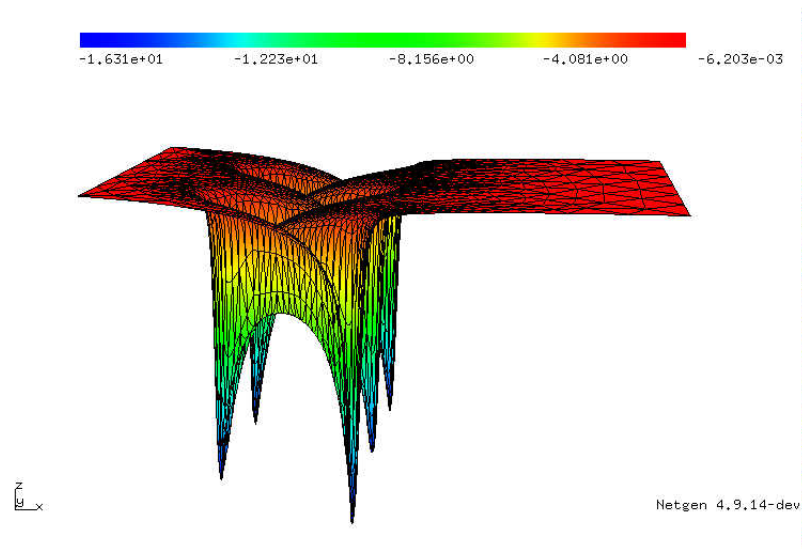
Figure 2.29: Dynamically generated form for the exciton trapping potential in a device with a sinusoidal interface. Note that the potential does not have a constant depth – this is a result of the small discrepancies between the theoretical sinusoidal interface and the versions which preserve the width of the interface.

NGSolve, but the extensive problems suggest that it might be more convenient to simply choose a new method. Although the programs are open-source and theoretically could be extended, this does not seem to be a feasible alternative, especially when easier simulations are possible. This especially holds for the case where we take the interface width to zero. In such a regime (as considered in the very recent work [20]), most of the problems presented in this section disappear. In particular, no work function is needed since there is simply a constant jump in the effective potential across the interface. In the simplest case, polaron pairs can be taken with fixed position on the interface, but it would also theoretically be possible to use the Laplace-Beltrami operator to model diffusion of the polaron pairs along the interface. If the interface is C^1, we can furthermore exactly define the direction of dipole moments.

2.10 Perspective: Proving Existence and Uniqueness

The reduced system of two reaction-convection-diffusion equations for n and p with Poisson's equation for V (also called the van Roosbroeck system) has received substantial
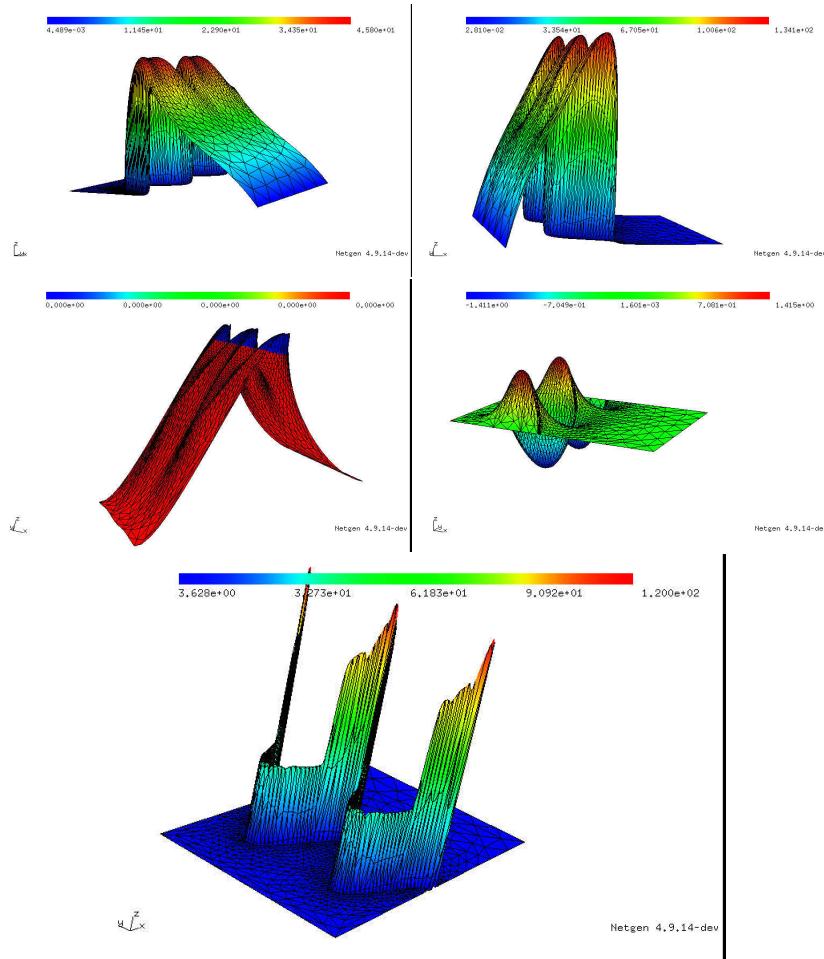
Figure 2.30: Simulation results for a 2D device with a sinusoidal interface for $V_{\mathrm{diff}} = -3$ corresponding to the optimal power point for the flat device (compare to Figure 2.9). The first row shows the electron and hole concentrations and the second row shows the electric field in the x and y directions (to the left and right respectively for both rows). The final image shows the exciton concentration in the device, which shows spikes corresponding to the lowest points in the trapping potential shown in Figure 2.29.

106

attention as a model for semiconductor devices. For substantial background on such systems, see Chapter 3 of [74]. Different methods have been used in various contexts to prove analytic results for systems in increasing generality. The system presented in Section 2.2 has several key extensions to the systems analyzed in the literature.

For convenience, we recall System (2.8) here (with the tilde notation suppressed):

$$\lambda^2 \nabla \cdot (\epsilon_r \nabla V) = n - p - \frac{h\chi_I}{L}\nabla_\nu X \tag{2.43a}$$

$$\frac{\partial n}{\partial t} = \nabla \cdot (\mu_n \nabla n - \mu_n n \nabla(U + V)) - c_r np + k_d X \tag{2.43b}$$

$$\frac{\partial p}{\partial t} = \nabla \cdot (\mu_p \nabla p + \mu_p p \nabla(U + V)) - c_r np + k_d X \tag{2.43c}$$

$$\frac{\partial X}{\partial t} = \nabla \cdot (\mu_X \nabla X) + c_r' np + G - k_d X - k_r X. \tag{2.43d}$$

One obvious extension is the addition of the exciton equation. Although a major component, the analysis for three drift-diffusion systems instead of two is more or less identical. We can also use a-priori estimates to show that X will only blow-up if n and p blow up simultaneously (see Section 2.10.3) and use the same machinery as presented in the literature [35]. The other structural change is the exciton dipole term in Poisson's equation. Assuming we use the same estimates for X as are used for n and p, we will lose one derivative on the right-hand-side of Poisson's equation. This can have dramatic consequences depending on the smoothness necessary for the proofs.

Finally, our given models for $R(n, p)$, μ, and k_d do not fit with the technical assumptions required by many analytical techniques. Most vitally, our mobilities do not saturate for large fields and are therefore a-priori unbounded. Since this requirement is generally used only technically for the analysis, cutting off μ for some value of the electric field E never physically expected should be sufficient. The other potential issue with the Poole-Frenkel [31] mobility is that it is not differentiable for $E = 0$. However, the model is based upon experimental data which results from conductance experiments. In particular, this means that they do not have data-points for $E \approx 0$ [31], and it should be reasonable to smooth μ in this regime. A similar problem may occur for k_d, which is also not differentiable for $E = 0$. In this case there are physical results for the dissociation at zero field, but it is unlikely that smoothing slightly in this regime will have any effect.

The constraints on $R(n, p)$ are potentially more limiting. Depending on the method used for the proofs, an assumption of sublinearity at infinity may be required. For our expression $R(n, p) = c_r np$ this clearly does not hold. However, since the exciton equation

107

contains the same term (up to a constant) with a different sign, it is likely that we could loosen this requirement, if not abolish it completely – see for instance [35].

A final quality is the dimensionality of the system. Many of the analytical techniques require a system in 2D. Although this is exactly the case considered for the numerical simulations presented above, it would be interesting to find a mathematical technique which was valid up to three dimensions in order to handle more complicated interfaces (particularly large-feature bulk heterojunction devices).

We now present three possible methods which could be used to establish new analytic results for the time-dependent organic photovoltaic system (2.8) and the corresponding steady state system (2.11).

2.10.1   Fixed Point Methods

Since the steady state system presented in Section 2.2.3 is of particular interest, it would be useful to establish the existence and uniqueness of this steady state. In particular, this could justify our numerical iteration scheme and quantify the damping results of Section 2.6.4. In addition, if we know that a steady state solution of specific smoothness exists, it might be possible to prove convergence to the steady state using other techniques. For the classical semiconductor system, there are extensive results for the steady state using fixed point methods. Of particular importance is the general fixed point result given in [55] for solutions in $H^1(\Omega) \cap L^\infty(\Omega)$. With several smoothing assumptions, our system satisfies most of the assumptions contained therein.

The most important missing requirement relates to the doping profile C, which is replaced by the dipole term $\frac{h\chi_I}{L}\nabla_\nu X$ in Poisson's equation. The requirement that $C \in L^\infty$ in [55] is not easily replicated for the exciton term. If we treat X analogously to n and p, then we will only have $X \in H^1 \cap L^\infty$, so $\nabla_\nu X \in L^2$ and we do not have obvious L^∞ estimates. This limitation is likely technical and more general existence and uniqueness theorems for elliptic PDEs (for instance in [34]) would likely provide a solution. The problem, however, is exasperated by further requirements on the smoothness of V. If $\mu(E)$ and $k_d(E)$ are not bounded from above for large E, then we require $E \in L^\infty$ and therefore $V \in W^{1,\infty}$. The lack of an $H^2 \to W^{1,\infty}$ embedding in two (or more) spatial dimensions means that even with full elliptic regularity results we will need $\nabla_\nu X \in L^p$ for $p > n$ (the spatial dimension).

Work on understanding the general case is ongoing, but at least some smoothing

assumptions on μ and k_d will likely be necessary to complete the proof. With k_d constant and $\mu_X = 0$, the exciton equation can be solved explicitly in terms of n and p. In such a case (with no dipoles and assuming μ_n and μ_p are bounded from above), the fixed point theorem from [55] can be applied directly to the organic model [21]. For more general results it may be easier to consider the dynamics of the time-dependent structure – we consider these possibilities in the following.

2.10.2 Entropy Methods

Entropy methods are a powerful tool for the analysis of partial differential equations [27], with such techniques being applied as early as 1985 for semiconductor equations [32]. Given specific simplifications to the structure of the equation, one can even use such entropy methods to prove exponential decay of the time-dependent solution to the steady state [23].

The first three terms in the entropy are the standard terms used for the semiconductor equations [23, 32], but we add the corresponding Boltzmann entropy of the exciton term:

$$F = \int_\Omega n \log \frac{n}{n^*} - (n - n^*) + p \log \frac{p}{p^*} - (p - p^*) + \frac{\lambda_D^2}{2}(\nabla V - \nabla V^*)^2$$
$$+ \int_\Omega X \log\left(X/X^*\right) - (X - X^*), \tag{2.44}$$

where $\lambda_D^2 := \lambda^2 \epsilon_r$.

Theorem 1. *Consider a local-in-time solution (V, n, p, X) to System (2.43) with no dipole term ($h = 0$), a field-independent dissociation rate ($k_d = k_d(x)$), no radiative losses ($c_r = c_r'$), and μ bounded from above. If the entropy functional F given by Equation (2.44) is bounded, then it stays bounded for all time and thus $V \in H^1$ and $n, p, X \in L^1$.*

Proof. We begin the proof by defining appropriate reference variables. We then show that the entropy F satisfies:

$$\frac{\partial F}{\partial t} \leq C(1 + F)$$

for some constant C independent of F and t. This will ensure via Gronwall inequality that at most the entropy grows exponentially, and therefore remains bounded for all time.

Reference Variables

Because the light generation term has the potential to create excitons (and therefore electrons and holes) from an otherwise stable system, we cannot expect exponential convergence to equilibrium. Instead, we choose arbitrary reference concentrations $n^*, p^*, V^* \in W^{1,\infty}$ which satisfy the same boundary conditions as n, p, and V. We require that the derivatives be bounded so that we have $J_n^* \in L^\infty$.

Finally, we can choose

$$X^* = \frac{k_d}{c_r n^* p^*},$$

so that $k_d X^* - c_r n^* p^* \equiv 0$ in Ω. Note that this requires that $c_r n^* p^* \neq 0$ everywhere, and in particular, we cannot take zero-boundary conditions. However, the exponentially small conditions discussed earlier [80] are still appropriate. This assumption is extremely important because it allows us to copy the analytical techniques for classical semiconductors. For a photovoltaic device, we expect constant generation via the photogeneration term G to drive extraction of charges from the device, and thus do not expect a steady state where $k_d X^\infty - c_r n^\infty p^\infty = 0$ (except in the $J = 0$ case – see Section 2.8).

Differentiating the Entropy

We differentiate F with respect to t and attempt to simplify and show that $\frac{\partial F}{\partial t}$ is bounded by some constant times F. Assuming that we can differentiate under the integral, we have:

$$\frac{\partial F}{\partial t} = \int_\Omega n_t \log \frac{n}{n^*} + p_t \log \frac{p}{p^*} + \lambda_D^2 (\nabla V - \nabla V^*) \cdot \nabla V_t + X_t \log \frac{X}{X^*}$$

$$= \int_\Omega (\nabla \cdot (J_n) - c_r np + k_d X) \log \frac{n}{n^*}$$

$$+ \int_\Omega (\nabla \cdot (J_p) - c_r np + k_d X) \log \frac{p}{p^*}$$

$$+ \int_\Omega \lambda_D^2 (\nabla V - \nabla V^*) \cdot \nabla V_t$$

$$+ \int_\Omega (\mu_X \Delta X + c_r' np + G - k_d X - k_r X) \log \frac{X}{X^*}.$$

Rearranging, we split the derivative into four terms:

$$\frac{\partial F}{\partial t} = F_1 + F_2 + F_3 + F_4$$

$$F_1 = \int_\Omega \nabla \cdot (J_n) \log \frac{n}{n^*} + \nabla \cdot (J_p) \log \frac{p}{p^*}$$

$$F_2 = -\int_\Omega (c_r np - k_d X) \left(\log \frac{n}{n^*} + \log \frac{p}{p^*} \right)$$

$$F_3 = \int_\Omega \lambda_D^2 (\nabla V - \nabla V^*) \cdot \nabla V_t$$

$$F_4 = \int_\Omega (\mu_X \Delta X + c_r' np + G - k_d X - k_r X) \log \frac{X}{X^*}$$

in order to simplify the notation for the following estimates. We proceed by simplifying each of these terms before recombining them.

We begin by integrating the first term by parts:

$$F_1 = \int_{\partial\Omega} J_n|_\nu \log \frac{n}{n^*} + J_p|_\nu \log \frac{p}{p^*} - \int_\Omega J_n \cdot \nabla \log \frac{n}{n^*} + J_p \cdot \nabla \log \frac{p}{p^*},$$

where the subscript ν represents a normal derivative. The boundary consists of two parts, with either Dirichlet or homogeneous Neumann boundary conditions for n, p, and V. At the Dirichlet boundaries $n = n^*$ and $p = p^*$ (by the definition of n^* and p^*) so the logarithms are zero. On the Neumann boundaries, $J_n|_\nu = J_p|_\nu = 0$, so the boundary term is zero. Expanding the gradients, we have

$$F_1 = -\int_\Omega J_n \cdot \nabla \log \frac{n}{n^*} + J_p \cdot \nabla \log \frac{p}{p^*}$$

$$= -\int_\Omega J_n \cdot \left(\frac{\nabla n}{n} - \frac{\nabla n^*}{n^*} \right) + J_p \cdot \left(\frac{\nabla p}{p} - \frac{\nabla p^*}{p^*} \right). \tag{2.45}$$

The second term we simply rearrange:

$$F_2 = -\int_\Omega (c_r np - k_d X) \log \frac{np}{n^* p^*}. \tag{2.46}$$

For the third term, again we integrate by parts:

$$F_3 = \int_{\partial\Omega} \lambda_D^2 (V - V^*) \nabla_\nu V_t - \int_\Omega (V - V^*) \lambda_D^2 \Delta V_t.$$

The boundary consists of two parts, with either Dirichlet or homogeneous Neumann boundary conditions for V. On the Dirichlet sections $V - V^* = 0$ (by the definition

of V^*) and on the Neumann sections $\nabla_\nu V_t = 0$. Thus the boundary term disappears. Assuming we can interchange the time derivative with the spatial derivatives,

$$F_3 = -\int_\Omega (V - V^*)(n_t - p_t)$$
$$= -\int_\Omega (V - V^*)(\nabla \cdot (J_n) - \nabla \cdot (J_p)).$$

We can now integrate by parts again:

$$F_3 = -\int_{\partial\Omega} (V - V^*)(J_n|_\nu - J_p|_\nu) + \int_\Omega (\nabla V - \nabla V^*) \cdot (J_n - J_p).$$

On the Dirichlet boundary sections, $V - V^* = 0$, and on the Neumann boundaries the normal fluxes are by definition zero. Thus the boundary term again cancels, yielding:

$$F_3 = \int_\Omega (\nabla V - \nabla V^*) \cdot (J_n - J_p). \tag{2.47}$$

Combining equations (2.45) and (2.47), we have

$$F_1 + F_3 = \int_\Omega \frac{J_n}{n} \cdot (n\nabla(U + V) - \nabla n) + \frac{J_n}{n^*} \cdot (-n^*\nabla(U + V^*) - n^*\nabla U + \nabla n^*)$$
$$+ \int_\Omega -\frac{J_p}{p} \cdot (p\nabla(U + V) + \nabla p) - \frac{J_p}{p^*} \cdot (-p^*\nabla(U + V^*) - \nabla p^*),$$

where we have added and subtracted both $J_n\nabla U$ and $J_p\nabla U$ to better illustrate the following step. Using the definitions of the current we can simplify this sum:

$$F_1 + F_3 = \int_\Omega -\frac{J_n^2}{\mu_n n} + \frac{J_n J_n^*}{\mu_n^* n^*} - \frac{J_p^2}{\mu_p p} + \frac{J_p J_p^*}{\mu_p^* p^*}.$$

We use Young's inequality on the steady state flux terms:

$$\frac{J J^*}{\mu^* u^*} \leq \frac{\epsilon J^2}{2\mu^* u} + \frac{\mu u J^{*2}}{2\epsilon \mu^{*2} u^{*2}}.$$

For any chosen $\epsilon < 2$, the negative flux terms will dominate the positive contribution, reducing the positive quadratic term to only the steady state currents. Choosing $\epsilon = 1$ for convenience, we have:

$$F_1 + F_3 \leq \int_\Omega -\frac{J_n^2}{2\mu_n n} + \frac{\mu_n n J_n^{*2}}{2\mu_n^{*2} n^{*2}} - \frac{J_p^2}{2\mu_p p} + \frac{\mu_p p J_p^{*2}}{2\mu_p^{*2} p^{*2}}.$$

However, by our choice of n^*, p^*, and V^* and our assumptions on μ, we have L^∞ bounds on all of the reference variables and μ, so we have:

$$F_1 + F_3 \leq \int_\Omega -\frac{J_n^2}{2\mu_n n} - \frac{J_p^2}{2\mu_p p} + \left\| \frac{J_n^{*2}\mu_n}{2\mu_n^{*2}n^{*2}} \right\|_\infty \int_\Omega n + \left\| \frac{J_p^{*2}\mu_p}{2\mu_p^{*2}p^{*2}} \right\|_\infty \int_\Omega p.$$

For F_4 we can integrate the Laplacian term by parts, yielding:

$$F_4 = \int_{\partial\Omega} \nabla_\nu X \log \frac{X}{X^*} + \int_\Omega -\frac{\mu_X |\nabla X|^2}{X} + \frac{\mu_X (\nabla X \cdot \nabla X^*)}{X^*}$$
$$+ \int_\Omega (c_r np + G - k_d X - k_r X) \log \frac{X}{X^*}.$$

The boundary term is zero because of the Neumann conditions on X. We once again use Young's inequality for the reference flux term, and again choosing $\epsilon = 1$, we have:

$$F_4 \leq \int_\Omega (c_r np + G - k_d X - k_r X) \log \frac{X}{X^*} - \frac{\mu_X |\nabla X|^2}{2X} + \frac{\mu_X X |\nabla X^*|^2}{2X^{*2}}$$
$$\leq \int_\Omega (c_r np + G - k_d X - k_r X) \log \frac{X}{X^*} + \left\| \frac{\mu_X |\nabla X^*|^2}{2X^{*2}} \right\|_\infty \int_\Omega X.$$

Noting the similarity of the logarithm term to F_4, we consider (temporarily moving the $\| \cdot \|$ term for notational convenience):

$$F_2 + F_4 - \| \cdot \| = \int_\Omega (c_r np + G - k_d X - k_r X) \log \frac{X}{X^*} - (c_r np - k_d X) \log \frac{np}{n^* p^*}$$
$$= \int_\Omega (G - k_r X) \log \frac{X}{X^*} - (c_r np - k_d X) \log \frac{X^* np}{X n^* p^*}$$
$$= \int_\Omega (G - k_r X) \log \frac{X}{X^*} - (c_r np - k_d X) \log \frac{c_r np}{k_d X},$$

where the last equality follows from our definition of X^*.

Closing the Entropy Dissipation Relation

Putting the expressions together, we have:

$$\frac{\partial F}{\partial t} \leq \int_\Omega -\frac{J_n^2}{\mu_n n} - \frac{J_p^2}{\mu_p p} - (c_r np - k_d X) \log \frac{c_r np}{k_d X} + (G - k_r X) \log \frac{X}{X^*}$$
$$+ \left\| \frac{J_n^{*2} \mu_n}{2\mu_n^{*2} n^{*2}} \right\|_\infty \int_\Omega n + \left\| \frac{\mu_p J_p^{*2}}{2\mu_p^{*2} p^{*2}} \right\|_\infty \int_\Omega p + \left\| \frac{\mu_X |\nabla X^*|^2}{2X^{*2}} \right\|_\infty \int_\Omega X. \qquad (2.48)$$

The first two terms are clearly strictly negative. The third term is also strictly negative because the logarithm is monotonic with a zero at the same point as the linear term.

Since all of the L^∞ terms are bounded, we can use the Csizsár-Kullback inequality:

$$\|u - u_\infty\|_{L^1}^2 \leq C \int u \log \frac{u}{u_\infty} - (u - u_\infty)$$

on the integrals to obtain

$$\left\| \frac{J_n^{*2} \mu_n}{2\mu_n^{*2} n^{*2}} \right\|_\infty \int_\Omega n + \left\| \frac{\mu_p J_p^{*2}}{2\mu_p^{*2} p^{*2}} \right\|_\infty \int_\Omega p + \left\| \frac{\mu_X |\nabla X^*|^2}{2X^{*2}} \right\|_\infty \int_\Omega X \leq CF$$

for some constant C which does not depend on t.

Finally, since $0 < G, k_r < \infty$, we can use the following trivial bounds:

$$\log \frac{X}{X^*} \le \log \frac{1}{X^*} + X - 1 \qquad \text{and} \qquad -X \log \frac{X}{X^*} \le \frac{X^*}{e}$$

to obtain:

$$\int_\Omega (G - k_r X) \log \frac{X}{X^*} \le \int_\Omega C + CX$$
$$\le C \left(|\Omega| + \|X\|_{L^1} \right)$$
$$\le C + CF$$

where the equalities result from the aforementioned bounds, the boundedness of the domain Ω, and the Csizár-Kullback inequality.

Using these inequalities with Equation (2.48) yields

$$\frac{\partial F}{\partial t} \le C(1 + F). \qquad (2.49)$$

Note that entropy decay cannot be expected because of the energy introduced to the system by the photogeneration G.

Using a Gronwall-type inequality, we see that:

$$\left(\frac{\partial F(t)}{\partial t} - CF(t) \right) \le C$$
$$\frac{\partial}{\partial t} \left(F(t) e^{-Ct} \right) \le C e^{-Ct}$$
$$F(t) e^{-Ct} - F(0) \le \int_0^t C e^{-C\tau} d\tau = \left[-e^{-C\tau} \right|_0^t = 1 - e^{-Ct}$$
$$F(t) \le F(t_0) e^{Ct} + \left(e^{Ct} - 1 \right)$$

and thus for a solution with initial entropy $F(t_0)$ bounded, the entropy remains bounded for all time. Therefore, again by the Csizár-Kullback inequality, we have that the L^1 norms of n, p, and X are bounded for all time, completing the proof.

$\square$

Conclusions

We have outlined using entropy methods to control any solutions with initially bounded entropy. Furthermore, following the argument presented in [35], it is possible to utilize

the bounded entropy and full elliptic regularity to obtain $L^{2+\epsilon}$ regularity for the 2D device (with no dipole term). Bootstrap arguments (presented for similar systems in [23, 35]) should then allow for even better regularity with the feasible goal of solutions in $L^\infty \cap H^1$ implying boundedness and finite energy. However, the argument is not valid in 3D and including the dipole term also causes the machinery to break down in 2D (for the reasons discussed in Section 2.10.1). For these cases existence of solutions is still an open problem.

One possibility for directly obtaining higher order estimates independent of the spatial dimension is using duality methods.

2.10.3 Duality Methods

Duality methods have shown great promise in deducing L^2 estimates directly for reaction-diffusion systems, especially positive systems with mass control [72]. Our system (modulo Poisson's equation), generally fits into this abstract setting if $c'_r \leq c_r$ (no carrier multiplication). In this case, the main difficulty is the mixed boundary conditions associated with a photovoltaic device. In particular, the boundary conditions of the dual problem cannot be chosen so that the boundary flux terms in the estimates all disappear.

The basic idea is as follows: given a weak solution to the system, we can take linear combinations of equations in order to obtain an inequality with only one variable on each side. Then, by a specific choice of our test function ϕ, we can obtain estimates which might otherwise have been quite difficult to obtain. In particular, since the dual method does not rely on the usual Sobolev embeddings, it is independent of dimension. Here we show estimates indicating that if blow-up in L^2 does occur, it must occur simultaneously for n, p, and X.

Bounding n and p by X

One direction can easily be calculated using the maximum principle (since n, p, $X > 0$).

Theorem 2. *For any weak solution to System* (2.43)

$$\|n\|_{L^2} \leq C\|k_d X\|_{L^2} \qquad and \qquad \|p\|_{L^2} \leq C\|k_d X\|_{L^2}.$$

Proof. For $u := n$ or $u := p$, we can define v such that

$$u_t - Lu = -c_r np + k_d X,$$

$$v_t - Lv = k_d X,$$

$$(u - v)_t - L(u - v) = -c_r np \leq 0.$$

If we choose v with the same boundary conditions and initial conditions as u, then by the maximum principle we have $u - v \leq 0$ and therefore $u \leq v$. Standard parabolic results on v yield L^p bounds in terms of $k_d X$. Since n and p are nonnegative, we have:

$$\|n\|_{L^2} \leq C\|k_d X\|_{L^2} \qquad \|p\|_{L^2} \leq C\|k_d X\|_{L^2} \tag{2.50}$$

where C is a constant depending on the coefficients of the equations (which are functions of ∇V but not n, p, or X) and the size of the domain. $\qquad\square$

Bounding X by n or p

The other direction is less straightforward, and we proceed using an argument based on the dual variable ϕ.

Theorem 3. *If we have L^2 bounds on the outgoing current, the initial data, and G for all time, then for any weak solution to System (2.43)*

$$\|X\|_{L^2} \leq C\left(1 + \|n\|_{L^2}\right) \qquad and \qquad \|X\|_{L^2} \leq C\left(1 + \|p\|_{L^2}\right).$$

Proof. Let us consider the combination:

$$(n + X)_t - \nabla \cdot (\mu_n \nabla n - \mu_n n \nabla V) + \mu_X \Delta X = -(c_r - c_r')np + G - k_r X$$
$$\leq G - k_r X$$

where the inequality holds because n and p are nonnegative and $c_r \geq c_r'$ (not all recombination yields excitons). We rearrange the terms and proceed to use a duality argument to bound X by n. Note that the same argument holds for p – only the sign of the convection term changes.

We proceed as follows:

$$X_t - \mu_X \Delta X + k_r X \leq -n_t + \nabla \cdot (\mu_n \nabla n - \mu_n n \nabla V) + G.$$

We multiply the equation by a test-function ϕ and integrate over the whole domain in space and up to time T:

$$\int_{\Omega_T} X_t\phi - \mu_X\Delta X\phi + k_r X\phi \leq \int_{\Omega_T} -n_t\phi + \nabla\cdot(\mu_n\nabla n - \mu_n n\nabla V)\phi + G\phi.$$

For notational simplicity, we rearrange the left hand side independently using integration by parts:

$$\int_{\Omega_T} X_t\phi - \mu_X\Delta X\phi + k_r X\phi$$

$$= \int_{\Omega_T} -X\phi_t - \mu_X X\Delta\phi + k_r X\phi + \left[\int_{\Omega} X\phi\right]_0^T - \int_{\partial\Omega_T}\mu_X\phi\nabla X\cdot\hat{\nu} + \int_{\partial\Omega_T}\mu_X X\nabla\phi\cdot\hat{\nu}$$

$$= \int_{\Omega_T} X\left(-\phi_t - \mu_x\Delta\phi + k_r\phi\right) + \int_{\Omega} X(T)\phi(T) - \int_{\Omega} X(0)\phi(0) + \int_{\partial\Omega_T}\mu_X X\nabla\phi\cdot\hat{\nu},$$

where the final identity occurs because X satisfies Neumann boundary conditions on the domain Ω. Since we have no a-priori knowledge about the future behavior of X, we will choose a specific test function ϕ which solves the problem:

$$-\phi_t - \mu_x\Delta\phi + k_r\phi = \Theta \text{ in } \Omega_T$$

$$\phi(T) = 0 \text{ in } \Omega \qquad \nabla\phi\cdot\hat{\nu} = 0 \text{ in } \partial\Omega$$

for some $\Theta \geq 0$. Moving the initial condition integral to the right hand side, our original inequality now reads

$$\int_{\Omega_T} X\Theta \leq \int_{\Omega} X(0)\phi(0) + \int_{\Omega_T} -n_t\phi + \nabla\cdot(\mu_n\nabla n - \mu_n n\nabla V)\phi + G\phi$$

Integrating the final integral by parts yields:

$$\int_{\Omega_T} -n_t\phi + \nabla\cdot(\mu_n\nabla n - \mu_n n\nabla V)\phi + G\phi$$

$$= \int_{\Omega_T} n\phi_t - (\mu_n\nabla n - \mu_n n\nabla V)\cdot\nabla\phi + G\phi + \left[\int_{\Omega} n\phi\right]_0^T + \int_{\partial\Omega_T}\phi(\mu_n\nabla n - \mu_n n\nabla V)\cdot\hat{\nu}$$

$$= \int_{\Omega_T} n\phi_t + n\nabla\cdot(\mu_n\nabla\phi) + \mu_n n\nabla V\cdot\nabla\phi + G\phi - \int_{\Omega} n(0)\phi(0) + \int_{\partial\Omega_T}\phi J_n\cdot\hat{\nu} - \int_{\partial\Omega_T}\mu_n n\nabla\phi\cdot\hat{\nu}$$

$$= \int_{\Omega_T} n\left(\phi_t + \nabla\cdot(\mu_n\nabla\phi) + \mu_n\nabla V\cdot\nabla\phi\right) + G\phi - \int_{\Omega} n(0)\phi(0) + \int_0^T\int_{\partial\Omega_D}\phi J_n\cdot\hat{\nu}.$$

The final boundary term is unusual in that it does not exactly cancel. On the Dirichlet boundary we only expect the current to be zero for specific operating conditions (i.e.

specific boundary conditions for V). Finally, using Hölder's inequality, we obtain:

$$\int_{\Omega_T} X\Theta \leq \|X(0) - n(0)\|_{L^p} \|\phi(0)\|_{L^q} + \|n\|_{L^p} \|\phi_t + \nabla \cdot (\mu_n \nabla \phi) + \mu_n \nabla V \cdot \nabla \phi\|_{L^q}$$
$$+ \|\phi\|_{L^q} \int_0^T \left(\|G(s)\|_{L^p} + \|J_n \cdot \hat{\nu}\|_{L^p(\partial \Omega_D)} \right) ds.$$

From standard estimates for Parabolic PDEs, we have:

$$\text{ess sup}_{0 \leq t \leq T} \left(\|\phi(t)\|_{H^2(\Omega)} \right) + \|\phi_t\|_{L^2(0,T;H_0^1(\Omega))} \leq C \left(\|f\|_{H^1(0,T;L^2(\Omega)} + \|\phi(T)\|_{H^2(\Omega)} \right)$$

([26], Theorem 7.1.5). This proof assumes that the non-divergence form coefficients are smooth. In particular, it would require $\mu \in C^1$ for the divergence case, but this technical assumption is probably not necessary. In particular, for Θ constant and $\phi(T) = 0$, we have:

$$\|\phi(0)\|_{L^2(\Omega)} + \|\phi_t\|_{L^2(\Omega_T)} + \|\nabla \phi\|_{L^2(\Omega_T)} + \|\Delta \phi\|_{L^2(\Omega_T)} \leq C\Theta.$$

Inserting these estimates into our inequality we have:

$$\int_{\Omega_T} X\Theta \leq C\Theta \|X(0) - n(0)\|_{L^2} + C\Theta \|n\|_{L^2} \left(1 + \|\nabla \mu_n\|_{L^\infty} + \|\mu_n\|_{L^\infty} \right)$$
$$+ C\Theta \int_0^T \left(\|G(s)\|_{L^2} + \|J_n \cdot \hat{\nu}\|_{L^2(\partial \Omega_D)} \right) ds$$

and by duality,

$$\|X\|_{L^2} \leq C \left(1 + \|X(0) - n(0)\|_{L^2} + \|n\|_{L^2} \left(1 + \|\mu_n\|_{W^{1,\infty}} \right) \right)$$
$$+ C \int_0^T \left(\|G(s)\|_{L^2} + \|J_n \cdot \hat{\nu}\|_{L^2(\partial \Omega_D)} \right) ds. \tag{2.51}$$

Therefore, if we have L^2 bounds on the outgoing current, the initial data, and G for all time, then with μ_n, μ_p Lipshitz, we have:

$$\|X\|_{L^2} \leq C \left(1 + \|n\|_{L^2} \right) \qquad \text{and} \qquad \|X\|_{L^2} \leq C \left(1 + \|p\|_{L^2} \right) \tag{2.52}$$

where the final constant C depends on T and the outgoing current $J_n \cdot \nu$ (as shown in Equation 2.51). $\qquad \qquad \square$

Combined with the results from above, we see that the L^2 norms of n, p, and X can each be controlled in terms of the others. In particular, if any one of the equations in the system is shown to be well-posed, Equations (2.50) and (2.52) give well-posedness for all of the variables.

We have presented several results, all of which show some promise for producing more general results in the future. Relaxing the assumptions for the fixed point or entropy results could produce interesting results for the full system, especially if they can be expanded to handle a field-dependent dissociation rate $k_d(E)$. Extensions to the duality estimates could also provide interesting results independent of dimension (unlike the entropy results presented in Section 2.10.2).

2.A   Explicit Calculations

2.A.1   Asymptotic $W^{1,\infty}$ Bounds for X

We seek to derive bounds for X and X_x which are valid in the asymptotic regime of Section 2.5.1.

We begin with the exciton equation:

$$-\mu_X X_{xx} = O(\epsilon)np + G - (k_d + k_r)X \qquad X_x(0) = X_x(x_L) = 0.$$

For notational convenience we define $K_{out} = k_{d,out} + k_{r,out}$ and $K_{in} = k_{din} + k_{rin}$. Furthermore, G is generally large $(O(1/\epsilon^2))$ so we assume that $G + O(\epsilon)np \approx G$ is constant for the working regime. Splitting the problem into three domains, we have:

$$X_{xx} = \begin{cases} \frac{K_{out}}{\mu_X}X - \frac{G}{\mu_X} & x_0 < x < x_l \\ \frac{K_{in}}{\mu_X}X - \frac{G}{\mu_X} & x_l < x < x_r \\ \frac{K_{out}}{\mu_X}X - \frac{G}{\mu_X} & x_r < x < x_L \end{cases}.$$

Each of these equations can be solved explicitly using hyperbolic trigonometric functions, and using the Neumann boundary conditions allows us to write:

$$X = \begin{cases} A_l \cosh\left(\sqrt{\frac{K_{out}}{\mu_X}}(x - x_0)\right) + \frac{G}{K_{out}} & x_0 < x < x_l \\ A_m \cosh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x - x_l)\right) + B_m \sinh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x - x_l)\right) + \frac{G}{K_{in}} & x_l < x < x_r \\ A_r \cosh\left(\sqrt{\frac{K_{out}}{\mu_X}}(x - x_L)\right) + \frac{G}{K_{out}} & x_r < x < x_L \end{cases}.$$
(2.53)

Using continuity at x_l yields:

$$A_l = \frac{1}{\cosh\left(\sqrt{\frac{K_{out}}{\mu_X}}(x_l - x_0)\right)}\left(A_m + \frac{G}{K_{in}} - \frac{G}{K_{out}}\right).$$

Continuity of the derivative (which follows from allowing only jump discontinuities in K_{out} and K_{in}) yields:

$$
\begin{aligned}
B_m &= \sqrt{\frac{\mu_X}{K_{in}}} A_l \sqrt{\frac{K_{out}}{\mu_X}} \sinh\left(\sqrt{\frac{K_{out}}{\mu_X}}(x_l - x_0)\right) \\
&= \sqrt{\frac{K_{out}}{K_{in}}} \tanh\left(\sqrt{\frac{K_{out}}{\mu_X}}(x_l - x_0)\right)\left(A_m + \frac{G}{K_{in}} - \frac{G}{K_{out}}\right) \\
&\approx \sqrt{\frac{K_{out}}{K_{in}}}\left(A_m + \frac{G}{K_{in}} - \frac{G}{K_{out}}\right),
\end{aligned}
$$

where the last approximation is exact to better than 10^{-100} since $\sqrt{\frac{K_{out}}{\mu_X}}(x_l - x_0) >> 0$ for typical device values.

The more difficult approximations occur at x_r. From continuity we have:

$$
A_r = \frac{A_m \cosh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right) + B_m \sinh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right) + \frac{G}{K_{in}} - \frac{G}{K_{out}}}{\cosh\left(\sqrt{\frac{K_{out}}{\mu_X}}(x_r - x_L)\right)}
$$

and from continuity of the derivative we have:

$$
A_r = \frac{A_m \sqrt{\frac{K_{in}}{\mu_X}} \sinh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right) + B_m \sqrt{\frac{K_{in}}{\mu_X}} \cosh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right)}{\sqrt{\frac{K_{out}}{\mu_X}} \sinh\left(\sqrt{\frac{K_{out}}{\mu_X}}(x_r - x_L)\right)}.
$$

Here we have a convenient cancellation of all dependence from the length of the right-hand side as long as the device size is physically reasonable $\sqrt{\frac{K_{out}}{\mu_X}}(x_r - x_L) << 0$. Note that this expression is negative because $x_L > x_r$. The hyperbolic tangent which results from setting the two equations for A_r equal to each other and dividing the hyperbolic trigonometric functions from the denominators is equal to -1 within many orders of magnitude. This yields the following equality:

$$
\begin{aligned}
A_m \cosh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right) + B_m \sinh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right) + \frac{G}{K_{in}} - \frac{G}{K_{out}} = \\
= -A_m \sqrt{\frac{K_{in}}{K_{out}}} \sinh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right) - B_m \sqrt{\frac{K_{in}}{K_{out}}} \cosh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right).
\end{aligned}
$$

We can further simplify the expression by noting that for our device values

$$
\tanh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right) \approx 1
$$

120

with errors on the order of 10^{-10}. Thus we can replace the hyperbolic sine functions by the equivalent cosines, yielding:

$$(A_m + B_m)\left(\sqrt{\frac{K_{in}}{K_{out}}} + 1\right)\cosh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right) = \frac{G}{K_{out}} - \frac{G}{K_{in}}.$$

We can now plug our value for B_m into this equation and solve for A_m. Algebraic manipulation yields

$$A_m = \frac{A'_m}{\cosh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right)} + \frac{G}{K_{in}}\left(\sqrt{\frac{K_{in}}{K_{out}}} - 1\right),$$

where we define

$$A'_m = \frac{\frac{G}{K_{out}} - \frac{G}{K_{in}}}{\left(1 + \sqrt{\frac{K_{in}}{K_{out}}}\right)\left(1 + \sqrt{\frac{K_{out}}{K_{in}}}\right)}$$

for notational convenience in the following calculations.

We can then immediately evaluate B_m and A_l:

$$B_m = \sqrt{\frac{K_{out}}{K_{in}}}\frac{A'_m}{\cosh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right)} + \frac{G}{K_{in}}\left(1 - \sqrt{\frac{K_{in}}{K_{out}}}\right),$$

$$A_l = \frac{\frac{G}{K_{out}}\left(\sqrt{\frac{K_{out}}{K_{in}}} - 1\right) + \frac{A'_m}{\cosh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right)}}{\cosh\left(\sqrt{\frac{K_{out}}{\mu_X}}(x_l - x_0)\right)}.$$

Evaluating A_r still requires some simplification and assumptions:

$$A_r = \frac{A'_m + \sqrt{\frac{K_{out}}{K_{in}}}A'_m\tanh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right) + \frac{G}{K_{in}} - \frac{G}{K_{out}}}{\cosh\left(\sqrt{\frac{K_{out}}{\mu_X}}(x_r - x_L)\right)}$$

$$+ \frac{\frac{G}{K_{in}}\left(\sqrt{\frac{K_{in}}{K_{out}}} - 1\right)\cosh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right)}{\cosh\left(\sqrt{\frac{K_{out}}{\mu_X}}(x_r - x_L)\right)}$$

$$+ \frac{\frac{G}{K_{in}}\left(1 - \sqrt{\frac{K_{in}}{K_{out}}}\right)\sinh\left(\sqrt{\frac{K_{in}}{\mu_X}}(x_r - x_l)\right)}{\cosh\left(\sqrt{\frac{K_{out}}{\mu_X}}(x_r - x_L)\right)}.$$

We can simplify the form for A_r by again using that the hyperbolic tangent is nearly one, and noting the identity $(\cosh x - \sinh x) = e^{-x}$. Furthermore, we can simplify the

following expression:

$$A'_m \left(\sqrt{\frac{K_{out}}{K_{in}}} + 1 \right) + \frac{G}{K_{in}} - \frac{G}{K_{out}} = \frac{G}{K_{out}} \left(\sqrt{\frac{K_{out}}{K_{in}}} - 1 \right).$$

Putting all of this together yields:

$$A_r = \frac{\frac{G}{K_{out}} \left(\sqrt{\frac{K_{out}}{K_{in}}} - 1 \right) + \frac{G}{K_{in}} \left(\sqrt{\frac{K_{in}}{K_{out}}} - 1 \right) e^{-\left(\sqrt{\frac{K_{in}}{\mu_X}} (x_r - x_l) \right)}}{\cosh \left(\sqrt{\frac{K_{out}}{\mu_X}} (x_r - x_L) \right)}.$$

Having explicitly calculated the coefficients A and B, Equation (2.53) now gives an explicit formulation for the exciton concentration. This expression is long and not very illuminating, but we can further eliminate small terms.

The hyperbolic secant term in the first expression of A_l and the exponential decay term in A_r are both small (by a factor of about 10^{-5}). Eliminating these terms yields:

$$X = \begin{cases} \frac{G}{K_{out}} \left(\sqrt{\frac{K_{out}}{K_{in}}} - 1 \right) \frac{\cosh \left(\sqrt{\frac{K_{out}}{\mu_X}} (x - x_0) \right)}{\cosh \left(\sqrt{\frac{K_{out}}{\mu_X}} (x_l - x_0) \right)} + \frac{G}{K_{out}} & x_0 < x < x_l \\[2em] \begin{aligned} &\frac{A'_m \cosh \left(\sqrt{\frac{K_{in}}{\mu_X}} (x - x_l) \right) + \sqrt{\frac{K_{out}}{K_{in}}} A'_m \sinh \left(\sqrt{\frac{K_{in}}{\mu_X}} (x - x_l) \right)}{\cosh \left(\sqrt{\frac{K_{in}}{\mu_X}} (x_r - x_l) \right)} \\ &+ \frac{G}{K_{in}} \left(\sqrt{\frac{K_{in}}{K_{out}}} - 1 \right) e^{-\left(\sqrt{\frac{K_{in}}{\mu_X}} (x - x_l) \right)} + \frac{G}{K_{in}} \end{aligned} & x_l < x < x_r \quad . \quad (2.54) \\[2em] \frac{G}{K_{out}} \left(\sqrt{\frac{K_{out}}{K_{in}}} - 1 \right) \frac{\cosh \left(\sqrt{\frac{K_{out}}{\mu_X}} (x - x_L) \right)}{\cosh \left(\sqrt{\frac{K_{out}}{\mu_X}} (x_r - x_L) \right)} + \frac{G}{K_{out}} & x_r < x < x_L \end{cases}$$

We can also write X_x explicitly:

$$X_x = \begin{cases} \dfrac{G}{\sqrt{\mu_X K_{out}}} \left(\sqrt{\dfrac{K_{out}}{K_{in}}} - 1 \right) \dfrac{\sinh\left(\sqrt{\dfrac{K_{out}}{\mu_X}} (x-x_0) \right)}{\cosh\left(\sqrt{\dfrac{K_{out}}{\mu_X}} (x_l-x_0) \right)} & x_0 < x < x_l \\[3em] \dfrac{\sqrt{\dfrac{K_{in}}{\mu_X}} A'_m \sinh\left(\sqrt{\dfrac{K_{in}}{\mu_X}} (x-x_l) \right) + \sqrt{\dfrac{K_{out}}{\mu_X}} A'_m \cosh\left(\sqrt{\dfrac{K_{in}}{\mu_X}} (x-x_l) \right)}{\cosh\left(\sqrt{\dfrac{K_{in}}{\mu_X}} (x_r-x_l) \right)} & x_l < x < x_r \\ \quad + \dfrac{-G}{\sqrt{\mu_X K_{in}}} \left(\sqrt{\dfrac{K_{in}}{K_{out}}} - 1 \right) e^{-\left(\sqrt{\frac{K_{in}}{\mu_X}} (x-x_l) \right)} & \\[3em] \dfrac{G}{\sqrt{\mu_X K_{out}}} \left(\sqrt{\dfrac{K_{out}}{K_{in}}} - 1 \right) \dfrac{\sinh\left(\sqrt{\dfrac{K_{out}}{\mu_X}} (x-x_L) \right)}{\cosh\left(\sqrt{\dfrac{K_{out}}{\mu_X}} (x_r-x_L) \right)} & x_r < x < x_L \end{cases}$$

$$(2.55)$$

Estimates

For our asymptotics, we have that $K_{out} > K_{in}$ and

$$\left(\sqrt{\frac{K_{out}}{K_{in}}} - 1 \right) > 0, \qquad \left(\sqrt{\frac{K_{in}}{K_{out}}} - 1 \right) < 0.$$

We conclude that X and $|X_x|$ are strictly increasing on $[x_0, x_l]$ and strictly decreasing on $[x_r, x_L]$. Therefore we find that the L^∞ estimates of X and X_x can be calculated from the interface region.

From equation (2.53) we can calculate that that the condition $X_x = 0$ for a maximum or minimum is achieved if and only if $A_m \tanh\left(\sqrt{\frac{K_{in}}{\mu_X}} (x - x_l) \right) + B_m = 0$. Since hyperbolic tangent is a one-to-one function, this equation has at most one solution. Therefore, since we know that $X_x(x_l) > 0$ and $X_x(x_r) < 0$, X obtains a maximum in the interface region of the device. Furthermore, since a maximum is obtained, we know that $X_{xx} < 0$ and so $X < \frac{G}{K_{in}}$ and we have a strict estimate $\frac{G}{K_{out}} < X < \frac{G}{K_{in}}$ everywhere.

Using this inequality, we see that $X_{xx} < 0$ for $x_l < x < x_r$, and thus that the maximum values for $|X_x|$ occur at the points x_l and x_r. Since X_x is continuous, we can easily calculate these values. Putting these results together, we have:

$$\|X\|_{L^\infty} < \frac{G}{K_{in}}, \tag{2.56}$$

$$\|X_x\|_{L^\infty} = \frac{G}{\sqrt{\mu_X}} \left(\frac{1}{\sqrt{K_{in}}} - \frac{1}{\sqrt{K_{out}}} \right). \tag{2.57}$$

123

2.A.2 Calculating J_{OC} for Nonsymmetric Devices

For the general non-symmetric case we proceed by combining the two current terms from Equation (2.42) over a common denominator:

$$
\begin{aligned}
J_{tot} &= \frac{\delta_1\eta\,(\delta_2-1)\,F(x_L)}{\delta_1\eta\,(\delta_2-1)+(\delta_1-1)} + \frac{(1-\delta_2)\,F(x_L)}{(\delta_1-1)\eta\delta_2+(\delta_2-1)} \\
&= F(x_L)\frac{\delta_1\eta^2\delta_2(\delta_2-1)(\delta_1-1)+\delta_1\eta(\delta_2-1)^2-\delta_1\eta(\delta_2-1)^2-(\delta_1-1)(\delta_2-1)}{\delta_1\eta^2\delta_2(\delta_2-1)(\delta_1-1)+(\delta_1-1)^2\eta\delta_2+\delta_1\eta(\delta_2-1)^2+(\delta_1-1)(\delta_2-1)} \\
&= F(x_L)\frac{(\delta_1\eta^2\delta_2-1)(\delta_2-1)(\delta_1-1)}{(\delta_1\eta^2\delta_2+1)(\delta_2-1)(\delta_1-1)+(\delta_1-1)^2\eta\delta_2+\delta_1\eta(\delta_2-1)^2}
\end{aligned}
$$

Note that the numerator has three apparent zeros. However, with a constant field approximation, we have a relationship between δ_1 and δ_2. If we define the symmetry ratio of the device, α, by $x_L-x_r=\alpha(x_l-x_0)$, then we have

$$
\delta_2 = e^{-E(x_L-x_r)} = e^{-\alpha E(x_l-x_0)} = \delta_1^\alpha.
$$

Inserting this into the expression for the current above gives:

$$
J_{tot} = F(x_L)\frac{(\delta_1^{1+\alpha}\eta^2-1)(\delta_1^\alpha-1)(\delta_1-1)}{(\delta_1^{1+\alpha}\eta^2+1)(\delta_1^\alpha-1)(\delta_1-1)+(\delta_1-1)^2\eta\delta_1^\alpha+\delta_1\eta(\delta_1^\alpha-1)^2} \tag{2.58}
$$

and we see that both the numerator and denominator are zero for $\delta_1=1$, corresponding to zero electric-field in the device. However, the current still exists in the limit. We find this most easily by considering J_n and J_p separately and applying L'Hopital's rule:

$$
\frac{J_p}{F(x_L)} = \frac{\delta_1\eta(\delta_1^\alpha-1)}{\delta_1\eta(\delta_1^\alpha-1)+(\delta_1-1)} \to \frac{\eta(\delta_1^\alpha-1)+\alpha\eta\delta_1^\alpha}{\eta(\delta_1^\alpha-1)+\alpha\eta\delta^\alpha+1} \to \frac{\alpha\eta}{\alpha\eta+1}
$$

$$
\frac{J_n}{F(x_L)} = \frac{1-\delta_1^\alpha}{(\delta_1-1)\eta\delta_1^\alpha+\delta_1^\alpha-1} \to \frac{-\alpha\delta_1^{\alpha-1}}{\eta\delta_1^\alpha+(\delta_1-1)\alpha\eta\delta_1^{\alpha-1}+\alpha\delta_1^{\alpha-1}} \to \frac{-\alpha}{\eta+\alpha}
$$

putting these together, we have:

$$
J_{E=0} \to F(x_L)\left(\frac{\alpha\eta}{\alpha\eta+1}-\frac{\alpha}{\eta+\alpha}\right) = F(x_L)\frac{\alpha\eta^2-\alpha}{\alpha\eta^2+\alpha^2\eta+\eta+\alpha}
$$

and in the small η limit, we have $J_{E=0}\to -F(x_L)=J_{\mathrm{SS}}$, identically to the symmetric device.

Finally, we can calculate the open circuit voltage by setting the final remaining term in the the numerator of the expression in Equation (2.58) to zero. We have:

$$
\delta_1^{1+\alpha} = \eta^{-2}
$$

By the definitions of δ_1, η, and α, we have

$$e^{V(x_L)-V(x_r)+V(x_l)-V(x_0)} = e^{2(\varphi(x_l)-\varphi(x_r))}$$

Thus we conclude using the definition of φ that

$$V(x_L) - V(x_0) = 2(U(x_l) - U(x_r)) + V(x_l) - V(x_r).$$

Since the field is assumed to be constant, in the small interface limit, $V(x_l) \approx V(x_r)$ and thus using the definition of V_{app} we recover

$$V_{\text{OC}} = 2\Delta U$$

and the results from Section 2.8.2 all continue to hold for a nonsymmetric device.

2.B   Coding Implementation Details

The numerical simulations in this chapter were done through modifications to the hybrid discontinuous Galerkin element code supplied with NETGEN-4.9.14-dev and NGSolve-4.9.14-dev [78]. In this section I will layout the necessary details to recreate the simulations.

2.B.1   Meshing

The meshing for all of the simulations was done with Netgen (NETGEN-4.9.14-dev) [78]. Netgen includes built-in functionality for drawing a closed object using lines or B-splines in 2D. In addition, subdomains of the object can be specified in order to separate the bulk from the interface. This is particularly useful since any coefficients can be defined independently on each of the subdomains. Thus for the simulations in Section 2.6 we need only to define the points and lines necessary to create Figure 2.31. We choose appropriate mesh tolerances to emphasize finer discretization near the interface and then refine uniformly once obtaining, for instance, Figure 2.5 automatically using Netgen.

Creating meshes with more complicated geometries is more difficult. As discussed in Section 2.9, we need an explicit formula for the position of the interface in order to appropriately calculate the necessary quantities in the modeling. For a mesh created with B-splines, this is not a trivial task. Instead, since the mesh will be triangulated
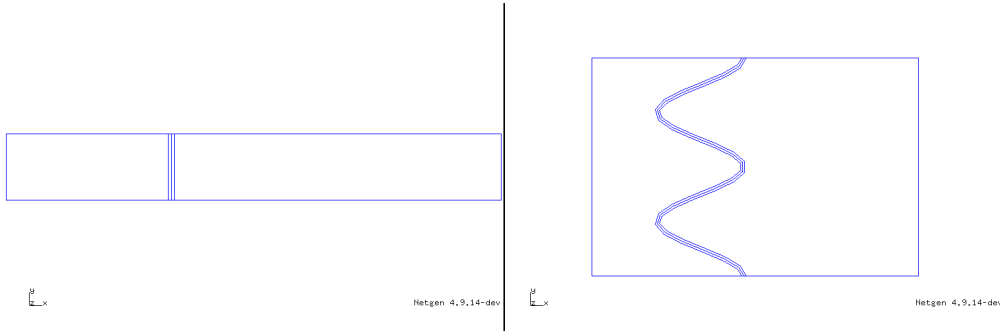
Figure 2.31: Schematic of the figure used by NetGen [78] to create the meshes shown in Section 2.6 and 2.9. The first structure yields the mesh given in Figure 2.5, and the second yields the mesh shown in the simulations of Figure 2.27.

at a specific element-size anyway, we can simply generate the appropriate set of points on our curve and create line-segments between the points with Netgen. By generating the appropriate points for the interface and the boundaries of the interface region, we can create an interface region fitting any functional form we desire. For instance, the sinusoidal boundary in Figure 2.27.

2.B.2 Solving the PDE

The simulations were all done using the hybrid discontinuous Galerkin (HDG) package of NGSolve (NGSolve-4.9.14-dev) [78]. We created inherited versions of the built-in integrators to generate bilinear forms using the nonlinear coefficients of the model. We then developed a solver which used these integrators iteratively, calculating the change in the variables between each step and proceeding until either the desired tolerance was reached or the maximum step count was exceeded.

Modified Integrators

One limitation of Netgen is that it is only able to pass one dependent variable into each integrator. For the model given in in System (2.12) this is not a stringent limitation because each term on the left-hand-side depends on at most one other variable. The convection and diffusion coefficients in the n and p equations depend only on V and the mass terms, $c_r np$ depend only on the other charge carrier. Likewise, the mass term for the exciton equation again depends only on V. In order to implement these new

126

integrators, we can create a new specialized HDG class which inherits all of the methods from the built-in HDG method. Because of the symmetry in the n and p equations, we are only need to create one new class for the charge carriers. Poisson's equation is standard and we can use the usual integrators.

For the charge carrier Laplace and convection integrators, we needed to add the nonlinear mobility depending on the electric field. The integrator calculates the mobility for each integration point (determined by the order of the system). We can loop over the degrees of freedom of the system and calculate the electric field at each integration point using the element vector for V and the known formula for the gradients of the reference element. This needs to be done twice, once for the bulk degrees of freedom and once for the facets. Once we have the value of the field, we can calculate the new value of μ for the bilinear form, completing the modifications.

The mass integrators proceed identically. For the exciton mass integrator we calculate the local field and use it to calculate the correct value of k_d at each integration point. For the charge carriers integrator we instead calculate the local value of the opposite charge carrier from the previous time step and use that (scaled by c_r) for the coefficient of the mass term.

For the reasons discussed in Section 2.9, we calculate U and V_X as the solution of individual PDEs before the first step for 2D simulations. Therefore they are treated by Netgen as dependent variables and converting them into a more useful form is difficult. Rather than trying to calculate derived quantities in the iteration procedure, we instead create two global pointers to the two solutions. Although global variables tend to be a terrible programming technique, we at least minimized the possible issues by never modifying them once they were set.

Iteration Procedure

For the most part, since our problem has been expressed in an appropriate form, Netgen will take care of the solution of each stage automatically. There are a few extra steps necessary to accommodate splitting our bilinear form into two parts, but since the solver will be linear, we can simply add the two forms together. Once again, the Netgen package takes care of the necessary details.

For each iteration stage k:

- Increment step number and set damping parameter

- Loop through all integration points for every element

    - Calculate and store the values of n_{k-1}, p_{k-1}, and $\nabla_x X_{k-1}$

- Calculate V_k using the above values for the right-hand-side

- Loop through all integration points for every element

    - Calculate and store the values of E_k, n_{k-1}, p_{k-1}, and X_{k-1}

    - Calculate and store the forcing terms: $k_d X_{k-1}$, $c'_r n_{k-1} p_{k-1}$

- Calculate and store the error $\|V_k - V_{k-1}\|_{L^2}$ (implemented by Netgen)

- Assemble the left-hand side for n using V_k

- Calculate and store both n_k and the error $\|n_k - n_{k-1}\|_{L^2}$

- Repeat previous two steps for p and X

- Loop through all facets for every element

    - If the facet lies along the boundary x_0

        * Calculate and store the values of n_k, p_k, ∇n_k, ∇p_k, and E_k

        * Add the flux through the facet to a running total of J

- Print the step number and current L^2 errors to a file

- Break the loop if the maximum steps or desired tolerance are met

Print whether the iteration was successful or failed
  - If it was successful, print the current
  - Otherwise print the L^2 error from the final step

Chapter 3

Graphene

This chapter is the result of collaborative work with Clemens Heitzinger and Peter A. Markowich. Most of the material appears in our preprint [11]. Professor Markowich was responsible for the initial direction of the project and Dr. Heitzinger was instrumental in deciding the final form.

3.1 Introduction

In the past few years, graphene has been receiving significant attention as a 2D material [2, 61, 62, 63, 75]. It characterizes a number of interesting electronic properties [59], and much recent work, for example, has focused on creating high frequency electronic components [54, 98]. In contrast to carbon nanotubes, where in-situ fabrication with controlled electronic properties is hard to achieve, graphene sheets can be fabricated relatively easily on substrates, chemically modified, and etched into the desired shape. Amongst other properties, this leads to interesting topological effects [6, 68].

Many of these interesting properties can be understood by considering the electrons as 2D Dirac fermions [2, 63, 68]. Although some work has been published addressing the Dirac equation with electromagnetic fields from a numerical perspective, it has largely focused on the 3D regime, often using a fluid-dynamics framework to speed computation [83, 84] or specifically in the semi-classical and non-relativistic regimes [45]. Further work has been done specifically in 2D, but it has largely focused on illustrating the phenomena associated with quantum mechanics [88] rather than the simulation of graphene as a 2D material.

In this chapter, we present a convergent 2D finite-difference scheme of second order in both space and time for the Dirac equation with a static electric potential and a magnetic field normal to the surface.

The chapter is organized as follows. After a discussion of the Dirac equation and its scaling in Section 3.2, the numerical scheme is described in Section 3.3. This includes a convergence proof based on its consistency and stability, and we further demonstrate how to include a self-consistent model for the electric potential via Gummel iteration. In Section 3.4, several numerical results are discussed. First the numerical scheme and its convergence order is verified in specific cases for which we can derive an explicit solution. We then consider numerical problems that can arise in the numerical approximation of solutions of the Dirac equation. Finally we consider Veselago lenses, beam splitting, and the influence of magnetic potentials and self-consistent fields.

3.2 Model Equations

3.2.1 The Scaled Equation

The Dirac equation is usually written

$$i\hbar\frac{\partial\psi}{\partial t} = (c\bar{\alpha}\cdot\bar{P} + \beta mc^2)\psi, \tag{3.1}$$

where the bar denotes a vector quantity in the spatial dimensions [79]. Here ψ is a wave-vector with four components. There are several possible choices for the α matrices. The usual choice uses the Pauli matrices in the top-right and lower-left corner, but the sign of the second Pauli matrix is not chosen consistently. We choose

$$\alpha_x := \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \qquad \alpha_y := \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix},$$

$$\alpha_z := \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \qquad \beta := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Using minimal coupling, we couple the equation to the magnetic field by considering the kinetic momentum operator

$$P \rightarrow -i\hbar\nabla + \frac{e}{c}A,$$

where A is the magnetic vector potential, and

$$i\hbar\frac{\partial}{\partial t} \rightarrow i\hbar\frac{\partial}{\partial t} - e\Phi,$$

where Φ is the electrical potential. With this substitution, we can explicitly write the equations as

$$\left(i\hbar\frac{\partial}{\partial t} - e\Phi\right)\begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \end{pmatrix} =$$

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\left(-i\hbar c\frac{\partial}{\partial x} + eA_x\right)\begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \end{pmatrix}$$

$$-\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}\left(-i\hbar c\frac{\partial}{\partial y} + eA_y\right)\begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \end{pmatrix}$$

$$-\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}\left(-i\hbar c\frac{\partial}{\partial z} + eA_z\right)\begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \end{pmatrix}$$

$$+mc^2\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \end{pmatrix}.$$

Writing each component individually, we have

$$\frac{\partial \psi^1}{\partial t} = -\left(\frac{\partial \psi^4}{\partial x} + i\frac{eA_x}{mc^2}\psi^4 - i\frac{\partial \psi^4}{\partial y} + \frac{eA_y}{mc^2}\psi^4\right)$$
$$-\left(\frac{\partial \psi^3}{\partial z} + i\frac{eA_z}{mc^2}\psi^3\right) - i\left(1 + \frac{e\Phi}{mc^2}\right)\psi^1, \qquad (3.2a)$$

$$\frac{\partial \psi^2}{\partial t} = -\left(\frac{\partial \psi^3}{\partial x} + i\frac{eA_x}{mc^2}\psi^3 + i\frac{\partial \psi^3}{\partial y} - \frac{eA_y}{mc^2}\psi^3\right)$$
$$+\left(\frac{\partial \psi^4}{\partial z} + i\frac{eA_z}{mc^2}\psi^4\right) - i\left(1 + \frac{e\Phi}{mc^2}\right)\psi^2, \qquad (3.2b)$$

$$\frac{\partial \psi^3}{\partial t} = -\left(\frac{\partial \psi^2}{\partial x} + i\frac{eA_x}{mc^2}\psi^2 - i\frac{\partial \psi^2}{\partial y} + \frac{eA_y}{mc^2}\psi^2\right)$$
$$-\left(\frac{\partial \psi^1}{\partial z} + i\frac{eA_z}{mc^2}\psi^1\right) + i\left(1 - \frac{e\Phi}{mc^2}\right)\psi^3, \qquad (3.2c)$$

$$\frac{\partial \psi^4}{\partial t} = -\left(\frac{\partial \psi^1}{\partial x} + i\frac{eA_x}{mc^2}\psi^1 + i\frac{\partial \psi^1}{\partial y} - \frac{eA_y}{mc^2}\psi^1\right)$$
$$+\left(\frac{\partial \psi^2}{\partial z} + i\frac{eA_z}{mc^2}\psi^2\right) + i\left(1 - \frac{e\Phi}{mc^2}\right)\psi^4. \qquad (3.2d)$$

From a physics perspective, we could eliminate the dimensional parameters by going to a different system of physical units, but this would leave the mass parameter as a variable. Instead, we rescale space and time to eliminate all of the parameters by setting

$$t =: \frac{\hbar}{mc^2}\tilde{t},$$
$$(x, y, z) =: \frac{\hbar}{mc}(\tilde{x}, \tilde{y}, \tilde{z}).$$

3.2.2 Dirac Equation in Two Spatial Dimensions

Since A_x and A_y always occur as coefficients of the same component of the wave vector, we can couple the equations on the xy-plane with a magnetic field in the z-direction. Constraining the particles to move in the plane and setting $B_x = B_y = 0$ yields $A_z = 0$

and we find

$$\frac{\partial \psi^1}{\partial \tilde{t}} = -\left(\frac{\partial \psi^4}{\partial \tilde{x}} + i\frac{eA_x}{mc^2}\psi^4 - i\frac{\partial \psi^4}{\partial \tilde{y}} + \frac{eA_y}{mc^2}\psi^4\right) - i\left(1 + \frac{e\Phi}{mc^2}\right)\psi^1,$$

$$\frac{\partial \psi^2}{\partial \tilde{t}} = -\left(\frac{\partial \psi^3}{\partial \tilde{x}} + i\frac{eA_x}{mc^2}\psi^3 + i\frac{\partial \psi^3}{\partial \tilde{y}} - \frac{eA_y}{mc^2}\psi^3\right) - i\left(1 + \frac{e\Phi}{mc^2}\right)\psi^2,$$

$$\frac{\partial \psi^3}{\partial \tilde{t}} = -\left(\frac{\partial \psi^2}{\partial \tilde{x}} + i\frac{eA_x}{mc^2}\psi^2 - i\frac{\partial \psi^2}{\partial \tilde{y}} + \frac{eA_y}{mc^2}\psi^2\right) + i\left(1 - \frac{e\Phi}{mc^2}\right)\psi^3,$$

$$\frac{\partial \psi^4}{\partial \tilde{t}} = -\left(\frac{\partial \psi^1}{\partial \tilde{x}} + i\frac{eA_x}{mc^2}\psi^1 + i\frac{\partial \psi^1}{\partial \tilde{y}} - \frac{eA_y}{mc^2}\psi^1\right) + i\left(1 - \frac{e\Phi}{mc^2}\right)\psi^4.$$

This couples ψ^1 with ψ^4 and ψ^2 with ψ^3.

We can also consider the standard representation of the 2D Dirac equation, in which we use the 2×2 Pauli matrices σ_k instead of the 4×4 Dirac matrices. This gives the equation

$$(i\hbar\frac{\partial}{\partial t} - e\Phi)U = \sigma_1(-i\hbar c\frac{\partial}{\partial x} + eA_x)U + \sigma_2(-i\hbar c\frac{\partial}{\partial y} + eA_y)U + mc^2\sigma_3 U$$

or, after expanding the matrices,

$$(i\hbar\frac{\partial}{\partial t} - e\Phi)\begin{pmatrix} u^1 \\ u^2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}(-i\hbar c\frac{\partial}{\partial x} + eA_x)\begin{pmatrix} u^1 \\ u^2 \end{pmatrix}$$

$$+ \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}(-i\hbar c\frac{\partial}{\partial y} + eA_y)\begin{pmatrix} u^1 \\ u^2 \end{pmatrix} + mc^2\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} u^1 \\ u^2 \end{pmatrix}.$$

Evaluating the last equation by component and dividing by i gives

$$\hbar\frac{\partial u^1}{\partial t} = -(\hbar c\frac{\partial}{\partial x} + ieA_x)u^2 - (-i\hbar c\frac{\partial}{\partial y} + eA_y)u^2 - i(mc^2 + e\Phi)u^1, \qquad \text{(3.3a)}$$

$$\hbar\frac{\partial u^2}{\partial t} = -(\hbar c\frac{\partial}{\partial x} + ieA_x)u^1 - (i\hbar c\frac{\partial}{\partial y} - eA_y)u^1 + i(mc^2 - e\Phi)u^2. \qquad \text{(3.3b)}$$

Dividing by mc^2 and scaling x, y, and t as before yields exactly the equations for ψ^1 and ψ^4 derived from the 3D Dirac equation constrained to 2 dimensions, i.e., System (3.2). The two formulations, (3.2) and (3.3), are equivalent except that the 4-component spinors include a second particle with the same dynamics up to a reflection in y. This effect is explained in the following in the case of graphene.

3.2.3   Derivation of Dirac Model for Graphene

Using the trigonal symmetry of the graphene lattice, it is well known [92] that the tight-binding model yields the Hamiltonian

$$E(k) = \pm\gamma\sqrt{3 + 2\cos\left(\sqrt{3}k_y a\right) + 4\cos\left(\sqrt{3}k_y a/2\right)\cos\left(3k_x a/2\right)} \qquad (3.4)$$

assuming symmetric energy levels for convenience [59]. Because each atom has one free electron, the negative branch of the square root is usually completely filled, and the minimum energy states will be the minimums of the positive branch. It is easy to verify that six such minimum momenta K exist, oriented at the corners of the hexagonal Brillouin zone of graphene. For wave vectors $k = K + q$, we find

$$E(k) = \frac{3a\gamma}{2}|q| + O(|q|^2).$$

This dispersion relation is identical to the dispersion relation of the 2D Dirac equation with the speed of light, c replaced by the velocity $v_f = 3at/2$. With this motivation, one can derive the Hamiltonian for the system and verify that indeed we obtain the Dirac Hamiltonian [82]. Thus for the low-energy wave vectors sufficiently close to the Dirac points K, we can use the 2D Dirac equation to simulate the behavior of electrons in graphene. Furthermore, depending on which K is chosen, we have time-reversal symmetry which corresponds exactly to the reflection in y illustrated in Section 3.2.2 [59].

Because of the emphasis on the 2D equation for 2D materials such as graphene, we restrict our attention to the 2D Dirac equations and call the unknown two-component wave-vector

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

from now on. Note, however, that the treatment below also holds when considering the coupled system ψ^1 and ψ^4 constrained to the plane as discussed above.

3.2.4   Adding Poisson's Equation

To further implement a consistent potential, we need to expand our system of equations to include an equation for the dependence of the electric potential on u (or ψ). This can be done in several ways, depending on the degree of accuracy required in the potential. The most accurate method would be to include Maxwell's equations for the

134

four-potential, but since the constraint of the particles to 2D is not a-priori relativistically invariant, this is probably not the most sensible approach. With the constraint of the particles, we expect that the magnetic field in the x and y direction should be zero, and thus that the magnetic vector potential term A_z will also be zero. If we assume that this is the case, then the magnetic contribution as a whole from the carriers should be small, since it will only provide a small counter-field according to Lenz's law (at least in the non-relativistic regime). This motivates considering only the electrical potential Φ, which (either in the non-relativistic regime or by choosing the Coulomb Gauge [48]) is modeled by Poisson's equation given the wave-vectors u^1 and u^2, i.e.,

$$-\epsilon\Delta\Phi = -e(|u^1|^2 + |u^2|^2) = -e|u|^2.$$

Note that for the 2D Dirac equation, only electrons appear in the system and the sign on the right-hand side is consistent with this case. In the case of the coupled ψ^1 and ψ^4 terms, the use of the Poisson equation is an approximation, since the electrons and positrons cannot be separated in the Dirac equation.

For graphene, the components are decoupled and only two components, namely u^1 and u^2, are involved for an electron near a single Dirac point. Due to its importance for 2D materials, we consider this case for the remainder of the paper. However, the numerical scheme continues to hold for simultaneously simulating the two coupled spinor systems (ψ^1, ψ^4) and (ψ^2, ψ^3) independently. For this case, the two systems couple implicitly through contributions to the right-hand side of the Poisson equation.

3.2.5   The Main Model Equations

To this point we have not chosen a scaling value for u. Because the continuity equations for u^1 and u^2 are independent with respect to rescaling, our continuity equations are invariant with respect to the scaling

$$u = \lambda\tilde{u}.$$

However, the coefficient of Poisson's equation will depend on λ^2, and thus our system can distinguish concentration-dependent effects.

For notational convenience, we define the rescaled quantities

$$\tilde{A} = \frac{eA}{mc^2},$$

$$V = \frac{e\Phi}{mc^2},$$

$$l^2 = \frac{e^2}{mc^2 \epsilon_0 \epsilon_r},$$

$$\tilde{\lambda}^2 = \lambda^2 l^2,$$

which give the following dimensionless system:

$$\frac{\partial \tilde{u^1}}{\partial \tilde{t}} = -\left(\frac{\partial \tilde{u^2}}{\partial \tilde{x}} + i\tilde{A}_x \tilde{u^2} - i\frac{\partial \tilde{u^2}}{\partial \tilde{y}} + \tilde{A}_y \tilde{u^2} \right) - i(1+V)\tilde{u^1}, \tag{3.5a}$$

$$\frac{\partial \tilde{u^2}}{\partial \tilde{t}} = -\left(\frac{\partial \tilde{u^1}}{\partial \tilde{x}} + i\tilde{A}_x \tilde{u^1} + i\frac{\partial \tilde{u^1}}{\partial \tilde{y}} - \tilde{A}_y \tilde{u^1} \right) + i(1-V)\tilde{u^2}, \tag{3.5b}$$

$$\Delta V = -\tilde{\lambda}^2 |\tilde{u}|^2. \tag{3.5c}$$

We use System (3.5) for $\tilde{u^1}$, $\tilde{u^2}$ and $\tilde{V}$ as the main model equations in the remainder of the chapter, where we will also suppress the $\sim$ notation for notational convenience.

3.3 The Numerical Scheme

In Section 3.3.1, a finite-difference scheme for System (3.5) is developed for the case of constant potentials and its convergence is shown. In Section 3.3.2, the numerical approximation of the self-consistent Dirac-Poisson system is discussed, and scaling suitable for graphene based devices is discussed finally in Section 3.3.3.

3.3.1 The Finite-Difference Scheme

We split the Dirac equations into two parts and evaluate using a Strang-splitting method [85]. The second term in both equations is purely rotational and explicitly solvable: the solution to $\frac{\partial f}{\partial t} = ikf$ is simply the exponential $f(t) = f(0)e^{ikt}$. We thus define the operator α by

$$\alpha_{\Delta t}(u^1, u^2)^\top = (u^1 e^{-i(1+V)\Delta t}, u^2 e^{i(1-V)\Delta t})^\top \tag{3.6}$$

as one term in our Strang splitting. The other term must account for the coupling of the equations through the magnetic vector potential and the spatial derivative terms.

136

We implement this term using the Crank-Nicolson method, using centered differences for the spatial derivatives

$$\frac{u^1_{k+1}(x_i, y_j) - u^1_k(x_i, y_j)}{\Delta t} = -\frac{1}{2}(iA_x + A_y)\left(u^2_k(x_i, y_j) + u^2_{k+1}(x_i, y_i)\right)$$

$$-\frac{1}{4\Delta x}\left(u^2_{k+1}(x_{i+1}, y_j) - u^2_{k+1}(x_{i-1}, y_j) + u^2_k(x_{i+1}, y_j) - u^2_k(x_{i-1}, y_j)\right)$$

$$+\frac{i}{4\Delta y}\left(u^2_{k+1}(x_i, y_{j+1}) - u^2_{k+1}(x_i, y_{j-1}) + u^2_k(x_i, y_{j+1}) - u^2_k(x_i, y_{j-1})\right), \tag{3.7}$$

$$\frac{u^2_{k+1}(x_i, y_i) - u^2_k(x_i, y_i)}{\Delta t} = -\frac{1}{2}(iA_x - A_y)\left(u^1_k(x_i, y_i) + u^1_{k+1}(x_i, y_i)\right)$$

$$-\frac{1}{4\Delta x}\left(u^1_{k+1}(x_{i+1}, y_j) - u^1_{k+1}(x_{i-1}, y_j) + u^1_k(x_{i+1}, y_j) - u^1_k(x_{i-1}, y_j)\right)$$

$$-\frac{i}{4\Delta y}\left(u^1_{k+1}(x_i, y_{j+1}) - u^1_{k+1}(x_i, y_{j-1}) + u^1_k(x_i, y_{j+1}) - u^1_k(x_i, y_{j-1})\right). \tag{3.8}$$

Note that we take $\Delta x = \Delta y =: h$ and define

$$\beta_{\Delta t, h}(u^1, u^2) := (u^1_{k+1}, u^2_{k+1}). \tag{3.9}$$

In addition to being second order convergent in space and time, we also prove below that this selection is also free of artificial dissipation.

We can also write the equations in matrix form, using the matrices M^{14} and M^{41} defined by

$$M^{14}u^2_k(x_i, y_j) = -\frac{1}{2}(iA_x + A_y)u^2_k(x_i, y_j)$$

$$-\frac{1}{4\Delta x}\left(u^2_k(x_{i+1}, y_j) - u^2_k(x_{i-1}, y_j)\right)$$

$$+\frac{i}{4\Delta y}\left(u^2_k(x_i, y_{j+1}) - u^2_k(x_i, y_{j-1})\right),$$

$$M^{41}u^1_k(x_i, y_j) = -\frac{1}{2}(iA_x - A_y)u^1_k(x_i, y_j)$$

$$-\frac{1}{4\Delta x}\left(u^1_k(x_{i+1}, y_j) - u^1_k(x_{i-1}, y_j)\right)$$

$$-\frac{i}{4\Delta y}\left(u^1_k(x_i, y_{j+1}) - u^1_k(x_i, y_{j-1})\right)$$

so that Equations (3.7) and (3.8) can be written as

$$u^1_{k+1}(x_i, y_j) - \Delta t M^{14}u^2_{k+1}(x_i, y_j) = u^1_k(x_i, y_j) + \Delta t M^{14}u^2_k(x_i, y_j),$$

$$u^2_{k+1}(x_i, y_j) - \Delta t M^{41}u^1_{k+1}(x_i, y_j) = u^2_k(x_i, y_j) + \Delta t M^{41}u^1_k(x_i, y_j).$$

We thus have the matrix equation

$$\begin{pmatrix} 1 & -\Delta t M^{14} \\ -\Delta t M^{41} & 1 \end{pmatrix} \begin{pmatrix} u_{k+1}^1 \\ u_{k+1}^2 \end{pmatrix} = \begin{pmatrix} 1 & \Delta t M^{14} \\ \Delta t M^{41} & 1 \end{pmatrix} \begin{pmatrix} u_k^1 \\ u_k^2 \end{pmatrix}. \tag{3.10}$$

for (3.7) and (3.8). Note that M^{14} and M^{41} are matrices which act on the discretized components of u_k^1 and u_k^2.

Our method thus consists of evaluating $U = (u^1, u^2)^\top$ and then calculating the next value U_{k+1} by

$$U_{k+1} = \alpha_{\Delta t/2} \circ \beta_{\Delta t,h} \circ \alpha_{\Delta t/2} U_k. \tag{3.11}$$

Note that the determinant of the left matrix in (3.10) is $1 + C\Delta t^2$ for a constant C independent of Δt. Thus for Δt small enough, the matrix will be invertible and the method is given explicitly by

$$U_{k+1} = \begin{pmatrix} e^{-i(1+V)\frac{\Delta t}{2}} & 0 \\ 0 & e^{i(1-V)\frac{\Delta t}{2}} \end{pmatrix} \begin{pmatrix} 1 & -\Delta t M^{14} \\ -\Delta t M^{41} & 1 \end{pmatrix}^{-1} \cdot$$

$$\cdot \begin{pmatrix} 1 & \Delta t M^{14} \\ \Delta t M^{41} & 1 \end{pmatrix} \begin{pmatrix} e^{-i(1+V)\frac{\Delta t}{2}} & 0 \\ 0 & e^{i(1-V)\frac{\Delta t}{2}} \end{pmatrix} U_k. \tag{3.12}$$

Consistency of the Crank-Nicolson Scheme

In order to establish convergence of the scheme, we start with its consistency.

Proposition 1. *The Crank-Nicolson scheme defined by (3.7) and (3.8) is second order in space and time and thus consistent under the assumption that A is time-independent.*

Proof. Beginning from Equations (3.7) and (3.8) we have

$$\frac{\Delta t \partial_t u^1 + \frac{1}{2}\Delta t^2 \partial_{tt} u^1}{\Delta t} + \frac{O(\Delta t^3)}{\Delta t} = -\frac{1}{2}(iA_x + A_y)\left(2u^2 + \Delta t \partial_t u^2 + O(\Delta t^2)\right)$$

$$- \frac{1}{4\Delta x}\left(4\Delta x \partial_x u^2 + 2\Delta x \Delta t \partial_{xt} u^2 + O(\Delta x^3, \Delta x \Delta t^2)\right)$$

$$+ \frac{i}{4\Delta y}\left(4\Delta y \partial_y u^2 + 2\Delta y \Delta t \partial_{yt} u^2 + O(\Delta y^3, \Delta y \Delta t^2)\right),$$

$$\frac{\Delta t \partial_t u^2 + \frac{1}{2}\Delta t^2 \partial_t t u^2}{\Delta t} + \frac{O(\Delta t^3)}{\Delta t} = -\frac{1}{2}(iA_x - A_y)\left(2u^1 + \Delta t \partial_t u^1 + O(\Delta t^2)\right)$$

$$- \frac{1}{4\Delta x}\left(4\Delta x \partial_x u^1 + 2\Delta x \Delta t \partial_{xt} u^1 + O(\Delta x^3, \Delta x \Delta t^2)\right)$$

$$- \frac{i}{4\Delta y}\left(4\Delta y \partial_y u^1 + 2\Delta y \Delta t \partial_{yt} u^1 + O(\Delta y^3, \Delta y \Delta t^2)\right),$$

138

where the subscripts $u_k^1(x_i, y_i)$ and $u_k^2(x_i, y_i)$ are implied.

For the terms proportional to Δt^0 we can use the original system (3.5) to cancel all of the terms. The terms proportional to Δt satisfy the time derivative of (3.5) and similarly cancel entirely. For the next order terms, we have

$$\frac{1}{6}\Delta t^2 \partial_{tt} u^1 + O(\Delta t^3) = -\frac{1}{2}(iA_x + A_y)\left(\frac{1}{2}\Delta t^2 \partial_{tt} u^2 + O(\Delta t^3)\right)$$
$$- \frac{1}{4}\left(\Delta t^2 \partial_{xtt} u^2 + \frac{2}{3}\Delta x^2 \partial_{xxx} u^2 + O(\Delta x^4, \Delta x^2 \Delta t^2, \Delta t^3)\right)$$
$$+ \frac{i}{4}\left(\Delta t^2 \partial_{ytt} u^2 + \frac{2}{3}\Delta y^2 \partial_{yyy} u^2 + O(\Delta y^4, \Delta y^2 \Delta t^2, \Delta t^3)\right),$$

$$\frac{1}{6}\Delta t^2 \partial_{tt} u^1 + O(\Delta t^3) = -\frac{1}{2}(iA_x - A_y)\left(\frac{1}{2}\Delta t^2 \partial_{tt} u^2 + O(\Delta t^3)\right)$$
$$- \frac{1}{4}\left(\Delta t^2 \partial_{xtt} u^1 + \frac{2}{3}\Delta x^2 \partial_{xxx} u^1 + O(\Delta x^4, \Delta x^2 \Delta t^2, \Delta t^3)\right)$$
$$- \frac{i}{4}\left(\Delta t^2 \partial_{ytt} u^1 + \frac{2}{3}\Delta y^2 \partial_{yyy} u^1 + O(\Delta y^4, \Delta y^2 \Delta t^2, \Delta t^3)\right)$$

so that the local truncation errors in u^1 and u^2 are

$$\tau_1(x, t) = -\frac{\Delta t^2}{12}\partial_{tt} u^1 - \frac{\Delta x^2}{6}\partial_{xxx} u^2 + \frac{i\Delta y^2}{6}\partial_{yyy} u^2,$$
$$\tau_2(x, t) = -\frac{\Delta t^2}{12}\partial_{tt} u^2 - \frac{\Delta x^2}{6}\partial_{xxx} u^1 - \frac{i\Delta y^2}{6}\partial_{yyy} u^1,$$

respectively. Thus we see that our Crank-Nicolson scheme is second order accurate in both space and time and is thus consistent. $\qquad\square$

Von Neumann Stability of the Crank-Nicolson Scheme

The second step is to establish stability.

Proposition 2. *The Crank-Nicolson scheme defined by (3.7) and (3.8) is von Neumann stable and non-dissipative.*

Proof. We must prove that the eigenvalues λ_g of the matrix $g(\xi_x, \xi_y)$ defined by the ansatz

$$U_{k+1}(x_p, y_q) = g(\xi)U_k(x_p, y_q) \quad \forall U_k(x_p, y_q) = (C_1, C_2)^\top e^{i(h\xi_x p + h\xi_y q)}$$

satisfy the bound $|\lambda_g| \leq 1$ to show von Neumann stability and furthermore that $|\lambda_g| = 1$ to show that the scheme is non-dissipative.

We want to rewrite (3.10) as

$$U_{k+1} = \begin{pmatrix} 1 & -\Delta t M^{14} \\ -\Delta t M^{41} & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & \Delta t M^{14} \\ \Delta t M^{41} & 1 \end{pmatrix} U_k,$$

but we must show that the first matrix is invertible.

With our given ansatz, we can write the action of M by

$$
\begin{aligned}
M_{14} e^{i(h\xi_x p + h\xi_y q)} &= -\frac{1}{2}(iA_x + A_y)e^{i(h\xi_x p + h\xi_y q)} \\
&\quad - \frac{1}{4\Delta x}\left(e^{i(h\xi_x(p+1)+h\xi_y q)} - e^{i(h\xi_x(p-1)+h\xi_y q)}\right) \\
&\quad + \frac{i}{4\Delta y}\left(e^{i(h\xi_x p + h\xi_y(q+1))} - e^{i(h\xi_x p + h\xi_y(q-1))}\right) \\
&= e^{i(h\xi_x p + h\xi_y q)}\left(-\frac{1}{2}(iA_x + A_y) - \frac{1}{4\Delta x}\left(e^{ih\xi_x} - e^{-ih\xi_x}\right) + \frac{i}{4\Delta y}\left(e^{ih\xi_y} - e^{-ih\xi_y}\right)\right) \\
&= e^{i(h\xi_x p + h\xi_y q)}\left(-\frac{1}{2}(iA_x + A_y) - \frac{i}{2\Delta x}\sin(h\xi_x) - \frac{1}{2\Delta y}\sin(h\xi_y)\right)
\end{aligned}
$$

and similarly

$$M_{41} e^{i(h\xi_x p + h\xi_y q)} = e^{i(h\xi_x p + h\xi_y q)}\left(-\frac{1}{2}(iA_x - A_y) - \frac{i}{2\Delta x}\sin(h\xi_x) + \frac{1}{2\Delta y}\sin(h\xi_y)\right).$$

We can thus equate M^{14} and M^{41} with their eigenvalues M_λ^{14} and M_λ^{41} (depending, of course, on $h\xi_x$ and $h\xi_y$) and explicitly calculate

$$\begin{pmatrix} 1 & -\Delta t M_\lambda^{14} \\ -\Delta t M_\lambda^{41} & 1 \end{pmatrix}^{-1} = \frac{1}{1 - \Delta t^2 M_\lambda^{14} M_\lambda^{41}}\begin{pmatrix} 1 & \Delta t M_\lambda^{14} \\ \Delta t M_\lambda^{41} & 1 \end{pmatrix}$$

for all sufficiently small Δt (such that $\Delta t^2 M_\lambda^{14} M_\lambda^{41} < 1$).

We can thus directly calculate

$$g(\xi_x, \xi_y) = \frac{1}{1 - \Delta t^2 M_\lambda^{14} M_\lambda^{41}}\begin{pmatrix} 1 + \Delta t^2 M_\lambda^{14} M_\lambda^{41} & 2\Delta t M_\lambda^{14} \\ 2\Delta t M_\lambda^{41} & 1 + \Delta t^2 M_\lambda^{14} M_\lambda^{41} \end{pmatrix}. \qquad (3.13)$$

The eigenvalues of this matrix are

$$\lambda_g = \frac{1 + \Delta t^2 M_\lambda^{14} M_\lambda^{41} \pm 2\Delta t \sqrt{M_\lambda^{14} M_\lambda^{41}}}{\Delta t^2 M_\lambda^{14} M_\lambda^{41} - 1} = \frac{(1 \pm \Delta t \sqrt{M_\lambda^{14} M_\lambda^{41}})^2}{\Delta t^2 M_\lambda^{14} M_\lambda^{41} - 1},$$

which give us the spectral radius of $g(\xi_x, \xi_y)$ and thus $|g|$.

140

We therefore calculate the product

$$M_\lambda^{14} M_\lambda^{41} = \left(-\frac{1}{2}(iA_x + A_y) - \frac{i}{2\Delta x} \sin h\xi_x - \frac{1}{2\Delta y} \sin h\xi_y \right) \cdot$$

$$\cdot \left(-\frac{1}{2}(iA_x - A_y) - \frac{i}{2\Delta x} \sin h\xi_x + \frac{1}{2\Delta y} \sin h\xi_y \right)$$

$$= -\frac{1}{4}(A_x^2 + A_y^2) - \frac{A_x}{2\Delta x} \sin h\xi_x - \frac{A_y}{2\Delta y} \sin h\xi_y - \frac{1}{4\Delta x^2} \sin h\xi_x{}^2 - \frac{1}{4\Delta y^2} \sin h\xi_y{}^2$$

$$= -\left(\frac{A_x}{2} - \frac{\sin h\xi_x}{2\Delta x} \right)^2 - \left(\frac{A_y}{2} - \frac{\sin h\xi_y}{2\Delta y} \right)^2 \le 0.$$

Since this term is real and non-positive, we have $\Delta t \sqrt{M_\lambda^{14} M_\lambda^{41}} = ib$ for some $b \in \mathbb{R}$. Thus the denominator terms are all strictly negative regardless of the size of Δt (eliminating the smallness assumption for the inversion of the first matrix). In addition, we have

$$|\lambda_g| = \left| \frac{(1 \pm ib)^2}{-b^2 - 1} \right| = \frac{|1 \pm ib|^2}{1 + b^2} = \frac{1 + b^2}{1 + b^2} = 1$$

and our Crank-Nicolson scheme is von Neumann stable for all Δt, Δx, and Δy. $\qquad \square$

Proposition 3. *For a constant magnetic vector potential A, the Crank-Nicolson scheme defined by equations (3.7) and (3.8) is convergent.*

Proof. This statement follows immediately from Propositions 1 and 2. $\qquad \square$

Consistency of the Splitting Scheme

Proposition 4. *The splitting scheme given by (3.11) is consistent for V and A constant in time and furthermore is second order in space and time.*

Proof. To determine the consistency of the splitting scheme, we follow the same procedure as Section 3.3.1, but must rotate the state vectors both before and after the Crank-Nicolson step. We define the factors

$$g^1 := e^{-\frac{i\Delta t}{2}(1+V)} = \left(1 - \frac{i\Delta t}{2}(1+V) - \frac{\Delta t^2}{8}(1+V)^2 + O(\Delta t^3) \right),$$

$$\frac{1}{g^1} = e^{\frac{i\Delta t}{2}(1+V)} = \left(1 + \frac{i\Delta t}{2}(1+V) - \frac{\Delta t^2}{8}(1+V)^2 + O(\Delta t^3) \right),$$

$$g^4 := e^{\frac{i\Delta t}{2}(1-V)} = \left(1 + \frac{i\Delta t}{2}(1-V) - \frac{\Delta t^2}{8}(1-V)^2 + O(\Delta t^3) \right),$$

$$\frac{1}{g^4} = e^{-\frac{i\Delta t}{2}(1-V)} = \left(1 - \frac{i\Delta t}{2}(1-V) - \frac{\Delta t^2}{8}(1-V)^2 + O(\Delta t^3) \right)$$

for simplicity. We can then write $u_{k+1}/g = \beta_{\Delta t,h}gu_k$, which follows since g is non-zero. Since the Crank-Nicolson scheme solves $u_{k+1} = \beta_{\Delta t,h}u_k$, we can use Equations (3.7) and (3.8) with the replacement $u_k \to gu_k$ and $u_{k+1} \to u_{k+1}/g$ to find

$$\frac{u_{k+1}^1(x_i, y_j)/g^1 - g^1 u_k^1(x_i, y_j)}{\Delta t} =$$
$$= -\frac{1}{2}(iA_x + A_y)\left(g^4 u_k^2(x_i, y_j) + u_{k+1}^2(x_i, y_i)/g^4\right)$$
$$- \frac{1}{4\Delta x}\left(u_{k+1}^2(x_{i+1}, y_j)/g^4 - u_{k+1}^2(x_{i-1}, y_j)/g^4 + g^4 u_k^2(x_{i+1}, y_j) - g^4 u_k^2(x_{i-1}, y_j)\right)$$
$$+ \frac{i}{4\Delta y}\left(u_{k+1}^2(x_i, y_{j+1})/g^4 - u_{k+1}^2(x_i, y_{j-1})/g^4 + g^4 u_k^2(x_i, y_{j+1}) - g^4 u_k^2(x_i, y_{j-1})\right),$$

$$\frac{u_{k+1}^2(x_i, y_i)/g^4 - g^4 u_k^2(x_i, y_i)}{\Delta t} =$$
$$= -\frac{1}{2}(iA_x - A_y)\left(g^1 u_k^1(x_i, y_i) + u_{k+1}^1(x_i, y_i)/g^1\right)$$
$$- \frac{1}{4\Delta x}\left(u_{k+1}^1(x_{i+1}, y_j)/g^1 - u_{k+1}^1(x_{i-1}, y_j)/g^1 + g^1 u_k^1(x_{i+1}, y_j) - g^1 u_k^1(x_{i-1}, y_j)\right)$$
$$- \frac{i}{4\Delta y}\left(u_{k+1}^1(x_i, y_{j+1})/g^1 - u_{k+1}^1(x_i, y_{j-1})/g^1 + g^1 u_k^1(x_i, y_{j+1}) - g^1 u_k^1(x_i, y_{j-1})\right).$$

The order $O(\Delta t^0)$ terms are

$$i(1 + V)u^1 + \partial_t u^1 = -\frac{1}{2}(iA_x + A_y)\left(2u^2\right)$$
$$- \frac{1}{4}\left(4\partial_x u^2 + O(\Delta x^2)\right) + \frac{i}{4}\left(4\partial_y u^2 + O(\Delta y^2)\right)$$
$$-i(1 - V)u^2 + \partial_t u^2 = -\frac{1}{2}(iA_x - A_y)\left(2u^1\right)$$
$$- \frac{1}{4}\left(4\partial_x u^1 + O(\Delta x^2)\right) - \frac{i}{4}\left(4\partial_y u^1 + O(\Delta y^2)\right),$$

which are both satisfied to order $O(\Delta x^2) + O(\Delta y^2)$ using system (3.5).

The order $O(\Delta t^1)$ terms are

$$\frac{\Delta t}{2}\partial_{tt}u^1 + \frac{\Delta t}{2}i\Delta t(1 + V)\partial_t u^1 = -\frac{\Delta t}{2}(iA_x + A_y)\left(\partial_t u^2\right)$$
$$- \frac{\Delta t}{2}\left(\partial_{xt}u^2 + O(\Delta x^2)\right) + \frac{\Delta t}{2}\left(i\partial_{yt}u^2 + O(\Delta y^2)\right)$$
$$\frac{\Delta t}{2}\partial_t t u^2 + \frac{\Delta t}{2}i\Delta t(1 + V)\partial_t u^2 = -\frac{\Delta t}{2}(iA_x - A_y)\left(\partial_t u^1\right)$$
$$- \frac{\Delta t}{2}\left(\partial_{xt}u^1 + O(\Delta x^2)\right) - \frac{\Delta t}{2}\left(i\partial_{yt}u^1 + O(\Delta y^2)\right).$$

Eliminating the factor of $\Delta t/2$, this is the time derivative of system (3.5) and thus (again assuming A and V are constant in time) satisfied to order $O(\Delta x^2) + O(\Delta y^2)$. As in the

Crank-Nicolson case, the $O(\Delta t^2)$ terms do not cancel, and the method is still locally second-order accurate in both space and time. $\square$

Von Neumann Stability of the Strang Splitting Scheme

Proposition 5. *The Strang splitting scheme (3.11) is von Neumann Stable and non-dissipative.*

Proof. The proof follows the same strategy as the proof of Proposition 2. By combining equations (3.12) and (3.13), we can directly calculate $g(\xi_x, \xi_y)$ for the splitting scheme, yielding

$$g(\xi_x,\xi_y) = \frac{1}{1+b^2} \begin{pmatrix} e^{-i(1+V)\frac{\Delta t}{2}} & 0 \\ 0 & e^{i(1-V)\frac{\Delta t}{2}} \end{pmatrix} \begin{pmatrix} 1-b^2 & 2\Delta t M_\lambda^{14} \\ 2\Delta t M_\lambda^{41} & 1-b^2 \end{pmatrix} \begin{pmatrix} e^{-i(1+V)\frac{\Delta t}{2}} & 0 \\ 0 & e^{i(1-V)\frac{\Delta t}{2}} \end{pmatrix}$$

$$= \frac{1}{1+b^2} \begin{pmatrix} (1-b^2)e^{-i(1+V)\frac{\Delta t}{2}} & 2\Delta t M_\lambda^{14} e^{-i(1+V)\frac{\Delta t}{2}} \\ 2\Delta t M_\lambda^{41} e^{i(1-V)\frac{\Delta t}{2}} & (1-b^2)e^{i(1-V)\frac{\Delta t}{2}} \end{pmatrix} \begin{pmatrix} e^{-i(1+V)\frac{\Delta t}{2}} & 0 \\ 0 & e^{i(1-V)\frac{\Delta t}{2}} \end{pmatrix}$$

$$= \frac{1}{1+b^2} \begin{pmatrix} (1-b^2)e^{-i(1+V)\Delta t} & 2\Delta t M_\lambda^{14} e^{-iV\Delta t} \\ 2\Delta t M_\lambda^{41} e^{-iV\Delta t} & (1-b^2)e^{i(1-V)\Delta t} \end{pmatrix}$$

$$= \frac{e^{-iV\Delta t}}{1+b^2} \begin{pmatrix} (1-b^2)e^{-i\Delta t} & 2\Delta t M_\lambda^{14} \\ 2\Delta t M_\lambda^{41} & (1-b^2)e^{i\Delta t} \end{pmatrix},$$

where we have once again used the negativity of the product $M_\lambda^{14} M_\lambda^{41}$ resulting in the constant $b \in \mathbb{R}$. We can again calculate the eigenvalues of this matrix explicitly to find

$$\lambda_g = \frac{e^{-iV\Delta t}}{1+b^2} \left((1-b^2)\cos\Delta t \pm i\sqrt{4b^2 + (1-b^2)^2 \sin\Delta t^2} \right)$$

and therefore

$$|\lambda_g|^2 = \frac{1}{(1+b^2)^2} \left((1-b^2)^2 \cos\Delta t^2 + 4b^2 + (1-b^2)^2 \sin\Delta t^2 \right)$$

$$= \frac{1}{(1+b^2)^2} \left((1-b^2)^2 + 4b^2 \right) = \frac{1}{(1+b^2)^2} \left(1 + 2b^2 + b^4 \right) = 1.$$

Hence our full splitting scheme is also von Neumann stable for all $\Delta x, \Delta y$, and Δt. $\square$

If we take the potentials A and V to be constant, then consistency and von Neumann stability guarantee convergence, and our numerical scheme for the 2D Dirac equation converges.

Theorem 4. *The splitting scheme given by (3.11) as implemented in Equation (3.12) is second order in space and time and convergent for A and V constant in time.*

Proof. This statement follows immediately from Propositions 4 and 5. □

3.3.2 Adding a Consistent Potential

To solve Poisson's equation, we use the standard 19-point stencil with modified right-hand side [33, 93] (taken directly from (14) in the latter):

$$-8(1 + \alpha^2 + \beta^2)V(x_i, y_i, z_i) + (4 - \alpha^2 - \beta^2)\left(V(x_{i+1}, y_i, z_i) + V(x_{i-1}, y_i, z_i)\right)$$

$$+(4\alpha^2 - 1 - \beta^2)\left(V(x_i, y_{i+1}, z_i) + V(x_i, y_{i-1}, z_i)\right) + (4\beta^2 - 1 - \alpha^2)\left(V(x_i, y_i, z_{i+1}) + V(x_i, y_i, z_{i-1})\right)$$

$$+\frac{1}{2}(1 + \alpha^2)\left(V(x_{i+1}, y_{i+1}, z_i) + V(x_{i+1}, y_{i-1}, z_i) + V(x_{i-1}, y_{i+1}, z_i) + V(x_{i-1}, y_{i-1}, z_i)\right)$$

$$+\frac{1}{2}(1 + \beta^2)\left(V(x_{i+1}, y_i, z_{i+1}) + V(x_{i+1}, y_i, z_{i-1}) + V(x_{i-1}, y_i, z_{i+1}) + V(x_{i-1}, y_i, z_{i-1})\right)$$

$$+\frac{1}{2}(\alpha^2 + \beta^2)\left(V(x_i, y_{i+1}, z_{i+1}) + V(x_i, y_{i+1}, z_{i-1}) + V(x_i, y_{i-1}, z_{i+1}) + V(x_i, y_{i-1}, z_{i-1})\right)$$

$$= \frac{\Delta x^2}{2}\left(6\rho(x_i, y_i, z_i) + \rho(x_{i+1}, y_i, z_i) + \rho(x_{i-1}, y_i, z_i) + \rho(x_i, y_{i+1}, z_i)\right)$$

$$+ \frac{\Delta x^2}{2}\left(\rho(x_i, y_{i-1}, z_i) + \rho(x_i, y_i, z_{i+1}) + \rho(x_i, y_i, z_{i-1})\right)$$

where

$$\alpha = \frac{\Delta x}{\Delta y}, \qquad \beta = \frac{\Delta x}{\Delta z},$$

and the local error is $O(h^5)$. We call the solution to this equation γ, i.e., $V = \gamma(\rho)$.

Note that the stencil is still compact, consisting only of grid points which are immediately adjacent to the center point (horizontally, vertically, and diagonally). Although the rest of our scheme is only second order, using the fourth order scheme for V does not increase the computational complexity of the step and was thus deemed to be a reasonable choice.

Since Poisson's equation is not time-dependent, we cannot use Strang-type splitting to evaluate it on half-time steps before and after the splitting scheme from the previous section. The standard method from semiconductor theory is to solve Poisson's equation at the beginning of each time step and then to use this new potential to solve the

transport equation [39]. Therefore our self-consistent numerical method is

$$V_{k+1} = \gamma(\rho_k),$$
$$(u_{k+1}^1, u_{k+1}^2) = \alpha_{\Delta t/2} \circ \beta_{\Delta t,h} \circ \alpha_{\Delta t/2}(u_k^1, u_k^2, V_{k+1}),$$
$$\rho_{k+1} = -\lambda^2 |u_{k+1}|^2,$$

where initial wave-vectors u_0^1 and u_0^2 must be specified. (Again, note that for the 2D Dirac equation u^2 and u^3 do not appear. For the 3D analog, this would be the equivalent of setting $u_0^2 = u_0^3 = 0$.)

3.3.3 Scaling for Graphene Devices

Regarding the simulation of graphene based devices, the equations arising for scalings different from those in Section 3.2 are advantageous. We can scale using characteristic values from a graphene based device. In order to obtain an equivalent non-dimensionalized system, the ratio of the distance and time scalings, i.e., L/T, must remain equal to $c = v_f$. For a device with a characteristic size $L = 1\mu m$, this corresponds to a time scaling of $T = 1ps$. In graphene, the particles are massless and hence the mass term disappears, so the scaling term, l^2 for the potential is $\frac{e}{T\hbar} \approx 1500$.

3.4 Numerical Results

We present and discuss several numerical results obtained from our implementation of the above numerical scheme in MATLAB. The first results are shown to verify the FD scheme and its second-order convergence in simple cases with $V = A_x = A_y = 0$ where 1D explicit solutions are known. Then we proceed by demonstrating the capabilities of our numerical scheme by applying it to situations with applied electrostatic and magnetic potentials such as a Veselago lens and a beam splitter.

The first numerical results in this section are performed on a square 100×100 grid with the scaling parameters introduced in Section 3.2 and a simulation domain of scaled size 1×1 so that $h = \Delta x = \Delta y = 0.01$; the time step is $k = 0.01$ as well. We furthermore take periodic boundary conditions in the x and y directions for numerical simplicity. After demonstrating several test cases, we begin altering the grid size to investigate the discretization errors and to better implement several physical examples.

3.4.1 Wave Equation Results

We first find a suitable wave equation with explicit solutions to compare the numerical results with. In the case $m = V = A_x = A_y = 0$, the Dirac equation reduces to

$$\frac{\partial u^1}{\partial t} = -\frac{\partial u^2}{\partial x} + i\frac{\partial u^2}{\partial y},$$

$$\frac{\partial u^2}{\partial t} = -\frac{\partial u^1}{\partial x} - i\frac{\partial u^1}{\partial y}$$

and therefore

$$\frac{\partial^2 u^1}{\partial t^2} = \left(-\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\left(-\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) u^1 = \frac{\partial^2 u^1}{\partial x^2} + \frac{\partial^2 u^1}{\partial y^2} = \Delta u^1$$

and analogously $\frac{\partial^2 u^2}{\partial t^2} = \Delta u^2$ holds as well.

However, for the charge density $\rho = |u^1|^2 + |u^2|^2$, we have

$$\begin{aligned}
\frac{\partial^2 \rho}{\partial t^2} =& \frac{\partial}{\partial t}\left(u^1\frac{\partial \bar{u^1}}{\partial t} + \bar{u^1}\frac{\partial u^1}{\partial t} + \bar{u^2}\frac{\partial u^2}{\partial t} + u^2\frac{\partial \bar{u^2}}{\partial t}\right)\\
=& u^1\frac{\partial^2 \bar{u^1}}{\partial t^2} + 2\frac{\partial u^1}{\partial t}\frac{\partial \bar{u^1}}{\partial t} + \bar{u^1}\frac{\partial^2 u^1}{\partial t^2} + \bar{u^2}\frac{\partial^2 u^2}{\partial t^2} + 2\frac{\partial \bar{u^2}}{\partial t}\frac{\partial u^2}{\partial t} + u^2\frac{\partial^2 \bar{u^2}}{\partial t^2}\\
=& u^1\Delta\bar{u^1} + 2\left(-\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) u^2 \left(-\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\bar{u^2} + \bar{u^1}\Delta u^1 +\\
&+ \bar{u^2}\Delta u^2 + 2\left(-\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\bar{u^1}\left(-\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) u^1 + u^2\Delta\bar{u^2}\\
=& \Delta\rho + i\frac{\partial u^2}{\partial x}\frac{\partial \bar{u^2}}{\partial y} - i\frac{\partial \bar{u^2}}{\partial x}\frac{\partial u^2}{\partial y} - i\frac{\partial u^1}{\partial x}\frac{\partial \bar{u^1}}{\partial y} + i\frac{\partial \bar{u^1}}{\partial x}\frac{\partial u^1}{\partial y},
\end{aligned}$$

meaning that the charge density ρ does not solve the 2D wave equation. On the other hand, in a one dimensional homogeneous case the additional terms cancel and ρ satisfies the 1D wave equation. We can thus verify our FD scheme and its implementation by comparing with solutions of the wave equation for which the exact traveling wave solution is known. We control the initial data and the velocity by taking $u^2(t = 0) = Cu^1(t = 0)$ for the y-uniform case, and then $\frac{\partial \rho}{\partial t}(t = 0) = -C\frac{\partial \rho}{\partial x}(t = 0)$ holds. The same procedure also applies for the x-uniform case after replacing C by $-iC$. In this manner, we can set the initial velocity by C and allow for left-traveling, right-traveling, and symmetric waves for $C = 1, -1, 0$, respectively.

The corresponding results are shown in Figure 3.1 for the case uniform in y and in Figure 3.2 for the case uniform in x. It is seen that the numerical solution indeed approximates the exact solution.
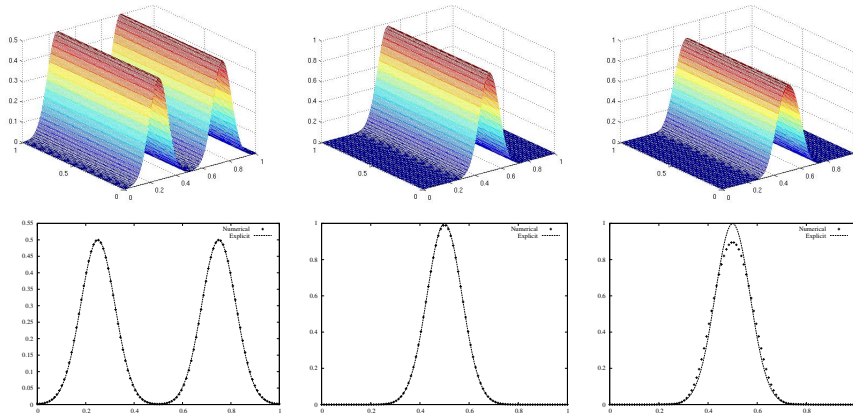
146

Figure 3.1: Plots of the charge concentration ρ computed by the FD scheme and of the exact solution of the 1D wave equation at times $t = 0.25$, $t = 1$, and $t = 4$ with $u^2(t = 0) = 0$ corresponding to a single traveling wave. Even after 4 periods, the wave form is nearly identical to the initial concentration.
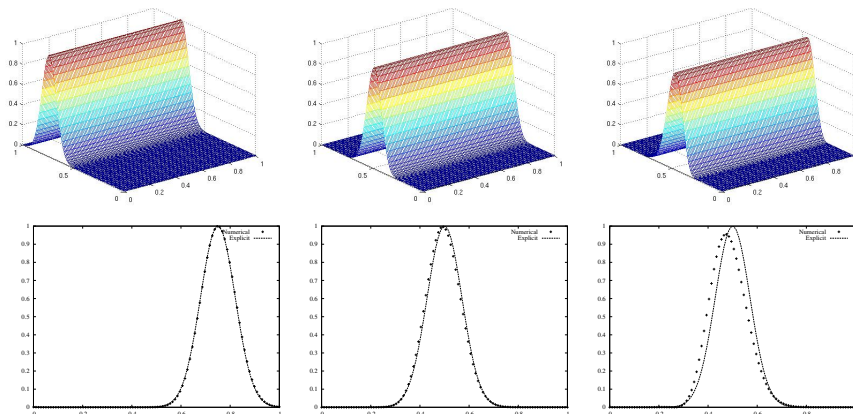

Figure 3.2: Plots of the charge concentration ρ computed by the FD scheme and of the exact solution of the 1D wave equation at times $t = 0.25$, $t = 1$, and $t = 4$ with $u^2(t = 0) = u^1(t = 0)$ corresponding to a single traveling wave. Even after 4 periods, the wave form is nearly identical to the initial concentration.

147

In the 2D case, we cannot use the exact solution of the wave equation for comparison. The rotationally symmetric solution obtained by the FD scheme for $u^2(t = 0) = 0$ is shown in Figure 3.3.
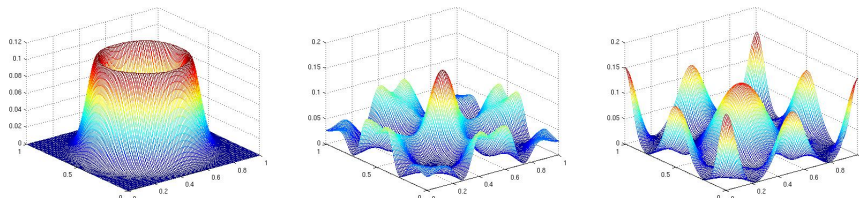

Figure 3.3: Plots of the charge concentration ρ obtained as the numerical solution of the Dirac equation at times $t = 0.25$, 1, and and 4.

Error estimates for the more complicated case $m \neq 0$ are shown in Figures 3.7 and 3.8 below.

3.4.2 Results for Particles with Mass

In the next step, we consider the Dirac equation in the $m \neq 0$ regime, while leaving the other parameters unchanged. Now the complex rotation term is added and the long-term behavior is completely changed. The initial behaviour of the system is close to the massless case, but a dragging wave slowly destabilizes the system and the eventual behavior is completely different.

We can again calculate explicit solutions by generating a second-order PDE and calculating

$$\begin{aligned}
\frac{\partial^2 u^1}{\partial t^2} &= -\frac{\partial}{\partial x}\left(-\frac{\partial u^1}{\partial x} + iu^2\right) - i\frac{\partial u^1}{\partial t} \\
&= \frac{\partial^2 u^1}{\partial x^2} - i\left(-\frac{\partial u^1}{\partial t} - iu^1\right) - i\frac{\partial u^1}{\partial t} \\
&= \frac{\partial^2 u^1}{\partial x^2} - u^1.
\end{aligned}$$

A similar calculation gives an identical result for y in place of x. As before in the massless case, the 2D case contains the additional terms mentioned above and there is no obvious explicit solution. Figure 3.4 shows the numerical solution; the different behavior at $t = 1$ and especially at $t = 4$ is seen compared to Figure 3.3.
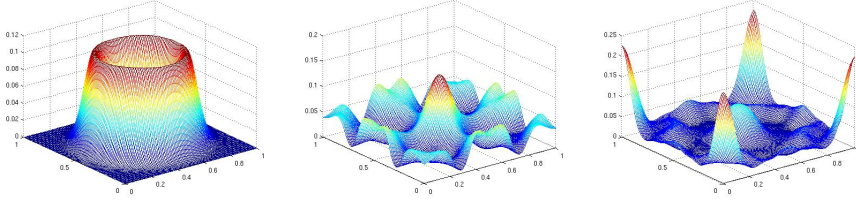
Figure 3.4: Plots of the charge concentration ρ for the $m = 1$ regime at times $t = 0.25$, $t = 1$, and $t = 4$.

The equation in 1D can be solved explicitly using, for instance, separation of variables, yielding

$$u^1 = \sum_{n=-\infty}^{\infty} C_n^{\pm} e^{i(\pm\sqrt{4\pi^2 n^2 + 1}t - 2\pi n x)}. \tag{3.14}$$

Thus wave packets move with the dispersion relation

$$\omega(k) = \sqrt{k^2 + 1}.$$

In this case, using 5 (symmetric) terms of the Fourier expansion yields a good approximation of the Gaussian and the explicit solution. The initial data are

$$u^1(t = 0) = \sum_{n=-4}^{4} \left(C_n^+ + C_n^-\right) e^{-2\pi i n x},$$

$$\frac{\partial u^1}{\partial t}(t = 0) = \sum_{n=-4}^{4} i\sqrt{4\pi^2 n^2 + 1} \left(C_n^+ - C_n^-\right) e^{-2\pi i n x}.$$

The sum $C_n^+ + C_n^-$ is calculated by the Fourier transform of the initial Gaussian distribution, and the equation for the initial velocity allows us to calculate the difference $C_n^+ - C_n^-$ and thus to determine the coefficients. This explicit solution makes it possible to verify the numerical scheme including the rotation step.

Figures 3.5 and 3.6 show results for waves traveling in the x- and y-directions, respectively. Figure 3.5 shows an initially traveling wave, whereas Figure 3.6 shows an initially stationary wave. Both Figure 3.5 and 3.6 show noticeable error at time $t = 4$. Since the errors in Figure 3.5 are larger, this case is used to demonstrate that the method indeed converges numerically – See Section 3.3.

Plots of the residual error for uniform refinements of space and time by a factor of 2 are given in Figure 3.7. It is seen that the error decreases by the expected factor of
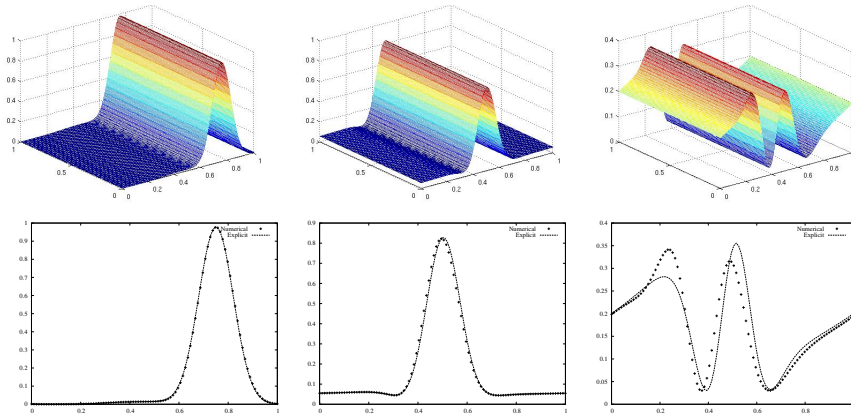
Figure 3.5: Plots of $|u^1|^2$ for the numerical simulation for the 1D $m = 1$ regime at times $t = 0.25$, $t = 1$, and $t = 4$ with the initial condition $u^2(t = 0) = u^1(t = 0)$ compared to the explicit solution given by Equation (3.14).
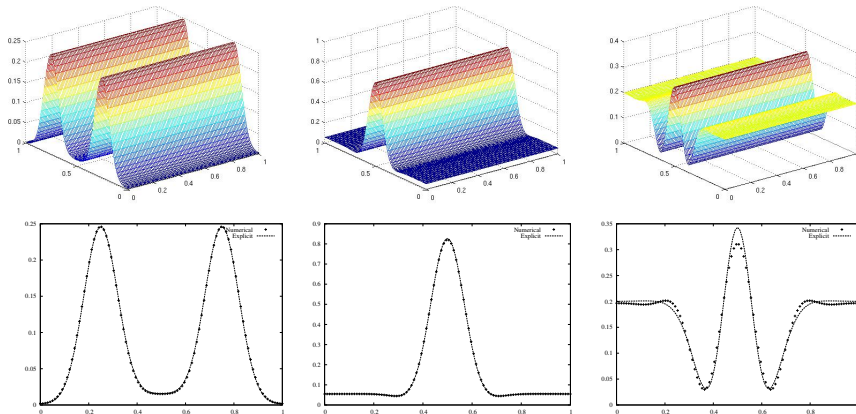

Figure 3.6: Plots of $|u^1|^2$ for the numerical simulation for the 1D $m = 1$ regime at times $t = 0.25$, $t = 1$, and $t = 4$ with the initial condition $u^2(t = 0) = 0$ compared to the explicit solution given by Equation (3.14).

≈ 4 for each refinement. The convergence order is quantified with more data by plotting the maximum value of the absolute value of the error in Figure 3.8. Each dotted line represents uniform refinement by factors of 2 (in both space and time) for different initial discretizations. The results in Figure 3.7 correspond to the central dotted line. The bold line indicates the expected second-order convergence with an arbitrary start point.
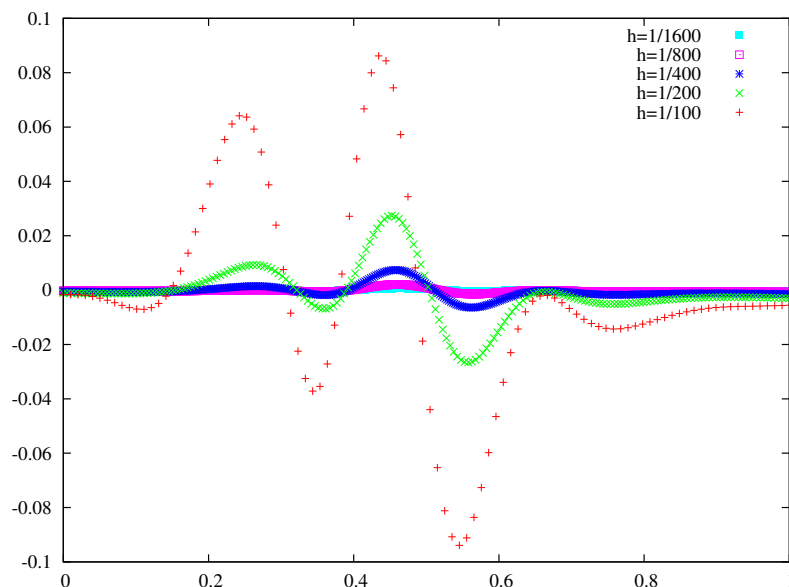

Figure 3.7: The error in calculating $|u^1|^2$ at time $t = 4$ for the case in Figure 3.5 for refined discretizations. Here $\Delta x = \Delta y = \Delta t = h = 100, 200, 400, 800, 1600$. The error decreases as the step size is reduced. The final two situations are indistinguishable on the plot, and hence more comparisons of the absolute error are shown in Figure 3.8.

A Note on Discretization Errors

A common error seen in similar numerical schemes for the Dirac equation is a loss of rotational symmetry, usually observed as a four-fold symmetric solution with errors following the directions of the square grid. In particular, if the original Gaussian wave packet is too tight with respect to the grid size, incorrectly broken symmetry is observed. In Figure 3.9, numerical approximations of solutions that are radially symmetric are shown at time $t = .5$ on a grid of size 100×100 with a time step of 0.01 for three different initial conditions given by $e^{-2\left(\alpha \frac{(x_i - 50)^2 + (y_i - 50)^2}{100^2}\right)}$ with $\alpha = 5, 10,$ and 20 (corresponding
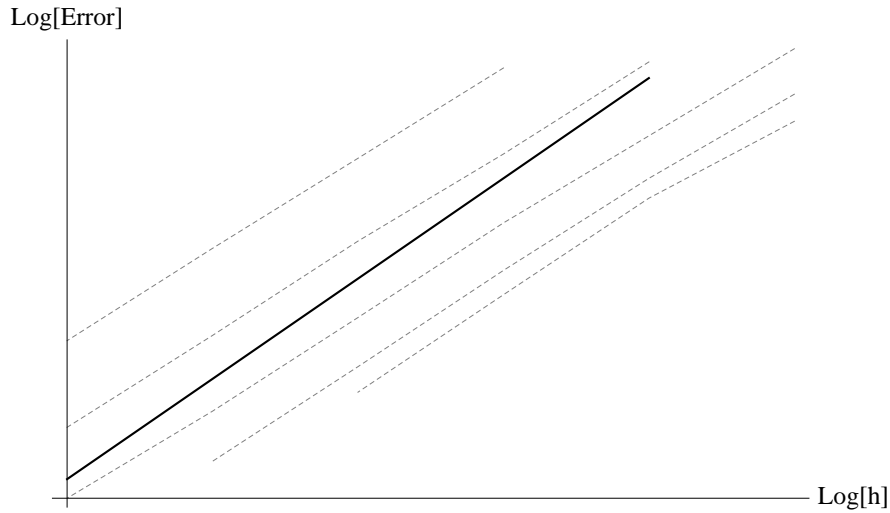
151

Figure 3.8: The maximal error in calculating $|u^1|^2$ at time $t = 4$ (as in Figure 3.5) for refined discretizations as a log-log plot. The expected second-order convergence is shown as a bold line. The other plots consist of uniformly refining by a factor of 2 both the grid $\Delta x = \Delta y = h$ and the time step $\Delta t = k$. The starting points $(1/h, 1/k)$ are (from top to bottom) the points $(50, 200)$, $(50, 100)$, $(50, 50)$, $(100, 50)$, and $(200, 50)$.

to MATLAB's $gausswin(N, \alpha)$ with $N = 100$). In addition, the solutions for the same initial conditions but with a homogeneous spatial refinement by a factor of 2 and a time-refinement by a factor of 10 are shown. The spatial refinement greatly reduces the non-symmetric errors that are seen as four peaks in the ring, whereas reducing the time step has little effect.

This symmetry-breaking effect can be explained by considering the truncation errors calculated in Section 3.3.1, since a sharply peaked Gaussian will have much higher second-order spatial derivatives than a broad profile. On the other hand, the change in time is relatively independent of the profile so that the spatial variations account for the majority of the error.

It is therefore recommended to use a suitable step size in the spatial discretization in order to avoid these numerical artifacts. The suitable step size depends on the smoothness of the initial condition.
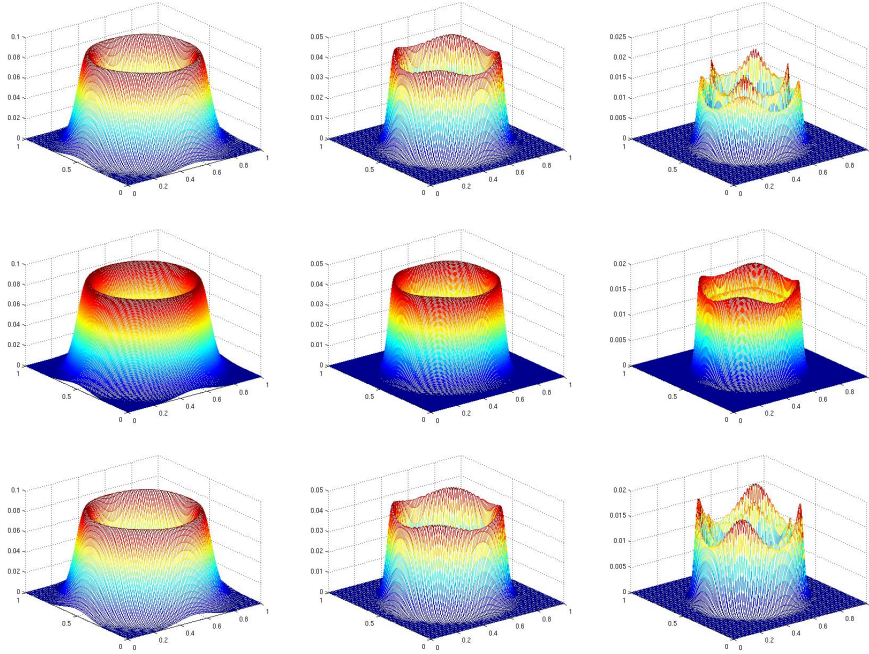
Figure 3.9: Plots of the charge concentration ρ for the numerical simulation for rotationally symmetric initial Gaussian wave packets at time $t = .5$. For narrower profiles, we see broken symmetry resulting from discretization errors along the directions of the discretizations. The columns correspond to Guassians of the form $e^{-2\left(\alpha\frac{(x_i-50)^2+(y_i-50)^2}{100^2}\right)}$ with $\alpha = 5$, 10, and 20. The second row shows the results for a 200×200 grid, whereas the third row shows the results for a time discretization of 0.001. Spatial refinement reduces the symmetry-breaking errors, whereas temporal refinement has little or no effect.

Plane Wave Simulation

We have already calculated the explicit solution for a wave in only one-direction in Equation (3.14), and we can use this result to write the explicit solution for plane-waves. As these plane waves are uniform in space, they result in a constant value for the charge concentration ρ, while u^1 and u^2 rotate in the complex plane. In order to examine such solutions more closely, we define the current densities J and the total currents F:

$$J_x := \bar{u^1}u^2 + u^1\bar{u^2}, \qquad\qquad F_x = \int_y J_x dy$$

$$J_y := -i\bar{u^1}u^2 + iu^1\bar{u^2} \qquad\qquad F_y = \int_x J_y dx.$$

We show plots of F_x and F_y for the simulations shown in Figures 3.2 and 3.5 in Figure 3.10.
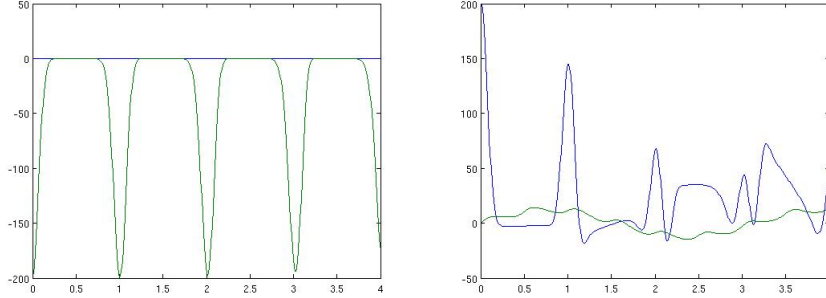

Figure 3.10: Plots of the currents F_x (blue) and F_y (green) for the simulations given by Figure 3.2 and Figure 3.5 respectively. For the massless case the current orthogonal to the wave-packet remains zero for all time and we see decay in F_y corresponding to the decay in ρ. For the second case, our original wave packet is not a pure-state and thus we see mixing in the current. Note that the total current in the domain is conserved, but that this plot gives the currents evaluated over lines at the center of the domain.

Although we see from Figure 3.7 that errors in $|u|$ occur over time, our method is very good at conserving currents. In Figure 3.11, plots of the currents F_x and F_y are shown after 1000 time steps up to time $t = 10$ in order to demonstrate the stability of the system regarding the long-time behavior of particles with mass for a uniform plane-wave. Effects such as these depend greatly on the chosen numerical scheme are are well known, for instance, in the case of the Schrödinger equation [43].
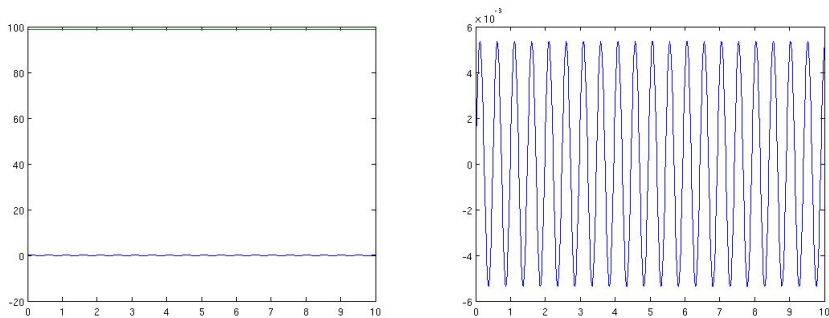
154

Figure 3.11: Plots of the currents F_x (blue) and F_y (green) for plane-wave solutions to the Dirac equation versus time up to time $t = 10$ with timestep $\Delta t = .01$ and a 100×100 grid. The first image shows a plane wave in the y-direction with associated momenta of 2π (to preserve periodicity). The second image shows a zoomed-in version of F_x indicating that the deviation from zero does not increase over time, but instead is completely periodic. Note that the scales are different so that the oscillations in the second image are four orders of magnitude smaller than the constant current in the first image.

3.4.3   Adding the Electric Potential

We can furthermore add an electric potential V. The most famous effect resulting from adding a potential to the Dirac equation is the Klein paradox. For the 3D Dirac Equation, particles of a specific energy regime are reflected from a potential barrier, but particles of higher and lower energy are transmitted. The 2D Dirac Equation does not exhibit the same behavior. In fact, the transmission is perfect for all energy barriers at normal incidence in the 2D Dirac model [59].

In order to demonstrate the effect of the electric potential, we consider two other interesting examples. First, we investigate the effect of a step potential for rotationally symmetric initial data. This mirrors the simulation of the effect of a P-N junction on electrons from a point source. In this case we hence expect the step potential to act as a a Veselago lens [14] with a symmetric focusing point on the opposite side of the barrier. Simulations showing the time evolution are given in Figure 3.12 and for potential barriers of various heights in Figure 3.13.

In the second example, we consider a triangle shaped potential and simulate a beam splitting potential. We show the potential used and a sequence of plots for various
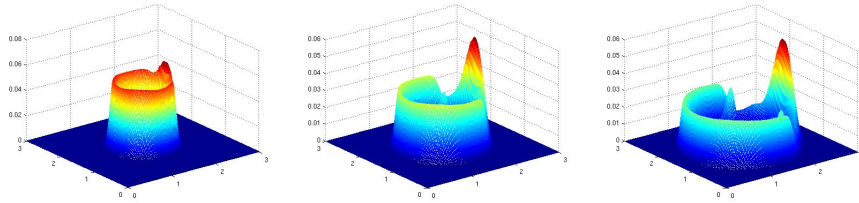
155

Figure 3.12: Plots of the charge concentration ρ on a $300{\times}300$ grid with $\Delta x = \Delta y = 0.01$ beginning with a rotationally symmetric Gaussian supported only on the middle $100{\times}100$ segment of the domain and set to zero outside. The potential is zero in the left two thirds of the domain and equal to 20 in the right segment. The solution is shown at times $t = 0.5$, $t = 0.75$, and $t = 1$, respectively, with a time step of $\Delta t = 0.01$. The phenomenon of Veselago lensing is clearly seen, whereby the charge concentration to the right concentrates at a point symmetrically opposite of the point source.

Figure 3.13: Plots of the charge concentration ρ on a $300{\times}300$ grid with $\Delta x = \Delta y = 0.01$ beginning with a rotationally symmetric Gaussian supported only on the middle $100{\times}100$ segment of the domain. The potential is zero for the left two thirds of the domain and equal to 5, 10, and 20 in the right segment. The solutions are shown at time $t = 1$ with a time step of $\Delta t = 0.01$. Veselago lensing is again observed as in Figure 3.12.

156

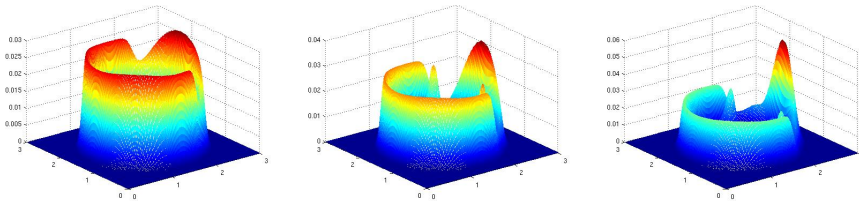time steps in Figure 3.14 and show the splitting capability and its dependence on the magnitude of the potential in Figure 3.15.

Figure 3.14: Plots of the charge concentration ρ on a 200×100 grid with $\Delta x = \Delta y = 0.01$ beginning with a traveling wave on the left half of the domain. The potential is shown in the bottom image. The solutions at times $t = 0.5$, $t = 0.75$, and $t = 1$, respectively, are shown for a time step of $\Delta t = 0.01$. The original wave packet is completely supported in the $V = 0$ region, and the peak coincides with the potential jump at the first time point, $t = 0.5$.

Figures 3.12 - 3.15 are shown for the massless case $m = 0$ to reflect the application to graphene, but similar results are obtained for particles with mass (see Figure 3.18 below).

3.4.4 Adding the Magnetic Potential

The next component to be considered is the magnetic potential. The main difficulty in appropriately applying a magnetic vector potential is that the potentials which correspond to constant magnetic fields are not periodic, and thus cause issues for the periodic boundary conditions of the Crank-Nicolson system. On the other hand, we can easily simulate toy models with constant A_x or A_y in the analogous 1D case to insure that
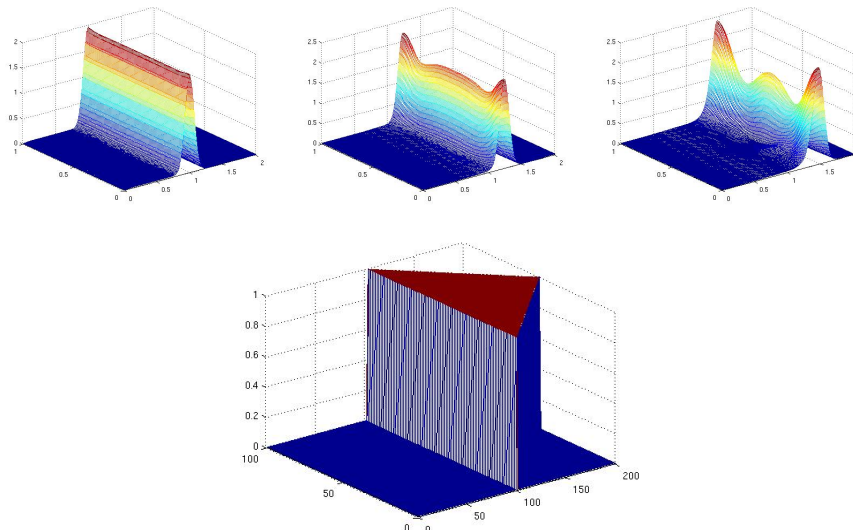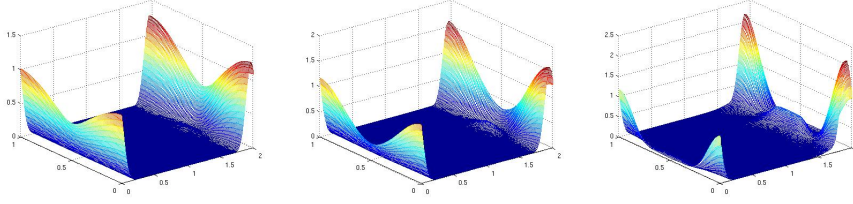
Figure 3.15: Plots of the charge concentration ρ on a 200×100 grid with $\Delta x = \Delta y = 0.01$ beginning with a traveling wave on the left half of the domain (as in Figure 3.1) for various heights of the potential barrier. The potential is shown in the bottom image from Figure 3.14 and has magnitude 1, 2, and 5, respectively, for the plots. The solutions are shown at time at $t = 1.5$ with a time step of $\Delta t = 0.01$.

they are implemented correctly. For $V = A_x = m = 0$ uniform in y, we have

$$\frac{\partial u^1}{\partial t} = -\left(\frac{\partial}{\partial x} + A_y\right) u_4,$$

$$\frac{\partial u^2}{\partial t} = -\left(\frac{\partial}{\partial x} - A_y\right) u_1$$

and thus for A_y constant,

$$\frac{\partial^2 u^1}{\partial t^2} = \left(\frac{\partial}{\partial x} + A_y\right)\left(\frac{\partial}{\partial x} - A_y\right) u^1 = \frac{\partial^2 u^1}{\partial x^2} - A_y^2 u^1. \tag{3.16}$$

If we take $A_y = 1$, then we get exactly the same case as $A_y = 0$, $m = 1$ above and we have already obtained the explicit solution in Equation (3.14). We obtain the same results for $V = A_y = m = 0$ uniform in x, and can thus do the same simulations as above to test the consistency of the implementation for A – see Figure 3.16.

3.4.5 Self-Consistent Fields

Note that up to this point we have not included the self-consistent field (V is always given) and thus have not chosen a value for λ^2. We show results for changing this parameter (the coefficient of ρ in the Poisson equation as discussed in Section 3.2.4) in Figure 3.17.

Since we consider a 2D Dirac model for confined transport embedded in 3D real space, the 3D Poisson's equation is solved with a thin sheet of charge representing the only contribution to the right-hand side. With this extension, we extend our domain
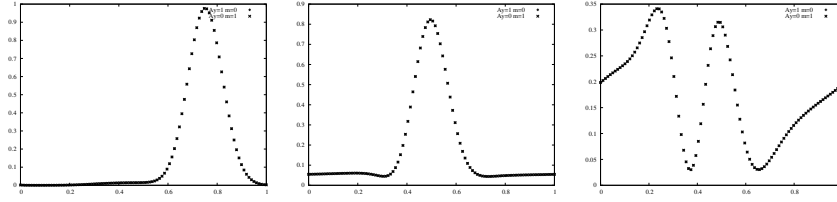
158

Figure 3.16: Plots of $|u^1|^2$ for the numerical simulation for the 1D in x massless regime with $A_y = 1$ uniformly plotted with the results for the 1D $m = 1$ regime at times $t = 0.25$, $t = 1$, and $t = 4$. Note that the two plots are nearly indistinguishable, verifying the results of Equation 3.16. See Figure 3.5 for plots of the explicit results. The same results are obtained in y but are not shown here.
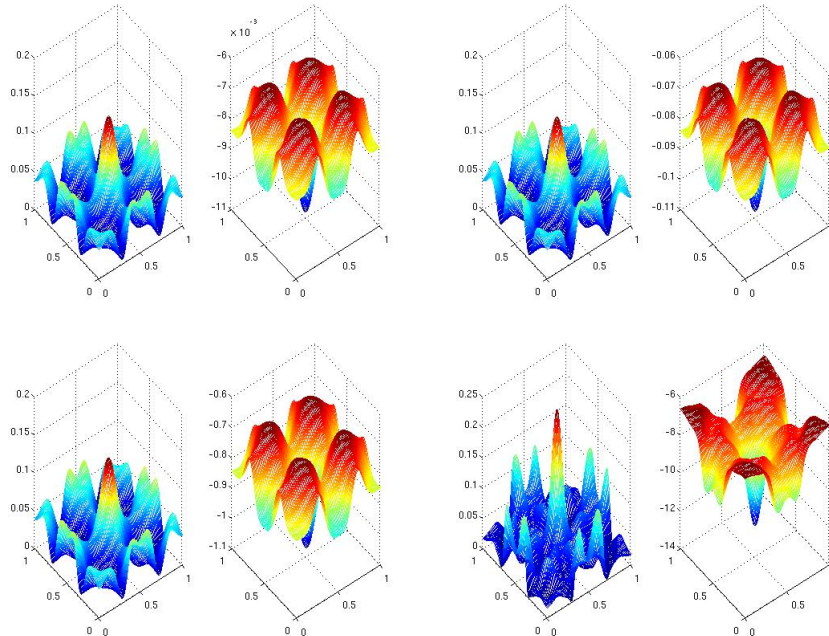
Figure 3.17: Plots of the charge concentration ρ and the electric potential V on a $100 \times 100 \times 19$ grid representing a $1 \times 1 \times 1$ box with a charge sheet parallel to the xy-plane in the center of the box and homogeneous Dirichlet boundary conditions for the potential at $z = 0$ and $z = 1$. The figures use $\lambda^2 = 1$, 10, 100, and 1000 respectively. The magnitude of the potential changes proportionally, but there is little impact on the charge density for the lower two values. Compare to the middle image in Figure 3.4 for the $V = 0$ case.

from a 2D rectangle to a 3D rectangular prism. In addition, we must now define boundary conditions for Poisson's equation in the z-direction (corresponding to the gate and backgate contacts for a graphene based device). Because Poisson's equation is linear, any uniform Dirichlet boundary conditions on the top and bottom of the box simply add a linear contribution to the potential in the z-direction. Thus adding nonzero boundary conditions at the top and bottom of the box acts in the same manner as adding a given potential in the plane. This roughly corresponds physically to the difference between doping a device and selectively applying an external potential. We simulate a beam splitter with a self-consistent potential in Figure 3.18, again illustrating the effect of changing λ^2.
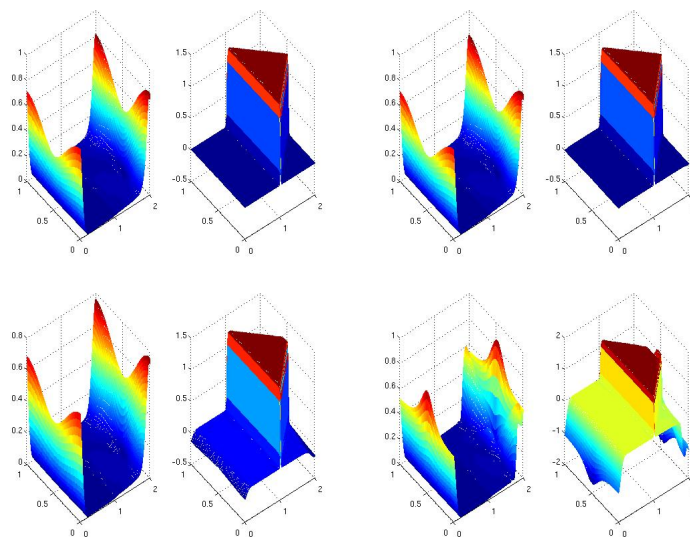
Figure 3.18: Plots of the charge concentration ρ on a $200 \times 100 \times 9$ grid with $\Delta x = \Delta y = 0.01$ and $L_z = .01$. The setup is identical to that in Figures 3.14 and 3.15, except that the potential is applied at the top of the domain and the potential at the bottom of the domain is zero. The plots correspond to $\lambda^2 = 1, 10, 100, 1000$ respectively. The large discrepancy in the fourth image is a result of the net potential difference being negative instead of positive. See Figure 3.19.

Decreasing the size of domain in the z-direction has a similar effect to increasing λ^2. Increasing the strength of the electric field in the z-direction reduces the potential differences in the xy-plane following Gauss' Law. This effect is illustrated in Figure 3.20.
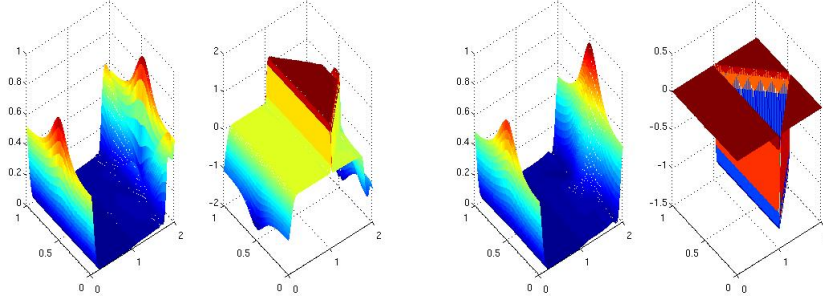
Figure 3.19: A reproduction of the final image from Figure 3.18 (left) and an image for the splitting case when the splitting potential is negative instead of positive. The Poisson scaling parameter $\lambda^2 = 1000$ for the left image and 1 for the right image. note that the concentrations are very similar, suggesting that the change in behavior observed in Figure 3.18 occurs because the net potential (from the external potential and the self-consistent potential) is negative and the particles react accordingly.
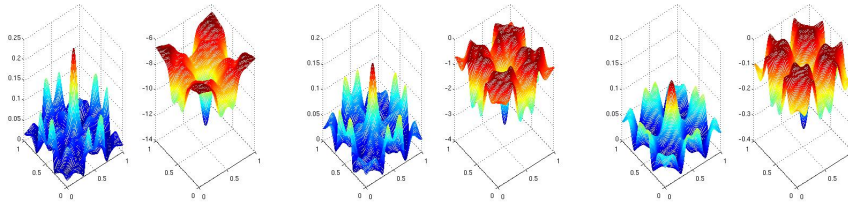
Figure 3.20: Plots of the charge concentration ρ and the electric potential V on a $100 \times 100 \times 19$ grid representing a box with a charge sheet parallel to the xy-plane in the center of the box. The boxes have dimensions $1 \times 1 \times 1$, $1 \times 1 \times 0.1$, and $1 \times 1 \times 0.01$, respectively, with homogeneous Dirichlet boundary conditions for the potential at the top and bottom of the box. Reducing the size of the domain counteracts the size of $\lambda^2 = 1000$. Compare to Figure 3.17.

Chapter 4

Conclusions

4.1   Photovoltaic Devices

Solar energy is an increasingly attractive alternative to fossil fuels and nuclear power. Classical inorganic semiconductor solar cells have been the mainstay of the industry for many years, but the more environmentally-friendly alternative of organic devices is growing more and more popular. Because the efficiency of such devices lags significantly behind their inorganic counterparts, understanding the simulation of the devices is incredibly important. The first step is to understand exactly which model characteristics are necessary in order to understand the operating characteristics of such a device.

In Chapter 2 we present a self-consistent model which considers independent dynamics for both excitons and polaron-pairs within a device. We develop the model based on extending, combining, and generalizing models discussed in the current literature. Of particular importance, we investigate the effect of consistently treating the minority carriers in the device. We additionally include dipole effects near the interface and diffusion of polaron pairs on the interface of the device – effects not widely considered in the literature.

We develop a 2D hybrid discontinuous Galerkin finite element scheme to simulate the model in the steady state and use the simulations to generate qualitatively realistic current-voltage characteristics. We furthermore develop an asymptotic scheme designed to accurately describe the behavior of the device in short circuit and optimal power point conditions. We also show that when the minority carriers are negligible, the PDE system can be solved explicitly in the bulk for nonzero currents – a result which had not

previously appeared in the literature.

We then discuss the physical ramifications of our model with regards to the simulation of actual devices. Our numerical simulations suggest that the consistent dipole model is not necessary for understanding the physical characteristics of the device. Our model also gives minority carrier concentrations which are significantly smaller than the majority carrier concentrations, but this does not guarantee that the minority current is zero. In fact, we show that a continuous model for charge carriers and polaron pairs in an interface region develops significant problems exist as the device nears open circuit. The presence of a minority carrier current corresponds loosely to naïvely allowing discontinuous currents in unipolar models. We give several conditions for models which will not exhibit this effect, which seems to be symptomatic of the smoothness enforced on the solutions by the reaction-diffusion framework.

We also show that a model with an interface region is extremely difficult to model numerically in a robust way. In particular, the exact two effects which have been largely ignored in the literature (interface dipoles and polaron pair diffusion) are extremely hard to generalize for complicated device interfaces. We conclude that numerical models with a thin interface provide the best opportunity for physically relevant numerical modeling for organic devices. Such a model would allow for dipole direction to be exactly defined via the interface normal, and allow for diffusion on the interface (which would be a manifold of one order lower than the domain).

Finally, we give perspectives on proving the existence and uniqueness of our model system. We demonstrate how our model fits into the fixed-point methods which have been used extensively in the mathematical analysis of semiconductor modeling and explain how dipole-interactions provide an interesting analytical challenge. We then discuss extending entropy methods to include an exciton equation. We must again ignore the dipole effects, but we can then develop a non-increasing entropy for the system with some mild additional assumptions. However, the smoothness provided by such entropy estimates is insufficient for the bootstrap argument necessary to prove global existence of reasonable solutions. We suggest that duality methods might provide a dimension-independent method for obtaining initial regularity results and show some sample calculations in this direction.

4.2 Graphene and the Dirac Equation

The Dirac equation is one of the fundamental equations in physics. It has received much less attention with regards to numerical methods compared to its closest relative, the Schrödinger equation. This is possibly due to its counter-intuitive properties and modeling complexity, providing complications for numerical methods. As a new and interesting 2D material, graphene represents an ideal system for numerical simulations with the 2D Dirac equation.

In Chapter 3 we develop a finite-difference scheme for the 2D electromagnetic Dirac equation and we show its convergence by demonstrating its consistency and stability. Electrostatic and magnetic potentials are included in the equation. The FD approximations are of second order in both space and time. The FD method, including its convergence order, is verified using several test cases with known results and further results are shown for cases where no explicit solution is known. We also discuss common numerical errors which result in a loss of rotational symmetry and explain how to avoid them.

Having established the capabilities of our numerical scheme, we use our method for several simulations of graphene. We consider the case of low energies, in which the particles have wave vectors sufficiently close to the Dirac points and thus the Dirac equation represents an accurate model system. We discuss the influence of the electrostatic potential, the magnetic potential, and self-consistent simulations. In particular we demonstrate that our method can simulate potentials which act as Veselago lenses or beam splitters.

Bibliography

[1] The nobel prize in physics 2010 - advanced information. *NobelPrize.org*, 2010.

[2] D. A. Abanin, S. V. Morozov, L. A. Ponomarenko, R. V. Gorbachev, A. S. Mayorov, M. I. Katsnelson, K. Watanabe, T. Taniguchi, K. S. Novoselov, L. S. Levitov, and A. K. Geim. Giant nonlocality near the Dirac point in graphene. *Science*, 332:328–330, 2011.

[3] V. I. Arkhipov, P. Heremans, and H. Bässler. Why is exciton dissociation so efficient at the interface between a conjugated polymer and an electron acceptor? *Applied Physics Letters*, 82(25):4605–4607, 2003.

[4] D. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM Journal on Numerical Analysis*, 39:1749–1779, 2002.

[5] J. A. Barker, C. M. Ramsdale, and N. C. Greenham. Modeling the current-voltage characteristics of bilayer polymer photovoltaic devices. *Physical Review B*, 67(7):075205, 2003.

[6] M. Bosson, S. Grudinin, X. Bouju, and S. Redon. Interactive physically-based structural modeling of hydrocarbon systems. *Journal of Computational Physics*, 231(6):2581–2598, 2012.

[7] C. L. Braun. Electric field assisted dissociation of charge transfer states as a mechanism of photocarrier production. *The Journal of Chemical Physics*, 80(9):4157–4161, 1984.

[8] F. Brezzi, L. Marini, S. Micheletti, P. Pietra, R. Sacco, and S. Wang. Finite elements and finite volumes discretizations of drift-diffusion type fluid models for

semiconductors. In W. Schilders and E. ter Maten, editors, *Numerical Methods in Electromagnetics*, volume 13 of *Handbook of Numerical Analysis*. Elsevier, 2005.

[9] D. Brinkman, K. Fellner, and P. A. Markowich. Asymptotic justification of a unipolar model for a photovoltaic bilayer. In prep.

[10] D. Brinkman, K. Fellner, P. A. Markowich, and M.-T. Wolfram. A drift-diffusion-reaction model for excitonic photovoltaic bilayers: Asymptotic analysis and a 2-D HDG finite-element scheme. *Mathematical Models and Methods in Applied Sciences*, 23(5), 2013.

[11] D. Brinkman, C. Heitzinger, and P. A. Markowich. A convergent 2D finite difference scheme for the Dirac-Poisson system and the simulation of graphene. Submitted to the *Journal of Computational Physics*, 2012.

[12] G. A. Buxton and N. Clarke. Predicting structure and property relations in polymeric photovoltaic devices. *Physical Review B*, 74, 2006.

[13] G. A. Buxton and N. Clarke. Computer simulation of polymer solar cells. *Modelling and simulation in Materials Science and Engineering*, 15:13–26, 2007.

[14] V. V. Cheianov, V. I. Fal'ko, and B. L. Altshuler. The focusing of electron flow and a Veselago lens in graphene p-n junctions. *Science*, 315(5816):1252–1255, 2007.

[15] D. Cheyns, J. Poortmans, P. Heremans, C. Deibel, S. Verlaak, B. P. Rand, and J. Genoe. Analytical model for the open-circuit voltage and its associated resistance in organic planar heterojunction solar cells. *Physical Review B*, 77:165332, 2008.

[16] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM Journal on Numerical Analysis*, 47(2):1319–1365, 2009.

[17] B. Cockburn, G. E. Karniadakis, and C.-W. Shu. The development of discontinuous Galerkin methods. In B. Cockburn, G. E. Karniadakis, and C.-W. Shu, editors, *Discontinuous Galerkin Finite Element Methods*, volume 11 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, 2000.

[18] J. Crank and P. Nicolson. A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Mathematical Proceedings of the Cambridge Philosophical Society*, 43(1):50–67, 1947.

[19] B. K. Crone, P. S. Davids, I. H. Campbell, and D. L. Smith. Device model investigation of bilayer organic light emitting diodes. *Journal of Applied Physics*, 87(4):1974–1982, 2000.

[20] C. De Falco, M. Porro, R. Sacco, and M. Verri. Multiscale modelling and simulation of organic solar cells. *Computer Methods in Applied Mechanics and Engineering*, 245-246:102–116, October 2012.

[21] C. De Falco, M. Verri, and R. Sacco. Analytical and numerical study of photocurrent transients in organic polymer solar cells. *Computer Methods in Applied Mechanics and Engineering*, 199:1722–1732, 2010.

[22] A. De Vos. The fill factor of a solar cell from a mathematical point of view. *Solar Cells*, 8(3):283–296, 1983.

[23] M. Di Francesco, K. Fellner, and P. A. Markowich. The entropy dissipation method for spatially inhomogeneous reaction-diffusion type systems. *Proceedings of the Royal Society A*, 46(2100):3273–3300, 2008.

[24] P. A. M. Dirac. The quantum theory of the electron. *Proceedings of the Royal Society of London*, 117:610–624, 1928.

[25] P. A. M. Dirac. A theory of electrons and protons. *Proceedings of the Royal Society of London*, 126:360–365, 1930.

[26] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 2008 reprint edition, 1998.

[27] L. C. Evans. A survey of entropy methods for partial differential equations. *Bulletin of the American Mathematical Society*, 41:409–438, 2004.

[28] A. Fischer, P. Pahner, B. Lüssem, K. Leo, R. Scholz, T. Koprucki, J. Fuhrmann, K. Gärtner, and A. Glitzky. Self-heating effects in organic semiconductor devices enhanced by positive temperature feedback. WIAS Preprint No. 1693, 2012.

[29] M. Flato, J. C. H. Simon, and E. Taflin. Asymptotic completeness, global existence and the infrared problem for the Maxwell-Dirac equations. arXiv:hep-th/9502061v1, 1995.

[30] J. Frehse and J. Naumann. On weak solutions to the non-stationary van Roosbroeck's system with discontinuous permittivities. *Mathematical Methods in the Applied Sciences*, 20:689–706, 1997.

[31] J. Frenkel. On pre-breakdown phenomena in insulators and electronic semiconductors. *Physical Review*, 54:647–648, 1938.

[32] H. Gajewski. On existence, uniqueness, and asymptotic behavior of solutions of the basic equations for carrier transport in semiconductors. *Zeitschrift für Angewandte Mathematik und Mechanik*, 65:101–108, 1985.

[33] Y. Ge. Multigrid method and fourth-order compact difference discretization scheme with unequal meshsizes for 3D Poisson equation. *Journal of Computational Physics*, 229:6381–6391, 2010.

[34] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, second edition, 1998.

[35] A. Glitzky. Analysis of electronic models for solar cells including energy resolved defect densities. *Mathematical Methods in the Applied Sciences*, 34:1980–1998, 2011.

[36] T. E. Goliber and J. H. Perlstein. Analysis of photogeneration in a doped polymer system in terms of a kinetic model for electric-field-assisted dissociation of charge-transfer states. *The Journal of Chemical Physics*, 80(9):4162–4167, 1984.

[37] M. A. Green, K. Emery, Y. Hishikawa, and W. Warta. Solar cell efficiency tables (version 37). *Progress in Photovoltaics: Research and Applications*, 19:84–92, 2011.

[38] C. Groves, R. A. Marsh, and N. C. Greenham. Monte Carlo modeling of geminate recombination in polymer-polymer photovoltaic devices. *Journal of Chemical Physics*, 129, 2008.

[39] H. K. Gummel. A self-consistent iterative scheme for one-dimensional steady state transistor calculations. *IEEE Transactions on Electron Devices*, ED-11:455–465, 1964.

168

[40] J. J. M. Halls, C. A. Walsh, N. C. Greenham, E. A. Marseglia, R. H. Friend, S. C. Moratti, and A. B. Holmes. Efficient photodiodes from interpenetrating polymer networks. *Nature*, 376:498–500, 1995.

[41] M. C. Hanna and B. A. Gregg. Comparing organic to inorganic photovoltaic cells: Theory, experiment, and simulation. *Journal of Applied Physics*, 93(6), 2003.

[42] X. He, F. Gao, D. Hasko, S. Huttner, U. Steiner, N. C. Greenham, R. H. Friend, and W. T. S. Huck. Formation of nanopatterned polymer blends in photovoltaic devices. *Nano Letters*, 10:1302–1307, 2010.

[43] C. Heitzinger, C. Ringhofer, and S. Selberherr. Finite difference solutions of the nonlinear Schrödinger equation and their conservation of physical quantities. *Communications in Mathematical Sciences*, 5(4):779–788, 2007.

[44] H. Hoppe and N. S. Sariciftci. Bulk heterojunction solar cells. In S.-S. Sun and N. S. Sariciftci, editors, *Organic Photovoltaics: mechanisms, materials, and devices*. 2005.

[45] Z. Huang, S. Jin, P. A. Markowich, C. Sparber, and C. Zheng. A time-splitting spectral scheme for the Maxwell-Dirac system. *Journal of Computational Physics*, 208:761–789, 2005.

[46] I. Hwang and N. C. Greenham. Modeling photocurrent transients in organic solar cells. *Nanotechnology*, 19, 2008.

[47] I. Hwang, C. R. McNeill, and N. C. Greenham. Drift-diffusion modeling of photocurrent transients in bulk herterojunction solar cells. *Journal of Applied Physics*, 106, 2009.

[48] J. D. Jackson. *Classical Electrodynamics*. John Wiley & Sons, Inc., second edition, 1975.

[49] A. K. Jonscher. Electronic properties of amorphous dielectric films. *Thin Solid Films*, 1:213–234, 1967.

[50] L. J. A. Koster, E. C. P. Smits, V. D. Mihailetchi, and P. W. M. Blom. Device model for the operation of polymer/fullerene bulk heterojunction solar cells. *Physical Review B*, 72(8):085205, 2005.

169

[51] S. Lacic and O. Inganäs. Modeling electrical transport in blend heterojunction organic solar cells. *Journal of Applied Physics*, 97, 2005.

[52] C. Lehrenfeld. Hybrid discontinuous Galerkin methods for solving incompressible flow problems. Master's thesis, RWTH Aachen University, 2010.

[53] R. J. LeVeque. *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems.* Society for Industrial and Applied Mathematics, 2007.

[54] Y.-M. Lin, C. Dimitrakopoulos, K. A. Jenkins, D. B. Farmer, H.-Y. Chiu, A. Grill, and P. Avouris. 100-GHz transistors from wafer-scale epitaxial graphene. *Science*, 327:662, 2010.

[55] P. A. Markowich. *The Stationary Semiconductor Device Equations.* Springer-Verlag/Wien, 1986.

[56] R. A. Marsh, C. Groves, and N. C. Greenham. A microscopic model for the behavior of nanostructured organic photovoltaic devices. *Journal of Applied Physics*, 101(8):083509–083509–7, 2007.

[57] A. B. F. Martinson, A. M. Massari, S. J. Lee, R. W. Gurney, K. E. Splan, J. T. Hupp, and S. T. Nguyen. Organic photovoltaics interdigitated on the molecular scale. *Journal of the Electrochemical Society*, 153(3):A527–A532, 2006.

[58] H. Najafov, B. Lee, Q. Zhous, F. C. Feldman, and V. Podzorov. Observation of long-range exciton diffusion in highly ordered organic semiconductors. *Nature Materials*, 9(11):938–943, 2010.

[59] A. H. C. Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim. The electronic properties of graphene. *Reviews of Modern Physics*, 81:109–162, 2009.

[60] V. Nikitenko and H. Bässler. An analytic model of electroluminescence in bilayer oganic light emitting diodes with ohmic injection of charge carriers. *Journal of Applied Physics*, 90(4):1823–1826, 2001.

[61] K. Novoselov, A. Geim, S. Morozov, D. Jiang, Y. Zhang, S. Dubonos, I. Grigorieva, and A. Firsov. Electric field effect in atomically thin carbon films. *Science*, 306:666–669, 2004.

[62] K. Novoselov, Z. Jiang, Y. Zhang, S. Morozov, H. Stormer, U. Zeitler, J. Maan, G. Boebinger, P. Kim, and A. Geim. Room-temperature quantum Hall effect in graphene. *Science*, 315:1379, 2007.

[63] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov. Two-dimensional gas of massless Dirac fermions in graphene. *Nature*, 438(10):197–200, 2005.

[64] K. S. Novoselov, D. Jiang, F. Schedin, T. J. Booth, V. V. Khotkevich, S. V. Morozov, and A. K. Geim. Two-dimensional atomic crystals. *Proceedings of the National Acadamy of Science*, 102(30), 2005.

[65] T. Offermans, S. C. J. Meskers, and R. A. J. Janssen. Monte-Carlo simulations of geminate electron-hole pair dissociation in a molecular heterojunction: a two-step dissociation mechanism. *Chemical Physics*, 308:125–133, 2005.

[66] R. S. Ohl. Light-sensitive electric device, 1941. Patent.

[67] L. Onsager. Deviations from Ohm's law in weak electrolytes. *Journal of Chemical Physics*, 2(599), 1934.

[68] J. K. Pachos. Manifestations of topological effects in graphene. *Contemporary Physics*, 50:375–389, 2009.

[69] S. H. Park, A. Roy, S. Beaupré, S. Cho, N. Coates, J. S. Moon, D. Moses, M. Leclerc, K. Lee, and A. J. Heeger. Bulk heterojunction solar cells with internal quantum efficiency approaching 100%. *Nature Photonics*, 3, 2009.

[70] W. F. Pasveer, F. Cottaar, C. Tanase, R. Coehoorn, P. Bobbert, P. W. M. Blom, D. M. de Leeuw, and M. A. Michels. Unified description of charge-carrier mobilities in disordered semiconducting polymers. *Physical Review Letters*, 94, 2005.

[71] N.-K. Persson and O. Inganäs. Simulations of optical processes in organic phtovoltaic devices. In S.-S. Sun and N. S. Sariciftci, editors, *Organic Photovoltaics: mechanisms, materials, and devices*. 2005.

[72] M. Pierre. Global existence in reaction-diffusion systems with control of mass: A survey. *Milan Journal of Mathematics*, 78:417–455, 2010.

171

[73] G. Richardson, C. Please, J. Foster, and J. Kirkpatrick. Asymptotic solution of a model for bilayer organic diodes and solar cells. *SIAM Journal on Applied Mathematics*, 72:1792–1817, 2012.

[74] C. A. Ringhofer, C. Schmeiser, and P. A. Markowich. *Semiconductor Equations.* Springer, 1990.

[75] F. Schedin, A. Geim, S. Morozov, E. Hill, P. Blake, M. Katsnelson, and K. Novoselov. Detection of individual gas molecules adsorbed on graphene. *Nature Materials*, 6:652–655, 2007.

[76] O. Schenk, M. Bollhöfer, and R. A. Römer. On large-scale diagonalization techniques for the Anderson model of localization. *SIAM Journal on Scientific Computing*, 28:2006.

[77] O. Schenk, A. Wächter, and M. Hagemann. Matching-based preprocessing algorithms to the solution of saddle-point problems in large-scale nonconvex interior-point optimization. *Computational Optimization and Applications*, 36(2-3):321–341, 2007.

[78] J. Schöberl. Netgen an advancing front 2d/3d-mesh generator based on abstract rules. *Computing and Visualization in Science*, 1:41–52, 1997. 10.1007/s007910050004.

[79] F. Schwabl. *Advanced Quantum Mechanics.* Springer, third edition, 2005.

[80] C. J. Scott and G. G. Malliaras. Charge injection and recombination at the metal-organic interface. *Chemical Physics Letters*, 299:115–119, 1999.

[81] S. Selberherr. *Analysis and Simulation of Semiconductor Devices.* Springer-Verlag, 1984.

[82] G. W. Semenoff. Condensed-matter simulation of a three-dimensional anomaly. *Physical Review Letters*, 53:2449–2452, 1984.

[83] J. V. Shebalin. Homogeneous quantum electrodynamic turbulence. *Physica D*, 66:381–391, 1993.

[84] J. V. Shebalin. Numerical solution of the coupled Dirac and Maxwell equations. *Physics Letters A*, 226:1–6, 1997.

[85] G. Strang. On the construction and comparison of difference schemes. *SIAM Journal on Numerical Analysis*, 5(3):506–517, 1968.

[86] C. W. Tang. Two-layer organic photovoltaic cell. *Applied Physics Letters*, 48(183), 1986.

[87] B. Thaller. Visualizing the kinematics of relativistic wave packets. arXiv:quant-ph/0409079.

[88] B. Thaller. *Advanced visual quantum mechanics.* Springer Science and Business Media, Inc., 2005.

[89] W. Van Roosbroeck. Theory of the flow of electrons and holes in germanium and other semiconductors. *Bell System Technical Journal*, 29, 1950.

[90] S. Verlaak, D. Beljonne, D. Cheyns, C. Rolin, M. Linares, F. Castet, J. Cornil, and P. Heremans. Electronic structure and geminate pair energetics at organic-organic interfaces: The case of pentacene/ C_{60} heterojunctions. *Advanced Functional Materials*, 19:3809–3814, 2009.

[91] A. B. Walker, A. Kambili, and S. J. Martin. Electrical transport modelling in organic electroluminescent devices. *Jorunal of Physics: Condensed Matter*, 14, 2002.

[92] P. R. Wallace. The band theory of graphite. *Physical Review*, 71:622–634, 1947.

[93] J. Wang, W. Zhong, and J. Zhang. A general meshsize fourth-order compact difference discretization scheme for 3D Poisson equation. *Applied Mathematics and Computation*, 183:804–812, 2006.

[94] E. Weston. Apparatus for utilizing solar radiant energy, 1888. Patent.

[95] J. Williams and A. B. Walker. Two-dimensional simulations of bulk heterojunction solar cell characteristics. *Nanotechnology*, 19, 2008.

[96] M.-T. Wolfram. *Forward and inverse solvers for electrodiffusion problems.* PhD thesis, University of Linz, 2008.

[97] H. Wu, P. A. Markowich, and S. Zheng. Global existence and asymptotic behavior for a semiconductor drift-diffusion-Poisson model. *Mathematical Models and Methods in Applied Science*, 18(3):443–487, 2008.

[98] Y. Wu, Y.-M. Lin, A. A. Bol, K. A. Jenkins, F. Xia, D. B. Farmer, Y. Zhu, and P. Avouris. High-frequency, scaled graphene transistors on diamond-like carbon. *Nature*, 472:74–78, 2011.

[99] G. Yu, F. Gao, J. C. Hummelen, F. Wudl, and A. J. Heeger. Polymer photovoltaic cells: Enhanced efficiencies via a network of internal donor-accepter heterojunctions. *Science*, 270(5243):1789–1791, 1995.