TiO_2 and NiO:Cu Carrier-Selective Barrier Layers for Heterojunction Solar Cells

Alexander Hillel Klaimitz Berg

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Electrical Engineering
Adviser: Professor James C. Sturm

June 2019

Abstract

Next-generation solar cell designs will figure significantly in the approaching green energy mix. This thesis makes contributions to both the modeling and the materials of these devices. External quantum efficiency-based measurements of PEDOT/n-Si/TiO_2 solar cells are used, in conjunction with modeling, to extract the surface recombination velocity of the n-Si/CVD-TiO_2 interface. It is also shown that even very thin layers of ALD-TiO_2 can effectively passivate crystalline silicon, though even the thinnest layers contribute measurably to the hole barrier. The increase in surface recombination velocity after ALD-NiO overlayer deposition is linked to changes in the TiO_2 stoichiometry as measured by XPS. We also present significant contributions to the transient electrical modeling of double-heterojunction silicon solar cells. It is shown how these measurements can be used to determine both the front- and back-interface recombination velocities in fully fabricated solar cells. Physics-based modeling is combined with full device simulations to account for geometric effects, and the full results are applied to measurements of PEDOT/n-Si and PEDOT/n-Si/TiO_2 heterojunction solar cells, showing that emitter efficiency is an increasingly significant problem in these devices as the back interface is made more ideal. Next, a thorough account is given of a novel ALD-NiO deposition process, outlining how changes in deposition parameters change the content of the resulting films. A process is also developed to fabricate ALD-CuO, and the two are integrated to give, for the first time, Cu-doped NiO deposited by ALD. The copper is shown to increase film conductivity, an effect corroborated by spectroscopic band measurements. ALD-NiO is shown to depin the interfacial Fermi level in NiO/c-Si diodes; integrated into perovskite solar cells, it gives results comparable to the more standard solution-deposited NiO films, with a small V_{OC} improvement. Finally, the last chapter of this thesis leverages patent data to examine how work or collaboration across economic sectors influences research choices and output. First, it is shown that university-industry collaborations

tend to produce more valuable patents. Next, using a novel data set of those who have patented in each of the military, commercial, and academic sectors, it is shown that these peoples' least influential contributions tend to come when they do military work.

Acknowledgements

I hope and believe that, if the front matter of this thesis were removed, it would still unmistakably come from the research group of my advisor, Professor James C. Sturm. Your grounded, fundamentals-focused approach to research is a model worth emulating. Over the last six months, especially, I often found myself considering how you would think about a problem and how you might try to solve it. Looking a level higher, you gave me, repeatedly, that rare freedom in grad school: the ability to choose what I would research and how I would do it. Realistically, this meant I spent a lot of time pushing, with great effort, straight into dead ends, but it also meant that my successes were my own, and it has taught me planning and analysis skills that I would otherwise lack. Thank you for your faith.

I must also thank my committee members, Professors Claire Gmachl and Barry Rand, and my thesis readers, Professors Sigurd Wagner and Michael Oppenheimer. You all contributed in meaningful ways to my time here, and I'm glad you were able to be there at the end of it. Special recognition goes to Prof. Oppenheimer, who has also been my PEI-STEP advisor. You took a chance on a random ELE student who wanted to do a macroeconomic research project without having any training in the subject, and you whipped that project into shape when it needed it the most.

By all rights, this paragraph should have been at the top: thank you to the cleanroom and IAC workers. You were all, of course, essential to the actual content of this thesis—that much goes without saying—but I'm also thanking you for something more. As people I knew in the department have graduated and left, your company has been a real treasure, and I will genuinely miss our various chats. Joe Palmer, Eric Mills, Zuzanna Lewicka, Conrad Silvestre, and Bert Harrop—you've handled the ebbing and flowing absurdities of the cleanroom with more grace than I ever could have. You all deserve, uncompromisingly, the best. Nan Yao, John Screiber, Paul

Shao, and Yao-Wen Yeh—thanks for being around at whatever weird hours I broke things, and for never grumbling under your breath at my clumsiness.

My research collaborators always set a high bar. Sushobhan Avasthi, we didn't overlap for too long, but, then and now, I marvel at your professionalism and your knowledge. Ken Nagamatsu, it's incredibly lucky that the nicest guy I met at Princeton also happened to be my mentor in the lab. You taught me not just how to "do research" but also how to be a grad student. Janam Jhaveri, thanks for keeping me company once it was down to just us two—you helped me out when I needed it, and the lab was a lonelier place without you around. Girija Sahasrabudhe, you were always up to learn new things and try them out, and I'll always admire that. Gabriel Man, there was nobody better at explaining techniques to me, and at the same time you always treated me as an equal. Levent Aygun, you were as much a friend as a colleague, but for the purposes of these acknowledgements I'll simply thank you for your help and company, in the cleanroom and elsewhere. Ross Kerner and Hannah Smith, my ALD buddies, you were down in the trenches with me—thanks for your help and patience. Nakita Noel, I wish we'd had more opportunities to collaborate, but what we did try was efficient and fruitful, thanks to you.

To my parents and sister: you clearly never really got a grip on what exactly I was doing here or why I came, but you still tried your hardest to care about my progress. That this was purely a love-derived obligation makes it, if anything, more meaningful. Koushiki, you, on the other hand, understand all too well what this experience has been like. You have been the light at the end of this tunnel. Thank you.

Thanks also must go to my friends, who both distracted and commiserated with me. My happiness has been, in significant part, due to you. Finally, all the admiration and respect in the world for my organizing comrades, from PCP to DSA to PGSU. Your dedication has always astounded me. A better world is possible.

Much of the work presented here was funded by the DOE Sunshot Grant #DE-EE0005315 and by Princeton's PEI-STEP fellowship program. I hope I gave a reasonable return on the investment.

Contents

List of Tables

List of Figures

List of Symbols

Symbol	Definition	Units
A	Device area	cm^2
$A^\star$	Richardson constant	C $sec^{-1}cm^{-2}K^{-2}$
a	Contact radius	cm
BE	Binding energy	eV
d	Polaron size	cm
D_n,D_p	Electron and hole diffusion coefficient	cm^2/sec
C	Atomic fraction;	—
	Capacitance	F
C_0	XPS film-on-substrate ratio constant	—
E	Energy	eV
EA	Electron affinity	eV
E_A	Activation energy	eV
E_C	Conduction band edge	eV
E_G	Bandgap	eV
E_{HOMO}	Highest occupied molecular orbital level	eV
E_i	Energy level halfway through bandgap	eV
E_{LUMO}	Lowest unoccupied molecular orbital level	eV
E_{photon}	Photon energy	eV
EQE	External quantum efficiency	—
E_t	Trap energy level	eV
E_V	Valence band edge	eV
f	X-ray flux	photons/(cm^2· sec)
FF	Fill factor	—
h	Planck's constant ($4.14{\cdot}10^{-15}$)	eV·sec
I	Signal intensity	counts/sec
IE	Ionization energy	eV
J	Current density	A/cm^2
J_0	Diode saturation current density	A/cm^2
J_{dark}	Dark current density	A/cm^2
J_{for}	Forward-phase current density in RR	A/cm^2
$J_{for,front}$	Front-interface forward-phase recombination current density in RR	A/cm^2

$J_{\text{for,inj}}$	Forward-phase minority carrier-diffusion current density in RR	A/cm^2
J_{ph}	Photogenerated current density	A/cm^2
J_{rev}	Reverse-phase current density in RR	A/cm^2
$J_{\text{rev,back}}$	Back-interface reverse-phase recombination current density in RR	A/cm^2
$J_{\text{rev,front}}$	Front-interface reverse-phase recombination current density in RR	A/cm^2
J_{SC}	Short-circuit current density	A/cm^2
k	Boltzmann's constant ($8.62 \cdot 10^{-5}$)	$\text{eV}\cdot\text{K}^{-1}$
KE	Kinetic energy	eV
L_{n},L_{p}	Electron and hole diffusion lengths	cm
n	Diode ideality Factor;	—
	Electron density;	cm^{-3}
	Curve labeling index in Chapter 4	—
n_{atom}	Atomic density	cm^{-3}
n_{i}	Intrinsic electron density	cm^{-3}
N_{A}	Acceptor concentration	cm^{-3}
N_{D}	Donor concentration	cm^{-3}
N_{t}	Trap concentration (bulk);	cm^{-3}
	Trap concentration (interface)	cm^{-2}
P	Power	W
p	Hole density	cm^{-3}
PCE	Power conversion efficiency	%
p_{in}	Incident photon rate	$\text{cm}^{-3}\text{sec}^{-1}$
q	Electron charge ($1.6 \cdot 10^{-19}$)	C (Coulomb)
Q_{DSD}	Stored charge density in a single-sided device	C/cm^2
Q_{FB}	Forward-bias stored charge density	C/cm^2
Q_{Recov}	Recovered charge density	C/cm^2
Q_{SSD}	Stored charge density in a single-sided device	C/cm^2
R_{eff}	Area-normalized effective series resistance	$\Omega\cdot\text{cm}^2$
R_{S}	Series resistance	Ω
$R_{\text{S,A}}$	Area-normalized series resistance	$\Omega\cdot\text{cm}^2$
S	XPS sensitivity factor	—
s	Surface/interface recombination velocity	cm/sec
s_{back}	Back (minority carrier-blocking interface) interface recombination velocity	cm/sec
s_{eff}	Effective interface recombination velocity	cm/sec

s_{front}	Front (minority carrier-injecting interface) interface recombination velocity	cm/sec
sim	Jaccard patent similarity measure	—
T	Temperature	°K
t	Time;	sec
	Thickness	cm
T_{det}	XPS detection efficiency	—
t_{r}	RR recovery time	sec
V	Voltage	volts
V_{bi}	Built-in voltage	V
V_{D}	Device voltage	V
V_{OC}	Open-circuit voltage	volts
V_{r}	Applied reverse voltage in RR	V
v_{th}	Thermal velocity	cm/sec
W	Base/substrate thickness;	cm
	Film width	cm
x	Spatial coordinate	cm
y	Photoelectric conversion efficiency	—
α	Absorption coefficient	cm^{-1}
γ	Minority carrier injection efficiency	—
γ_{area}	Area efficiency factor	—
γ_{DSD}	Minority carrier injection efficiency in a double-sided device	—
γ_{front}	Front (minority carrier-injecting) interface injection efficiency	—
γ_{SSD}	Minority carrier injection efficiency in a single-sided device	—
δ	NMR shift	ppm
ϵ	Permittivity	F/cm
θ	XPS angular efficiency factor	—
λ	Wavelength	cm
λ_{mfp}	Mean free path	cm
μ_{n},μ_{p}	Electron and hole mobility	$\text{cm}^2/(\text{V}\cdot\text{sec})$
ν	Frequency	sec^{-1}
ρ	Resistivity	Ω·cm
σ	Conductivity	Siemens/cm
$\sigma_{\square}$	Sheet conductance	Ω^{-1}
σ_{cap}	Trap capture cross-section	cm^2

σ_{ph}	Photoelectric cross-section	cm^2
τ_{b}	Minority-carrier bulk recombination lifetime	sec
τ_{eff}	Effective minority-carrier lifetime	sec
τ_{s}	Effective surface recombination lifetime	sec
ϕ	Work function	eV
χ	Electron mean free path	cm

Chapter 1

Introduction

This is a thesis about semiconductors, thin films, and the devices you can make out of them—the high-tech stuff of the past seventy years. But that atomic-age sheen disguises an approaching catastrophe. The fossil-fuel plants that power our technical equipment and the vehicles that transport our research supplies are, in so doing, constantly contributing to the medium-term deaths and immiseration of many millions of people.

According to the UN, carbon emissions have already contributed about 1.0°C of global warming [1]. By about 2040, this number will be at 1.5°C, and, absent major policy changes, may hit 3.0°C, or even 4.0°C, by 2100. The UN's dryly-worded assessment contains horrors. "Negative impacts on average crop yields and increases in yield variability due to climate change"— famine and starvation. "Urban risks associated with housing and energy systems"— whole cities decimated by floods and hurricanes, or widespread grid failures affecting society's most vulnerable. "Displacement associated with extreme events"— millions of refugees seeking safety in the Global North, brutalized and turned away at increasingly-militarized borders. "Reduced access to water for rural and urban poor people due to water scarcity and

increasing competition for water"— "water wars," transported out of science fiction and into the real world.

The development of less expensive, more environmentally friendly, and more efficient solar cells will do absolutely nothing to stop this. Climate change is not the result of a technical failure, nor something that can be solved by being "smarter." It is the natural endpoint of the industrialized world's journey over the past two centuries. Humanity already has the technological tools it needs to beat back climate change; what it lacks is the political organization and will. Confronted by the First World War, Rosa Luxemburg wrote:

> Violated, dishonored, wading in blood, dripping filth— there stands bourgeois society. This is it [in reality]. Not all spic and span and moral, with pretense to culture, philosophy, ethics, order, peace, and the rule of law— but the ravening beast, the witches' sabbath of anarchy, a plague to culture and humanity. Thus it reveals itself in its true, its naked form. [2]

That clear accounting holds just as true for the long-anticipated climate tragedy that awaits us.

But soon, hopefully, society will, on a large scale, begin to implement the steps necessary to beat back climate change, and, when that day comes, scientists and engineers must be ready with the necessary tools. Better solar cells will not make the transformation *possible*, but they may well make it a bit easier. A 2015 study [3] laid out an ambitious roadmap for getting the US to 100% renewable-energy by 2050. It projects a total capital cost of $13.4 trillion. Assuming proportional costs, the cost for the utility and rooftop photovoltaics (PV) it recommends would be about $5 trillion. Material and device advancements enabling a 5% cost reduction, therefore, would be projected to save about $250 billion, and significant PV efficiency gains (realized, for instance, with silicon/perovskite tandem cells) could save even more.

PV-focused materials development could be important for reasons beyond direct costs, too. A robust transition could well require redundancy in PV material, with devices fabricated differently based on place- and time-fluctuating material availability and price. Putting all of society's eggs in a single PV-device-basket, in other words, should be avoided, and explorations of new materials and deposition techniques help diversify manufacturers' options.

Nor will future PV-related technological advancements necessarily occur without intervention. The modern American research scene has been called a "triple-helix" [4] of knowledge diffusion and collaboration between academia, private companies, and the government, and efficient technology development requires an understanding of these dynamics. Recognizing this, the recently-introduced "Green New Deal" resolution [5] says that "a Green New Deal must be developed through transparent and inclusive consultation, collaboration, and partnership with frontline and vulnerable communities, labor unions, worker cooperatives, civil society groups, **academia, and businesses**." Such collaboration may be formal—contracted collaborations between sectors— or less formal, naturally arising from personnel movement between various positions. Therefore, this thesis also explores these relationships, first studying how inter-sector collaboration affects research productivity and then examining the class of workers who have worked in all three parts of the triple-helix.

Solar cells are, compared to other semiconductor devices, rather simple: just stacks of different materials on top of one another (with some basic patterning added in at the end). The development of better (ie, less expensive to produce, or less material- or energy-intensive to produce, or more efficient) solar cells, therefore, is mostly a study of materials: how to make them themselves better, and how to make their interfaces behave the way we want them to. This thesis will present new work on both sides of that development: both the materials and the interfaces.

In Chapter 2, we introduce and analyze the silicon heterojunction solar cell (SHJC). SHJCs are likely to be less energy-intensive to manufacture and more efficient than traditional silicon pn-junction solar cells, but their effectiveness depends fundamentally on the quality of their contact interfaces. We overview the carrier mechanics and process of heterojunction interfaces and use a combination of analytical calculations and device simulations to relate those processes to solar cell quality metrics. We then review a number of interface and material characterization techniques that will be used later in the thesis.

In Chapter 3, we turn to titanium dioxide, TiO_2, and its application in SHJCs. Precise characterization of the TiO_2/Si interface is important for gaining an understanding of these cells. We first present a method for optically extracting the recombination characteristics of the Si/TiO_2 interface; a simple numerical model is used to relate that recombination to the measured external quantum efficiency of PEDOT/n-Si(/TiO_2) devices. We also experimentally investigate the use of TiO_2 as a passivating tunnel layer in SHJC anodes. Ultrathin TiO_2 films are deposited by atomic layer deposition and studied by optical, electrical, and spectroscopic methods, both by themselves and as part of test-device stacks. Even very thin TiO_2 films are shown to provide excellent passivation, but, for the thinnest films, that passivation is mostly not maintained after NiO overlayer deposition.

In Chapter 4, we present a method for the extraction of recombination characteristics via transient electrical measurements. The reverse-recovery analysis technique dates, in its simplest form, back to the 1950s, but here we extend it to the SHJC with non-ideal interfaces, including geometric effects. Front- and back-interface effects are considered separately, with an explicit procedure given for their deconvolution in the context of actual experiments. Numerical modeling and 2D device simulations are combined to give results in terms of reduced units that will be useful to researchers studying generalized SHJC configurations. The method is also applied to the specific

case of the PEDOT/n-Si/TiO_2 cell; we measure the reverse-recovery characteristics of those devices and from that derive values for the injection efficiency of the PEDOT/n-Si emitter and the recombination velocity of the n-Si/TiO_2 cathode.

In Chapter 5, we present the development and refinement of nickel oxide deposition via ALD. Nickel oxide is an increasingly-important material in organic and perovskite solar cells, but its deposition by ALD has barely been covered in the literature. The key process parameters and how they affect final film composition are reviewed. Spectroscopic and electrical measurements paint a consistent picture of the ALD-NiO both as-deposited and doped with copper. The ALD-NiO is also demonstrated to perform well as a hole-selective layer in a perovskite solar cell, increasing V_{OC} compared to a device made with solution-NiO.

In Chapter 6, we shift our focus from technical research to a meta-study on the research process itself, investigating how research collaboration between economic sectors, or personnel movement between those sectors, affects research output as measured by patents. It is shown that inter-sector collaboration seems to benefit research outcomes, but that it does not appear to have a measurable effect on which subsequent research choices are made. The military-technology dual-use claim is found to be false, with military work having less effect than comparable work done by the same individuals working in other sectors. However, a patent-similarity-based search for a mechanism behind this pattern is unsuccessful.

Finally, Chapter 7 reiterates the most important contributions of the thesis and reflects on some of the underlying themes that tie it together.

Chapter 2

Silicon Heterojunction Solar Cells: Interface Processes and Methods

2.1 Introduction

A generalized solar cell needs to contain only two features. First, there must be an absorber, some material in which light is absorbed and converted into charge carriers. In addition, there must be carrier-selectivity, some mechanism by which electrons and holes are separated and brought through opposite sides of the device. These are core; everything else simply enhances the processes enabled by those main features.

In traditional p-i-n and related solar cells, a strong electric field in the absorber drives charge separation. In others, such as in classic pn-junction solar cells, a localized electric field acts as a valve, setting up net carrier separation after a diffusion process. And sometimes the critical separation happens at an *effective* electric field brought about by band offsets. As Peter Würfel phrased it,"The direction of the short-circuit charge current in a pn solar cell agrees with the direction of [the junction] field. This seems to be sufficient to believe that it is also causing this current. To exaggerate somewhat, this is mere coincidence. It would be a completely unnecessary

restriction to exclude structures for solar cells in which no electric field [excluding *effective* electric fields] is present, but which have the membrane function incorporated." [6] This compelling physical insight finds its full expression in the silicon heterojunction solar cell (SHJC).

This chapter first introduces the SHJC, with special attention paid to the heterojunction interfaces that determine these devices' ultimate worthiness. The primary interfacial parameters are outlined, and guidelines are offered for approximately where these parameters should lie for reasonable performance. Next, a number of common analysis techniques are described; these methods will be used later in the thesis in the analysis of fabricated films and devices.

2.2 Silicon Homojunction Solar Cells

The standard silicon solar cell (Figure 2.1(a)) is based on a silicon homojunction, either p^+-n or n^+-p. Light absorption in the silicon generates free electrons and holes, which proceed to diffuse. At the p-n interface, a space-charge region, and thus electric field, is formed, and this field selectively sweeps holes one way and electrons the other, generating a photocurrent. The contacts on either side serve a somewhat minor role: they sometimes feature their own passivation scheme (such as a backside field), but conceptually are there just to collect carriers, and one can make efficient cells even using simple Ohmic contacts to the n-type and p-type sides of the semiconductor.

Absent light, the device's "dark current" density follows the standard diode equation $J_{\text{dark}} = J_0(e^{qV/nkT} - 1)$, where J_0 is the saturation current density, q is the electron charge, V is the applied voltage, k is Boltzmann's constant, n is an ideality factor that accounts for current mechanisms besides minority-carrier bulk diffusion, and T is the operating temperature. Dark current is, ideally, due purely to minority-carrier flow, either to the opposite side of the substrate or until bulk recombination, though

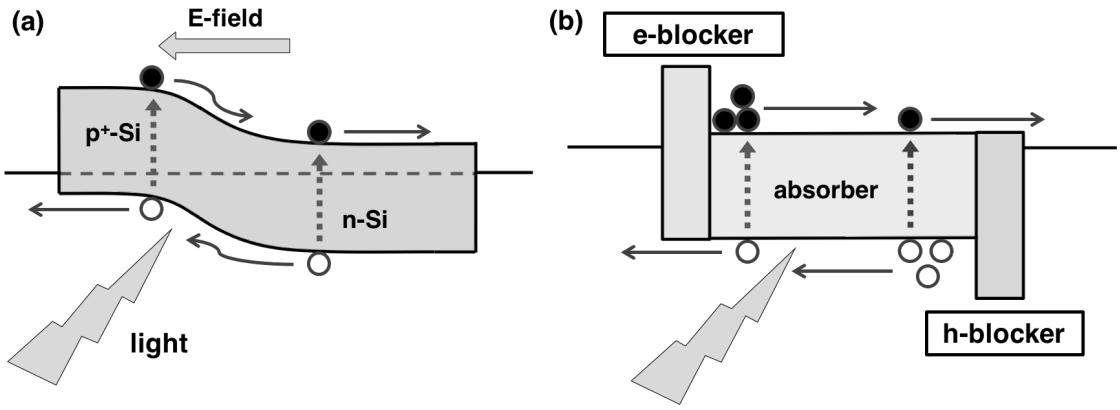

Figure 2.1: Conceptual schematic of illuminated solar cells under short-circuit condition. In the traditional silicon cell, (a), an electric field selectively drives carrier flow, while in the SHJC, (b), band offsets at the selective contacts drive carrier gradients. Metal contacts or the selective layers themselves may or may not induce space charge regions in the silicon.

majority-carrier effects like interface recombination can add additional terms. Illumination adds photogenerated current in the opposite direction to the dark forward current; with no voltage applied, this current is the short-circuit current density, J_{SC}. The full equation for the current density J of an ideal solar cell is just the sum of these two processes (Figure 2.2),

$$J = -J_{\mathrm{SC}} + J_0(e^{qV/nkT} - 1). \tag{2.1}$$

For some region of positive applied voltage, the total current density is negative, as is the power per unit area consumed $P/\mathrm{area} = J \cdot V$— in this range, power is being *generated*. Eventually, at $V = V_{\mathrm{OC}}$, the processes balance out and the total current density is zero. Solar cell development is really a process of trying to increase J_{SC} and V_{OC} while minimizing non-idealities such that the maximum value of P/area along the J-V curve is itself maximized.

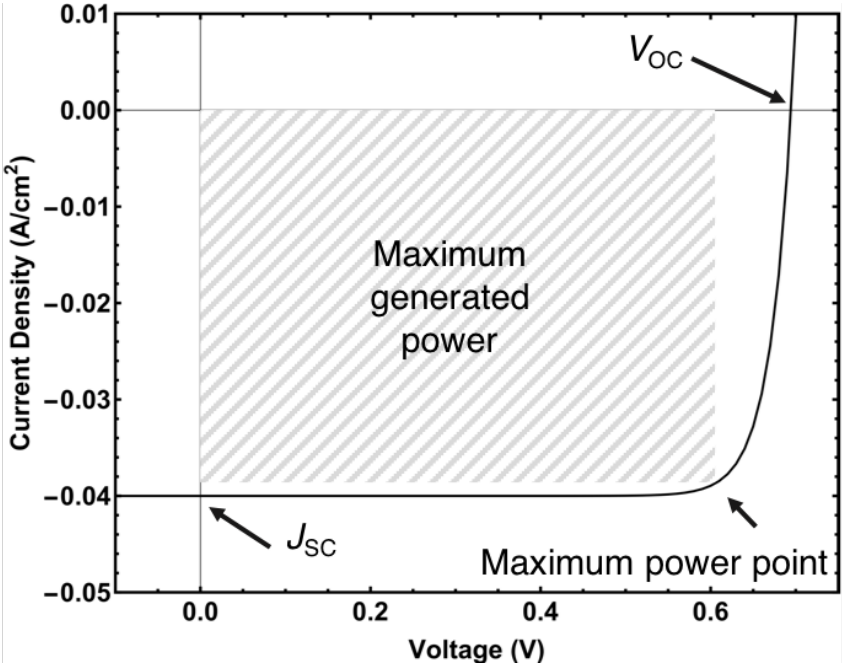

Figure 2.2: Basic solar cell current density-voltage plot, showing the short-circuit current density, the open-circuit voltage, and the visualized maximum power/area point product.

2.3 Silicon Heterojunction Solar Cells

SHJCs can offer the efficiencies of traditional silicon cells along with a decreased thermal budget and a potentially simpler manufacturing process. In these cells, the heavily-doped side of a typical p^+-n or n^+-p junction is replaced by an a-Si/Si [7], organic/Si [8],[9],[10], or metal oxide/Si heterojunction [11],[12] to create a carrier selective contact to lightly-doped crystalline silicon. This contact blocks majority carriers from the lightly-doped side from crossing to the contact in forward bias (thus greatly lowering dark current compared to a Schottky device), and the work function difference between the contact and the silicon creates an electric field that collects minority carriers generated by the sun. Ideally, this contact serves as a minority carrier injector into the silicon. On the other side of the device, at the Ohmic contact to the lightly doped semiconductor, a complementary heterojunction is used to block minority carriers in order to further reduce dark current. This effectively replaces the

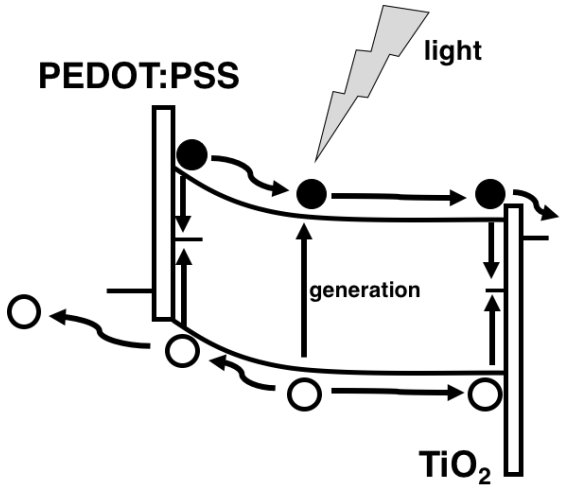

Figure 2.3: PEDOT/n-Si/TiO_2 solar cell in operating mode, as modeled in the Sentaurus simulations of this chapter. The included processes are carrier generation, drift/diffusion, and interface recombination. Net carrier flows are shown by arrows ("net" because, on the level of individual carriers, diffusion is omni-directional).

"backside field" implemented as an n/n^+ or p/p^+ junction. Such a device is shown schematically in Figure 2.1(b).

For much of this work, the model SHJC under study will be the PEDOT/n-Si/TiO_2 cell, shown in operating mode in Figure 2.3. PEDOT:PSS offers a large conduction band offset to silicon [13], making it an electron-blocker/hole-collector, and its highly p-doped nature (and resulting large work function) induces strong upward bandbending in the silicon, enhancing that carrier selectivity. TiO_2, meanwhile, has a large *valence* band offset with silicon, and a conduction band offset of only 0.1-0.2 eV, making it an effective hole-blocker/electron-collector.

2.4 Heterojunction Interface Processes

The interfaces between the silicon and the selective contacts are the core features of these devices, and their heterostructure nature determines many of the design priorities. The four primary under-illumination interface processes (for an electron-blocking interface) are shown in Figure 2.4; they are: (i) interface recombination, (ii)

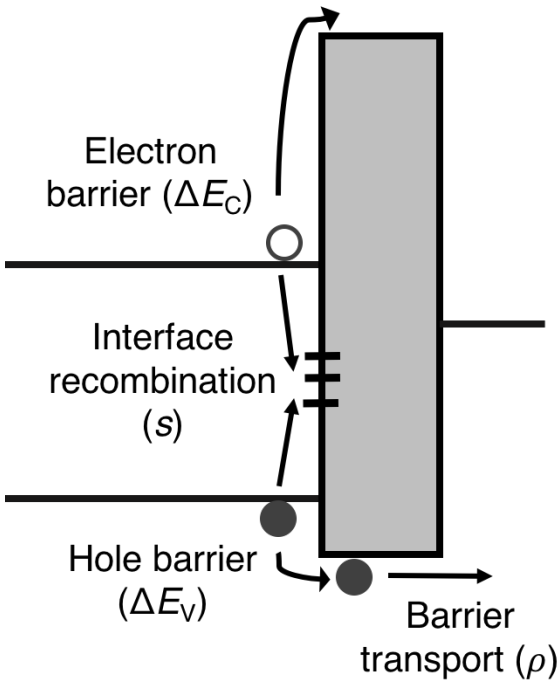

Figure 2.4: The four critical processes at a silicon heterojunction interface. The ideal electron-blocking interface would have low ΔE_V, ρ, and s, and high ΔE_C

and (iii) band offsets at the conduction and valence bands, and (iv) transport through the barrier layer. They are each discussed in a subsequent subsection. An effective barrier must meet minimum parameter requirements for each of these processes.

2.4.1 Interface recombination

In a crystalline silicon pn-junction cell, the driving field is at a homojunction between differently-doped pieces of silicon. The silicon crystal structure is maintained across the boundary, with no chemical bonds left unaccounted for. Silicon heterojunctions are messier. To see why, consider the nature of how chemical bonds give rise to semiconductor band structures (Figure 2.5). In a single atom, electrons occupy discrete energy levels. A covalent bond splits that level into two discrete levels, one higher than the initial level and one lower. In an infinite lattice, those split levels are expanded into continuous bands of possible electron states, with forbidden bandgaps

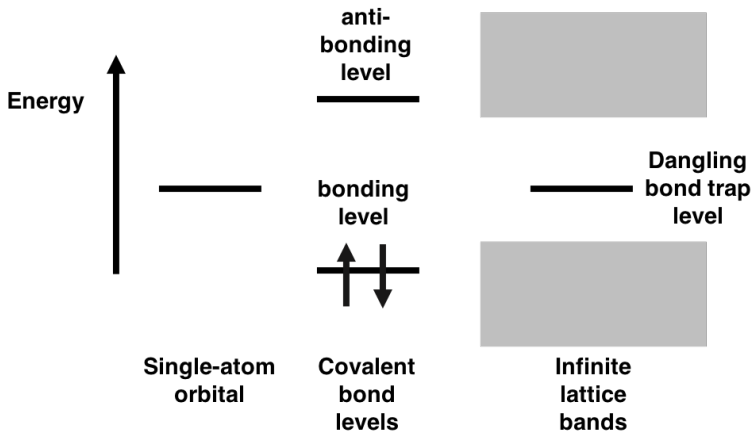

Figure 2.5: Energy-level diagram for single atoms, covalent bonds, and a crystal lattice. "Dangling bonds" in the lattice give rise to midgap states at approximately the unsplit energy level.

in between. At a heterojunction interface, though, the contrasting lattice constants and crystal structures mean that not every bond can be satisfied. The unsatisfied "dangling bonds" have energies similar to single-atom energy levels, meaning that they may lie in the midgap.

Midgap trap states are uniquely problematic for carrier recombination. Because trap states can provide or absorb differences in momentum necessary to enable momentum conservation in recombination processes, they enable high recombination rates. A process-balance approach to electron-hole recombination results in the following Shockley-Read-Hall (SRH) expression [14],[15] for recombination rate per unit volume (or, in the case of interfacial recombination, per unit area) U_{SRH}:

$$U_{\mathrm{SRH}} = N_t v_{\mathrm{th}} \sigma_{\mathrm{cap}} \Big[\frac{pn - n_{\mathrm{i}}^2}{p + n + 2n_{\mathrm{i}} \cosh(\frac{E_{\mathrm{i}} - E_{\mathrm{t}}}{kT})} \Big] \tag{2.2}$$

where p and n are the hole and electron concentrations, n_{i} is the intrinsic electron concentration, E_{i} is the energy level halfway through the bandgap, E_{t} is the trap energy level, N_t is the trap concentration (either cm^{-3} for bulk recombination or cm^{-2} at an interface), v_{th} (cm/s) is the carrier thermal velocity, and σ_{cap} (cm^2) is the trap

scattering capture cross-section. The cosh term causes recombination rates to peak when traps are close to the middle of the bandgap— exactly where we expect dangling-bond levels to appear. Making matters worse, the high concentration of dangling bonds at unpassivated heterojunction interfaces means a large N_t. Heterojunction interfaces in effective minority-carrier devices, therefore, simply must feature some kind of passivation.

It useful to consider the case of interface recombination under conditions of low-level injection in a doped semiconductor (say, n-Si). In that case, the n_i terms in Equation 2.2 drop out and the result is

$$U_\mathrm{SRH} = N_t v_\mathrm{th} \sigma_\mathrm{cap} \Delta p \tag{2.3}$$

$$\equiv s \Delta p \tag{2.4}$$

where Δp is the excess hole concentration and we have defined s, the interface/surface recombination velocity.

s is a key parameter in determining the quality of SHJCs, as this section will explore. The highest-quality SiO_2, SiN, and Al_2O_3 interfaces with float-zone lightly-doped n-Si have been reported to have $s \sim$ 1–5 cm/s [16], though a non-optimized thermal oxide growth will typically give $s \sim$ 100–300 cm/s. Wet-chemical SiO_2 growths put s in the 1000s of cm/s [17], and a completely unpassivated silicon interface (bare, say, or with a metal giving Ohmic contact) has s above 10^6.

To get an idea of how s impacts SHJC performance, 1D simulations were run using Synopsys Sentaurus, approximately modeling the PEDOT/n-Si/TiO_2 solar cell depicted in Figure 2.3. The substrate was a 300µm n-Si wafer with bulk lifetime 1ms and doping level $N_D = 10^{15}$ cm^{-3}. The anode electron-blocking layer consisted of a 70-nm-thick heavily-p-doped ($N_A = 3 \cdot 10^{20}$ cm^{-3}) layer with electron affinity 3.3 eV, bandgap 1.86 eV, and carrier mobility 20 cm^2/(Vs); on the other side, a 2 nm-

thick undoped oxide layer with electron affinity 3.95 eV, bandgap 4.6 eV, and carrier mobility 1 $cm^2/(Vs)$ served as a hole-blocker. The heavy p-doping and low valence band in the electron blocker induced approximately 0.75 eV of upward bandbending in the n-type silicon, while the cathode-metal work function was chosen so as to induce zero bandbending in the silicon.

The simulation parameters were relatively simple: SRH, Auger, and surface recombination were all included, but carrier tunneling was not, such that even the 2 nm-thick oxide was "impenetrable." With the AM1.5G spectrum incident on the anode, the Sentaurus software solved the coupled drift/diffusion and Poisson equations for electrons and holes as the voltage was swept and currents were recorded for a range of s values for the front and back interfaces. Because there were no antireflection layers/coatings besides that provided by the electron-blocking layer (which itself contributed to absorption), short-circuit current densities were generally less than 30 mA/cm^2. Optical generation was calculated by the software, using the AM1.5G spectrum and silicon's optical parameters (for simplicity, the PEDOT layer was given silicon's optical characteristics as well). The following discussion refers to the PEDOT/n-Si interface as the "front" side and the n-Si/TiO_2 interface as the "back" side.

The first notable result was that the solar cell's power efficiency was relatively insensitive to modulations of front-interface recombination velocity only. Even $s_{\text{front}} = 10^6$ only reduced the solar cells's V_{OC} by tens of mV compared to the best-case device— a measurable difference, to be sure, but not so bad considering just how extremely unpassivated this interface would be. This insensitivity, though, is particular to the specific situation being examined: it can be explained by the strong bandbending at this interface. With 0.75 eV of bending, the interface is in strong inversion, such that *electrons*, not holes, are the minority carrier, and the recombination velocity becomes a much smaller *effective* velocity modulated by the bandbending. Effective

passivation by this mechanism is called charge passivation; it is passivation that remains small even with a large number of dangling bonds at the interface, the one caveat being that a very high number of interfacial trap states can pin the substrate's bands, limiting the bandbending. Note that recombination, even in this case, is still proportional to N_t; the small carrier density, though, means that the $\Delta n \cdot N_\mathrm{t}$ product is small. Chaper 4 introduces an effective recombination velocity s_eff to capture this effect.

To give these simulation results more universal applicability, therefore, the following presentation and discussion talks in terms of not s_front but rather γ_front, the front-interface minority-carrier injection efficiency– not all heterojunction emitters will have the fortuitous bandbending enabled by PEDOT! A more thorough discussion of γ and how it depends on s_front, as well as an experimental method for measuring it in SHJCs, is given in Chapter 4.

The full simulation results are shown in Figures 2.6 and 2.7, from which it is clear that both front- and rear-interface properties can have significant effects on open circuit voltage and efficiency. Front-interface emitter efficiency has drastic effects throughout its range, which makes conceptual sense: as the device becomes less and less minority-carrier dominated (ie, as current becomes more due to interfacial recombination or majority-carrier injection), the dark current rises and the open-circuit voltage drops. Meanwhile, the rear-interface s has a strong effect in the range of $s < 1000$ cm/s; above that, recombination at that interface is so dominant that it essentially becomes non-selective. The case of $\gamma_\mathrm{front} = 1$, $s_\mathrm{back} > 10^5$ cm/s corresponds to that of a silicon solar cell without any backside field or passivation, while the case of $\gamma_\mathrm{front} = 1$, $s_\mathrm{back} = 5$ cm/s gives V_OCs approaching 0.7 V, near state-of-the-art in silicon solar cells. The simulated device's lack of antireflection coating/texturing, and thus relatively poor external quantum efficiency, accounts for much of this difference. Given ideal heterojunction selective contacts on the back (implying low s_back) and

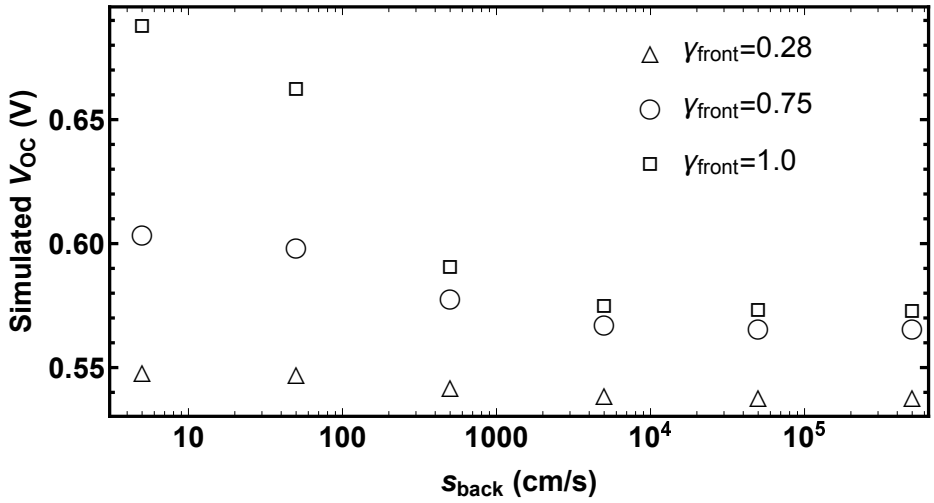

Figure 2.6: Simulated AM1.5G V_{OC} for a silicon heterojunction solar cell as a function of s_{back} and γ_{front}.

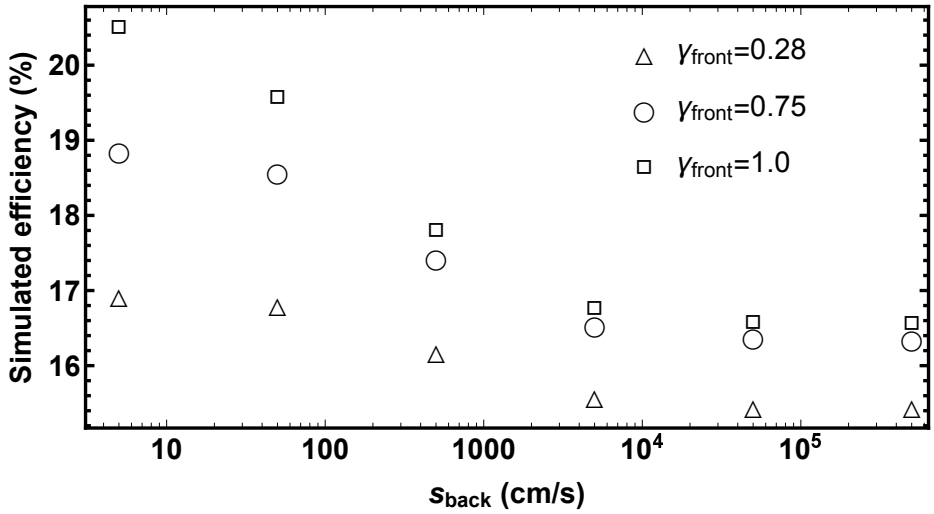

Figure 2.7: Simulated AM1.5G efficiencies for a silicon heterojunction solar cell as a function of s_{back} and γ_{front}.

front (implying $\gamma = 1$), the limiting process ends up being Auger recombination. These results make clear how important s is in the performance of SHJCs.

2.4.2 Band Offsets

In the absence of space-charge regions, heterojunction band offsets are the driving forces of carrier separation in SHJCs. An interface with a change in the conduction-

band level contributes to the total open-circuit voltage in the amount [18]

$$|\Delta V_{\mathrm{OC}}| = |\int \Big(\frac{q\mu_{\mathrm{n}}\Delta n}{\sigma}\Big)\frac{dE_{\mathrm{C}}}{dx}dx| \tag{2.5}$$

where E_{C} is the conduction-band energy level, Δn is the excess electron density, σ is the conductivity, and μ_{n} is the electron mobility. Thus band offsets contribute to photovoltage even in the absence of a true electric field. An electron-blocking interface, then, ought to have a large ΔE_{C} and, if one wishes it to allow transport of holes, a small ΔE_{V}. The ΔE_{C} requirement was previously examined in modeling by Jhaveri [19], who found that a useful barrier should have an offset of at least 400 meV. Presumably, candidate blocker materials can be chosen on the basis of meeting that minimum, but then the ΔE_{V} requirement must simultaneously be met.

For the electron-blocker, ΔE_{V} issues will not affect dark-current characteristics. This is because electron-blocking anodes in the dark feature hole current *into* the device (ie, from the blocking material into the silicon), and, ignoring possible metal-to-oxide charge injection barriers, even a significant positive ΔE_{V} presents no barrier to holes moving in that direction. Under illumination, though, charge extraction requires holes to flow from the silicon into the blocker; in that case, these band offsets matter.

Several SHJC Sentaurus simulations, similar to those in Section 2.4.1, were run in order to examine the effect of undesired band offsets. s at both heterojunction interfaces in an n-Si SHJC was set at 100 cm/s and the backside ΔE_{C} varied. As shown in Figure 2.8, an undesired barrier of 100 meV has almost no effect on the simulated illuminated J-V characteristic, but at around 200 meV the effect becomes significant (the "s-shaped" curve displayed is characteristic of parasitic barriers in solar cells). This, then, is an approximate upper limit on how much of the "wrong" band offset can be tolerated in SHJCs.

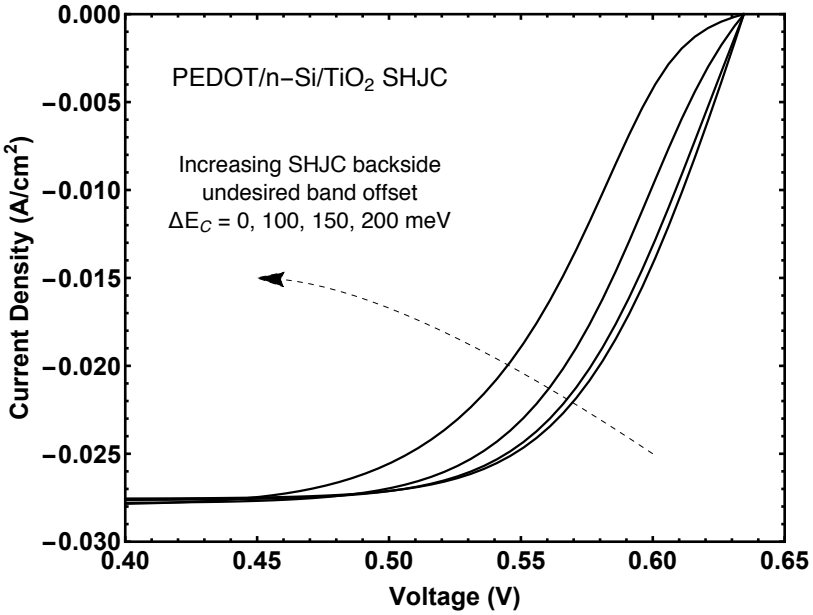

Figure 2.8: Simulated AM1.5G $J-V$ characteristics of a SHJC with varying backside (electron-selective) undesired band offset. At around $\Delta E_{\mathrm{C}} = 200$ meV the curve becomes noticeably s-shaped, leading to an efficiency loss. Recombination parameters were set at $s_{\mathrm{front}} = s_{\mathrm{back}} = 100$ cm/s, with other simulation details given above in Subsection 2.4.1.

2.4.3 Blocker Transport

The fourth important process inherent in the carrier-selective contact is carrier transport through the blocker itself. That layer contributes to total cell series resistance, and too-high values affect the cell's illuminated current-voltage characteristics. Selective-contact layers are typically thin, of course, so they have to be quite resistive for their resistance to matter. However the wide-bandgap oxides that constitute many of the candidate materials for these layers can indeed be very resistive.

This thesis focuses mostly on thin TiO_2 and NiO as blocking layers. In that context, one can consider a 10 nm-thick oxide layer as part of a cell that otherwise has no significant series resistances. The current density-voltage equation for such a cell (including photocurrent), assuming ideality otherwise, is

$$J = J_{\mathrm{SC}} - J_0 \exp\Big[\frac{q(V + (J \cdot A)R_{\mathrm{S}})}{kT}\Big] \tag{2.6}$$

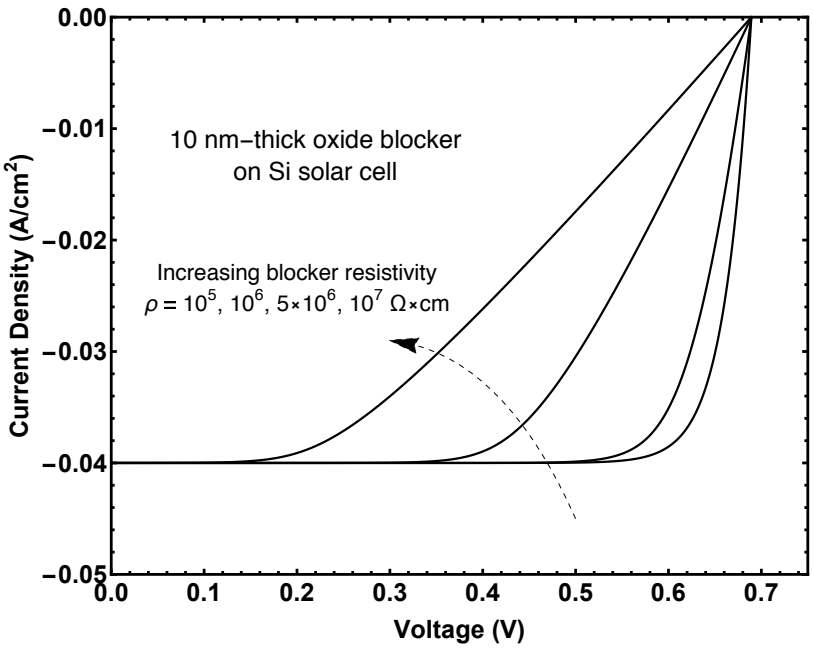

Figure 2.9: Silicon solar cell illuminated (AM 1.5G) current density-voltage characteristics for a device with a 10 nm-thick blocking layer contributing to the series resistance as in Equation 2.6.

where J_{SC} is the short-circuit photocurrent density, J_0 is the diode dark saturation current density, A is the device area, and R_S is the series resistance in Ω. Assuming a 10 nm-thick oxide layer, this equation was solved to give the J-V curves shown in Figure 2.9, using photocurrent density 40 mA/cm^2 and $J_0 = 10^{-13}$ A/cm^2. As is evident from the curves, for this oxide thickness the resistivity starts to become a problem at around $\rho = 10^6$ Ω·cm (varying the thickness changes this value proportionally, and other device resistances also being present would mean this requirement would get stricter). This value is less than would be expected for most intrinsic wide-bandgap oxides, but the presence of defects or dopants in the oxide can reduce that resistivity, making oxide layers feasible.

At this point, it is possible to combine the just-developed material parameters into a concise table (Table 2.1) giving the broad requirements for SHJC blocker materials at either side of the cell. These numbers are only approximate, and in the end the list of useful materials will depend not only on these but also on film thickness and

Parameter	Requirements
Effective s	$s \sim 100$ cm/s
Resistivity	$\rho < 10^6$ Ω·cm
Blocking offset	$\Delta E > 0.4$ eV
Passing offset	$\Delta E < 0.2$ eV

Table 2.1: Approximate requirements for SHJC selective-contact materials

morphology, optical considerations, etc, but they at least are a useful starting point for analysis.

2.5 Interface and Material Characterization Techniques

This thesis characterizes interfaces and materials using a variety of methods. Some, like current-voltage measurement, need no elaboration. The reverse recovery method, by contrast, is given an entire chapter of development (see Chapter 4). Here, a few less-standard techniques are reviewed in broad strokes, with further details given later as necessary.

2.5.1 Quasi-Steady-State Photoconductivity Decay

Quasi-Steady-State Photoconductivity Decay (QSSPCD) is a contactless method for the estimation of minority-carrier lifetime and thus, indirectly, interface quality. During a QSSPCD measurement [20], a silicon sample (which may or may not have films on its top/bottom) is exposed to a pulse of light that varies over the course of several ms. This pulse excites electrons and holes, generating a $\Delta n = \Delta p$ that temporarily increase the sheet conductance of the sample according to $\Delta\sigma_\square = qW(\Delta p)(\mu_n + \mu_p)$, where $\sigma_\square$ is the wafer sheet conductance in Ω^{-1}, W is the sample thickness, and μ is the carrier mobility. An inductive coil positioned below the sample measures this

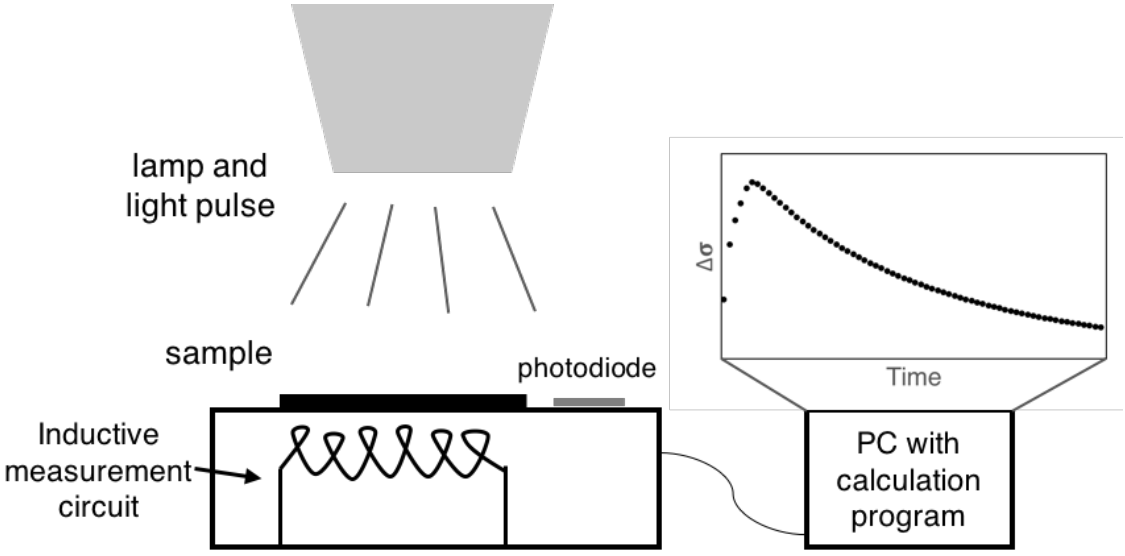

Figure 2.10: Quasi-Steady-State Photoconductivity Decay (QSSPCD) apparatus. This contactless technique allows for determination of minority-carrier lifetime in uncontacted devices.

time-dependent conductivity. Simultaneously, a photodiode next to the sample tracks the light intensity over time. The experimental setup is shown in Figure 2.10.

If the carrier lifetime is significantly less than the pulse variation time, then the system can be taken to be in steady-state at any given moment. Given this, one can calculate the effective carrier lifetime with

$$\tau_{\text{eff}} = \frac{\Delta\sigma_{\square}}{J_{\text{ph}}(\mu_n + \mu_p)}, \tag{2.7}$$

where J_{ph} is the photogenerated current density, which can be estimated using basic optical characteristics. Note that the mobility used in this calculation is the *average* mobility during the light pulse—this introduces a small element of estimation.

This effective lifetime, for low-level photocarrier excitation, is composed of two separate elements: a bulk recombination lifetime τ_{b} and a surface recombination component that depends on the passivation of the two interfaces. It can be shown [21],[22] that these two elements are related to the effective lifetime by $\frac{1}{\tau_{\text{eff}}} = \frac{1}{\tau_{\text{b}}} + a_1^2 D$, where a_1 is the solution to a transcendental equation that includes the relative

interface recombination properties as well as the minority-carrier diffusion constant. Looking only at the fundamental mode of the conductivity decay, one can alternatively say that $\frac{1}{\tau_{\rm eff}} = \frac{1}{\tau_{\rm b}} + \frac{1}{\tau_{\rm s}}$, where $\tau_{\rm s}$ is an effective surface lifetime.

One can think of $\tau_{\rm s}$ as including two processes: optically-generated carriers must diffuse to the interfaces, and they must then recombine. In the case where the two silicon interfaces are identical, a useful approximation [23] is

$$\tau_{\rm s} = \frac{W}{2s} + \frac{1}{D}(\frac{W}{\pi})^2 \tag{2.8}$$

where W is the wafer thickness, D is the diffusion coefficient, and s is the surface recombination velocity. Equation 2.8 is used throughout this thesis in the calculation of s for various systems. Note also that high-lifetime float-zone wafers often have bulk lifetimes $\tau_{\rm b} > 1$ ms, while a reasonably good s value of 100 cm/s leads to a $\tau_{\rm s} \approx 100$ μs; thus, in such systems, it is often reasonable to make the approximation $\tau_{\rm eff} \approx \tau_{\rm s}$.

2.5.2 X-ray Photoelectron Spectroscopy

X-ray photoelectron spectroscopy (XPS) is a useful and versatile method of material characterization. Properly interpreted, XPS spectra reveal not only the elemental composition of a sample but also those elements' chemical states/valency. This makes the technique indispensable for analysis of novel materials or materials deposited via new methods.

During an XPS measurement, the sample is irradiated with x-rays, kicking off electrons. These electrons are collected and their kinetic energy measured, giving a count rate versus KE plot that serves as the primary output. Each photon kicks off one electron with resulting kinetic energy $KE = E_{\rm photon} - BE$, where BE is the electron binding energy. Because these x-rays have energies greater than 1 keV (the Al K-alpha X-rays used in this thesis have $KE = 1.487$ keV), the electron states

that get probed are the core levels; each occupied level corresponds to a peak in the XPS spectrum. The measured kinetic energies depend on the incident photon energy and the detector work function, and accounting for those allows the extraction of electron binding energies. Possible surface charging can be alleviated by the use of a "flood gun" which supplies a steady flow of low-energy electrons to the sample, compensating for the x-ray-induced electron loss.

Chemical interactions shift electron peak locations or add new peaks; comparison with literature or theory allows for the determination of the chemical state. Peaks can also arise from other processes: "shake-up" peaks occur when an incoming photon excites an ion to some characteristic ion energy before then using its remaining energy to eject an electron; this electron has less kinetic energy than it normally would (due to the ion excitation), and thus one observes a peak a few eV above the primary peak on a binding energy scale. It should be noted that XPS is a surface technique: while x-rays themselves penetrate deep into a material, ejected electrons can only escape from the top 10 nm or so. This surface sensitivity is often an asset rather than a drawback, though, in that it allows for the probing of thin films. Furthermore, angle-dependent XPS can give information about composition as a function of depth without the need for destructive techniques (although ion milling for depth profiling *is* often combined with XPS).

The most basic use of XPS is for basic elemental-composition information; a more advanced use is for chemical state determination; and careful peak-fitting can allow for quantitative composition determination. The intensity of the signal (ie, the background-subtracted area under the XPS counts-per-second curve) from a given peak is theoretically predicted to be [24]

$$I = n_{\mathrm{atom}} f \sigma_{\mathrm{ph}} \theta y \lambda_{\mathrm{mfp}} A T_{\mathrm{det}} \tag{2.9}$$

where n_{atom} is the atomic density (in cm^{-3}) of the element responsible for the signal, f is the x-ray flux, σ_{ph} is "the photoelectric cross-section for the atomic orbital of interest," θ is an instrument-dependent angle factor, y is the photoelectric conversion efficiency, λ_{mfp} is the mean free path of photoelectrons in the sample, A is the sample area subject to illumination/collection, and T is the detection efficiency. While the interaction cross-section and mean free path vary from material to material, "the ratio of the two quantities [between two materials]... remains nearly constant." [24] Thus, the various terms in Equation 2.9 can be condensed into a material- and instrument-specific sensitivity factor S, such that the ratio of the atomic proportions of two elements 1 and 2 can be written as

$$\frac{n_1}{n_2} = \frac{I_1/S_1}{I_2/S_2}. \tag{2.10}$$

Generalizing to more than two signal contributors, the atomic fraction of constituent x is a sample is

$$C_x = \frac{I_x/S_x}{\Sigma I_i/S_i}. \tag{2.11}$$

Inelastic electron scattering is also present in the x-ray interactions, and it contributes to the background that exists in XPS spectra. Several methods exist for defining this background; this thesis uses a modified Shirley background. In a Shirley background for a given energy-region of interest [25], the background intensity at any given energy is proportional to the areas of the peaks at lower binding energies, adjusted to account for the magnitude of the signal at the edges of the given energy range; an iterative procedure adjusts peaks and background to make the Shirley relation true. The *modified* Shirley method used here adds the condition that the background can at no point be higher than the actual data in the region of interest.

Matters are made more complicated when multiple peaks stem from the same species. Transition metals like Ni and Cu, for instance, exhibit both significant spin-orbit splitting and the aforementioned shake-up peaks. In that case, one must include shake-up peaks in the area measurement, and S must be chosen to correspond to the particular peak included—that is, there is one S for just the $2p_{3/2}$ peaks, another for just the $2p_{1/2}$ peaks, and a third S for analyses based on the total area.

2.5.3 Ultraviolet Photoelectron Spectroscopy

XPS measurements identify the discrete core levels of materials, but the electrons most important for conduction are the low-binding-energy electrons. These valence levels can be probed with Ultraviolet Photoelectron Spectroscopy (UPS). Such measurements use ultraviolet light (He I 21.2 eV light, in the case of this study) to eject electrons from the material under study's valence band and reconstruct the density of states. For the purposes of this study, the important experimental outputs are the material ionization energy (valence band edge location relative to vacuum) and work function.

A typical UPS spectrum (Figure 2.11(b)) contains two signal-sources. First, there are the primary electrons, those that are ejected elastically directly from the material into the detector. There are also secondary electrons, which form a broad, generally featureless spectrum caused by various inelastic processes inside the material. That spectrum is continuous until the electrons in it do not have enough energy to escape the material; at that point on the energy scale, called the cutoff, there is a sharp decline in measured carriers.

The analysis of a metal UPS spectrum can be understood as follows (see Figure 2.11), following [26]. Take the binding energy BE to be zero at the sample Fermi level, and denote the sample work function ϕ_s. Electrons, then, are ejected into the vacuum with kinetic energy $KE_{\text{initial}} = h\nu - BE - \phi_s$. They then accelerate towards

the detector, which has its own work function ϕ_{d}, such that their observed kinetic energy is $KE_{\mathrm{observed}} = h\nu - BE - \phi_{\mathrm{s}} + (\phi_{\mathrm{s}} - \phi_{\mathrm{d}}) = h\nu - BE - \phi_{\mathrm{d}}$.

ϕ_{d} is not known, but (for metals) we can easily observe the $BE = 0$ point on the detector output and change from a KE_{observed} scale to a KE_{surface} scale, which measures primary electrons' energy right before they leave the sample. Electrons right at the Fermi level *started* at the sample surface, so they have $KE_{\mathrm{surface}} = h\nu$. Note that this transformation holds across multiple electrically-connected samples, since the previous KE_{observed} calculation showed that that quantity does not depend on the sample work function.

Given this shift, what is the KE_{surface} of electrons that only barely reach the detector— ie, the electrons at the cutoff edge? Well, they must reach the sample surface with just enough kinetic energy to escape: ϕ_{s}. That is, on this scale, the cutoff energy is $KE_{\mathrm{cutoff}} = \phi_{\mathrm{s}}$. This reproduces the standard expression

$$\phi_{\mathrm{s}} = h\nu - (KE_{\mathrm{Fermi}} - KE_{\mathrm{cutoff}}), \tag{2.12}$$

but note that physically that expression is justified by the explanation given here— by itself it is just a shortcut.

Practically, the valence-band energy or cutoff energy are identified by fitting a line to the spectrum in that region of the data and finding the intersection of that line with the background signal, generally also fit by a line. Metal energy bands are filled right up to the Fermi level, and so that point is identified by finding the midpoint of the Fermi-Dirac density-of-states falloff in that region. Here, the cutoff region was fit to the point of steepest slope.

The calculation of the ionization energy of a semiconductor follows the same steps as the metal work function calculation, with ϕ_{s} in Equation 2.12 replaced by the ionization energy IE. Semiconductor work function, meanwhile, can be measured

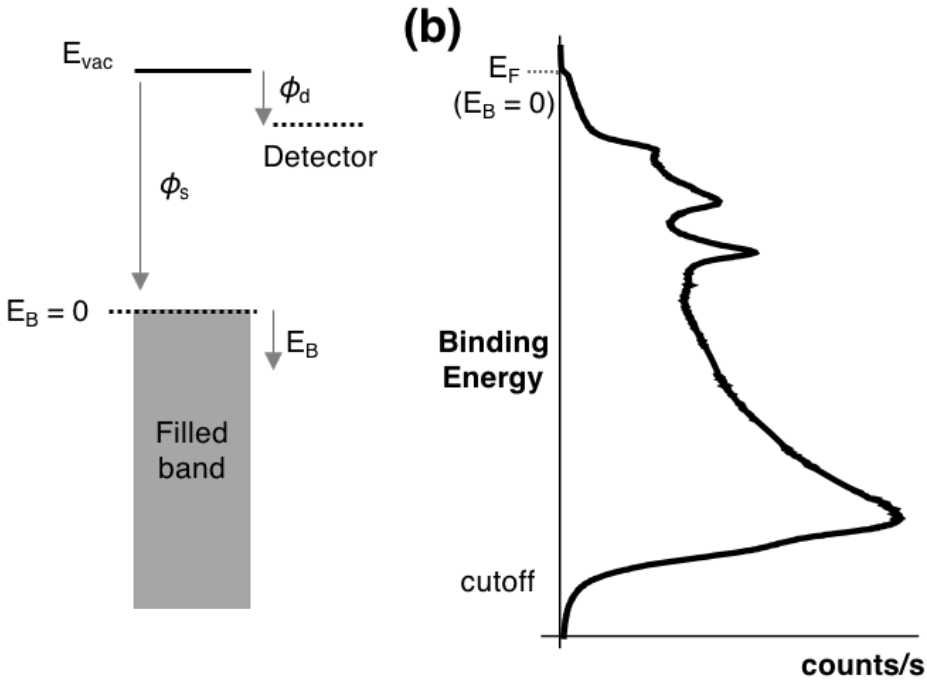

Figure 2.11: a) Conceptual plot of a UPS measurement and b) typical metal UPS spectrum (in this case, gold) with Fermi level and cutoff region labeled.

by placing the semiconducting sample in good electrical contact with a metal: that metal's Fermi level defines the Fermi level for the system and can serve as a reference for placing the semiconductor Fermi level relative to the valence band edge.

2.6 Conclusion

The SHJC is conceptually elegant, but strong practical requirements must be met for efficient operation. The work of developing these cells is first a process of candidate material identification and then an extended period of fabrication/process development. This chapter took a systematic approach, examining the interfacial processes that most determine heterojunction solar cell performance. Based on considerations of these processes, the chapter then outlined the most important material parameters to consider when choosing materials for SHJC experimentation. Simple models were used along with numerical simulations to set useful limits on these parameters.

Finally, a number of experimental techniques were briefly outlined— these can be referred to in the following chapters when experimental data is presented.

Chapter 3

Characterization of Recombination Rate and Structure at Crystalline Silicon/TiO_2 and Crystalline Silicon/TiO_2/NiO Interfaces

3.1 Introduction

Titanium dioxide is a prominent electron transport material/hole-blocker in solar cells of types from silicon [27],[28],[29] to InP [30] to organic [31] to perovskite [32]. With an electron affinity of about 4 eV and a bandgap of about 3.0 to 3.3 eV, depending on structure [33], the material's band characteristics are broadly favorable for electron transport/injection and hole-blocking.

TiO_2's effectiveness in silicon devices is particularly interesting. ALD-TiO_2 has been shown to form interfaces with s down to 14 cm/s on p-Si for 1.5 nm films [34], and few-nm annealed CVD-TiO_2 films demonstrate s with n-Si below 100 cm/s [19]. The exact mechanism(s) of this remarkably good passivation are still a matter of

investigation. Sahasrabudhe et al. [35] provided convincing evidence for the formation of Si-O-Ti bonds at the CVD-TiO_2/Si interface, implying some level of chemical passivation, and Man et al. [33] show the preservation of H-termination on the silicon surface even after 100°C CVD-TiO_2 deposition. At the same time, there is clearly some element of charge-passivation occurring, as well, since TiO_2 is able to enhance passivation quality even when separated from the silicon by a thin SiO_x layer [34].

TiO_2 has already been incorporated into promising silicon solar cell designs. Research here at Princeton [29] demonstrated it as a standalone selective contact on the backside of SHJCs, as shown in Figure 3.1. In these devices, the TiO_2 mostly acts to decrease the dark current by decreasing the gradient of hole density across the substrate (serving as a hole-blocker); see Section 4.4 for a fuller discussion of this. There is also a small positive J_{SC} effect due to the hole-blocking, which will be exploited in Section 3.3.

The material has also been incorporated into HIT cells [36] that reached efficiencies of over 20%, though those devices also included a number of other performance-enhancing features. Various researchers [37],[38],[39],[40] have also suggested and modeled TiO_2 as an ETL in potential perovskite/c-Si tandem cells. Studies of the TiO_2/Si interface, therefore, can inform current research with short- to medium-term application.

This chapter presents two distinct studies of the TiO_2/Si heterojunction. In the first, external quantum efficiency (EQE) measurements of completed TiO_2-containing SHJCs are used, alongside modeling, to extract an effective s for the interface. In the second, the use of ultrathin ALD-TiO_2 as a potential passivating tunnel layer in selective contacts is investigated. We find that even very thin TiO_2 can passivate silicon, but also that the correct balance of thickness and low resistance necessary for proper passivating tunnel action will be difficult to achieve.

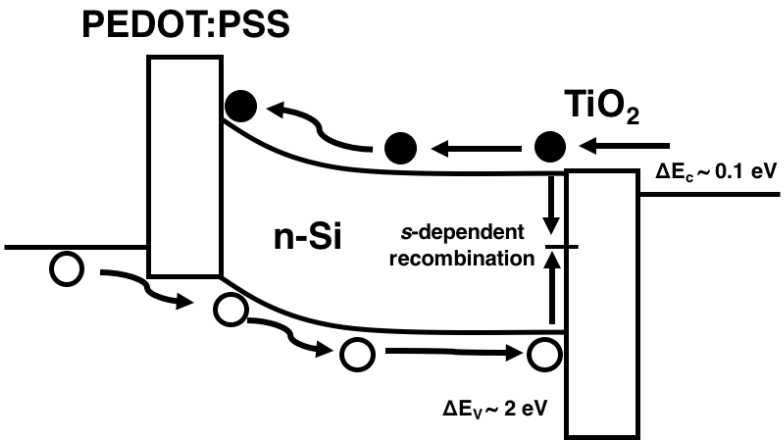

Figure 3.1: Bands and carrier flows in a PEDOT/n-Si/TiO_2 SHJC, shown for forward bias without illumination, neglecting front-interface non-idealities. The TiO_2's band offsets with silicon cause it to block holes, decreasing the minority-carrier-diffusion dark current.

3.2 Previous Work on TiO_2 on Crystalline Silicon As a Passivating Hole Blocker

As mentioned in the introduction, TiO_2 has already been demonstrated to be an effective hole-blocker/electron-transport-layer on crystalline silicon. Researchers in our group at Princeton pioneered the use of CVD-TiO_2 formed by thermolysis of *tert*-butoxide[28],[29], though CVD using other precursors has also been shown to passivate c-Si [41]. More recently, ALD-TiO_2 has emerged as a common choice [42],[43],[36]. Other methods, including in-situ annealing of e-beam deposited Ti films, have also worked [44].

The analysis in Chapter 2 suggested that selective contact layers should have surface recombination velocity $s < 100$ cm/s, and this standard has been met by many of the methods just indicated. Annealed CVD-TiO_2 on silicon achieved $s \sim 75$ cm/s [45]. An alternative CVD method was reported to give $s = 30$ cm/s [41]. As-deposited ALD-TiO_2 on n-Si gave $s = 38$ cm/s, an excellent value that dropped to 11 cm/s upon annealing and 2.8 cm/s after light-soaking [46]; these latter two value

approach the highest-quality passivation possible. In another study, ultrathin ($\sim$ 1.5 nm-thick) ALD-TiO_2 had s with silicon of $\sim$ 100 cm/s before annealing and $\sim$ 100 cm/s after [34].

These positive interface-quality results have led to high-quality devices. The addition of a TiO_2 layer increased V_{OC} by 45 mV in PEDOT/n-Si/TiO_2 devices compared to PEDOT/n-Si devices [45]. p^+/n silicon cells with selective back ALD-TiO_2 contact achieved PCE= 19.8%, indicating significant passivation over the metal-contact case; however, these devices were still worse than p^+/n/n^+ cells, a difference attributed to n-Si/TiO_2 recombination [42]. In another study, ALD-TiO_2 was integrated into the cathodes of doping-free HIT cells [36]. The thermal stability of the cells was improved without a significant increase in contact resistivity.

3.3 Recombination Velocity via Wavelength-Dependent External Quantum Efficiency Measurement

QSSPCD (see Section 2.5.1) is useful for measurements of simple film/substrate structures, and the reverse-recovery procedure detailed in Chapter 4 is an electrical technique that can give more detailed information in completed devices. This section outlines an *optical* technique for interface recombination velocity (s) measurement. Such a method could be useful in the case of optically-enhanced passivation (light-induced trap filling), or when EQE measurements have already been taken and researchers want to leverage them to extract as much information as possible. After a qualitative explanation and quantitative overview, the technique is applied to the n-Si/TiO_2 backside junction of a fully-fabricated SHJC. The core of this section was previously published in [29].

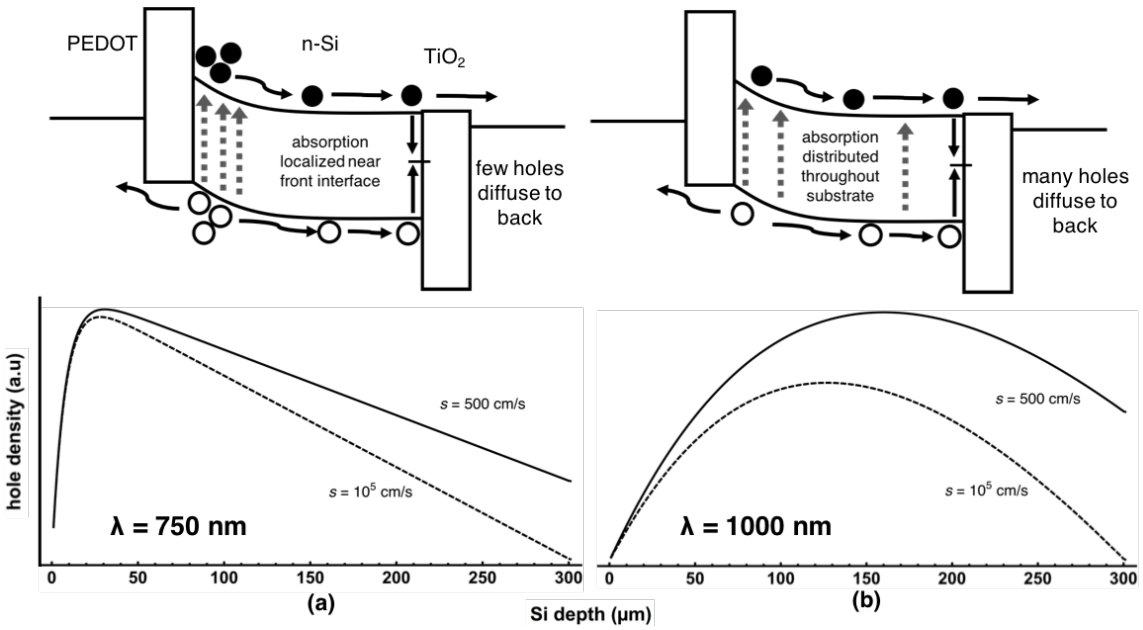

Figure 3.2: Hole profiles and short-circuit band schematics, for two different s_{rear} values at the Si/TiO_2 interface, created by light of wavelengths (a) 750nm and (b) 1000nm incident on a 300μm-thick n-Si heterojunction solar cell with ideal front interface. Profiles have been normalized by incident photon intensity at each wavelength to elide differences in AM1.5G illumination intensity and absorption efficiency.

In an EQE measurement (Figure 3.3), solar cell short-circuit current is measured as a function of incident light wavelength. EQE at a given wavelength λ is defined as

$$\text{EQE}(\lambda) = \frac{\text{rate of collected photogenerated carriers at } \lambda}{\text{rate of incident photons at } \lambda}. \tag{3.1}$$

EQE measurements can expose differences in device optical characteristics— for example, cells with antireflection layers should, other things held equal, have higher EQEs. Or, if optical characteristics are held fixed, EQE data can expose device-to-device differences in carrier generation and collection. This latter application is what will be exploited in this section.

3.3.1 Effect of Backside s on Photocarrier Collection

Because photogenerated minority carriers that recombine at the SHJC rear interface do not contribute to photocurrent, we can use measured photocurrent to evaluate the interface quality. Consider light incident on a double-sided silicon heterojunction cell with substrate thickness W, treated in one dimension, with only the silicon itself directly modeled, the selective contacts being incorporated as boundary conditions. Such a device might have the band structure depicted in Figure 3.2; the "front" side is the hole injector/collector, and the "back" or "rear" side is the electron collecting junction. Light with absorption coefficient $\alpha(\lambda)$ is absorbed at depth x with probability $P(x) = e^{-\alpha x}$, so, for constant illumination, we have generation rate $G_{\text{photo}}(t) = p_{in}e^{-\alpha x}$, where p_{in} is the rate of incident photos at the given wavelength. After absorption, generated holes, with excess concentration Δp, diffuse according to the standard diffusion equation $\frac{\partial \Delta p}{\partial t} = D_p \frac{\partial^2 \Delta p}{\partial^2 x}$, where D_p is the diffusion coefficient for holes; the complete equation for this system is thus

$$\frac{\partial \Delta p}{\partial t} = p_{in}e^{-\alpha x} + D_p \frac{\partial^2 \Delta p}{\partial^2 x}. \tag{3.2}$$

Assuming a hole-selective PEDOT/Si interface with zero recombination, we have at that interface the condition $\Delta p = 0$; all excess holes that reach this interface are collected and contribute to photocurrent. Finally, at the rear Si/TiO_2 interface with finite recombination, we have

$$D_{\text{p}} \frac{\partial \Delta p}{\partial x} = s_{\text{rear}} \cdot \Delta p, \tag{3.3}$$

where s_{rear} is the Si/TiO_2 interface recombination velocity. Any holes that recombine here represent a loss of potential photocurrent.

Photons of different wavelengths have different α values and thus different absorption profiles. This difference in absorption profile can, via the above PDE, be correlated with an observed wavelength-dependent change in EQE. We would expect EQE at low wavelengths to be relatively insensitive to rear interface quality, since those photons are absorbed near the front and are thus much more likely to be collected at the front interface. Conversely, higher-wavelength photons are (comparatively) more likely to be absorbed near, and thus create carriers that recombine at, the rear interface. This can be visualized by looking at the steady-state short-circuit profiles of carriers created by light of different wavelengths, as shown in Figure 3.2. Note that although the lower s value leads to to an increased backside hole density, the overall backside recombination rate (proportional to the slope of the hole profile going to that interface) is indeed lower.

3.3.2 Application to Experimental PEDOT/n-Si(TiO_2) Devices

PEDOT/n-Si(/TiO_2) devices, schematically shown in Figure 3.3, were fabricated by Ken Nagamatsu as described in [29]. Wavelength-dependent EQE measurements were taken from 750 to 1000 nm. s_{rear}—in this case, $s_{\text{Si/TiO}_2}$— was varied in the simulations described in Section 3.3.1 over the range possible for the n-Si/TiO_2 junction. Experimental and simulated results are compared in Figure 3.4. As expected, significant positive ΔEQE is observed at higher wavelengths but not at lower ones, a product of the different absorption coefficients of light at the varying wavelengths.

Comparing the experimental and simulated results, a good fit is given by $s_{\text{Si/TiO}_2} = 300$ cm/s. A method based on the TiO_2-induced change in cell V_{OC} gave the value $s_{\text{Si/TiO}_2} = 400$ cm/s [29], the reverse-recovery method in Chapter 4 gives $s_{\text{Si/TiO}_2} = 330$ cm/s, and annealed interfaces were measured by QSSPCD (Section 2.5.1) to have

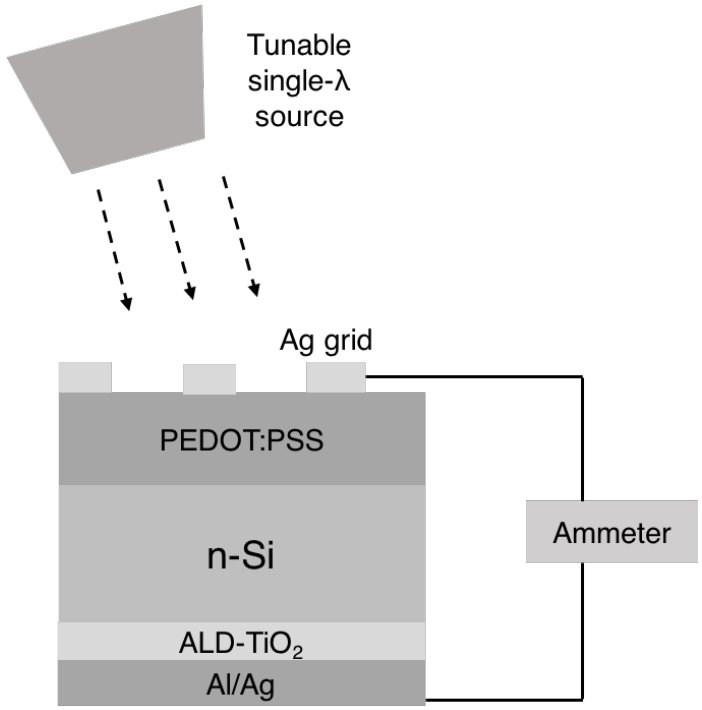

Figure 3.3: External quantum efficiency measurement schematic for the PEDOT/n-Si/TiO_2 device.

$s_{Si/TiO_2} = 200$ cm/s [47], so the EQE method seems to give accurate results as judged against other techniques.

In Chapter 2, it was suggested that selective contacts in SHJCs should have s with silicon of about 100 cm/s in order to produce cells of high efficiencies. The EQE method reveals that s_{Si/TiO_2} is indeed of that order of magnitude, though a bit higher. It should not be surprising, then, that as-deposited CVD-TiO_2 did indeed lead to $\Delta V_{OC} = 35$ mV and ΔPCE= 0.9% in PEDOT/n-Si(/TiO_2) SHJCs. TiO_2 should continue to be explored as part of next-generation silicon-including solar cells, and the EQE s-measurement method should continue to prove useful in characterizing those devices.

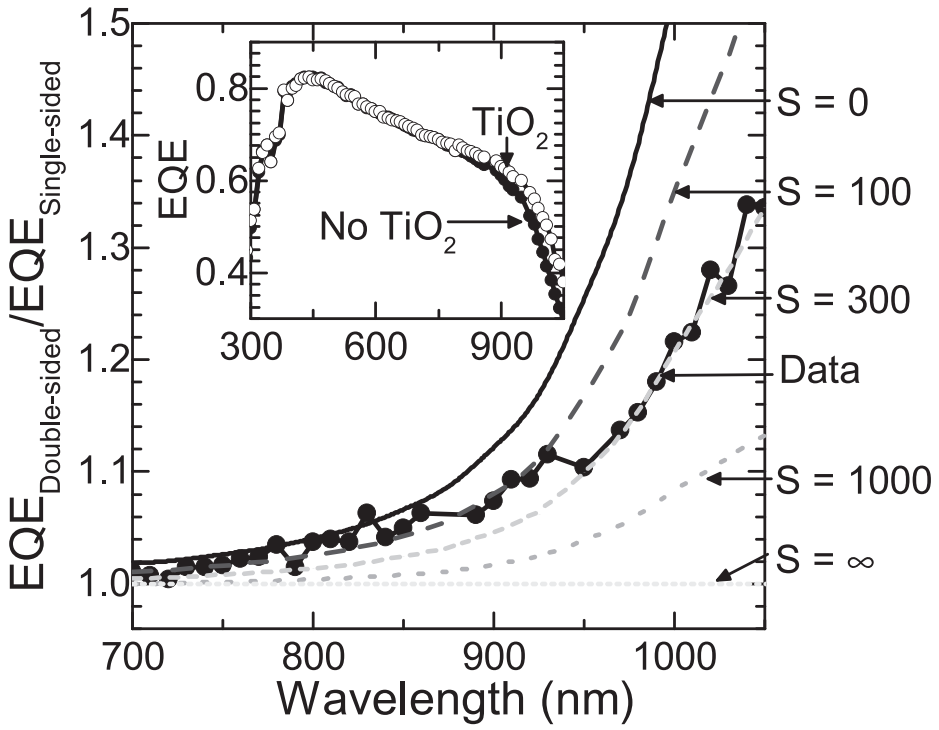

Figure 3.4: EQE (inset) and relative EQE gain for PEDOT/n-Si devices with and without backside TiO_2 films. Labeled lines show simulated ΔEQE for a device with ideal front interface and varying s_{rear}. Figure reproduced from [29]

3.4 Ultrathin TiO_2 As a Passivating Tunnel Layer

3.4.1 Introduction

The case of CVD-TiO_2, as just presented, is something of a happy coincidence—it is a selective contact material that also happens to have silicon passivation abilities. In general, though, the best silicon passivators, like SiO_2, SiN_x, and Al_2O_3, are non-selective insulators. Conversely, other carrier-selective materials may be poor silicon passivators. It is natural, therefore, to ask whether these two functions might be separated— whether a structure combining an ultrathin passivating tunnel layer with a thicker selective film could be a feasible contact for heterojunction solar cells. This structure is schematically illustrated, with TiO_2 serving as the tunnel layer, in Figure 3.5(b). Even without a novel selective layer, tunneling passivation layers could, for instance, replace the intrinsic a-Si layers found in conventional HIT cells.

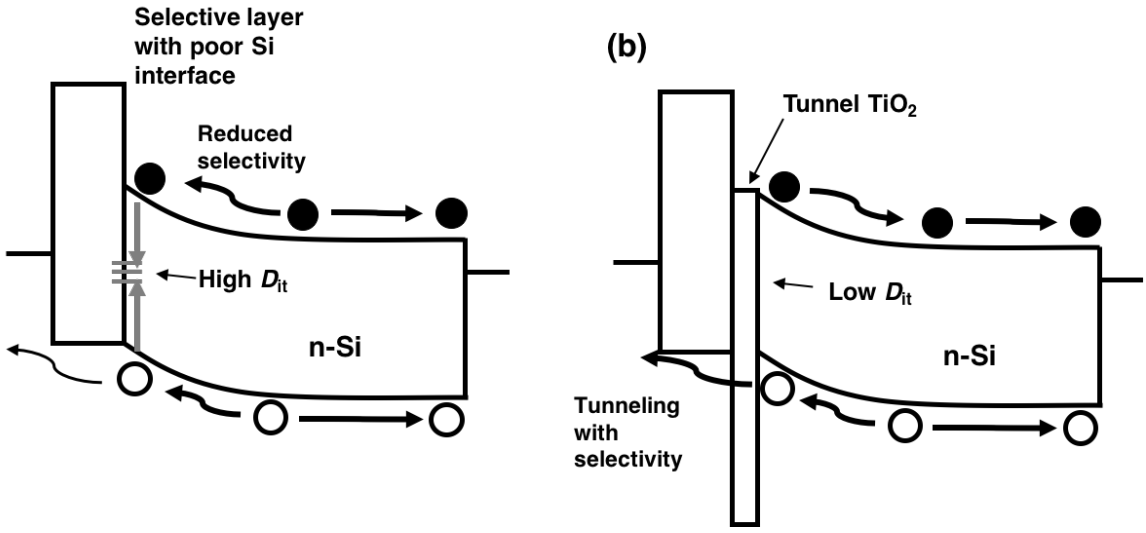

Figure 3.5: Schematic showing proposed combined tunnel layer + selective layer concept on silicon, in short-circuit condition under illumination. In (a), the selectivity of the blocking layer is wasted, because carrier tend to recombine at its interface with silicon. In (b), an intermediate tunnel layer prevents interface recombination, allowing the band offsets to impose selectivity.

The literature contains numerous attempts to fabricate tunnel layers for silicon solar cell applications, with SiO_x the most common material. Such layers can be grown chemically, via warm deionized water [48], hydrochloric acid [49], or nitric acid [50], or alternatively by plasma oxidation [49] or by rapid thermal oxidation [51],[52],[53]. However, surprisingly few papers on this topic report the *tunneling* characteristics of their SiO_x layers, generally focusing instead on the passivation ability of the films. Clearly, tunnel oxides must not only passivate but also have low resistances– carriers must be able to tunnel through easily.

Besides studies focused on solar cells, there is also a long history of studying ultrathin oxide tunneling in the context of gate dielectrics in Si-based FETs. This effect has become significant as dielectric thickness has approached the few-nm regime. According to one study [54], tunneling current densities for 100 mV applied bias on an Al/SiO_2/n-Si diode exceed 100 mA/cm^2 for 1 nm-thick SiO_2, dropping to 10 mA/cm^2 for a 1.5 nm-thick oxide and 2 mA/cm^2 for a 2.0 nm-thick oxide. Similar

results were found for devices with highly-doped polysilicon gates on 1.5–2.0 nm-thick SiO_2/inversion-layer structures [55]. Metal oxide films with these thicknesses are not out of the question if ALD is used for growth.

While SiO_2 is the standard material studied for ultrathin tunneling, there is no reason that another material could not be used, given sufficient passivation ability. Previous studies [29],[35],[56] have shown that thin (≈few nm) CVD-TiO_2 is an effective passivator of crystalline silicon, possibly via the formation of Si-O-Ti bonds that chemically passivate the surface. Although TiO_2, with its large valence-band offset and small conduction-band offset, is a hole-blocker on Si, if it were thin enough holes would theoretically be able to tunnel through, such that carrier-selection could be carried out by a separate layer. The traditional cathode material would thus find a place in *anodes*. As with the SiO_2 tunneling layers, this ultrathin TiO_2 must passivate effectively, present little resistance, and maintain its passivation ability even after overlayer deposition.

One concern specific to TiO_2 is the passivation mechanism. A fully chemical passivation mechanism would be ideal: theoretically, a monolayer of material would then be enough to fully passivate, and the degree and direction of bandbending would be determined by the selective-contact overlayer and/or metal contact. However, some papers point to electrical passivation as the dominant mechanism for TiO_2 on c-Si [41], with negative charge in the TiO_2 or at the interface leading to a low effective s (see Figure 3.13). Charge-induced bandbending could present a barrier to carriers even if the oxide itself were thin enough to tunnel through.

3.4.2 Ultrathin TiO_2 Experiments

Though our group's previous work was based on CVD-TiO_2 using thermolysis of titanium(IV) *tert*-butoxide without an oxidizing agent, TiO_2 deposition via ALD is also common, generally using the precursor titanium(IV)-isopropoxide (TTIP) [57],[34].

Ideally, layer-by-layer deposition would enable the growth of the extremely thin films that allow for tunneling. The ultrathin TiO_2 films in this section were grown at temperatures of either 125°C or 175°C, depending on the experiment, using water vapor as oxidizer. The precursor was held at a constant temperature of 70°C during processing. The ALD chamber had pressures in the hundreds of mTorr, and N_2 was kept continuously flowing during precursor/oxidizer pulses. Analyses of thick ALD-TiO_2 films revealed constant growth rates of about 0.35 Å/cycle, so this ALD process, as observed, did not provide true layer-by-layer film growth, but rather deposited only a fraction of a monolayer each cycle.

TiO_2 Tunnel Layer Uniformity

At cycle counts of less than a few dozen, full coverage becomes a major concern, since ALD processes can require the formation of an initial wetting layer before deposition becomes uniform. Making the testing more difficult, ellipsometry can become unreliable at scales of a few nm, particularly when the material's optical parameters are unclear. Therefore, XPS was used to test the uniformity of low-cycle-count ALD-TiO_2 films. For a single thin overlayer of thickness t_o deposited on a thick substrate, the ratios of XPS intensities I are expected to scale as [24]

$$I_o/I_s = C_0 \frac{1 - \exp(-t_o/\chi_o)}{\exp(-t_o/\chi_s)} \tag{3.4}$$

where C_0 is a material- and system-dependent constant and the χ_i are the energy-dependent inelastic mean free paths for electrons in the materials.

175°C TiO_2 depositions of 5, 10, 20, 40, 60, and 80 cycles on lightly-doped H-terminated n-Si were carried out and the Ti2p and Si2p XPS spectra measured. The measurements were performed in Princeton's Imaging and Analysis Center (IAC). For the purposes of intensity ratios as used here, the specific peaks used do not matter,

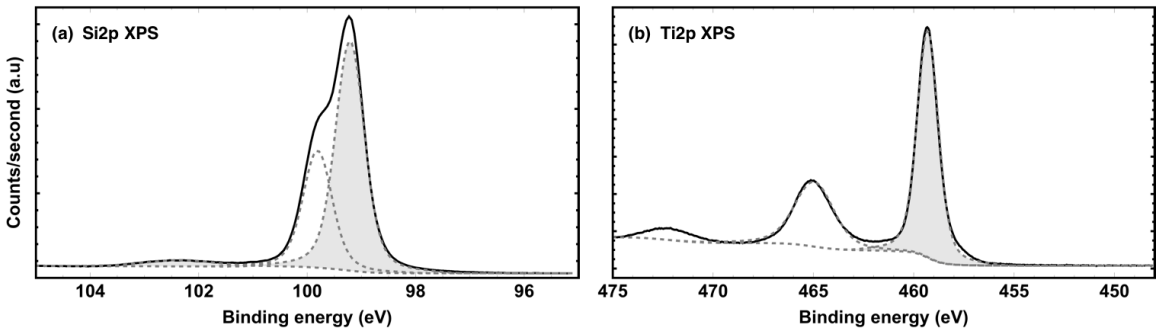

Figure 3.6: $2p$ XPS spectra for (a) Si and (b) ALD-TiO_2 (60 cycles), along with fitted peaks and modified Shirley background (Section 2.5.2). The shaded areas were used for the computation of the intensity ratios in this section.

as long as they are consistent from calculation to calculation. Here, for simplicity, ratios were taken of the large primary peaks indicated in Figure 3.6.

The resulting ratios are plotted against cycle count in Fig 3.7. The 40, 60, and 80-cycle points were used along with Equation 3.4 to fit for χ_o and χ_s, with C taken to equal 1 and thickness assumed to scale proportionally with cycle count. The 20-cycle Ti$2p$ signal is slightly lower than would be expected given uniform ALD growth, and the 10-cycle Ti$2p$ signal is significantly lower. The 5-cycle point does lie on the expected curve, though at that point the expected thickness would only be about 0.15 nm, such that uniform monolayer formation seems extremely unlikely. This data appears to indicate layer uniformity problems for thicknesses below about 20 ALD cycles.

Tunneling of Holes From Silicon Through Thin TiO_2 Layers

Another test of TiO_2 as a tunnel-oxide was its actual tunneling properties. Specifically, because a high or thick effective barrier would lower the fill factor of a completed solar cell, any tunnel layer must not significantly impede carrier flow. Test devices were fabricated to measure the effective barrier as a function of TiO_2 cycles. First, ALD-TiO_2 was deposited onto p-Si as already described. Direct contact evaporation onto ultrathin TiO_2 layers might damage them, making measurement of the true

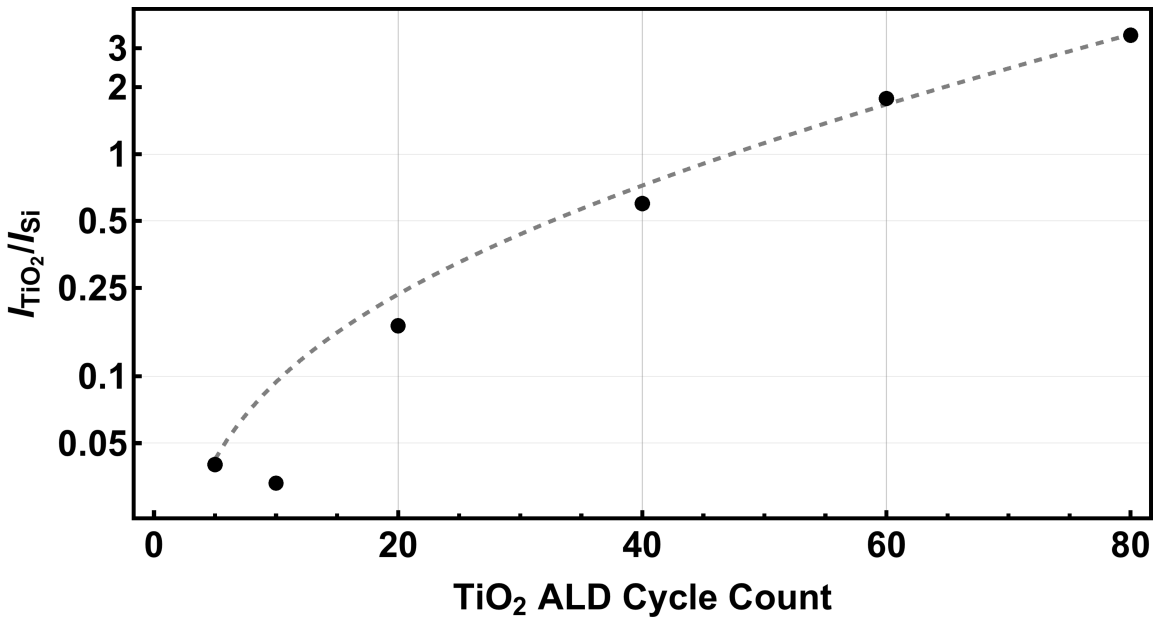

Figure 3.7: Experiment (circles) and model fit (dashed line, Eqn 3.4) for the ratio of the $2p$ XPS peaks of ALD-TiO_2 on Si. The model fit only uses the three highest-cycle points.

interface/film properties impossible; thus, prior to metallization, a $\sim$ 70nm-thick PEDOT:PSS layer was deposited via spin coating, acting as a conductive yet protective layer prior to metallization. PEDOT is a high-work-function conductor that should present only a small barrier to hole conduction out of silicon. Finally, silver was deposited onto both sides, and individual 16x16mm devices were defined via manual scribing of the PEDOT. The device schematic and typical resulting $J-V$ curves are shown in Figure 3.8.

As shown in Figure 3.9(a), these devices are majority-carrier dominated, with holes traveling from the p-Si through/over the TiO_2, through the heavily p-type PEDOT, and then finally out the top cathode contact. The question is how to characterize the magnitude of the hole barrier. In Chapter 2, it was suggested that a 10-nm-thick selective contact in a SHJC should have a resistivity of no more than 10^6 Ω·cm. Such a layer would present an area-normalized series resistance of $R_{S,A} = 1$ Ω·cm^2, corresponding to a 0.04 V drop at currents of 40 mA/cm^2. This limit is somewhat

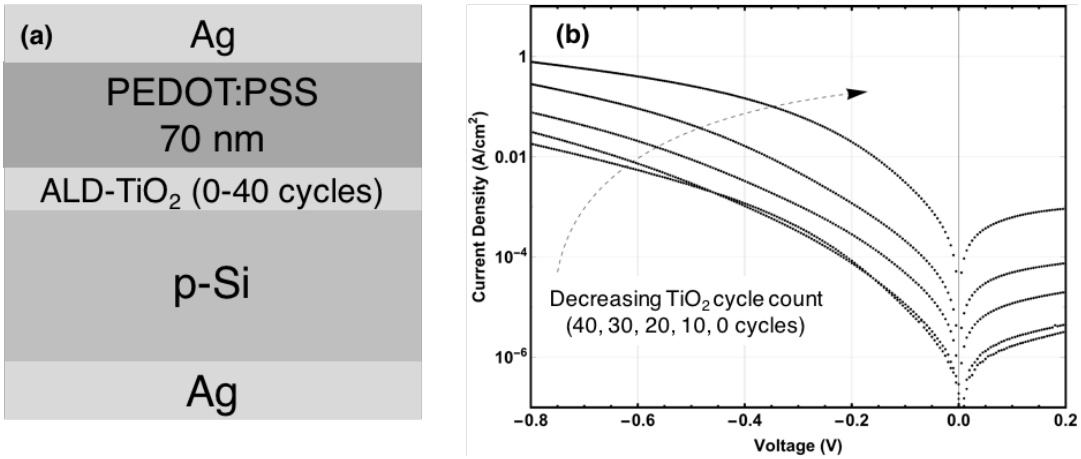

Figure 3.8: (a) Device structure and (b) $J-V$ curves for ALD-TiO_2/PEDOT/p-Si devices. Thicker TiO_2 layers present increasing barriers to hole transport. Note that forward bias on the TiO_2/p-Si cathode means applying a negative voltage.

generous, since it ignores that real solar cells will have other, unavoidable, sources of series resistance, but nevertheless the $R_{S,A} = 1\ \Omega{\cdot}cm^2$ point provides a reasonable standard against which to compare the ultrathin TiO_2. As shown in Figure 3.8, the device JV characteristics were diode-like (though with n not close to 1) with significant series resistances. Rather than fitting these highly non-ideal curves, we instead define effective resistances $R_{eff} = -0.6/J(-0.6V)$. For the PEDOT/p-Si device, $R_{eff,PEDOT} = 0.67\ \Omega{\cdot}cm^2$, below the desired limit. For the TiO_2-containing devices, the relevant expression is then $R_{eff,TiO2} = -0.6/(J(-0.6V) - R_{eff,PEDOT})$. Clearly, these devices are not in fact Ohmic, but this effective resistance tells us roughly to what extent the TiO_2 layer impedes hole flow, which is the important point.

Figure 3.9(b) shows these extracted effective resistances. While PEDOT by itself does not present an unacceptable effective resistance, all TiO_2 thicknesses have effective resistances above the desired limit. For instance, the 10-cycle layer has $R_{eff,TiO2} = 5.8\ \Omega{\cdot}cm^2$, corresponding to a voltage drop of over 200 mV at 40 mA/cm^2. The effect saturates for the thickest TiO_2 films, when holes are fully unable to tunnel through the oxide layer. Therefore, it appears as though TiO_2 as a passivating tun-

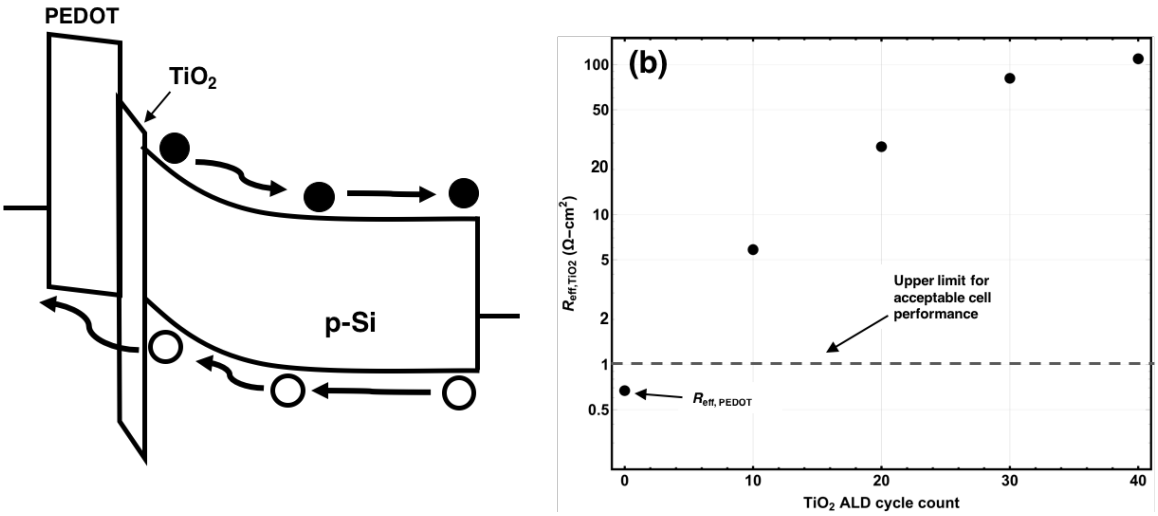

Figure 3.9: (a) Device band structure under forward bias (negative voltage on the PEDOT) and (b) extracted effective resistances for the TiO_2 layer, calculated at $V = -0.6$ V.

nel layer, at least in the thickness range tested, presents too-high a barrier to hole transport.

TiO_2 Tunnel Layer With ALD-NiO Overlayer

Actually-existing TiO_2 tunnel layers would have to maintain their passivation ability even after overlayer deposition. Even with no change in processing temperature, chemical reactions or charge transfer could occur that would render the tunnel layer useless. To test this, ALD-NiO/ALD-TiO_2/FZ-n-Si devices were fabricated, with the ALD processes taking place at 175°C (further details on the NiO process can be found in Chapter 5). 150 cycles, or about 6 nm, of NiO were deposited. The device version of this structure is essentially the one depicted in Figure 3.5(b), with NiO being the electron-blocking layer, but in this case the ALD processes deposited the oxide layers on *both* sides. For device application, the oxide stack would just be on the anode side, with the tunnel-TiO_2 passivating the n-Si and the NiO on top blocking electrons; here, the symmetric structure, without metal contacts, allows for QSSPCD

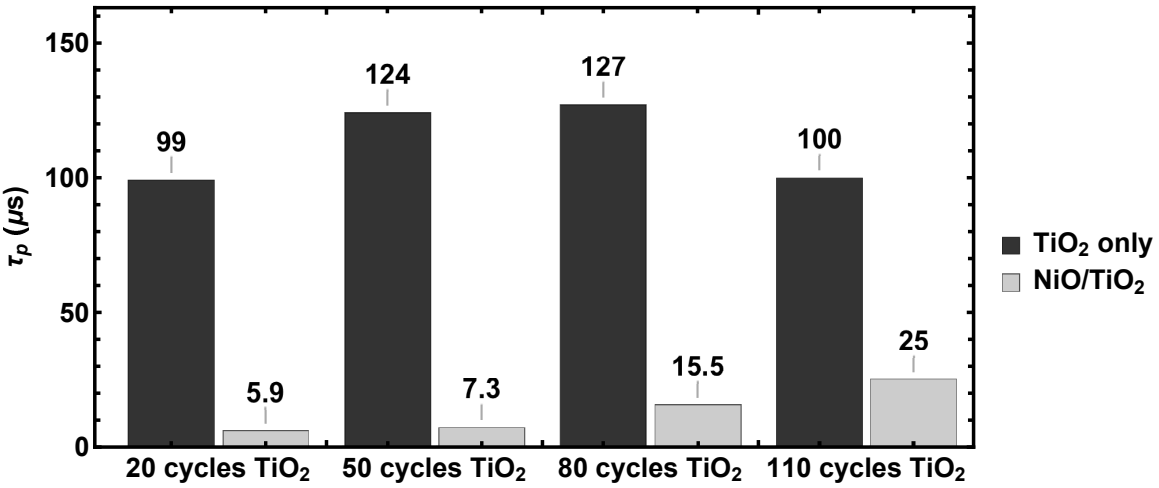

Figure 3.10: Minority-carrier lifetimes for various ALD-TiO_2 films on n-Si before and after 150 cycles of ALD-NiO, measured by QSSPCD.

measurements. QSSPCD measurements were taken before and after NiO deposition, allowing for the extraction of hole recombination times τ_p, as shown in Figure 3.10.

In all cases, the NiO deposition greatly decreases the observed carrier lifetimes, with thicker TiO_2 films shown to be more robust. The pre-NiO lifetimes are quite good—$\tau_p = 100$ μs corresponds to $s \sim 300$ cm/s (see Equation 2.8), so even the thinnest film is a good passivator. But the post-NiO recombination velocities are in the range of $\sim 10^3$ to 10^4 cm/s, values inadequate for high-performing solar cells (see Chapter 2). Interestingly, TiO_2 passivation as a function of cycle did not monotonically increase but rather rose and then slightly fell, mirroring the results in [19] and implying a charge effect. This is discussed further in the next subsection.

Mechanism of Silicon Surface Passivation in $NiO/TiO_2/Si$ Structures

To further investigate this effect, XPS was used to examine the chemistry of thin TiO_2 layers in oxide stacks. 20-cycle ALD-TiO_2 (H_2O oxidizer) films were deposited on H-terminated n-Si substrates at 125°C. One sample then received a 150-cycle ALD-NiO deposition at 175°C using hydrogen peroxide vapor as the oxidizer. Detailed scans were taken of the Ti2p and Si2p regions, and the resulting curves were analyzed.

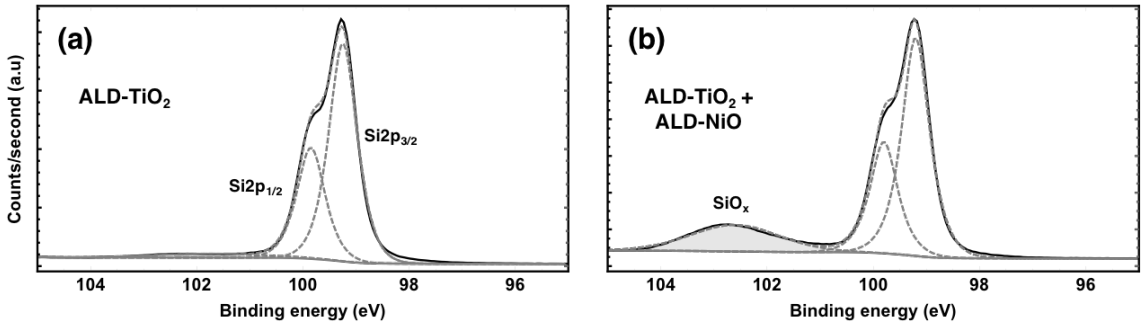

Figure 3.11: Substrate Si2p XPS signal for a 20-cycle ALD-TiO_2 deposition, without (a) and with (b) a 150-cycle ALD-NiO overlayer deposited.

For the as-deposited TiO_2 film (Figure 3.11), the Si2p spectrum shows very little SiO_2 formation. Presumably, the silicon, hydrogen-terminated after an HF dip, maintained that passivating H-termination after the ALD-TiO_2 deposition, possibly along with the formation of Si-O-Ti bonds [35]. Although high-temperature thermal oxides form well-passivated junctions with crystalline silicon, the disordered oxides that would form at these low temperatures would not be expected to passivate effectively; therefore, the lack of SiO_x at the interface is a good sign for passivation.

After the ALD-NiO deposition, though, a significant SiO_x peak exists. This could be due to the higher temperature (175°C vs 125°C) encouraging oxide formation, or because the stronger oxidant used in the NiO growth (H_2O_2 vs H_2O) is, unsurprisingly, more able to form low-temperature silicon oxide. The resulting thin SiO_x layer could be behind the much lower lifetimes observed after NiO deposition (Figure 3.10).

The NiO deposition also alters the chemistry of the TiO_2 layer itself, as shown in Figure 3.12. The XPS spectra can be decomposed into signals corresponding to the Ti(IV), Ti(III), and other titanium oxidation states [58]. Before NiO deposition, the Ti2p spectrum is dominated by peaks assigned to Ti(IV), ie, the peaks expected in fully stoichiometric TiO_2. There is also a considerable Ti(III) peak. After NiO deposition, the Ti(III) peak is not visible, but a significant shoulder can be observed on the higher-BE side of the primary Ti peak; deconvolution reveals that the processing has increased the prominence of a peak of uncertain origin. We tentatively assign it

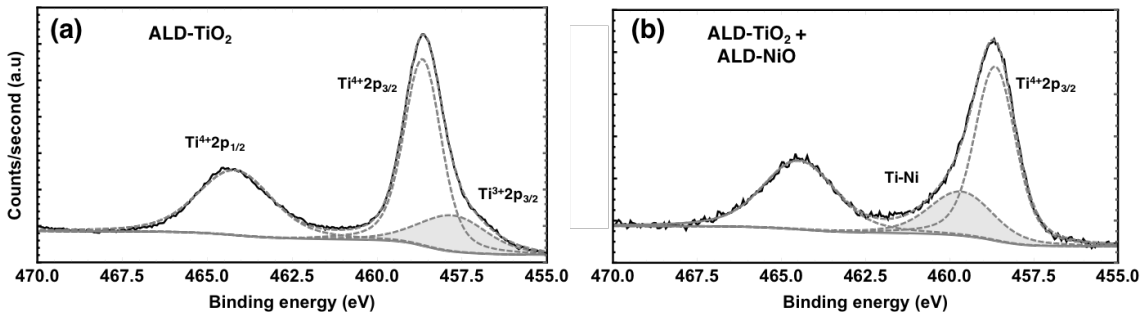

Figure 3.12: Ti2p XPS signal for a 20-cycle ALD-TiO_2 deposition, without (a) and with (b) a 150-cycle ALD-NiO overlayer deposited. The shaded peaks are the ones that differ between treatments.

to formation of NiTi [59], with the $Ti^{3+} \rightarrow$ Ti-Ni modification altering the film charge density and thus its passivation properties. In this formulation, the Ti^{3+} are not fully compensated and represent positive fixed charge in the oxide.

For TiO_2 passivation of crystalline silicon, it is worthwhile to distinguish between interfacial dipole effects and fixed-charge effects. Interfacial dipoles are expected to be present whenever ionic bonds are present at an interface. According to [60], atomic Si has an electron affinity (EA) of 1.39 eV, O has $EA = 1.44$ eV, Ti has $EA = 0.33$ eV, and H has $EA = 0.75$ eV. Electrons tend to move from low-EA to high-EA atoms, so dipole fields point (positive charge to negative charge) in the direction of higher EA. At the Si-O-Ti interface, therefore, a dipole field would form pointing from the Ti towards the O, and the TiO_2 bands would be shifted down relative to the Si bands. In the case of H-terminated silicon with TiO_2 on top of it (effectively Si-H-O-Ti), the same Ti-O dipole exists, but it is compensated by a smaller opposing dipole field (smaller because the ΔEA is smaller) pointing from the H to the O; the band shift in this case would be somewhat less than in the Si-O-Ti case (at least when considered with this EA heuristic). Finally, in the case of a thicker SiO_x on top of Si, as was observed by XPS after ALD-NiO deposition, the Si-O bonds are covalent; the only dipole is at the SiO_2/TiO_2 interface. In all of these cases, both the interfacial trap

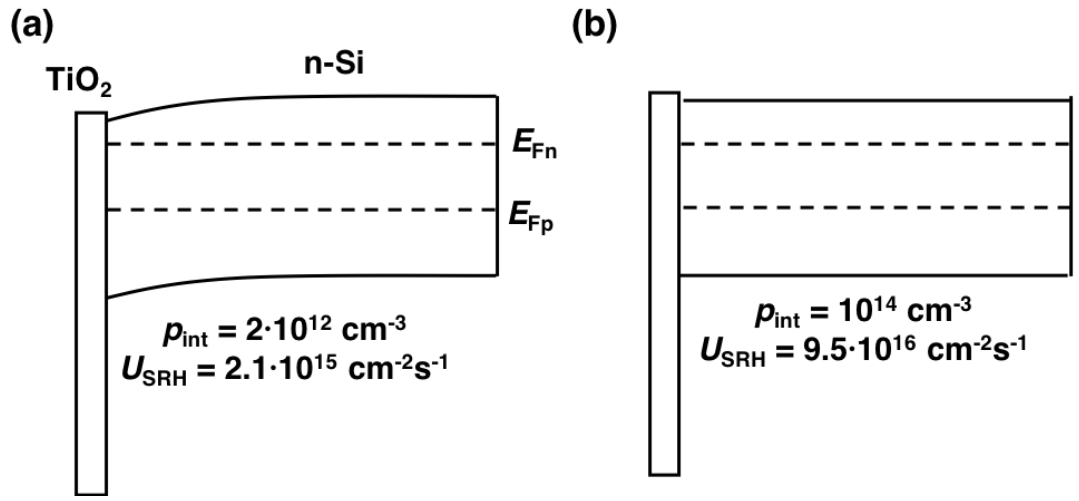

Figure 3.13: Hole density and interface recombination rate for a hypothetical illuminated TiO_2/n-Si junction (a) with and (b) without 100 meV downward bandbending induced by the oxide.

density N_t and the trap energy position in the gap E_t could be different, altering the observed recombination characteristics.

These cases are distinct from the question of fixed charge in the TiO_2. According to this model, fixed positive charge (Ti^{3+}) exists in the oxide, putting the silicon into accumulation and increasing n and decreasing p at the interface. In the presence of bandbending, n and p in Equation 2.2 are modified by factors of $e^{-\Delta E/kT}$ and $e^{\Delta E/kT}$, respectively, where ΔE is the amount of bandbending, with upward bandbending taken to be positive. If illuminated quasi-Fermi levels split to allow, say, $\Delta p = 10^{14}$ cm^{-3} in the silicon bulk, against a background doping of $n = 2 \cdot 10^{14}$ cm^{-3}, with $s = 100$ cm/s, then the predicted recombinative current density is 15 mA/cm^2 with flat bands but only 0.3 mA/cm^2 with bands bent down by $\Delta E = -100$ meV. The hypothetical TiO_2/n-Si interface with and without charge-induced bandbending is shown in Figure 3.13, with recombination figures calculated using Equation 2.2 and the parameters listed above, showing how a change in the TiO_2's charge properties could significantly change recombination amounts.

NiO deposition on top of the TiO_2 would modulate this fixed-charge-induced band-bending. It is unlikely that the ozone itself would be responsible for the change: the highly-oxidizing ozone would be expected to neutralize *negative*, not positive, charges. But some other form of neutralization or charge transfer could occur, reducing the charge density and thus the silicon bandbending.

3.5 Conclusion

The EQE-based analysis method presented in this chapter proved effective for extraction of Si/TiO_2 interface recombination. Given that EQE measurements are a standard tool in most PV labs, this method could fit well into research groups' analysis routines. The method could be made more accurate by accounting for front-interface non-ideality or by including more detailed optical calculations for the top layers. Regardless, these results provide further support for the use of TiO_2 on crystalline silicon for solar cell applications—whether measured optically, electrically, or as part of a full solar-cell efficiency analysis, the TiO_2/c-Si junction is of anomalously high quality given its heterojunction nature.

This chapter also featured a multipronged investigation into the possibility of using TiO_2 as a passivating tunnel layer in hole-selective contacts. The experiments presented send mixed messages. Although even very thin ALD-TiO_2 layers are able to passivate n-Si, that passivation, as measured by QSSPC, was not maintained during ALD-based overlayer deposition. An XPS-based analysis showed that the NiO overlayer deposition altered the chemistry of both the Si/TiO_2 interface and the TiO_2 film itself. At the same time, it is possible that gentler overlayer deposition techniques (lower temperatures, weaker oxidizers, etc) could provide better results. However, IV analysis showed that even 10 cycles (< 3.5 nm) of TiO_2 presents a too-high barrier

for holes, and pushing the thickness down further may not be possible, given the XPS data showing that thinner films may not be planar.

Further work on TiO_2 should focus on determining unequivocally and exactly how it is able to passivate crystalline silicon—the relative importance of chemical and electrical passivation are still not clear. The ALD-TiO_2 process should be further studied to make films as thin as possible while maintaining uniformity: even if the material cannot be used as a tunnel-layer in anodes, multi-material cathode stacks including ultrathin TiO_2 could be superior to single-layer cathodes. Finally, just as Chapter 5 studies doped ALD-NiO films, doping could be incorporated into the ALD-TiO_2 process to improve its conductivity and band alignments.

Chapter 4

Extraction of Interface Recombination Parameters Via the Reverse-Recovery Method

4.1 Introduction

The previous chapters explained the importance of s, the interface recombination velocity, in determining the ultimate efficiencies of heterojunction solar cells. At low s values, recombination is suppressed, meaning that heterojunction contacts' band-offsets with the silicon are given a chance to matter, the resulting effective electric fields driving carrier gradients, photovoltage, and photocurrent. When s is high, "selective" contacts behave like non-selective carrier sinks, and cell efficiency decreases.

The previous chapters also presented two optical methods for determining s. The first, based on transient photoconductivity measurements, can be used only on the very simplest test samples: the silicon substrate, plus the perhaps-passivating films. This is useful, but it overlooks the fact that recombination characteristics often change

after further processing and metallization. These steps can induce or diminish interfacial bandbending, modify the charge characteristics of the passivating layer, or even physically or chemically damage that layer. It is useful, therefore, to have s-measurement techniques that work in contacted, completed devices. Indeed, the previous chapter also suggested one such method, wavelength-dependent EQE measurements, and applied it experimentally to the case of the n-Si/TiO_2 junction in a completed solar cell. That technique uses the different absorption lengths of different light wavelengths to model carrier diffusion in the silicon bulk and extract interface recombination information. As shown above, the s found by the EQE method closely matched the s extracted by modeling the dark current of the same finished device.

Still, ideally one would be able to extract that recombination information from electrical, rather than optical, measurements, which are usually simpler to run and which can offer more flexibility, in that they can allow you to test the device not under short-circuit conditions but rather at the actual operating points of a solar cell. Furthermore, both the EQE method and the dark current method only allowed the extraction of the cathode-side s in an n-Si-based SHJC (they could also be used on the anode side of a p-Si-based SHJC); the minority-carrier-injecting interface's effective s was not independently calculated, but rather simply assumed to be very low. This chapter provides an electrical method for the extraction of either interface's recombination parameters, extending the classic "reverse-recovery" technique to the case of double-sided silicon heterojunction devices. Theory, simulation, and experiment are combined to successfully demonstrate the techniques efficacy and usefulness.

4.2 Background

A diode in steady-state forward bias has some elevated number of minority carriers distributed throughout its semiconductor regions. When such a device, which would

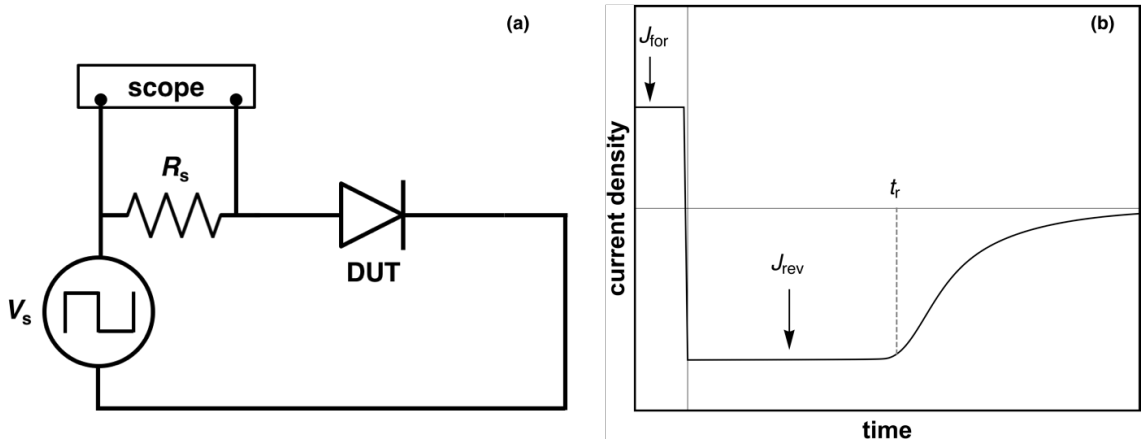

Figure 4.1: a) Basic circuit schematic of a reverse recovery experiment and b) a resulting current waveform.

normally have low reverse current density J_0, is switched quickly from forward to reverse bias, those minority carriers are able to flow back out from where they came, leading to a temporarily elevated reverse current $|J_{\rm rev}| > |J_0|$ that is sustained for some "recovery time" $t_{\rm r}$. This is the reverse recovery (RR) process; the basic experimental circuit and a resulting current waveform (neglecting capacitance, etc) are shown in Figure 4.1. Because the device maintains a forward bias as long as excess minority carriers maintain their forward-diffusing profile across the base, the resistor in the circuit controls current levels during the recovery process. Examples of stored hole profiles both in forward bias and during the recovery process are shown in Figure 4.3.

A simple analysis of this problem is as follows. Forward bias in a diode creates some stored minority carrier charge $\Delta Q_{\rm FB}$, which depends on the current mechanisms and recombination levels in the device. Imposing a reverse current $J_{\rm rev}$ at $t = 0$ draws those carriers out until they are depleted at time

$$t_{\rm r} = \frac{\Delta Q_{\rm Recov}}{J_{\rm rev}}. \tag{4.1}$$

The recovered charge Q_{Recov} can be lower than Q_{FB} due to recombination during the recovery process. Thus, an RR experiment should allow us to directly probe the recombination mechanisms in the device.

Because solar cells can be understood as fundamentally minority-carrier devices, this RR process has clear usefulness in measuring their internal physical processes. In particular, as we have seen, SHJC performance is largely determined by interface quality. Given the right assumptions and device models, then, an RR analysis should enable the measurement of interface recombination characteristics under the same forward-bias conditions that occur during actual solar cell operation.

Although Equation 4.1 embodies the fundamental physical reasoning underpinning RR experiments, it is unsuitable for actual quantitative work. Here, we will develop the diode RR idea from its simplest manifestation in bulk-recombination-based planar diodes to its form in the complex situation of small-area SHJCs. We will then use this final model to measure the interfaces of the PEDOT/Si/TiO_2 solar cell. Much of this chapter was first published in [61] and [62].

4.3 Previous RR Work

4.3.1 Long-Base Diode

The first paper to look at the RR question was published by Kingston in 1954 [63]; his approach defined the space in which future researchers would work. He examined the "long-base p^+-n diode, in which a left-side p^+ hole-injector sends holes into an infinitely long n-type base for a forward current density J_{for}. Injected holes are lost only through a bulk recombination process defined by a bulk lifetime τ_{p}, and 100% injection efficiency is assumed. Such a device has forward-bias injected hole profile $\Delta p(x) \propto e^{-x/(D_{\mathrm{p}}\tau_{\mathrm{p}})^{1/2}}$. For the rest of this chapter, we say $\Delta p \equiv p$, with the understanding that the equilibrium number of holes in an n-Si base will be much

smaller than any reasonable non-equilibrium number during forward bias and during the recovery process.

At $t = 0$, one sets a constant reverse current density $J_{\rm rev}$. This sets a left-side boundary condition of $\frac{\partial p}{\partial x} = \frac{J_{\rm rev}}{qD_{\rm p}}$, where q is the electron charge and $D_{\rm p}$ is the hole diffusion coefficient. The governing equation is

$$\frac{\partial p}{\partial t} = D_{\rm p}\frac{\partial^2 p}{\partial x^2} - \frac{p}{\tau_{\rm p}}. \tag{4.2}$$

Kingston's contribution was to observe that this constant current could be maintained *not* until all stored charge was depleted (ie, until $\int_0^\infty p(x)\mathrm{d}x = 0$), but rather until the left-side hole density $p(0) = 0$. At that time, which is the recovery time $t_{\rm r}$, the hole profile cannot sustain the same slope, and the current, by necessity, begins to decline, eventually settling at the saturation current of the diode.

Kingston solved this PDE and found the simple solution

$$\mathrm{erf}\Big(\sqrt{\frac{t_{\rm r}}{\tau_{\rm p}}}\Big) = \frac{1}{1 + J_{\rm rev}/J_{\rm for}}. \tag{4.3}$$

This formula is the one commonly given in elementary device textbooks.

4.3.2 Short-Base Diode

Equation 4.3 breaks down when the device fails to be solidly in the "long-base" regime—ie, when the diffusion length $L_{\rm p}$ approaches the base width W. Devices with longer bulk lifetimes are generally in the short-base regime $L_{\rm p} >> W$. Since it is now assumed that all carriers diffuse to the back contact, the nature of that contact matters. The most basic such contact, in the absence of a backside field or selective

contact, is a non-selective carrier sink that imposes the condition

$$p(W) = 0. \tag{4.4}$$

Lax and Neustadter [64] and Byczkowski and Madigan [65] both attacked the more general non-infinite-base problem, but an excellent short-base description was obtained by Grove[1] and Sah [66], who derived simple and accurate approximations for the short-base diode in reverse recovery.

Their recovery time approximation can be divided into two regimes:

$$t_{\mathrm{r}} \approx \begin{cases} \frac{4W^2}{\pi^2 D_{\mathrm{p}}} \ln\left[\frac{1+J_{\mathrm{for}}/J_{\mathrm{rev}}}{\pi^2}\right], & J_{\mathrm{rev}} \ll J_{\mathrm{for}} \\ \frac{\pi W^2}{4D_{\mathrm{p}}} \frac{1}{(1+J_{\mathrm{rev}}/J_{\mathrm{for}})^2}, & J_{\mathrm{rev}} \gg J_{\mathrm{for}}. \end{cases} \tag{4.5}$$

Even outside those two limits, choosing the appropriate equation for $J_{\mathrm{rev}} > J_{\mathrm{for}}$ or $J_{\mathrm{rev}} < J_{\mathrm{for}}$ provides a reasonably good approximation; the maximum error of about 20% occurs at $J_{\mathrm{rev}} = J_{\mathrm{for}}$. Additionally, the approximation is within about 20% for the entire range $L_{\mathrm{p}} > W$, not just in the asymptotic limit.

Other notable papers from this era are [67] and [68].

4.4 Reverse-Recovery in Non-Ideal Heterojunction Cells

Unfortunately, none of these results are strictly applicable to SHJCs. Charge-carrier flows in both the forward- and reverse-bias portions of the RR experiment are shown in Figure 4.2; these dynamics require the incorporation of two additional factors. We use the flow terms indicated in that figure throughout this section.

[1] This is the same Andrew Grove who would go on to found Intel.

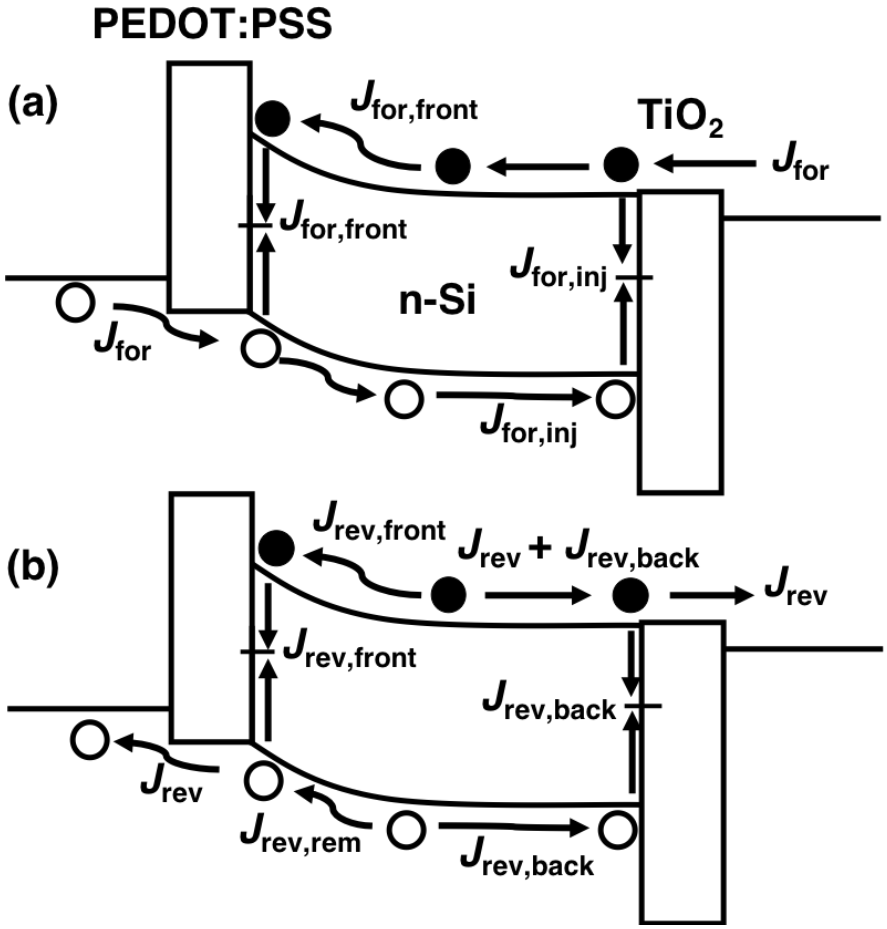

Figure 4.2: (a) Band diagram of device in forward bias, with total current density J_{for}. Arrows show carrier flux directions. The PEDOT/Si serves as a minority-carrier injecting (major-carrier blocking) selective contact. Ignoring bulk recombination, injected holes can either diffuse through the bulk to the back ($J_{\text{for,inj}}$) or recombine at the front ($J_{\text{for,front}}$). The rear heterojunction (Si/TiO_2 in our case) serves as a selective contact that blocks minority carriers from recombining at the cathode contact in the ideal case, but in practice will allow a finite amount of interface recombination. (b) Bands and carrier flows during reverse-current phase, again showing front- and rear-interface recombination. Because the device voltage is approximately constant (and still positive) during this phase, the bands are essentially unchanged from the forward-current case.

4.4.1 Recombination at back heterojunction interface

We first consider the case where the only recombination is at the right interface in Figure 4.2 (the heterojunction that blocks injected minority carrier holes from recombining at the cathode contact). It will be shown that the reverse recovery time can be increased substantially over the short-base case if the back interface recombination rate is low. We assume no recombination at the interface which blocks electrons from leaving the device, or equally the interface which injects minority carriers into the substrate.

We assume low-level injection current levels and no Auger recombination, so that the hole current in the base is due only to diffusion and thus the local concentration gradient. We also assume that equilibrium hole values are insignificant compared to the excited value. Furthermore, assuming a hole barrier sufficiently thick to prevent direct tunneling and zero bulk recombination (high lifetime float-zone substrates were used in our experiments), the hole recombination rate at the back interface (and thus the transport rate across the substrate in steady state) is determined by the hole density and recombination velocity at that interface, such that

$$J_{\text{for,inj}} = q \cdot s_{\text{back}} \cdot p_{\text{back}}, \tag{4.6}$$

where p_{back} is the hole density at the back interface. We assume a neutral back contact (flat silicon bands); band-bending up or down would increase or decrease the apparent recombination velocity, respectively.

Under forward bias, for the same voltage this contact reduces hole current by decreasing the magnitude of the hole gradient across the substrate, which, in the photovoltaic context, leads to an improvement in V_{OC}. In the RR context, for the same forward current, the layer (in a "double-sided device," DSD) leads to increases

in both applied voltage and total stored holes in forward bias over the "single-sided" heterojunction device (SSD), as shown in Figure 4.3.

Assuming no bulk recombination, low-level injection, and no recombination at the hole-injecting interface, one can calculate the effect of the back interface on total stored charge. It is a geometric argument based on the $t = 0$ curves in Figure 4.3. For a single-sided device (no backside passivation), the total stored charge is a triangle of area

$$\frac{1}{2}Wp(0) = \frac{J_{\text{for}}W^2}{2qD_{\text{p}}} \tag{4.7}$$

where the right-hand side came from the diffusion-current density definition $J = qD_{\text{p}}\mathrm{d}p/\mathrm{d}x$. For a double-sided device with backside recombination velocity s_{back}, the area is that same triangle plus a rectangle of area

$$Wp(W) = W\frac{WJ_{\text{for}}}{qs_{\text{back}}}. \tag{4.8}$$

Combining these values, one finds that the total number of stored charges for constant current (before switching to reverse bias) increases as

$$Q_{\text{DSD}} = Q_{\text{SSD}}\big(1 + \frac{2D_{\text{p}}}{Ws_{\text{back}}}\big). \tag{4.9}$$

This can be a significant difference: for parameters $W = 0.03$ cm, $D_{\text{p}} = 11.6$ cm^2/s, and $s_{\text{back}} = 250$ cm/s, the calculated increase is over 4x. Additionally, during the reverse recovery period itself, recombination at the back interface will consume holes faster at interfaces with higher recombination velocity. Because of the confluence of these two factors, a high quality back interface can increase recovery times by an order of magnitude for sufficiently low s_{back}.

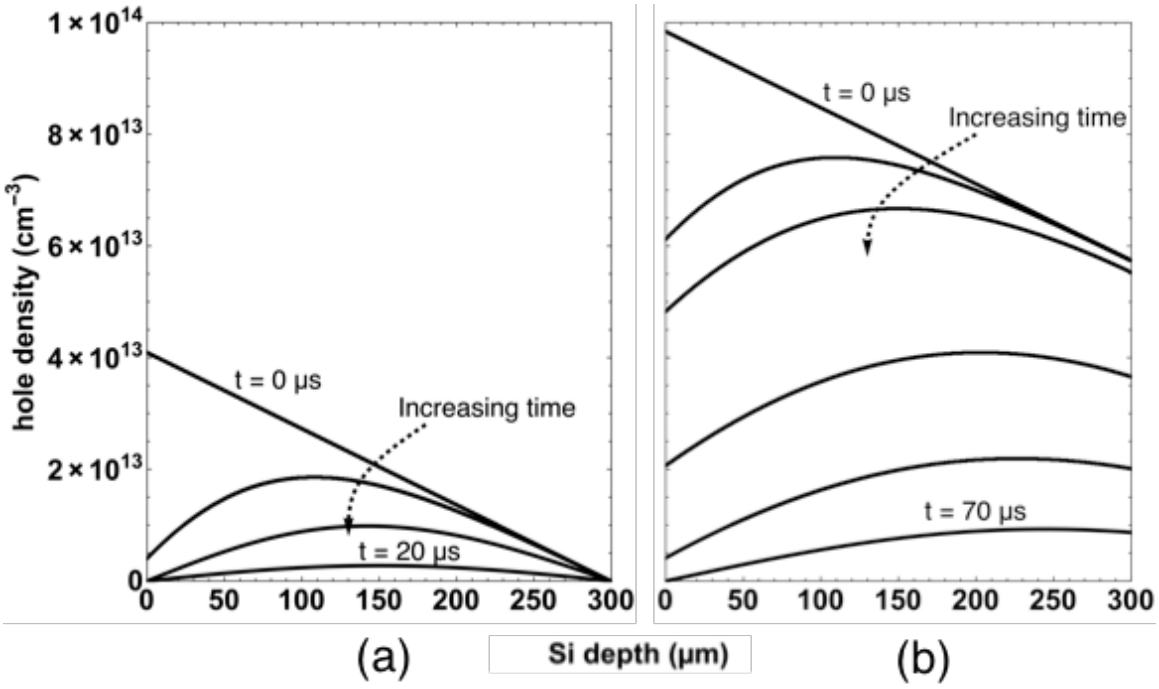

Figure 4.3: Modeled hole profile across quasi-neutral region over time after reverse bias was applied for (a) a PEDOT/n-Si device and (b) a PEDOT/n- Si/TiO_2 device with the same forward bias current density, assuming all current is carried by holes injected into the n-type base. The addition of the hole-blocking backside contact leads to a significantly increased number of holes in forward bias, and thus increased reverse-recovery times. Parameters are $J_{\text{for}} = 10$ mA/cm^2, $J_{\text{rev},t=0} = 15$ mA/cm^2,and $s_{\text{Si/TiO2}} = 275$ cm/s, with $W = 0.03$ cm and base doping $n = 2 \cdot 10^{15}$ cm^{-3}. This modeling is based on Equations 4.33 and 4.17.

4.4.2 Front heterojunction recombination in a single-sided device

We now consider the effect of recombination at the minority-carrier-injecting heterojunction (left interface in Figure 4.2). Recombination current at this interface affects reverse recovery experiments in two distinct ways. First, during the forward-bias phase, any component of the current that does not lead to holes injected into the base does not create a stored charge to later be recovered. Therefore, in setting up the hole distribution in the substrate in forward bias, the hole current density injected

across the base is

$$J_{\text{for,inj}} = \gamma_{\text{SSD}} \cdot J_{\text{for}} \tag{4.10}$$

with the emitter injection efficiency $\gamma_{\text{SSD}} \equiv J_{\text{for,inj}}/\big(J_{\text{for,front}} + J_{\text{for,inj}}\big)$, so that the $J_{\text{for,front}}$ reduces stored holes and thus the recovery time. We can incorporate this change into modeling by simply reducing the $J_{\text{for,inj}}$ as in Equation 4.10.

Second, during the reverse recovery phase, recombination ($J_{\text{rev,front}}$) adds to the external density J_{rev} flowing through the device terminals to give the total rate at which holes are being removed from the bulk, further reducing the recovery time compared to the case without such interface recombination. The hole-removal current density from the bulk at the front interface during the reverse transient is

$$J_{\text{rev,rem}} = J_{\text{rev}} + J_{\text{rev,front}}. \tag{4.11}$$

To model recombination at the front interface during the reverse-bias phase, we begin by writing down the Shockley-Read-Hall recombination current at the front interface, assuming a single dominant trap level, given by

$$J_{\text{rev,front}} = q s_{\text{front}} \Big[\frac{p_{\text{front}} n_{\text{front}} - n_{\text{i}}^2}{p_{\text{front}} + n_{\text{front}} + 2 n_{\text{i}} \cosh(\frac{E_{\text{i}} - E_{\text{t}}}{kT})} \Big] \tag{4.12}$$

where s_{front} is the front-interface recombination velocity (including implicit terms for carrier/trap interaction velocity, trap cross-section, and trap density), p_{front} and n_{front} are the hole and electron densities, respectively, at that interface, n_{i} is the intrinsic electron density, E_{i} is the silicon midgap energy level, and E_{t} is the trap level.

We first note that $p_{\text{front}} n_{\text{front}} = n_{\text{i}}^2 \exp[q V_{\text{pn,front}}/kT]$, with $V_{\text{pn,front}}$ being the quasi-Fermi level splitting at the interface. We also note that, assuming a relatively small electron current across the silicon (due mainly to this front-side recombination), the

electron quasi-Fermi level reflects the contact to the backside metal or TiO_2 layer, and we can say that the quasi-Fermi levels are effectively flat from the PEDOT/Si interface to the edge of the quasi-neutral bulk: in other words, $V_{\mathrm{pn,front}} = V_{\mathrm{pn,QNR}}$, the splitting at the edge of the quasi-neutral region. Since this splitting determines everywhere the pn product, we can say that $p_{\mathrm{front}} n_{\mathrm{front}} = p(0)n(0)$, where $p(0)$ and $n(0)$ are the hole and electron density at the edge of the quasi-neutral region. However, $n(0)$ is just the silicon doping N_{D}; furthermore, because the PEDOT lowest occupied molecular orbital (HOMO) and the silicon valence band maximum are closely aligned [69], p_{front} will be high and constant throughout the process, pinned at a number determined by the relative locations of the PEDOT and silicon valence band levels and reflecting a quasi-Fermi level set by the PEDOT contact. This allows us to rewrite the SRH expression as

$$J_{\mathrm{rev,front}} = q s_{\mathrm{front}} \Big[\frac{p(0)N_{\mathrm{D}} - n_{\mathrm{i}}^2}{p_{\mathrm{front}} + n_{\mathrm{front}} + 2n_{\mathrm{i}} \cosh(\frac{E_{\mathrm{i}} - E_{\mathrm{t}}}{kT})} \Big] \tag{4.13}$$

$$\approx q s_{\mathrm{front}} \Big[\frac{p(0)N_{\mathrm{D}}}{p_{\mathrm{front}}} \Big]. \tag{4.14}$$

All of the parameters in this expression are constant except for p(0)— it is a simple proportionality equation. Therefore, during time-evolution, given a certain forward-bias recombination value $J_{\mathrm{for,front}}$, we can accurately include recombination via the expression

$$J_{\mathrm{rev,front}}(t) = J_{\mathrm{for,front}} \frac{p(0,t)}{p(0,0)} \tag{4.15}$$

even without knowing the values of the other parameters in Equation 4.14 (note that, because tunneling or shunt currents do not follow this relationship, the presence of those would lead to a slightly worse fit between the model and experiment. However, because tunneling and shunt currents are very much a second-order effect, we would

not expect the change to be substantial). Equation 4.15 is useful because it allows us to couple an interfacial phenomenon dependent on several unknowns to a time-evolution that models holes only in the silicon quasi-neutral bulk. Equation 4.14 also suggests the definition $s_{\mathrm{front,eff}} = s_{\mathrm{front}} N_{\mathrm{D}}/p_{\mathrm{front}}$, an effective recombination velocity for minority-carrier holes that accounts for the effects of depletion at the junction. Given γ_{SSD}, a consideration of the hole profile across the base yields the expression

$$s_{\mathrm{front,eff}} = (D_{\mathrm{p}}(1-\gamma_{\mathrm{SSD}}))/(\gamma_{\mathrm{SSD}} L). \tag{4.16}$$

4.4.3 Reduction to dimensionless units

The reverse recovery process depends on the wafer thickness and minority-carrier diffusion coefficient D. To make our numerical modeling results more general, we define normalized variables for distance and interface recombination velocity. These normalized variables must be consistent for the device bulk and at each silicon interface. In the device bulk (for n-Si), the time-dependent minority carrier diffusion equation is

$$\frac{\partial p}{\partial t} = D_{\mathrm{p}} \frac{\partial^2 p}{\partial x^2} \tag{4.17}$$

$$= D_{\mathrm{p}} \frac{\partial}{\sqrt{D_{\mathrm{p}}}\partial(x/\sqrt{D_{\mathrm{p}}})} \frac{\partial}{\sqrt{D_{\mathrm{p}}}\partial(x/\sqrt{D_{\mathrm{p}}})} p \tag{4.18}$$

$$= \frac{\partial^2 p}{\partial (x')^2} \tag{4.19}$$

which is independent of D_{p}. Thus we define a normalized length scale $x' \equiv x/\sqrt{D_{\mathrm{p}}}$, implying a normalized wafer thickness $W' \equiv W/\sqrt{D_{\mathrm{p}}}$.

At the back interface, where the hole current towards the contact equals the recombination rate at the interface,

$$D_{\mathrm{p}}\frac{\partial p}{\partial x} = -s_{\mathrm{back}}p \tag{4.20}$$

$$\sqrt{D_{\mathrm{p}}}\frac{\partial p}{\partial x'} = -s'_{\mathrm{back}}\sqrt{D_{\mathrm{p}}}\cdot p \tag{4.21}$$

$$\frac{\partial p}{\partial x'} = -s'_{\mathrm{back}}p \tag{4.22}$$

where we defined a normalized $s'_{\mathrm{back}} \equiv s_{\mathrm{back}}/\sqrt{D_{\mathrm{p}}}$.

Finally, at the front interface,

$$J = -qD_{\mathrm{p}}\frac{\partial p}{\partial x} + J_{\mathrm{front}} \tag{4.23}$$

$$= -q\sqrt{D_{\mathrm{p}}}\frac{\partial p}{\partial x'} + J_{\mathrm{front}} \tag{4.24}$$

$$J' = -q\frac{\partial p}{\partial x'} + J'_{\mathrm{front}} \tag{4.25}$$

where we defined normalized current densities $J'_{\mathrm{i}} \equiv J_{\mathrm{i}}/\sqrt{D_{\mathrm{p}}}$ for both currents in the equation. Numerical modeling can then be done in terms of these normalized variables and ratios of normalized currents, as presented later in Figures 4.7 and 4.8. Given the wafer thickness and appropriate diffusion coefficient, the results of the figures can be scaled to the specific experiment of interest.

4.5 Computational Approach to Double-Sided RR

With this preliminary analysis completed, most of the work necessary to model the RR process in SHJCs had been completed. However, a few additional points remained to be addressed. First, real-world devices are not simply boxes of stored carriers; rather, the emitter space charge region provides an additional capacitance that would have to be included before modeled current vs time curves would match experimental ones.

The time-dependent space-charge capacitance is

$$C(t) = \Big[q\epsilon_{\rm s} N_{\rm D}/2(V_{\rm bi} - V_{\rm D}(t)) \Big]^{1/2} \tag{4.26}$$

where $\epsilon_{\rm s}$ is the silicon permittivity, $V_{\rm bi}$ is the built-in voltage of the junction, and $V_{\rm D}$ is the device voltage (as opposed to the externally-applied bias voltage).

In an actual RR experiment, a reverse voltage $V_{\rm r}$ is applied to a resistor $R_{\rm s}$ in series with the device, which besides its stored holes has the aforementioned capacitive component that contributes to the current density J according to the standard equation $J = C/A \cdot ({\rm d}V/{\rm d}t)$, where A is the device area. The governing equation during reverse recovery is thus

$$V_{\rm r} = AR_{\rm s}J_{\rm r} + V_{\rm D} \tag{4.27}$$

$$= AR_{\rm s}[J_{\rm rev,rem} - J_{\rm rev,front} + J_{\rm cap}] + V_{\rm D} \tag{4.28}$$

where the current terminology of Figure 4.2 was employed. Changing our view from macroscopic to microscopic variables, we can modify that last line to write, at the front interface,

$$V_{\rm r} = AR_{\rm s}\Big[- qD_{\rm p}\frac{\partial p(t)}{\partial x} - J_{\rm rev,front}(t=0)\frac{p(t)}{p(0)} + C(V_{\rm D})\frac{\partial V_{\rm D}}{\partial t}\Big] + V_{\rm D} \tag{4.29}$$

where we incorporated Equation 4.15. Note that $V_{\rm D}$ is itself related to the front-interface hole density by $p_{\rm front} = \frac{n_{\rm i}^2}{N_{\rm D}} e^{qV_{\rm D}/kT}$, such that $\frac{\partial V_{\rm D}}{\partial t} = \frac{kT}{qp_{\rm front}}\frac{\partial p}{\partial t}$.

This boundary condition is somewhat complicated. It was incorporated explicitly using a finite-difference ghost-point implementation [70]. Given a 1D lattice of constant Δx and time-steps Δt, and notation $p(x,t) = p(n\Delta x, m\Delta t) \equiv p_n^m$ this method

relies on the discretizations

$$\frac{\partial p}{\partial t} = \frac{p^{m+1} - p^m}{\Delta t} \tag{4.30}$$

$$\frac{\partial p_n}{\partial x} = \frac{p_{n+1} - p_{n-1}}{2\Delta x} \tag{4.31}$$

$$\frac{\partial^2 p_n}{\partial x^2} = \frac{p_{n+1} - 2p_n + p_{n-1}}{\Delta x^2}. \tag{4.32}$$

Equations 4.30 and 4.31 were used along with Equation 4.29 to solve explicitly for the time-evolution of the ghost point p_{-1}; this solution was then used along with the overall governing equation $\frac{\partial p}{\partial t} = D_{\mathrm{p}} \frac{\partial^2 p}{\partial x^2}$ to come up with the explicit time-evolution of the left-side boundary:

$$p_0^{m+1} = p_0^m + \frac{\Delta t D_{\mathrm{p}}}{\Delta x^2}\Big[2p_1^m - 2p_0^m + \frac{2\Delta x}{qD_{\mathrm{p}}}\big(\frac{V_{\mathrm{r}} - V_{\mathrm{D}}}{AR_{\mathrm{s}}} - \frac{J_{\mathrm{rev,front}}(t=0)}{2}\frac{p_0^m}{p(0)}\big)\Big] - \frac{2}{q\Delta x}C(V_{\mathrm{D}})\frac{kT}{qp_0^m}(p_0^{m+1} - p_0^m)$$

The bulk was subject to the standard diffusion equation (with bulk recombination neglected), and the backside interface followed the recombination-determined boundary condition

$$D_{\mathrm{p}}\frac{\partial p}{\partial x} = s_{\mathrm{back}} p. \tag{4.33}$$

Thus the system was completely determined, and transient modeling could be carried out.

4.6 Current-Spreading Effects

The analysis has so-far taken devices to be one-dimensional. However, when bulk recombination lifetimes are long, significant current-spreading can occur, such that appreciable numbers of holes are found outside the column defined by a circular contact area, throwing off 1D-derived results. In this section, a simple method is presented that allows one to reduce these 3D reverse-recovery problems to 1D equivalents. This method is tested successfully on the results of 3D simulations of RR in p^+n diodes.

4.6.1 3D Device Simulations

The simulation results presented in Section 2.4.1 were for 1D devices. The computational construction of those devices is simple: one explicitly defines the thicknesses of the various layers and instructs the Synopsys Sentaurus simulation software to treat the system as one-dimensional. One could try taking the same explicit-definition approach for 3D device construction and meshing, but results would probably be poor: the corners of, say, a sharply-defined rectangular doped region could lead to failed numerical solutions, and results could well be non-physical. Therefore, the devices used for the 3D simulations were themselves constructed using Sentaurus Process simulation software.

From the point of view of minority carriers in the substrate, p^+n devices are similar to short-base heterojunction devices with near-perfect carrier injection efficiency, making them applicable for a study specifically of current spreading. To make these silicon p^+n diodes, the following process steps were simulated, starting with a 300 μm-thick, 10 mm-wide n-Si wafer with phosphorus doping concentration $2 \cdot 10^{15}$ cm^{-3}:

1. Oxide growth via a 90 minute wet oxidation at 1100°C

2. Oxide etch to define the p-Si doped regions

3. Boron implantation (dose $5 \cdot 10^{15}$ cm^{-2}, energy 50 keV)

4. Dopant activation via 1000°C anneal for 20 minutes

5. Silver contact deposition (200nm)

This process simulation resulted in physically-realistic doping profiles, with approximately $p^+ = 10^{20}$ cm^{-3} doping density. Doping-field-dependent meshing was used to ensure simulation convergence. Though the resulting devices were two-dimensional, running simulations in cylindrical mode converted the contacts into circles (viewed from above) and the simulation to, effectively, 3D.

Simulated RR experiments were carried out for various $J_{\text{rev}}/J_{\text{for}}$ ratios ($J_{\text{for}} = 10.1$ mA/cm^2) for devices with circular top contacts of radii 2.5 mm, 1.5 mm, 0.5 mm, and 0.25 mm. To further isolate the effect of current-spreading, zero bulk recombination, zero top-surface recombination outside the device area, and an Ohmic back contact on the entire backside were assumed. Series resistances of 2.5 kΩ, to mimic the series resistors present in actual RR experiments, were also included.

Devices with smaller contacts and emitter regions had uniformly shorter recovery times (Figure 4.5, filled circles) for the same current densities; in those devices, the current is significantly spread out from the contacted device area, and spread-out carriers are less likely to be recovered during the reverse transient. If not accounted for, this effect would lead researchers to conclude that their interfaces are worse than they actually are.

4.6.2 Current-Spreading Model and Results

To quantify this effect, the hole distribution in the substrate during the reverse transient was examined. Figure 4.4(a) shows the forward-bias hole density (in number per cm^2, arbitrary units) obtained by integrating hole density from the top to the bottom of the wafer for a single slice along the device diameter. The total number of stored

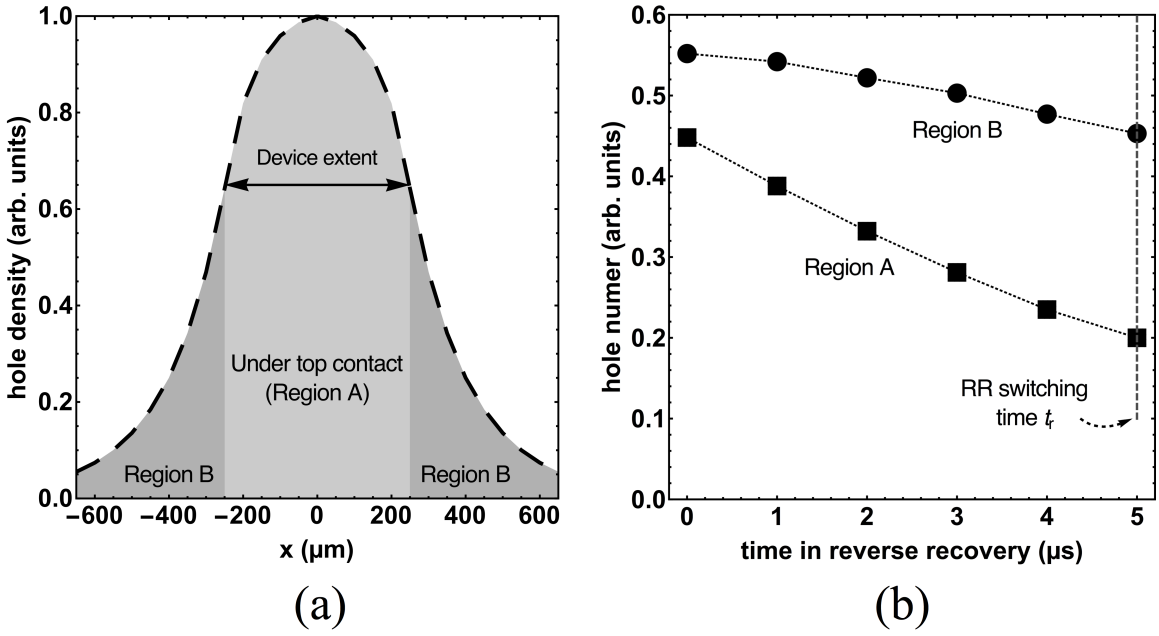

Figure 4.4: (a) Hole density (cm^{-2}, arbitrary units) in forward bias, obtained by integrating hole density from the top to the bottom of the wafer for a single slice along the device diameter, of a simulated 250 μm-radius 3D p^+n-diode, showing region A (under the contacted device area) and region B (the rest of the substrate) and (b) 3D-integrated hole number (normalized to total hole number), in each of the two regions, over time in a RR simulation of this device with $J_{for} = 10.1$ mA/cm^2 and $J_{rev} = 13.4$ mA/cm^2. Despite the fact that initially region B contains 55% of all stored holes, significantly more holes are extracted from region A than from region B during the RR process.

holes could be broken up into two parts: those directly under the device (region A) and those that had "spread out" (region B); totals in each region were obtained by integrating in all three dimensions. During the recovery transient (Figure 4.4b), the number of holes in region B decreases only slightly before the RR time t_r, when the reverse current begins to drop; the recovered holes that supply the reverse transient current come almost entirely from region A. Proportionally, the difference is even greater.

Thus, to a good approximation, the relevant forward current for purposes of RR is that directly under the contact area: those are the carriers that are later collected. In other words, besides the emitter efficiency γ_{front}, one must also include an area

efficiency factor $\gamma_{\text{area}} \equiv \frac{\text{current in region A}}{\text{total current}}$, applying the relation

$$J_{\text{for,1D,eff}} = J_{\text{for,experiment}} \cdot \gamma_{\text{area}}. \tag{4.34}$$

A 1985 paper by Chen et al. [71] investigates the problem of current-spreading in "shallow p-n junctions," the exact scenario of interest here. That paper solves the boundary-variable problem in cylindrical coordinates to obtain series solutions for what they denote the ratio I_{1D}/I_{2D}, the equivalent of γ_{area}. The solution itself is lengthy enough that it is not reproduced here; the important point is that "a significant two-dimensional current spreading occurs" when $a/L_{\text{p}} < 10$, where a is the contact radius and L_{p} is the hole diffusion length. Given that typical test silicon devices are on the millimeter scale, and that diffusion lengths in monocrystalline silicon are themselves on the order of millimeters, it is unsurprising that this effect would find its way into SHJC RR characteristics. Spreading effects are also more significant when W is comparable to or greater than L_{p}; in typical devices, L_{p} and W are of the same order.

To test the γ_{area} method, calculations based on [71] were used to adjust the 3D-simulated total forward current density J_{for} according to Equation 4.34 (J_{rev} remains unchanged, since nearly all reverse current is coming from the region under the contact). The 1D-reduced RR experiments were then modeled according to the method presented in Sections 4.4 and 4.5.

The resulting 1D-modeling curves (Figure 4.5) line up closely with the fully-simulated 3D curves; agreement is nearly exact for $J_{\text{rev}} \geq J_{\text{for}}$, with small deviations for $J_{\text{rev}}/J_{\text{for}}$ (within 10%), showing that this simple procedure can be used to effectively reduce experimental results to their 1D equivalents, which can be more easily analyzed. In summary, then, when devices are small, the modeling described in Section 4.4 must also include this area-modification factor γ_{area}.

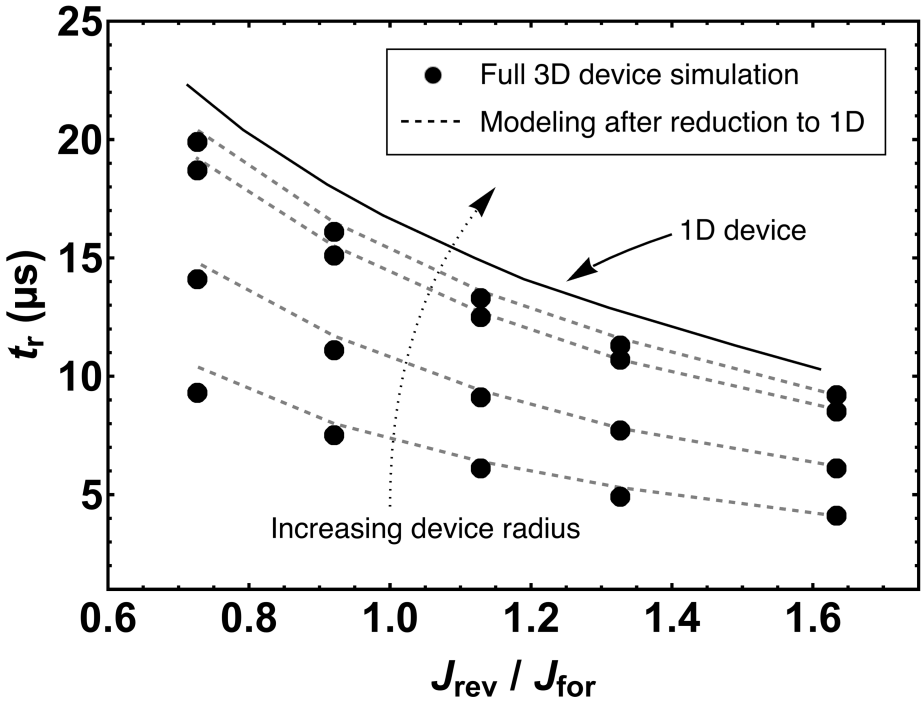

Figure 4.5: Reverse recovery times for short-base p^+n-diodes as a function of $J_{\rm rev}/J_{\rm for}$ by full 3D device simulation and by modeling after reduction to 1D. Also shown is the modeled time for a fully 1D device, which is the limiting case of increasing device radius. Devices are circular with radii (top to bottom) of 2.5 mm, 1.5 mm, 0.5 mm, and 0.25 mm. $J_{\rm for} = 10.1$ mA/cm^2. The size-dependent reduction allows us to analyze the 3D RR problem using its 1D equivalent.

4.7 Generalized Modeling Results

The reverse-recovery process was modeled, and the recovery time $t_{\rm r}$ was calculated, for single-sided and double-sided devices of varying quality. To better mimic the characteristics of real-world devices, and for consistency, for the rest of this chapter the convention will be to define $t_{\rm r}$ as the time when a line fit tangent to the J vs t at $J_{\rm rev}(t) = 0.75 \cdot J_{\rm rev,t=0}$, where $J_{\rm rev,t=0}$ is the reverse current immediately after the bias is switched from forward to reverse, intersects with the horizontal line through $J_{\rm rev,t=0}$, as shown in Figure 4.6. This worked well for the experimental choices illustrated here, but there are regimes in which $t_{\rm r}$ definition proves difficult. For instance, if $V_{\rm r}$ is too small, then the slight decreases in $V_{\rm D}$ over time will significantly change the reverse-current value even during the "constant current" phase of the RR process. Reliable

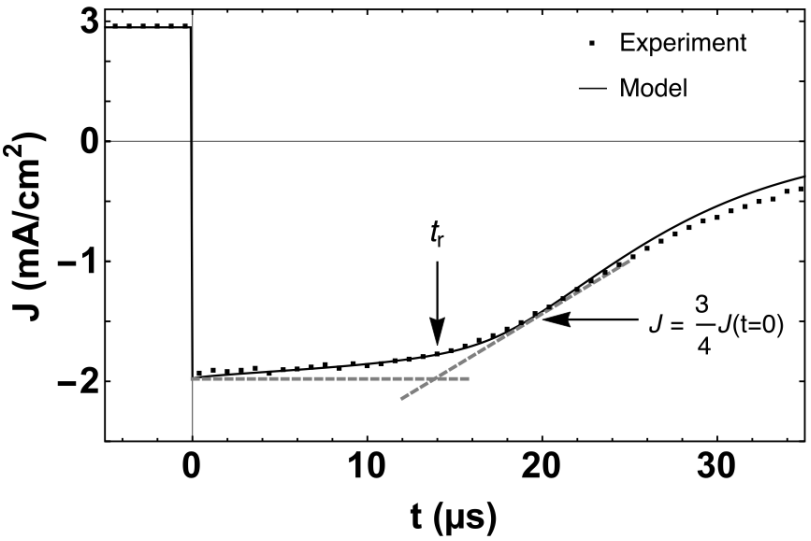

Figure 4.6: Transient experimental and modeled results for a room-temperature PEDOT/n-Si device. Parameters are $J_{\text{for}} = 2.85$ mA/cm^2, $J_{\text{rev,t=0}} = 1.98$ mA/cm^2, $J_{\text{for,front,t=0}} = 0.63$ mA/cm^2, $\gamma_{\text{area}} = 0.88$, and $qV_{\text{bi}} = 0.7$ eV. Note the differing scales for positive and negative current density.

t_r identification could also be confounded if t_r approaches the underlying switching times of the experiment, either because of too-high parasitic capacitances or because of very low stored-charge densities. In those cases, a re-evaluation of the experimental methods would be advised.

The method in Section 4.5 was used to model RR in devices with a perfect front interface ($s_{\text{front}} = 0$) for a variety of base thicknesses and $J_{\text{rev}}/J_{\text{for}}$ values, with results presented in Figure 4.7. For universal applicability, the normalized units derived in Section 4.4.3 were used. This figure (along with Figure 4.8) also uses $n = L/(\sqrt{D} \cdot 2.9 \cdot 10^{-3})$ to label curves; modeling parameters were chosen such that n takes integer values, as shown. This definition is purely for ease-of-reference in these specific figures; the value of the constant has no deeper meaning. Higher values of n require higher values of s_{back} to have the same resulting t_r, since devices using thicker wafers (or devices with smaller diffusion coefficients) store comparatively more charge for the same forward current.

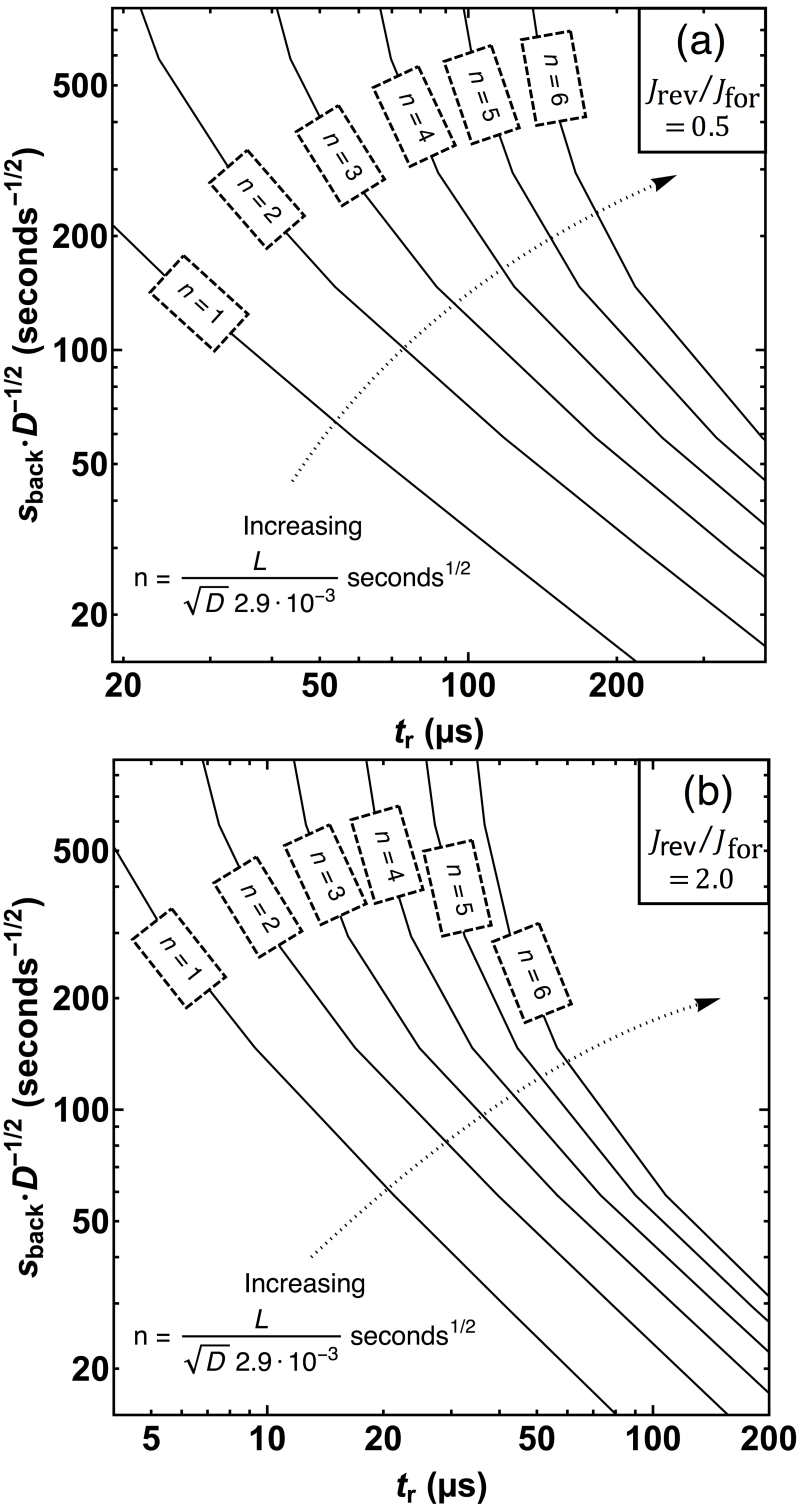

Figure 4.7: Backside recombination velocity s_{back} as a function of RR time t_{r} in a double-sided device with ideal front interface, with (a)$J_{\text{rev}}/J_{\text{for}} = 0.5$ and (b) $J_{\text{rev}}/J_{\text{for}} = 2.0$ for a variety of $L/\sqrt{D}$ ratios.

Given the recovery time t_r, these curves allow the extraction of s_{back} in any device with an ideal minority-carrier injector and a back-interface heterojunction. For example, if one were testing a novel selective back contact to replace the back-surface field of a 0.02 cm-thick n-Si HIT cell, one might measure $t_r = 118$ μs at $J_{rev}/J_{for} = 0.5$; in this case, $n = 2$, and a consultation of Figure 4.7 shows that this result corresponds to $s_{back}/\sqrt{D_p} = 59$ s$^{-1/2}$, or $s_{back} \approx 200$ cm/s. Because thinner devices store comparatively fewer carriers, a 0.01 cm-thick device with the same recovery time would have $s_{back} \approx 100$ cm/s.

Figure 4.8 shows modeled recovery times and their corresponding γ_{SSD} in short-base devices with an Ohmic back contact. To test, for example, the properties of a novel oxide/silicon heterojunction injector with an inverted interface, one could use this information to analyze RR measurements and obtain both injection efficiency and the effective recombination velocity of the metallized junction. Recently, researchers showed how trap-saturation qualitatively affects the dynamics of PEDOT/n-Si junctions during reverse recovery [72]; this method would allow the quantitative interpretation of their results.

4.8 Double-Sided Device with Two Non-Ideal Interfaces

The most complex device situation is a double-sided device with non-ideal interfaces on both sides. The value of γ_{SSD} has significant effects on the extracted value of s_{back}. The two stored-charge loss mechanisms operate in parallel, and a single set of data may be explained by many pairs of their respective magnitudes. An analysis that omits front-interface recombination (i.e., assumes $\gamma_{SSD} = 1$) will be able to be "fit" to experimental RR data, but the extracted s_{back} might be many times larger than the true value obtained by incorporating an accurate $\gamma_{SSD} \neq 1$.

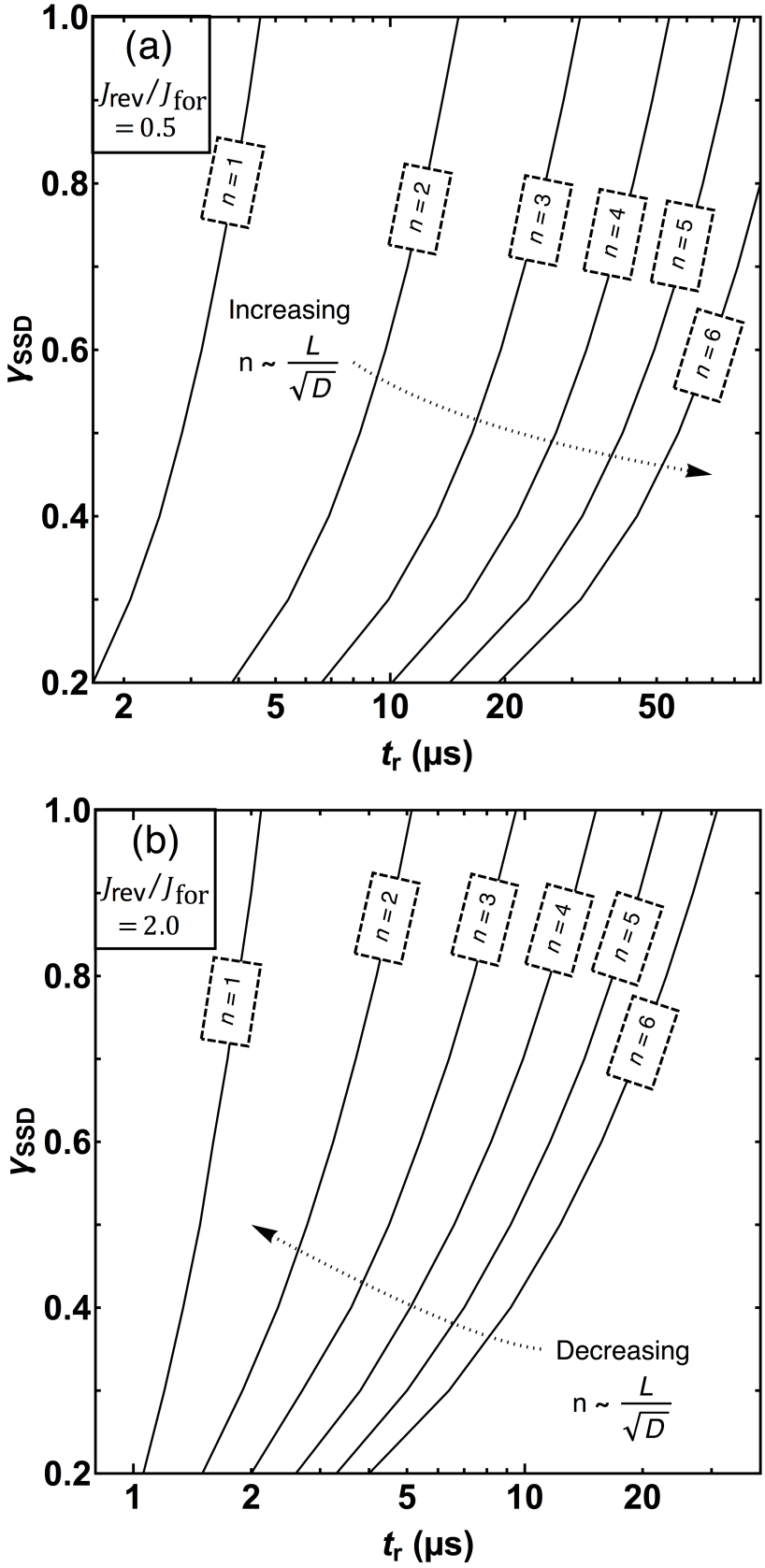

Figure 4.8: Injection efficiency $\gamma_{\rm SSD}$ as a function of RR time $t_{\rm r}$ in a single-sided device, with (a)$J_{\rm rev}/J_{\rm for} = 0.5$ and (b) $J_{\rm rev}/J_{\rm for} = 2.0$, with $n = L/(\sqrt{D} \cdot 2.9 \cdot 10^{-3})$.

To determine the parameters of such devices, the following method may be used:

1. Measure RR in a single-sided device (eg, PEDOT/n-Si) with Ohmic back contact.

2. Fit to 1D short-base diffusion modeling (using Figure 4.8) to find γ_{SSD} and thus $s_{\mathrm{front,eff}}$ as described in Section 4.4.2 and using Equation 4.16.

3. Measure RR in a double-sided device (eg, PEDOT/n-Si/TiO_2) and fit to numerical simulations of the diffusion equation with recombination at both sides (Section 4.5), subject to the condition

$$\gamma_{\mathrm{DSD}} = \left[1 + \frac{s_{\mathrm{front,eff}}}{s_{\mathrm{back}}}\left(1 + \frac{L s_{\mathrm{back}}}{D_{\mathrm{p}}}\right)\right]^{-1} \tag{4.35}$$

   which can be derived by considering the relationship between base hole profile, steady-state current, and the definition of s. This is a coupled fitting; each s_{back} fitting attempt necessitates a new value for $J_{\mathrm{for,front}}$. As an example, this method is applied to experimental data in Section 4.9.

4.9 RR Measurement of Silicon Heterojunction Solar Cells

This section describes the application of the above method to real-world devices. Both single-sided and double-sided devices were fabricated using lightly doped ($N_{\mathrm{D}} \approx 2{\cdot}10^{15}$ cm^{-3}) 0.03 cm-thick float-zone substrates. On the front, 70 nm of PEDOT:PSS was spun on to serve as a hole-selective injector; to make double-sided devices, 3 nm TiO_2 was additionally deposited on the back via CVD to make a hole-blocking contact, with Al contacts deposited via thermal deposition. The device size was 0.4 x 0.4 cm; they were defined by physically scribing the PEDOT. The top contacts were thermally

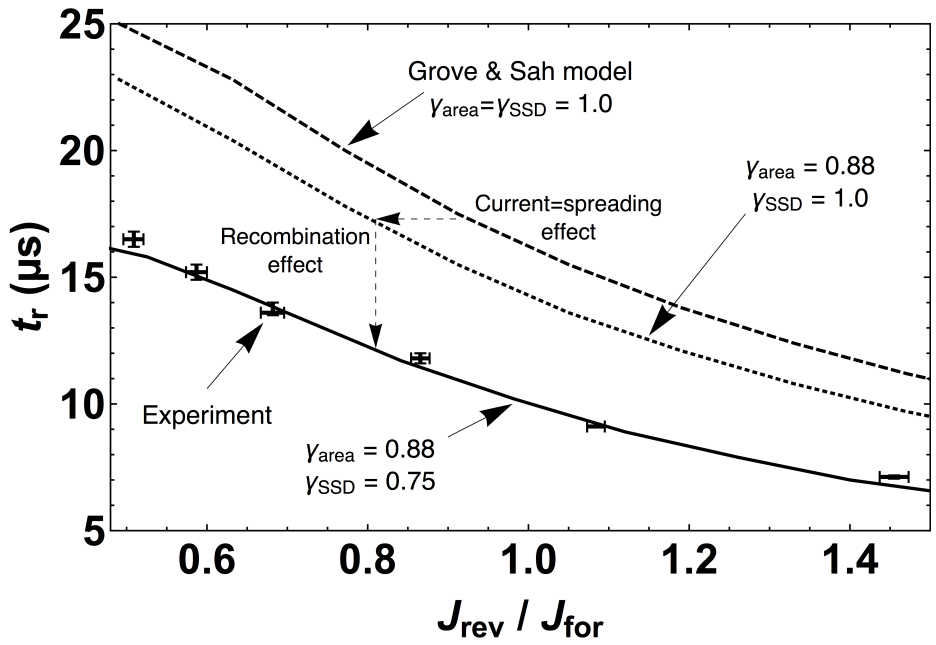

Figure 4.9: Data and model fits for a PEDOT/n-Si device, with $J_{\text{for}} = 2.85$ mA/cm^2 for all data points. The best fit line accounts for both current spreading ($\gamma_{\text{area}} = 0.88$, found by the method in [71]) and front-interface recombination ($\gamma_{\text{SSD}} = 0.75$).

evaporated silver, but note that the electrically-important contact, in terms of defining the anode work function for the device, was the PEDOT layer itself.

Reverse-recovery measurements were carried out over a range of reverse biases for each of several forward biases. The experimental setup was as depicted schematically in Figure 4.1, with connections made using standard coaxial cables. The signal source was a standard function generator set to square-wave mode, and the output signal was read over the series resistor using a two-channel oscilloscope. A variable resistor box was used to allow for resistance variation depending on the needs of the particular device—for instance, to maintain the same low-level-injection forward current densities requires a larger resistor for smaller devices. Despite the ad-hoc nature of the setup, the voltage switching times and parasitic capacitances were relevant only on time scales less than 1 μs, much less than the recovery times of the devices, which were typically 10-100 μs. Based on [9] and [13], the model's built-in voltage was set to $V_{\text{bi}} = 0.7$ V.

We first present results for single-sided devices. Figure 4.9 shows the experimental results for t_r as a function of J_{rev}/J_{for} for a PEDOT/n-Si device with $J_{for} = 2.85$ mA/cm^2. The Grove and Sah prediction of t_r (Equation 4.5), which has no adjustable parameters and implicitly assumes $\gamma_{area} = \gamma_{SSD} = 1$, is also plotted. The results of the Grove and Sah model, which significantly overestimates t_r for all J_{rev}/J_{for}, were then corrected for current spreading (which causes many injected carriers to not be recovered) using the approach of Section 4.6. Including the resulting factor ($\gamma_{area} = 0.88$) to account for current spreading, essentially shifting the t_r vs J_{rev}/J_{for} curve left, is not enough to account for all the overestimation.

The Grove and Sah model also implicitly assumes that all forward current injects minority carriers which are later collected—in other words, that $\gamma_{SSD} = 1$. We adjusted γ_{SSD} to achieve a best fit of the model to experiment, with $\gamma_{SSD} = 0.75$ giving excellent agreement. This means that 25% of the forward-bias current is due to non-ideal recombination at the PEDOT/n-Si interface. This fact has important ramifications for the improvement of V_{OC} via a lowering of the dark current. A strategy that aims only at selectively passivating the backside contact will be able to reduce the hole diffusion current but will not affect the front-interface recombination current, limiting the possible V_{OC} improvement.

Based on this result, Equation 4.16 was used to find $s_{PEDOT/Si,eff} = 130$ cm/s. This value is, it should be emphasized, not due solely to chemical passivation at the interface, but rather to chemical passivation combined with the significant band-bending imposed by the high PEDOT:PSS work function.

Next, reverse-recovery was measured in double-sided PEDOT/n-Si/TiO_2 devices, and the data was modeled using the results of single-sided experiments ($s_{PEDOT/Si,eff} = 130$ cm/s) to include the effects of emitter inefficiency as in Section 4.8 (Figure 4.10). J_{for} was also adjusted to account for the effects of current-spreading as in Section 4.6. There was a single fitting parameter, $s_{Si/TiO2}$, which was found to have a value of

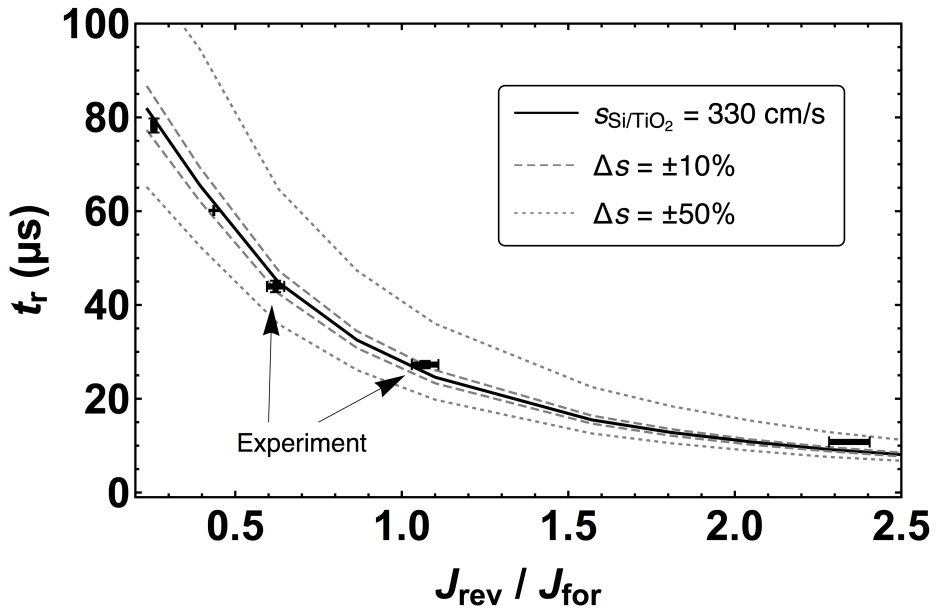

Figure 4.10: Data and model fits for a PEDOT/n-Si/TiO_2 device, with $J_{\text{for}} = 6.35$ mA/cm^2 for all data points, $\gamma_{\text{area}} = 0.79$, and $J_{\text{for,front,t=0}} = 2.11$ mA/cm^2. Dashed and dotted lines show the effect of changing $S_{\text{Si/TiO2}}$ by 10% and 50%, respectively.

$s_{\text{Si/TiO2}} = 330$ cm/s, a number in line with other measurements of that interface (300 cm/s for the EQE method in Chapter 3 and 400 cm/s for the dark-current method found in the same chapter). This demonstrates that the interface remains relatively stable, even after contact deposition. Additional curves in Figure 4.10 show the effects on the model of changing $s_{\text{Si/TiO2}}$ by ±10% and 50%, indicating sensitivity to the fitting parameter. These results suggest that the transient reverse-recovery method can fix the backside recombination velocity of complete, contacted devices to within approximately 10%.

It is important to note that, because of the reduced minority carrier density current in the double-sided vs single-sided device for a given forward bias, the minority-carrier injection ratio is decreased in the double-sided vs single-sided device, from $\gamma_{\text{SSD}} = 0.75$ to $\gamma_{\text{DSD}} = 0.58$; in other words, nearly half of the forward-bias current in this device is due to front-interface recombination.

4.10 Conclusion

This chapter presented a flexible reverse-recovery method to identify the effects of interface recombination at both silicon surfaces in silicon heterojunction solar cells. The analyses shown here allow the direct extraction of the recombination characteristics of novel heterojunction emitters and selective contacts; also detailed is a method for parameter extraction in cases where both front and back interfaces are non-ideal. The analysis is further extended to include the effects of current-spreading in small-area devices. This method is applied to investigate the current mechanisms in PEDOT/n-Si(/TiO_2) devices. The PEDOT/n-Si junction under study was found to be a significantly non-ideal minority carrier injector, with injection efficiencies as low as 75% in devices without a TiO_2 hole-blocking cathode at room temperature. The effective interface recombination velocity at the TiO_2 interface was found to be 330 cm/s, similar to that achieved by back-surface fields created via rapid thermal processing of screen-printed Al layers [73].

These results suggest that the PEDOT/Si junction, while of remarkably high quality for being an organic/inorganic and amorphous/crystalline interface, may become a limiting factor as devices built around it are optimized. If recombination at that interface remains significant, it will be imperative for researchers to develop alternative hole-injecting interfaces that more closely approach the quality of high-temperature diffused junctions.

There is still room to further develop the use of reverse-recovery as an analytical technique. One possible route is to do RR at varying forward voltages, allowing one to track injection efficiency as a function of applied bias. Such a method could potentially allow the mapping of interface states or the identification of multiple current mechanisms. Similarly, one could try temperature-dependent RR: as temperature varies, the relative importance of current mechanisms changes, changing the RR-measured injection efficiency. Illuminated RR could similarly be helpful: the ex-

periments presented here were carried out in the dark, but illuminated devices can exhibit trap-passivation and changed recombination rates that RR could potentially measure. Finally, on the development side, the 1D simulation tools developed here would be even more useful if put together into an easy-to-use software package.

Chapter 5

Development and Characterization of Copper-Doped Nickel Oxide Deposited By Atomic Layer Deposition

5.1 Introduction

Nickel oxide is a prominent electron-blocking material in organic and perovskite devices and of interest for future silicon-based photovoltaic devices. Its usefulness lies in its band structure. NiO, as shown in Figure 5.1, has an electron affinity of about 2.1 eV and an ionization energy of about 5.1 eV (this latter value commonly varies by a few hundred meV) [74]. Compare this to silicon's electronic energy levels (EA= 4.1 eV, IE= 5.2 eV), perovskites' (variable, but approximately EA= 3.9 eV and IE= 5.4 eV) [75], or organics' LUMO and HOMO levels (again, variable, but P3HT has E_{LUMO} = 3.0 eV and E_{HOMO} = 5.0 eV [76]) and it is immediately clear why. The significant ΔE_C and small or negative ΔE_V presented by the NiO will tend

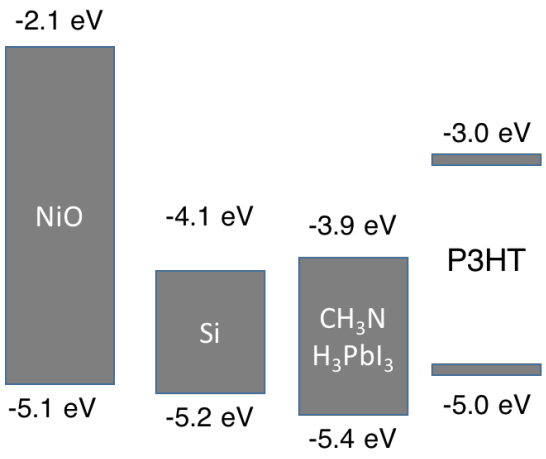

Figure 5.1: Standard conduction and valence band edges (or LUMO and HOMO values) for NiO and some common PV absorber materials. Favorable band offsets allow NiO to serve as a hole transport layer in these devices. Values taken from [74],[75],[76]

to block electrons from flowing out of the active layers but allow holes to pass, setting up a device photovoltage.

Taking advantage of these band alignments, researchers have used NiO in a variety of device architectures and deposited in multiple ways. Pulsed-laser-deposited NiO served as an interlayer between ITO and P3HT in ITO/(NiO)/P3HT:PCBM/LiF/Al solar cells, increasing V_{OC} by up to a volt and significantly increasing device efficiency [77], and sputtered NiO was found to enhance hole injection in ITO/(NiO)/TPD/Alq/LiF/Al LEDs [78]. In more recent work, though, NiO is routinely solution-deposited and used as an important layer in perovskite solar cells and OPVs [79],[80],[81],[82]. Following the theme of earlier chapters of this thesis, some researchers have also investigated the potential of NiO as a selective contact layer in Si-based devices. Notably, sputtered NiO on n-Si was found to somewhat depin the Fermi level at that interface [83]. Sputtered-NiO/c-Si solar cells reached efficiencies of 4.3% without an interfacial SiO_x layer and 10.8% with one [84].

This chapter introduces NiO films deposited by atomic layer deposition (ALD). After an overview on the theory of conduction in NiO, the ALD-NiO process is devel-

oped in terms of precursor choice, oxidizer choice, and deposition temperature. Film stoichiometry as a function of those deposition parameters is explored in detail via XPS measurements.

The resistance to holes flowing in the NiO would be a source of series resistance in a solar cell. Because copper doping of NiO is known to reduce its resistivity, an ALD-CuO process is developed, and the two materials are co-deposited via ALD to produce Cu-doped NiO films. XRD, electrical, optical, and XPS/UPS measurements are combined to give a consistent picture of how copper incorporation affects film properties. Finally, ALD-NiO is incorporated into both silicon-based and perovskite devices, where it proves itself to be an effective electron-blocking material, increasing V_{OC} over a control device. Some preliminary results in this chapter were previously presented in [85].

5.2 Conduction in NiO

Electrical conduction in nickel oxide is a complex topic. Stoichiometric, crystalline NiO has the simple cubic crystal structure of Ni^{2+} and O^{2-}, and a naive analysis of the nickel's unfilled d orbital would suggest that it should conduct. In reality, though, intrinsic NiO is a so-called Mott-Hubbard insulator, a material in which strong electron-electron repulsions between neighboring sites introduce an energy gap [86]. In NiO specifically, d-d interactions ($d^8 + d^8 \rightarrow d^7 + d^9$) introduce a large splitting of the d band into sub-bands, with a gap of 7-9 eV [87]. The conduction band lower edge states are Nid^9. However, the actual bandgap is less than this, some 3-4 eV, because O2p states end up lying in the middle of this gap [87],[88]. This would make it really a charge-transfer insulator rather than a strict Mott-Hubbard insulator. However, there are also Ni3d^8 states in the gap [87], overlapping with the O2p level, complicating matters; as Hufner puts in, "there is very strong hybridization between

the metal d- and ligand p-orbital in the first half of the $3d$ transition metal compounds which makes their classifications in the standard schemes as charge transfer or Mott-Hubbard system not very meaningful."[89] For the purposes of this thesis, the difference is not critical.

Non-stoichiometry in NiO, invariably in the form of nickel vacancies, introduces measurable p-type conduction at room temperature. Conductivity is

$$\sigma = q \cdot p \cdot \mu, \tag{5.1}$$

where q is the electron charge, p is the carrier (hole) density, and μ is the mobility. In general, these latter two terms both vary with temperature, and we can write the conductivity as a function of activation energy E_{A} with

$$\sigma = \sigma_0 e^{-E_{\mathrm{A}}/kT}. \tag{5.2}$$

When a vacancy is introduced into the lattice, charge neutrality requires nearby nickel atoms to change their charge state according to $\mathrm{Ni}^{2+} + \mathrm{O}^{2-} \xrightarrow{\mathrm{V}_{\mathrm{Ni}}^{2-\ (\text{effectively})}} 2\,\mathrm{Ni}^{3+} + \mathrm{O}^{2-}$. If one takes Ni$3d^8$ states to be the valence band top edge, then the hole on the Ni^{3+} is mobile, giving rise to conduction. If, alternatively, one takes O$2p$ to be the valence band, then this step is followed by $\mathrm{Ni}^{3+} + \mathrm{O}^{2-} \longrightarrow \mathrm{Ni}^{2+} + \mathrm{O}^{-}$ [90], in which case the hole on the oxygen is mobile. Regardless, more precisely, the carrier is a polaron, a hole coupled to the lattice around it, "dig[ging] its own potential hole." [86] In materials where the dielectric constant is not too high, and thus in which polaron size is small, the binding energy to free holes from charged acceptors is [86]

$$E = \frac{1}{4\pi\epsilon}\frac{q^2}{d}, \tag{5.3}$$

where ϵ is the material's permittivity and d is the polaron size. In NiO, $\epsilon \sim 12\epsilon_0$ and $d \sim 3\,\text{Å}$ (the approximate unit cell size), giving an estimated conduction activation energy of 0.4 eV.

Because this non-stoichiometry often arises unintentionally during fabrication, reported room-temperature resistivities for so-called "NiO" vary widely, with discrepancies not always acknowledged. Even a slight non-stoichiometry is enough to overwhelm intrinsic conduction, which is closest to being observed in high-temperature-fabricated samples. For example, with significant difficulty, Morin [91] was able to fabricate near-stoichiometric NiO by decomposing $NiCO_3$ in a slightly reducing atmosphere at 1150°C. This gave material with reported room-temperature resistivity around 10^{10} Ω·cm.

In practice, then, reported resistivities for NiO are a function of how many nickel vacancies the fabrication process happened to introduce. For example, one sputtering study reported $\rho_{\text{NiO}} = 1.4 \cdot 10^{-1}$ Ω·cm [92], while another sputtering study found $1.4 \cdot 10^5$ Ω·cm [93]. Reported numbers also vary somewhat for solution-deposited NiO films, from $4.5 \cdot 10^5$ Ω·cm [94] to $7.6 \cdot 10^4$ Ω·cm [95]. The 10^4–10^5 Ω·cm range can be taken to be generally "normal," with variation depending on the particulars of the process.

Besides nickel vacancies, NiO can be intentionally doped to produce a similar effect. Lithium is the most-used dopant; in that case, an Li^+ takes the place of an Ni^{2+}, leading to an Ni^{3+} bound to it and thus a single hole. Copper has also been used, as will be discussed later. Li-doping has been shown to decrease the activation energy of conduction, often significantly, with the effect varying from over 100 meV [96] to only a few 10s of meV [97]. This has been explained [96] by a calculation that views the Li^+-Ni^{3+} system as a dipole that contributes to the system's field response and increases the net dielectric constant.

This might help explain why there has been such significant variation in the reported activation energy for "undoped" NiO films, anywhere from 0.31 eV [98] to 0.6 eV [96]. Just as for the experimentally-reported resistivity numbers, we should view this variation as reflecting the density of nickel vacancies in the film, with higher vacancy densities corresponding to lower activation energies; the 0.3 to 0.4 eV range is typical.

5.3 Motivation for Atomic Layer Deposition

There are a number of reasons to prefer some sort of vapor process for NiO deposition over solution methods. First, atomic layer deposition allows for precise, sub-nm thickness control. This is not always necessary, of course, but in the context of blocking layers on SHJCs it can be quite important. As already discussed, the quality of CVD-TiO_2/Si devices was shown to exhibit a strong dependence on the TiO_2 layer thickness [56], an effect also observed for ALD-TiO_2/Si interface passivation as presented in Chapter 3. In that context, being able to reliably control the thickness is less a bonus than a necessity.

Vapor-deposited films are also, theoretically, conformal, able to remain uniform, with full coverage, even when deposited on rough substrates. The importance of conformality depends on the nature of that substrate. For the simplest possible NiO on c-Si cells, for instance, one would expect the substrate to be smooth already. But more complex cells, such as textured c-Si devices, would require that their blocking layers be able to adapt to some roughness. This is to say nothing of possible applications in non-Si devices, in which the NiO might be deposited on top of rough solution-deposited films.

Temperature considerations are a further reason to explore vapor deposition for NiO. Most solution-deposited NiO is annealed at about 400°C— not a problem for

deposition on ITO/glass stacks, but potentially prohibitive for devices with the anode on top of organic or perovskite layers. Furthermore, even confining the discussion to SHJCs, a 400°C heating requirement removes significant low-temperature design space, potentially problematic when attempting to carefully modulate interface characteristics, and the anneal step could also lead to the formation of SiO_x at the interface, necessarily affecting recombination and charge-transfer characteristics. Finally. a more long-term concern is that lower fabrication temperatures lead to lower costs, again pointing to the potential mass applicability of a low-temperature NiO process.

5.4 Atomic Layer Deposition

While ALD processes vary in their particulars, they generally follow the same general outline. Vapors of the metal precursor are pulsed into the chamber and allowed to attach themselves to oxygen atoms on the substrate, but by design not to other adsorbed metal oxide precursors; the chamber is purged of excess precursor gas; oxidizer vapor is pulsed in and, hopefully, reacts with the precursor; and the chamber is again purged. Beyond that core flow, pulse times, precursor temperatures, chamber and gas-line temperatures, purge times and times between pulses, and the background nitrogen flow can all vary from process to process.

For the processes discussed here, precursors were held in heated stainless-steel bottles connected to a chamber by manifolds that were themselves heated to avoid precursor condensation. Precursor pulse times were in the range of 250–500 ms, while oxidizer pulse times were in the tens of ms. N_2 was kept continuously flowing into the chamber, with the rate varying depending on process step; chamber pressures during deposition were in the 10s to 100s of mTorr. In order to encourage complete and uniform reactions, precursor/oxidizer pulsing was followed by a thirty second hold

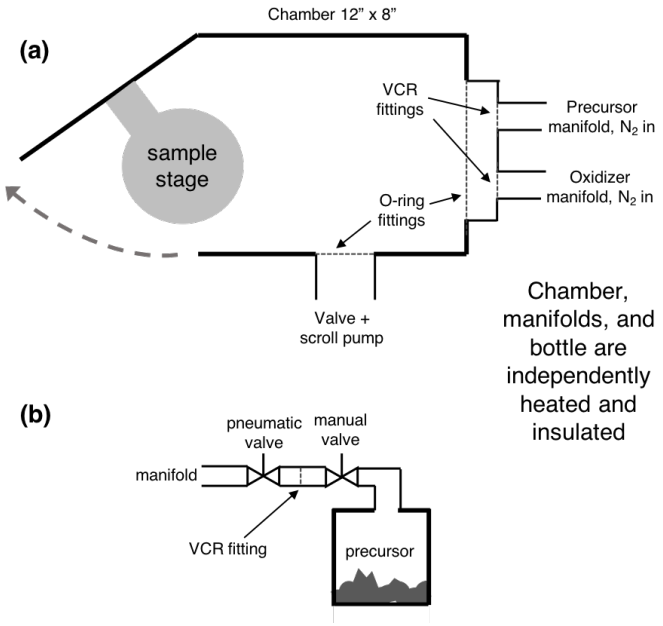

Figure 5.2: (a) ALD tool schematic and (b) precursor bottle schematic, with the two diagrams joining at the manifolds. Pulsing each precursor with the manual valve closed allows that volume to be degassed. The system was manufactured by Arradiance.

time during which the pump valve was closed. The ALD setup is shown schematically in Figure 5.2. The full NiO process can be found in Appendix 8.

It should be noted that the "ALD-NiO" process presented here is probably not "true" ALD. The prototypical example of the latter is the TMA-H_2O process used to grow Al_2O_3. In that process, growth is necessarily layer-by-layer; the chemistry of the film and the precursor means that only a single layer of precursor molecules can attach itself at one time, and furthermore that attachment is highly favored by the thermodynamics of the system. Thus, growth rate is stable within a wide range of process parameters. Here, over the lifetime of these experiments (approximately two years) growth rates were observed to vary from 0.4 Å/cycle to 1.0 Å/cycle without a significant change in process parameters. Furthermore, growth within the chamber was non-uniform. These problems may have been caused by precursor-injection problems, but regardless they would not have occurred in the first place if the process were an ultra-precise, ultra-uniform ALD growth. Therefore, while we use "ALD"

throughout this thesis, that term should not be interpreted to elide the very real process variability that was observed.

5.5 Previous ALD-NiO Work

Even now, and more so at the start of this project, ALD-NiO was a largely unexplored research area, with only a few papers approaching the question. Reactions of bis-methylcyclopentadienyl nickel with H_2O_2 at a substrate temperature of 400°C were used to grow amorphous NiO as early as 1996 [99]; as explained above, that temperature was not desirable for our applications. A wide-ranging study in 1998 [100] tested a number of precursor/oxidizer combinations, with the former including nickel acetylacetonate; only O_3 was shown to produce NiO in any of these experiments, and in general significant precursor decomposition was observed, a problem exacerbated by the low vapor pressures of all the tested precursors.

By the 2000s, researchers were conducting ALD-NiO experiments using the $Ni(dmamp)_2$/$Ni(dmamb)_2$ precursor class (dmamp = 1-dimethylamino-2-methyl-2-propanolate, dmamb = 1-dimethylamino-2-methyl-2-butanolate). $Ni(dmamp)_2$ was used along with H_2O to deposit amorphous NiO at between 90°C and 150°C [101], while $Ni(dmamb)_2$ was used along with O_3 to deposit polycrystalline NiO at between 140°C and 200°C [102]. Unfortunately, these precursors were not available for commercial purchase outside of special orders.

5.6 Process Development

5.6.1 Precursor Choice and Analysis

It was decided that a nickel amidinate (NiAMD, Figure 5.3) solid would be the precursor of choice for the ALD-NiO experiments done here. This air- and moisture-sensitive

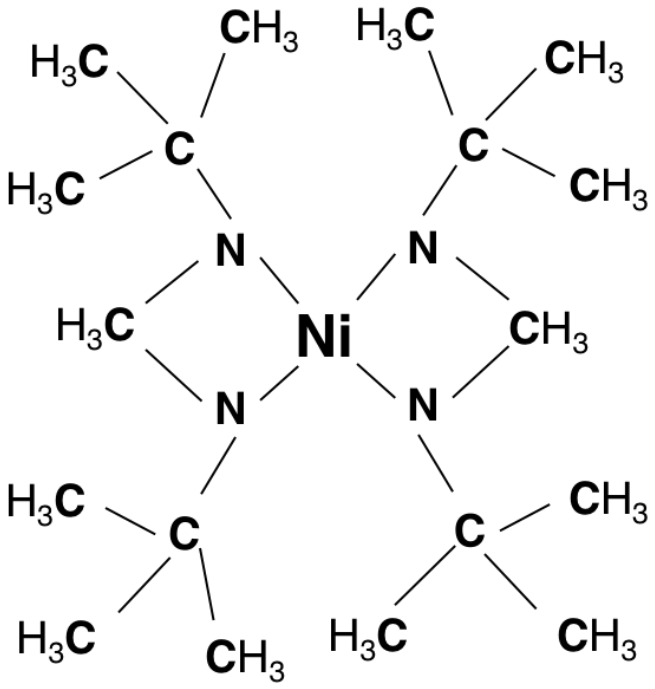

Figure 5.3: Chemical structure of the nickel amidinate precursor used for the deposition of ALD-NiO.

black powder was purchased first from Dow Chemical and later from Strem Chemicals. Industrial test data [103] showed it to exhibit thermal decomposition at 220°C, with vapor pressures of more than 1 Torr at 90°C. Though there was little process data available specifically for NiO deposition, the fundamentals were in place for this precursor to be a reliable and stable source.

To validate the structure of this recently-introduced material, nuclear magnetic resonance (NMR) experiments were carried out. ^{1}H NMR measurements measure the radio-frequency spectra emitted by the spin-split energy-level transitions of hydrogen nuclei held in an external magnetic field. The intensity of those emissions is plotted against the proportional shift of the emissions versus the operating frequency of the NMR tool (δ). Nuclei in a given chemical situation emit at a particular shift (due to the varying effects of electron shielding, magnetic induction, etc), and the ratios of

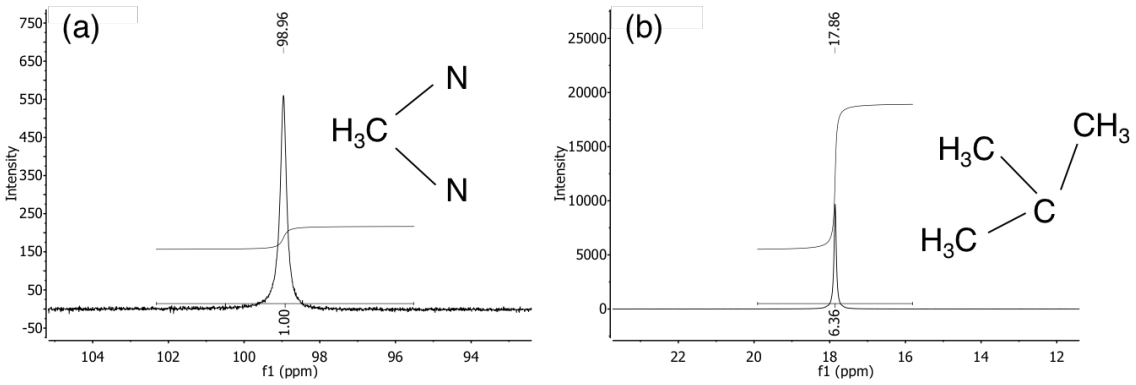

Figure 5.4: Detailed scans of the two notable peaks in the NiAMD-in-C_6D_6 NMR spectrum, with (a) corresponding to the nitrogen-bonded methyl protons and (b) corresponding to the tert-butyl groups.

the signals should be proportional to the ratios of the nuclei count, so NMR can give specific information about how atoms are bonded and configured in a molecule.

The NiAMD was readily dissolved in deuterated benzene (C_6D_6) and ^{1}H NMR measurements taken in the Princeton NMR facility's NB300 tool. The NiAMD molecule is quite symmetric, with ^{1}H nuclei existing in only two situations: as part of the two nitrogen-bonded methyl groups or as part of the four tert-butyl groups. There are six of the former and 36 of the latter, so, to first order, one would expect to see two peaks, with a 6:1 area ratio between them. As shown in Figure 5.4, non-solvent peaks were found at $\delta = 17.86$ and $\delta = 98.96$, with an intensity ratio of 6.36:1, close to the predicted value. The chemical composition of the precursor used, then, does appear to match what would be expected—with, presumably, some impurities as well.

Precursor temperature determines the vapor pressure of the material and thus the amount of precursor vapor injected into the chamber during each pulse. This can have a significant effect, as depicted in Figure 5.5, which shows non-normalized XPS spectra of nickel-containing films deposited via 100 cycles of ALD from precursor temperatures of 95°C and 105°C at a chamber temperature of 175°C and using H_2O as the oxidizer. While XPS signal intensity is expected to vary somewhat experiment-

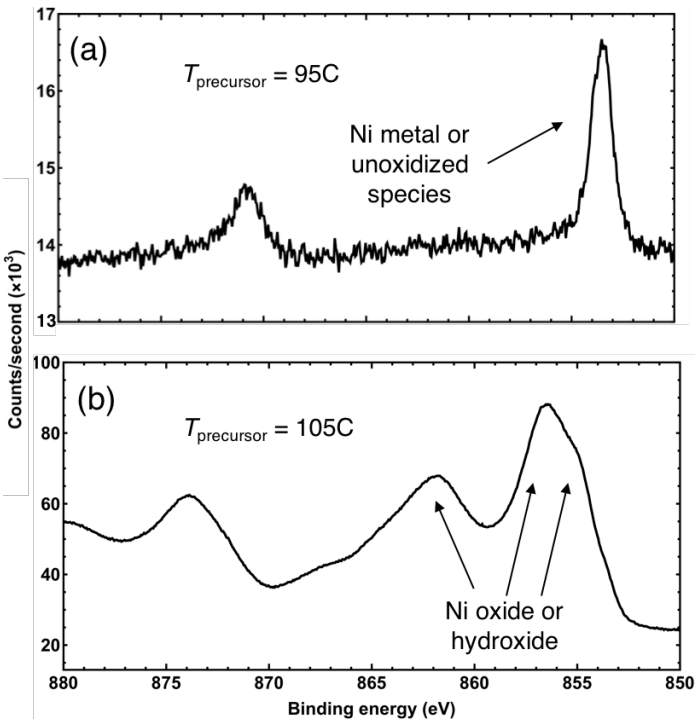

Figure 5.5: Ni2*p* XPS spectra of ALD-Ni-containing films, with water vapor as the oxidizer, deposited at a chamber temperature of 175°C and precursor temperatures of (a) 95°C and (b) 105°C. The low precursor temperature in (a) leads to small precursor vapor pressure in the chamber, and the resulting film generates a signal of smaller magnitude. The change in signal structure is explained by arguing that, at the lower temperature, the "precursor" pulsed into the chamber is mostly composed of higher-vapor-pressure impurities that do not react with H_2O.

to-experiment, Figure 5.5 shows a signal ratio of about 5x, a significant difference—we can conclude that there is significantly less nickel present for the film in Figure 5.5(a) than for the film in Figure 5.5(b). Furthermore, the signals themselves look different, with the 95°C-precursor film resembling metallic nickel or an unoxidized species in a more complex bonding configuration, while the 105°C-precursor film is more like nickel oxide or hydroxide (see Figure 5.6). A possible explanation for this is that, at 95°C, only precursor impurities are pulsed into the chamber, and these molecules (of unknown form) do not react significantly with the water vapor. At 105°C, by contrast, the injected precursor vapor is dominated by NiAMD, which does react with H_2O.

5.6.2 Oxidizer Choice

With a nickel precursor chosen, it was next necessary to determine the oxidizer to be used in the ALD process. The reaction between this oxidizer and the NiAMD would, hopefully, produce pure NiO films with reasonable efficiencies (conversion of NiAMD to NiO) and reaction speeds. The most common ALD-process oxidizers are water, oxygen, hydrogen peroxide, and ozone. Water and oxygen are the easiest to implement, with the former requiring just a bottle filled with water (assuming the room-tepmerature vapor pressure is enough for the ALD reaction) and the latter theoretically requiring just an open valve to atmosphere, though for a more controlled process one could also supply pure O_2. H_2O_2 is a stronger oxidizing agent than oxygen, but its continuous decomposition means that it must often be replaced and can also overpressure whatever closed bottle it is kept in if there is no release valve. Finally, O_3 gas is a very strong oxidizer, but it must produced by an ozone generator during use. For this work, ozone generation was implemented, first, using a system homemade by Ross Kerner, and afterwards using a commercial system. In both systems, pure oxygen passed through a small chamber containing a strong UV lamp that created a significant ozone concentration in the gas. Water, peroxide, and ozone were all tested during the development of the NiO process.

It was not enough to make non-specific Ni-containing films, as in Figure 5.5(b)—the objective was to specifically make *nickel oxide*. Nickel in the film could also be there as part of unreacted precursor, metallic nickel, or nickel hydroxide ($Ni(OH)_2$). Literature Ni2*p* XPS spectra for metallic Ni, γ-NiOOH, $Ni(OH)_2$, and NiO, reproduced with permission from [104], are shown in Figure 5.6. One can see that the curve in Figure 5.5(b) most closely matches the γ-NiOOH spectrum and is not similar to the desired NiO curve.

Therefore, hydrogen peroxide and ozone were also used as oxidizers to attempt ALD growth at 175°C; intensity-normalized XPS results of those depositions, along

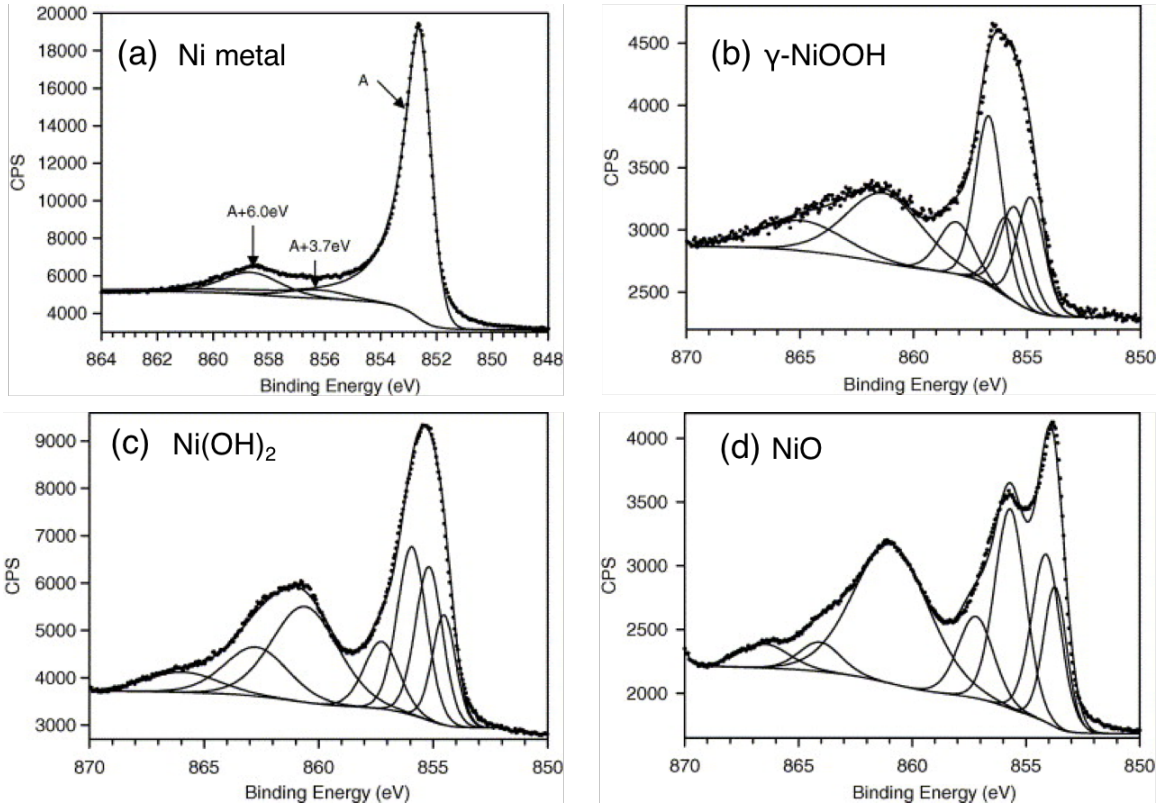

Figure 5.6: Ni2p XPS spectra of (a) Ni metal, (b)γ-NiOOH, (c) $Ni(OH)_2$, and (d) NiO, along with fits. Reproduced from [104] with permission.

with the water deposition, are shown in Figure 5.7. The absolute shifts in the spectra are probably due to the particular charge-correction method used here, as described in the Figure 5.7 caption. More details on Ni2p spectrum interpretation are given in Subsection 5.6.3; generally-speaking, though, stoichiometric NiO is marked by a particular ratio of the double-peak at approximately 855.5 eV binding energy. Only the O_3-deposited film shows the lower-binding-energy (BE) composite peak, located at around 854.3 eV, taller than the higher-BE one at 856.0 eV, as it ought to be for good NiO (cf. Figure 5.6(d)). The other two films show increased hydroxide prominence.

The H_2O-deposited film is apparently closer to a Ni hydroxide species. Given that the H_2O_2 film was closer to NiO than the H_2O film, it is possible that H_2O_2 might produce more stoichiometric NiO at higher temperatures, though that was not

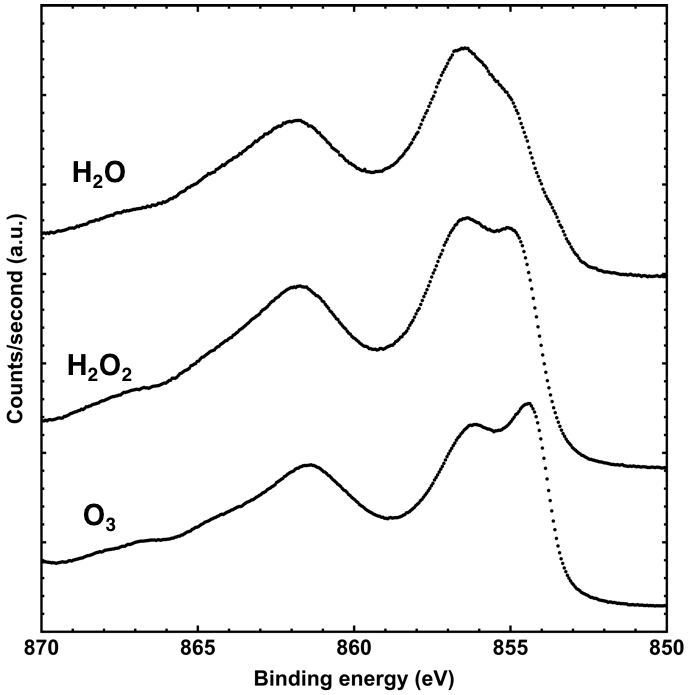

Figure 5.7: Ni2p XPS spectra of films deposited at 175°C using H_2O, H_2O_2, and O_3 as oxidizers. Spectra have been vertically shifted and normalized for easier viewing. Note that, because fine C1s scans were not carried out for these depositions, the standard C1s peak at 284.8 eV could not be used for this purpose, and so spectra were shifted (charge-normalized) using the substrate Si2p peak at 99.4 eV; thus, differences in silicon band-bending at the interface might cause artificial shifting of the spectra.

tested. Finally, the O_3 produced a film quite similar to the reference NiO. The H_2O_2 vs. O_3 results make physical sense: hydroxide species are perhaps more likely if the oxidizer contains hydrogen atoms, as water and peroxide do. Thus, for the rest of our experiments, we chose ozone as the oxidizer for NiO deposition.

5.6.3 Effect of ALD Temperature On Ni Valency

With ozone decided upon as the oxidizer in the ALD-NiO process, it was necessary to define a reasonable process temperature.

Assignment of NiO peaks to specific chemical species is not straightforward. Ideally, there would be clear differentiation between Ni^{2+} and Ni^{3+} peak locations, allowing for stoichiometric analysis similar to that conducted for the NiO-covered TiO_2

in Chapter 3. Unfortunately, there is no such simple correspondence. Of particular note is the characteristic double-peak structure of the NiO $2p^{3/2}$ main peak at approximately 855.5 eV binding energy. While calculations of NiO_6 clusters show just one peak around that location, calculations for Ni_7O_{36} clusters closely reproduce the experimentally-observed asymmetric double-peak [105]. Those authors attribute the difference to non-local electron screening, which is taken into account only in the multicluster simulation. The double-peak structure, therefore, is intrinsic to NiO and not attributable to defects leading to Ni^{3+}, notwithstanding the many papers that make that assignment. Making matters worse, Ni^{3+} is itself theoretically predicted [106] to have a double peak at extremely similar energies.

At the same time, the relative peak prominences *do* seem to differ between Ni^{2+} and Ni^{3+}, with Ni^{3+} exhibiting "a predominance of intensity at higher BE in the main line." [104] In other words, although neither the 854.3 eV composite peak nor the 856.0 eV peak can be attributed either to Ni^{2+} or Ni^{3+}, the 856.0 eV peak *does* seem to increase in prominence when there is more Ni^{3+} in the film. However, quantitative fits/interpretations of this are difficult, and seem to require some flexibility in terms of peak locations and FWHM.

Grosvenor et al. suggest, theoretically, a coupled twelve-peak fitting to the NiO $2p^{3/2}$ XPS curve, with six peaks each from Ni^2 and Ni^{3+} species, based on the calculations in [106]. They successfully used this method to fit the β-NiOOH and $Ni(OH)_2$ spectra. However, they had less success applying this approach to NiO; they achieved a reasonable fit to the NiO spectrum only after allowing both peak positions and FWHM to vary. This approach is fine if one simply wants to fit a curve, but it cannot be the basis of a trustworthy quantitative fitting with precise conclusions drawn about species stoichiometry.

Therefore, following the final fitting in Figure 5.5(d), a seven-peak fitting to the Ni2p spectra was used, as shown in Figure 5.8 and Table 5.1. FWHM was the same for

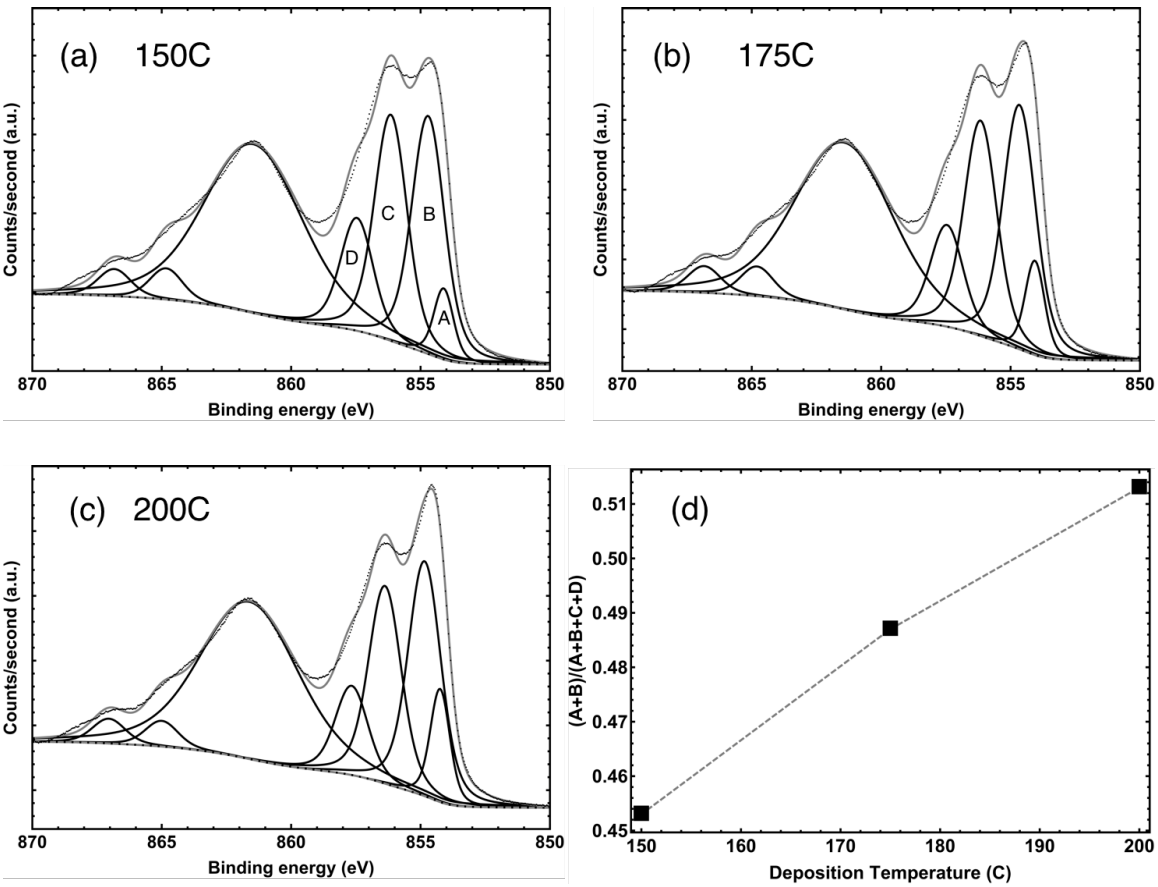

Figure 5.8: Ni2p XPS spectra for O_3-deposited films at (a) 150°C, (b) 175°C, and (c) 200°C. As temperature increases, the lower-binding-energy peaks grow more prominent, indicating increased Ni^{2+} concentration over Ni^{2+} and hydroxides, as shown in (d)

all peaks except for the lowest-BE peak and one peak "associated with other intrinsic losses."[104] As a general analysis, the ratios of peak areas A and B (as labelled in Figure 5.8(a)) to the sum of peak areas A-D was calculated (the remaining three peaks were necessary to include in order to fit peaks A-D properly). As process temperature increased, that ratio decreased (Figure 5.8(d)). This fitting cannot be used to directly calculate a stoichiometric ratio, but the general trend it implies—a decrease in the Ni^{3+}/hydroxide concentration relative to Ni^{2+} as process temperature increases—should hold true.

Peak Label	ΔE (eV)	FWHM (eV)
A	0	0.9
B	0.60	1.5
C	2.10	1.5
D	3.40	1.5
E	7.37	4.5
F	10.74	1.5
G	12.78	1.5

Table 5.1: Relative peak positions and FWHM for the seven-peak Ni2p XPS fit. Acceptable fits were achieved only when curves A and E were allowed to vary their FWHM.

5.7 ALD-NiO:Cu Development

5.7.1 Copper Doping of Nickel Oxide

Conductivity Improvement and Mechanism

As mentioned in Section 5.2, copper has been used to successfully dope NiO, both sputtered and solution-deposited, for device applications. When a Cu^{+} takes the place of a Ni^{2+} in the NiO lattice, it behaves similarly to a nickel vacancy, inducing the formation of Ni^{3+} as described in Section 5.2. Holes, either on that Ni^{3+} or on O^{-}, conduct through band transport in the NiO:Cu valence band.

A co-sputtering study [107] found that resistivity decreased exponentially with Cu%, down to less than 0.1 Ω-cm at Cu%=18 (they did not report the resistivity of undoped NiO because "it is difficult to measure the electrical resistivity using the four point probe because of an ultra high ρ value when the Cu content in NiO:Cu film is lower than 6.97 at.%"). However, this drastic effect could be due to Cu metal precipitating into percolation paths in the film, as indicated by their conductive-AFM data. More typical is the result found in the spray-pyrolysis study of Moghe et al. [98], who observed about an order of magnitude decrease in resistivity upon 10% copper

doping (that paper does not provide absolute resistivity numbers). Notably, they observe a decrease in electrical activation energy with Cu-doping, from $E_A = 0.31$ eV for NiO to $E_A = 0.28$ eV for 10%-NiO:Cu. Measurements carried out in another NiO:Cu sputtering study [92] show a 2x resistivity decrease, from $1.4{\cdot}10^{-1}$ Ω-cm for NiO to $7{\cdot}10^{-2}$ Ω-cm for 18%-NiO:Cu. It should be noted that that undoped resistivity number is extremely small relative to what it should be in the natively-insulating NiO, so there must be high numbers of nickel vacancies in the film. This is a common feature of the NiO:Cu literature—conductivity changes are rarely referenced to high-resistivity NiO, which, as was said above, has resistivities of something like 10^{10} Ω-cm in the absolutely stoichiometric case, or 10^5 Ω-cm in the case of only small numbers of nickel vacancies.

In other examples, a 2015 study [94] found that Cu doping decreased solution-deposited film resistivities from $4.5 \cdot 10^5$ Ω-cm to $1.2 \cdot 10^3$ Ω-cm. Cu-doping was also shown to increase grain size. In another study [93], resistivity in sputtered films was modified from $1.4 \cdot 10^5$ Ω-cm to $3.8 \cdot 10^3$ Ω-cm. They also observed an ionization energy decrease of about 150 meV (the same as the work function decrease) with doping as compared to without. Finally, a recent paper [108] compared theoretical band calculations with experimental data for solution-deposited NiO:Cu films. They observed no change in the Ni2p XPS signal but found a significant Cu^+ component; the latter is consistent with substitutional $Ni^{2+} \rightarrow Cu^+$ doping. They also report that both carrier concentration and mobility improved by about 10x upon copper doping, though it should be noted that their experimental IV curve shows a conductivity increase of only 6x (rather than the extrapolated 100x from those components of conductivity). Their oxide work function increased by 100 meV with copper doping.

The aforementioned results raise reasonable questions about the Cu-doping mechanism itself. Previously-reported copper-doping effects are summarized in Table 5.2. The multiplicative conductivity increase from Cu-doping varies from 2x to 300x, with

Method (reference)	NiO ρ (Ω-cm)	Cu%	NiO:Cu ρ (Ω-cm)	ρ Reduction Factor
Spray pyrolysis [98]	—	10%	—	20
Sputtering [92]	$1.4{\cdot}10^{-1}$	18%	$7{\cdot}10^{-2}$	2
Solution deposition [94]	$4.5{\cdot}10^{5}$	5%	$1.2{\cdot}10^{3}$	375
Solution deposition [93]	$1.4{\cdot}10^{5}$	5.7%	$3.8{\cdot}10^{3}$	37
Solution deposition [108]	9.8	5%	$3.4{\cdot}10^{-2}$	288

Table 5.2: Deposition methods and effect of copper doping of NiO as reported in previous work.

little consistency even when looking only at the solution-deposited films. If NiO:Cu is an extrinsic doped semiconductor then the final conductivity should not depend on the initial conductivity, modulo a constant addition. One possibility is that the substitutional doping sought by copper inclusion can only happen for films that would already have a high density of nickel vacancies, and thus a high "intrinsic" conductivity.

Application of Copper-Doped NiO in Perovskite Solar Cells

Most recently, NiO:Cu has found application in perovskite solar cells. As discussed in the introduction to this chapter, NiO(:Cu) is a good HTL in these devices. It is generally deposited on the bottom ITO electrode, above which the perovskite layer is spin-deposited. [94] reported that their perovskite devices with 5%-NiO:Cu as HTL exhibited improvements in both fill-factor (from 58% to 72%) and V_{OC} (from 1.08 V to 1.11 V) over NiO devices. Researchers in [93] observed a similar effect: fill-factor improved from 57% to 67%, and V_{OC} went from 0.99 to 1.06 V. Similar results were found in [108].

The exact mechanism behind device improvement upon copper doping is unclear. Papers sometimes focus on changes in film conductivity, but, given the thickness of their NiO:Cu films and the changes in conductivity reported, it seems unlikely that this alone would make the difference. Besides, according to standard theory, series-

resistance changes should be reflected in device fill-factors, not V_{OC}. It is more likely that doping-induced band-structure modification is causing these changes. Work-function deepening that accompanies p-doping could improve device performance by increasing the internal electric field. Simultaneously, if doping reduces the oxide ionization energy, then charge extraction from the absorber into the oxide would be improved, another potential reason for improvement.

NiO:Cu deposited via ALD has not been previously demonstrated and is an important novel contribution of this thesis.

5.7.2 ALD-CuO

Before attempting the deposition of NiO:Cu, we first developed a process for the deposition of the simpler CuO. CuO was deposited by ALD using $Cu(dmap)_2$ (Bis(dimethylamino-2-propoxy)copper(II)), purchased from Strem, as precursor and O_3 as an oxidizer. The precursor was held at 105°C, and substrate temperatures of both 175°C and 200°C were used at different times, with similar results. When H_2O was tested as an oxidizer, no copper deposition was observed. In the final process, ozone was used (with source as described in Section 5.6), pulse times were 250 ms for the $Cu(dmap)_2$ and O_3, and 30 second hold times were also used after pulsing (see the Appendix for details). This process produced films with a growth rate of approximately 0.13 Å/cycle.

Because of the process described in Section 5.2, the precise form and valency of the copper, as incorporated into NiO:Cu, is crucial for determining the doped film's conductive properties. Cu^{2+} can simply take the place of Ni^{2+} in the lattice without any electrical changes, whereas Cu^{+} enables the charge-transfer that causes p-doping. To serve as a reference, XPS spectra of CuO, Cu_2O, and $Cu(OH)_2$ powders (purchased from Sigma-Aldrich, and put under vacuum only a few minutes after opening) were measured, as shown in Figure 5.9. As described in the thorough reference [58], CuO

and $Cu(OH)_2$ are notable for their distinctive satellite peaks, but differentiation between those species requires more careful analysis; the $Cu(OH)_2$ primary feature is shifted to slightly higher binding energies than that in CuO. The Cu_2O exhibits a strong, narrow single peak (along with a trace CuO component). A 150-cycle deposition of ALD-CuO at 175°C, along with the measured reference spectra, all referenced to the C1*s* peak at 284.8 eV, are shown in Figure 5.9.[1]

ALD-CuO at 175°C (Figure 5.9(d)) and 200°C (not shown) is a mixed copper oxide with a dominant CuO contribution and a small hydroxide contribution. The primary peak in the distribution is centered in-between the CuO peak and the Cu_2O peak, and has significantly lower-BE than the $Cu(OH)_2$ peak. At the same time, the distinctive shelf at $\sim$ 935 eV is probably from a small hydroxide contribution, something not uncommon in ALD-deposited films, particularly at the surface. Finally, the distinctive shape and size of the satellite region is further evidence that the film is dominated by CuO, with some Cu_2O and an even smaller $Cu(OH)_2$ component.

5.7.3 ALD-NiO:Cu

With an ALD-CuO process defined, co-ALD of CuO and NiO was carried out at 200°C in an attempt to copper-dope the NiO films. Each independent process' parameters were preserved, except than an extra-long (two minute, as opposed to the usual thirty second) purge was carried out in between copper/nickel or nickel/copper transitions. Copper cycles were spread out evenly throughout the process: for example, in the 25%-by-cycle NiO:Cu process, the meta-cycle was three NiO cycles followed by a CuO cycle. Furthermore, the final cycle was never a copper cycle; rather, a few extra NiO cycles always composed the very top of the films. This was done to minimize

[1] Experimentalists should note that x-ray exposure apparently causes rapid decomposition of CuO and $Cu(OH)_2$ films, so Cu2*p* scans should be taken quickly and should be first in the measurement queue.

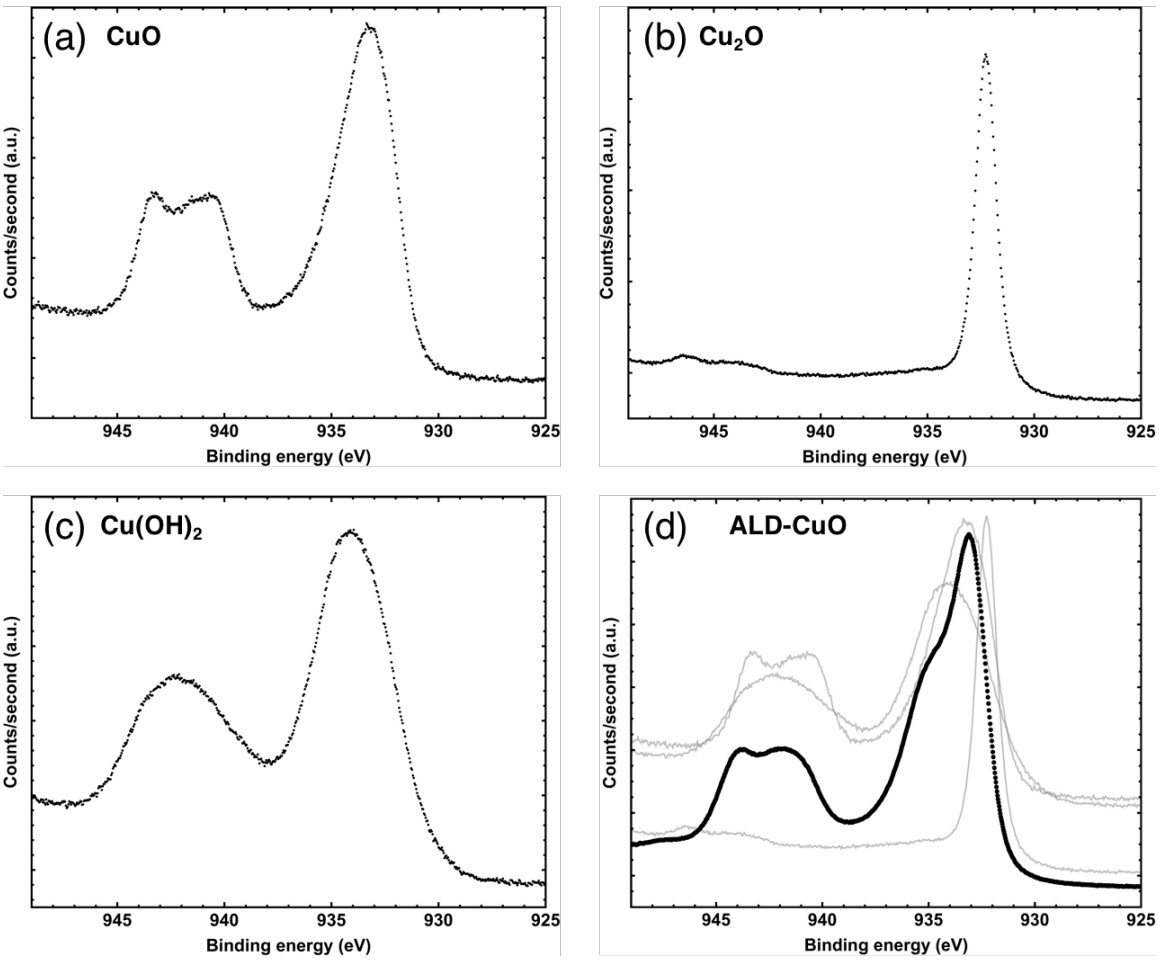

Figure 5.9: Cu2p XPS Spectra of (a) CuO, (b) Cu_2O, (c) $Cu(OH)_2$ (all three purchased from Sigma Aldrich), and (d) CuO film ALD-deposited at Princeton University. The position of the main peak in (d), the relative prominence of the satellite structures, and the higher-BE shelf on the primary structure at approximate 935 eV indicate that the film is a mixed oxide dominated by CuO and containing some hydroxide impurities, perhaps at the surface. (d) spectra have been intensity-normalized by their peak near 933 eV.

the chances that unexpected reactions could take place between surface copper and whatever material would later be deposited on top.

Copper incorporation could affect either or both of the copper and nickel valencies. XPS analysis was again used to determine what effects, if any, occurred. We focus first on the incorporated Cu2p signal. The co-ALD depositions always produced Cu2p spectra closely resembling the reference CuO curve, as shown in Figure 5.10(a). As previously discussed, Cu is thought to take the place of Ni atoms in the NiO cubic lattice; it may be that that process is so energetically favorable that it dominates impurity and mixed-oxide processes. Regardless, we can say confidently that the copper, as incorporated into ALD-NiO, is overwhelmingly present as Cu^{2+}.

Note that this approach thinks of the structure as, generally, homogenous. Another approach would conceive of the NiO and CuO as separate layers, with "doping" resulting from charge transfer between those layers. Because each ALD cycle deposits less than a monolayer, and we therefore expect the oxide to be significantly mixed, we did not use the latter approach; however, incorporating it into the analysis could be useful as NiO:Cu is further researched.

Any Cu^{+} incorporation into the NiO lattice could, as previously discussed, lead to the formation of Ni^{3+}. Therefore, the fits of Figure 5.8 were repeated for NiO:Cu films, and the Ni2p peak intensity ratio $(A+B)/(A+B+C+D)$ were calculated, as indicated in Figure 5.8(a). Those ratios for several films of various stoichiometry are shown in Figure 5.10(b). Results are inconclusive: the only films showing significant deviation in the ratio include copper dopants, but two different 25%-Cu-by-cycle films have peak ratios lying within the range set by the NiO results. Either these results are anomalous or there is a non-unidirectional effect occurring as Cu% increases. More experiments would be useful for determining what is actually happening.

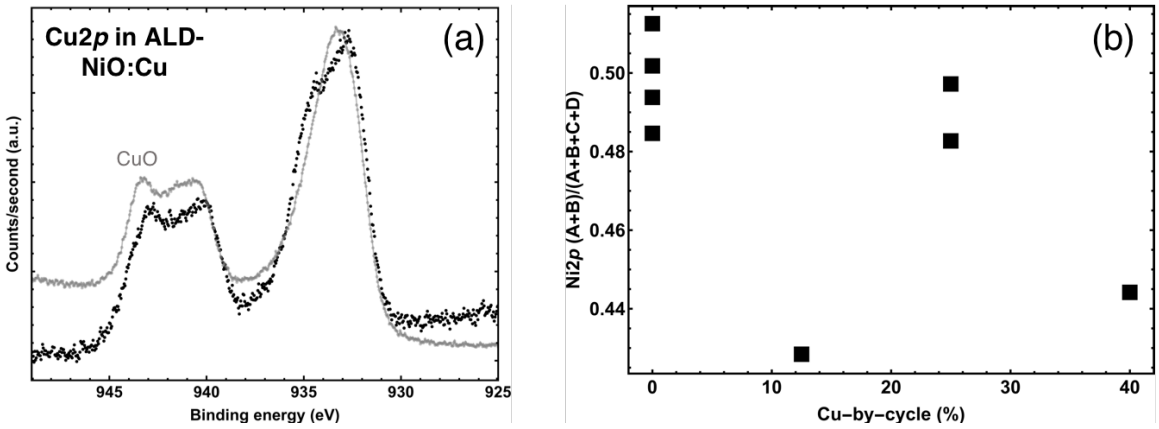

Figure 5.10: (a) Cu2p XPS Spectrum for a 25%-Cu-by-cycle ALD-NiO:Cu film, and reference CuO spectrum. The close fit indicates the copper is being incorporated as Cu^{2+}. (b) Ni2p peak-area ratios, with peaks defined as in Figure 5.8(a).

5.8 NiO:Cu Stoichiometry and Structure

One could, from known deposition rates and cycle ratios, try to predict Cu% in the NiO:Cu films. However, deposition rates can easily differ from their standard values when doing co-deposition, so it was desirable to directly measure Cu% in the films. XPS, with its ability to probe surfaces, was ideal for this. As discussed in Section 2.5.2, proportions of XPS intensities (i.e., areas under curves) can be used to determine proportions of elements or chemical species. In this case, fits were made to the O1s, Ni2$p_{3/2}$, and Cu2$p_{3/2}$ features. Peak fits used a Shirley background [109] modified to avoid the background intensity ever being higher than the actual data, and sensitivity factors were taken from the Avantage "ALTHERMO1" library.

Film stoichiometries are shown in Table 5.3. All films have O% slightly above 50%; this is typical of NiO, regardless of fabrication method, and in the NiO film the resulting Ni vacancies give rise to conduction. As Cu%-by-cycle rises to 12.5% and 25%, the copper content in the film itself increases linearly, at a rate of about $\text{Cu\%}/(\text{Cu\%} + \text{Ni\%}) = 0.4 \cdot \text{Cu-cycle\%}/(\text{Cu-cycle\%} + \text{Ni-cycle\%})$, and O% remains constant. At Cu%-by-cycle = 40%, though, the Cu% is significantly higher than it should be based on that linear trend. XRD measurements of the films, as described

Cu cycle%	Cu%	Ni%	O%	Cu%/(Cu% + Ni%)
0	0	45.9	54.3	0
12.5	2.5	43.5	54.0	5.4
25	4.6	41.2	54.2	10.0
40	13.8	33.5	52.7	29.1

Table 5.3: NiO:Cu stoichiometry, as extracted from quantitative XPS analysis. Cu% increases linearly up to 25%-by-cycle, but by 40%-by-cycle the Cu% increases faster than cycle%.

in the paragraphs immediately following this one, show that polycrystalline CuO precipitates out by Cu%-by-cycle = 40%; perhaps a "critical mass" of Cu in the film enhances the $Cu(dmap)_2$-O_3 reaction, leading to the observed stoichiometric and crystalline effects.

It was also important to examine how exactly the copper was being incorporated into the film. Possibly the copper could be precipitating out, with the film being a mixed CuO-NiO oxide rather than a true copper-doped NiO film. To probe the films' crystal structure, x-ray diffraction experiments were carried out in a Bruker D8 Discover tool in the Princeton PRISM Imaging and Analysis Center (IAC). The ∼20 nm films were too thin for standard XRD; instead, grazing-incidence diffraction (GIXRD) measurements were used. The sample was held at 0.5° from the incoming beam, such that the incoming x-rays were localized in the film. The resulting XRD spectra (Figure 5.11) were low-intensity and noisy, but they were enough to reveal clear crystalline characteristics.

The ALD-NiO is polycrystalline, with (111), (200), and (220) peaks all visible. At 10% Cu%/(Cu%+Ni%) that crystallinity seems to break down somewhat, with peaks growing broader (as expected) and less distinct, but no Cu-specific peaks are visible. Finally, at 30%-Cu, the NiO signal is significantly suppressed and several CuO peaks ((110), (002), and (200)) appear, indicating that at least some copper oxide is precipitating into its own polycrystalline phase. Note that the broad peak

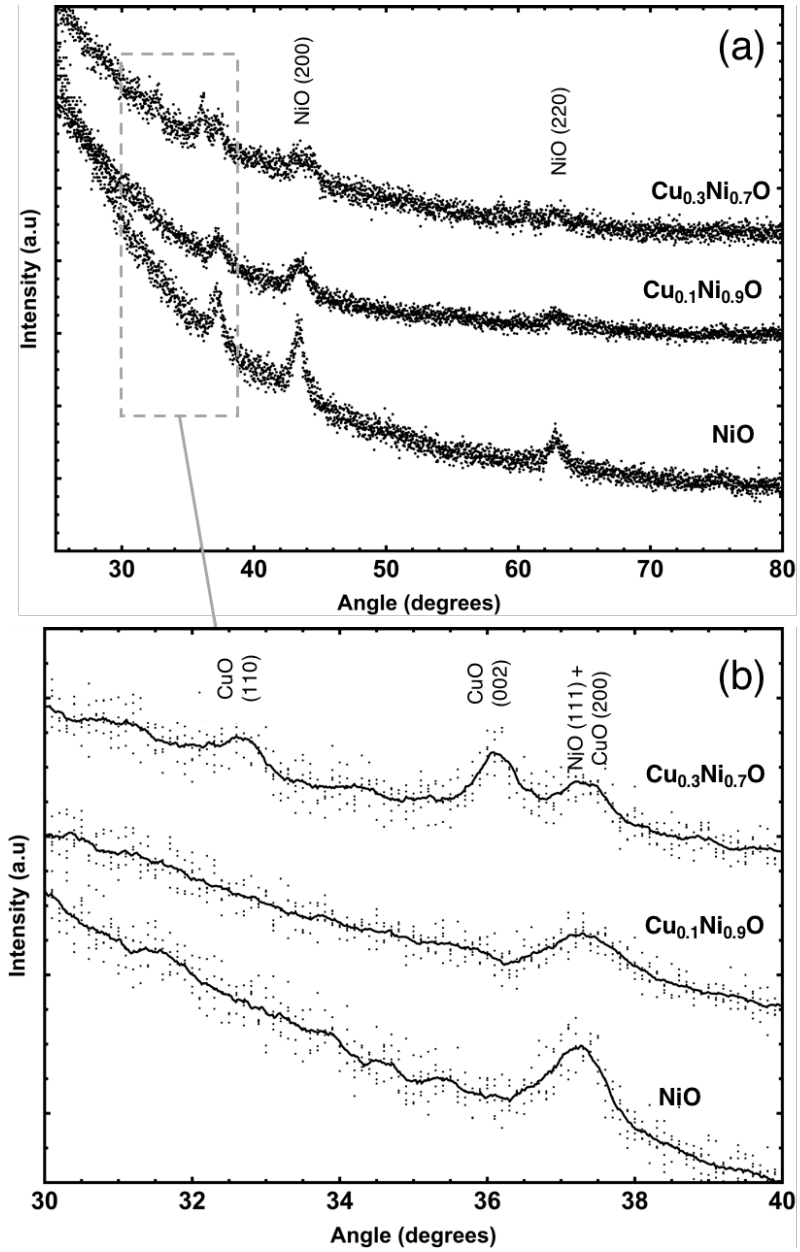

Figure 5.11: Full (a) and zoomed-in (b) GIXRD data of NiO and NiO:Cu films. The NiO has a polycrystalline structure, with no CuO peaks visible at lower concentrations. At 30%-Cu NiO:Cu, the CuO precipitates out and CuO peaks are visible.

at approximately 37.5° is likely a combination of NiO and CuO peak signals, as indicated, especially for $Cu_{0.3}Ni_{0.7}O$.

5.9 NiO:Cu Optical Measurements

In order to measure the dependence of the films' optical bandgap on copper doping, optical absorption (absorptance) measurements were taken of ALD-NiO:Cu films deposited at 200°C with O_3. All films were between 23 and 28 nm thick and deposited on top of glass slides with minimal absorption into the UV range. Absorption spectra, taken with a Cary 5000 spectrometer in the Princeton PRISM IAC, are shown in Figure 5.12. From this absorption data, extinction coefficients α were calculated with $\alpha = (1/t)\ln(1 - \text{abs})$, where t is the film thickness and abs is the absorptance. The Figure 5.12 inset shows the Tauc plot of $(\alpha h\nu)^2$ vs photon energy in eV, where h is Planck's constant and ν is the photon frequency; fits to the linear portion of this curve have an energy-intercept corresponding to the direct optical bandgap [86]. Those fits, along with a linear fit to the extrapolated bandgap as a function of Cu% in the films, are shown in Figure 5.13.

In the visible range, absorption increases with increasing Cu%, as is typical in doped-oxide samples. At the lowest wavelengths measured, though, the $Cu_{0.1}Ni_{0.9}O$ film is most absorptive, possibly corresponding to its increased conductivity (and increased hole density), as described in Section 5.10. The optical bandgaps decrease almost linearly with increasing Cu%; a best-fit line has the formula being $E_G = 3.66 - 0.01\text{Cu}\%$. Note that this relation implies that a pure CuO film would have a bandgap of 2.66 eV; this is significantly higher than reported measurements of the bandgap of CuO, which are typically in the range of 1.6 to 2.0 eV. This implies that the Cu is being incorporated directly into the NiO film rather than existing alongside it as a separate oxide. This is consistent with the x-ray diffraction experiments

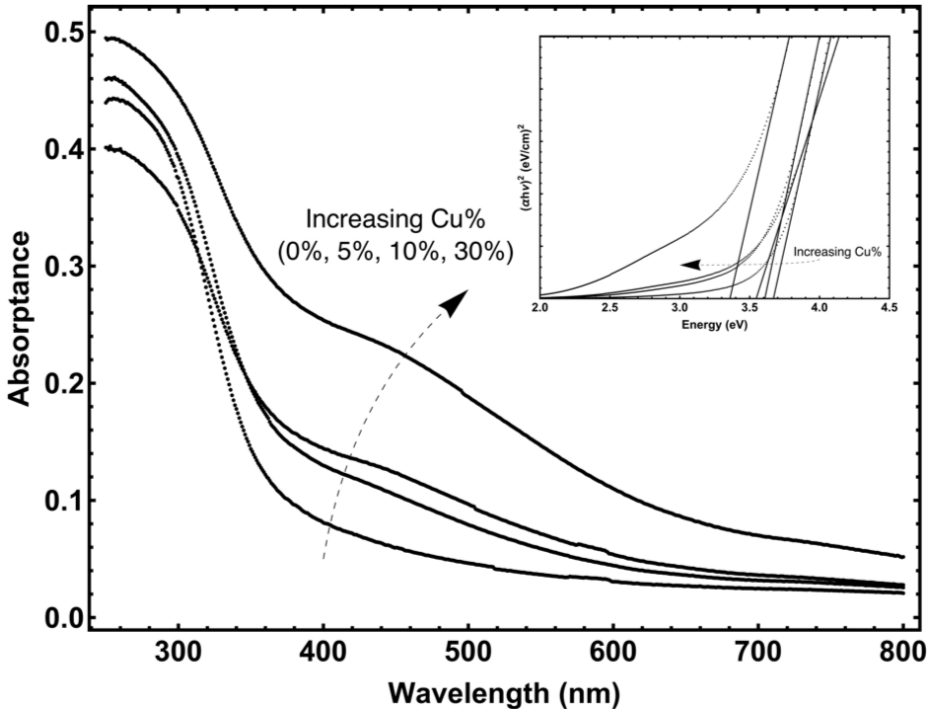

Figure 5.12: UV-Visible absorption spectra for NiO:Cu. At higher wavelengths, increased Cu% corresponds with increased absorption, though at the lowest wavelengths the 10%-NiO:Cu film is most absorptive. As shown in the inset Tauc plot, though, optical bandgaps trend directly negative with increasing Cu%.

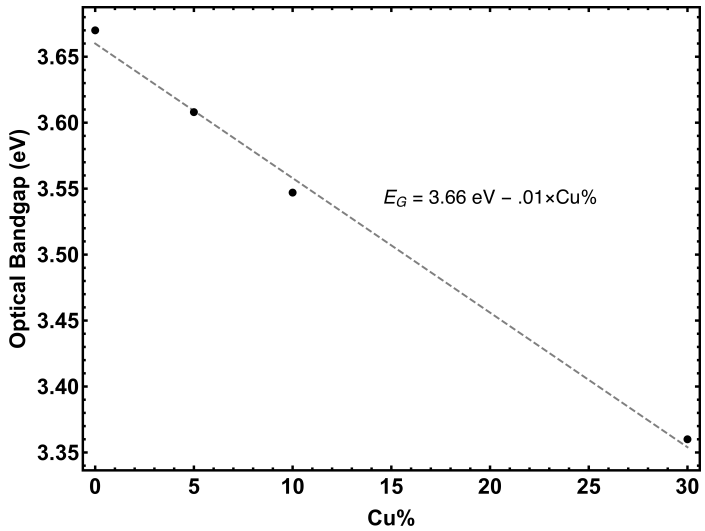

Figure 5.13: Optical bandgaps of NiO:Cu films as a function of Cu%, along with a linear fit. The extrapolated bandgap for Cu%= 100% is 2.66 eV, significantly higher than literature values, implying copper incorporation into NiO rather than separate-phase mixed oxides.

presented above, where films with Cu%= 10% showed no copper precipitation, while such precipitation is seen at Cu%= 30%. However, increasing disorder in the lattice might also be contributing to the bandgap changes. Iincreased absorption in the sub-bandgap part of these UV-Visible absorptance curves implies a high defect density falling off from a band edge into the gap. UPS measurements, to be discussed later (Section 5.11), also showed band tails.

5.10 Conductivity Enhancement by Copper Doping and UV-Ozone Treatment

The NiO and NiO:Cu films were analyzed by temperature-dependent IV measurement. The deposited films were both very thin and highly resistive. That, plus the expectation that making Ohmic contact to the oxide films might be a problem, necessitated the fabrication of dedicated IV test samples. A two-point bottom-contact arrangement was chosen, with film "resistors" of three different lengths on each sample allowing for contact resistance vs film resistivity separation, as shown in Figure 5.14.

Contact-sample fabrication followed a standard photolithography process. An approximately 140 nm-thick wet silicon oxide was grown on silicon wafers at 1050°C. Next, AZ5214 image-reversal photoresist was spun on and patterned, after which titanium/gold contacts (10 nm / 100 nm) were deposited, and finally a liftoff step left only the desired metal contacts on the chip. The resulting patterned wafer was then diced into individual samples for growth of the NiO or NiO:Cu. The oxide layer thicknesses were $\sim$ 20–25 nm.

Besides the copper-doping, UV-ozone post-treatment was also investigated as a way to boost NiO film conductivity. Such a method makes good physical sense: this treatment could further introduce nickel vacancies, thus p-doping the film as

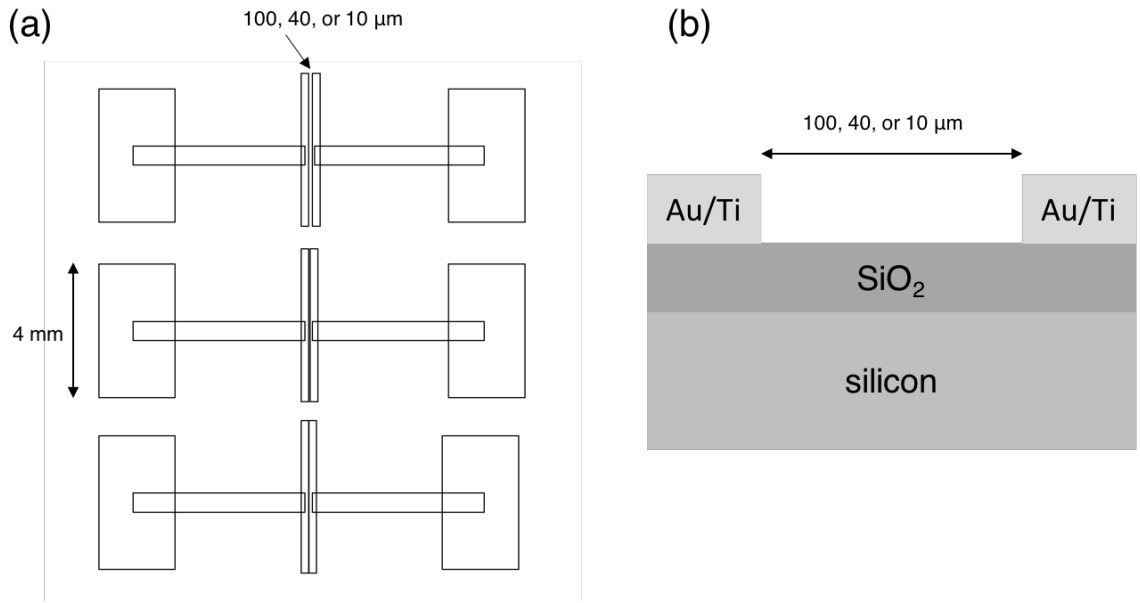

Figure 5.14: (a) Top view drawing of IV test-sample contact pattern and (b) side view of the same. The three different effective films lengths allows for the extraction of film resistivity numbers from two-point measurements. NiO or NiO:Cu was deposited on top of this structure.

previously discussed. The efficacy of this treatment was previously reported in [110]. These films were treated to 20 minutes of UV-ozone exposure in a UVOCS tool in the Princeton MNFL.

IV measurements were carried out on a temperature-controlled stage under vacuums of about 10^{-4} Torr from room temperature to about 100°C in the voltage range of about -1 to $+1$ V. A linear fitting to the IV curve gave the resistance for each two-point contact pair; this resistance is $R = R_{\text{contact}} + R_{\text{square}} * L/W$, where R_{contact} is the contact resistance, R_{square} is the sheet resistance, W is the film width, and L is the film length. Thus, the film sheet resistance could be extracted from the slope of the R vs L line, and, given the ellipsometrically-determined thickness, the film resistivity could be extracted. Example room-temperature IV data and temperature- and length-dependent film resistances for a 25%-by-cycle NiO:Cu film are shown in Figure 5.15(a) and (b), respectively.

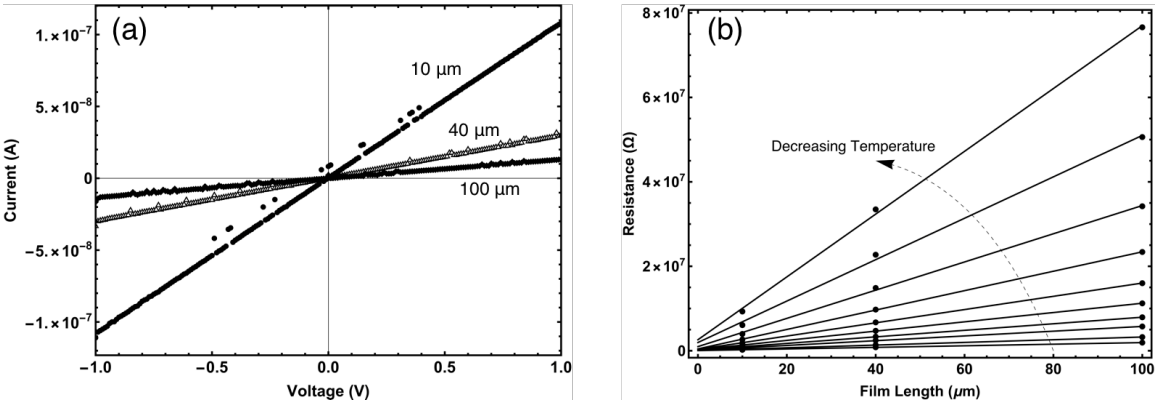

Figure 5.15: (a)Room temperature IV measurements of the 25%-by-cycle NiO:Cu film, with film lengths of 10, 40, and 100 µm, and (b) measured resistances, as a function of film length, for that sample for each temperature. Extrapolation to the R axis gives the contact resistance, and the slope can be used to calculate the film-specific resistivity.

Film conductivities, as a function of inverse temperature, are shown in Figure 5.16. Copper doping successfully increased the film conductivity by about an order of magnitude over the undoped case, with the 25% and 12.5%-by-cycle films more conductive than the 40%-by-cycle film. Compared to the undoped sample ($\rho = 5.2{\cdot}10^4$ Ω-cm), the 25%-by-cycle, or $Cu_{0.1}Ni_{0.9}O$, film had a resistivity of $5.4 \cdot 10^3$ Ω-cm at room temperature. This order-of-magnitude conductivity increase is similar to the improvement previously reported for solution-deposited NiO:Cu films; however, this still places the final conductivity multiple orders of magnitude above what has been reported for some Li- or Cu-doping studies. As expected, UV-ozone treatment also significantly decreased the resistivity without the incorporation of copper, from $5.2{\cdot}10^4$ Ω-cm to $2.4{\cdot}10^3$ Ω-cm.

In Chapter 2 it was suggested that, given 10 nm oxide layers, resistivities could be, at the most, 10^6 Ω·cm, assuming no other significant series resistances in the device. Even our undoped NiO has resistivity significantly less than this ($\sim 10^4$ Ω·cm), meaning that the as-deposited ALD-NiO, even without copper-doping, could be a viable layer in solar cells.

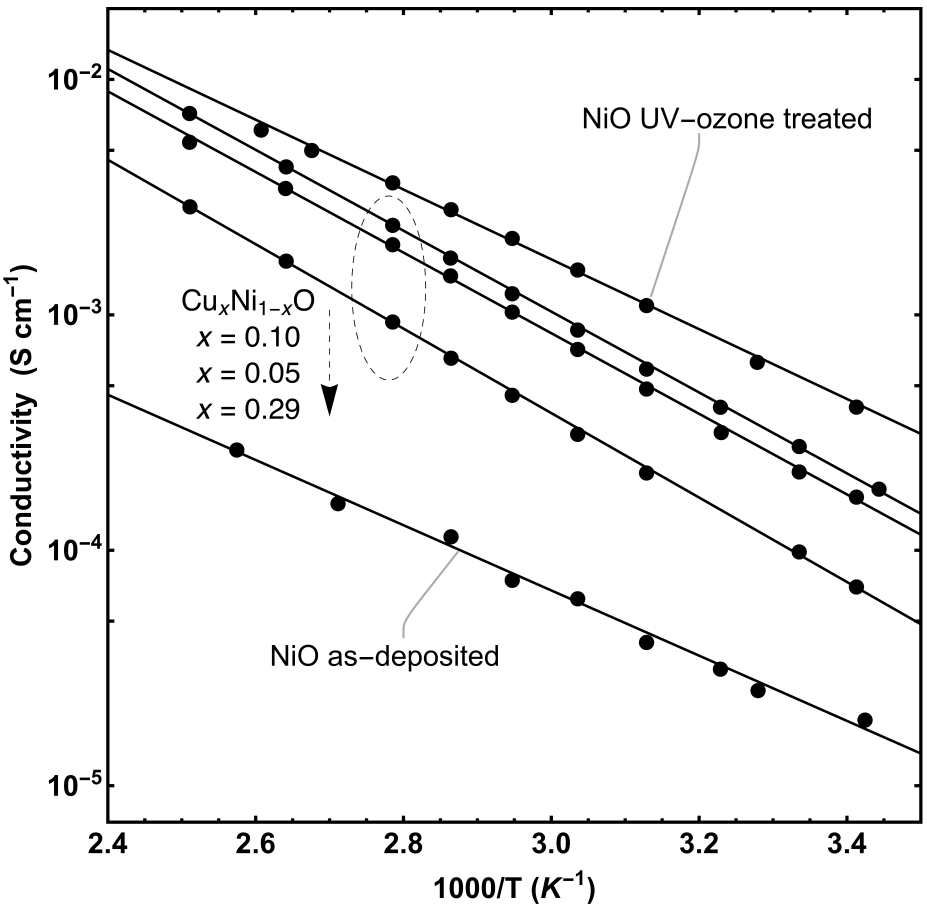

Figure 5.16: Temperature-dependent conductivities for NiO, NiO:Cu, and UV-ozone-treated NiO, along with Arrhenius fits. Both UV-ozone treatment and copper doping increase the conductivity, but the differing activation energies point to some difference in transport mechanism or film dielectric environment.

To determine the physical mechanisms underlying this conductivity change, Equation 5.2 was used to determine the activation energy E_{A} for each of the films. As shown in Figure 5.17(a), NiO and UV-ozone-treated NiO that I grew have very similar activation energies (0.275 and 0.29 eV, respectively); as their conductivities should both derive from Ni vacancies, this makes sense. As previously discussed, theoretical estimates for vacancy-induced hole conduction in NiO put E_{A} at 0.4 eV, but experiments comparable to ours report it being 0.31 eV [98]; our results, then, are consistent with prior work. The NiO:Cu films, though, have activation energies of $0.34-0.36$ eV. Recall that typically Li- and Cu-doped NiO films have an E_{A} that goes

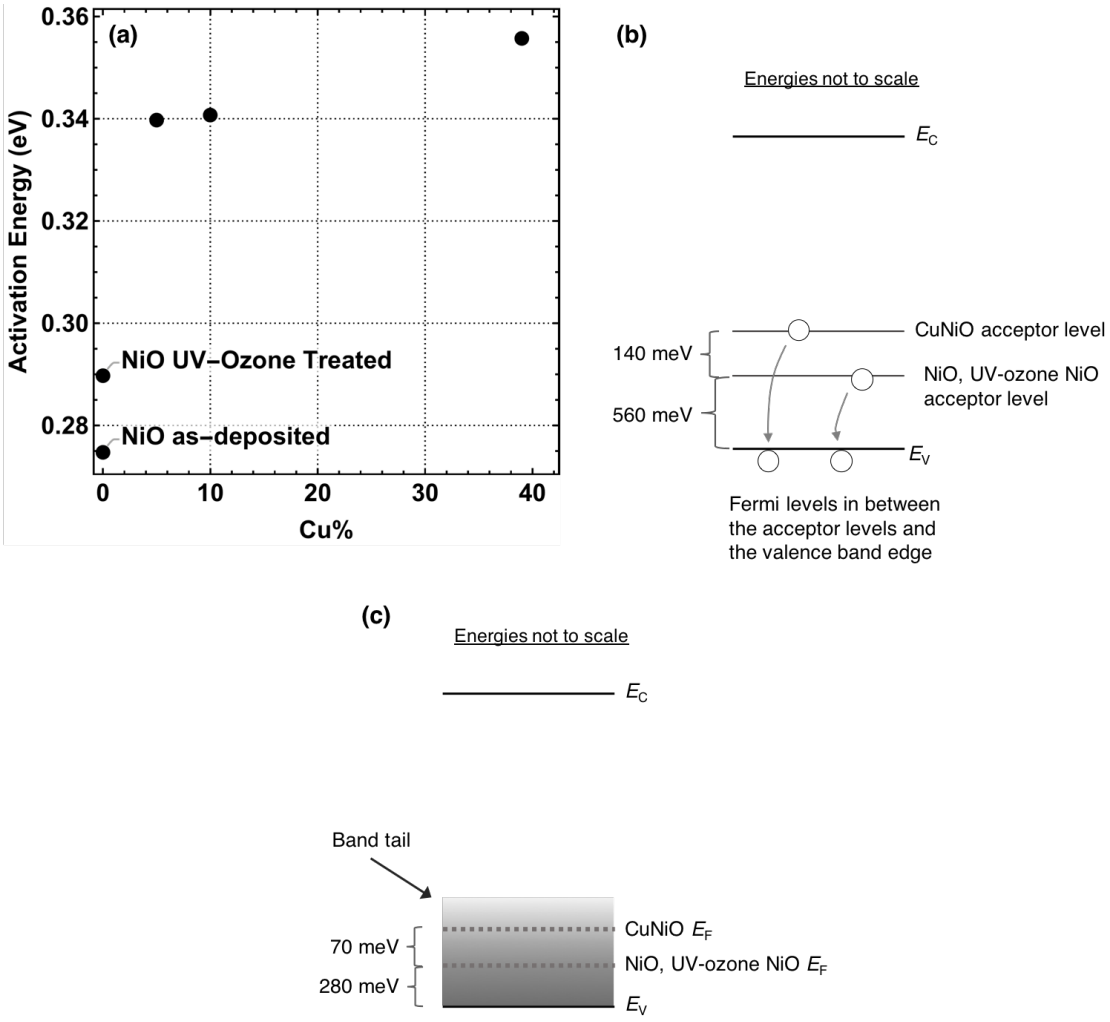

Figure 5.17: (a) Extracted activation energies E_A for undoped, UV-ozone-treated, and Cu-doped ALD-NiO films. Interestingly, Cu-doping increases the activation energy. (b) Band-energy picture of hole conduction in NiO(:Cu), using the measured activation energies, assuming the system is in "freeze-out" with acceptor levels contributing to band conduction. (c) Band-energy picture assuming conduction occurs through localized states in a band tail extending from the valence band.

down upon doping; as explained in Section 5.2, this is interpreted as due to dipole screening.

At room temperature, typical extrinsic semiconductors, like acceptor-doped silicon, have a conduction activation energy deriving almost entirely from mobility's temperature dependence. Besides high-temperature intrinsic conduction, it is only at low temperatures ("freeze-out") that activation energy becomes strongly temperature-

dependent, with $E_A = (E_{acceptor} - E_V)/2$ [111]. The Fermi level in the case of freeze-out is located between the acceptor level and the valence band edge. Freeze-out occurs at higher temperatures when the acceptor level is further from the band edge; a schematic of band energies assuming that case is shown in Figure 5.17(b). Alternatively, the gap states observed in both the UV-Visible absorption data (Figure 5.12) and the UPS data below (Figure 5.18) make it probable that the Fermi level is to be found in a band tail of localized states extending out from the valence band. In that case, because states fill/unfill as the Fermi level moves above/below them, the activation energy is $E_A = E_F - E_V$ (assuming the Fermi level does not move significantly in the measured temperature range). By this interpretation, the *IV* activation-energy measurements are actually revealing the Fermi level location relative to the valence band, as depicted in Figure 5.17(c).

The unexpected result that activation energy goes down upon copper doping may be related to a breakdown in crystallinity, with polaron binding energies changing as the structure changes. On the other hand, polarons in NiO are small (approximately unit-cell sized), and so the structural breakdown would probably have to be greater than what was observed by XRD in Section 5.8. An alternative hypothesis is that the Cu-doped films do not feature significant Cu^+-Ni^{3+} screening. Indeed, Austin and Mott's calculation [96] suggests that the decrease in the activation energy scales as $N^{1/3}$, where N is the acceptor/dipole density. Given an approximate hole mobility of 1 $cm^2/V{\cdot}s$, the room temperature conductivity of the $Cu_{0.1}Ni_{0.9}O$ film corresponds to a carrier density of 10^{15} cm^{-3}, a number very low compared to the copper density. The doping is apparently highly inefficient, with the vast majority of copper atoms incorporated as inactive Cu^{2+}. The aforementioned calculation predicted a ΔE_A of 30% for 1% Li^+ density; therefore, the E_A lowering for our extremely low Cu^+ density would be essentially nothing. The only flaw in this line of reasoning is that the undoped NiO and UV-ozone-treated NiO films would then be interpreted to themselves

feature screening, something not predicted by the ion-dipole theory. Screening could possibly exist in those films in the form of Ni^{2+}-Ni^{3+} dipoles, though the densities do not seem to match what would be required and nowhere in the literature is this effect mentioned.

Finally, film anneals in nitrogen, on separate samples, were also tested. For 175°C-deposited films, both NiO and NiO:Cu (25% by cycle), resistivity was shown to increase with temperature after annealing in N_2 at 250°C and 300°C; for the 300°C anneal, the measured resistance increase for NiO:Cu was 5x, and the resistance increase for NiO was 1.5x. Nitrogen annealing is expected to slightly reduce the NiO, decreasing the Ni:O ratio and thus reducing the intrinsic conductivity. Separate experiments that attempted nitrogen anneals at 400°C and 600°C produced films with very high resistivities that approached the measurement limits of the system.

5.11 NiO:Cu Band-Energy Measurements

The conductivity changes in the previous section ought to correlate with changes in band structure. Indeed, the ability of copper doping to modify band energies is a core part of the motivation for developing it in the first place. As discussed in Subsection 5.7.1, Cu-doping has been reported to have a mix of effects, both on the valence-band energy and on the work function.

UPS measurements (discussed theoretically in Subsection 2.5.3) were taken of ALD-NiO:Cu films on top of highly-doped p^{++} silicon. To ensure good electrical contact, the silicon backside was metallized with 100 nm of evaporated gold, and the samples were pressed down onto a conductive stage with a separate gold film on one section of it; UPS measurement of that gold film defined the work function for the entire system. The sample stage was kept at a relative bias of -8 V in order to fully separate out the sample signal [26].

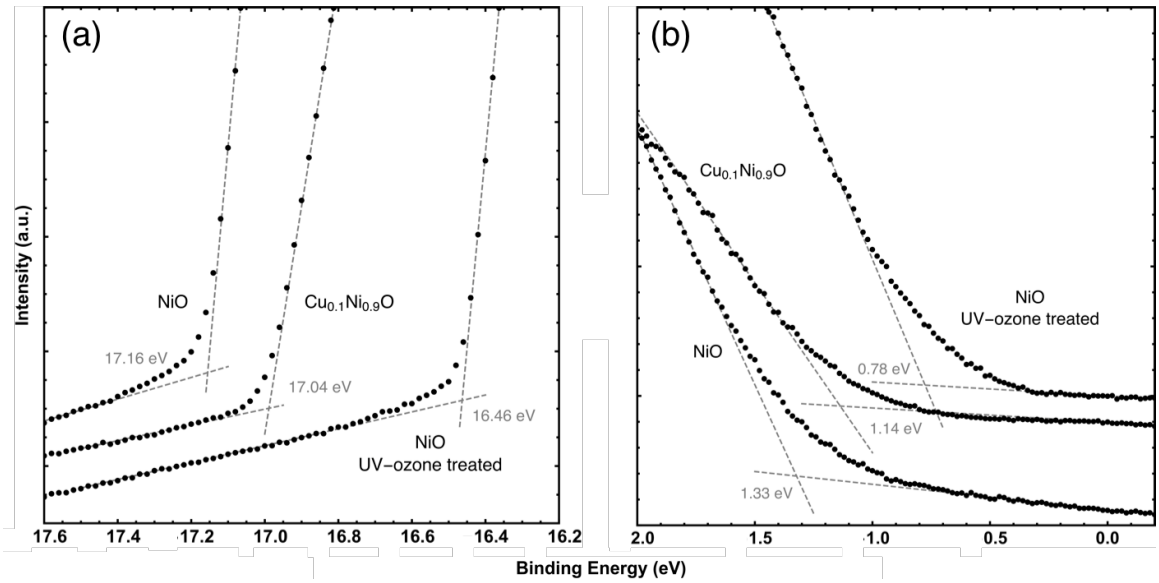

Figure 5.18: (a) UPS cutoff spectra and (b) UPS valence-band spectra for NiO, UV-ozone-treated NiO, and $Cu_{0.1}Ni_{0.9}O$, with linear fits shown. Binding energy has been zeroed to a point defined by an Au UPS Fermi-level scan (not shown). Spectra have been vertically shifted for visibility.

UPS data for NiO, $Cu_{0.1}Ni_{0.9}O$, and UV-ozone-treated NiO are shown in Figure 5.18. The binding-energy scale was zeroed using an Au film, as previously discussed. Linear fits to the background and the signal were used to define the valence-band edge and the cutoff energy, after which Equation 2.12 was used to calculate the ionization energies; the work function was then calculated using the positions of the intersection points in 5.18(b). Small valence band tails, extending approximately 0.5 eV from the band edge, are also visible.

The extracted work functions and ionization energies are shown in Figure 5.19. The NiO ionization energy $IE = 5.39$ eV ($E_F - E_V = 1.33$ eV) is within the normal range for NiO, as is the measured work function $\phi_s = 4.06$ eV. Upon copper-doping, IE remains nearly the same, but the work function increases to 4.23 eV, a significant shift, with $E_F - E_V = 1.14$ eV. Meanwhile, UV-ozone treatment causes a more significant downward valence-band shift, 0.14 eV, compared to pristine NiO, and the Fermi level shifts downwards by 0.7 eV. Fermi level movement towards the valence band,

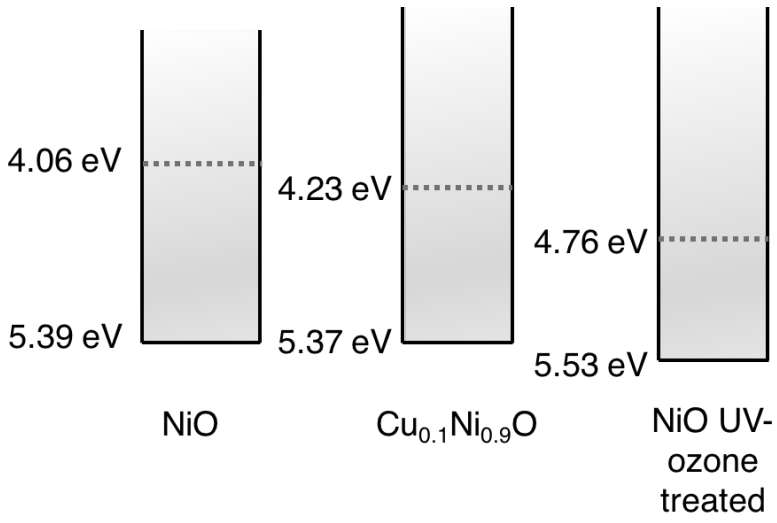

Figure 5.19: Valence band and Fermi level positions for NiO, UV-ozone-treated NiO, and $Cu_{0.1}Ni_{0.9}O$, as extracted from the fits in Figure 5.18.

as observed in both the copper-doped and UV-ozone-treated films, corresponds to increased hole density, an effect in agreement with the IV measurements previously presented. However, while the *direction* of Fermi-level movement matches what would be expected, and while these spectroscopic results are in agreement with what is reported for UPS measurements in the literature, quantitatively the measured $E_{\mathrm{F}} - E_{\mathrm{V}}$ is not in agreement with the activation-energy derived Fermi level locations in Section 5.10.

5.12 ALD-NiO/Silicon Test Devices

With a reasonable understanding of the NiO films established, NiO/Si test devices were fabricated to see how the material and its interface with silicon affected diode characteristics. Ag/n-Si, Ag/NiO/n-Si, and Au/NiO/n-Si devices were fabricated on 0.03 cm-thick float-zone n-Si with resistivity $\sim 2\Omega$-cm. For the Ag device, the ALD-NiO was ~ 5 nm thick and deposited at 175°C; the top contacts were uniform (~ 100 nm-thick) silver films deposited by thermal evaporation. For the Au device, the NiO was ~ 10 nm thick and deposited at 200 C; that device received an ultrathin blanket

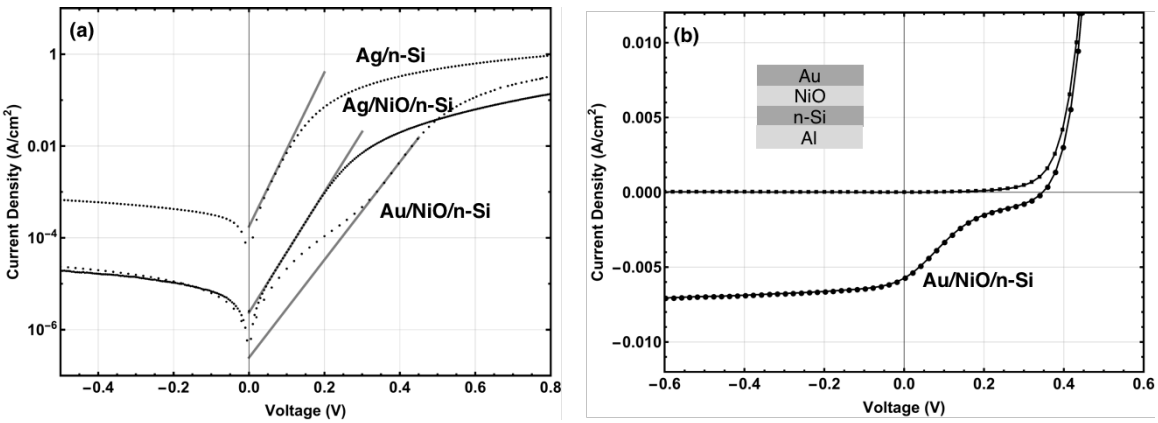

Figure 5.20: (a) Dark, room-temperature current-voltage curves for three n-Si-based diodes, and (b) dark and illuminated current-voltage curves for the Au/NiO/n-Si device, showing the barrier caused by the NiO/Si valence band offset.

Au deposition, to serve as a semi-transparent high-work-function anode, while thicker gold fingers deposited afterwards allowed for damageless probing and decreased the series resistance. All devices received aluminum/silver contacts deposited on the back. Dark, room-temperature JV curves, along with fittings to extract saturation currents, are shown in Figure 5.20(a). The forward-bias band structure and proposed dark-current mechanism is shown in Figure 5.21(a).

Metal-silicon junctions generally feature a large amount of Fermi level pinning, meaning that midgap interface states pin the interfacial Fermi level to a certain energy regardless of the metal's work function. This causes, for instance, aluminum/n-Si devices to be Schottky diodes rather than the Ohmic devices one would expect from their relative work functions. Schottky diodes with saturation current J_0 and barrier height ϕ_{B} follow the relationship

$$\phi_B = -kT\ln[J_0/(A^\star T^2)]. \tag{5.4}$$

where T is the temperature and $A^\star$ is the Richardson constant.

Polycrystalline silver is generally taken to have a work function of around 4.3 eV [112]; the Schottky barrier should be $\phi_{\mathrm{B}} = \phi_{\mathrm{m}} - EA \sim 0.2$ eV, where EA is the

silicon electron affinity (conduction band level) and ϕ_{m} is the metal work function. In reality, as is shown in Figure 5.20(a), that device is significantly diode-like, with an extrapolated barrier (using Equation 5.4) of 0.626 eV; the ideality factor is exactly 1.0, implying ideal device behavior. The addition of an ALD-NiO film changes that extrapolated barrier to 0.735 eV, a significant difference, with an ideality factor n of 1.282. Theory says that Fermi depinning should have actually *decreased* the Schottky barrier, but the barrier in fact increased. This fact leads us to the conclusion that the NiO is in fact behaving as an electron-blocking layer on n-Si, though the ideality factor $n > 1$ implies a significant level of interface recombination as well. The Au/NiO/n-Si device reaches an implied barrier of 0.792 eV, presumably as a result of both Fermi level depinning and NiO blocking. The ideality factor in this device was even higher, $n = 1.58$; as the charge-injection-current decreases, recombinative current is proportionally more important. These results echo those in [83], which found that sputtered NiO was effective at depinning the Fermi level in p-Si Schottky devices.

Illuminated (non-AM1.5G) measurements were also taken of the Au/NiO/n-Si device, as shown in Figure 5.20(b). The gold layer blocked much of the incoming light, so short-circuit currents were small, but really the aim of the experiment was to check for the presence of a parasitic barrier. Indeed, the pronounced s-shaped JV characteristic strongly suggests that a hole barrier exists at the interface. The UPS measurements previously presented in this thesis suggest that there is an approximately 0.2 eV barrier for holes at the NiO/Si interface. As was discussed in Chapter 2 and defined in the Blocker Parameter Requirements Table (Table 2.1), a 0.1 eV barrier does not significantly impair the transport of photogenerated minority carriers, but 0.2 eV *does* have an effect, causing an s-shaped illuminated IV curve. The barrier measured for NiO/n-Si is apparently enough to significantly impair transport, as shown conceptually in Figure 5.21(b), matching our expectations from the UPS band-offset measurements. While this barrier is somewhat discouraging in the

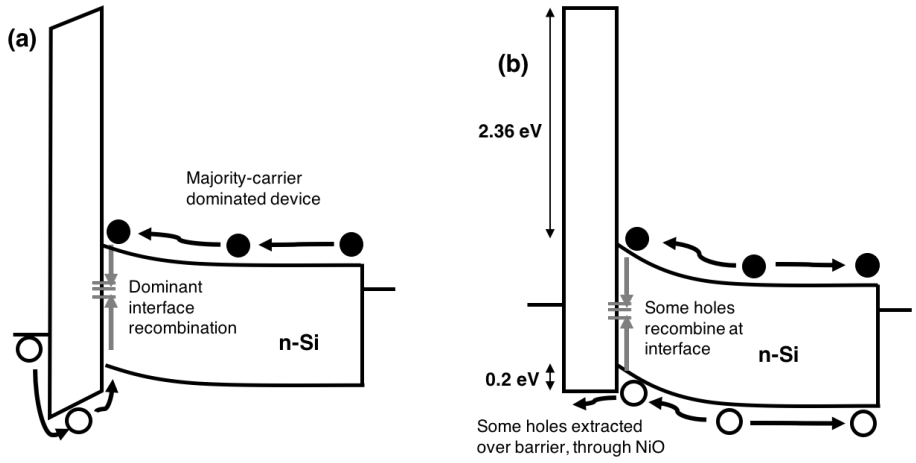

Figure 5.21: NiO/n-Si diode band diagrams, for (a) forward-bias dark current and (b) short-circuit illuminated conditions. The dark current in the silicon is majority-carrier dominated. Under illumination, the nonzero valence band offset presents a barrier to photogenerated holes.

context of PV applications, it is another suggestion that the interface is not fully dominated by recombination, since in that case holes at the surface would immediately recombine, leading to a Schottky device without a parasitic barrier to hole photocurrent.

5.13 Application of ALD-NiO in Perovskite Solar Cells

NiO finds important application in perovskite solar cells. To test ALD-NiO in this context, ITO/NiO/$CH_3NH_3PbI_3$/PCBM/BCP/Ag cells were fabricated and tested. In these devices [113],[114],[115],[116], the NiO is the hole-selective layer on the anode side; on the other (top) side of these inverted devices, the PCBM is the electron transport layer, the BCP is a buffer layer, and Ag is the cathode contact. The perovskite itself acts as the light absorber, and photogenerated carriers in that layer are able to diffuse up to a micron [117]. The structure is shown in Figure 5.22(a), with band diagram (energy values from [114] and [75]) in Figure 5.22(b). Although perovskite

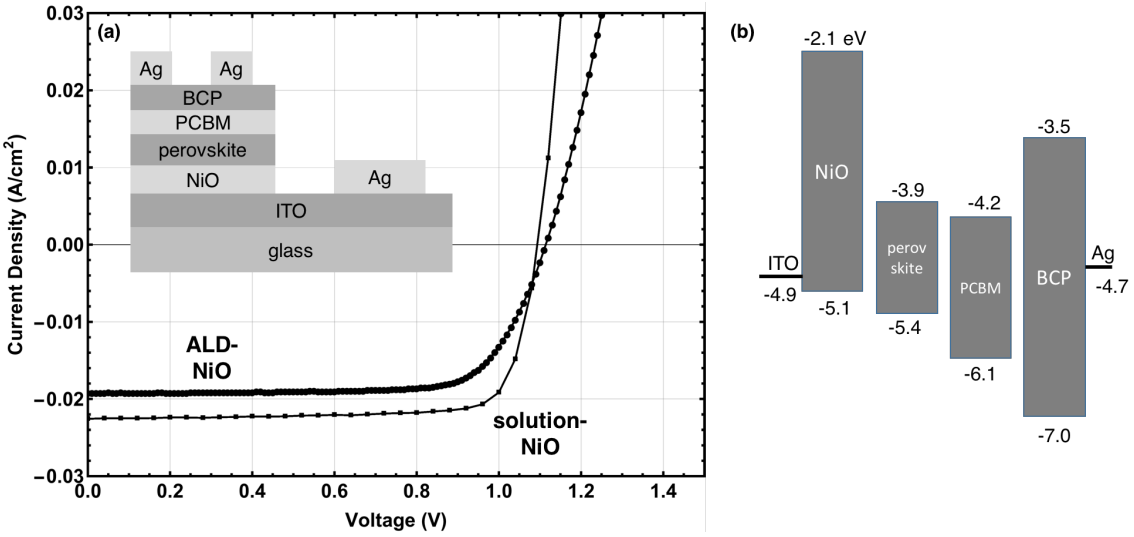

Figure 5.22: (a) Experimental JV curves for ITO/NiO/$CH_3NH_3PbI_3$/PCBM/BCP/Ag solar cells, using both ALD-NiO and solution-deposited NiO, and (b) band diagram for this device (energy values from [114] and [75]).

cells comprise an ongoing research area and present many unresolved questions, the role of the NiO in a perovskite cell is not particularly novel: it is a selective contact to which photogenerated free carriers diffuse, and into which holes are free to travel, with ideally a minimum of recombination at the NiO/perovskite interface.

Some devices used a 10 nm ALD-NiO layer deposited at 200°C with O_3; others used a solution-deposited NiO layer annealed at 400°C. Sample fabrication, besides the ALD-NiO growth, was conducted by Nakita Noel in the Rand Lab at Princeton. AM1.5G IV testing was done without exposure to atmosphere, as shown in Figure 5.22; separate ITO/NiO/perovskite samples also received SEM and XRD measurement.

The SEM and XRD measurements (not shown) showed that perovskite morphology (both grain size and crystallinity) did not change significantly with NiO deposition method; grain sizes were 500 nm to 1 μm and the crystal orientations matched.

In Chapter 2, it was discussed in detail how high interface recombination in solar cells reduces V_{OC}. That discussion focused on silicon-based cells, but the same general

	ALD-NiO	**Solution-NiO**
V_{OC} (V)	1.114	1.096
J_{SC} (mA/cm^2)	19.2	22.6
FF (%)	74.3	80.3
PCE (%)	15.9	19.9

Table 5.4: ITO/NiO/$CH_3NH_3PbI_3$/PCBM/BCP/Ag Solar Cell Properties

principle applies to perovskite cells. As shown in Table 5.4 along with other cell parameters, the ALD-NiO layer improved V_{OC} by 18 mV compared to the solution-NiO case. This is a significant improvement in the world of photovoltaics, where results are constrained by strict physical limits and every small increase is a hard-fought victory. While we cannot definitively point to a specific reason for the improvement, it seems likely that the ALD-NiO has a more favorable band structure for hole extraction from the perovskite, or alternatively that the ALD-NiO surface is able to form a less-defective interface with the perovskite. Either case would improve hole injection and reduce interface recombination, increasing V_{OC}.

Not all the cell parameters improved, though. Compared to the solution-NiO device, the ALD-NiO device exhibits a lower J_{SC} and worse fill factor. In cells of this design, J_{SC}, V_{OC}, and FF seem to scale positively with increasing NiO conductivity and work function, a change which is attributed to improved charge extraction [108],[113]; here, V_{OC} was observed to move against the other values. As previously discussed, ultrathin HTL resistivity becomes a serious problem at around $\rho = 10^6$ Ω-cm, a mark not hit even by our undoped ALD-NiO (cf. Section 5.10). However, our previous IV measurements were of *lateral* transport, whereas transport in the solar cell is vertical. It is possible that conduction in the ALD-NiO is anisotropic, possibly because of non-uniform defect densities over the depth of the film or because of an insulating layer formed by the ALD process at the ITO/NiO interface: in that case,

the effective HTL resistance in the perovskite cell could be much higher, explaining the reduced fill-factor.

A hypothetical "control" device without any HTL was not fabricated; such a device would be expected to perform much more poorly, with only the ITO anode providing carrier selectivity at that perovskite interface. Though the structures are not completely analogous, these devices could be compared to those in [118], which had the non-inverted structure FTO/TiO_2/perovskite/Au. It used TiO_2 as ETL rather than PCBM, and the cathode/anode materials are different, but otherwise that design is quite similar to a hypothetical ITO/perovskite/PCBM/BCP/Ag cell. [118] reports $J_{SC} = 7$ mA/cm^2, FF= 0.69, and PCE= 6.53% for that cell, significantly worse than the results for our ALD-NiO-containing device. Therefore, even if it did not quite match the solution-NiO, the ALD-NiO was indeed shown to act as an effective electron-blocker, giving cell figures of merit comparable to those found in the literature with little optimization and without UV-ozone treatment or copper doping. Implementation of those into the process would be a good next step for ALD-NiO as a core tool in perovskite cell fabrication.

5.14 Conclusion

This chapter presented a range of unified process-development data for and measurements of ALD-deposited NiO and, for the first time, ALD Cu-doped NiO films. Copper doping proved to increase film conductivity, an effect supported by the other data shown. However, neither the conductivity increase nor the valence-band/Fermi-level movement—all parameters that matter greatly in determining final PV device quality—show the sorts of effects one would expect from an efficient doping process: it appears that the Cu hole-doping is highly inefficient, given the significant atomic density it has in the films. However, even the undoped ALD-NiO served as an effec-

tive HTL in perovskite solar cells, exhibiting an 18 mV V_{OC} enhancement over control devices made with solution-NiO.

Many more experiments immediately present themselves as worthwhile and potentially educational. The ALD-NiO/perovskite solar cells shown here should be compared to ALD-NiO:Cu devices to look for fill-factor and V_{OC} enhancement. ALD-NiO:Cu/Si devices, both n-Si and p-Si should be fabricated, for comparison to the undoped devices, and anode contact metal should also be varied to test more quantitatively for Fermi-level depinning. More material-focused studies could be conducted, as well: AC-conductivity measurements, Hall measurements, and low-temperature IV measurements would likely give more information on the doped conduction mechanism, possibly resolving the mystery of the increased activation energies. The ALD process itself could perhaps be modified to encourage a higher efficiency of Cu^+ incorporation. Finally, the possible lateral vs. vertical conductivity disparity should be investigated; perhaps ALD process parameters could be changed to avoid the creation of highly-insulating layers of the film, decreasing R_S in completed solar cells.

Chapter 6

Cross-Sectoral Dynamics in Patenting

6.1 Introduction

The previous chapters of this thesis approached things from a technical perspective: experiments produced data that, in a small way, expanded human understanding of materials relevant to next-generation photovoltaics. But it should not have to be noted, in the year 2019, that scientific research is not socially neutral. A series of uneven filters determines who gets to conduct research; a complicated dance between funders, institutions, and practitioners determines which topics get researched and which scientific goals are priorities; and the dissemination and social effect of research findings depend on the desires of an entrenched ownership class. As discussed in the introductory chapter, the combined research excellence of thousands of climate scientists has been greatly counteracted by powerful fossil-fuel interests, and the funding of that research is itself often in the hands of politicians whose ideology forces them to reject climate science regardless of its objective persuasive power. But the situation is usually considerably less black-and-white. Practicing scientists look at their

research interests and abilities, weigh possible funders and hirers, and come to some decision about where or on what to work. This chapter is an attempt to analyze that research-decision process and its broader effects.

The first section looks at research collaboration between universities and private industry. Are desperate professors, perpetually searching for funding, reorienting their research trajectories to fit the needs of corporations? Or is corporate collaboration generally a non-disruptive phenomenon that helps academics research what they would want to research anyway? And are these collaborations scientifically worthwhile, or do they seek narrowly-focused advances with profit in mind? In fact, a patent citation/similarity analysis finds that these fears are misplaced: academic scientists who collaborate with industry do work that is more influential when they collaborate, and the analysis finds no indication that industrial partners drive drastic changes in research focus.

The second section interrogates the military-industrial-academic complex through the lens of its most high-achieving participants: those who have worked and patented in each of the three sectors. The stakes here are high: Diego Rivera's side-by-side portrayal, in the Detroit Industry Murals, of some scientists synthesizing poison gas while others develop medicines is not at all an over-dramatization. Military R&D, though, is not short of defenders who assert that military research is itself socially beneficial and does not represent a substitutive undertaking. This section uses a novel patent dataset to find that, indeed, these scientists' military work diffuses less than does their other work. However, a subsequent patent-similarity analysis finds that military work is not *uniquely* different in subject from non-patent work.

6.2 Academic-Industrial Patenting Collaboration

6.2.1 Introduction

Collaboration between the academic and industrial sectors has become a significant element of the American research environment. In 2017, private companies contributed $4.5 billion to university science and engineering R&D, or about 6% of total funding, a number which has remained approximately constant over the last decade [119]. When one includes complementary federal research funds, 20-25% of all university research funding can be traced to these collaborations [120], and such cross-sectoral work is subjectively viewed as ubiquitous by graduate students [120]. To understand the modern research landscape, then, one must have a good grasp of how these collaborations come about, their benefits and drawbacks for both their participants and society at large, and the dynamics of how they structure research choices.

University-industrial collaborations, or UICs, are often discussed in terms of their potential ability to negatively influence the research agendas of participating academics. According to this theory, scientists might alter their research agenda in order to access UICs, or, after participating in UICs, shift their future research directions with profit in mind. Indeed, there is some evidence that this process does occur, though its extent is difficult to measure. But assigning normative values to these possible effects is difficult. A scientist who shifts their research to more UIC-desirable work might be abandoning broadly useful work in favor of financial gain, true—but UIC-desirability might *itself* be a good signal of what work is socially useful. And a post-UIC shift in research focus could well be in line with researchers' "true" desires, desires now lucky enough to have a corporate benefactor. Therefore, this study will avoid calling UIC-related effects "good" or "bad:" rather, the goal is to attempt to measure those effects, leaving ethical analysis to possible industry-by-industry or case-by-case determination.

Regardless of broad normative value, it is important to determine how UICs can affect research agendas. Most critically, this information can inform individual decision-making in situations where that value *is* accessible. For instance, if university stakeholders know for a fact that work on a UIC project makes similar research lines more likely in the future, they might step in to stymie collaborations on projects they find objectionable, or they might specifically encourage UICs focused on "good" projects. In addition, at present, even if the general effect is unknown, surely many individual companies involved in UICs know what the research-shaping power of their collaborations has been; this project, then, can be seen as trying to extend that knowledge to the public at large.

Here, academic patenting is used as a lens via which to examine UICs and research direction. By tracking the patents of individual academic workers over time, it is possible to observe the rate of "research subject drift" among academic patenters. A statistical analysis of a newly-constructed dataset of high-patenting academic inventors, with special notice made of patents coming out of UICs, allows the testing of three hypotheses about UIC patent generation. First, do UICs significantly shift their academic participants' research topics, as reflected in the similarities between their UIC patents and their previous patents? Second, are UICs exceptionally powerful at "research-agenda-setting" for participating scientists, as reflected in how similar their future work is to those patents? And, finally, do UICs tend to result in patents that diffuse more to the rest of society, implying a higher value? The analysis finds that UIC patents have no significant effect on same-inventor research choices, but that UIC patenting is indeed correlated with higher knowledge diffusion.

6.2.2 Previous Work

While the patent-based methodology used here has not previously been applied to the aforementioned questions, the literature does contain articles seeking to answer them

in other ways. Mirroring the range of structural forms taken by UICs, methodological approaches in the academic-industrial research literature have varied from large-scale econometric analyses [121],[122],[123] to participant surveys and case studies [124],[125],[120],[126],[127]. Results point to UICs having a variety of positive effects on participating academics (the effects that collaborations have on *firms* are beyond the scope of this thesis, but those effects have also been found to be positive, with the primary benefit being, unsurprisingly, direct increased access to new research [125],[128]. Collaborations also seem to give rise to socially-beneficial local spillovers [129]). Researchers, obviously, gain funding that they can then use for equipment, more students/staff, etc: this is a primary motivation [125]. Researchers with industrial links also appear to publish more often than their peers [123], and, though the link to academic *citations* is nontrivial and dependent on the capacities of the participating university, there, too, the linear effect of industrial collaboration is positive [130]. A study of Italian researchers found that, while UICs are quantitatively less likely to occur than intra-sector collaborations and produce fewer patents than other collaboration types, the patents they *do* produce tend to garner more citations [131]. Looking to researchers' future work, collaboration on applied projects tends to expose academics to new ideas and possible research paths [132] and, for participating students, offers opportunities for careers in the private sector.

Along with these "output-focused" effects, UICs have the potential to qualitatively change academic research, either *how* or *on what* it is conducted. This is often discussed in terms of intangible negative effects, although such changes could also be positive. A persistent worry in the literature is that industrial collaborations will negatively shift the behavior of collaborating professors. Specifically, they may choose to work on more applied/less basic research, or on research that is specifically profitable as opposed to broadly socially valuable; devote less time to departmental work or teaching; disregard ethical constrains; leave their academic posts for the

private sector; delay research publication or avoid the broad dissemination of their findings; lose the trust of the public; or push for university-wide policies that enable their work [133],[134]. That many of these possible negative effects could just as easily be motivated by non-UIC research, or that some of those outcomes might not be negative at all, is not always satisfactorily addressed. Still, the direct incorporation of private industry into the academic research process *is* a potentially harmful thing, and its ramifications ought to be fully explored.

There have been comparatively few empirical studies looking at this problem, with sometimes contradictory conclusions reached. A study based on the publication data of UK engineers found that industrial collaboration skews research towards the applied and away from the basic, without making any normative claims [123]. A survey of graduate students [120] found no negative effects in terms of research freedom (research freedom was affected slightly by several other factors, including the basicness of the research); on the other hand, two studies by Blumenthal et al. [135],[136] found that industrial collaboration correlated with publishing delays and caused researchers to "[take] commercial considerations into account when choosing research topics." Similarly, a survey of graduate students and postdocs found that privately-funded researchers sometimes delay publication or face restrictions in their research [137]. So the risks of industrial collaboration cannot be discounted: they can certainly manifest, even if their effects are slight or their occurrence uncommon. And, insofar as UICs remain net positives for their academic participants or for society, that may be due in part to vigilance on the part of academia, which is cognizant of the possible problems and has worked to head them off [134].

The literature, then, gives us many reasons to believe that industrial collaboration can be pivotal for researchers; that it can, at least subtly, change scientists' research behavior and priorities; that it can boost per-patent research impact; and that, in general, positive effects of it are considerably easier to find than negative effects.

6.2.3 Background

American universities have a long tradition of collaboration with industry, so it is wrong to think of UICs as a "trend" that can be simply traced in one direction. Early collaborations were often with the agricultural industry, though even in the 19th century non-agricultural collaborational research was widespread [138]. As the country industrialized and industrial research become a more pressing concern, collaboration based on personnel exchange "through faculty consulting, faculty rotations to and from industry, and placement of graduates" [139] became the predominant form. After the Second World War, as government sponsorship of research soared, UICs became comparatively less important, but over the past few decades they have regained their previous proportional importance [140].

Federal and state governments have explicitly worked to encourage company-university research collaborations. In some cases, this support has taken the form of research centers operating with significant government oversight. Notably, the NSF's Industry-University Cooperative Research Centers (IUCRC) program, started in the 1970s, provides small amounts of funding to support the administrative and management costs of university-company collaborative research centers. The NSF's Engineering Research Centers program tries to build deeper and more lasting collaborations: it provides millions of dollars in direct research funding to its centers.

The broad area of industrial-academic collaboration covers a diverse array of modern organizational structures and outcomes. Perkmann and Walsh [141] categorized arrangements according to the level of inter-sector linkage: by this metric, collaborations range from formal centers of collaborative R&D ("research partnerships") and the direct commissioning of university research ("research services") to shallower interactions like IP transfer and knowledge diffusion through scientific publications. These arrangements can alternatively be mapped onto an axis of "finalization;" low-finalization partnerships like sponsored research centers lead to products that are

further from commercialization, while high-finalization collaborations, such as contract research and consulting, are more important for short-term gain. Schartinger et al. [124] categorized interactions in terms of their level of formalization, their level of transfer of tacit knowledge, and their amount of face-to-face contact, with a focus on differentiation between nine different "knowledge channels." Taking another tack, Powell et al. [121] viewed interactions in terms of their positions within learning networks, with a strong emphasis on time-evolution.

It makes logical sense that UICs could shape research choices. With research funding on the line, PIs may well undertake significant shifts of their research areas, particularly when their industrial partners are likely to fund not only personnel but also new equipment. UICs, then, may represent special opportunities to move a lab's work into a new area. If so, this might be represented not only in comparisons between UIC-work and previous work but also in comparison between UIC-work and *future* work. Having received an infusion of new funding and incentives, academic researchers may be slower to transition their future research choices— they may choose to stay longer-than-usual on the research topic defined by their UIC. These previous-patent and future-patent statements represent the first two hypotheses of this section.

The third hypothesis of this section is that UIC patents diffuse more readily than non-UIC patents, from which we can infer that they are more valuable. Nearly all patents, of course, have *some* value— otherwise, they would not be worth the expense of obtaining. This is likely to be especially true for patents originating in universities, for which the decision to pursue patenting is based not in individual whims but in institutional processes and decision-making. There is reason, though, to think that UIC patents might be especially valuable. These patents arise out of specific, highly-coordinated collaborations that are generally aimed at specific problems. They are probably less likely to be the "incidental" patents that might arise semi-randomly out of any sufficiently applied research; rather, they come about because a private

firm has decided that a specific problem is worth investing money into solving. If this hypothesis were proven correct, then it implies that industry's allocation of research funds into UICs is, at least on a broad, first-order basis, socially efficient, leading to outcomes more beneficial (on a per-patent basis) than would otherwise be obtained absent the collaboration. Here, we measure patent diffusion propensity/value using patent citations, a decision explained in Subsection 6.2.4.

6.2.4 Academic Patenting

Universities as institutions and university workers as individuals have increasingly sought to patent their research products, with university-associated patent applications rising sharply over the last thirty years, as shown in Figure 6.1. This rise, by all accounts, had several separate causes: the rise of biomedical technology as a major part of universities' research portfolios deserves special mention, as does the Bayh-Dole Act of 1980, which gave universities the rights to IP developed using federal funds [138],[140]. Other contributing causes were court cases like *Diamond v. Chakrabarty* (which extended patenting to manufactured organisms) and the general desire of universities to increase revenue [142].

After the development of a possibly patentable product, a school's Technology Transfer Office works with the associated students and PIs to develop a patent application. If the application is granted, the university gains the rights to the invention, while the inventors are allocated some share of royalties arising from the patent (this is in contrast to the situation at most private companies, where the inventors of patents settle instead for one-time bonuses, or for nothing). Inventors who wish to found companies based on their research must negotiate for those rights with their universities. This gets to an important feature of patents: the distinction between *applicants*— those entities, generally institutions, that are given ownership of the patent— and *inventors*, the individuals who actually developed the invention. A UIC

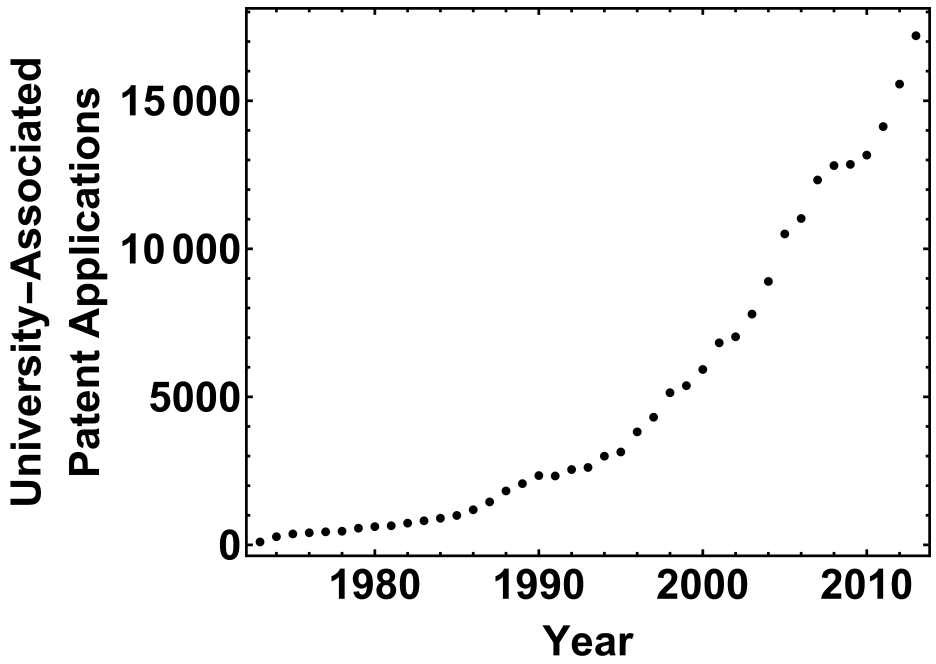

Figure 6.1: USPTO patent applications with university applicants from 1973 to 2013. Data extracted from PATSTAT.

patent, as defined here, is connected to at least one inventor and two applicants (one university and one company).

Standard procedure is complicated in the case of a UIC, with the particular arrangement depending on the collaboration agreement. In the standard IUCRC agreement, for instance, universities own patents that originate in the joint research, but participating companies are given nonexclusive royalty-free licenses. Alternatively, industrial partners may share IP rights directly with universities. It should be emphasized here that patents, per se, are not themselves the main motivators for companies or academic researchers to enter into UICs; the top motivators are, according to a survey of faculty scientists and engineers engaged in industry-sponsored research at major research universities, to secure research funding and to test applications of theory [125]. At the same time, patenting is not a marginal factor, and patents are by no means rare outcomes: in the same study, about 50% of surveyed professors said

that industrial collaboration influenced their patenting, and on average they gained about 1 patent from the experience [125].

This study uses academic patent applications, rather than academic publications, as the source of its data, a decision that requires some justification. After all, there are fewer academic patents than academic publications, and a given scientific advance is more likely to be reflected in a research paper than in a patent: the restricted dataset necessarily means a lower n, and possibly a skewed dataset as well, in that only some kinds of advances are patentable. But patent applications offer some advantages, as well. The primary one is that, because patent applicants must list what sector ("UNIVERSITY," "COMPANY," etc) they belong to, UICs can be identified by looking for applications that share those sectors in their applicants. There is no analogue of this in academic papers— the closest thing might be an analysis of the acknowledgements/funding section of papers, but large-scale use of that would be prohibitively difficult. One could also look for industrial co-authors, but that would redefine UICs into the particular and unusual form of actual personnel-level collaboration. Besides that main advantage, using patent applications also allows us to incorporate several ancillary public-domain datasets to aid in analysis.

Patent Citations as a Measure of Value

The use of patent citations to measure patent value is a widespread and well-justified, though not totally uncontroversial, method. Patents must cite relevant prior patents, so there is reason to think that the count of forward citations should correlate with importance/value. More precisely, forward citations are directly linked to the diffusion of the ideas embodied in a given patent, and that diffusion should be linked, in the end, to the importance of the patent. This idea has been quantitatively confirmed: expert evaluation of patent importance has been shown to correlate with citation count [143], and citation count has even been shown to correlate with the level of

technical improvement expected to be observed within particular technological domains [144]. Tests done as part of this study also showed a very strong correlation between (domestic) citation count and whether a given patent is "triadic"– that is, whether it has been filed in all three of the USA, Europe, and Japan, a traditional test of patent importance. The citations-diffusion-value idea is reviewed and interrogated in a review by Jaffe and Rassenfosse [145]; that review broadly supports the link between forward citations and patent value, though points out that more work is required to fully understand the mechanisms at play.

6.2.5 Methodology

Patent Similarity Metric

Previous studies on academic patenters and research-topic-selection have either surveyed researchers and used their subjective self-evaluation or quantitatively graded researchers on single scales of basic/applied or pure/commercial research [146],[147],[132],[123]. Here, the aim is to analyze UIC effects in terms of specific research topics, whether or not those topics are applied or basic. Essentially, the problem demands a metric to determine how similar, subject-wise, two patent applications are to each other.

One approach would be to use applications' International Patent Classification (IPC) codes. IPC codes are a nested, hierarchical classification scheme that assigns applications to one or more of about 70,000 specific subject areas, and they can be quite specific. For instance, code **H01L 31/0296** refers hierarchically to an electrical device/that is a basic electrical element/that is a semiconductor device/that is sensitive to light/that is characterized by a material/that is an inorganic material/that is a compound similar to CdS, ZnS, HgCdTe, etc. Specific levels of the hierarchy can be compared to each other, so that a patent for such a solar cell would be found to be totally dissimilar to a medical patent, barely similar to an incandescent lightbulb

patent, and quite similar to an organic solar cell patent. As mentioned before, a patent can be assigned multiple IPC codes to account for specific notable aspects or novel features. Unfortunately, in practice, these assignments are subjective and arbitrary; for a more consistent metric, one can also incorporate the IPC codes of all patents cited by the patent in question, as implemented by [148].

In that vein, one could avoid IPC codes altogether and simply measure patent similarity via citation networks. Applications with overlapping citation webs would be counted as more similar. One could also incorporate co-authorship, co-assignment, etc, into the calculation. Like the previous method, this would be a purely metadata-based approach that would ignore the actual content of the patent application.

An alternative method is to use the application text itself, identifying keywords in the two patent applications and using something like a Jaccard index to measure similarity, and for our purposes this method presents an important advantage. If one wishes to incorporate patent value as measured by patent citation counts, then any similarity metric that incorporates citations will necessarily entangle similarity with patent value in ways that are undesirable. For example, say one uses the IPC method with weighted citing-patent IPCs, and one compares patents A (highly-cited) and B (slightly-cited) each to patent C. C is more likely to cite A than it is to cite B, and, if it does, A's citing-IPC overlap with C will artificially inflate the similarity score; in other words, this method would lead to high covariance between citation-count and similarities. A naive analysis based on this method could well find nonexistent links between patent similarity and patent value. Using text-matching neatly separates the data-driven similarity scores from all other metadata-driven information.

For this reason, the method chosen was the text-matching method, incorporating the patent-keyword dataset constructed by Arts et al. [149]. There, they "cleaned" patent titles and abstracts, removing stop words and unique words, to assign each patent a keyword list that is, on average, 37 words long. This dataset can be found

at https://dataverse.harvard.edu/dataverse/patenttext. The similarity score between two patents i and j is then calculated using

$$sim = \frac{|\text{wordlist}_i \cap \text{wordlist}_j|}{|\text{wordlist}_i| + |\text{wordlist}_j| - |\text{wordlist}_i \cap \text{wordlist}_j|}. \tag{6.1}$$

Similarities ranges from 0 to 1; 1 corresponds to an exact match, while 0 means the keyword lists have no overlap.

Dataset Construction

The main dataset was constructed using PATSTAT 2016 Autumn Edition. PATSTAT is a purchasable database put out twice-yearly by the European Patent Office. It contains comprehensive patent information along with industry classification, patent family classification, concordance tables, etc, for patents from every significant patent office worldwide. This database was installed onto a MySQL server hosted by Princeton's computing services.

The first step was to construct a table of standard (ie, non-design, etc) USPTO patent applications from 1973 to 2013. Each application in this table was linked directly to its inventors and applicants, each of which was assigned a sector as described above (some of the code used to do this initial extraction was a heavily-modified version of the program published along with [150]). The next step was to construct a usable list of university-invented patents, along with with all accompanying metadata. Finally, the data was organized on the level of individual inventors: the basic outlines of the full process are shown in Figure 6.2. The consistent identification of individuals across many patents is not a trivial problem: the presence of various name formats and shared names make simple text-matching a poor strategy. Here, the tracking was done using the geolocated name disambiguation data compiled by Morrison et al. [151].

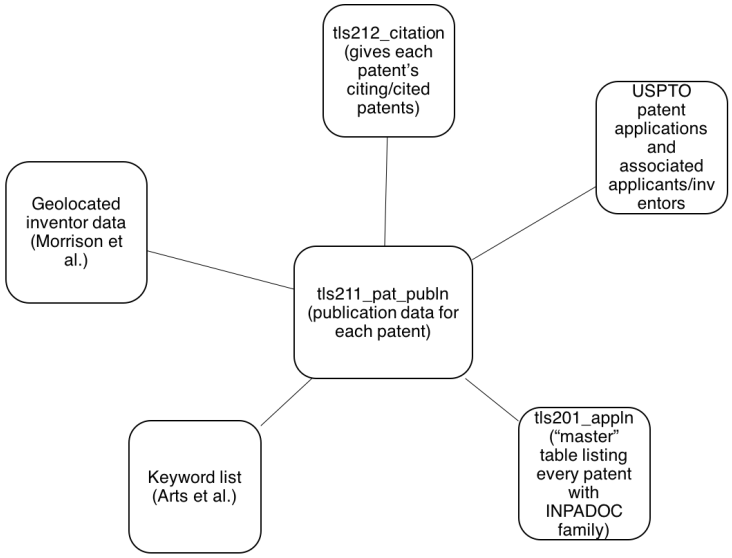

Figure 6.2: Relational chart showing primary joins and PATSTAT/other tables used in creation of university-inventors list

Of special note is the fact that all final data was compiled on the basis of INPADOC patent families, as opposed to on the individual-patent level. Patents belong to the same INPADOC family if they share a common priority patent application, and each patent belongs to exactly one family, such that the presence of continuations, divisions, international extensions, etc, do not result in "double-counting." Citations, too were done on a family basis, so that, for instance, two patents from the same family citing multiple patents in another single family would only count as one total family-family citation. Note that self-citations (outside of intra-family citations) were not filtered out. Undesired effects of self-citations cannot be ruled out, but, as patents are legal documents in which citations have a specific, narrowly-defined purpose, it makes sense to treat self-citations as legitimate.

The end result of this was a university inventors dataset. This comprised 678 university-based inventors, with per-inventor patent counts ranging from 7 to 128, with a mean of 15 and a median of 13. The study tracks research focus during the length of careers, so inventors with fewer than 7 patents to their name were excluded;

otherwise, possible early/late career artifacts would be more likely to corrupt data interpretation (the maximum of 128 patents was the natural record, not an imposed limit). This change does mean, of course, that the study is not really of UIC-patents in general but rather of the UIC-patents of high-patenting-propensity academic inventors; still, for the sake of the analysis, it was determined that the filtering was worthwhile.

To this dataset was then added information about dual-sector university-industry patents. 197 of the 678 aforementioned academic inventors also worked on at least one joint university-industry patent (identified, to be clear, by sectoral metadata). Several strict conditions were then imposed. To ensure that the analysis captured only primarily-academic patenters who sometimes collaborated with industry, researchers for whom UIC patents were more than 30% of their total patents were excluded; academic patenters with a higher UIC% than that probably have a more complicated relationship with industry than that being probed here. Imposing these conditions reduces our UIC-specific sample size to 161 academic researchers; 481 academic-only patenters remained in the dataset to help model general intra-career patent similarity dynamics.

For the university-only dataset, all same-inventor pairwise patent similarities were calculated. For the combined dataset, the following variables were defined for all applicable patents:

- UIC (binary): is this patent the result of a university-industry collaboration?
- $year$ (categorical): year of application filing
- $citations$: family-family citations of this patent
- $person$ (categorical): unique identifier of the inventor of the patent
- $prev\_sim$: the Jaccard similarity of the patent's keywords to the combined keyword list of all that inventor's previous patents

- *post_sim*: the Jaccard similarity of the patent's keywords to the combined keyword list of all that inventor's following patents
- *t_prev*: time, in years, since the inventor's first patent
- *pat_prev*: count of that inventor's previous patents
- *t_post*: time, in years, from this patent to the inventor's final patent
- *pat_post*: count of that inventor's subsequent patents

Some general information about the inventor dataset is shown in Figures 6.3 and 6.4. The peak in patents around the year 2000 is due to the convolution of an increase in patents over time and the imposed minimum patent count for each inventor. A significant number of late-2000s patents were invented by people who have not, by now, accumulated seven patents, so those patents are excluded from the analysis, causing the later-year falloff. The citations follow a highly skewed distribution; a few patents have more than 1000 citations, but the vast majority have between 0 and 100.

Data Analysis

The pairwise similarity data, while not specifically related to our UIC hypothesis, is useful for understanding the general question of patent similarity trajectories over time. Over 100,000 university-inventor intra-inventor pairwise similarities are plotted in Figure 6.5. The longer the time difference between patents, the lower the expected similarity is. This represents the natural research drift of inventors: having patented A, they, on average, move on to patent B, a followup of A, followed by C, a followup of B, etc. Eventually, many years later, their first and last patents have low expected similarity values. A good area for future study would be the phenomenology of this research-subject drift, comparing its rate across time periods, industries, sectors, etc. Here, though the focus is on narrower questions.

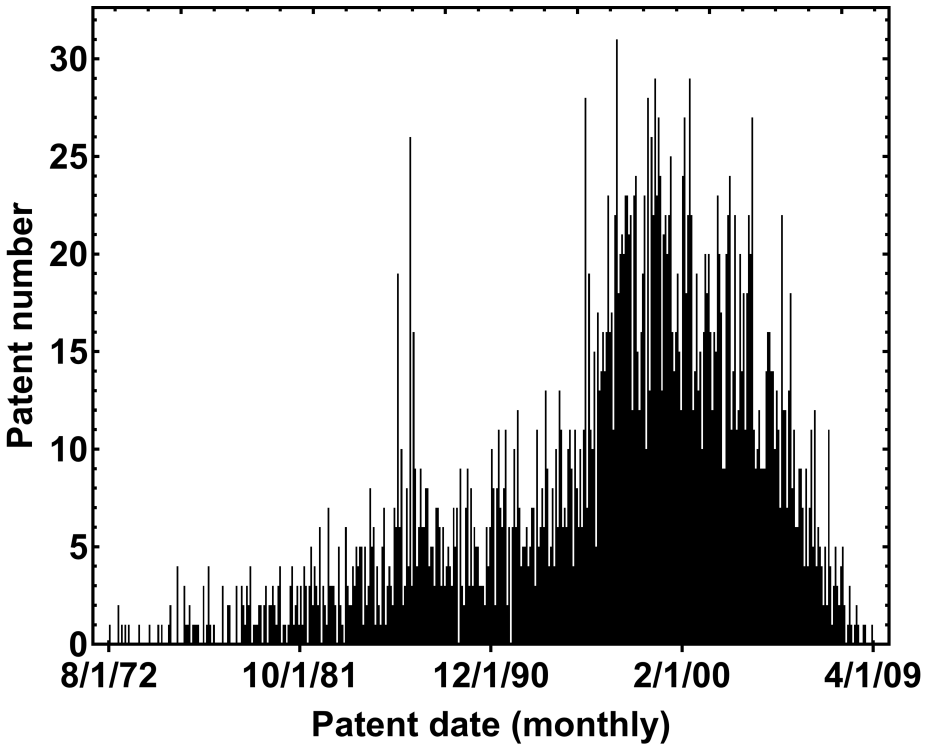

Figure 6.3: Time-wise histogram (monthly binnings) of patents in the UIC/academic-inventor dataset.

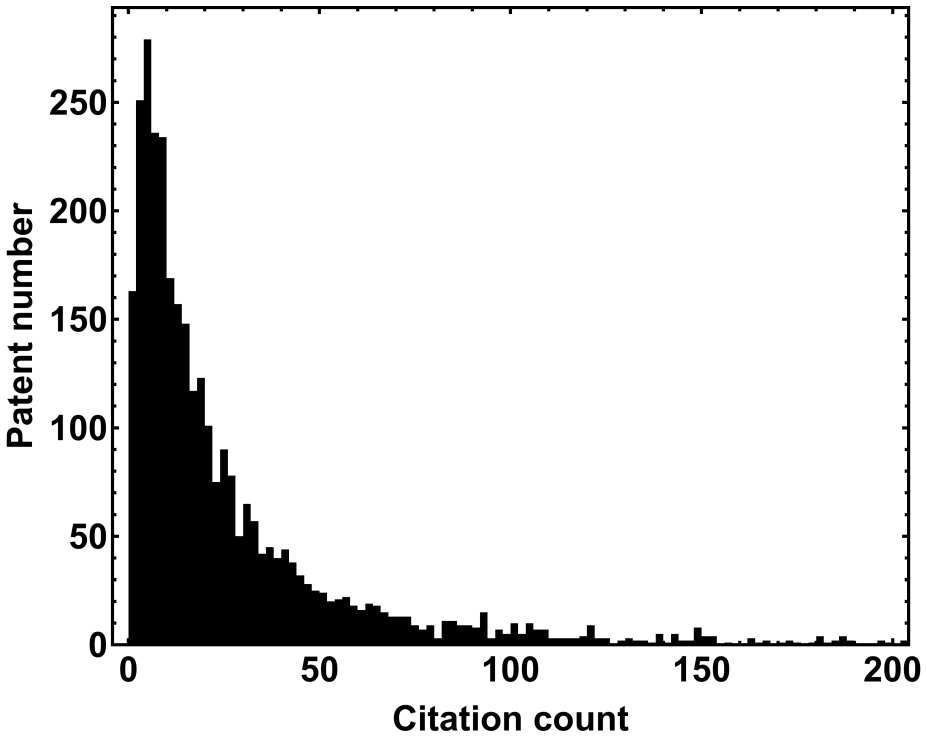

Figure 6.4: Citation-wise histogram of patents in the UIC/academic-inventor dataset.

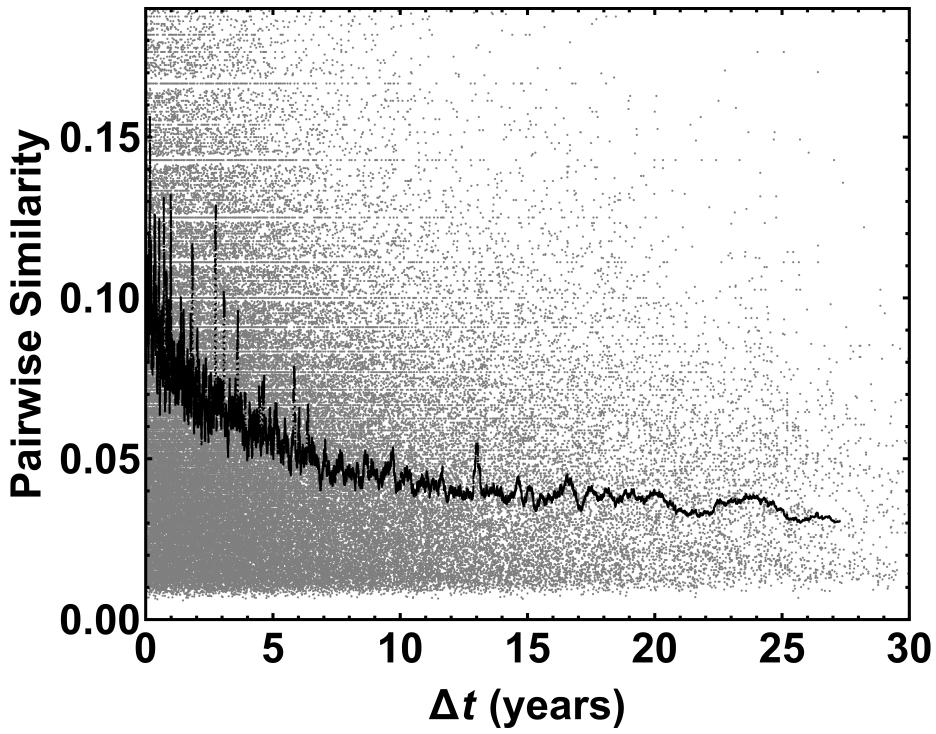

Figure 6.5: Approximately 10^5 same-inventor Jaccard patent similarities, plotted against the time difference between patents. Also shown is the rolling one-year average of all similarities.

The UIC-related hypotheses can be formulated in this language of patent similarities. If UIC-patents are more likely to be on novel research topics (*for their inventors*), then the similarities of those patents with the inventors' previous bodies of patent work should be lower. Similarly, if UIC-patents have a unique ability to shape future research paths, meaning that future patents are more likely to be similar to those patents, then the similarities of UIC patents with the inventors' future bodies of patent work should be higher. This full analysis compares single patents to collections of patents, as opposed to the pairwise similarities plotted in Figure 6.5, but the underlying concept is the same.

For Hypotheses 1 and 2, each patent was similarity-compared to all patents before and after it. For example, to test Hypothesis 1 for a particular inventor, their second patent's wordlist was compared to their first patent's, their third patent's was com-

pared to the combined list from patents one and two; their fourth's to the combined first, second, and third, and so on.

There were three hypotheses to be tested.

Hypothesis 1. *UIC arrangements cause researchers to significantly alter their research choices and represent larger-than-expected departures from their previous work.*

Similarities, by definition, lie between 0 and 1 (inclusive); in practice, they are heavily skewed towards lower values. A logit model was used, with a small epsilon applied to $sim = 0$ values. The similarity here, then, is thought of as a proportion of matching between the two keyword sets, a not-unreasonable interpretation.

To test this first hypothesis, the following logit model was estimated:

$$\text{logit}(prev\_sim) = \beta_1(UIC) + \beta_2(t\_prev) + \beta_3(t\_prev)^2 + \beta_4(pat\_prev) + \beta_0. \quad (6.2)$$

As inventors' bodies of patents grow, one expects to see the similarity of a given patent with all previous work decrease; thus time and patent-count variables are included.

Hypothesis 2. *UICs are especially effective at defining inventors' future research plans; future patents are more likely to be similar to a given patent if it is a UIC patent.*

To test this, the following model was estimated for every applicable patent in the dataset:

$$\text{logit}(post\_sim) = \beta_1(UIC) + \beta_2(t\_post) + \beta_3(t\_post)^2 + \beta_4(pat\_post) + \beta_0. \quad (6.3)$$

Here, post-patent time/count variables were included, to account, again, for natural research-subject divergence over time.

Hypothesis 3. *UIC patents tend to diffuse more than non-UIC patents published by the same academic inventor, as measured in patent citations.*

This was tested via a negative binomial regression of patent citation count, using the model below. Citation data is count data bounded at zero; this, plus the fact that it is significantly overdispersed, point to the negative binomial model being a correct one. Some researchers [152] have further applied a zero-inflation step; here, it was decided that there is no special process sorting citation counts into zero and nonzero values, so the pure negative binomial model was used. The fixed-effects (*year-person*) negative-binomial panel model, with dependent variable *citations*, was estimated for every applicable patent (having year i and inventor j) in the dataset, with coefficients and independent variables

$$\beta_1(UIC) + \beta_2(t\_prev) + \beta_3(t\_prev)^2 + \beta_4(pat\_prev) + \beta_{5,i}(year) + \beta_{6,j}(person) + \beta_0. \quad (6.4)$$

Year variables were included to account for the combined influence of there being more patents over time and there being less time to accumulate citations for newer patents. Person variables were included to account for person-specific differences in expected citation counts. It is possible, for instance, that the specific scientific fields featuring more UICs also draw more citations for each patent, or that ambition drives the same academic scientists to both participate in UICs and to do work that gets more citations: the per-person fixed-effects approach accounts for both of these possibilities, assuming that scientific field remains broadly constant during each career.

6.2.6 Results

For the first two hypotheses, the null could not be rejected. As shown in Tables 6.1 and 6.2, the expected time-dependent research-subject divergence/drift was observed; however, in neither case was it found to be significant whether or not a given patent was the result of a UIC.

Table 6.1: Hypothesis 1 (UIC Patents' Similarity to Previous Patents)

	Dependent variable:
	Similarity to prev. patents
Is a UIC patent	-0.062
	(0.284)
Previous career time	-0.065^{***}
	(0.022)
(Previous career time)^2	0.002^{*}
	(0.001)
No. previous patents	-0.016^{***}
	(0.006)
cons.	-2.095^{***}
	(0.096)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 6.2: Hypothesis 2 (UIC Patents as Trend-Setters)

	Dependent variable:
	Similarity to subsequent patents
Is a UIC patent	-0.005
	(0.204)
Subsequent career time	-0.0004^{***}
	(0.0001)
(Subsequent career time)^2	0.00000^{***}
	(0.000)
No. subsequent patents	-0.040^{***}
	(0.007)
cons.	-1.060^{***}
	(0.066)
Note:	*p<0.1; **p<0.05; ***p<0.01

For the third hypothesis, as shown in Table 6.3, the null was weakly rejected and it was found that UIC patents are more likely to have a higher citation count. This result was significant only with $p = 0.019$, and the residuals QQ-plot was observed to be slightly skewed, with a heavy tail at the high-citation end; however, the UIC coefficient remained positive and significant even after removal of the top-end fat tail, so there is some reason to be confident in the result. It is not definitive, but it is suggestive.

The analysis, incidentally, finds a drop-off in expected citations over the course of a career. Previous studies focusing on academic papers (rather than patents) have observed number-of-paper curves that rise and then fall [153],[154], but [154] also observed that citations per paper tended to immediately decrease and then increase later in careers. This study tracks most of its inventors for less than twenty years, so it is not surprising that the model would not reflect changes that happen thirty or more years into careers.

Table 6.3: Hypothesis 3 (Comparative Diffusion of UIC Patents)

	Dependent variable:
	Citations
Is a UIC patent	0.152^{**}
	(0.065)
Previous career time	-0.124^{***}
	(0.037)
(Previous career time)^2	0.001^{***}
	(0.0002)
No. previous patents	-0.002
	(0.002)
cons.	1.689
	(1.385)
Note:	*p<0.1; **p<0.05; ***p<0.01

6.2.7 Discussion

Neither hypothesis about UICs and research-subject choices was proven correct. While one cannot conclude from this that UICs have zero special significance with regard to research-subject trajectories, given the result it is worth considering what such a conclusion would mean. If UIC patents do not represent significant shifts in research topic, it would perhaps mean that companies are efficient at identifying researchers whose work is already "on-topic:" instead of significantly diverting scientists' work for their own purposes, companies instead collaborate with researchers who are already close to the problem at hand. As for the null result on the second hypothesis, it would mean that UIC-related funding is basically no different in its effects than other funding sources; this is in line with the result in [120], which found that the important differences in research approach are between funded- and non-funded projects, not between funded projects of different sources. In sum, these results should give decision-makers and stakeholders no reason to approach UIC projects differently than non-UIC projects.

The analysis does indicate that being UIC-connected is correlated with higher patent importance. Perhaps private companies are only interested in funding high-value projects, or perhaps academic scientists are better able to leverage UIC funding than other funding. Given the previous results shown here, that UIC patents seem to be "natural" outgrowths of inventors' previous work, it appears that these high-value UIC patents do *not* result from significant intra-inventor changes in research focus. Rather, industry may simply be better at predicting which particular next research step will be the most important. Academic researchers, then, should perhaps give extra consideration to research collaborations proposed by industry: companies may know what academics do not.

6.2.8 Approach Robustness

As mentioned in the previous section, of the three hypotheses, only one was supported, and even for that one the significance was not overwhelming. Several methodological barriers might have inhibited the effectiveness of the intra-inventor comparison approach. First, the testing of Hypotheses 1 and 2 incorporated text-comparison between single patents and sets of patents. This introduces an asymmetry into the career-long comparisons that is only partially accounted for by including the previous/subsequent patent counts into the regression: later patents are compared to a larger wordlist, artificially pushing similarity scores down towards long-term average values and suppressing the effects of other factors, including UIC incidence. An ideal approach would have used only single-patent to single-patent similarity scores, as in Section 6.3, but it is hard to square that desire with the need to look at similarity trajectories.

The previous/subsequent method, and the concept of "research topic trajectories" in general, may also be the wrong approach. Rather than patent subjects exhibiting gradual divergence over time, patents may be better explained by a clustering model, with each discrete research topic producing one or several closely-related patents. The divergence rate within and between these clusters could be significantly different than the average rate observed for all patent pairs. This would not be noticed by the models used here, which only go to second order in time, but it would, if true, help explain why no significant UIC effects on research topic were observed.

There were two cutoffs used in data collection. First, a minimum number of patents (seven) were required for an inventor to be included in the dataset. Second, because we desired to examine primarily-academic patenters who occasionally collaborate with industry (rather than non-academic workers doing research as part of some long-term joint arrangement), collaborational inventors were required to have at least 70% of their patents assigned purely to universities, with the rest (up to 30%)

University-only% Cutoff	# UIC inventors	Coeff. on "Is a UIC patent"	p
0	167	0.077	0.18
30	158	0.078	0.17
40	155	0.11	0.058
50	154	0.12	0.046
70	135	0.15	0.019
80	121	0.23	0.001

Table 6.4: Effect (using the Hypothesis 3 regression) of varying the cutoff of how many of an inventor's patents must be university-only (as opposed to university-company). As the cutoff is made more stringent, the effect of collaboration with industry increases.

composed of university-company patents. This latter cutoff is somewhat arbitrary, so the robustness of the non-null result of Hypothesis 3 was tested by varying that cutoff. Results of this exercise are shown in Table 6.4 As the cutoff% is lowered, the effect becomes less significant; raising the cutoff% increases both the effect size and the significance. It therefore seems as though the UICs have a greater positive effect on citation count when they are *rarer* in academics' careers.

The panel method also requires some reflection. While the person-panel is, for reasons already explained, extremely useful in that it should account for variables that would otherwise be difficult to include, the time-panel suffers from low per-year numbers as well as field-field variation within those numbers. It might be better to account for year effects via a much larger dataset, and also to include IPC codes so as to include time-variation within specific fields.

Finally, it should be noted that UIC (ie, whether or not a patent is the result of a university-industrial collaboration) is itself highly correlated with career time, reducing the accuracy of the estimated coefficient for Hypothesis 3. Likelihood-ratio tests of models with both UIC and career time, with only UIC, and with only career time, fail to reject the null that either of the simpler models is better; a way to disentangle the two factors would be highly useful for further analysis.

6.3 Intra-Career Military-Industrial-Academic Patenting

6.3.1 Introduction

The institutional manifestations of the "military-industrial-academic complex" are perhaps clearest to practicing scientists. Academic research objectives spun to appeal to military or industrial grant-awarders, university jobs fairs attended by a host of military agencies and well-paying defense companies, and the omnipresence of military uses (often euphemized) in the "applications" sections of research talks all speak to the tight integration of those three sectors in the American research scene. Military R&D has for decades composed at least 50% of all federal R&D spending [119], meaning that the priorities of the military necessarily find their way into labs (public, private, and academic) across the country. The situation demands clear-eyed analysis.

The most well-defined economic argument against military R&D is the crowding-out hypothesis. According to this hypothesis, military R&D funding replaces, rather than adds to, other types of research, on-net hurting society. Tests of this hypothesis are usually macroscopic, looking economy-wide at how military research funding affects economic and R&D indicators. Implicit in the hypothesis, and sometimes going unsaid, is the assumption that military research is less useful for society, on a per-dollar basis, than civilian research. Connected to this latter issue is the question of dual-use of military technology: to what extent does military research find useful application in the civilian sphere?

This study reframes the question from the macroscopic to the microscopic, focusing on individual scientists who have patented as members of each of universities, private companies, and government military research labs. These researchers, here denoted "military triple-patenters," represent a unique opportunity to study how sector shapes research focuses. Instead of an economy-wide crowding-out, could military

funding lead to individual-level crowding out, shifting the research areas of individual patenters into less-beneficial realms?

The primary hypothesis to be tested is that the military work these triple-patenters do diffuses less than does their other work— that is, that it is less valuable for society's technological development. This hypothesis, which the analysis here concludes is correct, is related to the dual-use question. A second, related hypothesis is that military research represents a significant detour from these scientists' other work— that the kind of work researchers do while working for the military is measurably different from the work they do in the academic and industrial sectors. This hypothesis is related to the crowding-out question, and the analysis ends up not supporting it.

6.3.2 Previous Work: Crowding-Out

The broadest version of the crowding-out hypothesis looks at the economic effects of all military spending. Buck et al. [155] formulate the question in two ways. First, given full employment and utilization of existing capital, *all* "expenditure decisions result in some form of crowding-out." This interpretation "refers to the crowding-out of valuable investment in the civil sector and suggests an inelastic supply of qualified scientists and engineers." An alternative version of the crowding-out hypothesis focuses on "the magnet effect, attracting scarce manpower with adverse effects on industry's international competitiveness. This suggests not so much a direct crowding-out effect but rather the belief that defense R&D was acting as a magnet, attracting firms, especially in the engineering sector, to safe, protected, cost-plus defense markets. The result may be a harmful externality with defense firms operating in a culture of dependency rather than of enterprise, which adversely affects their international competitiveness." These two formulations are not equivalent: in the first, nearly any military allocation would have negative effects on the economy, while the second version implies that good policy can make military spending efficient. In terms of

the analysis presented here, the first version would be reflected in the patent-related crowding-out tests (Hypothesis 5 below), while the second version is closer to what is investigated with Hypothesis 4.

A smaller literature looks specifically at military R&D, and its conclusions are ambivalent. One early study had to remove military fields from their dataset in order to find a positive correlation between federal R&D intensity and economic growth [156], implying that there is some negative effect but not specifically focusing on that conclusion. Later, though, a series of studies by Lichtenberg et al. [157],[158] struggled to link federal R&D (largely military-related) to either positive or negative growth effects. A significant recent study [159] *did* tentatively find a negative crowding-out effect on growth, but also found an even greater positive supply-side growth correlation, for a net positive effect; however, they noted strong country-to-country variation in this result and saw it mostly as a starting point. A few years later, a study [160] used patent counts to conclude that military R&D reduced the productivity of industry-financed R&D, though their model was not particularly extensive. Taking a wide view, Cowan et al. [161] point out some of the methodological problems intrinsic to this kind of sector-wide econometric approach. The question clearly requires more research, and one possible conclusion is that a different approach is needed.

A better approach might be to focus on individual-level career choices. Indeed, while most studies tend to treat labor as simply another factor of production that might be "crowded-out," a few papers have looked specifically at how labor fares in a world of significant military R&D. A given scientist has some skills within a certain research field, but the area of application of their skills depends on where the greatest demand is, and overwhelming military R&D funding accordingly shifts the research landscape. For example, Princeton University is a center of study on quantum cascade lasers, which have applications in, among other places, both civilian gas detection and surface-to-air missile countermeasures. A given QCL researcher

could easily move between these research topics depending on their source of funding and place of employment. Kistiakowsky [162] cites a DoD memo as suggesting that prospective university collaborators "modify [their] approach to fit DoD needs." She is concerned about "the change in emphasis within fields," but notes that "this is a difficult problem because one cannot know the course that scientific inquiry would follow if it were free from constrains," and, anticipating the obvious critique, admits that "both balance and emphasis are shaped by policy decisions on what will be funded, either by the federal government or by other sectors." A military grant is not *inherently* worse than an industrial grant— unless it can be shown that the research products would be less beneficial. Thus the necessity of the present study.

An interesting study on this subject was conducted by Lerner [163], who tracked a group of engineers between 1982 and 1986. He observed extensive mobility between the military and non-military sectors, though non-military to military was the dominant move direction, consistent with the military buildup of that period. The broad message of the study is that scientists and engineers are responsive to incentives and often switch back and forth between military and non-military work; the labor supplies for the two are substantially overlapping. The geography of this phenomenon was examined in [164], which looked at migration of scientists and engineers into "centers of defense procurement" during the same period. Unusually large numbers of "computer scientists, physicists, biologists, physical scientists, chemical engineers, and psychologists" were observed to migrate (this is, of course, the population most likely to alternatively seek work in universities, and thus to be included as triple-patenters in this study, so studies of the triple-patenters are getting right to the heart of this phenomenon). One can see, then, that there is reason to think that labor-constrained crowding-out is a real possibility, and that such crowding-out happens not only at the training/early-career level but also later in careers.

The crowding-out question as examined here looks not at aggregate populations but at individual scientific careers, tracking research topics using patent records. If military research does not represent a subject-wise diversion for scientists — if, as some pro-military-R&D talking points imply, the source and director of funding has little significance, "research being research" regardless of those particulars — then the work a given scientist does for universities should be similar to the work they do in military labs. Or, allowing for *some* inter-sector differences, the work that scientists do at universities should be equally dissimilar to the work they do at either companies or for the military. The triple-patenters dataset allows us to directly measure these similarities/dissimilarities and evaluate this justification.

6.3.3 Previous Work: Dual-Use

Studies of the economics of military R&D have coincided with an increasingly broad pessimism in the current ability of military technologies to diffuse to the civilian sector and thus provide broad social benefit. Foundational work on both semiconductor devices and the internet, among other areas, was conducted under military purview, and defenses of military research funding, fairly, never fail to cite these examples. And, of course, technologies get transferred between applications all the time; that problem alone is not unique to the military. But there are some special aspects of military technology development that make the transfer, potentially, harder. First, technology transfer in general often happens within companies, but primarily-military manufacturers can find it difficult to adapt their cultures and practices to the civilian-production sphere [165], thus restricting dual-use likelihood. Second, military technologies are often highly tied to specific processing methods, such that, even if the end-product would be dual-use, the manufacturing method cannot be adapted to civilian applications [165]. Finally, many argue that modern military technology is

too far removed from civilian use to have dual-use possibilities. As Cowan and Foray put it [161],

> "[previously] technologies... could be applied directly to civilian needs, so a direct and almost immediate spillover took place... Defense R&D was highly relevant to commercial industry, and defense R&D created breakthroughs that advanced US industrial performance. Over time, however, the size and even the direction of spillovers appear to have changed. The generic similarity of civilian and military technologies has declined... induc[ing] a significant diminution of the possibilities of direct and indirect transfer from military to civilian products.

However, they also complicate this simple picture, arguing instead that analyses should be made on an organizational basis with particular attention paid to technology lifecycle. Of course, one can *always* do more exact analyses if one looks at specific cases or narrows down the window of applicability, so their framework, while useful, does little to get at the larger question.

It is possible to look at the dual-use potential of a particular technology, but, at their core, concerns about dual-use are really concerns about whether the advances of military R&D are properly diffusing throughout the economy. Patents, then, provide a good substrate for this study: patent citations allow the tracking of specific technological developments and also serve as a reasonable proxy for overall patent value, as explained in the previous section. In this vein, a study of patents from NASA and other federal labs (many of them military-related) [166] found that by the 1980s those patents had converged in diffusability to that of a random sample of all other patents, as measured by patent citations. Another study [167] used IPC codes to classify patents as civil, military, or mixed civil and military, then looked at citation likelihood. They found that mixed patents have more citations, on average, than purely military patents, and they are also more likely to be cited by purely civil

patents; however, their classification of what counted as "military" was restricted to weapons and munitions, a categorization that does not come close to encompassing the true breadth of militarized R&D. Most recently, a 2017 article [152], taking an approach similar to the one found here, used citation counts to examine the diffusion of military patents as defined by the Derwent technology classification system. Among other findings, that study found no evidence that military patents diffuse less readily than do other patents. Within the existing literature, this is the result most relevant to this study. As with many of the questions surrounding military R&D, though, there is not yet a clear consensus.

To sum up, then, military crowding-out has mostly been tested economy- or sector-wide, and results are not in agreement with each other. Dual-use studies based on patents have generally found that military patents are as valuable as their civilian counterparts, although theoretical/qualitative papers are more skeptical. As of yet, nobody has analyzed the nature of dual-use, technology diffusion, and crowding-out by looking within individual careers. That approach is the main contribution of this section.

6.3.4 Methodology

Dataset Construction

Intra-inventor patent comparisons were used to evaluate research diffusion and research-subject effects. PATSTAT was used, similarly to in Section 6.2, to create a dataset of all inventors from 1971 to 2010 who could be identified as patenting at least once each with a company, a university, and for the US military. US military patents were identified by looking for patent assigned to the Army, Navy, or Air Force (contributions from each were approximately equal). This led to a dataset with 260 inventors and 2970 patents, of which 1363 were assigned to companies, 1125 to universities, and 613 to the military (including patents assigned to multiple sectors).

No patent minimum was imposed, except that the inventor had to have at least one patent representing each sector. Family-family forward citation counts were also collected; as done previously, the datasets compiled by [151] and [149] were used to match inventors across patents and identify keywords from each patent's abstract.

Clearly, this is a restricted definition of "military R&D." Both private companies and universities are also involved in military-funded research, so the sector assignment is only a partially useful method of differentiation—what an observer would call "military research" can happen in *any* sector. That said, as the studies referenced above show, researchers *do* constantly move back and forth between military and civilian work, not just between government military work and non-government military work. Thus, the sectoral method should capture at least some of this movement. While some amount of any effect is "lost" by this analysis, because a certain number of military patents are included in the university and industry bins, the opposite problem (seeing an effect where none exists) should not happen.

It is also necessary to consider the fact that military research is less likely to be patented in general, or is only patented up until it has to be classified, potentially confounding the analysis. In 2003, the US military had an R&D budget of $58 billion [168], an effort that resulted in about 700 government military patents, as defined here. Also in 2003, Raytheon spent $500 million on R&D [169] and filed about 500 patents, while, in the same year, Intel Corporation spent $4.4 billion on R&D [170] and registered over 4000 patents (patent numbers calculated for this study). Even given that Intel's particular industry tends to churn out a large number of patents, the military undoubtedly ranks low on a patents/dollar basis, even compared to non-governmental defense R&D. Assuming its research process is not simply less efficient, it appears that yes, government military R&D is under-patented. Either military research focuses on topics that, for whatever reason, are not patentable, or the government chooses not to patent inventions that would be patented if invented

for another sector. Because the analysis in this chapter is done on a per-patent basis, the first possibility would not cause a problem for the results. The second reason, though, could be problematic, because it would mean that there would be a selection bias in the military-patents dataset. This possibility is discussed further in Section 6.3.7 below.

As a point of comparison, a dataset of "medical triple-patenters" was also constructed. As with the military triple-patenters, these medical inventors all patented in both the university and the private industrial sectors; however, these inventors also patented as members of non-profit or government medical labs. The analogy between government military labs and non-profit medical labs is not exact, of course, but they at least both represent a third economic sector somewhat insulated from concerns of profit and outside the standard academic research sphere. The medical institutions/labs included (selected by looking at the most-highly-patenting "GOV NON-PROFIT" applicants) were: the US Department of Health and Human Services, the Scripps Research Institute, the Mayo Foundation, the Ludwig Institute, the Whitehead Institute, the Salk Institute, the Fred Hutchinson Cancer Research Center, the Oklahoma Medical Research Foundation, and Cold Spring Harbor Laboratory. The resulting dataset totaled 179 inventors, with 589 patents assigned to companies, 700 patents assigned to universities, and 569 patents assigned to the medical non-profits.

The following variables were defined for each patent:

- *sector* (categorical): is the patent assigned to a university, a company, or a non-profit?
- *year* (categorical): year of application filing
- *citations*: family-family citations of this patent
- *person* (categorical): unique identifier of the inventor of the patent

- *t_prev*: time, in years, since the inventor's first patent
- *pat_prev*: count of that inventor's previous patents

In addition, pairwise patent similarity values (as explained above) were calculated for each pair of patents invented by the same person. Each pair was assigned a categorical variable depending on the nature of the two patents: university-university, company-university, etc. The temporal separation Δt was also calculated for each pair; in order to avoid making comparisons between very similar inventions patented at the same time, only pairs with Δt of at least two months were included.

Data Analysis

The triple-patenter datasets were first analyzed in a search for general career patterns. Military triple-patenter patent distributions according to both year and career time-fraction are shown in Figure 6.6 (inventors' first and last patents were excluded from plot (b), as the fractional-time definition includes them). The by-year plot shows approximately similar distributions for all three types of patents, though military patents are somewhat more strongly distributed in the 1980s, perhaps corresponding to the Cold War-era military focus. By contrast, while company and university patenting are approximately equally distributed across careers, military-patenting propensity peaks early in careers and declines from there. This seems to indicate a propensity for these triple-patenters to move away from military labs and into other sectors, though it could also be the result of the year-bias previously mentioned.

There were two hypotheses to be tested. The first is the "what" of this section, focused on patent diffusion propensity:

Hypothesis 4. *Military triple-patenters' military-sector patents diffuse less than their patents from universities or companies.*

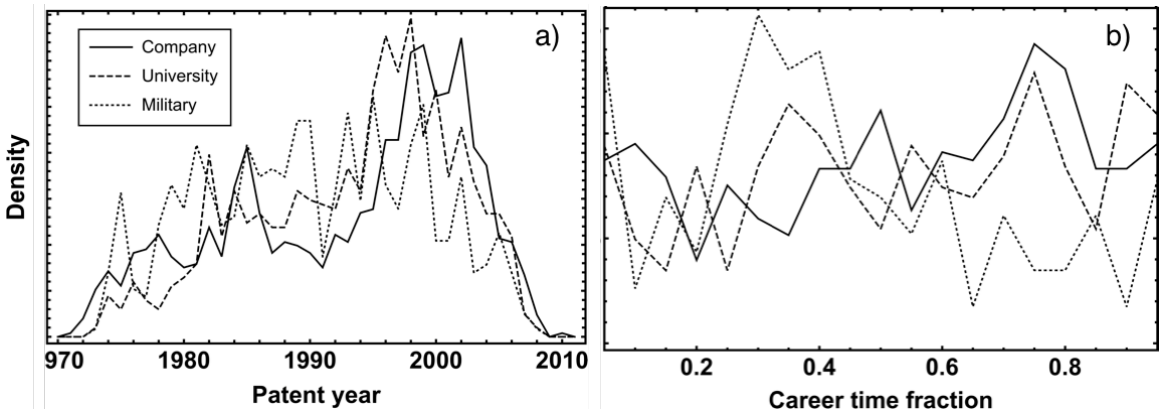

Figure 6.6: Normalized distributions of patents for military triple-patenters, according to (a) application year and (b) time since the inventor's first patent

To test this, a fixed-effects negative-binomial panel model for *citations* was estimated for every applicable patent (having year i and inventor j) in the dataset, with independent variables and fit coefficients

$$\beta_1(company) + \beta_2(military) + \beta_3(t\_prev) + \beta_4(t\_prev)^2 + \beta_5(pat\_prev) + \beta_{6,i}(year) + \beta_{7,j}(person) + \beta_0. \quad (6.5)$$

The time and previous-patent variables are included in order to account for career-time citation/productivity patterns. The person-specific factor variable is included to account for person-specific differences in research field and productivity; one could try to use, for example, IPC codes to account for the former, but otherwise accounting for the latter without taking this panel approach would be difficult. As far as sectoral variation goes, this regression compares company and military patents, implicitly, to university patents.

The second hypothesis is the "why." Given a non-null result from the first hypothesis, can it be explained by intra-inventor research-subject variations between the different sectors?

Hypothesis 5. *Military patents represent an unusually large departure from researchers' other work— military patents are more dissimilar to university and industrial patents than those two sectors' patents are to each other.*

This hypothesis was tested using pairwise similarities. The model is conceptually simple: person-specific research topics have a natural "drift," with researchers' patents getting less similar to each other as the time between them increases. This natural drift can be modulated by the sectoral differences between the patents. The model, then, for the similarity between each pair of patents invented by the same person, is (again using a logit approach, as in the previous section)

$$\begin{aligned}\text{logit}(sim) = \beta_1(\textit{company-company}) + \beta_2(\textit{military-military}) + \\ \beta_3(\textit{company-military}) + \beta_4(\textit{company-university}) + \\ \beta_5(\textit{military-university}) + \beta_6(\Delta t) + \beta_7(\Delta t)^2 + \beta_0. \quad (6.6)\end{aligned}$$

and the analogous regression for the medical triple-patenters dataset. University-university patent pairs are the case to be compared to. Again, because this method only compares *within careers*, control variables that would otherwise be important are implicitly included.

6.3.5 Results

As shown in Table 6.5, for both the military and medical triple-patenter groups, patents invented for industry were found to diffuse more readily than those invented for universities. There was a divergence when comparing to government/non-profit patents, though: government military patents were shown quite clearly to diffuse less than patents invented by the same people for other sectors, with the effect both substantial and significant (in agreement with Hypothesis 4), while the analogous medical nonprofit patents actually diffused *more*.

Interestingly, temporal effects also differed between the two datasets: medical patenters received comparatively fewer citations as their careers progressed, while military patenters received citations equally throughout their careers.

Table 6.5: Hypothesis 4 (Comparative Diffusion of Triple-Patenter Patents From Different Sectors)

	Dependent variable:	
	Citations	
	military	medical
Is a company patent	0.115^{**}	0.279^{**}
	(0.050)	(0.129)
Is a military/medical-nonprofit patent	-0.646^{***}	0.332^{***}
	(0.062)	(0.127)
No. previous patents	-0.003	-0.037^{***}
	(0.003)	(0.011)
Previous career time	-0.019	-0.373^{***}
	(0.065)	(0.139)
(Previous career time)^2	0.0003	-0.001
	(0.0003)	(0.001)
cons.	1.975^{**}	-10.441^{***}
	(0.999)	(4.050)
Note:	*p<0.1; **p<0.05; ***p<0.01	

The similarity analysis, shown in Table 6.6, confirms that patents tend to differ in subject area between sectors, but it provides no evidence to suggest that Hypothesis 5 is correct. For military triple-patenters, there are more differences within scientists' company-patents bodies of work than within their university or military patent sets, but their inter-sector comparisons are about equal. In contrast, for medical triple-patenters, company-company similarities were generally *higher* than their university-university or nonprofit-nonprofit similarities. Furthermore, the medical triple-patenters' similarity calculations showed that their work for nonprofit labs and their work for companies was significantly different (more than their other inter-sector differences).

Table 6.6: Hypothesis 5 (Comparative Similarities of Triple-Patenter Patents From Different/Same Sectors)

	Dependent variable:	
	Similarity	
	military	medical
Time difference	-0.081^{***}	-0.129^{***}
	(0.013)	(0.020)
(Time difference)^2	0.002^{***}	0.003^{***}
	(0.001)	(0.001)
(Company)-(Company)	-0.118^{*}	0.273^{***}
	(0.071)	(0.104)
(Military/Medical-nonprofit)-(Military/Medical-nonprofit)	0.156	-0.102
	(0.099)	(0.122)
(Company)-(Military/Medical-nonprofit)	-0.290^{**}	-0.307^{**}
	(0.116)	(0.156)
(Company)-(University)	-0.275^{***}	-0.142
	(0.088)	(0.141)
(Military/Medical-nonprofit)-(University)	-0.278^{**}	-0.254
	(0.127)	(0.162)
cons.	-2.466^{***}	-2.142^{***}
	(0.071)	(0.099)
Note:	*p<0.1; **p<0.05; ***p<0.01	

6.3.6 Discussion

Most importantly, the above analysis found that, compared to researchers' other work, patents produced for government military labs tend to diffuse less than their other work. This is evidence against the dual-use assertion that work done with military applications in mind will naturally find broader social application. It is important to note that this conclusion is not obvious or self-evident. Rather, the broad worth of military-aimed research is a constant assertion among those whose self-interest relies on the perpetuation of the military-academic-industrial complex. To explain away this result by saying that, well, military research is specialized, of course those patents receive fewer citations, is to give up quite a lot of ground! Note that this is not a comparison between all military patents and all other patents, but between

military patents and other patents invented by the same person—this result suggests that these people could have been doing more broadly useful work if they had been employed elsewhere.

The limits of this result should also be understood clearly. It is not a broad finding that military research, writ-large, diffuses less; it is not a refutation of positive dual-use claims. It is a result derived from a highly specific triple-patenter dataset, analyzed in a particular intra-career way, and it cannot speak to the economy as a whole. But it is a clear result within its domain: high-achieving scientists who work in multiple sectors apparently do less influential work as members of government military labs. The considerable talents of these researchers can be used in any number of ways, but the military utilization is less widely valuable.

Along with this result, it was found, tangentially, that the triple-patenters (both military and medical ones) do more highly-diffusing work when they invent for private companies than when they invent for universities. This aligns with the result in Section 6.2 that industrial collaborations produce more impactful research for academic patenters. At the same time, for medical patenters that high-value-patenting is matched by their output at the nonprofit labs, so this study does not imply that there is some special free-market magic at work.

It is also worth considering how the military-patents result compares to the finding in [152] that civilian and military technologies do not diffuse at different rates. That analysis also uses patent citation counts as the dependent variable, but instead of an inventor-panel approach incorporates more variables related to the specifics of each patent: technological domain, jurisdictional coverage (a proxy for patent quality), etc. Many of those control variables are found to be significant, but whether or not the patent corresponds to a military technology does not appear to have an effect. One explanation of these contradictory results is that military patents may be strongly marked by characteristics accounted for in the control variables of the other study,

such that the military dummy variable ends up being insignificant. For instance, if military patents are generally of lower quality, and lower quality patents get fewer citations, and the study (as it does) includes a separate variable for patent quality, then the military variable may end up not mattering in a regression analysis, even though a more important point — military patents being correlated with lower patent quality and thus fewer citations

The similarity analysis failed to identify any way in which the triple-patenters' military research represented an especially large break from their work in other sectors. Inter-sectoral similarities are all lower than intra-sector ones—unsurprisingly—but those differences are about equal regardless of sector. Military labs clearly change their scientists' research priorities, but the effect is no greater than universities' or companies'. This seems to strike against the more extreme versions of anti-military-R&D sentiment, that military research is *so* alien that it often bears no relation to other areas of study. Apparently it only takes a small shift in research focus to dramatically change applications and patent outcomes.

6.3.7 Approach Robustness

The robustness of the confirmation of Hypothesis 4 is buttressed by the parallel medical-triple-patenters analysis. The two datasets were constructed in the same way and are broadly similar, but only for the military dataset was a negative effect found for work in government labs—in fact, the analogous effect for medical patenters was *positive*. The one observation that should give pause is that there was no career-time effect found for the military patenters; this makes that dataset different from the medical dataset, as well as from the academic-patenters dataset in the previous section, all of which showed significant negative correlations between career time and citation counts. Note that there is a positive correlation between career time and likelihood of a patent being a military patent, perhaps because researchers tend

to move from graduate school to military labs, rather than the opposite. However, removing the sectoral categorical variables does not cause the career time coefficient to become significant, so we can rule out multicollinearity. Tentatively, then, the career-trend observation appears to be an actual result, perhaps something particular to the careers of the "superstars" that represent military triple-patenters.

It is valid, though, to question some of the study's assumptions. It was assumed that the intra-person citation comparison method accounts for scientific field differences, but at a certain point this logic can become circular. Because patents in some fields naturally garner more citations, we don't want to compare citation counts between fields, but this study also shows that military patents receive fewer citations—is this not just as "natural" an occurrence as, say, medical patents targeting a widespread disease being cited more than ones targeting a rare disease? And, come to think of it, why *should* we seek to normalize patent counts in that hypothetical medical case: the first patent seems to be genuinely more important than the second! Acknowledging the reasonableness of this line of questioning, we merely point out that doing the normalization on an inventor-by-inventor basis hopefully avoids as much bias as possible, and that the more conceptual criticism is mostly related to what exactly one is hoping to measure.

Above, it was discussed that there may be a selection bias in which some class of government-researched military patents are less likely to be patented than their counterparts invented in other sectors. More precisely, if the military's propensity to patent a given invention is correlated with the number of citations that patent would have received, then the coefficients found in our regression would be inaccurate. Recall that military patents were found to garner, on average, *fewer* citations. This result could be incorrect if the military preferentially patents its least impactful inventions. This would make some sense—perhaps the military would prefer to simply leave its most important research fully secret. On the other hand, a just-as-reasonable case

could be made for the other possibility: that the military, with no incentive to engage in "patent-trolling" behavior, would only patent its *most* important research; in that case, the result found here would actually be *smaller* than in reality. The literature does not appear to specifically discuss this question, so it would be a good subject for future research.

Finally, some of the caveats of Subsection 6.2.8 still apply here: it may have been better to attack Hypothesis 5 with clustering, particularly given that most people are not constantly moving from sector to sector but rather are are staying in each sector for some period of time—sector membership, therefore, is intertwined with time, probably in a more extreme manner than is capture in the $time$ and $(time)^2$ variables. Similarity-clustering, sector-clustering, and time-clustering could be independently carried out, for example, and the results compared for overlap. At the same time, all similarity values in this section were calculated between single patents, so the previous concern about comparing patents to groups of patents does not hold, and the existence of a possibly-superior method of analysis does not mean the one used here is invalid.

6.4 Conclusion

This chapter presented two separate studies of inter-sector research dynamics. In the first, a search for a "career-defining" effect of university-industrial collaborations (UIC) was not successful; patents that were produced by those collaborations did not tend to particularly resemble their future patents, nor do they stand out, subject-wise, from researchers' previous patents. UIC-derived patents *were* shown, tentatively, to diffuse more than non-UIC patents. The second study presented convincing evidence against the positive military-industrial-academic dual-use hypothesis: scientists who worked in multiple sectors, including government military labs, tended to do less

impactful work for those labs compared to the same people's work in industry or academia. However, no evidence was found for a strong crowding-out claim wherein military work was disproportionally different from non-military work. These results, while obviously not definitive, can inform stakeholders and policymakers as they consider what role American military R&D infrastructure can play in meeting current and future challenges.

Two themes unite these two sections. The first is the method used: intra-career patent comparisons with similarity indexes. Comparing patents while accounting for person-specific factors seems as though it should be an extremely fruitful route. It can eliminate the need for many of the control variables that are otherwise required for comparing patents, and it can also capture effects that other analyses would miss. The search for similarity-trajectory modifiers, though not applied to a non-null result here, should, it must be believed, give fruitful results when applied to more significant effects. As an example, it would be interesting to examine whether similarity metrics can capture the effect of the NSF's major early-career grants in shaping scientists' future research. The intra-career similarity results from Section 6.3, interesting in their own right, also suggest future work: this study used sectoral "splits," but other splits, like pre-tenure/post-tenure or home country/other country, could also give interesting results.

The other unifying theme is that both sections examined the extremely idiosyncratic, variable process by which scientists end up researching some topics instead of others and producing research output of variable importance and quality. Passing over inaccessible determinants ("number theory is beautiful," "I had a fever that day and the experiment went poorly", etc), we searched for *external* nudges as captured in patent data. In the first section, these nudges were university-industrial collaborations, which, it seemed, had the potential to disproportionally influence research choices and quality. In the second section, these nudges were peoples' sector of work,

which have the effect of imposing a significant Δ onto whatever topic a researcher happens to personally most desire to research. The contemporary research climate is extremely complex, and these nudges probably have a large effect in the aggregate—new technologies are not simple functions of allocated funding. The introduction to this thesis alluded to the large-scale public interventions necessary to stave off the very worst disasters of climate change, but that investment could well be wasted if it ends up siphoned off down unproductive paths, into military technologies without a strong dual-use argument, or into university-industry collaborations that are not well thought-out and understood. For these reasons, the research dynamics explored here should continue to be an area of study.

Chapter 7

Summary

This thesis described contributions in multiple technical areas. In Chapter 2, the silicon heterojunction solar cell was reviewed, and guidelines, developed through a combination of simulation and theory, were developed in order to inform device engineers and material scientists. In Chapter 3, the TiO_2/Si interface, important for photovoltaic applications, was explored in detail. In one study, external quantum efficiency measurements were compared to simulations in order to measure the interface quality; in another, experiments tested the possibility of using TiO_2 films as tunnel layers in solar cell anodes, with the conclusion that the barrier TiO_2 presents to holes on n-Si probably precludes maintaining both passivation and tunneling simultaneously.

In Chapter 4, the method of diode reverse recovery—transient electrical measurement—was extended to the double heterojunction silicon solar cell. After a detailed theoretical development, simulation results were presented in reduced units applicable to a wide range of devices. Non-idealities at both interfaces, along with geometric effects, were included in the analysis and validated by comparison to full 3D device simulations. This modeling was successfully applied to experimental

data from PEDOT/n-Si/TiO_2 devices, and interface parameters for both of those films on silicon were extracted.

In Chapter 5, the development of a process for ALD nickel oxide films was presented. The details of the process development, including precursor, oxidizer, and temperature selection, were explored in detailed, with those choices related to film properties. X-ray photoelectron spectroscopy was utilized heavily for extraction of film stoichiometries. A copper oxide ALD process was also developed, and the two processes were combined to produce Cu-doped NiO films. Copper doping was found to increase film conductivity by about 10x, a result that correlated with spectroscopic measurements. The effects of copper doping on the conduction mechanism and the device structure were also measured. ALD-NiO was integrated both into NiO/silicon test devices, where it was found to depin the interface Fermi level, and into perovskite solar cells, where it was found to give device figures of merit comparable to those in cells utilizing the more standard solution-deposited NiO films.

In Chapter 6, dual studies of inter-sector patenting dynamics were presented. One section examined university-industrial collaboration, finding that patents derived from that collaboration tend to diffuse more than university-only patents but not finding any pattern in how those patents shift scientists' research choices. In the second section, it was found that researchers who have worked in government military labs, in universities, and in private companies tend to do their least impactful work for the military labs. Furthermore, both that group and an analogous group of medical researchers tend to do their most impactful work while at private companies. However, it was not found that military research, topic-wise, is especially different from research done in other sectors.

Possible narrowly-focused future research was already suggested within each individual chapter, but there are still some points left to be made on the broader themes of the thesis. First, material development is fickle: results can vary, experiment-

to-experiment, not just between labs but even between runs on the same tool. A seat-of-the-pants approach, focused each day on getting short-term results, will not work well in the long run: it is better to develop the most stable processes possible and then build on those, even if initial progress is slower. SHJC researchers, even as they work with novel materials and as their style of discourse is shaped by emerging-materials scientists, should try to maintain a strong link to their heritage in the more formal, grounded style of traditional silicon semiconductor development.

Second, Chapter 4 was motivated largely by the desire to develop a technique that not only worked but was also simple to set up, and key results in other chapters would not have been possible without the arrival, midway through the graduate work presented here, of a public, accessible XPS/UPS tool. Analysis technique choice should prioritize ease of implementation, with allowance made, of course, for the critical importance of some more difficult methods.

Finally, the thought processes and research paths shown in Chapters 2 through 5—somewhat idiosyncratic and unpredictable—are reflected in the choice, in Chapter 6, to focus on intra-career comparisons and methods. Sometimes, highly-directed, centralized research infrastructures can obtain specific desired results, but on an individual level research is messy and hard to predict, motivated from many directions at once and leading often in unexpected directions. The American research and governance community should keep this picture in mind as they turn their attention to our ever-more-pressing objectives.

Chapter 8

ALD-NiO(:Cu) Process

8.1 ALD-NiO

A typical ALD-NiO process had the following recipe:

1. Open pump valve
2. Set N_2 flow to 200 sccm
3. Set manifolds' $T_{setpoint}$ to 130°C.
4. Set NiAMD precursor $T_{setpoint}$ to 120°C.
5. Set chamber and door to $T_{process}$ (typically 150°C to 200°C)
6. After reaching setpoints, wait for 600 seconds
7. Supercycle (repeat n times)
   (a) Ni cycle
      i. Set N_2 flow to 40 sccm
      ii. Close pump valve
      iii. Pulse NiAMD for 500 ms
      iv. Wait for 30 seconds
      v. Open pump valve
      vi. Set N_2 flow to 100 sccm
      vii. Wait for 20 seconds
   (b) O_3 cycle
      i. Set N_2 flow to 40 sccm
      ii. Close pump valve
      iii. Pulse O_3 for 250 ms
      iv. Wait for 30 seconds

v. Open pump valve
vi. Set N_2 flow to 100 sccm
vii. Wait for 20 seconds

8. Return chamber to standby

8.2 ALD-NiO:Cu

A typical 25%-by-cycle ALD-NiO:Cu process had the following recipe:

1. Open pump valve
2. Set N_2 flow to 200 sccm
3. Set manifolds' $T_{setpoint}$ to 130°C.
4. Set NiAMD precursor $T_{setpoint}$ to 120°C.
5. Set $Cu(dmap)_2$ precursor $T_{setpoint}$ to 105°C.
6. Set chamber and door to $T_{process}$ (typically 150°C to 200°C)
7. After reaching setpoints, wait for 15 minutes
8. Supercycle (repeat n times)
   (a) Ni supercycle (repeat 3 times)
      i. Ni cycle
         A. Set N_2 flow to 40 sccm
         B. Close pump valve
         C. Pulse NiAMD for 500 ms
         D. Wait for 30 seconds
         E. Open pump valve
         F. Set N_2 flow to 100 sccm
         G. Wait for 20 seconds
      ii. O_3 cycle
         A. Set N_2 flow to 40 sccm
         B. Close pump valve
         C. Pulse O_3 for 250 ms
         D. Wait for 30 seconds
         E. Open pump valve
         F. Set N_2 flow to 100 sccm
         G. Wait for 20 seconds
   (b) Wait for 120 seconds

(c) Cu supercycle

i. Cu cycle

A. Set N_2 flow to 40 sccm

B. Close pump valve

C. Pulse $Cu(dmap)_2$ for 250 ms

D. Wait for 30 seconds

E. Open pump valve

F. Set N_2 flow to 100 sccm

G. Wait for 20 seconds

ii. O_3 cycle

A. Set N_2 flow to 40 sccm

B. Close pump valve

C. Pulse O_3 for 250 ms

D. Wait for 30 seconds

E. Open pump valve

F. Set N_2 flow to 100 sccm

G. Wait for 20 seconds

9. Repeat Ni Supercycle 5 times

10. Return chamber to standby

Bibliography

[1] "Global Warming of 1.5C: An IPCC Special Report," Intergovernmental Panel on Climate Change, , 2018. Online: https://www.ipcc.ch/sr15/

[2] R. Luxemburg, "The Crisis of German Social Democracy (trans. by Dave Hollis)," 1915. Online: https://www.marxists.org/archive/luxemburg/1915/junius/

[3] M. Z. Jacobson, M. A. Delucchi, G. Bazouin, Z. A. Bauer, C. C. Heavey, E. Fisher, S. B. Morris, D. J. Piekutowski, T. A. Vencill, and T. W. Yeskoo, "100% clean and renewable wind, water, and sunlight (WWS) all-sector energy roadmaps for the 50 United States," *Energy and Environmental Science*, vol. 8, no. 7, 2093–2117, 2015.

[4] H. Etzkowitz, "The Triple Helix of university - industry - government relations," *Social Science Information*, vol. 42, no. 3, 293–337, 2003.

[5] 116th United States Congress, "H. Res. 109 - House Simple Resolution - Green New Deal," 1–14, 2019. Online: https://www.congress.gov/116/bills/hres109/BILLS-116hres109ih.pdf

[6] P. Wurfel, *Physics of Solar Cells: From Principles to New Concepts*. Wiley, 2007.

[7] S. de Wolf, A. Descoeudres, Z. C. Holman, C. Ballif, S. D. Wolf, A. Descoeudres, Z. C. Holman, and C. Ballif, "High-efficiency silicon heterojunction solar cells: A review," *Green*, vol. 2, no. 1, 7–24, 2012.

[8] S. Avasthi, S. Lee, Y. L. Loo, and J. C. Sturm, "Role of majority and minority carrier barriers silicon/organic hybrid heterojunction solar cells," *Advanced Materials*, vol. 23, no. 48, 5762–5766, 2011.

[9] A. S. Erickson, A. Zohar, and D. Cahen, "n-Si-organic inversion layer interfaces: A low temperature deposition method for forming a p-n homojunction in n-Si," *Advanced Energy Materials*, vol. 4, no. 9, 1–4, 2014.

[10] K. A. Nagamatsu, S. Avasthi, J. Jhaveri, and J. C. Sturm, "A 12% efficient silicon/PEDOT:PSS heterojunction solar cell fabricated at <100C," *IEEE Journal of Photovoltaics*, vol. 4, no. 1, 260–264, 2014.

[11] C. Battaglia, X. Yin, M. Zheng, I. D. Sharp, T. Chen, S. Mcdonnell, A. Azcatl, C. Carraro, B. Ma, R. Maboudian, R. M. Wallace, and A. Javey, "Hole selective MoOx contact for silicon solar cells," *Nano Letters*, vol. 14, no. 2, 967–971, 2014.

[12] L. Gerling, S. Mahato, C. Voz, R. Alcubilla, and J. Puigdollers, "Characterization of transition metal oxide/silicon heterojunctions for solar cell applications," *Applied Sciences*, vol. 5, no. 4, 695–705, 2015.

[13] S. Jäckle, M. Mattiza, M. Liebhaber, G. Brönstrup, M. Rommel, K. Lips, and S. Christiansen, "Junction formation and current transport mechanisms in hybrid n-Si/PEDOT:PSS solar cells," *Scientific Reports*, vol. 5, 13008, 2015.

[14] W. Shockley and W. T. Read, "Statistics of the recombinations of holes and electrons," *Physical Review*, vol. 87, no. 5, 835–842, 1952.

[15] R. N. Hall, "Electron-hole recombination in germanium," *Physical Review*, vol. 87, no. 6, 387, 1952.

[16] R. S. Bonilla and P. R. Wilshaw, "On the c-Si/SiO2 interface recombination parameters from photoconductance decay measurements," *Journal of Applied Physics*, vol. 121, no. 13, 135301, 2017.

[17] Z. P. Ling, Z. Xin, C. Ke, K. J. Buatis, S. Duttagupta, J. S. Lee, A. Lai, A. Hsu, J. Rostan, and R. Stangl, "Comparison and characterization of different tunnel layers, suitable for passivated contact formation," *Japanese Journal of Applied Physics*, vol. 56, no. 8S2, 08MA01, 2017.

[18] S. Fonash, *Solar Cell Device Physics*. New York: Academic Press, 1981.

[19] J. Jhaveri, "Interface recombination in TiO2 / silicon heterojunctions for silicon photovoltaic applications," Ph.D. dissertation, Princeton University, 2018.

[20] R. Sinton, A. Cuevas, and M. Stuckings, "Quasi-steady-state photoconductance, a new method for solar cell material and device characterization," in *Conference Record of the Twenty Fifth IEEE PVSC*, 1996, 457–460.

[21] Y. I. Ogita, "Bulk lifetime and surface recombination velocity measurement method in semiconductor wafers," *Journal of Applied Physics*, vol. 79, no. 9, 6954–6960, 1996.

[22] K. L. Luke and L. J. Cheng, "Analysis of the interaction of a laser pulse with a silicon wafer: Determination of bulk lifetime and surface recombination velocity," *Journal of Applied Physics*, vol. 61, no. 6, 2282–2293, 1987.

[23] A. B. Sproul, "Dimensionless solution of the equation describing the effect of surface recombination on carrier decay in semiconductors," *Journal of Applied Physics*, vol. 76, no. 5, 2851–2854, 1994.

[24] J. Moulder, W. Stickle, P. E. Sobol, and K. Bomben, *Handbook of X-ray Photoelectron Spectroscopy*. Eden Prairie: Perkin-Elmer Corporation, 1992.

[25] D. A. Shirley, "High-resolution x-ray photoemission spectrum of the valence bands of gold," *Physical Review B*, vol. 5, no. 12, 4709–4714, 1972.

[26] R. Schlaf, "Calibration of photoemission spectra and work function determination," Online: http://rsl.eng.usf.edu/Documents/Tutorials/PEScalibration.pdf

[27] X. Yang, Q. Bi, H. Ali, K. Davis, W. V. Schoenfeld, and K. Weber, "High-performance TiO2-based electron-selective contacts for crystalline silicon solar cells," *Advanced Materials*, vol. 28, no. 28, 5891–5897, 2016.

[28] S. Avasthi, W. E. McClain, G. Man, A. Kahn, J. Schwartz, and J. C. Sturm, "Hole-blocking titanium-oxide/silicon heterojunction and its application to photovoltaics," *Applied Physics Letters*, vol. 102, no. 20, 203901, 2013.

[29] K. A. Nagamatsu, S. Avasthi, G. Sahasrabudhe, G. Man, J. Jhaveri, A. H. Berg, J. Schwartz, A. Kahn, S. Wagner, and J. C. Sturm, "Titanium dioxide/silicon hole-blocking selective contact to enable double-heterojunction crystalline silicon-based solar cell," *Applied Physics Letters*, vol. 106, no. 12, 123906, 2015.

[30] X. Yin, C. Battaglia, Y. Lin, K. Chen, M. Hettick, M. Zheng, C.-Y. Chen, D. Kiriya, and A. Javey, "19.2% efficient InP heterojunction solar cell with electron-selective TiO2 contact," *ACS Photonics*, vol. 1, no. 12, 1245–1250, 2014.

[31] S. Ito, S. M. Zakeeruddin, R. Humphry-Baker, P. Liska, R. Charvet, P. Comte, M. K. Nazeeruddin, P. Péchy, M. Takata, H. Miura, S. Uchida, and M. Grätzel, "High-efficiency organic-dye-sensitized solar cells controlled by nanocrystalline-TiO2 electrode thickness," *Advanced Materials*, vol. 18, no. 9, 1202–1205, 2006.

[32] W. Ke, G. Fang, J. Wang, P. Qin, H. Tao, H. Lei, Q. Liu, X. Dai, and X. Zhao, "Perovskite solar cell with an efficient TiO2 compact film," *ACS Applied Materials and Interfaces*, vol. 6, no. 18, 15 959–15 965, 2014.

[33] G. Man, J. Schwartz, J. C. Sturm, and A. Kahn, "Electronically passivated hole-blocking titanium dioxide/silicon heterojunction for hybrid silicon photovoltaics," *Advanced Materials Interfaces*, vol. 3, no. 15, 1–8, 2016.

[34] K. M. Gad, D. Vossing, A. Richter, B. Rayner, L. M. Reindl, S. E. Mohney, and M. Kasemann, "Ultrathin titanium dioxide nanolayers by atomic layer deposition for surface passivation of crystalline silicon," *IEEE Journal of Photovoltaics*, vol. 6, no. 3, 649–653, 2016.

[35] G. Sahasrabudhe, S. M. Rupich, J. Jhaveri, A. H. Berg, K. Nagamatsu, G. Man, Y. J. Chabal, A. Kahn, S. Wagner, J. C. Sturm, and J. Schwartz,

"Low temperature synthesis of a TiO2/Si heterojunction," *Journal of the American Chemical Society*, vol. 137, no. 47, 14 842–14 845, 2015.

[36] J. Bullock, M. Boccard, Y. Wan, A. Javey, S. Essig, M. Hettick, H. Wang, Z. Xu, A. Cuevas, W. Ji, and C. Ballif, "Stable dopant-free asymmetric heterocontact silicon solar cells with efficiencies above 20%," *ACS Energy Letters*, vol. 3, no. 3, 508–513, 2018.

[37] R. Mishima, M. Hino, K. Yamamoto, H. Uzu, T. Meguro, R. Santbergen, M. Zeman, and J. Blanker, "Minimizing optical losses in monolithic perovskite/c-Si tandem solar cells with a flat top cell," *Optics Express*, vol. 24, no. 18, A1288, 2016.

[38] K. Jäger, L. Korte, B. Rech, and S. Albrecht, "Numerical optical optimization of monolithic planar perovskite-silicon tandem solar cells with regular and inverted device architectures," *Optics Express*, vol. 25, no. 12, A473, 2017.

[39] C. D. Bailie, M. G. Christoforo, J. P. Mailoa, A. R. Bowring, E. L. Unger, W. H. Nguyen, J. Burschka, N. Pellet, J. Z. Lee, M. Grätzel, R. Noufi, T. Buonassisi, A. Salleo, and M. D. McGehee, "Semi-transparent perovskite solar cells for tandems with silicon and CIGS," *Energy and Environmental Science*, vol. 8, no. 3, 956–963, 2015.

[40] A. J. Bett, P. S. Schulze, K. Winkler, J. Gasparetto, P. F. Ndione, M. Bivour, A. Hinsch, M. Kohlstädt, S. Lee, S. Mastroianni, L. E. Mundt, M. Mundus, C. Reichel, A. Richter, C. Veit, K. Wienands, U. Würfel, W. Veurman, S. W. Glunz, M. Hermle, and J. C. Goldschmidt, "Low temperature perovskite solar cells with an evaporated TiO2 compact layer for perovskite silicon tandem solar cells," *Energy Procedia*, vol. 124, 567–576, 2017.

[41] A. F. Thomson and K. R. McIntosh, "Light-enhanced surface passivation of TiO2-coated silicon," *Progress in Photovoltaics: Research and Applications*, vol. 20, no. 3, 343–349, 2012.

[42] X. Yang and K. Weber, "N-type silicon solar cells featuring an electron-selective TiO2 contact," in *IEEE 42nd Photovoltaic Specialist Conference*, 2015, 13–16.

[43] W.-C. Wang, M.-C. Tsai, J. Yang, C. Hsu, and M.-J. Chen, "Efficiency enhancement of nanotextured black silicon solar cells using Al2O3/TiO2 dual-layer passivation stack prepared by atomic layer deposition," *ACS Applied Materials & Interfaces*, vol. 7, no. 19, 10 228–10 237, 2015.

[44] Z. Ling, J. He, X. He, M. Liao, P. Liu, Z. Yang, J. Ye, and P. Gao, "Excellent passivation of silicon surfaces by thin films of electron-beam-processed titanium dioxide," *IEEE Journal of Photovoltaics*, vol. 7, no. 6, 1551–1555, 2017.

[45] J. Jhaveri, K. A. Nagamatsu, A. H. Berg, G. Man, G. Sahasrabudhe, S. Wagner, J. Schwartz, A. Kahn, and J. C. Sturm, "Double-heterojunction crystalline

silicon solar cell with electron-selective TiO2 cathode contact fabricated at 100C with open-circuit voltage of 640 mV," in *IEEE 42nd Photovoltaic Specialist Conference*, 2015, 1–4.

[46] B. Liao, B. Hoex, A. G. Aberle, D. Chi, and C. S. Bhatia, "Excellent c-Si surface passivation by low-temperature atomic layer deposited titanium oxide," *Applied Physics Letters*, vol. 104, no. 25, 253903, 2014.

[47] J. Jhaveri, S. Avasthi, G. Man, W. E. Mcclain, K. Nagamatsu, J. Schwartz, and J. C. Sturm, "Hole-blocking crystalline-silicon / titanium-oxide heterojunction with very low interface recombination velocity," in *IEEE 39th Photovoltaic Specialists Conference*, 2013, 3292–3296.

[48] W. Lu, C. Leendertz, L. Korte, J. A. Tofflinger, and H. Angermann, "Passivation properties of subnanometer thin interfacial silicon oxide films," *Energy Procedia*, vol. 55, 805–812, 2014.

[49] B. Stegemann, K. M. Gad, P. Balamou, D. Sixtensson, D. Vossing, M. Kasemann, and H. Angermann, "Ultra-thin silicon oxide layers on crystalline silicon wafers: Comparison of advanced oxidation techniques with respect to chemically abrupt SiO2/Si interfaces with low defect densities," *Applied Surface Science*, vol. 395, 78–85, 2017.

[50] Asuha, T. Yuasa, O. Maida, and H. Kobayashi, "Effects of postmetallization annealing on ultrathin SiO2 layer properties," *Applied Physics Letters*, vol. 80, no. 22, 4175–4177, 2002.

[51] J. B. Heng, J. Fu, B. Kong, Y. Chae, W. Wang, Z. Xie, A. Reddy, K. Lam, C. Beitel, C. Liao, C. Erben, Z. Huang, and Z. Xu, ">23% high-efficiency tunnel oxide junction bifacial solar cell with electroplated Cu gridlines," *IEEE Journal of Photovoltaics*, vol. 5, no. 1, 82–86, 2015.

[52] P. Mur, M. N. Semeria, M. Olivier, A. M. Papon, C. Leroux, G. Reimbold, P. Gentile, N. Magnea, T. Baron, R. Clerc, and G. Ghibaudo, "Ultra-thin oxides grown on silicon (100) by rapid thermal oxidation for CMOS and advanced devices," *Applied Surface Science*, vol. 175-176, 726–733, 2001.

[53] Y. Lee, W. Oh, V. A. Dao, S. Q. Hussain, and J. Yi, "Ultrathin oxide passivation layer by rapid thermal oxidation for the silicon heterojunction solar cell applications," *International Journal of Photoenergy*, vol. 2012, 753456, 2012.

[54] A. Schenk and G. Heiser, "Modeling and simulation of tunneling through ultra-thin gate dielectrics," *Journal of Applied Physics*, vol. 81, no. 12, 7900–7908, 1997.

[55] N. Yang, W. K. Henson, J. R. Hauser, and J. J. Wortman, "Modeling study of ultrathin gate oxides using direct tunneling current and capacitance-voltage measurements in MOS devices," *IEEE Transactions on Electron Devices*, vol. 46, no. 7, 1464–1471, 1999.

[56] J. Jhaveri, A. H. Berg, and J. C. Sturm, "Isolation of hole versus electron current at p-Si/TiO2 selective contact using a heterojunction bipolar transistor structure," *IEEE Journal of Photovoltaics*, vol. 8, no. 3, 726–732, 2018.

[57] A. K. Chandiran, A. Yella, M. Stefik, L.-P. Heiniger, P. Comte, M. K. Nazeeruddin, and M. Grätzel, "Low-temperature crystalline titanium dioxide by atomic layer deposition for dye-sensitized solar cells." *ACS Applied Materials and Interfaces*, vol. 5, no. 8, 3487–93, 2013.

[58] M. C. Biesinger, B. P. Payne, A. P. Grosvenor, L. W. Lau, A. R. Gerson, and R. S. C. Smart, "Resolving surface chemical states in XPS analysis of first row transition metals, oxides and hydroxides: Cr, Mn, Fe, Co and Ni," *Applied Surface Science*, vol. 257, no. 7, 2717–2730, 2011.

[59] P. F. Luo, T. Kuwana, D. K. Paul, and P. M. Sherwood, "Electrochemical and XPS study of the nickel-titanium electrode surface," *Analytical Chemistry*, vol. 68, no. 19, 3330–3337, 1996.

[60] A. Lagunin, A. Zakharov, D. Filimonov, and V. Poroikov, "QSAR modelling of rat acute toxicity on the basis of PASS prediction (Supporting Information)," *Molecular Informatics*, vol. 30, no. 2-3, 241–250, 2011.

[61] A. H. Berg, K. A. Nagamatsu, and J. C. Sturm, "Extraction of front- and rear-interface recombination in silicon double-heterojunction solar cells by reverse bias transients," *IEEE Transactions on Electron Devices*, vol. 64, no. 11, 4518–4525, 2017.

[62] A. H. Berg and J. C. Sturm, "Extracting interface recombination velocities from double-heterojunction solar cell reverse-recovery characteristics," *Device Research Conference - Conference Digest*, vol. 106, no. 5, 1–2, 2017.

[63] R. Kingston, "Switching time in junction diodes and junction transistors," *Proceedings of the IRE*, vol. 42, no. 5, 829–834, 1954.

[64] B. Lax and S. F. Neustadter, "Transient response of a p-n junction," *Journal of Applied Physics*, vol. 25, no. 9, 1148–1154, 1954.

[65] M. Byczkowski and J. R. Madigan, "Minority carrier lifetime in p-n junction devices," *Journal of Applied Physics*, vol. 28, no. 8, 878, 1957.

[66] A. Grove and C. Sah, "Simple analytical approximations to the switching times in narrow base diodes," *Solid State Electronics*, vol. 7, no. 1, 107–110, 1964.

[67] L. Davidson, "Simple expression for storage time of arbitrary base diode," *Solid-State Electronics*, vol. 9, no. 1965, 1145–1147, 1966.

[68] D. Scharfetter, "Minority carrier injection and charge storage in epitaxial Schottky barrier diodes," *Solid-State Electronics*, vol. 8, no. 3, 299–311, 1965.

[69] G. Greczynski, T. Kugler, M. Keil, W. Osikowicz, M. Fahlman, and W. R. Salaneck, "Photoelectron spectroscopy of thin films of PEDOT-PSS conjugated polymer blend: A mini-review and some new results," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 121, no. 1-3, 1–17, 2001.

[70] J. Thomas, *Numerical Partial Differential Equations: Finite Difference Methods*. New York: Springer-Verlag, 1995.

[71] P. Chen, K. Misiakos, A. Neugroschel, and F. Lindholm, "Analytical solution for two-dimensional current injection from shallow p-n junctions," *IEEE Transactions on Electron Devices*, vol. 32, no. 11, 2292–2296, 1985.

[72] A. B. Prakoso, L. Ke, J. Wang, Z. Li, C. Jiang, and Rusli, "Reverse recovery transient characteristic of PEDOT:PSS/n-Si hybrid organic-inorganic heterojunction," *Organic Electronics*, vol. 42, 269–274, 2017.

[73] S. Peters, "Rapid thermal processing of crystalline silicon materials and solar cells," Ph.D. dissertation, University of Konstanz, 2004.

[74] K. X. Steirer, P. F. Ndione, N. E. Widjonarko, M. T. Lloyd, J. Meyer, E. L. Ratcliff, A. Kahn, N. R. Armstrong, C. J. Curtis, D. S. Ginley, J. J. Berry, and D. C. Olson, "Enhanced efficiency in plastic solar cells via energy matched solution processed NiOx interlayers," *Advanced Energy Materials*, vol. 1, no. 5, 813–820, 2011.

[75] T. Salim, S. Sun, Y. Abe, A. Krishna, A. C. Grimsdale, and Y. M. Lam, "Perovskite-based solar cells: Impact of morphology and device architecture on device performance," *Journal of Materials Chemistry A*, vol. 3, no. 17, 8943–8969, 2015.

[76] S. Sista, M.-H. Park, Z. Hong, Y. Wu, J. Hou, W. L. Kwan, G. Li, and Y. Yang, "Highly efficient tandem polymer photovoltaic cells." *Advanced Materials*, vol. 22, no. 3, 380–383, 2010.

[77] M. D. Irwin, D. B. Buchholz, A. W. Hains, R. P. H. Chang, and T. J. Marks, "p-type semiconducting nickel oxide as an efficiency-enhancing anode interfacial layer in polymer bulk-heterojunction solar cells," *Proceedings of the National Academy of Sciences*, vol. 105, no. 8, 2783–2787, 2008.

[78] I. M. Chan, T. Y. Hsu, and F. C. Hong, "Enhanced hole injections in organic light-emitting devices by depositing nickel oxide on indium tin oxide anode," *Applied Physics Letters*, vol. 81, no. 10, 1899–1901, 2002.

[79] B. Mustafa, J. Griffin, A. S. Alsulami, D. G. Lidzey, and A. R. Buckley, "Solution processed nickel oxide anodes for organic photovoltaic devices," *Applied Physics Letters*, vol. 104, no. 6, 063302, 2014.

[80] P. Schulz, S. R. Cowan, Z. L. Guan, A. Garcia, D. C. Olson, and A. Kahn, "NiOX/MoO3 bi-layers as efficient hole extraction contacts in organic solar cells," *Advanced Functional Materials*, vol. 24, no. 5, 701–706, 2014.

[81] E. L. Ratcliff, J. Meyer, K. X. Steirer, A. Garcia, J. J. Berry, D. S. Ginley, D. C. Olson, A. Kahn, and N. R. Armstrong, "Evidence for near-surface NiOOH species in solution-processed NiOx selective interlayer materials: Impact on energetics and the performance of polymer bulk heterojunction photovoltaics," *Chemistry of Materials*, vol. 23, no. 22, 4988–5000, 2011.

[82] J. R. Manders, S.-W. Tsang, M. J. Hartel, T.-H. Lai, S. Chen, C. M. Amb, J. R. Reynolds, and F. So, "Solution-processed nickel oxide hole transport layers in high efficiency polymer photovoltaic cells," *Advanced Functional Materials*, vol. 23, no. 23, 2993–3001, 2013.

[83] R. Islam, G. Shine, and K. C. Saraswat, "Schottky barrier height reduction for holes by Fermi level depinning using metal/nickel oxide/silicon contacts," *Applied Physics Letters*, vol. 105, no. 18, 182103, 2014.

[84] X. Yang, W. Liu, J. Chen, and Y. Sun, "On the annealing-induced enhancement of the interface properties of NiO:Cu/wet-SiOx/n-Si tunnelling junction solar cells," *Applied Physics Letters*, vol. 112, no. 17, 173904, 2018.

[85] A. H. Berg, G. S. Sahasrabudhe, R. A. Kerner, B. P. Rand, J. Schwartz, and J. C. Sturm, "Electron-blocking NiO/crystalline n-Si heterojunction formed by ALD at 175C," in *Device Research Conference - Conference Digest*. IEEE, 2016.

[86] N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials*. Oxford: Clarendon Press, 1979.

[87] G. A. Sawatzky and J. W. Allen, "Magnitude and origin of the band gap in NiO," *Physical Review Letters*, vol. 53, no. 24, 2339–2342, 1984.

[88] V. I. Anisimov, M. A. Korotin, and E. Z. Kurmaev, "Band-structure description of Mott insulators (NiO, MnO, FeO, CoO)," *Journal of Physics: Condensed Matter*, vol. 2, no. 17, 3973–3987, 1990.

[89] S. Hüfner, "Electronic structure of NiO and related 3d-transition-metal compounds," *Advances in Physics*, vol. 43, no. 2, 183–356, 1994.

[90] D. Adler and J. Feinleib, "Electrical and optical properties of narrow-band materials," *Physical Review B*, vol. 2, no. 8, 3112–3134, 1970.

[91] F. J. Morin, "Electrical properties of NiO," *Physical Review*, vol. 93, no. 6, 1199–1204, 1954.

[92] W.-Y. Chen, J.-S. Jeng, K.-L. Huang, and J.-S. Chen, "Modulation of Ni valence in p-type NiO films via substitution of Ni by Cu," *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 31, no. 2, 021501, 2013.

[93] A. Huang, L. Lei, Y. Chen, Y. Yu, Y. Zhou, Y. Liu, S. Yang, S. Bao, R. Li, and P. Jin, "Minimizing the energy loss of perovskite solar cells with Cu+doped NiOx processed at room temperature," *Solar Energy Materials and Solar Cells*, vol. 182, 128–135, 2018.

[94] J. H. Kim, P. W. Liang, S. T. Williams, N. Cho, C. C. Chueh, M. S. Glaz, D. S. Ginger, and A. K. Y. Jen, "High-performance and environmentally stable planar heterojunction perovskite solar cells based on a solution-processed copper-doped nickel oxide hole-transporting layer," *Advanced Materials*, vol. 27, no. 4, 695–701, 2015.

[95] W. Guo, K. Hui, and K. Hui, "High conductivity nickel oxide thin films by a facile solgel method," *Materials Letters*, vol. 92, 291–295, 2013.

[96] I. G. Austin and N. F. Mott, "Polarons in crystalline and non-crystalline materials," *Advances in Physics*, vol. 18, no. 71, 41–102, 1969.

[97] T. Dutta, P. Gupta, A. Gupta, and J. Narayan, "Effect of Li doping in NiO thin films on its transparent and conducting properties and its application in heteroepitaxial p-n junctions," *Journal of Applied Physics*, vol. 108, 083715, 2010.

[98] S. Moghe, A. D. Acharya, R. Panda, S. B. Shrivastava, M. Gangrade, T. Shripathi, and V. Ganesan, "Effect of copper doping on the change in the optical absorption behaviour in NiO thin films," *Renewable Energy*, vol. 46, 43–48, 2012.

[99] H. Kumagai, M. Matsumoto, K. Toyoda, and M. Obara, "Preparation and characteristics of nickel oxide thin film by controlled growth with sequential surface chemical reactions," *Journal of Materials Science Letters*, vol. 15, no. 12, 1081–1083, 1996.

[100] M. Utriainen, M. Kröger-Laukkanen, and L. Niinistö, "Studies of NiO thin film formation by atomic layer epitaxy," *Materials Science and Engineering: B*, vol. 54, no. 1-2, 98–103, 1998.

[101] T. S. Yang, W. Cho, M. Kim, K.-S. An, T.-M. Chung, C. G. Kim, and Y. Kim, "Atomic layer deposition of nickel oxide films using Ni(dmamp)2 and water," *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 23, no. 4, 1238, 2005.

[102] P. Antony Premkumar, M. Toeller, C. Adelmann, J. Meersschaut, A. Franquet, O. Richard, H. Tielens, B. Brijs, A. Moussa, T. Conard, H. Bender,

M. Schaekers, J. a. Kittl, M. Jurczak, and S. Van Elshocht, "NiO thin films synthesized by atomic layer deposition using Ni(dmamb)2 and ozone as precursors," *Chemical Vapor Deposition*, vol. 18, no. 1-3, 61–69, 2012.

[103] H. Li, D. Shenai, L. Zhefeng, and R. G. Gordon, "ALD and CVD Ni using Ni Amidinate Precursor," Dow Electronic Materials, , 2009.

[104] A. P. Grosvenor, M. C. Biesinger, R. S. C. Smart, and N. S. McIntyre, "New interpretations of XPS spectra of nickel metal and oxides," *Surface Science*, vol. 600, no. 9, 1771–1779, 2006.

[105] M. A. Van Veenendaal and G. A. Sawatzky, "Nonlocal screening effects in 2p x-ray photoemission spectroscopy core-level line shapes of transition metal compounds," *Physical Review Letters*, vol. 70, no. 16, 2459–2462, 1993.

[106] R. P. Gupta and S. K. Sen, "Calculation of multiplet structure of core p-vacancy levels," *Physical Review B*, vol. 12, no. 1, 15–19, 1975.

[107] S. C. Chen, T. Y. Kuo, Y. C. Lin, and H. C. Lin, "Preparation and properties of p-type transparent conductive Cu-doped NiO films," *Thin Solid Films*, vol. 519, no. 15, 4944–4947, 2011.

[108] W. Chen, Y. Wu, J. Fan, A. B. Djurišić, F. Liu, H. W. Tam, A. Ng, C. Surya, W. K. Chan, D. Wang, and Z. B. He, "Understanding the doping effect on NiO: Toward high-performance inverted perovskite solar cells," *Advanced Energy Materials*, vol. 8, no. 19, 1–10, 2018.

[109] J. Végh, "The Shirley background revised," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 151, no. 3, 159–164, 2006.

[110] R. Islam, G. Chen, P. Ramesh, J. Suh, N. Fuchigami, D. Lee, K. A. Littau, K. Weiner, R. T. Collins, and K. C. Saraswat, "Investigation of the changes in electronic properties of nickel oxide (NiOx) due to UV/ozone treatment," *ACS Applied Materials and Interfaces*, vol. 9, no. 20, 17 201–17 207, 2017.

[111] B. Streetman and S. Banerjee, *Solid State Electronic Devices*, 7th ed. Pearson, 2014.

[112] A. W. Dweydari and C. H. B. Mee, "Work function measurements on (100) and (110) surfaces of silver," *Physica Status Solidi (a)*, vol. 27, no. 1, 1973.

[113] J. Liu, H. Chen, E. Bi, X. Yang, Y. Yue, I. Ashraful, M. Gratzel, Y. Wu, W. Zhang, L. Han, and W. Chen, "Efficient and stable large-area perovskite solar cells with inorganic charge extraction layers," *Science*, vol. 350, no. 6263, 944–948, 2015.

[114] C. He, F. Zhang, M. Ye, C. Lin, and X. Zhao, "Interface engineering of BCP buffer layers in planar heterojunction perovskite solar cells with NiOx hole transporting layers," *Frontiers in Physics*, vol. 6, 99, 2018.

[115] J. Zheng, L. Hu, J. S. Yun, M. Zhang, C. F. J. Lau, J. Bing, X. Deng, Q. Ma, Y. Cho, W. Fu, C. Chen, M. A. Green, S. Huang, and A. W. Y. Ho-Baillie, "Solution-processed, silver-doped NiOx as hole transporting layer for high-efficiency inverted perovskite solar cells," *ACS Applied Energy Materials*, vol. 1, no. 2, 561–570, 2018.

[116] J.-Y. Jeng, K.-C. Chen, T.-Y. Chiang, P.-Y. Lin, T.-D. Tsai, Y.-C. Chang, T.-F. Guo, P. Chen, T.-C. Wen, and Y.-J. Hsu, "Nickel oxide electrode interlayer in CH3NH3PbI3 perovskite/PCBM planar-heterojunction hybrid solar cells," *Advanced Materials*, vol. 26, no. 24, 4107–4113, 2014.

[117] M. J. P. Alcocer, T. Leijtens, L. M. Herz, A. Petrozza, and H. J. Snaith, "Electron-Hole Diffusion Lengths Exceeding Trihalide Perovskite Absorber," *Science*, vol. 342, 341–344, 2013.

[118] W. Peng, L. Wang, B. Murali, K. T. Ho, A. Bera, N. Cho, C. F. Kang, V. M. Burlakov, J. Pan, L. Sinatra, C. Ma, W. Xu, D. Shi, E. Alarousu, A. Goriely, J. H. He, O. F. Mohammed, T. Wu, and O. M. Bakr, "Solution-grown monocrystalline hybrid perovskite films for hole-transporter-free solar cells," *Advanced Materials*, vol. 28, no. 17, 3383–3390, 2016.

[119] "Historical Trends in Federal R&D," American Association for the Advancement of Science, , 2018. Online: https://www.aaas.org/programs/r-d-budget-and-policy/historical-trends-federal-rd

[120] T. R. Behrens and D. O. Gray, "Unintended consequences of cooperative research: Impact of industry sponsorship on climate for academic freedom and other graduate student outcome," *Research Policy*, vol. 30, no. 2, 179–199, 2001.

[121] W. W. Powell, K. W. Koput, and L. Smith-doerr, "Interorganizational Collaboration and the Locus of Innovation : Networks of Learning in Biotechnology," *Administrative Science Quarterly*, vol. 41, no. 1, 116–145, 1996.

[122] R. Ponds, F. van Oort, and K. Frenken, "Innovation, spillovers and university-industry collaboration: An extended knowledge production function approach," *Journal of Economic Geography*, vol. 10, no. 2, 231–255, 2010.

[123] A. Banal-Estanol, M. Jofre-Bonet, and C. Meissner, "The impact of industry collaboration on research: evidence from engineering academics in the UK," 2010.

[124] D. Schartinger, C. Rammer, M. M. Fischer, and J. Fröhlich, "Knowledge interactions between universities and industry in Austria : sectoral patterns and determinants," *Research Policy*, vol. 31, 303–328, 2002.

[125] Y. S. Lee, "The Sustainability of University-Industry Research Collaboration: An emperical assessment," *Journal of Technology Transfer*, vol. 25, no. 2, 111–133, 2000.

[126] M. D. Santoro and A. K. Chakrabarti, "Firm size and technology centrality in industry-university interactions," *Research Policy*, vol. 31, no. 7, 1163–1180, 2002.

[127] P. D'Este and M. Perkmann, "Why do academics engage with industry? The entrepreneurial university and individual motivations," *Journal of Technology Transfer*, vol. 36, no. 3, 316–339, 2011.

[128] A. Wirsich, A. Kock, C. Strumann, and C. Schultz, "Effects of universityindustry collaboration on technological newness of firms," *Journal of Product Innovation Management*, vol. 33, no. 6, 708–725, 2016.

[129] A. Jaffe, "Real Effects of Academic Research," *The American Economic Review*, vol. 79, no. 5, 957–970, 1989.

[130] J.-y. Lin, "Balancing industry collaboration and academic innovation: The contingent role of collaboration-specific attributes," *Technological Forecasting & Social Change*, vol. 123, 216–228, 2017.

[131] R. Crescenzi, A. Filippetti, and S. Iammarino, "Academic inventors: collaboration and proximity with industry," *Journal of Technology Transfer*, vol. 42, no. 4, 730–762, 2017.

[132] M. Perkmann and K. Walsh, "The two faces of collaboration: Impacts of university-industry relations on public research," *Industrial and Corporate Change*, vol. 18, no. 6, 1033–1065, 2009.

[133] G. W. Prigge, "University-industry partnerships : what do they mean to universities? A review of the literature," *Industry and Higher Education*, vol. 19, no. 3, 221–229, 2005.

[134] J. B. Martin, "Academic-industrial collaboration: the good, the bad, and the ugly," *Transactions of the American Clinical and Climatological Association*, vol. 113, no. Figure 1, 227–240, 2002.

[135] D. Blumenthal, "Participation of life-science faculty in research relationships with industry," *Science*, vol. 335, no. 23, 1734–1739, 1996.

[136] D. Blumenthal, E. G. Campbell, M. S. Anderson, N. Causino, and K. S. Louis, "Withholding research results in academic life science," *Journal of the American Medical Association*, vol. 277, no. 15, 1224–1228, 1997.

[137] M. E. Gluck, D. Blumenthal, and M. A. Stoto, "University-industry relationships in the life sciences: Implications for students and post-doctoral fellows," *Research Policy*, vol. 16, no. 6, 327–336, 1987.

[138] D. C. Mowery, R. R. R. Nelson, B. N. Sampat, and A. A. Ziedonis, "The effects of the Bayh-Dole Act on U.S. university research and technology transfer: An analysis of data from Columbia University, the University of California, and Stanford University." *Research Policy*, vol. 30, no. 1, 99–119, 2001.

[139] D. C. Mowery, "The evolving structure of university-industry collaboration in the United States: Three cases," in *Research Teams and Partnerships: Trends in the Chemical Sciences: Report of a Workshop*. Washington (DC): National Academies Press, 1999, 7–20.

[140] B. H. Hall, "University-industry research partnerships in the United States," 2004.

[141] M. Perkmann and K. Walsh, "University-industry relationships and open innovation: Towards a research agenda," *International Journal of Management Reviews*, vol. 9, no. 4, 259–280, 2007.

[142] C. P. Ailes, I. Feller, and H. R. Coward, "The impact of engineering research centers on institutional and cultural change in participating universities," The National Science Foundation, , 2001.

[143] M. B. Albert, D. Avery, F. Narin, and P. McAllister, "Direct validation of citation counts as indicators of industrially important patents," *Research Policy*, vol. 20, no. 3, 251–259, 1991.

[144] C. L. Benson and C. L. Magee, "Quantitative determination of technological improvement from patent data," *PLoS ONE*, vol. 10, no. 4, 1–23, 2015.

[145] A. B. Jaffe and G. de Rassenfosse, "Patent Citation Data in Social Science Research: Overview and Best Practices," NBER, , 2016. Online: https://www.nber.org/papers/w21868

[146] P. Azoulay, W. Ding, and T. Stuart, "The determinants of faculty patenting behavior: Demographics or opportunities?" *Journal of Economic Behavior and Organization*, vol. 63, no. 4, 599–623, 2007.

[147] P. Azoulay, W. Ding, and T. E. Stuart, "The impact of academic patenting on (public) research output," *J. Indus. Econ.*, vol. 57, no. 4, 637, 2009.

[148] M. Verma and V. Varma, "Patent search using IPC classification vectors," in *Proceedings of the 4th workshop on patent information retrieval*, 2011.

[149] S. Arts, B. Cassiman, and J. C. Gomez, "Text matching to measure patent similarity," *Strategic Management Journal*, vol. 39, no. 1, 62–84, 2017.

[150] G. De Rassenfosse, H. Dernis, D. Guellec, L. Picci, and B. Van Pottelsberghe De La Potterie, "The worldwide count of priority patents: A new indicator of inventive activity," *Research Policy*, vol. 42, no. 3, 720–737, 2012.

[151] G. Morrison, M. Riccaboni, and F. Pammolli, "Disambiguation of patent inventors and assignees using high-resolution geolocation data," *Scientific Data*, vol. 4, 1–21, 2017.

[152] J. Schmid, "The diffusion of military technology," *Defence and Peace Economics*, vol. 29, no. 6, 595–613, 2017.

[153] D. K. Simonton, "Creative productivity: A predictive and explanatory model of career trajectories and landmarks," *Psychological Review*, vol. 104, no. 1, 66–89, 1997.

[154] Y. Gingras, V. Larivière, B. Macaluso, and J. P. Robitaille, "The effects of aging on researchers' publication and citation patterns," *PLoS ONE*, vol. 3, no. 12, 2008.

[155] D. Buck, K. Hartley, and N. Hooper, "Defence research and development, crowding out and the peace dividend," *Defence Economics*, vol. 4, no. 2, 161–178, 1993.

[156] W. N. Leonard, "Research and development in industrial growth," *Journal of Political Economy*, vol. 79, no. 2, 232–256, 1971.

[157] F. R. Lichtenberg, "The relationship between federal contract R&D and company R&D," *The American Economic Review*, vol. 74, no. 2, 73–78, 1984.

[158] F. R. Lichtenberg and D. Siegel, "The impact of R&D investment on productivity-new evidence using linked R&D-LRD data," *Economic Inquiry*, vol. 29, no. 2, 203–229, 1991.

[159] E. Morales-Ramos, "Defence R&D expenditure: The crowding-out hypothesis," *Defence and Peace Economics*, vol. 13, no. 5, 365–383, 2002.

[160] J. W. Ward, "Does defense R&D boost or bust innovation? An examination of defense R&D budget and patenting," Ph.D. dissertation, Georgetown University, 2008.

[161] R. Cowan and D. Foray, "Quandaries in the economics of dual technologies and spillovers from military to civilian research and development," *Research Policy*, vol. 24, no. 6, 851–868, 1995.

[162] V. Kistiakowsky, "Military Funding of University Research," *The Annals of the American Academy of Political and Social Science*, vol. 502, 141–154, 1989.

[163] J. Lerner, "The mobility of corporate scientists and engineers between civil and defense activities: Implications for economic competitiveness in the postcold war era," *Defence Economics*, vol. 3, no. 3, 229–242, 1992.

[164] S. Campbell, "Interregional migration of defense scientists and engineers to the Gunbelt during the 1980s," *Economic Geography*, vol. 69, no. 2, 204–223, 1993.

[165] J. Molas-Gallart, "Which way to go? Defence technology and the diversity of 'dual-use' technology transfer," *Research Policy*, vol. 26, no. 3, 367–385, 1997.

[166] A. B. Jaffe, M. S. Fogarty, and B. A. Banks, "Evidence from patents and patent citations on the impact of NASA and other federal labs on commercial innovation," *The Journal of Industrial Economics*, vol. 46, no. 2, 183–205, 2003.

[167] M. Acosta, D. Coronado, and R. Marín, "Potential dual-use of military technology: Does citing patents shed light on this process?" *Defence and Peace Economics*, vol. 22, no. 3, 335–349, 2011.

[168] "Historical Tables, Budget of the United States Government, Fiscal Year 2015," Government Publishing Office, , 2015. Online: www.govinfo.gov

[169] "Trend in spending on research and development of Raytheon from 2001 to 2017," 2019. Online: www.statista.com

[170] "Intel 2004 Annual Report," Intel Corporation, , 2004. Online: www.intel.com