Progress in Low-Cost Gallium Arsenide Solar Cells

Thesis by
Phillip Jahelka

In Partial Fulfillment of the Requirements for the
Degree of
Doctorate of Applied Physics

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2022
Defended 6 May 2022
ACKNOWLEDGEMENTS

Writing acknowledgements is tricky when one is uncertain about the metaphysical status of causality. It is hard to say, who, and to what extent, exactly, is responsible for me being able to share this work. What I do know, though, it that I have been unbelievably fortunate in receiving kindness and support from an enormous number of people over the years. There are certainly some people that I feel especially indebted to, but what I most clearly recognize is an unbroken chain of good-will being forged every day from my childhood to now. I most look forward to getting to be a part of this chain for other people so that they too may achieve ends beyond their wildest expectations.
Humanity has created, by degrees, a gordian knot of incentives that no one person or even country has the ability to cut through. It’s no one individual or country. It is a system. No one governs this system. It is governed by webs of incentives acting across individuals, nations, and corporations which reward and have normalized the very actions that will accelerate the process of climate destruction.

/u/TheBirminghamBear

Hello, babies. Welcome to Earth. It’s hot in the summer and cold in the winter. It’s round and wet and crowded. At the outside, babies, you’ve got about a hundred years here. There’s only one rule that I know of, babies—God damn it, you’ve got to be kind.

Kurt Vonnegut
In order to prevent disastrous global warming the manufacturing capacity of renewable energy power sources must grow rapidly. Solar photovoltaics will likely be one of humanity’s main sources of energy in the future due to the enormous available resource but increasing the manufacturing capacity of solar panels is hamstrung by both the limited profit margins of the highly-competitive renewable energy market and the enormous capital cost of building the factories that convert sand into semiconductor-grade silicon. Gallium arsenide is a material that can potentially help with the capital bottleneck because it absorbs light much more strongly than silicon and so the capital cost per unit weight of making the semiconductor can be spread over a larger number of devices and therefore effectively reduced. We present a number of results aimed at enabling low capital-cost GaAs solar cell manufacturing. First is a technique for open-tube, vapor phase zinc diffusion in GaAs. This method is dramatically simpler than its historical counterparts. Second, we use this technique to fabricate solar cells with $V_{oc}$’s greater than 960 mV and uncertified efficiencies over 23%, large improvements over the state of the art. We further demonstrate a base-metal, air-tolerant ohmic contact to n-type GaAs which is an improvement over traditional contacts that require noble metals and inert atmospheres. We also found the existence of melt-grown n-type GaAs with minority carrier diffusion length comparable to vapor grown material which helps with the economic viability of these devices. We also performed a techno-economic analysis on our proposed devices and find that they satisfy the desired properties of both the capital and electricity being cheaper than silicon solar cells. We also demonstrate the first n-on-p diffused junction GaAs solar cells.

As a parallel path to low capital intensity GaAs solar cells we also investigated non-epitaxial heterojunction devices. In the course of this work we both developed and characterized passivation chemistries for GaAs. Results in include the first use of a carbene and dithiothreitol for GaAs passivation and achieving surface recombination velocities comparable to GaInP passivation. With passivated organic heterojunction solar cells we were able to achieve a $V_{oc}$ of 840 mV which is a record for this class of devices, but its unclear how to improve the result to make them competitive with diffused junctions.
We also explored nanowire solar cells as an alternative strategy to reducing material usage by exploiting their strong light-absorption. We developed a computational model for a non-epitaxial GaAs heterojunction nanowire solar cell and predict an optimized efficiency over 30%. Towards fabrication we used metal-assisted-chemical-etching to make nanowire arrays and found we were able to cleanly cleave the nanowires embedded in a polymer from a 110 oriented wafer.

We also share some preliminary work on using total internal reflection in a solar cell encapsulant to mitigate shading loss due to the contacts on the front of a solar cell. We developed a computational model arguing that these structures could increase energy yield by 8% and demonstrated proof-of-principle experiments.

Finally, we share work on designing solar cells for operation on Venus. We developed models for the optical properties and recombination that correctly model the temperature dependence of a reference solar cell and using that model predict that a GaInP single-junction solar cell is a good solar-cell design for general usage in the atmosphere.
PUBLISHED CONTENT AND CONTRIBUTIONS

P.R.J. performed device physics simulations and validation and contributed to writing and editing the manuscript.

P.R.J. conceived of the project, performed luminescence and device measurements, fabricated devices, and wrote the manuscript.

P.R.J. conceived of the project, performed device physics simulations, and wrote the manuscript.

P.R.J. performed device physics simulations and validation and contributed to writing and editing the manuscript.

P.R.J. conceived of the project, performed optical simulations, performed proof-of-principle experiments, and wrote the manuscript.
# TABLE OF CONTENTS

Acknowledgements .................................................. iii
Abstract ............................................................... v
Published Content and Contributions ............................ vii
Table of Contents ................................................... vii
List of Illustrations ................................................ xi
List of Tables ........................................................ xxi
Chapter I: Introduction ............................................. 1
  1.1 Brief Overview of the Thesis ................................. 1
  1.2 Scaling Photovoltaics ........................................ 2
      How Much Electricity, and How Quickly .................... 6
      Only Solar Can Meet the Demand ........................... 6
      Climate-Motivated Growth Models of Photovoltaic Manufacturing 8
      Economics of Photovoltaic Manufacturing Growth .......... 12
      Opportunities for Photovoltaic Research .................. 16
Chapter II: Diffusion Doped GaAs Solar Cells ............... 18
  2.1 Introduction .................................................. 18
  2.2 Historical and Recent Context on GaAs Solar Cells .......... 18
  2.3 Deducing GaAs Solar Cell Design and Fabrication .......... 22
  2.4 Device Performance Prediction .............................. 26
      Detailed Balance Model ..................................... 26
      Drift-Diffusion Model ...................................... 30
  2.5 Low Temperature GaAs Zinc Diffusion Doping ............. 33
      Practical GaAs Zinc Diffusion ............................. 39
  2.6 Zinc Diffusion Doped GaAs Solar Cells .................... 52
      Minimal Cell Fabrication ................................. 53
      Optimizing Fabrication .................................... 56
  2.7 Air Tolerant, Abundant Metal Ohmic Contacts ............. 76
  2.8 Diffusion Doped GaAs Technoeconomic Analysis ............ 79
      GaAs Ingot Synthesis ....................................... 81
      Ingot Processing ........................................... 82
      Diffusion Doping ........................................... 82
      Spalling, Lamination, and Chrome Removal ................. 83
      ARC, Laser Opening, Metal Printing, and Testing .......... 84
      TEA Commentary ............................................. 84
  2.9 $n^+/p$ Homojunction GaAs Solar Cells ................... 88
      Liquid Phase Epitaxy ....................................... 88
      Monolayer Dopant Diffusion ................................ 88
      Spin on Glass Diffusion Doping ............................ 89
  2.10 Closing Thoughts ........................................... 91
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solar Wind</td>
<td>213</td>
</tr>
<tr>
<td>Geothermal Heating</td>
<td>213</td>
</tr>
<tr>
<td>Overall Power Inflow</td>
<td>213</td>
</tr>
<tr>
<td>B.3 A Simple Model</td>
<td>213</td>
</tr>
<tr>
<td>B.4 More Complicated Model</td>
<td>218</td>
</tr>
<tr>
<td>B.5 Conclusion</td>
<td>219</td>
</tr>
<tr>
<td>B.6 Dediscretizing Second Climate Model</td>
<td>221</td>
</tr>
<tr>
<td>B.7 Directional vs Isotropic Optical Absorption</td>
<td>223</td>
</tr>
<tr>
<td>Number</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.1</td>
<td>How various risks of global warming depend on temperature.</td>
</tr>
<tr>
<td>1.2</td>
<td>Development pathways compatible with 1.5 °C provided by the IPCC. We reproduce the figure in full so readers can understand what, exactly, policy makers are exposed to.</td>
</tr>
<tr>
<td>1.3</td>
<td>Map of solar resource in the US.</td>
</tr>
<tr>
<td>1.4</td>
<td>Illustration of manufacturing capacity over time for photovoltaics that meets energy demands that either does or does not incorporate sustainable demand consideration.</td>
</tr>
<tr>
<td>1.5</td>
<td>Maximum sustainable growth rate as a function of capital intensity and operating margin. 1.19 on the horizontal axis corresponds to the silicon PV supply chain. Reproduced from [200] under license CC BY 3.0.</td>
</tr>
<tr>
<td>1.6</td>
<td>Operating margin of various publicly traded PV companies. Notably these companies do not reliably make money.</td>
</tr>
<tr>
<td>1.7</td>
<td>Capital expenditure by fabrication step for silicon PV and grouped by broad process category. Reproduced from [200] under license CC BY 3.0.</td>
</tr>
<tr>
<td>2.1</td>
<td>Epitaxial Liftoff Process for thin-film III-V solar cell fabrication. The grown solar cell includes a sacrificial AlAs layer that rapidly etches in HF to free the cell from the substrate. Reproduced from [99]</td>
</tr>
<tr>
<td>2.2</td>
<td>Illustration of the spalling process.</td>
</tr>
<tr>
<td>2.3</td>
<td>Cost overview for modern silicon solar cells. [185]</td>
</tr>
<tr>
<td>2.4</td>
<td>Proposed low-cost, diffused-junction GaAs solar cell. Cheap tinned-plated steel is the back contact with a GaAs absorber formed by spalling and p-n junction made by diffusion doping with simple evaporated anti-reflection coating and screen printed silver top contacts.</td>
</tr>
<tr>
<td>2.5</td>
<td>Proposed process for industrial cell fabrication.</td>
</tr>
</tbody>
</table>
2.6 Optical properties of GaAs versus doping. a) Absorption coefficients of n and p-type GaAs as a function of doping. b) Absorption coefficient of GaAs multiplied by the blackbody spectrum to estimate luminescence spectrum. 28

2.7 Specifications for the detailed balance model. a) Optical model for cell efficiency calculation. b) Electrical model for cell efficiency calculation. 28

2.8 Cell performance metrics as a function of Urbach energy. Performance begins degrading at 40 meV. 31

2.9 Electrical properties of heavily doped p-type GaAs. a) Minority electron lifetime and diffusion length as a function of doping and minority electron diffusion coefficient and intrinsic carrier concentration enhancement factor as a function of doping. b) Calculated electron contribution to the diode diffusion current. c) Calculated PV performance as a function of doping using the diffusion-current diode model. 33

2.10 Cell efficiency as a function of doping predicted by either the detailed balance or electrical diode models. Performance switches from electrically limited to optically limited around $5 \times 10^{19}$ cm$^{-3}$ doping. 34

2.11 Vapor pressures of the metal and group-V elements above GaAs and InP. The horizontal line is the vapor pressure for maintaining satisfactory surfaces for 30 min. 35

2.12 Ga-As-Zn ternary phase diagram at 700 °C. Zinc arsenides must be used as diffusion sources to prevent liquid formation. Reprinted/adapted by permission from Springer Nature, Diffusion in the III–V Compound Semiconductors by H. C. Casey Jr., 1973. 37

2.13 Predicted solubility of zinc in GaAs as a function of temperature and zinc activity. 39

2.14 The Zn/Sn diffusion source. a) Schematics of the Zn/Sn diffusion source. b) Picture of the diffusion source with various elements highlighted with corresponding portions from a. 41

2.15 Box and whisker plots for effect of post-diffusion annealing on Hall measurement results. Annealing reduces sheet resistance and mobility while increasing carrier concentration. 46
2.16 Illustration of improved diffusion source that will allow improved uniformity and consistency by better controlling the atmosphere around the sample. ................................................. 47
2.17 Picture of brass sheet with GaAs and carbon paper spacer (center) for zinc diffusion doping. ................................. 48
2.18 SIMS measurement of zinc and copper in GaAs doped with the brass source. ......................................................... 49
2.19 Required time for ZnO reduction under hydrogen to effectively dope GaAs as a function of temperature. ..................... 50
2.20 Picture of setup used for making ZnO films on a hotplate. ........ 51
2.21 Picture of ZnO film after being used for GaAs diffusion doping. The regions covered by GaAs are clearly visible, demonstrating the considerable decomposition of the film during the process. 52
2.22 Schematic of fabrication process used for making the prototyped solar cells. .......................................................... 53
2.23 False color picture of solar cell with mesa etch and metal finished. 55
2.24 Comparison of cell performance depending on if the cells were singulated by cleaving or mesa etching. Cleaving badly reduces cell voltage. .............................................................. 56
2.25 $V_{oc}$ as a function of edge length if all recombination is due to the edge. ............................................................... 57
2.26 Electrical model for calculating cell size and metal dimensions for minimizing ohmic conduction losses. ......................... 58
2.27 Calculated amount of collected photocurrent from the bulk of the wafer as a function of the hole diffusion length. ............... 61
2.28 IQE and model fits for solar cells made using commercially available n-GaAs wafers by either the ZnO or Zn/Sn diffusion methods. The extracted hole diffusion length is 1 µm. ....................... 62
2.29 Schematic of vapor controlled Czochralski growth tool for making high-quality ingots of n-type GaAs. The lack of an encapsulant above the melt allows dynamically controlling the arsenic vapor pressure via the separate heated arsenic area. Reprinted/adapted by permission from Elsevier, Growth of GaAs crystals from Ga-rich melts by the VCz method without liquid encapsulation by F. M. Kiessling, P. Rudolph, M. Neubert, U. Juda, M. Naumann, W. Ulrici, 2004. ............................... 63
2.30 IQEs and model fits for cells made using either GaAs grown by HVPE or GaAs grown by vCZ with a low arsenic temperature. The extracted hole diffusion-lengths are far-superior to the commercial bulk material. .............................................. 64
2.31 Predicted $J_{sc}$ as a function of emitter thickness and dimensionless surface recombination parameter. ................................. 66
2.32 Cell performance metrics as a function of emitter thinning etch time. ............................................................. 67
2.33 IQE of solar cell before and after passivation demonstrating enhances short-wavelength response. ................................. 68
2.34 Dark IV curve of heavily doped solar cell exhibiting unusual ideality factor behavior as a function of bias. ................................. 70
2.35 Scatter plots correlating cell $V_{oc}$ and Hall measurement results taken after diffusion doping. ............................................... 71
2.36 Light and dark IV and ideality factor as a function of bias for the champion cell. ............................................................. 72
2.37 Measured quantum efficiency of the champion cell before and after ARC deposition. ............................................................. 73
2.38 IV curve of cell made with ZnCl$_2$, AlCl$_3$, NaCl diffusion source exhibiting low-voltage and unusual intermediate ideality factor. 75
2.39 Comparison of EQEs of possible AlGaAs cell and Zn/Sn cell doped at the same temperature and for the same duration. The cell with AlCl$_3$ has much better blue-response. ................................. 75
2.40 Illustration of solid-phase regrowth method for making ohmic contacts. IV is usually Si or Ge and M is typically Pd or Ni. Left to right is increasing temperature and time. Reprinted from [213], with permission from Elsevier ................................. 77
2.41 Tin solder ohmic contacts to n-type GaAs. a) Picture of GaAs chip soldered to copper foil with metallic tin and SnCl$_2$ flux. Also visible are gold Schottky contacts. b) IV curve of a Schottky contact demonstrating exponential diode behavior to over one-sun current density. .................................................. 78
2.42 Processing cost by step and cost category for thin-film GaAs solar cell fabrication. Costs are dominated by the gallium and silver metal. ............................................................. 86
2.43 Attempts at Sn-based LPE of GaAs. a) Illustration of LPE method for GaAs doping where Sn dissolved some GaAs and then redeposits it as heavily n-type. b) Microscope image of GaAs surface after tin removal showing extremely rough surface with crystallographic regrowth features.

2.44 Defects on GaAs after Sn diffusion from spin-on-glass. Left) Wide-area image of GaAs chip post-Sn doped glass diffusion showing defects. Right) Detailed image of defect cluster with striking large-scale symmetry and rectangular defect pits.

2.45 Best cell performance with tin-doped spin on glass diffusion source.

3.1 Residue from trying to passivate GaAs with functionalized thiols. a) From left to right, pictures of wax in jar after trying to passivate GaAs with functionalized thiols, a polymer film that grew on GaAs during functionalized thiol passivation, cloudy residue in passivation solution even when run under a nitrogen purge. b) Putative mechanism for base-catalyzed oxidative polymerization of organic thiols.

3.2 Oxidation scheme of DTT where it undergoes a ring-closing instead of polymerization reaction.

3.3 Illustration of 1,3-Diisopropylimidazolium. Note the electron pair on the bottom carbon.

3.4 AFM scan of GaAs etched by NHC solution.

3.5 Illustration of TRPL setup.

3.6 TRPL data for the same piece of GaAs, a) taken with pulse rate of 10 MHz, b) taken with a pulse rate of 500 kHz.

3.7 Collation of TRPL data as a function of surface treatment and hotplate annealing temperature.

3.8 Extracted SRV as a function of surface passivation chemistry and annealing temperature.

3.9 Overlaid Ga 2p XPS spectra for each passivation chemistry.

3.10 Ga 2p XPS spectra for each passivant decomposed into peaks with chemical identification.

3.11 Calculated structures of DTT (left) and NHC (right) on GaAs.

3.12 Non-epitaxial GaAs heterojunction solar cell design. a) General schematic of non-epitaxial heterojunction GaAs solar cell. b) Notional band diagram color coded to a.
3.13 IV curves for the highest $V_{oc}$ organic heterojunction solar cell. From left to right, illuminated, dark, and dark diode ideality factor. ................................................................. 112

3.14 IV curves for a lower $V_{oc}$ organic heterojunction solar cell that best demonstrates unusual ideality factor behavior. ................. 112

3.15 GaAs Schottky barrier solar cell performance. a) $V_{oc}$ of a GaAs Schottky solar cell as a function of barrier height and electron recombination velocity. b) Same as a except the difference between the electron and hole Fermi levels at the back contact. ....... 115

3.16 Band diagrams of GaAs Schottky solar cells for two different barrier heights under illumination and at $V_{oc}$ .............. 116

3.17 Solar-Cell $V_{oc}$ dependence on TCTA thickness. ................. 117

3.18 Collation of GaAs/Organic heterojunction solar cell IV curves. 118

3.19 IV curves of devices using doped PTAA as the hole transport layer as a function of passivation. ......................... 121

4.1 Schematics of simulated solar cells. a) Schematics of simulated thin-film solar cells. Left is thin-film cell with doped regions for carrier collection. Left is heterojunction cell with SnO$_2$ and CuSCN for carrier collection. b) Schematics of simulated nanowire solar cells. Left is doped nanowire solar cells. Right is heterojunction nanowire solar cell with variable thickness dielectric separating the SnO$_2$ and CuSCN. .............................. 126

4.2 Optical constants of SnO$_2$ and CuSCN measured by spectroscopic ellipsometry. .......................................................... 127

4.3 Bandstructure of 1$\mu$m thick thin-film device at short-circuit. Note the change in slope of the minority carrier electrochemical potentials midway in the quasi-neutral regions indicating their traveling in the wrong direction. .................................................. 129

4.4 Total minority carrier recombination currents at the contacts for the 1$\mu$m and 20$\mu$m thick homojunction devices. The thicker device suppresses the currents and illumination increases them. 131

4.5 Performance of nanowire homojunction as function of design parameters. The optimal tip length is 50nm and efficiencies are limited to 13%. ................................................................. 133
4.6 Current at each electrode versus bias for champion homojunction solar cell. The large electron current at the anode limits cell performance.

4.7 Plot of homojunction solar cell band structure near the contacts demonstrating large minority carrier concentrations near the contacts.

4.8 Hole current density and electric field near the cathode. Horizontal line at 1.45 µm denotes metallurgical junction. The hole current is in the correct direction up to 20 nm from the contact, which corresponds to where the electric field induced by the p-n junction weakens.

4.9 Current at each electrode versus bias for nanowire cell with planar top and bottom heterojunctions. The anode electron current is two orders of magnitude smaller and roughly constant. I-V characteristics are dominated by majority carrier currents. Cathode electron current obscures the anode hole current.

4.10 Plot of heterojunction solar cell band structure near the contacts demonstrating low minority carrier concentrations near the contacts.

4.11 Heterojunction solar cell performance dependence on dielectric infill thickness. Performance monotonically decreases with infill thickness.

4.12 Performance of nanowire heterojunction as function of design parameters. P-type devices perform better and efficiencies greater than 30% are possible for defect concentrations less than \( 1 \times 10^{12} \text{ cm}^{-2} \text{ eV}^{-1} \).

4.13 Illustrates of MACE process on silicon. a) cross section showing chemical reactions and geometry. b) Interfacial band diagram illustrating why substrate doping and Schottky barrier height matter. Reprinted (adapted) with permission from [132]. Copyright 2016 American Chemical Society.

4.14 Left, gold layer with holes defined by wet-etching. Right, resulting GaAs nanowires showing the rough edges of the gold mask badly deforming the GaAs.

4.15 Different orientations of GaAs MAC-etched for 30 min.
4.16 Cross section of GaAs etched with cold citric acid and hydrogen peroxide showing no undercutting and steep sidewalls. . . . . . 145

4.17 Cross section of GaAs etched with cold citric acid and hydrogen peroxide showing no undercutting and steep sidewalls. . . . . . 146

5.1 Illustration of how low-refractive index regions can improve photocurrent of solar cells. (a) Demonstration of why metal contacts reduce photocurrent. Light incident on the solar cell reflects off the cell’s front metal contact, reducing the photocurrent of the cell. (b) Example of how lower refractive index regions can improve photocurrent. A low refractive index triangle is placed above the metal contact. Assuming a favorable relationship between the angle of incidence, triangle geometry, and index contrast, the light reflects off the encapsulant/triangle interface and into the solar cell. . . . . . . . . . . . . . . . . . . . . . 150

5.2 Exploration of angle of incidence dependence of photocurrent change as a function of triangle geometry. (a) Schematic illustrating the simulated geometry and the scheme for encoding the angle of incidence. (b) Each subfigure illustrates the angular response of triangles with different geometries. The geometry is indicated in the lower left of each subfigure, the color represents the photocurrent change, and the axes are the angles of incidence defined in part a. . . . . . . . . . . . . . . . . . . . . . 156

5.3 Examining different structures and their performance as a function of angle of incidence. (a) Two possible non-triangular structures for the low-index region above the contact. The upper structure is a sawtooth, and the other is a ‘house.’ The sawtooth is 2.5mm tall and the walls and roof of the house are each 2mm tall. (b) The relative change in photocurrent due to the structures as a function of angle of incidence. . . . . . . . . 157
5.4 Experimental demonstration of functioning effectively transparent contacts for a solar cell busbar. (a) Photograph of solar cell busbar with triangular void in PDMS above the busbar. (b) Reflection microscopy image of a busbar with a triangular void above it and a finger contact without a triangle. (c) Spatially resolved laser beam induced current measurement of a busbar with a triangular void above it and a finger contact without a triangle. The dashed line is the approximate location of the busbar. The color indicates the measured photocurrent, scale is arbitrary units. (d,e,f) Images analogous to (a,b,c) except the PDMS layer is planar.

6.1 Solar spectrum on Venus as a function of altitude. a) Spectra of downward incident light measured by the Venera 13 probe as a function of altitude in km with CO\textsubscript{2} and H\textsubscript{2}O absorption lines highlighted. b) Total downward flux measurements. Curve 1 was measured by Pioneer-Venus, 2 by Venera 11, 13, and 14. 3 and 4 are global mean net fluxes measured by Pioneer-Venus and the Venus International Reference Atmosphere, respectively.

6.2 Simulated and experimental cell performance as a function of temperature. a) Comparison of GaAs band-edge EQE at 300°C for two experimental solar cells and simulations that do or do not include Urbach tail absorption. b) Validation of recombination model predicting V\textsubscript{oc} as a function of temperature.

6.3 Solar cell efficiency as a function of altitude and GaInP top cell thickness. The optimal GaInP thickness is plotted as a line following the maximum power at each altitude.

6.4 Summary of optimized solar cell design and performance as a function of altitude. a-c) Solar cell designs optimized for operation at high altitude, mid altitude, and low-altitude, respectively. d) Electrical power generated by each cell design as a function of altitude. e) Efficiency of each cell design as a function of altitude.

A.1 Sentaurus tool flow for a coupled FDTD/Sentaurus nanowire solar cell simulation.

B.1 Power Sources for the Earth including solar light and wind, thermal energy from the Earth’s core, radio from the cosmic microwave background (CMB), and starlight.
B.2 Normalized blackbody spectra for 300 K and 5800 K. What’s important is that they peak at very different wavelengths, with the 5800 K spectra highest at 0.5 µm, and the 300 K highest at 10 µm.

B.3 Power flows in the Ground/Atmosphere/Space system. a. Visible light power flows. b. IR light power flows. c. Power flows as a graph.

B.4 Ground and air temperature vs IR absorption and GHG concentration of the atmosphere.

B.5 Sketch of model that allows radiative coupling between layers of the atmosphere.

B.6 Ground temperature vs IR absorption and GHG concentration of the atmosphere.

B.7 Power being radiated to space as a function of altitude layer index.
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Number</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Urbach energy of GaAs as a function of doping extracted from literature absorption coefficient measurements.</td>
<td>27</td>
</tr>
<tr>
<td>2.2 Cell performance as a function of doping for detailed balance model that includes subgap absorption.</td>
<td>30</td>
</tr>
<tr>
<td>2.3 Zn/Sn Diffusion Source Hall Measurements</td>
<td>44</td>
</tr>
<tr>
<td>2.4 Effect of Annealing Temperature on Electrical Properties</td>
<td>44</td>
</tr>
<tr>
<td>2.5 Effect of Annealing Time on Electrical Properties</td>
<td>45</td>
</tr>
<tr>
<td>2.6 Device parameters extracted from IQE measurements including the emitter thickness, minority electron diffusion length, dimensionless surface recombination parameter and emitter sheet resistance.</td>
<td>65</td>
</tr>
<tr>
<td>2.7 Summary of cell performance for as-diffused cell, after emitter thinning, and finally after surface passivation.</td>
<td>68</td>
</tr>
<tr>
<td>2.8 vCZ GaAs Fixed Costs</td>
<td>81</td>
</tr>
<tr>
<td>2.9 vCZ GaAs Growth Variable Costs.</td>
<td>82</td>
</tr>
<tr>
<td>2.10 Itemized costs in proposed GaAs thin-film solar cell fabrication process.</td>
<td>85</td>
</tr>
<tr>
<td>2.11 Tin Doped Glass Hall Measurements</td>
<td>91</td>
</tr>
<tr>
<td>3.1 Energy splitting between Ga 2p GaAs peak and the Ga and additional surface peak.</td>
<td>106</td>
</tr>
<tr>
<td>4.1 Parameters used in device physics simulations. $\chi_e$; $\phi$: electron affinity or work function, eV. $E_g$: bandgap, eV. $\mu_e/g$: mobilities, cm$^2$V$^{-1}$s$^{-1}$. $m_e^*$: effective masses. $\tau_{e/h}$: lifetimes, ns. $\epsilon$: dielectric permittivity.</td>
<td>128</td>
</tr>
<tr>
<td>5.1 Simulated annual improvements in annual energy harvest for a solar panel equipped with perfect effectively transparent contacts</td>
<td>153</td>
</tr>
<tr>
<td>6.1 Device stack for comparing simulated and measured EQE measurements.</td>
<td>161</td>
</tr>
</tbody>
</table>
Chapter 1

INTRODUCTION

1.1 Brief Overview of the Thesis
The structure of the thesis is as follows. First, in this introduction, I will describe some of the environmental and economic factors constraining renewable energy. The progression is:

- global warming and human energy needs place demands on renewable energy generation
- geophysics and politics dictate solar must be the main renewable energy source of the future
- laissez-faire handling of current photovoltaic manufacturing is insufficient to meet the demands of global warming
- we need new photovoltaic technologies and policy interventions to handle global warming
- GaAs photovoltaics are a potentially useful technology towards this end.

I am taking a detailed, systems-oriented approach so that I can forcefully argue what, exactly, are the outstanding problems in photovoltaics qua renewable energy, and what fruitful contributions to these problems must look like. With an understanding of the central problems, I can then describe why low-cost GaAs solar cells are worth researching and share my progress in fabricating them, which is Chapter One. This will include new junction formation processes, discoveries about GaAs wafer fabrication, metallization, and cost modeling. Chapter Two will include my work on surface passivation and non-epitaxial heterojunction GaAs solar cells. The surface passivation is directly relevant to the first chapter and overall goal of the work, but the heterojunction work turned out to be a dead end. Chapter Three will describe my work on nanowire solar cell modeling and fabrication. I will describe why I thought this was a good idea at the time, and why I was wrong. Chapter Four is my work on effectively transparent contacts which is conceptually interesting, but
has large problems in implementation. Finally, Chapter Five is my work on designing solar cells for operating on Venus. While not related to my overall motivation, it is a) cool, and b) depends on physics germane to my main work such as Urbach tails. This introduction and Chapter One are, in my opinion, the most important and interesting sections.

There are also two appendices. First is code for importing semiconductor optical generation rates calculated by Lumerical FDTD into TCAD Sentaurus which allows premier optical and device solvers to leverage each other’s power. Second is a monograph I wrote that explains how global warming works at a level between ‘CO$_2$ is like a blanket’ and global circulation models. I wrote this because I figured that if I was going to try to convince people that we need to spend trillions of dollars to fix global warming, I had better convince myself global warming is real and understand some of how it works. Multiple graduate students and even a professor have found this useful and interesting.

1.2 Scaling Photovoltaics
There is an urgent need to limit the amount of greenhouse gases that we emit into the atmosphere in order to limit further global warming. Many countries acknowledge this and have, non-bindingly, agreed to try to limit global warming to 1.5$^\circ$C. The International Panel on Climate Change Special Report on 1.5$^\circ$C lays out numerous reasons for wanting to limit global warming to 1.5$^\circ$C [160]. Several examples include:

- Limiting warming to 1.5$^\circ$C instead of 2$^\circ$C results in 10.4 million fewer people being impacted by sea level rise. Over 100 million people are expected to be affected regardless.

- Above 1.5$^\circ$C of warming, the Sahara penetrates into the Mediterranean, “causing changes unparalleled in the past 10,000 years.”

- Even at 1.5$^\circ$C of warming, 70-90% of warm water coral will disappear which currently provide protein for about 500 million people. At 2$^\circ$C of warming, it is completely gone [35].

- Warming decreases the yield of vital cereal crops, with reductions of 6, 3.2, 7.4, and 3.1% per degree Celsius for wheat, rice, maize, and soybeans, respectively.
How the level of global warming affects impacts and/or risks associated with the Reasons for Concern (RFCs) and selected natural, managed and human systems

Five Reasons For Concern (RFCs) illustrate the impacts and risks of different levels of global warming for people, economies and ecosystems across sectors and regions.

- Beyond these concrete points, 1.5°C is also a tipping point where many impacts see a qualitative increase in severity or risk. This is encapsulated in Figure SPM.2 of the IPCC report, which is reproduced in Fig. 1.1. Particularly interesting is the increase in risk of ‘Large Scale Singular Events’ such as ice sheets collapsing or the Atlantic Meridional Overturning Current (AMOC) shutting down.

The emissions goals for limiting warming to 1.5°C are expressed as a budget for how much more CO₂ we can emit, and a time by which emissions must reach net zero. The IPCC in its 1.5°C warming report estimates that we can emit about 400 additional gigatons of CO₂, which is equivalent to 8 years at
current emission levels, and must reach net-zero emissions by 2050. After 2050, we will have to achieve net-negative emissions in order to stabilize the climate. A limitation acknowledged in the report is that it does not include positive climate feedback loops, chiefly methane released by melting permafrost. Recent work since the report suggests that permafrost melting may reduce the anthropogenic carbon budget to near-zero [69, 179].

The IPCC report also includes pathways to achieving climate goals which specify how much various indicators such as final energy consumption and energy from various sources will have to change by 2050. The pathways are reproduced in Fig. 1.2. The EIA predicts the world energy consumption will increase 50% between 2018 and 2050, suggesting that scenarios P3 and P4 are the more realistic in terms of energy demand. Both of these scenarios demand bold changes in how humanity produces energy. As one example, the scenarios require a five-fold increase in nuclear capacity when even the International Atomic Energy Agency is in its best-case scenario predicting only a doubling. The models also rely on large-scale deployment of Bio Energy with Carbon Capture and Sequestration (BECCS). BECCS plants use some of the energy they generate from burning grown fuel to capture and sequester the generated CO$_2$ to achieve net-negative emissions. Unfortunately, recent work on BECCS limits their sequestering capacity to 7.5 Gt of CO$_2$ per year even with converting pasture to fuel production due to limited ability to irrigate the fuel plants [5]. This is in tension with the P3 and P4 demands of about 15 Gt yr$^{-1}$. In terms of land usage, the Grantham Institute at Imperial College London estimates that deploying BECCS at a scale to limit warming to 1.5°C consumes 25 to 80% of current global cropland [21].

Motivating why we need new photovoltaic technology and why GaAs thin-films are a good candidate requires a multidisciplinary argument spanning climate science, economics, and device physics. Schematically, the argument goes:

1. Determining how much electrical generation capacity we need, and on what timescale.

2. Arguing solar is the only energy source that can meet this demand because other power sources have insufficient resources available.

3. Calculating how quickly silicon solar cell manufacturing must increase to hit the energy targets from the first point.
Characteristics of four illustrative model pathways

Different mitigation strategies can achieve the net emissions reductions that would be required to follow a pathway that limits global warming to 1.5°C with no or limited overshoot. All pathways use Carbon Dioxide Removal (CDR), but the amount varies across pathways, as do the relative contributions of Bioenergy with Carbon Capture and Storage (BECCS) and removals in the Agriculture, Forestry and Other Land Use (AFOLU) sector. This has implications for emissions and several other pathway characteristics.

Breakdown of contributions to global net CO₂ emissions in four illustrative model pathways

Figure 1.2: Development pathways compatible with 1.5°C provided by the IPCC. We reproduce the figure in full so readers can understand what, exactly, policy makers are exposed to.

### Global indicators

<table>
<thead>
<tr>
<th>Pathway classification</th>
<th>P1: No or limited overshoot</th>
<th>P2: No or limited overshoot</th>
<th>P3: No or limited overshoot</th>
<th>P4: No or limited overshoot</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO₂ emissions change in 2030 (% relative to 2010)</td>
<td>58</td>
<td>47</td>
<td>41</td>
<td>4</td>
</tr>
<tr>
<td>Primary energy from coal in 2030 (% relative to 2010)</td>
<td>78</td>
<td>81</td>
<td>63</td>
<td>70</td>
</tr>
<tr>
<td>Primary energy from nuclear in 2030 (% relative to 2010)</td>
<td>102</td>
<td>93</td>
<td>86</td>
<td>86</td>
</tr>
<tr>
<td>Forest biomass in 2030 (% relative to 2010)</td>
<td>150</td>
<td>142</td>
<td>131</td>
<td>122</td>
</tr>
<tr>
<td>Biomass to biofuels in 2030 (% relative to 2010)</td>
<td>51</td>
<td>49</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>Cumulative CCS until 2030 (GtCO₂)</td>
<td>131</td>
<td>127</td>
<td>113</td>
<td>117</td>
</tr>
<tr>
<td>Forest biomass in 2050 (% relative to 2010)</td>
<td>130</td>
<td>125</td>
<td>116</td>
<td>116</td>
</tr>
<tr>
<td>Land area of bioenergy crops in 2030 (million km²)</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Agricultural NOₓ emissions in 2030 (% relative to 2010)</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Agricultural NOₓ emissions in 2050 (% relative to 2010)</td>
<td>5</td>
<td>26</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>Nitrous oxide emissions in 2030 (% relative to 2010)</td>
<td>6</td>
<td>26</td>
<td>39</td>
<td>39</td>
</tr>
</tbody>
</table>

**NOTE:** Indicators have been selected to show global trends identified by the Chapter 2 assessment. National and sectoral characteristics can differ substantially from the global trends shown above.

* CO₂ emissions are based on IPCC Second Assessment Report (GWP-34).
* Changes in energy demand are associated with improvements in energy efficiency and behaviour change.
4. Showing that the needed growth rate is not possible within a classical economic framework.

5. Discussing the desirable properties of scalable photovoltaic technologies and arguing that III-V thin-films are a good alternative technology.

**How Much Electricity, and How Quickly**

To set the time scale for how quickly photovoltaic manufacturing capacity needs to grow, we can use the net-zero dates that limit global warming to 1.5 or 2 °C. According to the IPCC special report on 1.5 °C of warming, limiting warming to 1.5 and 2 °C requires reaching net-zero CO$_2$ by 2050 and 2075, respectively, leaving 30 years and 55 years to reach these goals [141]. We can also estimate a lower bound on how much energy carbon-neutral sources will have to be able to provide at those times. The estimate is a lower bound because additional energy will have to be expended to sequester CO$_2$ from hard-to-decarbonize sectors of the economy such as cement, heavy transport, and steel smelting. According to the International Energy Agency and the UN Development Program, total useful energy consumption in 2015 or 2018 was $1.6 \times 10^{20}$ J [283, 106]. Further, the US Energy Information Administration predicts an energy demand growth of 3.1% per year, so that total useful energy demand will be around $4.2 \times 10^{20}$ J and $9.1 \times 10^{20}$ J in 2050 and 2075, respectively [51].

**Only Solar Can Meet the Demand**

We can also roughly estimate how much of the world’s energy must be produced by solar as compared to hydroelectric, wind, biomass, and nuclear. Starting with hydro in the US, a pair of assessments from Oak Ridge National Lab estimates that there is 12 GW and 85 GW of electricity possible from retrofitting non-power-generating dams and new construction, respectively [9, 115]. Along with existing production of 291 TWh, this can meet about 6% of the US’s 2020 useful energy demand of $7 \times 10^{19}$ J yr$^{-1}$ [264, 277]. Globally, there is up to $1.2 \times 10^{19}$ J yr$^{-1}$ of ecologically sustainable hydroelectric generation available at a price below $0.1/(kWh)$, which comes out to about 3% of humanity’s projected 2050 energy needs [71].

---

1 The data used from the IEA in this claim is no longer available because the IEA has locked it behind a paywall since I visited. Considering the importance of this data, the IEA’s position is controversial [205].
Regarding wind, recent theoretical work in understanding how large windfarms couple to mesoscale atmospheric dynamics suggests large wind farms are limited in energy production to $1 \text{ W/m}^2$ [2, 169]. At this scale, wind power has substantial effects on the local climate. For example, harvesting $1.5 \times 10^{19} \text{ J}$ of wind power in the Midwestern US, some of the best wind resources in the world, is predicted to raise nighttime temperatures in the wind farm area by over $1^\circ \text{C}$ [168, 75]. On a more global scale, a recent NREL assessment indicates a correction for the limited power density of wind where the yearly wind resource is $4 \times 10^{20} \text{ J}$ and 90% of the installed capacity is low to mid quality [55]. This means that powering the world with wind in 2050 requires harvesting ALL available wind energy. Based on the climate impacts of harvesting much smaller quantities of wind in high-quality areas and the Dyson-sphere-esque, scale of the project I do not expect wind to generate most power on net-zero-carbon timescales. For the sake of argument, let us suppose we can capture 5% of available wind power for $2 \times 10^{19} \text{ J}$.

As for biofuels, there is large disagreement about the potential energy resource. This is because there are large questions involving the sustainability of the biofuel source, the land available for biofuel, and the productivity of that land. A commonly accepted reasonable-bound is $1 \times 10^{20} \text{ J}$, with an upper bound of $7 \times 10^{20} \text{ J}$ [41]. The upper bound assumes large changes in dietary patterns, no issues with irrigation or chemical runoff, and compounding increases in energy crop yields; all contentious assumptions. Another key point is that this estimate of biofuel energy is the primary energy, not the useful energy, which is degraded by the engine burning the fuel. BECCS to electricity is expected to be 10% efficient, while BECCS to biofuel is expected to be 26% efficient [21]. The biofuel is then anywhere from 50% to 100% efficient in its application depending on if it is being used in an engine or for heating. We can be generous and allow 20% efficiency in going to the final application, for a useful energy of $2 \times 10^{19} \text{ J}$ which is another 5% of humanity’s 2050 energy demand.

Reasoning about nuclear capacity is more difficult because of its political sensitivity. Regardless, the International Atomic Energy Agency in its high-side forecast for nuclear in 2050 predicts 800 GW of nuclear generating capacity, and the IPCC 1.5 report calls for pentupling nuclear capacity [54]. Assuming the IAEA’s own high side estimate, nuclear can produce $2.5 \times 10^{19} \text{ J}$, or 6%
of humanity’s 2050 energy demand. Thus, between wind, hydro, biofuel, and nuclear, we have about 20% of humanity’s 2050 energy demands satisfied, leaving $3.3 \times 10^{20}$ J for solar to deliver in 2050. If nuclear doubles again between 2050 and 2075 and there are no fundamental developments in wind, hydro, or biomass, then solar will have to provide $8.1 \times 10^{20}$ J.

There are two central reasons why solar can reach much larger scales than wind, hydro, or biomass. First, the available resource is enormous. Yearly insolation is $3 \times 10^{24}$ J, so powering humanity in 2050 requires only harvesting 0.01% of the available resource [279]. Second, as compared to wind, the areal energy density is much greater. For example, the Topaz solar farm in California, using relatively low performance CdTe panels, has an annualized power density of 7.7 W/m$^2$, and the Solar Star power stations achieve 14.7 W/m$^2$ by using high performance cells with single-axis tracking, demonstrating that photovoltaics achieve an order of magnitude greater power density than wind [259, 234, 235, 236, 53]. A further concrete example is that NREL solar resource data, reproduced in Fig. 1.3, shows that most of the continental US has a resource $>4 \text{kWh/(m}^2\text{d)}$, with the southwestern US resource exceeding $5 \text{kWh/(m}^2\text{d)}$[223]. To put this in perspective, covering New Mexico with a solar farm as good as Solar Star produces $1.5 \times 10^{20}$ J. While covering double the area of New Mexico in solar panels is daunting, it is less daunting than covering the entire land and continental shelves of the Earth in windmills.

**Climate-Motivated Growth Models of Photovoltaic Manufacturing**

The first thing we need to do is derive estimates for the growth-rates for the existing silicon photovoltaics industry compatible with climate goals. The three main variables we have to work with are the net-zero dates, the module lifetime, and the final manufacturing capacity. The module lifetime is important because it tells us which manufactured panels contribute to power on a given date. For example, with the 2050 target and 25-year module lifetime, only panels installed after 2025 contribute to the power. The 25-year module lifetime is motivated by most major manufacturers warranting the power output of their panels for that long. This is another point where we underestimate necessary manufacturing capacity because the power output of panels decrease over time, but we will assume they are constant. The final manufacturing capacity is the manufacturing capacity needed to keep the energy supply growing at the needed 3.1% after the net-zero date. We will see that this reveals large
problems with a naive growth model. We can derive two constraints based on installed capacity and manufacturing capacity. We express the constraint on generating capacity as:

\[ GC_{NZ} = \int_{NZ-\Delta t}^{NZ} MC(t) \, dt, \quad (1.1) \]

where \( GC_{NZ} \) is the generating capacity on the net-zero date \( (3.3 \times 10^{20} \text{ J for 2050}, \text{ for example}) \), \( NZ \) is the net-zero year, \( \Delta t \) is the module lifetime, and \( MC(t) \) is the manufacturing capacity in a given year. It simply says that the total installed capacity has to meet demand. We can express the constraint on final manufacturing capacity as:

\[ g_d \times GC_{NZ} + \int_{NZ-\Delta t}^{NZ-\Delta t+1} MC(t) \, dt = \int_{NZ}^{NZ+1} MC(t) \, dt \quad (1.2) \]

where \( g_d \) is the yearly growth in demand, the first term is end-of-life modules being taken offline, and the right is newly manufactured capacity. It is important to note that manufacturing capacities are usually specified in “GW_p” where the ‘p’ subscript means that the power was produced the solar panels under AM1.5g illumination at 25°C and it is implicitly understood that the value is per-year. To convert that to annualized useful power production, we
use the capacity factor, which corrects for the day-night cycle, seasonal variations, clouds, plant maintenance, etc., and is typically 25% [52]. An intuitive model to use for MC is exponential growth:

$$MC(t) = MC_0(1 + g)^t,$$  \hspace{1cm} (1.3)  

where $MC_0 = 1 \times 10^{18}$ J is the initial manufacturing capacity in 2020 and $g$ is the annual growth rate [194]. If we solve Eqns. 1.3 and 1.1 with a 25 yr lifetime, we find $g = 13\%$ and that $MC(2050) = 4 \times 10^{19}$ J. The issue comes from considering Eq.1.2. In this case, $MC(2050) = 1.1 \times 10^{19}$ J so that manufacturing capacity has been overbuilt by a factor of three. This means many of the manufacturing plants will never be paid off which is, prima facie, unacceptable on economic grounds. Let us also consider net-zero in 2075 with the exponential growth model. In this case, the growth rate is only 8%, but the overbuild is now 150%. What this means is that 13% and 8% are lower bounds on the needed growth rate. Satisfying Eqn. 1.1 while decreasing the value of $MC_{NZ}$ requires increasing the value of MC in earlier years.

It is unclear how well known this issue is in the literature. For example, in the IPCC report, CO$_2$ emissions follow a sigmoid, which is consistent with sustainable demand being integrated in the model. Another example is [205], where they use a constant manufacturing capacity of roughly 1 TW yr$^{-1}$ over all of renewable energy starting in 2025, so installed capacity will plateau around 2050 as old generating capacity is replaced. But this point may not be widely appreciated within materials science and PV. For example, in [129], in their baseline model, they assume 100 TW$_p$ of installed PV and 10 TW$_p$ yr$^{-1}$ of manufacturing capacity in 2050. Assuming demand grows 3%, then two-thirds of factories are unused. Photovoltaics currently employs nearly 4 million people, so extrapolating out this crash results in the instant unemployment of 200 million people.

To get around this problem, let us consider a growth model where MC grows exponentially and then plateaus:

$$MC(t) = \begin{cases} MC_0(1 + g)^t & t < t_p, \\ MC_0(1 + g)^{t_p} & t \geq t_p \end{cases},$$  \hspace{1cm} (1.4)  

where $t_p$ is the plateau date. The motivation for this is that manufacturing capacity grows quickly, and can then gracefully produce enough power without
Figure 1.4: Illustration of manufacturing capacity over time for photovoltaics that meets energy demands that either does or does not incorporate sustainable demand consideration.

large capital inefficiencies or layoffs. Solving this model for 2050, we find $g = 25\%$, $t_p = 2032$, and $MC(NZ) = 1.8 \text{ TW}_p$. Solving for 2075, we find $g = 9\%$ and $t_p = 2063$, and $MC(NZ) = 5 \text{ TW}_p$. This calculation illustrates how much more dramatically difficult net-zero in 2050 is compared to 2075, assuming monotonic growth and no glaring capital inefficiencies, because manufacturing capacity increases must be completed in ten years instead of forty years. The difference between the exponential and plateau models are illustrated in Fig. 1.4.

We can also reason about the desired growth rates for a novel photovoltaic industry. Let us assume it can reach $10 \text{ GW}_p \text{ yr}^{-1}$, or $8 \times 10^{16} \text{ J}$ in 2030 from initial investments (an extremely optimistic assumption). In the simple exponential growth model, we find $g = 44\%$ and $g = 18\%$ with overbuilds of 13 times and 4.5 times for 2050 and 2075, respectively. With the plateau growth model, net-zero in 2050 requires instant build-out to $1.3 \text{ TW}_p$ because the module lifetime is greater than the time between the start and end of scaling. For 2075, we find $g = 26\%$ and $t_p = 2057$.

The plateau model can also illustrate a perverse feature of energy markets. The longer the solar panel lifetime, the lower the needed long-term manufacturing capacity, but the time-integrated manufacturing capacity is a constant because there is a fixed amount of energy that solar needs to generate on a timescale comparable to the panel lifetime, so manufacturing needs to grow more quickly
the longer panels last. To see this, we can consider a fifty-year module lifetime, which is a goal of the DuraMAT collaboration [38]. Assuming they meet their goal, the growth rates and plateau times for silicon for 2050 net-zero become 65% and 2025, and for 2075 are 10.5% and 2053. Increasing module lifetime makes hitting climate targets more difficult!

In conclusion, simple exponential growth models for installed photovoltaic capacity are unacceptable because they result in large overbuilds of manufacturing capacity. Using a more reasonable model where manufacturing capacity grows and then plateaus to be compatible with long-term energy demands, we find that silicon photovoltaic manufacturing must grow 25% and 9% per year for net-zero in 2050 or 2075, respectively. A novel photovoltaic technology would have to grow overnight for net-zero in 2050, or 26% per year for 2075. Increasing module lifetimes increases the speed with which manufacturing must grow to prevent capital and labor inefficiencies.

**Economics of Photovoltaic Manufacturing Growth**

In the previous section, we found that silicon photovoltaic manufacturing needs to grow at least 25% per year for net-zero in 2050. In this section, we will examine the economic factors that influence this goal. This section is almost entirely due to [200], which has deeply affected how I think about opportunities in photovoltaic research.

How quickly an industry can grow is governed by three factors. First is the operating profit, which is the profit a company can use for things like building new factories. Second is the cost of a factory, or CapEx. With a greater operating margin, a company can buy more factories and expand manufacturing more quickly. Third is the revenue, which sets the relative scale for CapEx and the operating profit. With these three numbers, we can calculate the maximum sustainable growth rate of a firm, which is reproduced from [200] in Fig. 1.5.

Each panel corresponds to a different financing strategy, where in Panel a growth is financed at a constant debt ratio while in Panel b no additional debt is incurred. Growth rates with constant debt loads are generally larger. On the horizontal axis, PP&E_0 is property, plant, and equipment (CapEx) in year-zero dollars, divided by the first year revenue. The contours correspond to different operating margins, which are operating profits divided by revenue.
To get a sense of real-world operating margins, we plot the operating margins of Jinko Solar (JKS), SunPower (SPWR), Canadian Solar (CSIQ), and First Solar (FSLR) in Fig 1.6. We choose these companies because they are traded on US exchanges, so their financials are readily available [109]. The main takeaways are that operating margins are volatile, usually less than 10%, and occasionally negative. Because capital depreciates over roughly ten years and the volatility in margin has a quarter-to-quarter timescale, planning new factories is difficult and risky. The limited operating margins of a competitive, well-established industry are not surprising [231, 159].

To further illustrate the fragility of silicon photovoltaic operating margins, let us consider the effect of the increase in polysilicon prices from $11 kg^{-1}$ in February 2021 to approximately $30 kg^{-1}$ in June 2021 and remaining that high as of the time of writing [199]. Assuming cells are 150 µm thick and 20% efficient this corresponds to a price increase of $0.03 W_p$ which, according to the latest NREL cost benchmarks, is enough to completely consume manufacturer’s gross margin of approximately $0.01 W_p$ [232]. The increase is so severe that 56% of 2022 utility PV installations are projected to be postponed or canceled [209].

To gauge the horizontal axis in Fig. 1.5, we now need to introduce how silicon solar cells are made. As a gross simplification, making a silicon solar cell involves the following steps:
Figure 1.6: Operating margin of various publicly traded PV companies. Notably these companies do not reliably make money.

- Carbothermal reduction of quartz into metallurgical grade silicon.

- Purification of metallurgical silicon into semiconductor-grade silicon by the Siemens process.

- Crystallization of the semiconductor-grade silicon into large-grained poly crystalline silicon by directional solidification or mono-crystalline silicon by Czochralski growth.

- Slicing the ingot into wafers, modifying their surfaces, and then packaging them into a module.

Something odd about this description is that all the semiconductor processing (doping, passivization, etc.), the steps that device physicists generally find most interesting, are bundled into the last step. If we examine the step-by-step capital intensity of turning metallurgical-grade silicon into a solar panel, reproduced from [200] in Fig. 1.7, then this breakdown is natural from an
Figure 1.7: Capital expenditure by fabrication step for silicon PV and grouped by broad process category. Reproduced from [200] under license CC BY 3.0 economics perspective because the capital intensities of purifying metallurgical silicon, recrystallization, and cell processing are all comparable. In fact, the nitride passivation and anti-reflection-coating (ARC) made by plasma enhanced chemical vapor deposition (PECVD), a step we might intuitively expect to be capital intensive because it requires vacuum, plasma, and reactive precursors, is actually less capital intensive than screen printing the metal contacts; the technique used to mass produce graphic t-shirts. The vertical lines in Fig. 1.5 correspond to different starting points in the cell manufacturing process with PP&E\(_0\) = 0.79, and 1.19 corresponding to starting with poly-Si, and starting with metallurgical Si, respectively. Because we are interested in the entire supply chain, we think 1.19 is the relevant PP&E\(_0\), so achieving 25% yearly growth requires an operating margin of greater than 25%, which the photovoltaic industry has historically been unable to achieve.

We can also estimate the necessary profit margin for a new photovoltaic technology to scale for net-zero purposes. Using the result from the 2075 net-zero calculation in the previous section that a new technology must grow 26% per-year and assuming a PP&E\(_0\) of 0.5 which corresponds to a capital-free substrate, this then requires an annual operating margin of about 10%, which has been reliably achieved by tech companies like Intel and Apple, but not by commodity companies like US Steel or Ford.

So, there are two ways for photovoltaic manufacturing to scale more quickly: increase profit margins, or decrease capital intensity. Increasing profit margins is impossible in a competitive market because of the race-to-the-bottom in prices, but could hypothetically be solved by forming a photovoltaic cartel a-la OPEC, but this is highly speculative. By lowering PP&E\(_0\) to 0.5, 26% per year
growth is possible with 10% profit margins, which can be done by eliminating the capital cost associated with silicon ingot production and maintaining tech-giant operating margins for fifty years. Further, whatever company starts manufacturing low capital intensity photovoltaics needs to maximally exploit its first-mover advantage to preserve its profit margin from competition. There is a great irony here. On one hand, the solar panel needs to be low capital-intensity so that manufacturing can grow, and on the other, there needs to be a high barrier-to-entry to suppress competition and preserve profit margins, but low-capital-intensity and a high barrier-to-entry seem to be mutually exclusive properties. In light of this, it may be best to think of a new, low-capital, photovoltaic technology as reducing the capital investment necessary if policymakers decide to tackle global-warming, rather than a solution all on its own.

**Opportunities for Photovoltaic Research**

There have already been numerous attempts to lower the capital intensity of photovoltaics manufacturing. Within silicon photovoltaics, this is the strategy being explored by upgraded metallurgical silicon (UMG-Si), where the silicon is purified by zone-refining instead of Siemens CVD. Polycrystalline UMG-Si cells are near performance parity with Siemens process cells and preliminary modules have similar lifetimes, but it still requires two melt-processing steps so it is not clear if UMG-Si achieves the desired capital reductions [62, 214].

An alternative strategy is to use thinner wafers or no wafers at all. By using thinner wafers, the capital cost of the ingot is spread over a larger number of devices, reducing the capital cost per Watt. Unfortunately, this technique is not available for silicon because it being an indirect bandgap semiconductor mandates using approximately 100 µm of material per cell to absorb enough sunlight [126].

By using a direct bandgap semiconductor, though, the active layer of a cell only needs to be a few microns thick, meaning large reductions in the capital intensity of fabricating the semiconductor are conceptually possible. Some example materials with photovoltaic relevance are: CdTe, copper indium gallium selenide (CIGS), perovskites like methylammonium lead iodide (MAPI) and caesium lead iodide (CsPbI), GaAs, and InP, with CdTe and CIGS already being manufactured at large scales. First Solar and its CdTe technology are an encouraging example of low CapEx photovoltaic manufacturing. The solar cell is vapor-deposited directly on glass, with fabrication taking less than
four hours [60]. Indeed, First Solar in June 2021 announced a new plant with a CapEx of only $0.2\, \text{W}_{\text{acap}}$ [61]. Despite this capital advantage, First Solar has struggled to maintain positive operating margins, as seen in Fig. 1.6. Less is known about CIGS financials, but MiaSole, one of the leading CIGS manufacturers, claimed a CapEx as low as $0.5\, \text{W}_{\text{acap}}$ in 2010 [166]. Tragically, though, CIGS is no longer economical with Mia Sole having ceased production and Solar Frontiers shifting to silicon modules. It is not enough to have low CapEx; the end cost of electricity must also be competitive which CIGS struggled with due to its low efficiencies. To demonstrate the hypothetical value of low CapEx PV, consider that the total capital investment for 2 TW$_p$ of silicon PV manufacturing is about $2T$ (over three times the value of the US Navy’s fleet [8]), but only $0.4T$ for CdTe (about half of what was spent on COVID stimulus checks).

Despite the success of CdTe as a low-CapEx technologies, it is worth exploring alternatives because it relies on tellurium, an extremely rare element. Optimistic projections of Te availability limit CdTe manufacturing to 0.25 TW$_p$ and $3 \times 10^{19}$ J of energy in 2050. This leaves about 1.25 TW$_p$ of manufacturing capacity uncovered by existing low-CapEx technologies.

Notably, the existing commercial technologies do not dominantly use gallium or indium, which can be used to make GaAs and InP photovoltaics. GaAs is promising because it is the second most used semiconductor after silicon and already produces the highest efficiency solar cells. What has been missing, and what much of this thesis is devoted to, is a low CapEx GaAs solar cell process. Estimating if there is enough gallium to be useful is difficult because the USGS has not even calculated the resource available, but it estimates there is at least 100 kt extractable from the world’s bauxite [171]. Estimates using market-models and line of sight improvements in gallium extraction suggest at least 5 kt yr$^{-1}$ availability, and very crudely multiplying world bauxite production and the average gallium content of 50 ppm suggests a resource of 15 kt yr$^{-1}$ [66, 151]. Assuming cells are 20% efficient and use 3 µm of GaAs, these imply a total installed capacity of 2.7 TW$_p$ and manufacturing capacities of 0.13 TW$_p$ to 0.4 TW$_p$. These capacities are similar to CdTe and CIGS, and so GaAs merits a similar amount of consideration.
Chapter 2

DIFFUSION DOPED GAAS SOLAR CELLS

2.1 Introduction
In the introduction to the thesis, I motivated why low CapEx GaAs solar cells can be a valuable contribution to fighting global warming. The goals of this chapter are to:

- Deduce the design and fabrication process of a low-CapEx GaAs solar cell, with first providing the necessary context;
- Calculate the theoretical performance of the device;
- Describe advances in low-CapEx GaAs device processing;
- Demonstrate much of a high-performance, low-CapEx GaAs cell;
- Develop a cost model for the proposed solar cells;
- Outline the remaining work to be done.

At this point, I assume the reader is already familiar with device physics and solar cell physics. Better explanations already exist than what I can write here, so I refer the reader to [249] for general device physics and [286] for solar-cell physics in particular. Note though that [286] does not correctly handle the heterojunction boundary conditions, which should be treated by thermionic emission theory [291]. Two other resources that were immensely helpful were pveducation.org, which is both an easy to use reference and has a number of good calculators for things like contact geometry, and pvlighthouse.com.au, in particular the OPAL 2 optics calculator which is a fully featured transfer-matrix solver for optimizing solar cell coatings.

2.2 Historical and Recent Context on GaAs Solar Cells
GaAs and, more broadly, III-V device design and processing is a mature and complex field, with one of the early milestones being the first semiconductor laser using diffusion doped GaAs in 1962, and the compound semiconductor market having since grown to around $100B [84, 180, 37]. With regard to
photovoltaics, III-V devices hold the efficiency records for both single-junction and multi-junction designs at 29.1% and 39.5%, respectively [16]. III-V cells are the preferred devices for aerospace due to their high performance and radiation hardness, with multiple established companies fabricating cells, such as SpectroLab, MicroLink, and SolAero. Despite their performance and maturity, III-V photovoltaics are currently limited to aerospace due to their prohibitive cost of around $100 W_p^{-1} [281, 99]. This exorbitant cost, three orders of magnitude greater than silicon PV, can be traced to three factors. First is the cost of the wafers. A six inch GaAs wafer costs about $100. Assuming this wafer becomes a record multi-junction device it already costs $14 W_p^{-1}, two orders of magnitude greater than silicon. The second cost-driver is the growth of the cell. Unlike silicon where the wafer eventually becomes optoelectronically active, with III-V’s the wafer is merely a template for crystal growth. The optoelectronically active portion is grown on the wafer using metal-organic chemical vapor deposition (MOCVD) or hydride vapor-phase epitaxy (HVPE). This is advantageous because vapor-grown material is generally higher quality than bulk material, and it allows forming heterojunctions and precise doping profiles [105, 155]. Heterojunctions in particular are critical to the high performance of III-V solar cells, enabling features such as high-voltages and monolithic multi-junction cells [117, 3]. The epitaxial growth itself then costs $15 W_p^{-1}, and subsequent processing like metallization an additional $20 W_p^{-1}, and the remaining cost comes from R&D, administration, and profit.

There are ongoing projects to make III-V PV viable for terrestrial applications. They revolve around the concepts of substrate reuse, and decreasing epitaxy costs. By reusing the substrate, the initial wafer cost can be spread over a large number of devices, thus reducing its cost per Watt. There are two common ways of implementing this. First is a process called epitaxial liftoff (ELO). In ELO, a sacrificial layer is grown on the wafer, commonly high aluminum content AlGaAs, before the cell. The sacrificial layer is then removed, separating the cell from the substrate and so the substrate can be used as a template for more cells. The highest efficiency is achieved with ELO devices because replacing the wafer with a mirror allows using photon recycling to boost the voltage [268, 124, 272]. The process is illustrated in Fig. 2.1. Using current technologies, ELO costs around $12 W_p^{-1}, assuming five reuses per wafer and a $25 chemo-mechanical polish after each ELO. With free
polishing and hundreds of wafer reuses the ELO cost can come down to about 
$0.10 \, \text{W}_{p}^{-1}$.

A second strategy for wafer reuse is mechanical spalling. In spalling, the wafer is attached to a handle with tensile stress. To minimize energy, a crack forms and propagates through the semiconductor parallel to the handle/semiconductor interface [248]. The process is illustrated in Fig 2.2. The stressor film is usually electroplated nickel because of its internal stress up to 1 GPa. Some exciting results from spalling include: 28.7% efficient multijunction cells [224], growing cells on spalled, unpolished germanium [27], no loss of efficiency for spalling compared to ELO for GaAs [247], growth on differently oriented surfaces [164], and spall thickness less than 1 $\mu$m [42]. According to preliminary technoeconomic modelling of spalling, the capital cost can be as low as $0.02 \, \text{W}_{p}^{-1}$, right in line with the per-step capital costs shown in Fig. 1.7.
Assuming 5 \mu m of nickel and 3 \mu m of GaAs from a 500 \mu m wafer, the materials cost is 0.5 c W_p^{-1} of nickel and $0.17 W_p^{-1}$ of GaAs. We will revisit these numbers in more detail later, but they set the stage for the economic viability of spalling. One particular standout feature of spalling compared to ELO is that spalling does not need an epitaxial growth to work, while ELO requires an epitaxial sacrificial layer.

Reducing epitaxy costs is mainly being tackled by increasing tool throughput. The first component of this is increasing the growth rates, with GaAs grown as quickly as 528 \mu m h^{-1} having been demonstrated using HVPE [161]. To provide a sense of scale, a typical growth rate for vertical gradient freeze (VGF) GaAs is 2 mm h^{-1} to 4 mm h^{-1} [114]. A second advance being worked on is in-line growth instead of batch-growth to maximize the amount of time that the tool is actually growing and get the most value out of it. With such a tool and perfect precursor utilization, NREL expects they can reduce the epitaxy cost to about $0.25 W_p^{-1} [99]. With further improvements to backend processing and free 200 mm substrates, the expected final cost of electricity is $0.4 W_p^{-1}.

Unfortunately, even in this asymptotic limit, the cost-per-Watt is still twice as expensive as silicon cells.

There have also been historical efforts at making diffusion-doped GaAs solar cells [17, 118, 68]. In these efforts, the diffusion source is a CVD deposited doped-glass layer. In [68], their best devices achieve a $V_{oc}$ of 770 mV and efficiency up to 14%, and in [17] the best device reaches 894 mV and 13%. The reason we thought we could do better is that the voltages reported were implausibly low, and so we hoped that maybe something had changed in crystal
growth between the mid 80’s and now that allow us to make better devices. In section 2.4, we will advance a hypothesis for why these early attempts suffered from low performance.

2.3 Deducing GaAs Solar Cell Design and Fabrication

From the previous chapter, we determined that a low CapEx GaAs solar cell must use a GaAs thin-film, and in the previous section, we determined that epitaxy is not a viable manufacturing process. A couple other requirements to consider are:

- Longevity Solar panels need to last thirty years. This means both that the materials used need to be stable in contact with each other and that they need to be strongly bonded together. One of the tests used to assure this is called damp heat, where the panel must survive 85°C at 85% relative humidity for 1000 hours. Perovskites in particular struggle with this test because it either activates their intrinsic decomposition pathways or water ingress catalytically destroys their crystal structure [47, 163, 271].

- Durability A key failure mode with solar panels is delamination of the layers. With silicon modules, this leads to loss of photocurrent and metal corrosion. Empirically, it has been found that modules need a cohesion energy of 160 J/m² to meet their durability goals [20]. In comparison perovskites and organics have cohesion energies of only a couple J/m² [206, 207].

Having ruled out epitaxy and perovskite-style heterojunctions, the only remaining strategy, that I could think of, for forming the electrically active junction is diffusion doping. From longevity and durability standpoints, III-V concentrating photovoltaic (CPV) solar cells have already passed damp-heat testing, indicating existing contact metallurgy and encapsulation are sufficient [266].

To understand what processing steps are allowable, it is helpful to examine the step-by-step manufacturing costs of silicon solar cells, shown in Fig. 2.3. At a total cost of $6.3 e W_p^{-1}$, 20% is wet chemical processing for cleaning and texturing, 6% is the diffusion doping, 10% per PECVD process, and 30% from metallization. What this shows is that the cost drivers are not complicated
tools like PECVDs, but instead corrosive chemicals like KOH, HF, and HNO₃, and the metal used for the cells.

Therefore, as long as we can avoid corrosive, toxic chemicals and expensive metals, the design space for fabrication is fairly flexible. In fact, GaAs has an intrinsic advantage in metal usage compared to silicon. Because GaAs solar cells operate around 30 mA/(cm²), while silicon cells operate around 40 mA/(cm²), and Joule heating conduction losses scale quadratically with current, a GaAs cell can use half as much metal for the same amount of power loss which is an instant 0.7¢/W⁻¹ cost savings. It would also be desirable for the cell to be a drop-in replacement for silicon cells in module fabrication. The main requirements in this regard is that it be rigid and the front and back metallization be solderable.

There is also the question of how to do the diffusion doping. We favor a $p^+/n$ design because $p$-type diffusion doping can be done at relatively low temperatures with simple materials. We will revisit this in more detail in the next section, and we’re purposefully being vague about where, exactly, the GaAs comes from. But, with all these restriction is mind, we can now present the solar cell, illustrated in Fig. 2.4. The electronically active semiconductor is a $n$-type GaAs film with a $p$-type diffusion-doped surface. The back handle that provides rigidity and a solderable surface is piece of tin-plate steel. We
chose this because the tin can make an ohmic contact to n-type GaAs, is easily solderable, cheap, and available in large quantities [219]. At $500 \text{ t}^{-1}$, 150 $\mu$m thick, and coupled to a 20% efficient cell, the cost is only $0.3 \text{ c W}^{-1}$ and enough tinplate is manufactured per year for $4 \text{ TW}_p \text{ yr}^{-1}$ of solar cell manufacturing. The front of the cell is covered with a ZnS anti-reflection coating. We choose ZnS because it can be easily deposited by thermal evaporation and is compatible with our surface passivation techniques, which we’ll go into more detail about later. Finally, the top contact is screen-printed silver; just the same as silicon. An intriguing possibility is switching the silver for copper. While silicon cells plated with copper show enhanced degradation during damp heat testing, diffusing copper into GaAs has been shown to be benign [116, 18].

We illustrate a potential fabrication strategy in Fig. 2.5. Starting with an n-type GaAs wafer we first diffusion dope it to make a $p^+/n$ junction. The next step is electroplating chromium that is the stressor layer for spalling. We choose chromium over nickel for a few reasons. First, chromium is much easier to remove than nickel. In personal experience, chrome readily etches...
in HCl while nickel is very corrosion resistant [40]. Second, chromium is not recombination-active in GaAs like nickel is [193]. Third, chrome can be plated at just as high of internal stresses as nickel, making it just as viable for spalling [196]. The next step is grabbing the chrome with a vacuum chuck and peeling it off to form the spalled film. Next, the GaAs film is laminated to a tinplate sheet. This provides both a rigid handle and an ohmic contact. The contact is formed through liquid phase epitaxy and works at temperatures as low as 350°C [219]. After this, the chuck is released and the chrome is removed in HCl. At this point, the surface can also be etched or passivated. The next major step is evaporating a ZnS film to work as anti-reflection coating. To enable contacts, opening are laser-ablated in the ZnS. Finally, contacts are screen-printed and fired. Metal pastes for silicon are typically fired around 800°C to allow the lead frit in the paste to burn through the nitride and for the aluminum on the back of the cell to form a eutectic with the silicon, but silver pastes are commercially available which cure at 200°C and are therefore compatible with this cell design [70, 145].
2.4 Device Performance Prediction

Having conceptually outlined a low-CapEx GaAs solar cell, we will now predict its performance to check that it is actually worth trying to build. To be competitive with existing PV, it should be at least 20% efficient. We will calculate cell performance in two ways. First, using a detailed-balance model where the only variable is the optical properties. Second, using an analytic diode model that includes the electrical transport. From the detailed balance model, we will find effects that limit performance at high doping and help explain the poor performance of historical devices, and the drift-diffusion model will describe performance at lower doping where short-base diode effects become important. In the process of my own work, I fabricated cells before doing this calculation which, in hindsight, was rather brash but seems to have worked out OK.

Detailed Balance Model

In a detailed balance model, the idea is to keep track of thermodynamically mandated particle fluxes as a function of bias and extract cell performance characteristics. This concept was originally applied to solar cells by [229], and we will use an extension to it described in [282].

To make the model work, we need a simple device structure to consider and the optical properties of each material. Something that complicates this is that the optical properties of GaAs depend on its doping. To wit, we plot the absorption coefficients of n and p-type GaAs in Fig. 2.6a, which has been collated from [262, 222]. The plots have been truncated before free-carrier absorption dominates the long-wavelength response. In both n and p-type GaAs the above gap absorption decreases with doping which can be attributed to a Burstein-Moss shift. However, with p-type material, there is significant below-gap absorption that extends further into the IR with increasing doping. To understand why we need the more sophisticated model of [282], we plot the absorption coefficients times the energy dependence of the blackbody spectrum, Eq. 2.1, in Fig. 2.6b. The main thing to notice is that for n-GaAs the weighted absorption is relatively steady with doping, but with p-GaAs the absorption peak red-shifts and then starts diverging at high doping.

\[
\text{BB}(E) = \frac{E^2}{e^{(E/k_BT)} - 1}
\]  

To understand what is going on here, we can invoke the Urbach model of
Table 2.1: Urbach energy of GaAs as a function of doping extracted from literature absorption coefficient measurements.

<table>
<thead>
<tr>
<th>p-GaAs</th>
<th>Urbach Energy, meV</th>
<th>n-GaAs</th>
<th>Urbach Energy, meV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doping, cm(^{-3})</td>
<td></td>
<td>Doping, cm(^{-3})</td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>6.9</td>
<td>5 × 10(^{16})</td>
<td>6.4</td>
</tr>
<tr>
<td>2.2 × 10(^{17})</td>
<td>9.3</td>
<td>2 × 10(^{17})</td>
<td>11.3</td>
</tr>
<tr>
<td>1.2 × 10(^{18})</td>
<td>10.7</td>
<td>5.9 × 10(^{17})</td>
<td>10.7</td>
</tr>
<tr>
<td>1.6 × 10(^{19})</td>
<td>18.8</td>
<td>2 × 10(^{18})</td>
<td>13.5</td>
</tr>
<tr>
<td>6 × 10(^{19})</td>
<td>62.5</td>
<td>6.7 × 10(^{18})</td>
<td>23.0</td>
</tr>
<tr>
<td>1 × 10(^{20})</td>
<td>67.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sub-gap absorption where the absorption coefficient goes as:

\[
\alpha(E) = \alpha_0 e^{E/E_U},
\] (2.2)

where \(E\) is the photon energy, \(\alpha_0\) fixes the absorption scale, and \(E_U\) is the Urbach energy which sets the slope of the long-wavelength absorption. Multiplying the blackbody spectrum by the Urbach absorption and taking \(E \gg k_B T\), we get:

\[
\alpha_{BB} \propto E^2 e^{E/(E_U^1 - (k_B T)^{-1})}.
\] (2.3)

When \(E_U > k_B T\), the low energy behavior switches from exponential suppression to exponential enhancement, providing a mechanism for the divergence. Fitting the long-wavelength absorption coefficient to the Urbach model, we present the Urbach energy as a function of doping for GaAs in Table 2.1. As expected, the highest doped p-type material has \(E_U > k_B T\).

The reason we might want to work with such heavily doped material in the first place is that a desirable contact resistance for the front contacts of a solar cell is about 1 × 10\(^{-3}\) Ω cm\(^2\) [217]. Using Eqn. 127 of [249], we can calculate the contact resistance of p-type GaAs as a function of doping. Assuming the Fermi-level is at the band-edge, the surface is pinned midgap, and the Richardson constant is 7 A cm\(^{-2}\) K\(^{-2}\), we find the satisfactory carrier concentration is 2.5 × 10\(^{19}\) cm\(^{-3}\); awfully close to where the Urbach energy increases [172].

We also need a minimal model for the quasi-Fermi level (QFL) splitting throughout the device because the emission depends on it. The model is illustrated in Fig. 2.7b. We are using a simple, abrupt, one-sided diode model where the majority carrier Fermi-levels are constant in the quasineutral regions, the QFL’s are constant through the depletion region, and then decay linearly over
Figure 2.6: Optical properties of GaAs versus doping. a) Absorption coefficients of n and p-type GaAs as a function of doping. b) Absorption coefficient of GaAs multiplied by the blackbody spectrum to estimate luminescence spectrum.

Figure 2.7: Specifications for the detailed balance model. a) Optical model for cell efficiency calculation. b) Electrical model for cell efficiency calculation.
the width of the p-GaAs, or over the hole diffusion length in the n-GaAs. The QFL splitting in the depletion region is equal to the applied bias, $V$, and the depletion region is entirely in the n-GaAs. Taking $x = 0$ as the device surface, the chemical potential throughout the device is:

$$\Delta \mu(x) = \begin{cases} \frac{V x}{W_p} & x < W_p \\ V & W_n < x < W_p + W_d \\ V \left(1 - \frac{x - W_p - W_d}{L_p}\right) & x > W_p + W_d \end{cases}.$$  \hspace{1cm} (2.4)

We can now start calculating the current-voltage curve of the device. The central thing we need to calculate is the electroluminescent emission from the device as a function of bias:

$$J_{EL} = q \int S_{BB}(E, \Delta \mu(x))(f_v - f_c)dE dx,$$  \hspace{1cm} (2.5)

where $J_{EL}$ is the electroluminescent emission from the device as function of bias, $q$ is the elementary charge, $S_{BB}(E, \Delta \mu(x))$ is the volume emission given by [136], and $f_v - f_c$ encapsulates band-filling effects. $S_{BB}(E, \Delta \mu(x))$ is given by:

$$\frac{8\pi n^2(E)}{h^3 c^2} \frac{E^2 \alpha(E)}{\exp \left( \frac{E-\Delta \mu}{k_B T} \right) - 1},$$  \hspace{1cm} (2.6)

where $n(E)$ is the index of refraction. We also take the approximation that:

$$f_v - f_c = \tanh \left( \frac{E - \Delta \mu}{4k_B T} \right).$$  \hspace{1cm} (2.7)

At this point, we now have enough information to calculate the current-voltage curve of the solar cell using:

$$J(V) = J_{sc} - J_{EL}(V),$$  \hspace{1cm} (2.8)

where $J_{sc}$ is the zero-bias current under illumination. We assume $J_L = 30 \text{ mA cm}^{-2}$. From this model, we can calculate the efficiency, $\eta$, open-circuit voltage $V_{oc}$, and fill-factor (defined as $\eta/(V_{oc} * J_{sc})$. We additionally calculate the dark-current ideality factor. The results of this model are presented in Table 2.2.

There are several things to note. First, all the predicted efficiencies are greater than 20%, which is a good first sign for being able to make a cell. Second,
Table 2.2: Cell performance as a function of doping for detailed balance model that includes subgap absorption.

<table>
<thead>
<tr>
<th>Doping, cm$^{-3}$</th>
<th>$J_{sc}$, mA cm$^{-2}$</th>
<th>$V_{oc}$, V</th>
<th>Eff, %</th>
<th>FF, %</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1.2 \times 10^{18}$</td>
<td>30</td>
<td>1.085</td>
<td>29.0</td>
<td>89</td>
<td>1.02</td>
</tr>
<tr>
<td>$1.6 \times 10^{19}$</td>
<td>30</td>
<td>1.085</td>
<td>29.0</td>
<td>89</td>
<td>1.02</td>
</tr>
<tr>
<td>$6 \times 10^{19}$</td>
<td>30</td>
<td>0.975</td>
<td>22.9</td>
<td>78</td>
<td>2.16</td>
</tr>
<tr>
<td>$1 \times 10^{20}$</td>
<td>30</td>
<td>0.935</td>
<td>21.6</td>
<td>77</td>
<td>2.13</td>
</tr>
</tbody>
</table>

all device performance metrics decrease with doping; with the efficiency dropping from 29% to 21.6%. This illustrates the importance of taking non-ideal optical effects into consideration. A third interesting effect is the increase in ideality factor from one to two. It is common to assume in first-order device modeling that the effective diode has an ideality factor of one because it is consistent with both the ideal-diode equation and the most common dependence of radiative emission on bias, with ideality factors of two being attributed to depletion region recombination. In this case, an ideality factor of two comes from a purely thermodynamics and optics model. Finally, there seems to be an inflection point between $1.6 \times 10^{19}$ cm$^{-3}$ to $6 \times 10^{19}$ cm$^{-3}$ doping, which highlights the need for modern measurements of GaAs’s optical properties at these high carrier concentrations, especially because that is also near the critical doping range for making ohmic contacts. To figure out when the Urbach absorption becomes important, we plot the device performance versus Urbach energy assuming $6 \times 10^{19}$ cm$^{-3}$ doping in Fig. 2.8. From this plot, the critical Urbach energy when performance begins to severely degrade is 40 meV. Unfortunately, the literature data for absorption versus doping does not cover this range.

This model also suggests a hypothesis for why historical attempts at diffusion-doped GaAs solar cells did not work as well as expected. In both [17, 68], the authors found inexplicably low $V_{oc}$’s and high ideality factors. They also doped their GaAs to $1 \times 10^{20}$ cm$^{-3}$ or higher. This means they were both in regimes where Urbach absorption becomes important and the reasoning of this section presents a mechanism for their missing voltage.

**Drift-Diffusion Model**

In the previous subsection, we calculated the limiting efficiency of a $p^+/n$ GaAs solar cell based purely on its optical properties and a rudimentary model of position-dependent chemical potential. In this section, we will also calculate
limiting inefficiencies, but this time by using the carrier transport properties of the material. The central equation we will be working with is the drift-diffusion model for the diode saturation current:

\[ J_0 = q \left( \frac{D_p p_0}{L_p} + \frac{D_n n_0}{L_{\text{phys}}} \right), \]  

(2.9)

where \( J_0 \) is the saturation current, \( D \) is the diffusion coefficient, \( L_p \) is the hole diffusion length, \( p/n \) are the minority carrier concentrations in the quasi-neutral bulk, and \( L_{\text{phys}} \) is the physically relevant length-scale for the electrons: \( L_n \) or \( W_p \), whichever is smaller in case the device is best modeled as a short-base diode. We can also relate the diffusion lengths to minority carrier lifetimes by:

\[ L = \sqrt{D \tau}, \]  

(2.10)

where \( \tau \) is the lifetime. For the weakly \( (5 \times 10^{16} \text{ cm}^{-3}) \) doped n-type side of the diode, we can use common tabulated values to calculate the hole diffusion current with a worst-case scenario of \( \tau_p = 1 \text{ ns} \) to find \( J_p = 1 \times 10^{-15} \text{ mA cm}^{-2} \) \[195\]. By setting the diode diffusion current equal to the photogenerated current, we can estimate how the hole-diffusion current limits \( V_{oc} \) using:

\[ V_{oc} = \frac{k_B T}{q} \ln \left( \frac{J_{sc}}{J_0} \right), \]  

(2.11)

where \( J_0 \) is the diffusion current. Using this, the hole diffusion current limited \( V_{oc} \) is 980 mV.
The p-type side is more complicated because the bandstructure of GaAs changes at very high doping. Thankfully the heterojunction bipolar transistor community has determined the minority electron diffusion coefficient and intrinsic carrier concentration up to $1.4 \times 10^{20} \text{ cm}^{-3}$, and the lifetime up to $6 \times 10^{19} \text{ cm}^{-3}$ [156, 87]. The lifetime is well modeled by radiative and Auger recombination with:

$$\frac{1}{\tau} = B N_A + C N_A^2,$$  \hspace{1cm} (2.12)

where $B$ is the radiative recombination coefficient at $0.5 \times 10^{-10} \text{ cm}^3 \text{s}^{-1}$ and $C$ is the Auger coefficient at $4 \times 10^{-30} \text{ cm}^6 \text{s}^{-1}$.

In [87], they also include a correction factor, $\alpha$, for ballistic carrier effects and how doping increases the intrinsic carrier concentration by altering the bandstructure and being an effective correction for degeneracy effects. The collated material parameters are shared in Fig. 2.9a. Using all this data, we can then calculate the electron contribution to the diffusion current using 2.9, which is plotted in Fig. 2.9b. Using the sum of the electron and hole diffusion currents and the ideality-factor one diode voltage dependence, we can now calculate the cell performance metrics. These are displayed in Fig. 2.9c. The chief thing to notice is that the efficiency and $V_{oc}$ both increase with doping, in strong contrast to the detailed balance model. In this case, the increasing performance with doping makes sense because it suppresses the electron diffusion current, increasing the voltage. The reason the optical model misses the low-doping efficiency decrease is because it does not know about short-base diode effects. Therefore, we expect this electrical model to be accurate at low-doping, while the optical model is accurate at high doping. To illustrate how the models bracket theoretical efficiency, we plot the predicted efficiencies on the same axis in Fig. 2.10. We see that the theoretical device efficiency is bounded by about 25% and decreases rapidly after $5 \times 10^{19} \text{ cm}^{-3}$, which is close to the doping concentration needed for good ohmic contacts.

It would be interesting to do more detailed device physics calculations of performance using a tool like Sentaurus. It would be straightforward to incorporate the heavy-doping transport effect, but it is not obvious how to include the optical effects on radiative recombination. Having discussed the theoretical performance of $p^+ / n$ GaAs solar cells and finding that they can be greater than 20% efficient, we now turn to their fabrication.
2.5 Low Temperature GaAs Zinc Diffusion Doping

Of the steps illustrated in Fig. 2.5, we focused our efforts on the diffusion doping, cell fabrication, and metallization processes. We will discuss each in turn, with now focusing on the diffusion doping.

Despite diffusion doping of silicon and III-V’s being conceptually identical; the migration of impurity atoms from high to low concentration, the role of diffusion doping in device fabrication in each material system has diverged substantially. The divergence is due to differences in the chemistry, economics, final devices, and complementary fabrication techniques available to each material system. A concrete example of the consequences of this divergence is that diffusion doping is the standard technique used for making silicon solar cells, and there are numerous companies that supply parts for silicon diffusion doping, while the latest, semi-commercial, diffusion doped III-V photovoltaic research I can find is on InP from 1996 [178].
Chemical reasons for III-V diffusion doping being less common include surface decomposition from the high vapor pressure of the group-V component, toxicity issues with the group V element and doping elements, and complexities of working with a two-sublattice crystal. To illustrate the issues with vapor pressure, I plot the group-V dimer and tetramer, and group-III vapor pressures of GaAs and InP in Fig. 2.11, along with the vapor pressure that, in my experience, leaves an acceptable surface after 30 minutes in the particular diffusion apparatus I have developed [12, 244]. In contrast, silicon does not have a notable vapor pressure until 1000°C. The maximum reasonable working temperature is somewhere around 550°C and 450°C for GaAs and InP, respectively.

The reason why III-V semiconductors being a compound crystal complicates things is that the electrical properties of a dopant atom depend on which lattice site it occupies, and likewise the properties of native point defects depend on which component and sublattice they are interacting with. A great example of this is the amphoteric behavior of germanium and silicon in GaAs. One of these on a gallium site is a donor, while on an arsenic site it is instead an acceptor. This is not a purely academic issue. The doping behavior of these elements
depends on the growth parameters of the crystal in real experiments [289, 253]. To understand compound semiconductor doping, we need to understand the mechanisms that determine how impurities are incorporated into the crystal and how they depend on processing conditions. The thermodynamics of defects and dopants in GaAs is at this point well understood, and I suggest [104] for a comprehensive review and [26] for focusing on diffusion in particular.

The central point is that to diffusion dope III-V materials we need to control the creation of interstitial and vacancy defects on the appropriate sublattice. For example, for zinc doping of GaAs to work, we need to generate gallium vacancies for the zinc to occupy. In practice, the relevant equilibrium reaction is:

\[ \text{Zn}_i^+ \rightleftharpoons \text{Zn}_{\text{Ga}} + \text{I}_{\text{Ga}}^+ + h^+, \]

where \( \text{Zn}_i^+ \) is a zinc interstitial donor, \( \text{Zn}_{\text{Ga}} \) is a zinc acceptor on a gallium site, \( \text{I}_{\text{Ga}}^+ \) is a gallium interstitial donor and \( h^+ \) is a hole. This is known as the ‘kick-out’ mechanism, where a fast-diffusing interstitial displaces a regular atom in the lattice to form a self-interstitial. With this reaction in mind, we can now understand why historical III-V semiconductor doping techniques are the way...
that they are. Historically,\(^1\) the diffusion doping is done at 600°C or above. At this temperature, the arsenic pressure from the GaAs is high enough that it presents a safety risk, and so the diffusion is done in a sealed quartz ampule. Also, at these temperatures, the arsenic has a vapor pressure greater than the gallium, so the surface becomes dominated by arsenic vacancies. Due to the equilibrium between vacancies, codified by:

\[
V_{As} + V_{Ga} \rightleftharpoons 0 ,
\]

where \(V_{As}\) is an arsenic vacancy, \(V_{Ga}\) is a gallium vacancy, and 0 is an empty lattice site (the surface), there is a dearth of gallium vacancies to annihilate the gallium interstitials generated by Eqn. 2.13 through the reaction:

\[
I_{Ga} + V_{Ga} \rightleftharpoons Ga_{Ga} .
\]

To get around this, device fabricators add an arsenic source to the ampoule, usually some combination of \(Zn_3As_2\) or \(ZnAs_2\) \(^{227}\). These have a greater arsenic vapor pressure than GaAs, so the GaAs surface becomes dominated by gallium vacancies that the gallium interstitials can annihilate with. A second reason for the additional arsenic is prevent formation of a liquid phase. To see this, we reproduce the 700°C Ga-As-Zn phase diagram in Fig. 2.12 \(^{26}\). Without an arsenic source, the system would be held in zones I or II where a liquid phase exists, which will damage the sample surface.

The realization that allows dramatically simplifying the process is that the problems of arsenic vapor pressure and liquid formation disappear by working at lower temperatures. By working at a lower temperature, the gallium vapor pressure exceeds the arsenic pressure so that the GaAs surface becomes arsenic rich and the overall vapor pressure is low enough that decomposition is no longer an issue. This means that the sealed quartz ampule and toxic zinc arsenide precursor are no longer necessary, and we can instead use metallic zinc. We can see this in the phase diagram as well where perturbing the GaAs by adding zinc and losing gallium pushes it into zones III through VI where the coexisting phases are solids. The temperature where the surface switches from III-rich to V-rich is the congruent sublimation temperature and is 625°C for GaAs, and 365°C for InP.

\(^1\)And the way it is still done in APh109.
The GaAs-Zn phase diagram is sufficiently well understood that we can also predict the zinc concentration quantitatively for given experimental conditions \([112]\). The zinc concentration is given by:

\[
[Zn] = \frac{K a_{As}^2 a_{Zn}^2}{\gamma_h \left( \frac{n_i^2}{K} + a_{As} a_{Zn} \right)},
\]

(2.16)

where \([Zn]\) is the zinc concentration, \(a_x\) is the chemical activity of arsenic and zinc, \(\gamma_h\) is the hole activity, \(n_i\) is the intrinsic carrier concentration, and \(K\) is an equilibrium constant given by:

\[
K = 10^{-45} \exp \left( \frac{0.37 \text{ eV}}{k_B T} - 9.26 \right) \text{ cm}^{-6}.
\]

(2.17)

The arsenic-zinc activity product is given by:

\[
a_{As} a_{Zn} = \frac{p_{As}^{1/4} p_{Zn}}{p_{0,As}^{1/4} p_{0,Zn}}
\]

(2.18)
where \( p \) denotes the pressure of the relevant species, a zero subscript refers to the pressure above a liquid elemental source (even if it would be supercooled) and no subscript refers to conditions in the diffusion apparatus. To be consistent, we use the vapor pressures from [127]. To get the arsenic pressure, we consider that the vapor pressure of \( \text{As}_4 \) over solid arsenic is:

\[
p_{\text{solid,As}_4} \text{(mmHg)} = 10^{\frac{-6160}{T(K)} + 9.82}.
\]  

To convert this to the vapor pressure over the liquid, we can use the Clausius-Claperyon relation which, in this situation, is:

\[
\log_e(p_{\text{liquid}}) = \log_e(p_{\text{solid}}) + \frac{\Delta H_{\text{fus}}}{R} \left( \frac{1}{T_{\text{sub}}} - \frac{1}{T_{\text{fus}}} \right),
\]

where \( \Delta H_{\text{fus}} \) is the heat of fusion (111 kJ mol\(^{-1}\) for \( \text{As}_4 \)), \( T_{\text{sub}} \) is the sublimation temperature (887 K), and \( T_{\text{fus}} \) is the melting temperature (1089 K). Plugging Eqn. 2.20 into Eqn. 2.19, we get:

\[
p_{0,\text{As}_4} \text{(mmHg)} = 10^{\frac{-6160}{T(K)} + 11.04}.
\]

To estimate \( p_{\text{As}_4} \) and \( p_{\text{Zn}} \), we now need to describe the diffusion setup. Conceptually, we have a zinc source of variable activity and a GaAs source in close proximity. The composition of the zinc source determines \( p_{\text{Zn}}/p_{0,\text{Zn}} \), so we simply set \( p_{\text{Zn}}/p_{0,\text{Zn}} = a_{\text{Zn}} \). Limited decomposition of the GaAs provides \( p_{\text{As}_4} \). The GaAs predominately loses arsenic as \( \text{As}_2 \), but the thermodynamically favored form is \( \text{As}_4 \), which we assume dimerizes quickly so that:

\[
p_{\text{As}_4} = p_{\text{GaAs,As}_4} + 1/2 * p_{\text{GaAs,As}_2}.
\]

As a practical matter, we also know that \( K \gg n_i^2 \) and will assume that \( \gamma_h = 1 \), so that Eqn. 2.16 reduces to:

\[
[Zn]^2 = Ka_{\text{As}}a_{\text{Zn}}.
\]

At this point, we know how the above variables depend on temperature and zinc activity so we can now plot the zinc solubility, which is shown in Fig. 2.13. There are two main things to notice. First, the solubility is basically temperature independent with just a slight retrograde solubility. Second, the zinc solubility is high; up to \( 1 \times 10^{21} \text{ cm}^{-3} \). The high solubility may be problematic in light of the optically-limited efficiency at high doping, but it does bode well for being able to make a device at these low temperatures. Also,
the solubility only scales with the arsenic pressure to the one-eighth power so it is of little practical relevance.

We can also think about whether the kinetics will be reasonable. Assume that we want 100 nm of GaAs doped to $5 \times 10^{19}$ cm$^{-3}$ of zinc. This corresponds to a dose of $5 \times 10^{14}$ cm$^{-2}$. Using the Hertz-Knudsen equation and assuming that the zinc pressure comes from a pure metal source at 500°C, the zinc is supplied in just 2.3 µs. The necessary amount of gallium would be released in twenty minutes. This order-of-magnitude calculation shows that gallium release will be rate-limiting, but the timescale is still reasonable.

**Practical GaAs Zinc Diffusion**

In the previous subsection, we demonstrated the theoretical viability of low-temperature GaAs zinc diffusion doping. In this section, I will present three
techniques for doing the doping:

- a molten Sn/Zn source,
- a solid Cu/Zn source,
- a solid ZnO source.

For each technique I will share electrical measurements of the diffused layers and practical considerations about each method. All methods did work, but I have found the molten source the easiest to use and the most flexible.

For completeness, this is not the first time zinc diffusion has been done at lower temperatures [58]. In this previous work, they used a doped glass and they did not appreciate that at lower temperature the glass is not even needed. The reason we did not explore doped-glass methods of diffusion doping is that the cost is prohibitive for photovoltaics. Honeywell recommends 3 mL of liquid per six inch wafer, and the solution costs about $100 per ounce\(^2\), so the materials cost is $2.8 W_p^{-1}$, which is two orders of magnitude too high [240]. An interesting possibility would be to use in-furnace CVD, similar to how phosphosilicate glass is formed on silicon for phosphorous diffusion using POCl\(_3\), but we do not have a furnace readily amenable to this [46].

Before diffusion doping, the GaAs was first sonicated for three minutes each in acetone and isopropanol to remove dust and most organic contamination.

**The Sn/Zn Diffusion Source**

The diffusion system consists of a closed boat that contains a Zn/Sn melt and the GaAs to be doped. A conceptual schematic is shared in Fig. 2.14a., and a specific realization in part b of the same figure.

In the schematic, the premise is that zinc vaporizes from the melt and zinc vapor fills the region defined by the quartz boat as cover. At the same time, gallium leaves the GaAs surface, allowing the zinc to enter the GaAs. One possible sink for the gallium is the zinc/tin melt because it will absorb gallium until the partial pressure of gallium from the melt equals the partial pressure of gallium from the GaAs. Because the gallium pressure from GaAs is near that of gallium metal, the melt can absorb a lot of gallium before saturating.\(^2\)

\(^2\)According to the current price list from Desert Silicon.
Figure 2.14: The Zn/Sn diffusion source. a) Schematics of the Zn/Sn diffusion source. b) Picture of the diffusion source with various elements highlighted with corresponding portions from a.
There are a few reasons why I thought tin was a good choice for making the diffusion source alloy. First, tin has a very low vapor pressure below 600°C, so it should be basically inert with respect to the GaAs. Even if it does react with the GaAs, it will only be as a compensated n-type dopant. Secondly, tin and zinc form a eutectic that melts at only 200°C [177]. This is attractive because it means the diffusion source will remain well mixed. Finally, SnO₂ is readily reduced by hydrogen so the diffusion source will resist oxidation. We also found that the zinc HAS to be diluted otherwise the GaAs surface becomes badly degraded, likely due to the formation of zinc arsenides. This is supported by research on GaAs epitaxy where using zinc near its undiluted vapor pressure caused parasitic Zn₃As₂ growth [30].

In the experimental apparatus, the actual diffusion region is a quartz slider boat (MTI Corp.) and a piece of mineral wool (Lynn Manufacturing Kaowool). I prefer mineral wool because it is flexible enough to reasonably seal the boat to minimize zinc escape, and is inert. Inside the quartz boat, there is a smaller quartz boat that contains the Zn/Sn melt and the sample is placed somewhere near it. The melt was prepared by mixing metallic tin and zinc (Kurt Lesker) at approximately the eutectic mixture and pre-melting it under forming gas at 300°C. I also used graphite electrode paper as the boat cover but had a hard time making a seal. Zinc metal would noticeably coat the cold end of the tube furnace.

The diffusion is performed in a tube furnace under forming gas (10% H₂ in N₂, Airgas Certified Standard). The apparatus is loaded into the tube furnace when it is below 200°C, and then warmed up to the diffusion temperature over thirty minutes with a ramp starting at 0°C. We did this to make sure that there was time for the oxygen in the apparatus to be displaced or react with the hydrogen. It is then held at the diffusion temperature for the desired time, and then naturally allowed to cool to at most 200°C before we removed it. By letting the apparatus cool in the furnace we avoided oxidation by exposing hot metal or GaAs to air.

The main way we characterized the diffusion was by Hall measurements. We found processing between 450°C for at least an hour and 550°C for up-to thirty minutes left the surface mirror smooth and with sheet resistances less than 1 kΩ/□. The results from the Hall measurements are compiled in Table 2.3. In the table, Temp refers to the diffusion temperature, time to the time
with zero meaning reaching the diffusion temperature then cooling back down, Sh. Res. is the sheet resistance, $\mu_h$ is the hole mobility and Conc. is the sheet carrier concentration.

This data comes from 29 samples over a dozen runs. Every run successfully doped the surface p-type, indicating the robustness of the process. One thing to notice, though, is that the mobility is not consistent within temperatures. This is a clear sign of run-to-run variation. Several reasons for this could be inconsistent sealing of the quartz boat or control of oxygen in the boat. Another way to assess run-to-run consistency is to use the literature result that diffusion depth goes with the square-root of time so that the sheet resistance should obey:

$$\rho_{\text{Sh.}} \sqrt{t} = \text{Const.} \quad (2.24)$$

Calculating this for constant for 450, 500, and 500$\degree$C, we get:

$$(5300 \pm 1200) \, \Omega/\square \sqrt{\text{min}}, \quad (2300 \pm 1000) \, \Omega/\square \sqrt{\text{min}}, \quad \text{and} \quad (800 \pm 300) \, \Omega/\square \sqrt{\text{min}},$$

respectively, which suggests 50% run-to-run variation. Encouragingly, though, the values increase with temperature which is consistent with the diffusion coefficient increasing with temperature. As another test, we can use that:

$$\rho_{\text{Sh.}}^{-1} \propto \sqrt{D \ast t}, \quad (2.25)$$

where $D$ is the diffusion coefficient. Assuming the diffusion coefficient has an Arrhenius form, we fit it to the resistance-time products above and found an activation energy of 1.9 eV, which is consistent with literature [130].

We also explored the effects of post-diffusion annealing. This was motivated by the knowledge that the gallium interstitials generated during the diffusion create crystal defects such as dislocation loops and that annealing under arsenic pressure eliminates the defects [96]. The reason the arsenic pressure reduces the defect density is that it provides the gallium-rich defects a way to annihilate by reforming GaAs. We did the annealing by placing the chip face down on a quartz slab. To study the effect of annealing temperature, we diffused a chip at 500$\degree$C for 30 minutes, and then sequentially annealed it at 450, 500, and 500$\degree$C for 30, 30, and 10 minutes, respectively. The Hall measurement results are shared in Table 2.4. Most notable is that the 500$\degree$C anneal increases the carrier concentration by 20% relative to the baseline, and the mobility decreases by 14%, indicating a greater density of ionized dopants. At 550$\degree$C, the mobility increases suggesting that the zinc is now diffusing deeper into the GaAs.
### Table 2.3: Zn/Sn Diffusion Source Hall Measurements

<table>
<thead>
<tr>
<th>Temp., °C</th>
<th>Time, min.</th>
<th>Sh. Res., Ω/□</th>
<th>$\mu_h$, cm²/(V s)</th>
<th>Conc., cm⁻²</th>
</tr>
</thead>
<tbody>
<tr>
<td>450</td>
<td>60</td>
<td>780</td>
<td>43.2</td>
<td>1.84E+14</td>
</tr>
<tr>
<td>450</td>
<td>60</td>
<td>546</td>
<td>32.8</td>
<td>3.48E+14</td>
</tr>
<tr>
<td>450</td>
<td>60</td>
<td>852</td>
<td>38.1</td>
<td>1.31E+14</td>
</tr>
<tr>
<td>450</td>
<td>60</td>
<td>541</td>
<td>21.2</td>
<td>5.47E+14</td>
</tr>
<tr>
<td>475</td>
<td>30</td>
<td>307</td>
<td>32.2</td>
<td>6.31E+14</td>
</tr>
<tr>
<td>475</td>
<td>45</td>
<td>480</td>
<td>44.5</td>
<td>2.92E+14</td>
</tr>
<tr>
<td>500</td>
<td>15</td>
<td>483</td>
<td>37.1</td>
<td>3.49E+14</td>
</tr>
<tr>
<td>500</td>
<td>15</td>
<td>366</td>
<td>43.4</td>
<td>3.92E+14</td>
</tr>
<tr>
<td>500</td>
<td>30</td>
<td>321</td>
<td>50.7</td>
<td>3.82E+14</td>
</tr>
<tr>
<td>500</td>
<td>30</td>
<td>330</td>
<td>66.15</td>
<td>2.86E+14</td>
</tr>
<tr>
<td>500</td>
<td>30</td>
<td>377</td>
<td>13.5</td>
<td>1.22E+15</td>
</tr>
<tr>
<td>500</td>
<td>30</td>
<td>511</td>
<td>39</td>
<td>3.10E+14</td>
</tr>
<tr>
<td>500</td>
<td>30</td>
<td>338</td>
<td>44.4</td>
<td>4.16E+14</td>
</tr>
<tr>
<td>500</td>
<td>30</td>
<td>824</td>
<td>45.4</td>
<td>1.66E+14</td>
</tr>
<tr>
<td>550</td>
<td>0</td>
<td>340</td>
<td>64.3</td>
<td>2.85E+14</td>
</tr>
<tr>
<td>550</td>
<td>5</td>
<td>280</td>
<td>65</td>
<td>3.42E+14</td>
</tr>
<tr>
<td>550</td>
<td>15</td>
<td>154</td>
<td>47</td>
<td>8.64E+14</td>
</tr>
<tr>
<td>550</td>
<td>15</td>
<td>331</td>
<td>57.5</td>
<td>3.27E+14</td>
</tr>
<tr>
<td>550</td>
<td>15</td>
<td>154.7</td>
<td>59.14</td>
<td>1.40E+15</td>
</tr>
<tr>
<td>550</td>
<td>15</td>
<td>208</td>
<td>40</td>
<td>7.58E+14</td>
</tr>
<tr>
<td>550</td>
<td>15</td>
<td>238</td>
<td>51.7</td>
<td>5.06E+14</td>
</tr>
<tr>
<td>550</td>
<td>20</td>
<td>88.3</td>
<td>30.5</td>
<td>2.32E+15</td>
</tr>
<tr>
<td>550</td>
<td>20</td>
<td>171</td>
<td>41</td>
<td>8.90E+14</td>
</tr>
<tr>
<td>550</td>
<td>20</td>
<td>268</td>
<td>40.5</td>
<td>5.72E+14</td>
</tr>
<tr>
<td>550</td>
<td>20</td>
<td>118</td>
<td>37.9</td>
<td>1.40E+15</td>
</tr>
<tr>
<td>550</td>
<td>20</td>
<td>168</td>
<td>35.9</td>
<td>1.00E+15</td>
</tr>
<tr>
<td>550</td>
<td>20</td>
<td>305</td>
<td>28.7</td>
<td>7.13E+14</td>
</tr>
<tr>
<td>550</td>
<td>30</td>
<td>115</td>
<td>47</td>
<td>1.15E+15</td>
</tr>
</tbody>
</table>

### Table 2.4: Effect of Annealing Temperature on Electrical Properties

<table>
<thead>
<tr>
<th>$T_a$, °C</th>
<th>$t_a$, min.</th>
<th>Sh. Res., Ω/□</th>
<th>$\mu_h$, cm²/(V s)</th>
<th>Conc., cm⁻²</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>330</td>
<td>66.2</td>
<td>2.9 × 10¹⁴</td>
<td></td>
</tr>
<tr>
<td>450</td>
<td>30</td>
<td>332</td>
<td>63.4</td>
<td>3.0 × 10¹⁴</td>
</tr>
<tr>
<td>500</td>
<td>30</td>
<td>310</td>
<td>57</td>
<td>3.5 × 10¹⁴</td>
</tr>
<tr>
<td>550</td>
<td>10</td>
<td>267</td>
<td>63.8</td>
<td>3.7 × 10¹⁴</td>
</tr>
</tbody>
</table>
Table 2.5: Effect of Annealing Time on Electrical Properties

<table>
<thead>
<tr>
<th>$t_a$, min.</th>
<th>Sh. Res., $\Omega/\square$</th>
<th>$\mu_h$, cm$^2$/V s</th>
<th>Conc., cm$^{-2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>155</td>
<td>59.1</td>
<td>$1.4 \times 10^{15}$</td>
</tr>
<tr>
<td>15</td>
<td>121</td>
<td>16.4</td>
<td>$3.2 \times 10^{15}$</td>
</tr>
<tr>
<td>30</td>
<td>121</td>
<td>17.2</td>
<td>$3.0 \times 10^{15}$</td>
</tr>
<tr>
<td>45</td>
<td>117</td>
<td>22</td>
<td>$2.4 \times 10^{15}$</td>
</tr>
</tbody>
</table>

To understand the annealing rate, we diffused a chip at 550°C for 15 minutes, and then annealed it at 500 °C for three consecutive treatments of 15 minutes. The results of this experiment are shared in Table 2.5. After just 15 minutes, the carrier concentration has increased by 130% and the mobility has decreased by 72%, strongly suggesting an increase in dopant activation. For longer diffusion times, the mobility recovers and the carrier concentration decreases, possibly due to vaporizing zinc. Reaching such low mobilities is interesting because, in the literature, it corresponds to doping densities near $1 \times 10^{21}$ cm$^{-3}$ which is in turn similar to the zinc solubility predicted in the previous section [237].

Chronologically, we figured out this technique to increase dopant activation before realizing that such high carrier concentrations could be detrimental to device performance through optical effects, so we used it on all subsequent diffusions and have data on its effects for 19 different samples. The ratio of electrical properties before and after annealing is shared in Fig. 2.15. We see that the reduction in sheet resistance and mobility and increase in carrier concentration is a robust effect, with each changing by -14%, -13%, and +44%, on average.

We also experimented with this technique on InP. In a single diffusion into semi-insulating InP at 500°C for 30 minutes, we found a sheet resistance of 97 $\Omega/\square$, a mobility of 29 cm$^2$V$^{-1}$s$^{-1}$, a and sheet carrier concentration of $2.2 \times 10^{15}$ cm$^{-2}$. After annealing at 450°C for 15 minutes, these changed to 91 $\Omega/\square$, 18.4 cm$^2$V$^{-1}$s$^{-1}$, and $3.7 \times 10^{15}$ cm$^{-2}$. This demonstrates that the technique works similarly for InP.

There are a large number of ways that we would like to improve and better characterize this diffusion technique. To improve run-to-run consistency, we would like to remove all variability associated with the sealing and atmosphere inside the quartz boat. The easiest way to do this would be in a vacuum
Figure 2.15: Box and whisker plots for effect of post-diffusion annealing on Hall measurement results. Annealing reduces sheet resistance and mobility while increasing carrier concentration.

furnace where the air is evacuated and then back-filled with forming gas. We could also make sure the atmosphere around the sample is immersed in zinc by using a box for the diffusion with a gauntlet style seal so that air and forming gas can flow in and out, but the zinc is unlikely to diffuse out. A rough schematic of what this could look like is presented in Fig. 2.16. In this idealization, the mineral wool is a semi-permeable seal that can allow gas to flow under high pressure differentials like evacuation and backfill, but will block zinc from diffusing from the box. A graphite lid reliably holds the wool in place.

It would also be desirable to explore more dilute zinc sources. With our diffusion source, the zinc activity is roughly 20%, which means we are still on the far right side of Fig. 2.13 [184]. For the zinc solubility to be around $5 \times 10^{19} \text{cm}^{-3}$, we need zinc activities of 0.1%, a factor of 200 less.

The techniques that would be most helpful for characterizing the diffusion are SIMS, Polaron profiling, and TEM. SIMS would tell us the zinc diffusion profile and if any impurities are introduced during the diffusion. Polaron would provide the electrically active zinc concentration. Comparing SIMS and
Figure 2.16: Illustration of improved diffusion source that will allow improved uniformity and consistency by better controlling the atmosphere around the sample.

Polaron before and after a post diffusion anneal would provide strong evidence for or against enhanced dopant activation. TEM of samples before and after annealing would also check if the diffusion generates large-scale defects, and if annealing removes them.

To summarize, we have developed a new, simple, reliable technique for zinc diffusion doping of GaAs and InP. We readily achieve sheet resistances less than 1 kΩ/□. We have also found a process for post-diffusion annealing that enhances dopant activation. There is ample remaining work to optimize and characterize the diffusion process, but it plainly works.

Cu/Zn Source

A second diffusion source we tried was a solid copper/zinc source, embodied as a piece of 200-series brass (McMaster Carr). We thought brass would be interesting for a few reasons. First, brass is widely available. Second, it is resistant to oxidation. Third, by being solid at the diffusion temperature we could keep the GaAs and source in close proximity. The diffusion consisted of placing the GaAs to be doped face-to-face with the brass. A photo of the configurations we tried is shown in Fig. 2.17. From left to right, there is a piece
of GaAs a few mm away from the brass, with a piece of carbon paper between the brass and GaAs, and directly against the brass. With a diffusion at 550 °C for thirty minutes, we found that the piece near the brass was undoped, the piece near the carbon space had a sheet resistance of 558 Ω/□, and touching the brass was 735 Ω/□. The piece touching the brass had a degraded surface, but the piece with the carbon spacer looked good. Unfortunately, we did these experiments before having easy access to Hall measurements. We also did SIMS on a chip diffused like the carbon spacer one above. The reason we especially wanted to do SIMS on the brass source is that copper is a fast diffuser and electrically active in GaAs, so the diffusion source could be unintentionally doping the semiconductor. The results of the SIMS measurement are shown in Fig. 2.18 [84]. We see the classic flat diffusion profile for zinc and an extremely high concentration of copper. Based on all the copper, we stopped working with brass sources to simplify the considered metallurgy.

**ZnO Source**

The other major zinc source we tried was ZnO. This was inspired by [246]. We found a ZnO source attractive because it eliminates concerns about the diffusion source oxidizing because it, itself, is an oxide, and the GaAs can be held directly against the source. The mechanism is that hydrogen in the diffusion ambient reacts with the ZnO according to:

\[
\text{ZnO(s)} + \text{H}_2 \rightleftharpoons \text{H}_2\text{O(g)} + \text{Zn(g)},
\]

(2.26)

liberating Zn to dope the GaAs. In the literature, they worked at much higher temperatures than we are interested in, so we checked if the thermodynamics
of using this reaction makes sense. To use Eqn. 2.23, we need to calculate the equilibrium zinc pressure. The change in Gibbs free energy of the above reaction is:

\[ \Delta G^0 = 239 \, \text{kJ mol}^{-1} - T \times 175.4 \, \text{J mol}^{-1} \, \text{K}^{-1}. \tag{2.27} \]

Using the fact that water and zinc are liberated in equal quantities, the equilibrium constant is:

\[ \frac{p_{Zn}^2}{p_{H_2}} = e^{-\frac{239 \, \text{kJ mol}^{-1}}{R T} + \frac{175.4 \, \text{J mol}^{-1} \, \text{K}^{-1}}{R} \, T}. \tag{2.28} \]

To calculate the zinc activity, we use that the vapor pressure of liquid zinc is:

\[ p_{0, Zn}, \text{Pa} = 10^{10.384 - 6286/T}. \tag{2.29} \]

Assuming the diffusion ambient is 10% hydrogen, then \( p_{H_2} = 0.1 \, \text{bar} \). Calculating \( a_{Zn} = \frac{p_{Zn}}{p_{0, Zn}} \), we find that it is nearly temperature invariant over 400 °C to 600 °C with a value of \( a_{Zn} = 0.002 \) which corresponds to a doping concentration of \( 1 \times 10^{20} \text{ cm}^{-3} \). Admittedly, we did not take into account how the liberated water reacts with the GaAs so this calculation should be taken as an argument for plausibility rather than a strong quantitative prediction.
We can also estimate if the kinetics are reasonable. The reaction rate in these conditions is given approximately by [142]:

\[
R = 1.27 \times 10^{-2} \text{ cm s}^{-1} \times C_{\text{H}_2} \times e^{-eV/kB(1/T-1/700\,\text{°C})},
\]

where \(C_{\text{H}_2}\) is the \(\text{H}_2\) particle density which we calculate using the ideal gas law.

In Fig. 2.19, we plot the time required to release a zinc dose of \(5 \times 10^{14} \text{ cm}^{-2}\). Above 450 °C, the time is less than ten minutes, suggesting the plausibility of this process.

To actually do the diffusion, we first had to make a ZnO film on a silicon wafer. At the time the equipment we would normally use to do this was down, so we found a technique that was both so unusual and successful that we decided it was worth including. Conceptually, the process was pyrolytic CVD of zinc acetate, and we were able to do this process using only a hotplate.

The first step was grinding ZnO chunks (Plasmaterials) into a fine powder using a mortar and pestle. We then dissolved the powder in hot acetic acid (VWR) and then dried it to make zinc acetate crystals. Once we had the crystals, we put a small dish with zinc acetate and the silicon wafer to be
coated under a Pyrex dish on a hotplate. We started deposition by turning the hotplate to 400°C. Zinc acetate has a high vapor pressure and when its vapors contacted the hot silicon wafer it decomposed into ZnO. A picture of the setup mid-deposition is shown in Fig. 2.20. Thin film interference effects are visible as ZnO grows on the silicon wafer.

To run the diffusion, we just put the GaAs face-down on the ZnO and annealed it under forming gas. We did not do many diffusion this way, but as an example result, diffusing at 550°C for 30 minutes resulted in 40 Ω/□ sheet resistance. In hindsight, we should have pursued this further because of its much lower zinc activity than the metal sources, but we had to keep remaking ZnO films because they would noticeably etch during each diffusion. This is shown in Fig. 2.21, where we can see the shadows of where GaAs chips were during the diffusion because they prevented some ZnO loss. We also did not pursue this source because we were not sure about how the evolved water affects all the relevant thermodynamics and chemistry.

To review, we have demonstrated three techniques for the low-temperature zinc diffusion doping of GaAs that use a: molten zinc source, a solid zinc source, and a solid zinc oxide source. All sources work, but the solid zinc source had large problems with copper contamination. We also have demonstrated how post-diffusion annealing can enhance the activation of the zinc dopants. There is lots of work remaining to fully characterize the diffusion techniques in terms of doping profiles, activation, unintentional impurities, and defect generation.
Figure 2.21: Picture of ZnO film after being used for GaAs diffusion doping. The regions covered by GaAs are clearly visible, demonstrating the considerable decomposition of the film during the process.

However, we considered being able to make p-type layers at all enough of a success that we turned our attention to solar cell fabrication.

2.6 Zinc Diffusion Doped GaAs Solar Cells

This section will cover our progress in turning diffusion doped chips into solar cells and characterizing their performance. We will go through each fabrication step and try to highlight points where innocuous corner-cutting will instantly cost lots of performance. After covering the basics of fabrication, we will then cover the more complex issues of:

- front metal grid design,
- surface treatments,
- anti reflection coatings,
- and substrate choice.

With substrate choice, in particular, we find new relevance to old issues with GaAs metallurgy. Over the course of these experiments, our goal was to achieve a good solar cell efficiency rather than comprehensively characterize the fabrication parameter space, so there will be some loose ends and we will try to explain our rationale in changes between generations of devices.
Minimal Cell Fabrication

To turn a diffusion-doped wafer into a solar cell we, at a bare minimum, need to make separate electrical contacts to the n-type a p-type regions. For reference, our standard process for this is outlined in Fig 2.22. We first diffusion dope the GaAs using one of the techniques described in the previous section.

Rearside Etch

To make a good contact to the wafer bulk, we first had to etch the diffused material off the backside of the wafer which we did by pipetteing enough etchant on to the backside of the sample to coat it, letting the etch proceed, and then quickly rinsing the sample with water to remove the etchant and make sure it does not have an opportunity to attack the front of the chip. The typical etchant we used was 3 : 1 : 50 H₃PO₄ : H₂O₂ : H₂O by volume [175] for five minutes. In general, we have been pleased with the etchants described in [175].

Back Metal

We wanted to do the rear metal as quickly as possible in our fabrication process because heat treatments are unavoidable for making contacts to n-GaAs and we wanted to make sure it would not disturb other parts of the device. With this in mind, we also wanted a metallization that would work at temperatures well below the doping temperature so that it would be unlikely to perturb the diffused region. The default ohmic contact to n-GaAs is AuGeNi which
requires annealing at at least 400°C, so we rejected it outright [91]. It is also possible to make ohmic contacts below 200°C, but it requires palladium and takes several hours, which would be very inconvenient [269]. Instead, we decided use a silver/tin metallurgy [203]. Attractively, this process works at 350°C in a few minutes and only requires silver instead of extremely expensive gold or palladium.

To make the contacts, if the chip had been rearside etched within a few minutes, we loaded it directly into the evaporator. If it had been sitting in air, we first etched the native oxide by dipping the chip in concentrated ammonium hydroxide for a few moments before rinsing it with water and drying it with N₂. In the evaporator we sequentially deposited 50 nm of silver followed by 50 nm of tin by ebeam evaporation. We then annealed the contact at 350°C for five minutes on a hot plate in a glove box.

**Front Metal**

We typically made front metal contacts using liftoff of nLof-2020 photoresist in N-methylpyrrolidinone (NMP). During early proof-of-principle experiments we simply evaporated metal though holes in a shadow mask. The metal itself was typically 0.5 nm of Ti as an adhesion layer followed by 250 nm of Au. There is no electrical or semiconductor metallurgy reason why we usually used gold. It was just because we knew it would survive any experiments we did to the frontside of the cell. We used 250 nm because it seemed to be enough to survive probing reliably. In initial experiments we used silver to mirror silicon PV, but found that it would etch in phosphoric-acid based solutions we used for emitter thinning. We also found aluminum, chromium, titanium, and copper all made acceptable contacts to the diffusion doped p-GaAs, highlighting its high carrier concentration.

**Mesa Etch**

The final step in making the cell electrically active is etching a trench between cells and between cells and the edge of the chip. This eliminates shunts and allows us to fabricate a large number of devices on each chip to help get statistics about device performance. To help visualize what this looks like, we share a photo of a finished device in Fig. 2.23. In the image, the metal, an isolated cell, and the trench have been given false colors. We can see a number
Figure 2.23: False color picture of solar cell with mesa etch and metal finished.

of other adjacent cells. The black finger coming from the top is a tungsten probe. We defined the etched regions using photolithography, and did the mesa etch itself by pipetting the same etchant as the backside etch onto the front of the sample and etching for three minutes followed by rinsing under flowing water. We pipetted the etchant to make sure it didn’t attack the back contact.

In early experiments when it still was not clear if this diffusion doping scheme would work we, did cell singulation by scribing and cleaving instead of lithography. Intuitively, it seems like this should be fine because the cleaving results in mirror-smooth surfaces and it saves a lot of time if you do not care about precisely knowing the area of the device. However, we found that cells singulated by cleaving lost typically lost 100 mV of $V_{oc}$ compared to cells isolated lithographically. If we did not figure this out by trying one batch with etched isolation, this project may not have happened due to the uninspiring $V_{oc}$. To demonstrate this effect, we show the photovoltaic curves of two cells fabricated
Figure 2.24: Comparison of cell performance depending on if the cells were singulated by cleaving or mesa etching. Cleaving badly reduces cell voltage.

by ZnO source diffusion at 550°C for thirty minutes in Fig. 2.24. The cell singulated by etching has 110 mV better Voc. This is likely due to the scribing operation injecting a large number of dislocations around the perimeter of the device [274].

Optimizing Fabrication
Having described the basic process for fabricating cells, we will now cover the design aspects we put more thought into optimizing.

Front Grid Design
Designing the front-grid contacts for a solar cell is a complex optimization problem with interplay between semiconductor fabrication, metallization techniques, and economics. Before starting to design the metal, we first needed an idea of the cell size. On one hand, we want small cells so that we can make lots of devices per sample to test, but, on the other hand, small cells have to deal with parasitic edge recombination currents that can misleadingly handicap cell performance. The perimeter recombination of GaAs solar cells has been previously characterized and is \( J_{0p} = 2.5 \times 10^{-13} \text{ A cm}^{-1} \) with ideality factor 2 [242]. Assuming a \( J_{sc} \) of 30 mA cm\(^{-2}\), the edge limited \( V_{oc} \) of a square solar cell is:

\[
V_{oc} = \frac{2k_B T}{q} \ln \left( \frac{J_{sc} \ast l}{4J_{0p}} \right),
\]  
(2.31)
Figure 2.25: $V_{oc}$ as a function of edge length if all recombination is due to the edge.

where $l$ is the edge length. This is plotted in Fig. 2.25. From section 2.4, we found that the maximum expected $V_{oc}$ is around 1 V, so devices with edges longer than 200 $\mu$m will not be artificially limited by edge recombination. We ended up using edge-lengths of 350, 850, and 1350 $\mu$m, where the upper bound was set by wanting several of the largest cell per chip that was typically 5 mm on a side.

To calculate the thickness of metal needed to effectively contact these cells, we need to consider the sheet resistance, operating point, and cell size. The largest cell will wave the largest losses, so we consider only that case. In Fig. 2.26, we illustrate the simple model we used for calculating resistive losses. In the figure, $L_m$ is the length of the cell/metal strip, $W_m$ is the width of the metal contact, and the external current and voltage are defined at the end of the metal. The arrows depict the current flow, with the arrow length increasing along the x axis depicts the current density increasing towards the metal, and the arrow density increasing with $y$ depicting the increasing current density as it heads to the external circuit.

The current along the metal is $J(y) = (J_{ext}/L_m) \cdot y$, linearly increasing from
Figure 2.26: Electrical model for calculating cell size and metal dimensions for minimizing ohmic conduction losses.
zero at the far end of the cell to the external load. The power loss in the metal is then:
\[
P_{\text{metal}} = \int_0^{L_m} \frac{\rho_{\text{metal}}(J_{\text{ext}} \cdot y/L_m)^2}{W_m \cdot h_m} \, dy = \frac{\rho_{\text{metal}} J_{\text{ext}}^2 L_m}{3 W_m h_m},
\]
where \(\rho_{\text{metal}}\) is the metal resistivity, and \(h_m\) is the height of the metal. It is a similar calculation for the losses due to conduction along the surface of the semiconductor. The current as a function of \(x\) for the left-hand side of the cell is \(J(x) = J_{\text{ext}} \cdot L_m \cdot x/(2L_m)\), so the power dissipated is:

\[
P_{\text{semi}} = \int_0^{L_m/2} \frac{\rho_{\square}(J_{\text{ext}} \cdot x)^2}{L_m^3} \, dy = \frac{\rho_{\square} J_{\text{ext}}^2}{24}.
\]

To underestimate the amount of metal necessary to minimize resistive losses, let’s assume a worst-case scenario where the cell is operating at 30 mA cm\(^{-2}\) and 20% efficiency so that the PV power is 360 \(\mu\)W and the external current is 550 \(\mu\)A. For the metal, we chose \(W_m = 50 \mu\)m because that is reliably fabricatable and \(\rho_{\text{metal}} = 2.2 \times 10^{-8} \mu\Omega\) m (the resistivity of gold). The power loss is then 0.06 \(\mu\)W \(\mu\)m, normalized to the gold thickness. Therefore, with the 250 nm of gold we typically use, the metal conduction losses are negligible.

With a sheet resistance of 100 \(\Omega/\square\), the power loss is 1.26 \(\mu\)W, or less than 1%. Something we anticipated was being interested in cells with high sheet resistances (> 1 k\(\Omega/\square\)), in which case the power loss rises to 4%, which is why we added the cross to the largest cell contact design, which can be seen in Fig. 2.23. With the cross-piece, the shadowing fraction is 6%, which is on par with commercial silicon cells [226]. The contact design can be further improved by distributed circuit modeling and iterative optimization, but we think the current designs are good enough for proof-of-principle devices.

### Substrate Choice

The two most important characteristics of the substrate are its doping, \(N_D\), and minority carrier diffusion length \(L_p\). Both of these enter directly into the diffusion current calculation, and therefore influence the \(V_{oc}\), and the minority carrier diffusion length is important in determining the \(J_{sc}\). To understand the importance of \(L_p\) and other device parameters, we introduce the following model for the spectral response:

\[
\text{IQE}(\lambda) = \text{IQE}_p(\lambda) + \text{IQE}_D(\lambda) + \text{IQE}_n(\lambda),
\]

(2.34)
where IQE refers to the internal quantum efficiency, or, given that a photon is absorbed by the device, what is the probability that it contributes to the photocurrent, and the subscripts refer to the individual IQE’s of the p-type surface region, the depletion region, and the n-type bulk. For the moment, we will focus on IQE\textsubscript{n}:

$$\text{IQE}_n(\lambda) = \exp\left(-\alpha(\lambda)(W_p + W_D)\right) \frac{1}{1 + \left(\frac{L_p \alpha(\lambda)}{L_p \alpha(\lambda)}\right)^{-1}}, \quad (2.35)$$

where $\alpha(\lambda)$ is the wavelength dependent absorption, the first factor derates for the amount of light that reaches the bulk after traveling through the surface and depletion region, and the second is the hole collection probability. For completeness, the generated photocurrent is:

$$J_{sc} = \int \Phi(\lambda) \ast \text{IQE}(\lambda) d\lambda, \quad (2.36)$$

where $\Phi(\lambda)$ is the incident spectral photocurrent density. The photocurrent produced in the bulk as a function of $L_p$ under AM1.5G illumination and assuming $W_p + W_D = 100 \text{ nm}$ is plotted in Fig. 2.27. We can see that there are diminishing returns around 10 $\mu\text{m}$ at 18.5 mA cm$^{-2}$ out of a theoretical limit of 21 mA cm$^{-2}$. The two bulk recombination mechanisms that determine $L_p$ are radiative recombination and defect-mediated recombination. Using the doping/mobility relation from [237] and a radiative recombination coefficient of $1.7 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$, this is achieved at doping less than $4 \times 10^{16} \text{ cm}^{-3}$.

Based on this calculation, we made our first cells using commercially available (MTI Corp) wafers doped in the mid-16’s and grown by vertical gradient freeze (VGF). To analyze the performance of these cells, we measured the cell IQE and fit it. For the other contributions to the total IQE model, we used:

$$\text{IQE}_D = \exp(-\alpha(\lambda)W_p)(1 - \exp(-\alpha(\lambda)W_d)), \quad (2.37)$$

so all carriers generated in the depletion region are collected, and:

$$\text{IQE}_p = \left(\frac{\alpha L_n}{(\alpha L_n)^2 - 1}\right) \times \left(\frac{(K + \alpha L_n) - \exp(-\alpha W_p)(K \cosh(W_p/L_n) + \sinh(W_p/L_n))}{K \sinh(W_p/L_n) + \cosh(W_p/L_n)} - \frac{\alpha L_n \exp(-\alpha W_p)}{\alpha L_n}\right), \quad (2.38)$$
where $L_n$ is the minority electron diffusion length and $K \equiv S_n L_n / D_n$, where $S_n$ is the surface recombination velocity and $D_n$ is the electron diffusion coefficient. The IQE of a cell made by the ZnO method at $550 \, ^\circ\text{C}$ for 30 minutes and a fit to the above model is presented in Fig. 2.28a, and one by Zn/Sn diffusion at $550 \, ^\circ\text{C}$ for 20 minutes followed by an anneal at $500 \, ^\circ\text{C}$ for 15 minutes is shown in part b. According to the fit, the hole diffusion lengths for both are 1 $\mu$m, corresponding to a defect-mediated recombination lifetime of only about 1 ns. Even though the hole lifetime is only 1 ns, the hole-injection diffusion current limited $V_{oc}$ is still greater than 980 mV, which is comparable to the maximum $V_{oc}$ allowed by electron injection to the surface or Urbach-mediated emission.

We hypothesized that this extremely short lifetime was due to the EL2 defect in the wafers. EL2 in an arsenic antisite defect produced during cooling of GaAs [104]. It is typically present in melt-grown wafers at a concentration greater than $1 \times 10^{16} \, \text{cm}^{-3}$, and has a hole capture cross section of $5.5 \times 10^{15} \, \text{cm}^2$, leading to a lifetime of 0.9 ns, consistent with our diffusion length measurement [296]. Despite it badly limiting the hole-lifetime, device fabricators actually try
Figure 2.28: IQE and model fits for solar cells made using commercially available n-GaAs wafers by either the ZnO or Zn/Sn diffusion methods. The extracted hole diffusion length is 1 µm.

to maximize its concentration. This is because it is a midgap donor and when it has a concentration greater than shallow acceptors, the material becomes semi-insulating with the Fermi level pinned midgap. This is desirable for RF circuit design because it reduces parasitic capacitances and provides for easy device isolation [50]. It is also not a problem for common optoelectronic applications because the devices layers are typically grown by epitaxy and the substrate only functions as some combination of mechanical handle, electrical contact, and growth template.

There are several techniques for growing EL2-free GaAs or removing it. Because EL2 forms during crystal cooling when there is a supersaturation of $V_{Ga}$ and $As_i$ simply quickly cooling the crystal will avoid its formation. This has been found in annealing studies where heating a crystal to $1200^\circ C$ to dissociate the EL2 and then quenching the crystal eliminates the EL2 [187] by not allowing enough time for the vacancies and interstitials to find each other. EL2 can also be eliminated by annealing at low temperatures in an arsenic depleted ambient [32]. N-type doping also helps with suppressing EL2 by increasing the solubility of gallium vacancies [131]. Unfortunately the necessary electron concentration then limits the diffusion length through radiative recombination. GaAs grown by LPE or CVD also typically has low concentrations of EL2. Notably, the HVPE group at NREL has demonstrated GaAs grown faster than $300 \mu m h^{-1}$ with EL2 concentrations of $3 \times 10^{14} cm^{-3}$, which is compatible with our desired material properties [165]. A second way to reduce EL2 is by changing the melt composition that bulk GaAs is grown from because gallium rich melts will decrease the quantity of both gallium vacancies
Figure 2.29: Schematic of vapor controlled Czochralski growth tool for making high-quality ingots of n-type GaAs. The lack of an encapsulant above the melt allows dynamically controlling the arsenic vapor pressure via the separate heated arsenic area. Reprinted/adapted by permission from Elsevier, Growth of GaAs crystals from Ga-rich melts by the VCz method without liquid encapsulation by F. M. Kiessling, P. Rudolph, M. Neubert, U. Juda, M. Naumann, W. Ulrici, 2004.

and arsenic interstitials. A great example of this is in [119] where, by doing a Czochralski growth with a hot wall system, they were able to maintain arsenic vapor equilibrium between the melt and solid arsenic held at a lower temperature. Their apparatus is illustrated in Fig. 2.29. The key features are the hot wall pressure vessel and the lack of B$_2$O$_3$ encapsulant. In their growth ‘V2,’ with the arsenic at 590°C and using a quartz crucible that provides silicon to dope the crystal n-type, the EL2 concentration is below their detection limit.

We contacted the authors of [119] and [165] and they were gracious enough to
Figure 2.30: IQEs and model fits for cells made using either GaAs grown by HVPE or GaAs grown by vCZ with a low arsenic temperature. The extracted hole diffusion-lengths are far-superior to the commercial bulk material.

provide us with some of their low EL2 GaAs. The material from NREL was 15 µm of HVPE n-type GaAs with a donor concentration of $5 \times 10^{16} \text{cm}^{-3}$. The material from IKZ-Berlin was a 1 cm$^2$ chip taken from ingot V2 which has an electron concentration of $2.6 \times 10^{16} \text{cm}^{-3}$.

The IQEs and their model fits for these two devices that were fabricated by the Zn/Sn method at 550°C for 20 minutes followed by an anneal at 500°C for 15 minutes are shown in Fig. 2.30. A qualitative difference between these IQEs and those in Fig. 2.28 is that the long wavelength response stays high until the bandgap instead of falling off at much shorter wavelength. This is quantified from the fits where the hole diffusion length of the HVPE material is 11 µm and that of the vCZ material is 7.4 µm, corresponding to one-hundred and fifty-times better hole lifetimes than the traditionally grown bulk material.

We think the vCZ material is particularly exciting because its minority carrier properties are similar to vapor-phase epitaxial material and it was grown as a bulk ingot at 5 mm h$^{-1}$, demonstrating the possibility of growing high quality GaAs, quickly, using equipment far less complex than a CVD. In private communication with the authors, they were optimistic that a similar setup could be devised for VGF growth, but such a tool does not currently exist. We would also be interested to use GaAs annealed in an arsenic-depleted atmosphere, but we are not currently equipped for that experiment.

**Surface Treatments**

In Fig. 2.30, we can see that the devices have good long-wavelength response; not falling off until close to the bandgap, but the blue-response is lackluster,
Table 2.6: Device parameters extracted from IQE measurements including the emitter thickness, minority electron diffusion length, dimensionless surface recombination parameter and emitter sheet resistance.

<table>
<thead>
<tr>
<th>Sample</th>
<th>$W_p$, µm</th>
<th>$L_n$, µm</th>
<th>$K$</th>
<th>Sh. Res., Ω/□</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commercial</td>
<td>0.22</td>
<td>3.7</td>
<td>186</td>
<td>88</td>
</tr>
<tr>
<td>HVPE</td>
<td>0.14</td>
<td>4.0</td>
<td>130</td>
<td>135</td>
</tr>
<tr>
<td>vCZ</td>
<td>0.19</td>
<td>3.9</td>
<td>135</td>
<td>150</td>
</tr>
</tbody>
</table>

similar to even the low quality material in Fig. 2.28. This is due to surface recombination and the finite diffusion length of electrons at the surface. To understand how we can do better, the fitted diffused layer thickness, electron diffusion length, $K$, and measured sheet resistance of the Zn/Sn doped devices are shown in Table 2.6. Despite the different growth techniques used for each crystal, the fitted surface properties are remarkably similar with a diffusion depth of 200 nm and hole diffusion length of 4 µm. The long electron diffusion length is surprising and there is not consensus in the literature on electron diffusion length at the high doping we have [88, 278]. Unfortunately, the model was also able to provide satisfactory fits by presuming $L_n$ as short as 300 nm and only changing the diffusion depth by a few nm. Computing $K$ into something physical, the $K/L_n$ ratio is robust to fitting and assuming the electron diffusion coefficient is around 50 cm$^2$s$^{-1}$, the SRV is then $1 \times 10^7$ cm s$^{-1}$, which is consistent with unpassivated GaAs [152, 87].

To understand the opportunities and the limits of modifying the surface, we plot the photocurrent for a device with $L_p = 10$ µm, $L_n = 1$ µm, and $W_D = 200$ nm as a function of $W_p$ and $K$ in Fig. 2.31. The current devices are in the top-right of this plot. At the bottom left is 30.6 mA cm$^{-2}$. With these devices having sheet resistances of 100 Ω/□ and the metal having been designed for up to 1 kΩ/□, we can etch a lot of the emitter to improve photocurrent before series resistance becomes large.

To do emitter thinning practically, we prepared a very slow GaAs etchant and went back and forth between the solar simulator and the wet-chem hood over and over until performance saturated. The two etchants we used were the mesa etch diluted 10:1, and an alkaline etch that used 5.6 g of KOH, 2 mL of H$_2$O$_2$ and 100 mL of H$_2$O. Because the temperature in the lab noticeably varied day-to-day we did not bother to characterize the etch rates. To illustrate how emitter thinning affects performance, we show the performance versus
etch time for an HVPE chip doped by the Zn/Sn method at 500°C for thirty minutes followed by an anneal at 550°C for ten minutes in Fig. 2.32. The post-anneal sheet resistance of this chip was 267 Ω/□.

As expected, the $J_{sc}$ increases with etch time due to less absorption in the surface. However, the $V_{oc}$ decreases with etch time due to short-emitter diode effects. At the last data point there is a large drop in FF, likely do the increasing sheet resistance as the emitter is thinned. We terminated the experiment when the efficiency stayed constant after the last etch. The IQE fit and model for this device assuming $L_p = 4 \mu$m is shown in Fig. 2.33a. According to the fit, the emitter thickness is now 72 nm and $K = 170$. 

Figure 2.31: Predicted $J_{sc}$ as a function of emitter thickness and dimensionless surface recombination parameter.
The next treatment we used to improve performance was surface passivation. This subject will be covered in more detail in the next chapter. We did the passivation by submerging the chip in a 10 mM solution of dithiothreitol (DTT) composed of a 4:1 by volume mixture of ethanol to ammonium hydroxide for five minutes. The reason we chose DTT is that it has hydroxyl groups pointing away from the surface to help with ARC deposition. The IQE of the chip after passivation is presented in Fig. 2.33b. Holding the emitter thickness constant, $K$ is now 72, reduced by a factor of 2. The PV performance before and after passivation is shown in Table 2.7. The passivation uniformly improved device performance; gaining an extra 1.2 mA cm$^{-2}$, recovering most of the lost voltage, some of the fill factor, and overall efficiency improving by over 8% relative. Over the course of all surface treatments performance improved by 22%. Something that needs a better explanation, and will become are recurring theme, is the relation between the measure $J_{sc}$ and the IQE. According to the IQE model, increasing the $J_{sc}$ by 6.3%, as the passivation has, requires $K$ reducing to 30, which is in tension with IQE.
Table 2.7: Summary of cell performance for as-diffused cell, after emitter thinning, and finally after surface passivation.

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Etched</th>
<th>Passivated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J_{sc}$, mA cm$^{-2}$</td>
<td>16.3</td>
<td>19.1</td>
<td>20.3</td>
</tr>
<tr>
<td>$V_{oc}$, mV</td>
<td>949</td>
<td>933</td>
<td>943</td>
</tr>
<tr>
<td>FF, %</td>
<td>82.3</td>
<td>80.3</td>
<td>81</td>
</tr>
<tr>
<td>Eff, %</td>
<td>12.7</td>
<td>14.3</td>
<td>15.5</td>
</tr>
</tbody>
</table>

Figure 2.33: IQE of solar cell before and after passivation demonstrating enhances short-wavelength response.

Anti Reflection Coating

The final tool we used to improve device performance was an anti reflection coating (ARC). We designed a dual-layer ARC using the optimization feature of PV Lighthouse’s OPAL 2 optics solver. We chose to use ZnS and SiO$_2$ as the ARC layers; ZnS because the sulfur may help to passivate the exposed GaAs and SiO$_2$ because we had previous experience depositing it. OPAL 2 concluded 60 nm of ZnS and 100 nm of SiO$_2$ would be optimal. We deposited the ZnS by thermal evaporation and the SiO$_2$ by electron beam evaporation. While unloading samples after ZnS evaporation we could smell H$_2$S, suggesting the existence of reactive sulfur species during the evaporation. With the ARC the cells were nearly black.

With the ARC, the highest voltage cell on the chip achieved 947 mV $V_{oc}$, 28.4 mA cm$^{-2}$ $J_{sc}$, 79% fill factor, and 21.2% efficiency. The highest efficiency cell reached 944 mV $V_{oc}$, 28.7 mA cm$^{-2}$ $J_{sc}$, 80% fill factor, and 21.6% efficiency, thus we’ve recovered almost all the voltage and have demonstrated $V_{oc}$’s significantly better than literature. The $J_{sc}$, though, presents something of a puzzle. The spectrum-integrated EQE of the ARC coated cells is
25.1 mA cm\(^{-2}\), much less than measured. One possibility is that the solar simulator was miscalibrated, but the intensity was calibrated with a GaAs solar cell with known EQE. Another possibility is that the EQE measurement setup is miscalibrated. To resolve this, we are sending cells to NREL for third-party validation as soon as possible. Regardless, these cells exceed the performance of historical attempts by all metrics.

**Champion Cell Fabrication**

To figure out how to make the best possible cell, at this point we had the \(V_{oc}\) of over fifty pre-surface-modification devices that we could correlate with Hall measurements. This was also around the point in time where we realized from the detailed-balance calculation that includes Urbach-tails that heavy doping may actually be detrimental to device performance. To demonstrate the deleterious effects of heavy doping, we plot the \(V_{oc}\) versus mobility, sheet carrier concentration, and sheet resistance in Fig. 2.35, including what eventually becomes our champion device. From all three, we can see that \(V_{oc}\) suffers under low sheet resistance, low mobility, and especially high sheet carrier concentration; all of which are expected to correlate with the volume carrier concentration. The \(V_{oc}\) correlations presented are consistent with our theory of Urbach absorption limiting performance at high doping. The low voltage, high carrier-concentration cells were all fabricated with the primary doping step temperature being 550\(^\circ\)C. One possible competing explanation for the missing voltage is that the diffusion doping introduced lots of defects that enhanced depletion region recombination. As evidence for the Urbach interpretation, we present the dark-IV and ideality factory of a cell from sample V2 that had a sheet carrier concentration of \(1.7 \times 10^{15}\) cm\(^{-2}\) and a \(V_{oc}\) of 912 mV in Fig. 2.34. From the dark-IV, we see basically exponential diode behavior with some photovoltage due to stray light in the room. The ideality factor as a function of bias is interesting because it never reaches 2, which would be indicative of depletion-region recombination, and it does not look like it approaches one, which would indicate diffusion current. Also, while not quantitative, we could see infrared emission from this device with an infrared viewer, suggesting a significant amount of the current is due to luminescence. One way to resolve if the current is due to Urbach mediated emission or other recombination mechanisms would be to study the emission intensity versus bias. If the emission ideality factor mirrors the electronic ideality factor,
then that is evidence for the current being due to Urbach mediated emission.

With these correlations and the Urbach theory in mind, we decided to make a batch of devices using an HVPE substrate and the Zn/Sn source at 450°C for an hour followed by an anneal at 400°C for thirty minutes. This chip produced the devices at 960 mV in Fig. 2.35. Of the cells on this chip, the highest $V_{oc}$ was 962 mV with a $J_{sc}$ of 29.5 mA cm$^{-2}$, FF of 82%, and efficiency of 23.1% and area of 1.82 mm$^2$. The highest efficiency device and a $V_{oc}$ of 960 mV, a $J_{sc}$ of 29.8 mA cm$^{-2}$, FF of 82%, and efficiency of 23.5% and area of 0.723 mm$^2$. On this chip, twelve devices were over 21% efficient, and eight were over 22.7% efficient, which speaks to the good yield of the process. The light and dark IV curves and the bias dependent ideality factor of the champion device are plotted in Fig. 2.36. In the light and dark IV curves we see standard diode behavior. The ideality factor is more interesting. At low bias it is two, which is indicative of depletion-region recombination, and then decreases at high bias before increasing again due to series resistance. Despite having a greater ideality factor than the device in Fig. 2.34, which would normally indicate much worse recombination, this device has 50 mV better $V_{oc}$, which is another piece of evidence for Urbach limited performance of the high carrier concentration devices. Another interesting feature is that this $V_{oc}$ is within 20 mV of the hole diffusion current limit derived in section 2.4 which means these devices are operating very close to their theoretical voltage limits and that qualitative improvements will require detailed analysis of the electron and
Figure 2.35: Scatter plots correlating cell $V_{oc}$ and Hall measurement results taken after diffusion doping.
Figure 2.36: Light and dark IV and ideality factor as a function of bias for the champion cell.

hole diffusion currents, the radiative emission from the heavily doped surface, and the diffusion current into the bulk.

We also present the post treatment IQE and the EQE of a cell with an ARC from the champion chip in Fig. 2.37. The IQE integrated $J_{sc}$ is 27.9 mA cm$^{-2}$, and the EQE integrated $J_{sc}$ is 24.1 mA cm$^{-2}$. Both of these are much less than the measured $J_{sc}$’s. The ARC disagreement likely is not as bad as face value because it was difficult to tell when the light source was on the cell due to the lack of reflections. If this is the case, we can simply multiply the EQE by a constant to bring it into agreement with IQE around 750 nm where the ARC should be perfect. Doing this the corrected $J_{sc}$ is 25.8 mA cm$^{-2}$. This continuing disagreement between $J_{sc}$ and EQE further highlights the need for better measurements that NREL is capable of. Regardless, though, with 960 mV $V_{oc}$ and 82% FF, we only need a $J_{sc}$ of 25.4 mA cm$^{-2}$ to reach 20% efficiency. Again, though, these devices are unambiguously better than literature attempts at diffusion doped GaAs solar cells regardless of these $J_{sc}$ calibration issues.

Speculative Techniques for AlGaAs Heterojunction Cells

Achieving qualitatively higher efficiency will require a window layer to minimize absorptive losses in the cell. This is currently achieved with AlGaAs or GaInP surface layers where their higher bandgap of 1.8 eV transmit visible light and they form a high quality passivating interface with the GaAs with SRV’s as low as 210 cm s$^{-1}$ and 1.8 cm s$^{-1}$, respectively [191]. Cells with AlGaAs windows surpassed 1 V $V_{oc}$ as early as 1984 [121, 276]. Unfortunately these surface layers are typically fabricated by epitaxial processes, making their
economic viability unlikely. Therefore, we began investigating non-epitaxial processes for making heterojunctions.

Ideally what we would like to do is take a GaAs wafer and by low-cost means convert the surface to AlGaAs. Conceptually this means swapping gallium at the surface for aluminum. An intuitive way to try and do this is evaporating aluminum on GaAs and annealing it. Indeed, this does produce an extremely thin GaAs/AlGaAs heterojunction as evidenced by an increase in the Schottky barrier height [182]. Unfortunately, the resulting layer is extremely thin and almost pure AlAs. This is due to the extremely small diffusion coefficient of Al in GaAs and the large heat of formation of AlAs driving its exclusive formation [29, 290, 275]. We also focus on AlGaAs because it is lattice-matched to GaAs across the entire composition range and so the material composition is more forgiving. In comparison, making a GaInP surface layer requires swapping both III and V atoms at a specific ratio to maintain lattice matching which I, frankly, have no clue how to even start on.

An interesting fact that inspired us to try AlGaAs fabrication anyways is that zinc diffusion can drive the interdiffusion of GaAs and AlAs, enhancing it by over ten order of magnitude [251, 138, 133]. The mechanism for the
enhancement is the zinc in-diffusion generating a supersaturation of gallium interstitials. To test this, we prepared a diffusion source consisting of ZnCl$_2$, AlCl$_3$, and NaCl powder consistent with the stoichiometry of Na$_3$ZnAlCl$_8$. By adding NaCl we reduce the vapor pressures of the ZnCl$_2$ and AlCl$_3$ [143]. The reaction we hoped would take place is a chloride ligand exchange between gallium and aluminum:

$$\text{AlCl}_3 + \text{GaAs} \iff \text{GaCl}_3 + \text{AlAs}, \quad (2.39)$$

followed by zinc-enhanced diffusion:

$$\text{Zn} + \text{AlAs} + \text{GaAs} \rightarrow \text{Zn:AlGaAs}. \quad (2.40)$$

By adding chlorides, we radically increase the number of possible relevant chemical species so we didn’t bother checking the thermodynamic viability of this process. Doing a diffusion at 500°C for 30 min into an NREL HVPE chip, we found a sheet resistance of 115 $\Omega/^2$, mobility 9.6 cm$^2$V$^{-1}$s$^{-1}$, and carrier concentration $5.6 \times 10^{16}$ cm$^{-2}$. Interestingly the mobility and carrier concentration are lower and higher, respectively, that what we achieved with the zinc/tin source. The mobility is even lower than the minimum quoted in literature [237]. Considering that AlGaAs has lower hole-mobility than GaAs this is encouraging. We then made cells with these chips. The light IV and dark ideality factor are plotted in Fig. 2.38. The two key points are the extremely low V$_{oc}$ and ideality factor stubbornly stuck between one and two, both of which are indicative of Urbach dominated recombination. This suggests the ZnCl$_2$ is an extremely potent doping source.

We also measured the EQE of this sample. The EQE of this sample is compared to cell prepared using the Zn/Sn method at the same diffusion temperature and time with no surface preparation in Fig. 2.39. Intriguingly, the AlCl$_3$ cell has far superior short-wavelength response, consistent with there being a wider-bandgap window layer. However, the long wavelength response is degraded. We suspect this is due to contamination from the AlCl$_3$ which had visible brown chunks in it.

We also studied the chip using EDS, which did detect aluminum, but the quantitative composition results are badly non-physical, so we do not report them. We also studied the chip using spectroscopic ellipsometry and were unable to fit the results using a stack of GaAs / AlGaAs / oxide, suggesting
Figure 2.38: IV curve of cell made with ZnCl$_2$, AlCl$_3$, NaCl diffusion source exhibiting low-voltage and unusual intermediate ideality factor.

Figure 2.39: Comparison of EQEs of possible AlGaAs cell and Zn/Sn cell doped at the same temperature and for the same duration. The cell with AlCl$_3$ has much better blue-response.
something interesting happened. In conclusion, we have found that using a
doping source based on ZnCl$_2$ and AlCl$_3$ increases the dopant incorporation
as evidenced by Hall and solar cell IV measurements and that the solar cell’s
short-wavelength response is greatly enhanced compared to a control device.
Opportunities for further work include using a much more dilute ZnCl$_2$ source
to increase the solar cell $V_{oc}$ and using more sensitive techniques like XPS and
TEM to understand how the AlCl$_3$ and ZnCl$_2$ are interacting with the GaAs
to achieve that excellent blue response.

2.7 Air Tolerant, Abundant Metal Ohmic Contacts

At this point, we have now discussed simple, low-temperature zinc diffusion
doping of III-V’s and using that diffusion doping to make solar cells. The next
issue we turn to is making ohmic contacts to n-type GaAs. The metallurgy
of GaAs contacts is thoroughly reviewed in [213]. The basic mechanism is
forming a heavily doped region near the surface so that the depletion region
becomes thin enough to tunnel through.

Assuming the GaAs to be contacted is not already degenerately doped, there
are, broadly speaking, two metallurgical techniques for achieving the doping.
First is the solid-phase regrowth mechanism, which is exploited by metalliza-
tions that use nickel or palladium and germanium or silicon, and is illustrated
in Fig. 2.40. In these contacts, nickel or palladium are deposited first followed
by the group IV element, as illustrated in part a. Upon annealing, the Ni or
Pd reacts with the GaAs to form a ternary compound. At the same time,
the Si or Ge is reacting with the Ni or Pd to form a germanide or a silicide,
shown in part b. When the ternary and the silicide/germanide meet, the sili-
cide or germanide is more stable so GaAs falls out of the solid solution and
regrows on the substrate, which is part c. The regrown GaAs is heavily doped
with the group-IV element, making it degenerately n-type and enabling ohmic
conduction.

A particularly noteworthy version of this contact uses Pd/Ge. The diffusion
coefficients in this system are high enough and the energetics favorable enough
that it’s possible to make ohmic contacts at only 150 °C [269]. This is also
the technique used by the ubiquitous AuGeNi contact which was discovered
before the regrowth mechanism was understood and is generally satisfactory
but has problems with spiking and thermal stability.
The second major class of contact strategies uses liquid-phase regrowth. In this paradigm, the contact metal melts and dissolves some GaAs into the melt and then during cooling the GaAs regrows and is heavily doped with elements from the melt. This mechanism is strongly related to the crystal growth technique of liquid-phase-epitaxy (LPE) [267]. In LPE, the crystal grower uses a liquid that is benign with respect to solid GaAs; usually Ga or Bi. By dissolving GaAs and dopants into the melt high quality semiconductor layers can be grown upon cooling. The most well-known contact in this paradigm is the Au/Ge contact, where a poorly controlled Au/Ge/Ga eutectic forms and then upon cooling heavily Ge doped GaAs precipitates out.

In thinking about the contact n-type contact for our low-cost GaAs solar cells, we decided the above approaches were all unsatisfactory because they relied on noble metals or difficult to deposit materials like Si or Ge which need to be deposited under high vacuum due to their reactivity. Tin, though, is uniquely interesting as a contact metal because it is a good solvent for LPE, produces n-type GaAs layers, is abundant, and easy to solder [192]. After some digging, we found a report for an extremely simple way to make ohmic contacts to n-GaAs using only metallic tin and a tin halide flux [219]. In this process, the halide flux reacts with oxides to allow the tin and GaAs to react and undergo the LPE process. Without the flux, the tin just balls up on the GaAs and is
ineffective. In [219], the authors made their samples using tin solder balls and annealing in an inert ambient. We thought a contact along these lines was extremely promising, but we did not know if we could combine soldering to a substrate and contacting into one step and how oxygen-tolerant the process in.

To test the process, we soldered a piece of commercial n-type GaAs doped to $5 \times 10^{16} \text{ cm}^{-3}$ to a copper foil using a sliver of pure tin and SnCl$_2$ dissolved in acetone as a flux on a hot-plate in air at 350°C. To test the contact, we then evaporated gold pads on the GaAs surface through a shadow mask to use as Schottky contacts. A picture of the device is in Fig. 2.41a. We can see the tin has readily wet the copper and GaAs and the gold contact pads. The current-voltage curve for this device is shared in Fig. 2.41b. In the figure, a one-sun photocurrent of 30 mA cm$^{-2}$ is highlighted. The two key points are that the one-sun current is still in the exponential part of the current, so that series resistance is not relevant for solar cells and that fitting the series resistance results in a resistance of 0.01 Ω cm$^2$, which is consistent with all the series resistance coming from the wafer.

The two advances we demonstrate are combining soldering and making ohmic contacts, and that these contacts can be made in air, which represent a considerable simplification over the state of the art using evaporated contacts. Two details we do not know about this contact are the uniformity on the backside of the wafer and how much GaAs was consumed to make it. From
the GaAs/Sn pseudobinary phase diagram, we can estimate that only a few nanometers were likely consumed [192].

At this pointm I have finished describing my scientific work that directly contributes to low-cost, high-efficiency diffusion doped GaAs solar cells. Next we will study if these make economic sense to fabricate at scale according to the scheme in 2.5.

2.8 Diffusion Doped GaAs Technoeconomic Analysis

Having advanced a number of enabling technologies for low-cost GaAs processing, the next natural question is if these solar cells can be economically viable. The process of figuring this out is called technoeconomic analysis (TEA). The overall number that we would want to calculate is the levelized cost of energy (LCOE). The LCOE is the cost of energy that a grid-operator cares about and is usually measured in units of $\text{¢ kW}^{-1}\text{h}^{-1}$. NREL has an excellent LCOE calculator that takes as input: the cost of the system (cell, module, inverter, bypass diodes, etc.), the energy yield (which depends on the installation location), operating costs, the degradation rate, project lifetime, and the discount rate (for calculating present values) [10]. In the analysis here, we will focus on the cell cost and also argue why other parts of the LCOE cost calculation favor III-V cells. Calculating a full LCOE will be the subject of future study because it requires data we do not currently have available, namely energy yield and degradation parameters.

In calculating the cell cost, the two numbers we want are the cell cost, in units of $\text{W}^{-1}$, and the PP&E$_0$ ratio (capital investment divided by revenue), so that we can then calculate profit margin and growth relations. To calculate the cell cost, we need the fixed and variable costs associated with each manufacturing step. Fixed costs are the costs of the tools and facility. Variable costs are the materials, labor, electricity, and maintenance. Two additional fixed costs are research and development (R&D), and selling, general, and administrative expenses (SG&A). In discussing fixed costs, we also need to state how they’re measured. One way is normalized by the tool lifetime, so the fixed cost of a tool is:

$$\frac{\text{Tool Cost}}{\text{Throughput} \times \text{Tool Lifetime}},$$

(2.41)

where the throughput is measured in W yr$^{-1}$ and this is denoted with $\text{W}^{-1}$. The second way is to normalize the cost over one year of production, so the
formula is:
\[
\frac{\text{Tool Cost}}{\text{Throughput}}.
\]

which is denote by \$/W_{\text{acap}}. Working with \$/W_{\text{acap}} is useful because it is what enters into the PP&E ratio which is used for calculating growth rates.

To have an apples-to-apples comparison to silicon, we will use [185] as a roadmap because many of the steps we are interested in have parallels to silicon manufacturing. The manufacturing steps we need to calculate the costs of are:

1. GaAs ingot synthesis,
2. Ingot slicing,
3. Zinc diffusion doping,
4. GaAs film spalling,
5. Tinplate lamination,
6. Chrome etching,
7. ARC evaporation,
8. Laser opening,
9. Contact printing,
10. Contact firing,
11. Testing,
12. Module integration.

For each step, we will estimate the fixed and total costs and comment on any step-specific critical elements. As a baseline, we will grant that cells are 20% efficient and the spalled films are 3 µm thick. This leads to a specific power of 7.5 W/g_{GaAs}. 
GaAs Ingot Synthesis

We will assume that we are using GaAs grown by vCZ in a quartz crucible, the growth rate is \(5 \text{ mm h}^{-1}\), the ingot is 200 mm in diameter, and 1000 mm long \[221\]. We will also include an unused cycle-time of one day for heating and cooling. Assuming each circular ingot is subsequently trimmed into a \(242 \text{ cm}^2\) M0 wafer size for 77\% utilization, we can then calculate that each tool produces 62 MW worth of GaAs at 1.6 MW/ingot.

In terms of fixed costs, we have learned in private communication that these tools cost approximately $1M per unit. Using numbers NREL has found for silicon, we assume the tool depreciates over ten years, the facility price is half of the tool price and depreciates over 20 years, and R&D and SG&A are 13\% of the total tool, facility, and variable costs minus the gallium and arsenic. These are summarized in Table 2.8.

<table>
<thead>
<tr>
<th>Item</th>
<th>Unit Cost, $</th>
<th>Lifetime, yr.</th>
<th>$ W^{-1}$</th>
<th>$/W_{\text{acap}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grower</td>
<td>$1M$</td>
<td>10</td>
<td>0.0016</td>
<td>0.016</td>
</tr>
<tr>
<td>Facility</td>
<td>$500k$</td>
<td>20</td>
<td>0.0004</td>
<td>0.0008</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td></td>
<td>0.002</td>
<td>0.024</td>
</tr>
</tbody>
</table>

For the variable costs, the materials we need to account for are the gallium, arsenic, crucibles, and hotzone consumables like heating elements. NREL estimates each operator can run three crystal pullers, or 1 full-time-equivalent (FTE) per puller due to eight hour shifts, and that each wafer takes 0.74 kW h per 160 \(\mu\)m silicon wafer. Scaling by relative thickness, the electricity for GaAs is 0.02 kW h per GaAs film. We will assume the gallium is fully utilized by recycling the trimmed pieces from ingot shaping. We will also use NREL’s maintenance estimate of 4\% of initial facilities cost per year. The variable prices are collated in Table 2.9. The metals prices come from public metals markets, and the slashes denote the difference between US and Chinese manufacturing rates as collated by the BLS \[22\].

Even at this point there are several interesting comparisons to silicon. First, referring to Fig. 11 of \[185\], they estimate a capital cost of 4\(\epsilon\)/wafer for crystal growth. To convert per-wafer to per-watt, the NREL analysis is assuming they’re growing M2 wafers which have an area 245 cm\(^2\). With a cell efficiency of 22\%, this comes out 5.4 W/wafer, so a capital cost of 0.74 \(\epsilon\) W\(^{-1}\).
Table 2.9: vCZ GaAs Growth Variable Costs.

<table>
<thead>
<tr>
<th>Item</th>
<th>Unit</th>
<th>Qty per Ingot</th>
<th>Unit Price</th>
<th>$/ingot</th>
<th>¢ W⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gallium</td>
<td>kg</td>
<td>102</td>
<td>3000</td>
<td>30672</td>
<td>1.92</td>
</tr>
<tr>
<td>Arsenic</td>
<td>kg</td>
<td>110</td>
<td>10</td>
<td>1100</td>
<td>0.070</td>
</tr>
<tr>
<td>Crucible</td>
<td>pc.</td>
<td>0.5</td>
<td>350</td>
<td>175</td>
<td>0.011</td>
</tr>
<tr>
<td>Hot-zone</td>
<td>pc.</td>
<td>1</td>
<td>450</td>
<td>450</td>
<td>0.028</td>
</tr>
<tr>
<td>Labor</td>
<td>h</td>
<td>224</td>
<td>35/2</td>
<td>7840/448</td>
<td>0.5/0.028</td>
</tr>
<tr>
<td>Electricity</td>
<td>kW h</td>
<td>6700</td>
<td>0.1</td>
<td>670</td>
<td>0.042</td>
</tr>
<tr>
<td>Maintenance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td>R&amp;D, SG&amp;A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.005</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.7/2.2</td>
</tr>
</tbody>
</table>

In comparison, we estimate our capital cost at only 0.2 ¢ W⁻¹, a quarter of silicon. Secondly, they estimate the Si materials variable cost at 25 ¢/wafer (4.63 ¢ W⁻¹), while ours is 2.2 ¢ W⁻¹, a 50% savings. From the variable costs, we can also see why US manufacturing struggles. Even with a highly automated process like this, US labor makes the ingots 22% more expensive. For the rest of the analysis, we will assume Chinese manufacturing.

The two critical cost drivers for our proposed process are, on the capital side, the tool being used at a high capacity, and on the variable side, the price of gallium. However, even with only 50% yield, this process would still be competitive with silicon. There is also a good chance we are overestimating the crucible, hot-zone, and electricity costs because silicon melts at 1410 °C while GaAs melts at 1238 °C, 170 °C less, which significantly reduces the stress on the furnace components and the energy required to heat the furnace.

**Ingot Processing**

Now that we have a GaAs ingot, the next two steps are squaring it and slicing it into more manageable pieces. For silicon, the squaring costs 0.37 ¢ W⁻¹. Due to thickness difference of the final wafers (250 µm vs 3 µm), we ignore these costs as trivial. Sawing the ingot into wafers for silicon costs 16 ¢/wafer. For ease of use we are interested in “wafers” about 1 cm thick, so we again ignore these costs.

**Diffusion Doping**

The next step is the zinc diffusion doping. We need to do this step next because it happens at 450 °C and the materials subsequently used are either themselves
not stable or not stable in contact with GaAs at this temperature. To estimate the cost of this step, we compare it to the emitter diffusion process for silicon. In silicon emitter diffusion, the wafer is first coated with a phosphosilicate glass (PSG) from the decomposition of POCl$_3$ flowing through the furnace at 800°C and then is ramped under oxygen to nearly 900°C to drive the phosphorous into the silicon, with the whole process taking about an hour.

Because our process runs at only 450°C, takes less time, and can be done in a belt furnace instead of a tube furnace, we estimate that our diffusion variable costs will be half as much. The variable costs of materials, labor, and electricity then come out to 0.081¢ W$^{-1}$. For the fixed costs, we are still buying a similar furnace, but just running it at a lower temperature, so we assume the same depreciation and maintenance costs of 0.18¢ W$^{-1}$ and 0.052¢ W$^{-1}$, respectively. To decompose the depreciation into equipment and facility, we use that on a acap basis the facility (20 yr depreciation) is half as expensive as the tooling (5 yr depreciation), which results in the diffusion tools costing 0.8¢/W$_{acap}$ and the facility being 0.4¢/W$_{acap}$. The R&D and SG&A costs come out to 0.051¢ W$^{-1}$ for a total cost of 0.36¢ W$^{-1}$.

**Spalling, Lamination, and Chrome Removal**

Analyzing the cost of spalling films from the wafer is difficult because an industrial spalling tool does not exist. However, there are literature attempts to estimate the cost with the result being a price of $2M per tool with a throughput of one wafer-per-second. This hypothetical tool does both the electroplating and the spalling. Using the same assumption that the facility costs half the tool and with 5 and 20 year depreciation times, the fixed costs for spalling are 0.32¢ W$^{-1}$ and 2¢/W$_{acap}$. With maintenance costing 3% of the equipment investment per year this contributes 0.04¢ W$^{-1}$.

The chief variable inputs are the electroplating solution, tinplate, labor, and electricity. The electroplating solution consumable is mainly CrCl$_3$. To estimate its cost, CrCl$_3$ is available on Alibaba for $5 kg^{-1}$. To determine the amount of chrome per cell, Andrew Nyholm has calculated that it takes an upper bound of 2µm of chrome to spall 3µm of GaAs, which then comes to 0.1¢ W$^{-1}$. Again from Alibaba, tinplate costs roughly $600 t^{-1}$ for a 150µm thick sheet which corresponds to 0.4¢ W$^{-1}$.

It is difficult to find precise estimates for the cost of electricity that includes
the plating, ventilation, pumps, etc., but as a worst case scenario we can assume that it costs the same as all of the silicon’s wet-chemical processing electricity of 0.2 $\text{¢}/\text{W}^{-1}$. Likewise, we use silicon’s labor cost for wet-chemistry of 0.057 $\text{¢}/\text{W}^{-1}$. With all this we then estimate the R&D and SG&A costs to 0.15 $\text{¢}/\text{W}^{-1}$.

**ARC, Laser Opening, Metal Printing, and Testing**

Thankfully these processes have direct analogues to silicon manufacturing so we can simply copy their cost model. The non-metal cost is 2.47 $\text{¢}/\text{W}^{-1}$, the metal costs 1.51 $\text{¢}/\text{W}^{-1}$, and the acap basis capital cost is 3.56 $\text{¢}/\text{W}^{-1}$.

**TEA Commentary**

The cost by process step broken down into materials, labor, electricity, equipment, maintenance, and R&D and SG&A is presented graphically in Fig. 2.42, and numerically in Table. 2.10. The total cell cost comes to 7.8 $\text{¢}/\text{W}^{-1}$, and the acap capital to 9.16 $\text{¢}/\text{W}_{\text{acap}}$. We can compare the cell cost to NREL’s 2020 PERC benchmark price of over 12 $\text{¢}/\text{W}^{-1}$ [232]. In this analysis GaAs cells are 65% the price of a silicon cell. We can also compare the annualized capacity basis capital costs. For silicon, the cost from silicon metal to finished cell based on the 2018 analysis is about 24 $\text{¢}/\text{W}_{\text{acap}}$, so that the capital investment for GaAs is half as much.

There are a few interesting features of Fig. 2.42. First, the cell cost is dominated by the front metallization and the gallium metal for making ingots. It’s likely we’ve overestimated the contact metal cost because GaAs operates at lower current densities than silicon and so intrinsically needs less metal. To further decrease the cost associated with the contact, we can likely use copper contacts because of GaAs’s good tolerance to copper [18], which would essentially eliminate that element of the cost. Gallium, though, presents a serious threat because being a low-volume, byproduct metal it is vulnerable to price swings. Second is that the materials are more expensive than the tools. This is perhaps unsurprising given that modern silicon manufacturing tools run around five-thousand wafers per hour (which incidentally limits the minimum interesting factor size to around 200 MW), but it highlights the need for parsimony. An interesting example is that we can not use tape for the spalling because it is too expensive. From wholesale prices for 3M 9082 high temperature transfer tape we estimate it adds a variable cost of 20 $\text{¢}/\text{W}^{-1}$.
Table 2.10: Itemized costs in proposed GaAs thin-film solar cell fabrication process.

<table>
<thead>
<tr>
<th>Category</th>
<th>GaAs Ingot</th>
<th>Zn Diffusion</th>
<th>Spalling</th>
<th>ZnS ARC</th>
<th>Laser Opening</th>
<th>Screen Printing</th>
<th>Metal Firing</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Materials</td>
<td>2.025</td>
<td>0.032</td>
<td>0.500</td>
<td>0.751</td>
<td>0.000</td>
<td>0.049</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Labor</td>
<td>0.028</td>
<td>0.011</td>
<td>0.057</td>
<td>0.044</td>
<td>0.025</td>
<td>0.069</td>
<td>0.016</td>
<td>0.052</td>
</tr>
<tr>
<td>Electricity</td>
<td>0.042</td>
<td>0.038</td>
<td>0.200</td>
<td>0.084</td>
<td>0.035</td>
<td>0.036</td>
<td>0.054</td>
<td>0.063</td>
</tr>
<tr>
<td>Equipment</td>
<td>0.002</td>
<td>0.180</td>
<td>0.320</td>
<td>0.180</td>
<td>0.069</td>
<td>0.153</td>
<td>0.035</td>
<td>0.098</td>
</tr>
<tr>
<td>Maintenance</td>
<td>0.097</td>
<td>0.052</td>
<td>0.040</td>
<td>0.082</td>
<td>0.012</td>
<td>0.040</td>
<td>0.020</td>
<td>0.046</td>
</tr>
<tr>
<td>Metal</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.509</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>R&amp;D, SG&amp;A</td>
<td>0.005</td>
<td>0.051</td>
<td>0.150</td>
<td>0.148</td>
<td>0.018</td>
<td>0.241</td>
<td>0.016</td>
<td>0.034</td>
</tr>
<tr>
<td>Total</td>
<td>2.199</td>
<td>0.364</td>
<td>1.267</td>
<td>1.290</td>
<td>0.159</td>
<td>2.097</td>
<td>0.140</td>
<td>0.292</td>
</tr>
</tbody>
</table>

Grand Total 7.809
Figure 2.42: Processing cost by step and cost category for thin-film GaAs solar cell fabrication. Costs are dominated by the gallium and silver metal.

which immediately kills the economic viability of the project.

Also, we do not address process yield and uptime. However, it is unlikely they would change our results by over a factor of two because silicon manufacturing has yields over 90%. Also, by working with thin, flexible GaAs films we are not vulnerable to wafer breakage like silicon is.

Accounting for module fabrication, because these GaAs cells are drop-in replacements for silicon cells, we use silicon’s benchmark cost model of 13¢/W\(^{-1}\) which includes all costs, and an acap cost basis of 8¢/W\(_{acap}\). This brings the profit-free cost to 20.8¢/W\(^{-1}\) and capital to 17.2¢/W\(_{acap}\). To account for working capital and debt-financing NREL suggests a 15% gross margin for the company to be sustainable which raises the final cell price to 24¢/W\(^{-1}\). This can be compared to the silicon price in the same benchmark report of 37¢/W\(^{-1}\) for panels manufactured in urban China. In this analysis, GaAs solar panels are 65% the price of silicon solar panels.

We can now also estimate the growth-rate for a GaAs solar cell company. Using the silicon solar price of 37¢/W\(^{-1}\) as the market rate and assuming the difference between this and the sustainable price for GaAs of 24¢/W\(^{-1}\) can be captured as operating profit then the profit margin is 35%. For the PP&E\(_0\) ratio, the revenues are 37¢/W\(^{-1}\) and the capital cost is 17.2¢/W\(_{acap}\)
so $PP&E_0 = 0.46$, which roughly corresponds to our original estimate for a capital-free-substrate solar cell process from the introduction. With this profit margin and capital intensity the maximum sustainable growth rate is over 100%. It’s important to remember that as soon as competitors start making panels, though, that price competition will shrink profit margins and therefore growth rates.

So far we have found that it is plausible for GaAs solar cells to be cheaper than silicon and, assuming silicon sets the price, extremely high growth rates are possible. We can also argue that even if the cost-per-watt were the same, electricity from GaAs solar panels would be cheaper than silicon because GaAs is less sensitive to temperature than silicon is. Solar cell efficiency changes with temperature because with increasing temperature the bandgap narrows which allows the cell to absorb more light, increasing $J_{sc}$, but it also increases the intrinsic carrier concentration, which decreases $V_{oc}$. In real-world outdoor testing GaAs solar panels have extremely weak temperature dependence, where the $J_{sc}$ gain and $V_{oc}$ loss very nearly cancel out [230]. In contrast, according to Jinko’s solar panel datasheets, silicon solar cells typically lose 0.4%/°C (relative) conversion efficiency as they warm up past 25°C. As a real-world demonstration, outdoor testing of silicon panels has found they operate at about 80% of their nameplate output in desert climates due to the high temperatures [59, 94, 77, 49]. GaAs panels would not suffer the same extent of degradation due to their better temperature performance and therefore offer a lower cost of electricity. This is especially important during the summer where daytime energy is needed for air conditioning and daytime temperatures will degrade silicon’s performance at this critical time.

In conclusion, we have developed a technoeconomic model for diffusion doped, thin-film, GaAs solar cells. With this model we predict per-watt costs 35% lower than silicon, and upfront capital costs half that of silicon. If these costs can be realized and the price difference between these cells and silicon be converted into operating margin the maximum sustainable growth rate is over 100% per year. Further, all else being equal, GaAs solar panels will also have lower costs of energy due to their superior temperature performance.
2.9 $n^+/p$ Homo junction GaAs Solar Cells

Because of the poor minority carrier properties of commercially available n-type GaAs, we also explored $n^+/p$ homojunction cells. We expected the electron diffusion lengths of p-type wafers to be superior due to higher electron diffusion coefficients (the electron mobility is roughly ten times greater than the hole mobility) and there is a large asymmetry between the minority electron and hole capture coefficients of EL2 ($\sigma_p = 5.5 \times 10^{-15} \text{cm}^2$, $\sigma_n = 1.4 \times 10^{-16} \text{cm}^2$), so EL2 will not limit the minority carrier diffusion length \[150, 296\]. We tried four different techniques for making $n^+/p$ junctions: liquid phase epitaxy, monolayer dopant diffusion, gallium sulfide proximity diffusion, and spin-on-glass diffusion doping. The monolayer and glass diffusion doping work was done if collaboration with Dr. Rosaria Puglisi, an expert in silicon monolayer diffusion doping \[24, 197, 23\].

Liquid Phase Epitaxy

Liquid phase epitaxy (LPE) is one of the oldest and simplest techniques for growing high-quality semiconductors \[267, 6\]. The way we tried to implement it was evaporating a film of tin on GaAs and then heating and cooling it in a furnace. The proposed scheme is illustrated in Fig. 2.43a. When heated, the tin would dissolve some GaAs, and then upon cooling it would re-crystallize as Sn doped n-type GaAs. In a typical experiment, we evaporated 1 µm of Sn onto a semi-insulating GaAs chip and then annealed it at 550°C for ten minutes and then let it cool in the furnace. We then removed the tin film in warm HCl. A microscope image of the GaAs surface is shared in Fig. 2.43b. We can see both that the reaction was non-uniform and has crystallographic features (the rectangular features). Due to the extremely poor surface morphology we did not pursue this process any further. However, we think this continues to be interesting conceptually due to the similarity between LPE and wave-soldering, but we do not have the equipment to further pursue this.

Monolayer Dopant Diffusion

The idea of monolayer diffusion doping is to graft of monolayer of a dopant bearing molecule to the semiconductor surface and the diffuse the dopant into the semiconductor. This has been previously demonstrated on GaAs using laser driven doping or rapid-thermal-annealing (RTA) \[297, 149, 139\]. The procedure we tried was coating the GaAs with DTT as in the surface treatment
Figure 2.43: Attempts at Sn-based LPE of GaAs. a) Illustration of LPE method for GaAs doping where Sn dissolved some GaAs and then redeposits it as heavily n-type. b) Microscope image of GaAs surface after tin removal showing extremely rough surface with crystallographic regrowth features.

section and then coating it with 20 nm of electron beam deposited SiO$_2$, ALD deposited SiO$_2$, or ALD deposited Al$_2$O$_3$. The ALD SiO$_2$ was deposited using bis[diethylamino]silane (BDEAS) and oxygen plasma at 150°C and the Al$_2$O$_3$ using Al(CH$_3$)$_3$ (TMA) and H$_2$O also at 150°C. The samples were annealed at 850°C for 30 min. After annealing the oxide was stripped in HF. Of these, the only one that had measurable Hall results was the ALD SiO$_2$ which had a sheet resistance of 37 Ω/□, a mobility of 470 cm$^2$ V$^{-1}$ s$^{-1}$, and a sheet carrier concentration of $3.6 \times 10^{14}$ cm$^{-2}$. We did not make further devices with this process, but it is interesting that it worked at all. These processes may work better with an RTA which reduces the time spent at high temperatures, but unfortunately we did not have access to one.

**Spin on Glass Diffusion Doping**

We also explored using tin-doped spin on glass. There is literature precedent for using tin-doped glass to diffusion dope GaAs, but we could not find anything quantitative about the minority carrier properties or resulting junction [137, 93]. We performed tin diffusion using Filmtronics Sn970 spin on glass. We prepared the film by first cleaning the GaAs chip in acetone and isopropanol and then spin coating the glass for 30 s at 3000 rpm followed by baking on a hotplate at 150°C for 15 min. We then did the diffusion by placing the chip face-up on a quartz slider in a cold tube furnace, brought the furnace up to temperature over an hour, held it for a period of time, and then let it cool before removing the chip. We stripped the glass with a brief
HF dip. One thing we found is that the chips can not be annealed under forming gas. Chips annealed under forming gas had severely degraded morphology while chips annealed under house nitrogen remained mostly mirror smooth. We suspect hydrogen in the forming gas reduced the tin oxide to metallic tin which then formed a liquid phase with the GaAs. Despite the good appearance, the chips still had large densities of defects visible under a microscope and the outer \( \approx 0.5 \text{ mm} \) of the chip was still badly degraded. Pictures of a chip annealed at 850 °C for 30 min are shared in Fig. 2.44. The picture on the left shows the large-scale features. On the right hand side is the badly degraded edge of the chip. The gold dots are shadow-mask defined contact pads for Hall measurements. We can also see defects throughout the entire field. The right picture is a detailed view of a particular defect cluster. The symmetry of the cluster is particularly striking. Also, the defects are all rectangles, suggesting degradation consistent with crystallographic thermal etch-pitch formation, showing the glass does not perfectly prevent arsenic vaporization [167]. Also, from a technoeconomic standpoint, we have found that these devices may not be viable regardless of performance because the spin on glass alone costs roughly $1 \text{ W}^{-1}$.

We found two recipes that reliably produced doped layers that were suitable for Hall measurements. They are summarized in Table. 2.11. The mobility measurements are consistent with heavily doped n-type GaAs.

We also made solar cells using the 850 °C process on 3 μm thick p-type HVPE layers doped at \( 5 \times 10^{16} \text{ cm}^{-3} \) on top of a heavily doped substrate. The fabrica-
Table 2.11: Tin Doped Glass Hall Measurements

<table>
<thead>
<tr>
<th>Temp., °C</th>
<th>Time, min.</th>
<th>Sh. Res., Ω/□</th>
<th>(\mu_e), cm²/(V s)</th>
<th>Conc., cm⁻²</th>
</tr>
</thead>
<tbody>
<tr>
<td>755</td>
<td>30</td>
<td>197</td>
<td>1210</td>
<td>1.74E+14</td>
</tr>
<tr>
<td>850</td>
<td>30</td>
<td>35</td>
<td>1050</td>
<td>2.68E+13</td>
</tr>
</tbody>
</table>

The cell achieved a \(V_{oc}\) of 882 mV, but suffers from a number of unusual effects. First, it is shunted. Second, there is a kink in the IV curve near \(V_{oc}\). Third, it took thirteen minutes of using the mesa etch for emitter thinning to reach this performance, suggesting the junction was several microns deep and further supporting the idea that there was a thick dead layer of extremely defective GaAs. The first priority for improving the cell performance should be optimizing the thermal processing to preserve the surface morphology and minimize the dead layer thickness. Rapid thermal processing will likely be necessary. Also, for economic reasons, less expensive dopant sources need to be developed. In-situ grown doped glass sources may be satisfactory, but they are beyond our current ability to fabricate.

2.10 Closing Thoughts

In this chapter, I have presented a number of advancements in the low-cost processing of GaAs solar cells including: a model for heavily doped GaAs solar cell performance, new techniques for GaAs and InP diffusion doping, the existence of high minority carrier lifetime n-GaAs, high-performance \(p^+/n\) diffused junction solar cells, tentative steps towards simply-processed AlGaAs heterojunction, air-tolerant, low-temperature metallization, and the first, to our knowledge, \(n^+/p\) diffusion doped GaAs solar cells. We also presented an economic argument for the viability of thin-film GaAs solar cells using bulk-grown material in combination with spalling to produce thin films and the developed cell-formation techniques.

Within every topic above, there are multitude of details that could be worked out; much of the work flowed from ‘gut-feelings’ about what should work. However, the cells plainly work, and in terms of major steps forward there
Figure 2.45: Best cell performance with tin-doped spin on glass diffusion source.

are two obvious choice. First is fabricating a spalled thin-film diffusion doped cell. This is the global-warming and economically interesting device and all the enabling technologies are now in place. Second is qualitatively improving performance. The low-cost AlGaAs process is a step in this direction that right now just has the barest proof-of-principle validation.

This is the bulk of my work that I consider valuable. Interestingly, this was all done in a year, and would not have been possible without already having spent years meditating on GaAs processing and metallurgy and the freedom granted to me by Harry to work on, what started as a one-off-just-for-fun device.
Chapter 3

GAAS PASSIVATION AND NON-EPI TAXIAL HETEROJUNCTIONS

3.1 Introduction
Before working on diffusion doping, I was previously working on GaAs passivation and non-epitaxial heterojunctions. The idea was that silicon and perovskite solar cells can both work with non-epitaxial heterojunctions (amorphous silicon for Si [250], a whole zoo of organics and inorganics for perovskites [56, 39, 128, 245]), why not the same for GaAs? If we could figure out how to make good non-epitaxial heterojunctions, then that is one of the major cost-barriers for III-V solar cells resolved.

Making a useful non-epitaxial heterojunction has two basic requirements. First is passivating the surface states. This is required for the device to perform other than as a simple Schottky contact. Second is for the heteropartner to have the proper properties to conduct carriers to and from the GaAs. This involves issues like energy band alignment and conductivity. I will present my results on each of these problems in sequence, starting with surface passivation.

The main passivation results that I share are the passivation of GaAs by sulfides, thiols, and an N-heterocyclic carbene, 1,3-Diisopropylimidazolium (the NHC), achieving surface recombination velocities as low as 200 cm s$^{-1}$, the robustness of passivation up to 200°C in air, XPS measurements of the GaAs surface, and DFT calculations of how dithiothreitol, a thiol, and 1,3-Diisopropylimidazolium bond to GaAs.

We also made a variety of solar cells on the passivated GaAs using a wide-variety of organic and inorganic semiconductors, including: SnO$_2$, TiO$_2$, CuSCN, poly triaryl amine (PTAA), Tris(4-carbazoyl-9-ylphenyl)amine (TCTA) and polystyrene (PS). With these, we were able to achieve a $V_{oc}$ of 840 mV. There were also a large number of things we tried that did not work which we will share as well. The net result of this work was passivation strategies that enabled our champion diffusion-doped cell shared in the previous chapter, but the solar cells are categorically inferior to their homojunction counterparts in both $V_{oc}$ and reproducibility.
3.2 GaAs Passivation

Introduction

Devices based on passivated GaAs surfaces have been something of a holy grail of semiconductor device fabrication for decades [89, 110]. The reason passivation of GaAs for the purpose of making MOSFETs has been so intensely pursued is that the higher electron mobility and the existence of semi-insulating substrates allows higher-speed devices than silicon. What makes GaAs MOSFETs so difficult to fabricate is that GaAs does not have a simple, passive, oxide like silicon does with SiO$_2$ where surface state densities below $1 \times 10^{10}$ cm$^{-2}$ eV$^{-1}$ are readily achieved [67, 218]. Instead, researchers observe the GaAs native oxide, and the GaAs/O phase diagram predicts, a non-uniform, segregated, mix of GaAs, Ga oxides, As oxides, and ternary oxides [83, 220]. The complex quaternary thermodynamics renders reactive dielectrics like SiO$_2$ and Al$_2$O$_3$ unhelpful.

Despite the poor-quality of GaAs/oxide interfaces, GaAs readily reacts with sulfides to form high electronic quality surfaces [188, 288]. Based on photoluminescence measurements the passivation can be nearly as good as AlGaAs [152]. Further, sulfide passivation also increases the possible range of Schottky barrier heights on GaAs [25]. Photoconductivity measurements show that weak illumination can decrease the surface band-bending of a sulfur passivated surface by nearly 1 V [287].

We will discuss our results on surface passivation by first describing the chemical techniques we used to passivate the GaAs, second, measuring the performance of the various passivation strategies, and finally interpreting the measurements in terms of predicted solar cell performance.

GaAs Chemical Passivation Techniques

Over the course of this work, we developed and refined a number of passivation strategies. Frustratingly, many things passivated GaAs, but were unacceptable for making solar cells because they left some kind of visible residue. We will now discuss each broad class of passivants and the practical issues surrounding their use.
Inorganic Sulfide

GaAs is typically passivated using solutions of Na$_2$S or (NH$_4$)$_2$S. One problem we found with these solutions as described in the literature is that they tended to leave a hard-to-remove residue that was invulnerable to water and common organic solvents (alcohols, acetone, pyridine, benzene, THF, etc.). We suspected that the residue was due to complex poly-sulf-oxides and so added a reducing agent, hydrazine monohydrate, to our passivating solution which resulted in residue-free surface after water or alcohol rinsing.

The best solution we found consisted of an equivolume mixture of (NH$_4$)$_2$S, ammonia water, and hydrazine monohydrate ($N_2H_4\cdot H_2O$) along with 5% by mass elemental sulfur. The elemental sulfur increases the quality of the passivation by allowing it to very slowly etch through the defective GaAs surface [189]. When making the passivant we started with the sulfide and the ammonia. When adding the hydrazine the solution would lightly fume, indicating some reaction taking place. Finally adding the sulfur caused the solution to turn deep orange from yellow. To actually passivate a chip, we would first sequentially sonicate in acetone and IPA for three minutes each, followed by treating the chip with the sulfide for at least five minutes. After five minutes we would rinse the chip under flowing DI water or methanol. Occasionally there would be a residue that readily sublimated on a hot plate at 100°C, indicating there was sometimes leftover elemental sulfur.

Organic Thiols

One issue with the inorganic passivants is that there is only one atom separating the GaAs from the environment. Organic thiols can get around this by simply putting more atoms between the GaAs and the environment. There is a large existing literature of self-assembled monolayers on GaAs [44, 254, 100, 113, 102, 101, 157, 294]. In a typical experiment, GaAs is submerged in an alcoholic solution of alkanethiol and over the course of hours the alkanethiol bonds to the GaAs surface. Adding ammonia to the solution helps with passivation by removing the GaAs native oxide and and deprotonating the thiol which increases its reactivity.

The specific thiols we worked with were:

- 1-Octadecanethiol (OT)
- 1-Hexadecanethiol (HT)
- 1,8-Octanedithiol (ODT)
- 8-Mercapto-1-octanol (MO)
- 8-Mercaptooctanoic acid (MOA)
- Dithiothreitol (DTT)

We chose the first five because of their common usage in making self-assembled monolayers, and using DTT was informed by our experience working with ODT, MO, and MOA. The SAM forming solution typically consisted of 10 mL of 4:1 by volume ethanol-to-ammonia and one-drop of the monolayer-precursor. After cleaning the GaAs in solvents the chip was treated with the solution for at least 15 min and then rinsed under flowing DI water. The OT and HT treated samples were extremely hydrophobic, providing qualitative evidence for the treatment working.

The reason we worked with functionalized thiols (ODT, MO, MOA, DTT) is that the hydrophobicity of the alkanethiols is incompatible with durable solar cells because there is no cohesive force holding the cell together; the material is completely free to slide or delaminate. Indeed, we found that it was impossible to coat materials on the OT and HT treated samples. By adding polar groups to the head of the molecule dipole interactions enhance the sticking of molecules to the surface. The problem with strategy, though, is that we found the functionalized thiols would quickly polymerize and form a wax, as shown in Fig. 3.1a. We think the mechanism is base-catalyzed oxidative coupling of the molecules similar to how disulfide bonds are formed. The proposed reaction is shared in Fig. 3.1b. Something we tried was adding hydrazine to reduce the bonds back down to their monomers, but ironically the S–O bond is stronger than the C–O bond so reducing the chemical just results in more reactive sulfoxides and the wax remains [81]. Interestingly, despite the biochemical importance of thiol oxidation reactions, they remain poorly understood in general [215].

Serendipitously, I was discussing my frustrations with thiols with biochemist Siobhán MacArdle, and she taught me about dithiothreitol (DTT). DTT, shown on the left side of Fig. 3.2, is popular in biochemistry because when
Figure 3.1: Residue from trying to passivate GaAs with functionalized thiols. a) From left to right, pictures of wax in jar after trying to passivate GaAs with functionalized thiols, a polymer film that grew on GaAs during functionalized thiol passivation, cloudy residue in passivation solution even when run under a nitrogen purge. b) Putative mechanism for base-catalyzed oxidative polymerization of organic thiols.

it reduces the disulfide bond of a protein it undergoes a ring closing instead of coupling reaction, shown in the remainder of Fig. 3.2. What this means for GaAs passivation is that instead of forming a polymer when oxidized it will stay as small molecules. For experiments with DTT, we would add about 20 mg per 10 mL of ethanol/ammonia solution. We never saw DTT solutions form a wax, but small crystals would grow, presumably its oxidized form, that were easily removed from chips with water rinsing. The passivated GaAs was also strongly hydrophilic, which is why we used DTT as the pre-ARC passivation for diffusion doped cells.
After a group meeting presentation about the thiol work, one of my chemist labmates, Aidan Fenwick, introduced me to N-heterocyclic carbenes (NHCs), an additional class monolayer forming molecules. NHCs are a newer class of monolayer-forming molecules where the complexes are formed the the lone electron pair of a divalent carbon atom instead of a sulfur atom. The particular carbene we used, 1,3-diisopropylimidazolium, is illustrated in Fig. 3.3. NHCs form monolayers on gold similarly to thiols and are more electrochemically stable than their thiol counterparts [43]. In addition, the hydrogen carbonate salts of carbenes are air-stable and produce reactive carbenes in solution [57]. Conveniently, 1,3-diisopropylimidazolium hydrogen carbonate salt is directly available from TCI. To use the NHCs, we prepared a solution similarly to DTT except where the solvent is ammonia saturated methanol to avoid water protonating the carbene. GaAs treated with the NHC was hydrophilic. An additional interesting feature of the NHC is that it etched the GaAs at about 200 nm d⁻¹. We discovered this after accidentally leaving a chip soaking overnight. An AFM scan of the etched surface is shared in Fig. 3.4. The RMS surface roughness is 5.1 nm. Also, the texturing does not have an obvious crystallographic preference. A possible mechanism is that the NHC is such a potent complexing agent that it’s solvating the GaAs into solution.

Figure 3.3: Illustration of 1,3-Diisopropylimidazolium. Note the electron pair on the bottom carbon.

**N-Heterocyclic Carbenes**

After a group meeting presentation about the thiol work, one of my chemist labmates, Aidan Fenwick, introduced me to N-heterocyclic carbenes (NHCs), an additional class monolayer forming molecules. NHCs are a newer class of monolayer-forming molecules where the complexes are formed the the lone electron pair of a divalent carbon atom instead of a sulfur atom. The particular carbene we used, 1,3-diisopropylimidazolium, is illustrated in Fig. 3.3. NHCs form monolayers on gold similarly to thiols and are more electrochemically stable than their thiol counterparts [43]. In addition, the hydrogen carbonate salts of carbenes are air-stable and produce reactive carbenes in solution [57]. Conveniently, 1,3-diisopropylimidazolium hydrogen carbonate salt is directly available from TCI. To use the NHCs, we prepared a solution similarly to DTT except where the solvent is ammonia saturated methanol to avoid water protonating the carbene. GaAs treated with the NHC was hydrophilic. An additional interesting feature of the NHC is that it etched the GaAs at about 200 nm d⁻¹. We discovered this after accidentally leaving a chip soaking overnight. An AFM scan of the etched surface is shared in Fig. 3.4. The RMS surface roughness is 5.1 nm. Also, the texturing does not have an obvious crystallographic preference. A possible mechanism is that the NHC is such a potent complexing agent that it’s solvating the GaAs into solution.
GaAs Passivation Characterization
The two main ways we characterized the passivation were time-resolved PL (TRPL) and X-ray photoelectron spectroscopy (XPS). TRPL provided direct information about the minority carrier recombination properties of the sample and XPS told us about the chemical state of the passivated GaAs. We also did some brief density functional theory (DFT) calculations to evaluate the structure of DTT and the NHC on GaAs. We will also briefly discuss a hypothetical method for characterizing the surface enable by our simple diffusion-doping method and inspired by [287].

Time Resolved Photoluminescence Spectroscopy
For a review of transient photoluminescence, we refer the reader to section 13 of [4]. The main advantage of TRPL over static PL is that the injection level is high enough to force the layer into a flat-band regime and that results don’t depend on optical coupling efficiencies of the experimental setup. The
Figure 3.5: Illustration of TRPL setup.

Experimental setup we used for TRPL measurements consisted of a PicoQuant 405 nm pulsed laser coupled into an inverted microscope and the luminescence is measured by a Micro Photon Devices PDM-series single photon detector whose output is correlated with the laser by a PicoHarp. The setup is illustrated in Fig. 3.5. The 405 nm laser light is filtered by the short pass filter to remove harmonics before being focused by a 20X objective onto the GaAs. The luminescence is collected by the same objective and selectively reflected by a dichroic mirror before being further filtered by a longpass filter in front of the detector to make sure it is not overwhelmed by stray laser light.

Something we found was that the laser had to be externally driven at a lower frequency than the internal oscillator could provide in order to see the signal decay over several lifetimes. The experimental hint that let us know the pulse repetition was too high was that the ‘dark’ detector measurement before the pulse varied with the pulse rate indicating all the carriers had not decayed before the next pulse from the laser arrived. To show this problem is not entirely obvious, we show the unprocessed data for sulfide passivated GaAs taken at the internal clock pulse rate of 10 MHz in Fig. 3.6a, and one with an
Figure 3.6: TRPL data for the same piece of GaAs, a) taken with pulse rate of 10 MHz, b) taken with a pulse rate of 500 kHz.

external clock of at 500 kHz in Fig. 3.6b. For the fast rate, there appears to be an exponential decay down to the dark value after a few ns. For the slow rate, we see the same transient followed by a slow decay over a µs. Notice also the dynamic range of the fast pulse is 400, while it is 3,000 for the slow pulse, indicating the ‘dark’ level of the fast sample is in fact just a very long tail.

Samples for TRPL were an n-type GaAs substrate doped to $3 \times 10^{18}$ cm$^{-3}$ with a buffer layer, 50 nm GaInP confinement layer, 1 µm GaAs photoabsorber, and 50 nm of GaInP passivation, all doped to approximately $3 \times 10^{16}$ cm$^{-3}$ n-type. We also measured similar samples doped p-type, but the limited conduction band offset between GaInP and GaAs failed to confine the generated carriers, complicating quantitative analysis. To measure the effectiveness of chemical passivation we etched the GaInP top layer in concentrated HCl and then treated the samples as described in the previous section. We present data for OT, Sulfide, DTT, and NHC. We also measured the stability of the passivation by heating the samples on a hotplate in air for five minutes at increasing temperatures. Carrier lifetimes were determined by fitting the TRPL data to:

$$I(t) = I_0 e^{t/\tau} + \text{dark},$$

where $I(t)$ is the TRPL signal as a function of time, $I_0$ is the initial signal intensity, $\tau$ is the exponential decay constant, and dark is the dark count-rate of the detector. The TRPL data and fits to the above equation are shared in Fig. 3.7. The lifetimes of the as-treated samples are on the order of hundreds on ns. We can use this lifetime to estimate the one-sun intensity quasi-fermi level splitting in the passivated GaAs and therefore the maximum $V_{oc}$ of a
solar cell. Consider that the np product is given by:

\[ p \times n = n_i^2 e^{\frac{q\Delta \mu}{kT}}. \]  (3.2)

In this case, n is just the donor doping density, and p is given by:

\[ p = G \times \tau / d, \]  (3.3)

where \( G \) is the generation rate and \( d \) is a representative length-scale, around 10 \( \mu \)m. For one-sun generation of 30 mA cm\(^{-2}\), \( N_D = 5 \times 10^{16} \) cm\(^{-3}\), and \( \tau = 100 \) ns, we find \( \Delta \mu = 1.0 \) V, a very respectable \( V_{oc} \).

To analyze the data, if we assume the lifetime is dominated by non-radiative recombination at the surfaces, we can further infer the surface recombination velocity of the interfaces as:

\[ S_{\text{front}} = \frac{d}{\tau} - S_{\text{back}}, \]  (3.4)

where S is the surface recombination velocity of the corresponding surface, \( d \) is the thickness of the epilayer (1 \( \mu \)m in this case), and \( \tau \) is the PL lifetime. In using this expression we assume the carrier lifetime is large compared to the diffusion limited lifetime given by:

\[ \tau_{\text{diff}} = \frac{d^2}{D\pi^2}, \]  (3.5)

where \( D \) is the carrier diffusion constant, and results in a lifetime less than 1 ns. This is consistent with the lifetime of the unpassivated surface in Fig. 3.7, justifying this assumption.

The carrier lifetime and surface recombination velocity for each passivation type and hotplate temperature is presented in Fig. 3.8. First, we note that the lifetimes are all longer than the naive radiative-limited lifetime of approximately 150 ns, indicating the presence of photon recycling in all samples at room-temperature. The lifetime of the GaInP passivated structure implies a photon-recycling factor of at least 6.5, which is consistent with literature values for similar structures [154].

The inorganic sulfide provides the best initial passivation with an SRV of 240 cm s\(^{-1}\) and the OT provides the best initial organic passivation with an SRV of 250 cm s\(^{-1}\). The SRV of an untreated surface is at least 2.5 \( \times 10^5 \) cm s\(^{-1}\).
Figure 3.7: Collation of TRPL data as a function of surface treatment and hotplate annealing temperature.
limited by the diffusion lifetime. Interestingly, the quality of the DTT and NHC passivation improves with a 100°C anneal, reaching an SRV of 200 cm s\(^{-1}\), and the OT reaches 200 cm s\(^{-1}\) after a 150°C anneal, corresponding to a photon recycling factor of 2.7. This is to our knowledge the first report of photon-recycling in chemically passivated GaAs films. Passivation is completely lost by 250°C for all passivation chemistries with the NHC failing first, followed by the thiols, and finally the sulfide. The thiol passivation failing at temperature above 150°C agrees with previous work on the thermal desorption of thiol monolayers from GaAs [44]. One possible mechanism for the improvement of SRV with annealing is desorption of volatile sulfides.

We also measured the TRPL of samples after coating the passivated surfaces, but the measurements were done before we realized we were simply measuring the initial transient, and by the time we figured that out, we were nearly out of heterostructure samples to measure. However, we can qualitatively say that the passivation is preserved when it is coated with organics like TCTA by spin-coating or evaporation. Covering the sulfide-passivated surface with materials that form inorganic crystals like CuSCN killed the passivation, but the organic-treated samples survived, indicating the hydrocarbon chains did protect the surface. We also measured the absolute PL quantum yield of passivated, bulk, GaAs using a CW argon laser. From the measurement we
estimated a one-sun Fermi level splitting of 970 mV, in good agreement with the TRPL measurement.

**XPS of GaAs Surfaces**

To help understand the chemistry happening between the GaAs and passivant we also studied the passivated surfaces using XPS. The XPS measurements were performed and analyzed by Rebecca Glaudell. Samples were prepared as above and waited in air for around an hour before being loaded into the XPS.

We chiefly studied the Ga 2p state because it is more surface sensitive and has previously been well correlated with GaAs MIS capacitor performance [95]. The oxygen exposed gallium forms either $\text{Ga}_2\text{O}_3$ (Ga$^{3+}$) or $\text{Ga}_2\text{O}$ (Ga$^+$), where Ga$^{3+}$ indicated poor-performance devices and Ga$^+$ is electrically benign.

To begin, we share the collated Ga 2p spectra in Fig. 3.9. In this figure, zero binding energy is set to the valence band maximum and we are showing the spectra for GaAs with a native oxide and passivated with NHC, DTT, OT, and the sulfide. We can immediately see that the passivated spectra are qualitatively different from the native oxide and that there are differences within the passivated samples. The DTT and OT spectra are similar, which makes sense because is the same thiol functional group interacting with the GaAs, but the sulfide and NHC are distinct, with the NHC clearly having multiple large peaks.

The spectra of each passivant and its peak decomposition is shared in Fig. 3.10. The energy splitting from the GaAs to the Ga$^+$ and additional state are shared in Table 3.1. For each passivated spectra there is a conserved splitting between the GaAs bulk peak and a Ga$^+$ peak of 0.55 eV which corresponds well with the literature value[95]. Each passivated sample requires an additional peak for fitting that we identify as corresponding to the bond between the GaAs and the passivant. The native oxide has a state that we label as Ga$^+$, but whose binding energy is distinct, making its classification uncertain. The thiols have the smallest chemical shift, near 0.05 eV, the sulfide is next largest at 0.3 eV larger, and the NHC has a much larger shift at 1.36 eV. Interestingly, the NHC shift is nearly as large, but definitely distinct by 0.05 eV, from the native oxide’s highest bonding energy peak. That the NHC etches GaAs also suggests that this high binding energy peak is the NHC and not just undisturbed oxide. From the As 5d spectra (not shown) the passivated samples, including the
Figure 3.9: Overlaid Ga 2p XPS spectra for each passivation chemistry.

Table 3.1: Energy splitting between Ga 2p GaAs peak and the Ga and additional surface peak.

<table>
<thead>
<tr>
<th>Surface</th>
<th>Ga(^{2+}), eV</th>
<th>Other, eV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>0.84</td>
<td>1.41</td>
</tr>
<tr>
<td>DTT</td>
<td>0.56</td>
<td>0.04</td>
</tr>
<tr>
<td>OT</td>
<td>0.58</td>
<td>0.1</td>
</tr>
<tr>
<td>NHC</td>
<td>0.56</td>
<td>1.36</td>
</tr>
<tr>
<td>Sulfide</td>
<td>0.56</td>
<td>0.36</td>
</tr>
</tbody>
</table>

NHC, have no detectable oxide. This suggests the only remaining oxide is the low valence Ga\(_2\)O. In the case of DTT, the sulfur peak shows only one feature, suggesting both thiols are bonded to the GaAs.

**DFT Structure Prediction**

To predict the structure of the DTT and NHC SAMs on GaAs we did structure relaxation calculations using the Quantum Espresso DFT code, the BURAI frontend, and the Materials Cloud Standard solid-state efficiency pseudopotentials [183, 72, 73, 201, 140]. From XPS data, we inferred that the passivants
Figure 3.10: Ga 2p XPS spectra for each passivant decomposed into peaks with chemical identification.
bonded mainly to gallium atoms, and that the DTT had both of its sulfur atoms bonded to the GaAs surface. For initial placement of the DTT we noticed that the thiolate-thiolate distance, 3.9 Å, is remarkably similar to the gallium nearest neighbor distance on the 100 surface of 4.0 Å, and so placed the sulfurs above nearest neighbor gallium sites. In the NHC case, the size of the molecule means only one carbene can fit for each two nearest-neighbor gallium atoms, so we simply placed the carbene between nearest neighbors. To allow the GaAs surface to relax we simulated two layers of the 100 surface and fixed the bottom layer to the unperturbed structure. Illustrations of the relaxed structures are presented in Fig. 3.11. We find that the sulfur atoms of the DTT each coordinate two gallium atoms and that the carbon backbone lies along the valleys of the surface. With the NHC we find that the carbene strongly breaks the symmetry between the surface gallium atoms. To characterize the degree of symmetry breaking we can consider the nearest-neighbor gallium distances on the surface which are 4.01 and 4.17 Å for the DTT case, 2.59 and 4.35 Å for the NHC, and 4.00 Å for the unperturbed lattice, a change of 35% for the NHC treated surface. The large perturbation of the NHC treated surface qualitatively aligns well with the Ga 2p XPS spectra being substantially different from the native oxide, sulfide, or thiol passivated surfaces, and the NHC complexing so strongly with the GaAs that it etches the surface.
Conclusion
In this section, we have presented measurements of the passivation quality of an inorganic sulfide, 1-octanethiol, dithiothreitol, and 1,3-diisopropylimidazolium passivation chemistries on GaAs. This is the first report we are aware of using DTT or an NHC to passivate GaAs. This is also the most direct measurement of the organically passivated surface SRV. We find surface recombination velocities around $250 \text{ cm s}^{-1}$, which implies photon recycling in the passivated samples, and that the passivation is stable to around $150 \, ^\circ\text{C}$ in air. We also estimate the solar cell $V_{oc}$'s of 1 V should be possible. From XPS measurements we find the only remaining oxide is Ga$_2$O and that each passivant has additional bonding states associated with it, with the NHC bonding state having a notably larger bonding energy than the sulfides. We also use basic DFT calculations to predict the structure of the passivation/GaAs interface and find that the NHC strongly deforms the GaAs surface. Continued work in this project would involve more carefully analyzing the bonding and stability of the passivating molecules, and more fully exploring how NHC choice affects passivation quality.

3.3 Non-Epitaxial Heterojunction GaAs Solar Cells
Motivated by our ability to passivate GaAs within an order-of-magnitude as well as GaInP, we also worked on making solar cells exploiting these passivated surfaces. There is not much literature on non-epitaxial GaAs heterojunction solar cells due to the fantastic performance of epitaxial heterojunctions, but we are not the first group to try. Previous groups have made hole-selective contacts using graphene, PEDOT:PSS (an organic, p-type, ITO analogue), and amorphous silicon [225, 28, 31, 144, 146]. Among these devices, the highest reported $V_{oc}$ is 820 mV using doped graphene and an anti-reflection coating. Taking this result at face value (I have a number of concerns regarding the fabrication process and results), it is troubled by the cost of graphene. At the time of writing, for a 20% efficient solar cell, according to the Graphene Supermarket, graphene costs $5 \, \text{W}^{-1}$. Even if this were to decrease two orders of magnitude the graphene would still be cost prohibitive.

The device design we used was motivated by perovskite devices and the emerging use of high workfunction metal oxides as hole selective contacts in silicon PV [7, 252]. The general schematic of the devices we used is shown in Fig. 3.12a. The cell consisted of a bulk GaAs wafer with an ohmic back contact, a
passivated surface, a layer selective to hole transport, a metal oxide to set the surface Fermi level, and then a thin, semi-transparent contact for lateral charge transport and probing. The associated, notional, band diagram is shared in Fig. 3.12b. In the band diagram, we have drawn the conduction and valence bands and the Fermi level, and the background shading indicates the corresponding material from part a. To motivate the band-diagram, working from the bulk to the surface, we start with an n-type wafer with band-bending at the surface. On top of that is a passivation layer with a dipole. We have an idea of the dipole direction from past-measurements of how thiol and NHC monolayers change the work-function of gold [120, 285]. The dipole can be quite large; 1 eV for hydrocarbon thiols and 1.5 eV for NHCs. For the metal oxide / organic interface there is a universal 0.3 eV misalignment between the oxide and organic despite the oxide’s extremely deep work function of over 6 eV [125, 82]. From [125] we further know that the band-bending in the organic decays over just a few nm.

Discussing these devices is challenging because their reproducibility is poor and their performance does not have many obvious trends with device processing. The approach I’ve decided to take is discussing the highest performing devices and then covering a couple attempts to more systematically understand the device performance. To highlight the inadequacy of these devices, despite at least fifty different variations, I never broke 850 mV $V_{oc}$, while I broke 900 mV on my first try with the diffusion doped GaAs cells.
Highest $V_{oc}$ Cells

The highest $V_{oc}$ cells I fabricated used a wafer doped to $5 \times 10^{17} \text{ cm}^{-3}$ with a CuGe ohmic back contact, passivated with sulfide, a 5 nm thick layer of TCTA, 5 nm of WO$_3$ and PEDOT:PSS for lateral charge transport with gold contacts defined by a shadow mask. The chip was isolated into cells by cutting the PEDOT with a razor blade, so we can not normalize the photocurrent, but the main variable of interest is the voltage.

The CuGe back contact was fabricated according to [190]. After making the contact we etched the wafer surface with 3 : 1 : 50 H$_3$PO$_4$ : H$_2$O$_2$ : H$_2$O by volume for five minutes [175]. We then did sulfur passivation as in the previous section. The TCTA we spin coated from chlorobenzene solutions between 0.1% and 1% concentration and annealed in air at 100°C for 1 min. The TCTA thickness was calibrated on silicon control samples by ellipsometry. The WO$_3$ was deposited by thermal evaporation. The PEDOT:PSS was spun coat at 3000 rpm for 1 min followed by annealing at 130°C for 5 min. We used Heraeus Clevios Al 4083 doped to 10% by mass with dimethylsulfoxide (DMSO) and isopropanol (IPA). The DMSO increases the conductivity of the film [147, 34]. We found PEDOT on silicon control samples typically had sheet resistances of 500 $\Omega/$□.

The light, dark, and ideality factor curves for the highest $V_{oc}$ device, which used 5 nm of TCTA, are shared in Fig. 3.13. The good news is that this device achieved 840 mV $V_{oc}$ and the dark IV curve corresponds roughly to a diode. This $V_{oc}$ is a record for non-epitaxial GaAs heterojunction solar cells. The first confusing point is is that the device clearly does not obey superposition; where the light IV curve is the dark IV curve plus a constant accounting for the photogenerated current. This suggests that the dominant current sources are different in the light and the dark. As more evidence for something odd going on, extrapolating the exponential portion of the dark-IV curve suggests the device should have a $V_{oc}$ of over a Volt. Finally, the diode ideality factor is reliably greater than 2.

To further illustrate the unusual ideality factor behavior, we share the IV curves of a device with 820 mV $V_{oc}$ where a lower series resistance allows us to see more of the diode behavior at high-bias in Fig. 3.14. In particular looking at the dark IV and ideality factor, we see an extremely unusual behavior where the ideality factor is decreasing, as if a diffusion current is starting to take over,
Figure 3.13: IV curves for the highest $V_{oc}$ organic heterojunction solar cell. From left to right, illuminated, dark, and dark diode ideality factor.

Figure 3.14: IV curves for a lower $V_{oc}$ organic heterojunction solar cell that best demonstrates unusual ideality factor behavior.

but the IV curve abruptly kinks and has a constant ideality factor of 4.5 for the last 200 mV of bias. Intriguingly, extrapolating the portion of the low-ideality factor part of the curve to the short circuit current roughly predicts the correct $V_{oc}$. This suggests that something is happening at around 700 mV that fundamentally changes the conduction mechanism.

**Device $V_{oc}$ Analysis**

To understand why 700 mV is special, we refer to Fig. 2. of [287]. In this work, the authors measured the band bending as a function of illumination for GaAs passivated very similarly to our process. Because their device is not contacted, it is analogous to one of my solar cells operating at $V_{oc}$ because the external current is zero. In their experiment, the maximum light intensity approaches one-sun intensity, 0.1 W cm$^{-2}$, and the sulfide passivated GaAs is in the flat-band condition at one-sun intensity. Further, the dark band-bending is roughly 700 meV. If the hole quasi-Fermi level is pinned to the dark band-bending position then applied voltages greater than the dark band-bending
will put the surface into a flat-band condition. The reasons why the current saturates at this point is that there is no longer a potential barrier to induce a thermionic emission current and so other conduction mechanisms need to take over.

Another puzzling issue is the extremely low $V_{oc}$. Both the TRPL and PLQY measurements imply a $V_{oc}$ around 1 V, and we can also do a worst-case estimate assuming a 1 ns carrier lifetime, a 10 µm characteristic length, 30 mA cm$^{-2}$ light current, and a $1 \times 10^{17}$ cm$^{-3}$ background doping using the relation:

$$V_{oc} = \frac{kT}{q} \ln \left( \frac{J_L * \tau * N_D}{q * L_C * n_i^2} \right),$$

(3.6)

which results in 940 mV $V_{oc}$. A possible resolution to this problem is that the hole electrochemical potential at the surface is less than that in the bulk at $V_{oc}$. Both the PL measurements and the above equation are estimating the quasi-Fermi level splitting in the bulk, but what we measure is the difference between the electron and hole chemical potentials and the device surfaces. For the surface and bulk hole electrochemical potentials to be different, there must be a hole current flowing because the gradient of the hole chemical potential is:

$$\frac{dE_{Fp}}{dx} = \frac{J_p}{p\mu_p}.$$  

(3.7)

Charge neutrality and the zero external current boundary condition then also demands there be a balancing electron current. Assuming the organic effectively rejects electrons, the missing voltage is then due to poor hole conductivity from the bulk of the GaAs to the surface which makes sense if at applied biases near $V_{oc}$ the surface electric field is repelling holes. This opposing electric field is plausible because the measured $V_{oc}$ is greater than the dark band bending measured by [287]. This effect was in fact predicted much earlier and can been seen in the lower band diagrams of Fig. 1. in [80]. Attaining a good $V_{oc}$ with these devices, then, requires maximizing the band-bending to minimize the denominator of Eqn. 3.7 by increasing $p$.

To illustrate this point, we used AFORS-HET to calculate the $V_{oc}$ of a 5 µm thick GaAs Schottky barrier solar cell with $1 \times 10^{17}$ cm$^{-3}$ n-type bulk doping where the only recombination pathways are the bulk radiative recombination and at the front surface where $S_p = \text{inf}$, and $S_n$ and the Schottky barrier height are variables [265]. The results for this model are shared in Fig. 3.15a and we have included a particular contour for our best 840 mV $V_{oc}$. What this plot
shows is, that within the context of this model, if $S_n = 250 \text{ cm s}^{-1}$, as measured by TRPL, then the barrier-height is around 1 eV, which is among the largest measured Schottky barriers of GaAs surface devices [162]. If, alternatively, $S_n$ is large then the band-bending covers the entire bandgap. To illustrate how the $V_{oc}$ poorly correlates to the bulk Fermi-level splitting, we plot the difference between the electron and hole Fermi levels at the rear contact that has $S_p = 0$ in Fig. 3.15b. Distressingly, we do in fact find that the bulk QFL splitting can be large even when the solar cell $V_{oc}$ is low. For example, for a 0.5 eV barrier height, the $V_{oc}$ is less than 400 mV, while the bulk QFL splitting is 950 mV. This means that even within the context of this relatively simple device that PL CANNOT predict solar cell $V_{oc}$. Finally, we share the band diagrams for $S_n = 100 \text{ cm s}^{-1}$ and barrier heights of 0.5 eV and 1.2 eV in Fig. 3.16. Although the conduction and valence bands look nearly identical despite the 0.7 eV difference in initial band bending, the hole Fermi level behavior is very different. For the 0.5 eV barrier height device we can see the holes losing lots of potential from the device bulk toward the surface, whereas for the 1.2 eV barrier height device the hole chemical potential is flat.

From the PL data and the above analysis we determine that the highest priority for improving these devices is maximizing the dark band-bending as a function of surface treatment and device structure. An excellent way to measure the band-bending is measuring the conductance of an underlying channel that shrinks as the surface band-bending increases [15]. Unfortunately, this experiment requires an epi-wafer order to make the n-type channels on a semi-insulating substrate. It’s tempting to try to characterize these devices by Mott-Schottky analysis, but the several nm thick layers between the metal and semiconductor make the analysis intractable because we then need detailed knowledge about the band-bending, carrier concentrations, and potential drops across each layer.

**Device Design Exploration**

**TCTA Thickness**

We also did a couple more systematic explorations to try and understand how device design accepts performance. In the simplest we varied the TCTA thickness of the high-$V_{oc}$ cells. The results are shared in Fig. 3.17 for thicknesses ranging from zero to 30 nm. That the $V_{oc}$ peaks at 5 nm is indirect evidence
Figure 3.15: GaAs Schottky barrier solar cell performance. a) $V_{oc}$ of a GaAs Schottky solar cell as a function of barrier height and electron recombination velocity. b) Same as a except the difference between the electron and hole Fermi levels at the back contact.
Figure 3.16: Band diagrams of GaAs Schottky solar cells for two different barrier heights under illumination and at $V_{oc}$. 
for the Schottky model from the previous section because at 0 nm it’s just a defective Schottky diode, and then as the TCTA thickness increases the band-bending in GaAs decreases because the potential is being dropped over the TCTA instead of the GaAs.

**Passivation and Hole Transport Layer**

In a more detailed investigation, we made a series of diodes with varying passivation and hole-transport layer and measured the IV characteristics. We used the sulfide, DTT, and NHC passivation recipes from the previous section, and then spun-coat PTAA, TCTA, or polystyrene (PS) from 0.1% by weight solutions in chlorobenzene that was annealed in air at 100°C for 1 min. We then evaporated 5 nm of MoO$_3$ and then used a shadow-mask to define 20 nm thick silver contacts with an area of 0.135 mm$^2$. The choice of polystyrene was inspired by its usage as a passivating tunnel contact in some perovskite solar cells [270]. The collated current-voltage data for these devices is shared in Fig. 3.18. Each super-column corresponds to a passivation strategy and each super-row to a hole-transport layer. For each device we present the dark,
light, and ideality factors with both forward and reverse bias sweeps.

The first thing to notice with these solar cells is that none of them obey superposition. With $J_{sc} > 1 \text{ mA cm}^{-2}$, all of them should have $V_{oc}$ greater than their measured values. This indicates that the conduction mechanisms
that determine the dark-IV behavior are different from those that set $V_{oc}$ under illumination.

Commenting on the hysteresis, it is interesting that the polystyrene devices have minimal hysteresis while the PTAA and TCTA devices have $V_{oc}$ hysteresis up to half a volt. A piece of complementary information is that the polystyrene devices have minimal $J_{sc}$, suggesting that the polystyrene device current is dominated by electron transport. Taken together, this suggests that the hysteresis is controlled by holes in the hole-transport layer.

The lack of hole-transport through the polystyrene also allows us to put upper-bounds on the achievable $V_{oc}$ because $V_{oc}$ will have to be before the electron leakage current through the polystyrene is equal to the one-sun photocurrent. By extrapolating the exponential part of the dark IV curves, the electron limited $V_{oc}$ is 1.1 V, 1.0 V, and 1.4 V for the sulfide, DTT, and NHC passivated samples, respectively. The fact that all of these are much larger than the measured $V_{oc}$ suggests electron leakage current is not what is limiting these devices’ $V_{oc}$.

An interesting trend with the passivation is that for the PTAA and TCTA devices the $J_{sc}$ dramatically decreases when the passivation is the NHC, but the exact opposite is true for the polystyrene devices where sulfide and DTT passivation have tiny $J_{sc}$, but the NHC/polystyrene device has measurable (albeit an order of magnitude smaller than the others) $J_{sc}$. In addition, only the polystyrene sulfide and DTT devices have ideality factors of roughly 1.5, while the photoactive PS/NHC has one of 2.5, which is in line with the unusually high ideality factor of the other photoactive devices. The DTT polystyrene device is actually a pretty good dark diode in its own right. We also recognize that all the photoactive devices have ‘S’ shaped illuminated IV curves which is indicative of internal barriers to hole conduction [210]. There are not obvious trends of $V_{oc}$ with passivation or hole-transport layer. This is actually in line with some work in the perovskite community showing $V_{oc}$ does not depend on the hole-transport layer ionization energy [45]. The weak dependence of $V_{oc}$ on processing is not promising for making high $V_{oc}$ devices.

**Hole Transport Layer Doping**

Due to the S-shaped IV curves in the previous section we also explored doping the hole transport layers to try to reduce internal barriers to hole-transport.
We produced devices as above except we doped the PTAA using tris (pentfluoroophenyl) borane at 10% molar according to the method of [148]. We also tried doping TCTA but the films did not spin-coat cleanly and the solution did not turn orange like the PTAA did, suggesting the borane-amine adduct did not form due to the larger ionization energy of TCTA than PTAA. The IV curves of these devices passivated with sulfide, DTT, and NHC are shared in Fig. 3.19.

Comparing the doped PTAA devices to the devices in Fig. 3.18, the first thing we can notice is that the hysteresis is much smaller, with the smallest undoped hysteresis being 80 mV for the undoped devices compared to the DTT devices here being hysteresis free. This indicates that slow hole states were responsible for the hysteresis and increasing the hole conductivity reduced their relevance. Secondly, the NHC passivated devices here have much better $J_{sc}$ and the DTT passivated devices have a less pronounced S-shape. The sulfide device, though, has an unusual light-IV curve. One possible explanation is that the dopant is a strong acid and it may have attacked or otherwise reacted with the sulfide passivation while the low-reactivity of the hydrocarbons backs of the DTT and NHC protected the GaAs surface. Another interesting feature of these devices is that the $J_{sc}$ is much larger. This may be due to the doped PTAA helping with lateral transport and enabling carrier collection from beyond the metal. Laser beam induced current and mesa isolation are both experiments that would settle this issue. Despite the qualitatively better-shaped IV curves, though, the $V_{oc}$ remains basically the same as the undoped devices, suggesting nothing dramatic has changed with respect to the GaAs band bending or surface recombination velocity.

**N-type Contacts**

With this device architecture, we also explored n-type carrier selective contacts on passivated p-type GaAs. Using a variety of n-type layers including:

- TiO$_2$ nanoparticles
- SnO$_2$ nanoparticles
- TiO$_2$ deposited by ebeam or ALD
- C$_{60}$ by evaporation
Figure 3.19: IV curves of devices using doped PTAA as the hole transport layer as a function of passivation.
• PTCDA by evaporation

We were never able to exceed a couple hundred mV. This is easily rationalized by the Schottky model described in the previous section where the n-type band-bending will by the bandgap minus the p-type band-bending. If the n-type bending is 1 eV, then the n-type bending will be 0.4 eV, and the cells will be doomed to poor voltages.

**Conclusion**

In this section, we have explored a variety on non-epitaxial GaAs heterojunction solar cells. The highest \( V_{oc} \) we achieved was 840 mV; a record for this class of devices. Within the context of a simple passivated Schottky diode model this corresponds to excellent surface passivation and a Schottky barrier height of 1 eV, among the highest in the literature. Exploring a wide variety of passivation and hole-transport strategies we were unable to do qualitatively better. If this work were to be further pursued, the main objective should be maximizing the Schottky barrier height. A good way to do this is measuring the conductivity of a depleted n-type channel and trying to minimize its conductivity as a function of passivation, overlayer, and extreme work-function metal. If a Schottky barrier height greater than 1.15 eV or so can be achieved, then efforts to fabricate high-efficiency devices make sense. Two issues not addressed are device stability and large-scale fabrication. In particular with respect to large-scale fabrication we would need to use a TCO for lateral charge transport. We did experiment with ITO, but sputtering ITO killed the cells. One possible way forward could be using more gentle ITO deposition conditions where the depositing atoms have thermalized to room temperature [216]. A final lesson from the passivation results and the diode results is the PL does not imply \( V_{oc} \). The bulk quasi-Fermi level splitting and the Fermi-levels at the surface do not have a necessary relation, even for high quality materials like monocrystalline GaAs.
Chapter 4

GAAS NANOWIRE SOLAR CELLS

4.1 Introduction
In this chapter, I will share my work on simulating and fabricating GaAs nanowire solar cells. The goal of this work was to contribute to low-cost GaAs solar cells. Nanowire array solar cells are potentially low cost because nanowires are remarkably efficient optical absorbers and they can be exfoliated from wafers which provides an additional tool besides ELO and spalling to make thin III-V devices [63]. In the simulation work, we predict efficiencies over 30% are possible, but require excellent passivation and heterojunctions due to severe short-base diode effects. In the fabrication work, we made GaAs nanowire arrays using metal-assisted-chemical-etching and exfoliated them in polymer. The poor reproducibility, yield, general pain, and fact that I was never actually able to make a cell was not encouraging from the perspective of mass fabrication.

In the initial planning of the project, we thought GaAs passivation was well-enough understood that making non-epitaxial heterojunctions would be simple. So, once I was able to make nanowire arrays, I then turned to the heterojunction work of the previous chapter. The poor performance of those cells basically doomed this project because making the contacts was not going to be as remotely straight-forward as hoped.

Something of general use that come from this is code for importing charge generation profiles calculated by Lumerical FDTD into TCAD Sentaurus. This is helpful because Lumerical offers a fantastic FDTD solver but relatively limited device solver while Sentaurus has a fantastic device solver but a limited FDTD solver. The idea of the method has been previously presented by [202], but their provided code does not work and is uncommented, while here we will provide a working, commented example in Appendix A.

4.2 GaAs Heterojunction Nanowire Solar Cell Simulation
This section is largely a reprint of [107]. Reprinted with permission from the copyright holder, IEEE.
Introduction

There has been tremendous progress in III-V nanowire photovoltaics within the last decade [90, 97, 293, 1, 158]. The current experimental efficiency record is 17.8%, and computational work suggests efficiencies greater than 19% are imminently realizable [261]. However, in order for any novel solar cell design to be economically viable and rapidly scalable, it must be low cost in terms of material usage, processing cost, and capital investment [200]. Current experimental and modeled nanowire solar cells all consist of nanowires either epitaxially grown on wafers or etched into epitaxial layers grown on a wafer. The capital and materials cost contained in the substrate and the processing and capital cost of epitaxy preclude the economically viability of these solar cells [76].

With these economic constraints in mind, we propose a design space for economically viable III-V nanostructure solar cells and predict their performance. Central to determining the design space are the principles of producing many solar cells from one substrate and eliminating epitaxial processing.

Producing many solar cells from one substrate is essential to distribute the material and capital cost of the substrate over many solar cells. Eliminating epitaxy is vital for minimizing the processing and capital costs of each solar cell. Eliminating epitaxy and mandating wafer reuse implies that the nanostructures must be etched into a bulk III-V substrate and subsequently exfoliated, with the substrate being reused many times. The benefit of this is apparent when considering that a III-V nanostructure must only be 1-2 micrometers tall to efficiently absorb sunlight, but a wafer is 500 micrometers thick, implying hundreds of solar cells can be produced from one wafer [63]. Disallowing epitaxy also implies that charge carrier separation must instead be accomplished through diffusion doping or non-epitaxial heterojunctions.

To motivate and explore the nanowire design space we consider four design classes: thin film homojunction, thin film heterojunction, axial nanowire homojunction, and nanowire heterojunction. Starting with one-dimensional thin film devices allows us to cleanly lay out the device physics problems that limit the performance of planar homojunction devices [286]. These physics arguments then help us understand the limited performance of homojunction nanowire devices and why heterojunction nanowire devices can achieve much higher performance.
We show that homojunction devices are incapable of achieving high efficiencies due to the large number of photogenerated carriers throughout the nanowire supporting recombination at the back ohmic contact. The nanowire heterojunction, though, is predicted to achieve over 30% efficiency with realistic device properties.

**Model Details**

In the case of the homojunction thin-film cells we consider 20 µm and 1 µm thick GaAs slabs. Half of the slab is doped p-type and the other half is doped n-type, both with carrier concentration 1E18 cm\(^{-3}\). We assume uniform carrier generation in the entire GaAs slab with an integrated generation rate equivalent to 30 mA cm\(^{-2}\) of photocurrent. For the heterojunction thin-film device we consider a 1 µm thick p-GaAs slab with a 1 \times 10^{15} \text{cm}^{-3} background carrier concentration, but which is capped with on one side with SnO\(_2\) and on the other with CuSCN layers 10 nm thick. The cell geometries are illustrated in Fig 4.1a. In the case of doped nanowires cells, we consider a square array of GaAs nanowires with pitch 600 nm, height 1.5 µm, and radius 150 nm on a thick gold substrate. The array is infilled with a dielectric of refractive index 1.4, and covered with a planar layer of ITO 41 nm thick.

In the case of heterojunction nanowires cells, we consider a thick layer of gold covered with 15 nm of CuSCN as a hole transport layer. On the CuSCN is a square array of GaAs nanocones with pitch 600 nm, height 1.5 µm, lower radius 170 nm, and upper radius 25 nm. Nanocones are chosen due to their superior light absorption compared to nanowires [64]. The array is partially infilled with a dielectric of refractive index 1.4. The dielectric and GaAs cones are sequentially, conformally coated with 25 nm of SnO\(_2\), as an electron transport layer, and 41 nm of ITO for lateral charge transport. The device structures are illustrated in Fig. 4.1b. We choose CuSCN and SnO\(_2\) as charge transport layers due to their use in high-efficiency perovskite solar cells [108, 11].

To simulate the cells, we first used Lumerical FDTD to calculate the optical generation profile under AM1.5g illumination. We used the built-in optical constants for GaAs and gold, and assume the ITO has a carrier concentration of 5 \times 10^{20} \text{cm}^{-3} [98]. We measured the optical constants of SnO\(_2\) and CuSCN using spectroscopic ellipsometry. The SnO\(_2\) films were prepared by diluting Alfa-Aesar 15% SnO\(_2\) colloidal nanoparticle 1:4 with water and spin-coating
Figure 4.1: Schematics of simulated solar cells. a) Schematics of simulated thin-film solar cells. Left is thin-film cell with doped regions for carrier collection. Left is heterojunction cell with SnO2 and CuSCN for carrier collection. b) Schematics of simulated nanowire solar cells. Left is doped nanowire solar cells. Right is heterojunction nanowire solar cell with variable thickness dielectric separating the SnO2 and CuSCN.
on silicon at 3000 rpm for 30 s. CuSCN films were prepared by dissolving 35 mg mL$^{-1}$ Strem Chemicals 99% CuSCN in saturated ammonia water and spin-coating on silicon at 2000 rpm for 30 s. The measured optical constants are plotted in Fig. 4.2.

The optical generation profile is then loaded into Matlab where the generation profile is compressed to two dimensions by averaging over the azimuthal angle:

$$\bar{G}(x, z) = \frac{1}{2\pi} \int_0^{2\pi} G(x \cos(\theta), x \sin(\theta), z) d\theta,$$

(4.1)

where $\bar{G}$ is the 2-D generation profile, $G$ is the 3-D profile and $z$ is along the axis of the device. We compress to two dimensions to reduce the electrical simulation run-time. $\bar{G}$ is then loaded into TCAD Sentaurus for device physics simulation using cylindrical symmetry [202]. We ignore asymmetry due to op-
Table 4.1: Parameters used in device physics simulations. $\chi_e$; $\phi$: electron affinity or work function, eV. $E_g$: bandgap, eV. $\mu_{e/g}$: mobilities, cm$^2$ V$^{-1}$ s$^{-1}$. $m^*_e/h$: effective masses. $\tau_{e/h}$: lifetimes, ns. $\epsilon$: dielectric permittivity.

<table>
<thead>
<tr>
<th></th>
<th>SnO$_2$</th>
<th>CuSCN</th>
<th>Au</th>
<th>ITO</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi_e$; $\phi$, eV</td>
<td>4.5</td>
<td>1.8</td>
<td>5.4</td>
<td>4.2</td>
</tr>
<tr>
<td>$E_g$, eV</td>
<td>3.6</td>
<td>3.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_{e/g}$, cm$^2$ V$^{-1}$ s$^{-1}$</td>
<td>10/10</td>
<td>0.01/0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m^*_e/h$</td>
<td>0.4/0.4</td>
<td>1/1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_{e/h}$, ns</td>
<td>1/1</td>
<td>$1 \times 10^3 / 1 \times 10^5$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>12</td>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

tical interactions between the nanocones, which are small [65]. The generation in the SnO$_2$ and CuSCN is included.

The GaAs is modeled with radiative, bulk, surface, and Auger recombination. The radiative recombination rate is assumed to be the same for the bulk, planar case. We do not consider photon recycling, which avoids costly self-consistent calculations. The bulk non-radiative recombination is modeled by a midgap donor state with concentration $1 \times 10^{16}$ cm$^{-3}$, electron cross-section $2 \times 10^{-16}$ cm$^2$, and hole cross-section $4 \times 10^{-18}$ cm$^2$ [150]. This bulk state is intended to model the EL2 defect present in bulk GaAs wafers. The surface recombination is modeled by a uniformly distributed midgap donor state with width 0.2 eV, $1 \times 10^{-15}$ cm$^2$ capture cross-sections, and variable concentration.

We choose to model the surface recombination with a donor state because GaAs surface recombination velocity depends on injection level and because the GaAs MOSFET community, which produces the highest quality surface passivation, reports their surface states in terms of state densities [249]. The donor surface state models the arsenic double-donor dangling bond responsible for Fermi-level pinning [239]. The GaAs doping is variable.

For the doped nanowire cell, the gold and ITO are assumed to form ohmic contacts. The nanowire base is either n or p-type with carrier concentration $5 \times 10^{16}$ cm$^{-3}$, and the nanowire tip of variable length is doped p or n-type with $1 \times 10^{18}$ cm$^{-3}$ carrier concentration. For the heterojunction nanocone cell, the CuSCN and SnO$_2$ material properties are summarized in Table 4.1, with the gold and ITO forming Schottky contacts to the CuSCN and SnO$_2$. Thermionic emission is included at all heterointerfaces. The nanocone doping is variable.
Results

We first outline the processes that fundamentally limit the performance of thin-film homojunction devices. To motivate the inquiry, we first consider that a 20 µm thick homojunction device achieves a $V_{oc}$ of 1000 mV and a $J_{sc}$ of 8.3 mA cm$^{-2}$, while a 1 µm thick device achieves a $V_{oc}$ of 956 mV and a $J_{sc}$ of 15.9 mA cm$^{-2}$. These results suggest two interesting puzzles: why are the internal quantum efficiencies (IQE’s) so low, with IQE decreasing with thickness, and why does voltage increase with thickness?

To tackle the IQE puzzle, we first consider the band diagrams of the 1 µm device at $J_{sc}$, plotted in Fig. 4.3. To achieve high IQE, the electron/hole electrochemical potential would ideally decrease/increase monotonically as ap-

Figure 4.3: Bandstructure of 1 µm thick thin-film device at short-circuit. Note the change in slope of the minority carrier electrochemical potentials midway in the quasi-neutral regions indicating their traveling in the wrong direction.
proaching the cathode/anode except near the anode/cathode where the ohmic contact boundary condition pins the electrochemical potential to its equilibrium level. Examining the $J_{sc}$ band-diagram we instead see that the minority quasi-Fermi level slope changes sign halfway through the each doped layer, implying many of the carriers are traveling to the wrong contact or recombining on their way to the junction. This makes sense because minority carriers generated in the quasi-neutral bulk do not experience any driving forces besides diffusion to the ohmic contact or junction, both of which impose the same boundary condition on the quasi-Fermi level. Hence, the origin of low IQE is the ohmic contact boundary condition on the GaAs surface and the large quasi-neutral region. The reason the thick device has lower EQE is that minority carriers on average have to travel substantially further before becoming majority carriers on the other side of the junction.

Considering $V_{oc}$, we would naively expect the thin device to achieve higher voltages due to the carriers being confined in a smaller volume which would increase the quasi-Fermi level splitting. To understand why this does not occur, we consider the contributions to the carrier recombination current from the contacts versus applied bias in the dark versus the light for the 1µm and 20µm devices, which are plotted in Fig. 4.4. Ideally, the contact recombination current would be zero, but we instead see that recombination is dominated by the contacts, and that the situation is worsened by illumination. In the dark, the contact recombination currents have ideality factors of one, indicating their source as diffusion of holes/electrons from the p/n type regions to the cathode/anode that fail to recombine in the n/p type regions. Illumination increases the contact recombination current by generating minority carriers throughout the material that reduce the electrochemical potential gradient necessary to drive minority carriers to the contacts by putting a floor on the minority carrier conductivity through the relation:

$$\frac{dE_n}{dx} = \frac{J_n}{\mu_n n},$$  

where $E_n$ is the electrochemical potential, $J_n$ is the current density, $\mu_n$ is the mobility, and $n$ is the carrier concentration. The 1µm device supports much larger contact recombination currents than the 20µm device indicating that increasing the thickness is giving injected minority carriers enough time to recombine in the bulk.
Figure 4.4: Total minority carrier recombination currents at the contacts for the 1 µm and 20 µm thick homojunction devices. The thicker device suppresses the currents and illumination increases them.

We have now laid out the fundamental problem with 1-D homojunction devices: carriers generated in quasi-neutral regions are not efficiently collected, but quasi-neutral regions are needed to block minority carriers from the contacts. The doped nanowire will suffer the same problem.

To solve this tension between IQE and voltage, we need some way to block minority carriers from the contacts. Heterojunctions with wide bandgap semiconductors solve this problem by having much smaller intrinsic carrier concentrations so that larger electrochemical gradients are needed to drive minority diffusion currents through them and they absorb little light so that there is a lower floor on the minority carrier conductivity. By cladding an undoped 1 µm GaAs slab on one side with SnO$_2$ and on the other with CuSCN, each 10 nm thick and with 1 mA cm$^{-2}$ of photogeneration, the device achieves a $V_{oc}$ of 1081 mV and a $J_{sc}$ of 31.5 mA cm$^{-2}$, a substantial improvement over its homojunction counterpart.
We can now clearly interpret the performance of the doped nanowire solar cell. Fig. 4.5 summarizes n and p-type bodied nanowire cells’ dependence of short-circuit current ($J_{sc}$), open-circuit voltage ($V_{oc}$), fill-factor (FF), and efficiency on tip-region length and surface state density. Integrating the photo-generation in the GaAs nanowire suggests a $J_{sc}$ of 24.7 mA cm$^{-2}$. Two results are immediately striking. First, the $J_{sc}$ does not exceed 18 mA cm$^{-2}$, implying a maximum average IQE of 73%. Second, $V_{oc}$’s of 950 mV are possible with excellent passivation, qualitatively mirroring the performance of the planar homojunction. Fig. 4.6 plots the carrier currents at each contact as a function of bias for the champion simulated device (p-type body, 50 nm tip length, $1 \times 10^{10}$ cm$^{-2}$ eV$^{-1}$ defect concentration). The large electron current at the anode at zero-bias is responsible for the low IQE, and its voltage dependence pins it as the main recombination source determining $V_{oc}$. The source of this large minority carrier current can be understood by considering the band diagram along the axis of the nanowire at zero-bias, presented in Fig. 4.7. The electrons have over 800 meV of electrochemical potential to drop between the p-type bulk and the anode, and there is almost $5 \times 10^{10}$ cm$^{-3}$ electron concentration within the p-type region. This large electrochemical potential that must be dropped along with the large electron concentration supplying ample electron conductivity results in a large electron current headed toward the anode. A simple order-of-magnitude calculation using:

$$\Delta E_n = \frac{J_n \times L}{\mu_n \times n},$$

(4.3)

where $\Delta E_n$ is the change in electron electrochemical potential, $J_n$ is the electron current, $L$ is the wire length, $\mu_n$ is the electron mobility, and $n$ is the electron concentration, suggests an electron current greater than 100 mA cm$^{-2}$. This further highlights the role of the ohmic contact boundary condition and large bulk electron concentration in limiting IQE. The linear voltage dependence at low bias demonstrates that the electrons in the wire are functioning as a resistive shunt. In the exponential region at high bias the anode electron current has an ideality factor of 1.05, which suggests the electrons are diffusing from the cathode to the anode and are unable to recombine in the nanowire bulk.

The low hole current at the cathode can be understood by considering the y-component of the hole current and the electric field, plotted in Fig. 4.8. The hole current does not reverse until within 20 nm of the tip, which also
Figure 4.5: Performance of nanowire homojunction as function of design parameters. The optimal tip length is 50nm and efficiencies are limited to 13%.
Figure 4.6: Current at each electrode versus bias for champion homojunction solar cell. The large electron current at the anode limits cell performance.

corresponds to when the electric field induced by the p-n junction is near zero. Therefore, the electric field is efficiently protecting holes from the cathode. This suggests the utility of using an electric field to repel electrons from the anode, as in a p-i-n or p+-p-n design, but it is unclear how such a device could be fabricated nonepitaxially. Doping the backside of wires after exfoliation is infeasible because the temperature for diffusion or implant activation would destroy whatever material the wires are embedded in.

This large parasitic minority carrier current requires that there be both a large density of electrons in the nanowire bulk and that the electron’s electrochemical potential be pinned to its equilibrium level at the nanowire end. The first
condition is unavoidable. Even simply reducing the injection level does not solve the problem because the bulk electron concentration, and thus conductivity, and electron current are linear in optical generation rate and therefore will always flow as long as some small electrochemical potential is present. The second condition can be relaxed by using heterojunctions.

Wide bandgap semiconductors capping the nanowire will have minimal minority carrier concentration near the metal contacts by virtue of their weak light absorption. This small minority carrier concentration suppresses parasitic minority carrier currents that reduce IQE and $V_{oc}$ . To demonstrate this, we now consider a GaAs nanowire solar cell with uniform $5 \times 10^{16} \text{cm}^{-3}$

---

**Figure 4.7:** Plot of homojunction solar cell band structure near the contacts demonstrating large minority carrier concentrations near the contacts.
p-type doping, but capped with planar layers of SnO$_2$ on top and CuSCN on bottom with $1 \times 10^{10}$ cm$^{-2}$ eV$^{-1}$ densities of states at the heterointerfaces and surfaces. The SnO$_2$ is an electron selective contact, and the CuSCN is hole selective. The device’s integrated photogeneration is 25.1 mA cm$^{-2}$. The $J_{sc}$ is 25.0 mA cm$^{-2}$, the $V_{oc}$ is 1.05 V, the FF 80.7% and the efficiency is 21.2%; switching to heterojunctions has massively improved device performance. Fig. 4.9 plots the carrier currents at each contact as a function of bias. The anode minority carrier current is two order of magnitude smaller than in the doped wire case, and is roughly independent of bias. The small minority carrier currents mean that the recombination is happening inside the device as opposed to the contacts. Fig. 4.10 plots the band structure at short-circuit. The much greater distance from the minority carrier electrochemical potentials to the band edges highlight their radically reduced concentration. We then used a
Figure 4.9: Current at each electrode versus bias for nanowire cell with planar top and bottom heterojunctions. The anode electron current is two orders of magnitude smaller and roughly constant. I-V characteristics are dominated by majority carrier currents. Cathode electron current obscures the anode hole current.
genetic algorithm to optimize device performance by altering the upper and lower radii of the nanowire, its height, its doping, and the thickness of the infill between the SnO2 and CuSCN. The optimal device has a 170 nm lower radii, 25 nm upper radius, is 1.5 µm tall, and has $1 \times 10^{18}$ cm$^{-3}$ p-type doping. Interestingly, the optimization eliminated the dielectric spacer. The interface between the SnO2 and CuSCN does not form a shunt because the SnO2 and CuSCN form a wide bandgap diode whose turn-on voltage is greater than the open-circuit voltage of the solar cell, rendering it electrically benign. The optimal device achieves 32.3 mA cm$^{-2}$ $J_{sc}$, 1.13 V $V_{oc}$, and an overall efficiency of 32.6%, which would be a record single-junction device efficiency. This voltage corresponds to an external radiative efficiency of 30%. At 30% ERE, photon
Figure 4.11: Heterojunction solar cell performance dependence on dielectric infill thickness. Performance monotonically decreases with infill thickness.

recycling effects are non-negligible \[170, 268\]. Accurately accounting for them would require expensive self-consistent electro-optical simulations in addition to calculating the LDOS of the nanowire to determine the spatially varying radiative recombination rate. Because we do not perform these calculations we cannot precisely predict device efficiencies, but we can be certain that they are high.

We then investigated the sensitivity of performance to device parameters by varying the density of interfacial states, the dielectric infill thickness, and the nanowire doping. Fig. 4.11 illustrates the results of varying infill thickness. We find that increasing infill thickness uniformly decreases device performance,
but that thicknesses up to 25 nm are acceptable. After 25 nm device performance degrades primarily by losing $V_{oc}$ and FF. Fig. 4.12 shows the effects of varying doping and interfacial densities of states. p-type devices are generally higher performing, and both n and p-type device performance depends similarly on trap concentrations. In p-type devices, high doping is always preferred while in n-type devices, high doping is preferred only if the trap density is less than $3 \times 10^{12} \, \text{cm}^{-2} \, \text{eV}^{-1}$. The reason for performance increasing with doping is that doping shrinks the width of depletion regions created by charged surface states [186]. We expect that higher doped wafers would also have shorter carrier lifetimes, but without detailed studies on the doping dependence of non-radiative recombination in VGF GaAs wafers, it is difficult to say anything for certain. Lower densities of interfacial states are always preferred, but as long as they are lower than $3 \times 10^{12} \, \text{cm}^{-2} \, \text{eV}^{-1}$, greater than 25% efficiency is possible. To decide if this threshold is reasonable, we look to the GaAs MOSFET community. It is counterintuitive that MOSFETs are an enabling technology for photovoltaics because a MOSFET ideally has zero current flow through the gate insulator, while a solar cell must conduct significant current. However, the only conceptual difference between a MIS capacitor and heterojunction solar cell is the bandgap of the heteropartner ($> 7 \, \text{eV}$ for MIS, $\approx 3 \, \text{eV}$ for PV).

In the last decade, non-epitaxial and low-temperature GaAs MOSFET passivation schemes have been developed where sub-nanometer thick films produce MIS capacitors with defect densities around $1 \times 10^{11} \, \text{cm}^{-2} \, \text{eV}^{-1}$ [33, 238]. This density would enable devices with over 30% efficiency, suggesting the feasibility of fabricating a high-efficiency device.

Conclusions

In this paper, we first determined the device physics mechanisms that limit the efficiency of off-substrate GaAs nanowire homojunction solar cells. The high density of photogenerated carriers throughout the nanowire and the lack of an electric field near the back contact to repel minority carriers severely degrade IQE of the cells. To fix these issues, we then consider a heterojunction solar cell using SnO$_2$ and CuSCN for charge separation. An optimized device is predicted to achieve over 32% efficiency. The sensitivity of device performance to device design is analyzed and efficiencies over 30% should be possible by borrowing processing techniques from the GaAs MOSFET community.
Figure 4.12: Performance of nanowire heterojunction as function of design parameters. P-type devices perform better and efficiencies greater than 30% are possible for defect concentrations less than $1 \times 10^{12}$ cm$^{-2}$ eV$^{-1}$. 
4.3 GaAs Nanostructure Fabrication

We also worked on fabricating GaAs nanostructures for eventual device integration. The two fabrication techniques we explored were metal assisted chemical etching (MACE) and anisotropic chemical etching. In this work, we mainly focused on 110 oriented substrates so that etched structures could be cleanly cleaved from the substrate, which we demonstrate in our work on MACE. The idea was that we would develop a simple method for making GaAs nanostructures, rip them off the wafer, and then use that wafer over-and-over again to realize cost savings.

Metal Assisted Chemical Etching

Metal assisted chemical etching (MACE) is an electrochemical technique for highly-anisotropic semiconductor etching where the material directly underneath a metal layer is preferentially etched [103, 85]. The basic mechanism is that an oxidizer, typically $\text{H}_2\text{O}_2$, oxidizes the metal and then the hole in the metal migrates to the semiconductor where it manifests as a broken bond that the etchant can attack [132]. The etching depends on the Schottky barrier height and semiconductor doping and bend-bending where having the semiconductor surface Fermi level near the valence band is desirable because it makes it easier for the hole to inject from the metal to the semiconductor and n-type doping is preferable because it localizes the hole to the semiconductor surface. The process for silicon using gold, $\text{H}_2\text{O}_2$, and HF is illustrated in Fig. 4.13 which is adapted from [132]. MACE has already been used for GaAs device fabrication, where GaAs LEDs defined by MACE showed enhanced luminance compared to planar control devices [173].
Using the methods of [173] we explored MACE of GaAs. Something we found was the whether MACE worked at all depended strongly on the cleanliness of the Au/GaAs interface. Without paying special attention to keeping the Au/GaAs interface organic free the etch would fail. This makes sense because an organic residue to stop holes from being injected, or would create islands where the gold couldn’t sink into the semiconductor. The three techniques we had success with were:

- Oxygen plasma cleaning of the wafer followed by gold evaporation and defining the MACE pattern by etching the Au with Transene KI/I₂ based gold etchant.
- Gold liftoff where after developing resist the wafer was received an oxygen plasma descumming
- Lift off where we took an oxygen plasma cleaned GaAs chip, evaporated 300 nm of SiO₂ onto it, patterned the SiO₂ and then plasma etched it to make SiO₂ pillars followed by an oxygen plasma to remove the resist. We then evaporated the gold and removed the SiO₂ pillars in HF.

With the first method we struggled to reliably etch the gold and the gold feature had rough edges that imprinted onto the etched GaAs nanowire, as shown in Fig. 4.14. The second method had only roughly 50% yield. The third, despite its greater complexity, was reliable. We typically used gold layers 25 nm thick at evaporated at 5 Å s⁻¹ to ensure the gold would be porous and let the etchant uniformly access the semiconductor [263]. The most reliable etch mixture we found was 1:1 HF: 0.01 mol dm⁻³ KMnO₄ which etched at 2.4 µm h⁻¹. We used substrates doped in the low 1 × 10¹⁷ cm⁻³ n-type range.

In terms of MACE itself, the main result we found was that the etching was orientation independent. We share this if Fig. 4.15. With the same gold pattern and etching for 30 min we found the the etched wire arrays were identical for 100, 110, and 111 (unknown polarity) substrates.

Being able to successfully etch nanowires, we then worked on being peeling them off the wafer in a polymer handle. This has been previously demonstrated for InP using PDMS as the handle [63]. We wanted to develop an alternative process because removing PDMS requires hazardous tetra-n-butylammonium
Figure 4.14: Left, gold layer with holes defined by wet-etching. Right, resulting GaAs nanowires showing the rough edges of the gold mask badly deforming the GaAs.

Figure 4.15: Different orientations of GaAs MAC-etched for 30 min.

fluoride and the dissolved PDMS is not easily recycled. The process we developed involved coating the nanowires in PMMA by spin-coating, and then welding the PMMA to a bead of Instamorph thermoplastic by putting it on a hot plate at 100 °C. The final structure is illustrated in Fig. 4.16a. This is preferable to PDMS because the PMMA is easily dissolved in acetone and both can be recycled because no chemical reaction is happening. We could then remove the plastic by scraping the plastic-wafer interface with a razor blade or tweezer. Fig. 4.16b is an SEM of the 110 oriented GaAs chip post peel. The roughly square regions are the bases of GaAs nanowires that have been cleanly cleaved from the wafer. Inbetween the wires is the gold mask. There are also six nanowires that did not peel with the rest and remain on the wafer. We can immediately see that the gold has badly deformed during the etch, suggesting this is not a viable way to turn a wafer into nanowires without repatterning. Improving the surface preparation may reduce this issue. Fig. 4.16c is a microscope image of the nanowires embedded in plastic. Each point
Figure 4.16: Cross section of GaAs etched with cold citric acid and hydrogen peroxide showing no undercutting and steep sidewalls.

of light is the base of a nanowire scattering light. We can see the yield is not outstanding. Fig. 4.16d is an SEM of nanowires that we attempted to peel from a 100 oriented substrate. We can see that instead of cleanly cleaving as in part c, the wires instead plastically deformed before breaking, indicating the importance of using a crystallographically favorable substrate.

**Anisotropic Chemical Etching**

The work on anisotropic chemical etching was mainly done by Ariel Stiber for an undergraduate thesis under my direction [243]. Probably the most exciting result from the work was finding an etchant that has greater than 10:1 selectivity for etching 110 planes more quickly than 100 planes. The etchant is 3:1 citric acid to hydrogen peroxide, where the citric acid is then a 1:1 by mass ratio mixture of citric acid to water and we did the etching in a refrigerator at 2°C. The 110 etch rate is 600 nm h⁻¹. With traditional etching solutions
the 110 and 100 etch rate are fairly similar [175]. We suspect the citric is so selective because it effectively chelates, and passivates, the 100 GaAs surface when a gallium plane is exposed. A cross sectional picture of the etching is shared in Fig. 4.17, with the major crystal planes highlighted and looking at etched trenches of increasing width from left to right. We can see that the etch is making deep trenches instead of a standard, bowed, undercutting profile and that the etch eventually terminates in sharp triangular profiles. This indicates that the etch rate slows dramatically near the 100 etch plane instead of precisely at it. Unfortunately it was not clear to us how to use this etch to make photovoltaically useful structures.

**Conclusion**

In summary, we demonstrated the isotropic MACE of GaAs, the exfoliation of GaAs nanowires from a 110 oriented substrate, and a novel anisotropic GaAs etchant that reveals 100 planes. The main challenges that still need to be overcome for MACE are mainly related to process control, in particular more reliably preparing the GaAs/gold interface. From the silicon MACE literature gigantic aspect ratios are possible and there is no obvious reason that should not also be possible with GaAs. With the anistropic etch this is an interesting new tool for GaAs fabrication, but its application for a useful device is open.
Chapter 5

EFFECTIVELY TRANSPARENT CONTACTS

This is some work I did at the start of my PhD trying to figure out how to reduce optical losses due to the front metal contacts of solar cells. The reason this work is not particularly useful is that: a) it is not clear how to incorporate the idea at low-cost into a module and b) module manufacturers are, justifiably so, overwhelmingly conservative when it comes to changing module design. This is why I was so careful to propose a solar cell design for my diffusion doped cells that is a drop-in replacement for silicon cells. Solar panels need to last the advertised 25 years for companies to stay in business and even issues like spontaneous delamination of standard modules was not considered solved until well into the 2000’s [20, 208]. In addition, the problem here is not as dire as it used to be because cell manufacturers have done a good job of making the printed metal a Lambertian scattering surface so that much of the light is reflected back into the cell by total-internal-reflection at the glass/air interface.

5.1 Introduction

Increasing photovoltaic cell and module efficiency via improved photonic design is an important route to decrease the cost of electricity because increases in cell efficiency reduce the cost per Watt of all the components of a photovoltaic system. Most commercial silicon solar modules still operate at less than 18% efficiency; well below the detailed-balance efficiency limit of 29.8% for silicon.[255, 13]

A fundamental problem limiting the efficiency of front-contacted solar cells is reflection and photocurrent losses arising from the metal contacts.[153] Contacts are necessary to achieve low electrical resistance, but typically incur shadowing losses as they reflect light away from the solar cell, reducing the absorbed photocurrent. A number of strategies have been developed to minimize the reflection losses. One established cell design that eliminates front metal reflection losses are inter-digitated back contact cells.[134, 233, 284] While inter-digitated back contact cells completely eliminate the front metallization reflection losses and hold the current silicon efficiency records, they
requiring a more complex and costly fabrication process. For conventional front-contacted solar cells with screen printed contacts, mitigating photocurrent losses requires a judicious tradeoff between emitter sheet resistance and front contact grid finger and busbar spacing and area. For cells featuring low carrier mobility in the emitter layer, including thin film cells as well as high efficiency heterojunction intrinsic thin layer (HIT) Si cells, transparent conductive oxides (TCOs), are used to reduce emitter sheet resistance for these cells, and high aspect ratio effectively transparent contacts can significantly reduce photocurrent losses. [211, 123, 295, 122] High aspect ratio silver triangular contacts have been demonstrated to work by redirecting incident light impinging on grid fingers into the solar cell, and have demonstrated front surface transparency of over 99% with excellent sheet conductivities, but require an unconventional process to form the high aspect ratio front contact metallization and have not been demonstrated on textured silicon solar cells.

Another loss mitigation strategy is prismatic encapsulation in which the front surface of the module acts as a lens directing light away from the contacts, but this method reduces the acceptance angle of solar radiation by actively concentrating the sunlight.[19]

In this paper, we demonstrate a technique for eliminating solar cell front contact reflection losses by modifying the polymer encapsulation design. By forming triangular cross-section structures of suitable geometry and whose refractive index is lower (e.g. n=1) than the surrounding encapsulation, and which are aligned with respect to the contacts, incident light can undergo total internal reflection due to the refractive index contrast and be redirected into the absorbing material instead of being scattered out of the module. This proposed design is compatible with existing commercial solar cell contact designs and fabrication methods, is independent of the light incidence angle, and is relatively simple to implement. We first discuss an analytic ‘toy’ model that describes light interactions with such structures, followed by numerical ray-tracing simulation results that indicate an increase in photocurrent. We also demonstrate proof-of-principle experimental results demonstrating the efficacy of this strategy for reduction of busbar shadow losses for conventional screen-printed solar cells.
5.2 Theory

Fig. 5.1a illustrates the issue presented by photocurrent losses arising from front contact reflections. The solar module is modeled as a dielectric superstrate which mimics the optical performance of module glass laminated to the cell by a layer of encapsulant with index of refraction \( n_e \), a cell with planar metal contact lines with perfect reflectivity, and a solar cell absorber with perfect absorption. Light incident at an angle \( \theta_i \) on the module is refracted by the air/encapsulation interface, reflected by the metal, and then escapes from the module, giving rise to photocurrent loss due to reflection from the metal contact. Fig. 5.1b illustrates a proposed solution to this reflection problem.

By placing a triangle cross-section region with index of refraction \( n_t < n_e \), height \( h \), and width \( w \), and sufficient aspect ratio, incident light is reflected from the triangular structure/encapsulant interface by total internal reflection (TIR). These reflected rays are redirected into and absorbed by the cell, instead of escaping the module, thus increasing the absorbed photocurrent. Using Snell’s law, simple constraints can be derived for the total internal reflection condition. For a given triangular structure geometry, with indices of refraction, and incident angle, the angle of incidence of the light on the triangle/encapsulant interface, \( \theta_t \) is:

\[
\theta_t = \tan^{-1} \left( \frac{2h}{w} \right) - \sin^{-1} \left( \frac{n_e \sin(\theta_i)}{n_t} \right).
\]  

(5.1)

Setting \( \theta_t \) to the critical angle for the triangle/encapsulant interface and solving for \( \theta_i \) returns the maximum angle of incidence for which TIR is supported:

\[
\theta_{i,\text{max}} = \sin^{-1} \left( n_e \sin \left( \tan^{-1} \left( \frac{2h}{w} \right) - \sin^{-1} \left( \frac{n_t}{n_e} \right) \right) \right).
\]  

(5.2)

An interesting feature of this equation is that for \( n_e/n_t > \sqrt{2} \), it is possible to choose the aspect ratio \( 2h/w \) to be large enough that \( \theta_{i,\text{max}} = \pi/2 \), implying that incident light can be redirected by total internal reflection for any angle of incidence. The physical interpretation of this result is that it is the index contrast at which a vertical boundary will reflect light traveling along the air/encapsulant interface’s escape cone. We thus refer to triangular structures that satisfy the constraints for TIR at \( \theta_t = \pi/2 \) as ‘perfect’ triangular structures.
Figure 5.1: Illustration of how low-refractive index regions can improve photocurrent of solar cells. (a) Demonstration of why metal contacts reduce photocurrent. Light incident on the solar cell reflects off the cell’s front metal contact, reducing the photocurrent of the cell. (b) Example of how lower refractive index regions can improve photocurrent. A low refractive index triangle is placed above the metal contact. Assuming a favorable relationship between the angle of incidence, triangle geometry, and index contrast, the light reflects off the encapsulant/triangle interface and into the solar cell.

5.3 Simulation

To study the light management efficacy of triangular structures, we performed two types of ray tracing simulations. First, we modeled a single contact and studied how the absorbed photocurrent changed with angle of incidence and geometry of the structure above the contact. Second, we modeled light interaction with a full solar cell with a grid-and-busbar front contact structure, and estimated the achievable improvement in annual energy harvest for fixed-axis and tracking module orientations. In all simulations, the solar cell was modeled as a perfect isotropic absorber, the contacts as specular reflectors, the triangular structures were assumed to consist of air voids formed in an encapsulant whose index of refraction is 1.55 as a proxy for standard solar cell encapsulants.

In the single-contact simulations, the solar cell is 20mm on each side with encapsulant 6mm thick and 40mm per side on top. The contact is a 2mm wide strip running down the center. The ray source is a large sheet just above the cell whose angle of emission is variable. We first simulated the photocurrent of such a single-contact cell coated with a conventional encapsulant layer with homogeneous refractive index, and then simulated triangular cross section structures with various aspect ratios and normalized the absorbed
photocurrents calculated for these structures to the response of a cell with a conventional encapsulant layer.

Fig. 5.2a illustrates the setup and our convention for angle of incidence. A line parallel to the z axis is first rotated in the y-z plane by $\psi$ and then rotated in the x-z plane by $\omega$. The line is then parallel to the direction of incidence. The intent is that $\psi$ is the angle of incidence along the contact and $\omega$ is the angle of incidence perpendicular to the contact. Fig. 5.2b contains the results of the simulations. In each subplot, the horizontal axis is the angle of incidence along the contact, $\psi$, and the vertical axis is the angle of incidence perpendicular to the contact, $\omega$. The color represents the relative photocurrent change compared to the plain cell. The geometry of the triangle is defined in the bottom left. The red region in the false-color map indicates a relative photocurrent improvement of approximately 10%. Since the contact area comprises 10% of the cell area, a 10% improvement is photocurrent density indicates complete elimination of reflection-related photocurrent losses by the triangular cross-section structures. However in other regions of this plot, the performance abruptly degrades due to the breakdown of the total internal reflection condition at the encapsulant/void interface, in some regions leading to lower performance than for a cell with homogeneous encapsulation. In the bottom-right figure, we see that decreasing aspect ratio to 1.5:1 reduces the range of incidence angles for which the triangular structure is effective. However in regions where total internal reflection occurs, we observe the same improvement in photocurrent as in the bottom-left figure. The top-right right figure indicates that by increasing aspect ratio, the angular performance is recovered. The top-left figure indicates a triangular structure with the minimum aspect ratio to performance as a perfect triangle. The simulation confirms that a perfect triangle will enhance the photocurrent for all angles of incidence. The deviations at extremely oblique angles of incidence are artifacts caused by light incidence at the edges of the ray tracing simulation cell, which would not be observed at similar angles for cells with realistic geometries.

We also studied the properties of other triangular and polygonal cross section structures. We describe here two particular cases: an assymetrical ‘sawtooth’ triangular cross section structure in Fig. 5.3a and a polygon ‘house-shaped’ cross section structure in Fig. 5.3b. Fig. 5.3a depicts the geometries and some sample light rays for the sawtooth and house-shaped cross sections. The
response of the sawtooth cross sectional structure depends on which side of the contact the light is incident from, owing to its broken symmetry. When light is incident from the left, impinging on the vertical sidewall, total internal reflection occurs for a very large range of incident angles, while light incident on the right side will experience total internal reflection over a much smaller range of incident angles. The house-shaped cross section structure is intriguing because the ‘walls’ exhibit total internal reflection over a wide range of incident angles while the ‘roof’ supports total internal reflection over a narrower angular range. Fig. 5.3b shows simulation results for the sawtooth structure and for the ‘house-shape’ structure. The upper subplot illustrates the asymmetry of the sawtooth behavior where photocurrent is enhanced when light is incident from the left side, but degraded when incident from the right side. The lower subplot indicates reflectance variation with angle for the house-shaped structure, indicating that while total internal reflection breaks down for the roof, the overall photocurrent response is not degraded compared to the homogeneously encapsulated reference structure, whereas the photocurrent response is degraded for low aspect ratio triangular cross section structures. We also observe a smoother transition between the regions where total internal reflection is supported and where it breaks down. These shapes could be advantageous for situations for which the solar irradiance is known to exhibit unequal intensities for incidence from the left and right sides or a weaker angle variation of photocurrent loss desired and perfect triangles are not possible.

To estimate the impact on annual energy production for a module with cells featuring triangular cross section contacts, we compared results of simulations for the yearly energy harvest of solar cells with and without triangular cross section structures for fixed axis illumination and for single-axis and two-axis tracking strategies. Using the solar simulator built into the commercial LightTools ray tracing software package, we utilized historical data for direct and diffuse radiation at different locations and times of day and year. The simulated solar cell is a square with area $100\,cm^2$, three 1.25mm busbars spaced 25mm apart, and 75$\mu m$ finger contacts with a pitch of 2mm, leading to a total metal coverage of 7.5%. The triangular void structures are located within the polymer encapsulation layer at a height 15$\mu m$ above the contacts in order to simulate a structure in which the metal contact layers are coated by a thin polymer layer rather than exposed to air. The busbar triangles are 2.95mm wide and 8.7mm tall. The finger triangles are 245$\mu m$ wide and 730$\mu m$ tall.
overall encapsulation thickness is 9mm. The cell was simulated in Phoenix, AZ for a year. In the stationary and single-axis tracking simulations the cell is assumed to be inclined 26 degrees to the south. The results are compiled in Table 5.1. We find that significant increases in yearly energy production could be achieved by implementing triangular voids above the contacts and that the enhancement is not substantially different for fixed axis and tracking modules. In addition, the increase in energy harvest was independent of cell orientation. The increase in energy production being greater than the metalized area and the east/west and two-axis tracking performing differently is a simulation artifact due to use of a ray tracing simulation cell where the edges of the module are important.

### Table 5.1: Simulated annual improvements in annual energy harvest for a solar panel equipped with perfect effectively transparent contacts

<table>
<thead>
<tr>
<th>Tracking Strategy</th>
<th>% Increase In Annual Energy Relative to Plain Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>7.6</td>
</tr>
<tr>
<td>East-West</td>
<td>8.0</td>
</tr>
<tr>
<td>Two Axis</td>
<td>8.1</td>
</tr>
</tbody>
</table>

5.4 **Experiment**

We also performed proof of principle experiments demonstrating reduction of reflection losses from the busbar of a commercially available solar cell. Monocrystalline Si front-contacted solar cells with busbar widths of 1.5mm were used, and a negative mold for the encapsulation sheet structure was printed using a 3D printer. The mold is a square basin in a polymer sheet 2.5mm deep and 25mm across with a triangle running down the middle of width 1.5mm and height 1.5mm. The encapsulation sheet structure is formed by casting and curing polydimethylsiloxane (PDMS), and then brushing a thin layer of uncured PDMS onto the interior of the triangular grooves to smooth out roughness arising to the finite pixel resolution of the printed mold. The sheet structure is then placed on top of a solar cell that has been coated with a thin layer of uncured PDMS to form a good optical contact; and following this step, the PDMS was cured. Fig. 5.4a is a photograph of the completed structure. A thin white line is all that can be seen in reflection from the busbar. This indicates light from the busbar is unable to reach the camera due to scattering by the triangle. Reciprocity then implies that normally incident light is incident on the silicon instead of the busbar. The finger contacts that
appear to run through the triangular structures are reflections of the contacts from the triangular structure surface. Fig. 5.4b is a reflection micrograph taken by a microscope equipped with a 5X lens and numerical aperture = 0.13. A strong reflection is observed from the finger contact while reflection from the busbar on the left is eliminated by the triangular cross section structure. This is further evidence that light is being redirected into the silicon. Light striations in the image along the busbar are visible and are due to the finite pixel size features printed by the 3D printer that defined the mold. Fig. 5.4c is a spatially-resolved laser beam induced current (LBIC) measurement utilizing a 543nm laser. The image is normalized so that the grid finger contacts are not photoactive and that the illuminated silicon is just at a saturated intensity level. The grid finger contact appears to run through the triangle because light reflecting off the triangle then reflects off the finger contact. The photocurrent improvement is complete to within the measurement accuracy of our apparatus. Fig. 5.4d,e,f are the analogous images taken of a cell with conventional homogeneous encapsulation. The contrast between these images clearly demonstrates the efficacy of the triangular cross section structures for light harvesting.

5.5 Discussion and Conclusion

The presented results suggest triangular cross-section structures above the contacts could be an effective way to reduce reflection losses from the front busbar and grid finger contacts of a solar cell. However many questions must be addressed before this technique can be implemented as a viable and manufacturable process. Foremost, high-quality master molds for defining aligned triangular cross section structures in the polymer encapsulation must be developed. Triangular structures for busbars can be easily defined using traditional machining techniques. Effective transparent triangular structures for finger contacts are more difficult to machine, since they are about 100 microns wide, and the required mold height is several hundred microns tall, which is too large to be defined using, for example, inductively coupled plasma etching of silicon due to the etch always operating in the infinitely-wide-trench regime, resulting in uncontrollably re-entrant etch profiles.[92] Anisotropic wet-etching of silicon is likely to excessively limit the achievable aspect ratio. Non-conventional lithographic techniques such as grayscale and inclined exposure may be promising ways forward.[86, 204] Another important question is how triangular cross
section structures would affect module reliability. For example, would the gases in the triangular cross section structures degrade the metallic contacts over time, or possibly ease delamination of the module?

In conclusion, we have demonstrated an approach to eliminate reflection losses from solar cell front contacts. By defining triangular cross section structures with an index of refraction less than the surrounding encapsulation above the contacts, total internal reflection can direct the light away from the contact and into the solar cell absorber. Simulations of triangle structures and other shapes show that the photocurrent can be enhanced over a wide and range of incidence angles, and simulation of the energy harvest by a cell integrated over over a year demonstrates that large gains in yearly energy harvest are possible. Proof of principle optical characterization experiments utilizing commercially available components demonstrate that this technique does enhance photocurrent and can be extremely simple to implement.

5.6 Acknowledgments
This material is based upon work primarily supported by the Engineering Research Center Program of the National Science Foundation and the Office of Energy Efficiency and Renewable Energy of the Department of Energy under NSF Cooperative Agreement No. EEC-1041895. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the National Science Foundation or Department of Energy.

The authors would like to thank John Lloyd for helpful discussions while setting up the LightTools simulations.
Figure 5.2: Exploration of angle of incidence dependence of photocurrent change as a function of triangle geometry. (a) Schematic illustrating the simulated geometry and the scheme for encoding the angle of incidence. (b) Each subfigure illustrates the angular response of triangles with different geometries. The geometry is indicated in the lower left of each subfigure, the color represents the photocurrent change, and the axes are the angles of incidence defined in part a.
Figure 5.3: Examining different structures and their performance as a function of angle of incidence. (a) Two possible non-triangular structures for the low-index region above the contact. The upper structure is a sawtooth, and the other is a ‘house.’ The sawtooth is 2.5mm tall and the walls and roof of the house are each 2mm tall. (b) The relative change in photocurrent due to the structures as a function of angle of incidence.
Figure 5.4: Experimental demonstration of functioning effectively transparent contacts for a solar cell busbar. (a) Photograph of solar cell busbar with triangular void in PDMS above the busbar. (b) Reflection microscopy image of a busbar with a triangular void above it and a finger contact without a triangle. (c) Spatially resolved laser beam induced current measurement of a busbar with a triangular void above it and a finger contact without a triangle. The dashed line is the approximate location of the busbar. The color indicates the measured photocurrent, scale is arbitrary units. (d,e,f) Images analogous to (a,b,c) except the PDMS layer is planar.
DESIGNING SOLAR CELLS TO OPERATE ON VENUS

The final piece of my work I’d like to discuss is my work on designing solar cells for Venus. While this is not directly related to my work on low-cost PV, being able to accurately simulate the high-temperature EQE did require taking into account Urbach tails, thus illustrating another situation where they became important. The main conclusion of this work is that the best solar cell for Venus is a GaInP single junction cell. The results are shared in [79, 78]. Reprinted with permission from the copyright holder, John Wiley and Sons and IEEE.

6.1 Introduction

There is a growing interest in long-duration mission on Venus [74, 14]. Venus is at the opposite extreme of how the greenhouse effect can influence climate (with Mars at the opposite end with a thin atmosphere and Earth in the middle) and the origin of its extreme greenhouse effect is not well understood. In addition, the source of UV-blue absorption that is responsible for about 50% of the planet’s solar spectrum absorption is a complete mystery [256]. Resolving these questions will require long-duration missions in the atmosphere, with one proposed concept being a PV-powered balloon. There has been previous, preliminary work on evaluating the suitability of commercial triple-junction solar cells for operation on Venus [135].

Venus presents a number of challenges for photovoltaics. The surface temperature of 465 °C is, on its own, a large problem. First, the group-V volatility at this temperature means the cells need to be well-encapsulated to not degrade. Second, it is warm enough for dopants to start diffusing. Third, even GaAs has an extremely large intrinsic carrier concentration, \( n_i \), at this temperature. At 700 K \( n_i = 1 \times 10^{15} \text{ cm}^{-3} \), which we can compare to its room temperature value of \( 2 \times 10^6 \text{ cm}^{-3} \). Because the diode saturation current goes as \( n_i^2 \), it increases by a factor of \( 2.5 \times 10^{17} \). Very roughly plugging this into a diode model the \( V_{oc} \) at 30 mA cm\(^{-2}\) is 150 mV; the cell is basically dead.

The incident solar spectrum also changes dramatically in both spectrum and
intensity as a function of altitude. In orbit, the solar intensity is 2.6 kW m$^{-2}$, while it is only 50 W m$^{-2}$ on the ground. The total downward flux as a function of altitude and its spectral decomposition are shared in Fig. 6.1 and is reproduced from [257]. The large changes in spectrum mean that it will be difficult to design a multijunction solar cell that operates well at all altitudes because one of the subcells will always be current-starving the other, and the low surface intensity makes eliminating shunts critical. Incidentally, the atmosphere also has a large density of mineral acids (HF, HCl, H$_2$SO$_4$) that need to be encapsulated against.

My portion of the work consisted of two main pieces. First, building a solar-cell model that accurately reproduced experimental data for cell performance at high temperatures. And second, using the validated model to predict the performance of existing devices on Venus and modify the design to optimize their performance.
Table 6.1: Device stack for comparing simulated and measured EQE measurements.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Material</th>
<th>Thickness, µm</th>
<th>Doping, cm$^{-3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARC</td>
<td>Al$_2$O$_3$</td>
<td>0.035</td>
<td></td>
</tr>
<tr>
<td>ARC</td>
<td>TiO$_2$</td>
<td>0.075</td>
<td></td>
</tr>
<tr>
<td>Window</td>
<td>AlInP</td>
<td>0.025</td>
<td>$n-1 \times 10^{17}$</td>
</tr>
<tr>
<td>Top Cell</td>
<td>GaInP</td>
<td>0.1</td>
<td>$n-2 \times 10^{18}$</td>
</tr>
<tr>
<td>Top Cell</td>
<td>GaInP</td>
<td>0.7</td>
<td>$p-3 \times 10^{16}$</td>
</tr>
<tr>
<td>BSF</td>
<td>AlInP</td>
<td>0.03</td>
<td>$p-1 \times 10^{17}$</td>
</tr>
<tr>
<td>Tunnel</td>
<td>AlGaAs</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Tunnel</td>
<td>GaAs</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Windows</td>
<td>AlGaAs</td>
<td>0.03</td>
<td>$n-1 \times 10^{18}$</td>
</tr>
<tr>
<td>Bottom Cell</td>
<td>GaAs</td>
<td>0.1</td>
<td>$n-1 \times 10^{18}$</td>
</tr>
<tr>
<td>Bottom Cell</td>
<td>GaAs</td>
<td>3.5</td>
<td>$p-5 \times 10^{16}$</td>
</tr>
<tr>
<td>BSF</td>
<td>AlGaAs</td>
<td>0.1</td>
<td>$p-1 \times 10^{17}$</td>
</tr>
<tr>
<td>Buffer</td>
<td>GaAs</td>
<td>0.3</td>
<td>$p-1 \times 10^{17}$</td>
</tr>
<tr>
<td>Wafer</td>
<td>GaAs</td>
<td>500</td>
<td>$p-5 \times 10^{18}$</td>
</tr>
</tbody>
</table>

6.2 High Temperature Solar Cell Model Validation
Before being able to predict solar cell performance on Venus we had to first be able to reproduce existing measurements of solar cells operating at high-temperatures. The two PV parameters we focused on reproducing were the solar cell’s EQE and $V_{oc}$ as a function of temperature.

Focusing first on the EQE, we worked on reproducing the 300$^\circ$C EQE measured by [241] and by JPL using cells provided by MicroLink Devices Inc. especially for this project. The device stack we modeled is presented in Table 6.1. The AlGaAs is the 33% alloy and the GaInP and AlInP are the lattice-matched compositions to GaAs. For the room-temperature electrical and optical properties we used the values provided by Sentaurus. To calculate the optical constants as a function of temperature we first calculated the bandgap narrowing and then shifted the dispersions in energy by the corresponding amount. For absorption coefficients less than 8000 cm$^{-1}$, we inserted by hand an Urbach tail. For the temperature dependence of the Urbach energy we used the results from [111] for GaAs which found a linear relationship between temperature and Urbach energy and used the same value for all materials in the device.

The GaAs band-edge EQE at 300$^\circ$C for the MicroLink cell, the [241] cell, and two versions of our model are shared in Fig. 6.2a. Without Urbach
absorption the model predicts both a shorter wavelength and more steep EQE at the band-edge. With the Urbach absorption we predict an EQE between the two experimental measurements, which we consider a success. The reason we focused on 300°C is because the stated goal of the project was to optimize a solar cell for operating at 21 km altitude which is 300°C. We also calculated the AM0, room-temperature $J_{sc}$ and predicted a value of 15.3 mA cm$^{-2}$ while the experimental value is 15.5 mA cm$^{-2}$, another excellent agreement between model and experiment.

We then worked on correctly calculating the cell $V_{oc}$. In Fig. 6.2b, we show the $V_{oc}$ measured by JPL as a function of temperature and in purple the $V_{oc}$ predicted by Sentaurus using its built-in Shockley-Reed-Hall temperature dependence model. We were able to bring it into line with experiment by modifying the lifetime temperature dependence to:

$$\tau(T) = \tau(T = 300 \text{ K}) \times \left(\frac{T, \text{ K}}{300 \text{ K}}\right)^{3.85}.$$  \hspace{1cm} (6.1)

The $V_{oc}$ versus temperature with this lifetime model is the yellow curve in 6.2b which trends well with the experimental data.

### 6.3 Venus Solar Cell Optimization

Now having a model for the optoelectronic properties of the solar cell as a function temperature validated against measurements at JPL we turned to optimizing the solar cell design for Venus. We modeled illumination on Venus by using the spectrum of [176] and multiplying it by 0.68 to derate the ARC for diffuse illumination because the light will be scattered by the clouds.

We explored three main strategies for designing the solar cell. First, figuring out the optimum GaInP thickness as a function of altitude. Second, modifying the windows and BSF layer thicknesses. Third, exploring a GaInP-only design.

The reason the optimal GaInP thickness for the tandem changes a function of altitude is that the shape of the incident light’s spectrum changes. In particular, at low altitudes, molecular absorption removes a lot of red light so a thick GaInP cell will photon-starve the GaAs cells and reduce the photocurrent of the tandem. To find the optimal GaInP thickness we swept the thickness of the GaInP base from 0 µm to 1 µm and calculated the solar cell power at each altitude. The results are shared in Fig. 6.3. As expected, the optimal thickness changes with altitude, starting at 300 nm at 54 km, reducing to 200 nm between
Figure 6.2: Simulated and experimental cell performance as a function of temperature. a) Comparison of GaAs band-edge EQE at 300°C for two experimental solar cells and simulations that do or do not include Urbach tail absorption. b) Validation of recombination model predicting $V_{oc}$ as a function of temperature.
Figure 6.3: Solar cell efficiency as a function of altitude and GaInP top cell thickness. The optimal GaInP thickness is plotted as a line following the maximum power at each altitude.

10 and 50 km, and finally approaching the thick cell at the surface. There is no single tandem design that is optimal at all altitudes, but it is interesting that 200 nm works over such a wide range; basically from the bottom of the clouds down to ground haze.

We also explored if increasing the thickness of the BSF and window layers would improve the high-temperature performance. We suspected that at high-temperatures the thermionic emission barrier that confined minority carriers would be ineffective and that it would take thicker layers to suppress minority carrier currents. Indeed, we found that by changing the window thickness from 25 nm to 100 nm and the BSF thickness from 25 nm to 200 nm, we could improve the power generated by a 2J GaInP/GaAs at the surface with 800 nm thick GaInP from 3.31 W m\(^{-2}\) to 4.10 W m\(^{-2}\), a 24% improvement.

In light of this improvement and the optimal GaInP thickness vs. altitude calculation we considered the performance vs altitude for three different solar cell designs, presented in Fig. 6.4a-c. The first is a tandem with 200 nm thick GaInP and thin windows and BSF layers specifically optimized for 10-
50 km. Next is a mid-altitude solar cell with a 3.5 µm single-junction GaInP cell with thick window / BSF (100 nm/200 nm) layers. The final one is a surface/low-altitude solar cell, a 2J cell with 800 nm GaInP and thick window/BSF (100 nm/200 nm) layers. The 2J cell with 200 nm GaInP and thin window/BSF layers is best at high altitude, because the two subcells are current matched and the temperature is low enough that the GaAs subcell is producing substantial power. The dominance of the single-junction GaInP cell at moderate altitudes is perhaps surprising because from Figure 6.3 we found a single GaInP thickness is optimal for a tandem over the whole mid-altitude range. However, it can be made sense of by considering that the current-matched tandem requires a very thin GaInP cell to achieve current matching by allowing high-energy photons to reach the GaAs cell. Because the performance of the GaAs subcell is rapidly decaying with temperature, it becomes parasitic compared to simply having a single GaInP cell absorbing all the light. At low altitudes, the 2J cell with thick window/BSF layers is more efficient because the temperature is high, and for the surface spectrum, a 2J with a 800 nm thick GaInP subcell is current matched with a GaAs bottom subcell, so the GaAs subcell is providing an additional modest amount of power without absorbing photons that are more efficiently converted by the GaInP subcell.

Fig. 6.4d displays the power and Fig. 6.4e displays the efficiency of each solar cell design as a function of altitude. The efficiency is what would be measured under a solar simulator with only normally incident light. The best performance we find at 21 km/300 °C is 22% from the optimized single-junction GaInP cell. There are two main opportunities to improve this result. First, the anisotropic lighting means we could take advantage of the upward flux and put two cells back-to-back in a bifacial configuration and significantly increase the power generated [258]. Second, the combination of diffuse illumination and the ARC’s angle dependence means we lose 30% of the light due to reflection. This could be improved by surface texturing or by using a module design that more efficiently utilizes diffuse illumination, such as a luminescent solar concentrator which is currently a topic of research [181].

6.4 Conclusion
In this work, we developed a device physics model for simulating solar cells operating in the atmosphere of Venus. Correctly simulating the cells required
Figure 6.4: Summary of optimized solar cell design and performance as a function of altitude. a-c) Solar cell designs optimized for operation at high altitude, mid altitude, and low-altitude, respectively. d) Electrical power generated by each cell design as a function of altitude. e) Efficiency of each cell design as a function of altitude.
including the Urbach absorption, which is typically ignored, and modifying TCAD Sentaurus’s built-in carrier lifetime temperature dependence models.

With a validated model, we then developed a number of solar cell designs suitable for operation on Venus. We found that thickening the window and BSF layers of III-V heterojunction cells improved the high-temperature performance. The best cells at ground, mid, and high-atmosphere were a tandem with optically thick GaInP, a single junction GaInP cell, and a 200 nm thick GaInP tandem, respectively. These results will help mission planners design power systems for future Venus missions.

With the proposed solar cells being fairly standard III-V designs, the critical remaining challenge is testing these cells at high-temperature in corrosive environments similar to Venus’s atmosphere. In preliminary testing (Fig. 8 of [79]) the solar cell thermally decomposed and traditional organic encapsulants will not survive on Venus, highlighting the need to develop and test high-temperature and corrosive resistant solar cell encapsulants.
I would like to share some closing thoughts about PV research. First, I would like to caution young scientists about getting into this field if their goal is to improve the world. That was how I went in, and learning in detail all the economic problems with renewable energy and what global warming has in store has left me a sadder, angrier person. The best way to get into this field is to be excited for the science for the science’s sake, not for the potential applications.

That being said, there are still opportunities in solar cell research. Any useful research must be laser-focused on fabricating or enabling devices that use low-capital cost processing and can achieve electricity costs comparable to or lower than silicon. Perovskites are still a promising candidate for this, but their stability and yield issues must be resolved with the greatest urgency. An additional way forward is still using monocrystalline material but going even thinner. 2D materials such as $(\text{Mo/W})(\text{S/Se})_2$ are promising because of their strong light absorption and low defect concentrations, but there is lots of fundamental work remaining revolving around fabricating large areas of material, achieving absorption over the whole solar spectrum, and finding high-performance device designs. I think it would be really interesting to see a detailed theoretic analysis of 2D solar cell performance limits. One extension I would suggest is using f-sum rules to first calculate required material thicknesses for light absorption, independent of complex optical design [292, 228]. This can then guide choices about number of layers and how that affects the bandgap and overall cell performance.
BIBLIOGRAPHY


Daniel Cuypers et al. “Sacrificial Self-Assembled Monolayers for the Passivation of GaAs (100) Surfaces and Interfaces”. In: Chemistry of Materials 28.16 (Aug. 2016), pp. 5689–5701. ISSN: 0897-4756. DOI: 10.1021/acs.chemmater.6b01732. URL: http://pubs.acs.org/doi/10.1021/acs.chemmater.6b01732.


[48] R. O. Dendy. *Plasma physics : an introductory course*. Cambridge University Press, 1995, p. 513. ISBN: 9780521484527. URL: https://books.google.com/books?id=puuQM4Dx0zYC%7B%5C%7Dprintsec=frontcover%7B%5C%7Ddq=isbn:9780521484527%7B%5C%7Dhl=en%7B%5C%7Dsa=X%7B%5C%7Dved=2ahUKEwieqde3rLrkAhWUIQIHoEAMQ6AEwAhECAAQA%7B%5C%7Dv=onepage%7B%5C%7Dq%7B%5C%7Df=false.


[52] Electric Power Monthly - U.S. Energy Information Administration (EIA). URL: https://www.eia.gov/electricity/monthly/epm%7B%5C_%7Dtable%7B%5C_%7Dgraphe.php?t=epm%7B%5C_%7D6%7B%5C_%7D07%7B%5C_%7Db (visited on 01/12/2022).

[53] Electricity data browser - Topaz Solar Farm. URL: https://www.eia.gov/electricity/data/browser/%7B%5C%7Dplant/57695?freq=A%7B%5C%7Dctype=linechart%7B%5C%7Dltype=pin%7B%5C%7Dmaptype=0%7B%5C%7Dlinechart=ELEC.PLANT. GEN.57695-ALL-ALL.A%7B%5C%7Dcolumnchart=ELEC.PLANT.GEN. 57695-ALL-ALL.A (visited on 01/12/2022).


[64] Katherine T. Fountaine, Christian G. Kendall, and Harry A. Atwater. “Near-unity broadband absorption designs for semiconducting nanowire arrays via localized radial mode excitation”. In: Optics Express 22.S3 (May 2014), A930. ISSN: 1094-4087. DOI: 10.1364/OE.22.00A930. URL:


179


[106] IEA Sankey Diagram. URL: https://www.iea.org/sankey/%7B%5C# %7D?c=IEA%20Total%7B%5C&%7Ds=Balance (visited on 10/06/2021).


[164] Wondwosen Metaferia et al. “(110)-Oriented GaAs Devices and Spalling as a Platform for Low-Cost III-V Photovoltaics”. In: Conference Record of the IEEE Photovoltaic Specialists Conference (June 2021), pp. 1118–1120. ISSN: 01608371. DOI: 10.1109/PVSC43889.2021.9518754.


Moon Facts. 

10 Things You Didn’t Know About the Moon — Space. 


V. I. Moroz et al. “Spectrum of the Venus day sky”. In: Nature 284.5753 (Mar. 1980), pp. 243–244. issn: 0028-0836. DOI: 10.1038/284243a0. url: http://www.nature.com/articles/284243a0.


[211] Rebecca Saive et al. “Effectively Transparent Front Contacts for Optoelectronic Devices”. In: Advanced Optical Materials (June 2016). ISSN:


[235] SolarStar1. Electricity data browser - Solar Star 1. URL: https://www.eia.gov/electricity/data/browser/%7B%5C#%7D/plant/58388?freq=A%7B%5C&%7Dctype=linechart%7B%5C&%7Dltype=pin%7B%5C&%7Dlinechart=ELEC.PLANT.GEN.58388-SUN-ALL.A%7B%5C&%7Dpin=ELEC.PLANT.GEN.58388-SUN-ALL.A%7B%5C&%7Dmaptype=0%7B%5C&%7Dcolumnchart=ELEC.PLANT.GEN.58388-ALL-ALL.A (visited on 01/12/2022).

[236] SolarStar2. Electricity data browser - Solar Star 2. URL: https://www.eia.gov/electricity/data/browser/%7B%5C#%7D/plant/58389?freq=A%7B%5C&%7Dctype=linechart%7B%5C&%7Dltype=pin%7B%5C&%7Dlinechart=ELEC.PLANT.GEN.58389-SUN-ALL.A%7B%5C&%7Dpin=ELEC.PLANT.GEN.58389-SUN-ALL.A%7B%5C&%7Dmaptype=0%7B%5C&%7Dcolumnchart=ELEC.PLANT.GEN.58389-ALL-ALL.A (visited on 01/12/2022).


[263] Han-Don Um et al. “Versatile control of metal-assisted chemical etching for vertical silicon microwire arrays and their photovoltaic applications”. In: Scientific Reports 5.1 (Sept. 2015), p. 11277. ISSN: 2045-2322. DOI: 10.1038/srep11277. URL: http://www.nature.com/articles/srep11277.


[269] L. C. Wang, P. H. Hao, and B. J. Wu. “Low-temperature-processed (150–175 °C) Ge/Pd-based Ohmic contacts (ρc~1×10-6 Ω cm2) to n-GaAs”. In: Applied Physics Letters 67.4 (June 1998), p. 509. ISSN: 0003-


IMPORTING OPTICAL GENERATION DATA FROM LUMERICAL FDTD TO TCAD SENTAURUS

The purpose of this appendix is to describe the process of importing optical generation data from Lumerical FDTD to TCAD Sentaurus. A number of group members have successfully used this code for simulating a variety of nano electro-optical devices. Assuming you already have an optical generation data file from Lumerical and a meshed structure in Sentaurus, the outline of the process is:

- Get the coordinates of the device regions that need optical generation rates assigned to them.
- Interpolate the optical generation rate from the FDTD mesh to the device physics mesh.
- Import the interpolated generation rates into a Sentaurus compatible file.

To help illustrate the process, the tools with annotations for one of my projects is shared in Fig. A.1. From left to right, the first tool runs a Lumerical FDTD simulation, the second meshes the matching structure in Sentaurus, the third generates a blank optical generation file, the fourth extracts the coordinate list from that blank generation file, the fifth interpolates the generation profile onto the list of coordinates, the sixth then writes the values to an optical generation input file, the seventh then runs the device simulation, and the eighth analyzes the device performance.

Before that though, you need to know the region numbers of the Sentaurus device regions that you want optical generation rates assigned to. Unfortunately SVVisual does not supply this information, so I use the below tcl script:

```#tcl command file the prints to the terminal the region information of structure #invoke with $tdx -tcl get_info.tcl```
Figure A.1: Sentaurus tool flow for a coupled FDTD/Sentaurus nanowire solar cell simulation.

```python
# specify a mesh file that you want the region information of
set f n23_060.tdr
puts "file: $f"
TdrFileOpen $f
# loop through geometries
set ng [TdrFileGetNumGeometry $f]
puts "#geometries: $ng"
for {set ig 0} {($ig < $ng)} {incr ig} {
    set gname [TdrGeometryGetName $f $ig]
    set ns [TdrGeometryGetNumState $f $ig]
    set nr [TdrGeometryGetNumRegion $f $ig]
    puts " geometry $ig: $gname"
    puts " type: [TdrGeometryGetType $f $ig]"
    puts " dimension: [TdrGeometryGetDimension $f $ig]"
    puts " transform: [TdrGeometryGetTransform $f $ig]"
    puts " shift: [TdrGeometryGetShift $f $ig]"
    puts " #states: $ns"
    # loop through states
    for {set is 0} {($is < $ns)} {incr is} {
        set sname [TdrStateGetName $f $ig $is]
        puts " state $is: $sname"
    }
    puts " #regions: $nr"
    # loop through regions
    for {set ir 0} {($ir < $nr)} {incr ir} {
        set rname [TdrRegionGetName $f $ig $ir]
        puts " region $ir: $rname"
    }
}
TdrFileClose $f
```
To get an empty optical generation file, I use the following sdevice command:

```plaintext
#setdep @node|sde@
File {
    Grid = "n@node|sde@_msh.tdr"
    Current = "@plot@"
    Plot = "@tdrdat@"
    #The empty optical generation file
    OpticalGenerationOutput = "n@node@_OGO.tdr"
}

Electrode{
}

Physics {
    Optics(
        OpticalGeneration(
            SetConstant(
                Value=0
            )
        )
    )
}

Plot{}

Math{-CheckUndefinedModels}
Solve{poisson}
```

where getting the file requires at least solving the Poisson equation. We then extract the coordinate list of the regions we care about with the below TDX code. This example extracts the coordinates of regions 0, 2, and 3, and formats them as a csv file.

```plaintext
#setdep @node|dark@
#Specify the empty generation file
set inp n@previous@_OGO.tdr
#specify the file that lists the coordinates of the relevant regions
set outf [open "n@node@_OGO_COORDINATES.txt" "w"]
TdrFileOpen $inp
```
#set the device region
set ig 0
set is 0
set ir 0

#loop through datasets until finding the optical generation one and
#set coords to its coordinates
set nd [TdrRegionGetNumDataset $inp $ig $ir $is]
for {set id 0} {$id < $nd} {incr id} {
    if { [string match OpticalGeneration [TdrDatasetGetQuantity
        $inp $ig $ir $is $id] ] } { set coords [TdrDataGetAllCoordinates $inp $ig $ir $is $id] } }

#repeat for region 2
set ir 2
#concatenate the coordinates of region 2 to those of region 0
set nd [TdrRegionGetNumDataset $inp $ig $ir $is]
for {set id 0} {$id < $nd} {incr id} {
    if { [string match OpticalGeneration [TdrDatasetGetQuantity
        $inp $ig $ir $is $id] ] } { set coords [concat $coords [TdrDataGetAllCoordinates $inp $ig $ir $is $id]] } }

#same as above but for 3
set ir 3
set nd [TdrRegionGetNumDataset $inp $ig $ir $is]
for {set id 0} {$id < $nd} {incr id} {
    if { [string match OpticalGeneration [TdrDatasetGetQuantity
        $inp $ig $ir $is $id] ] } { set coords [concat $coords [TdrDataGetAllCoordinates $inp $ig $ir $is $id]] } }

#covert coords to csv format
regsub -all "\} \n" coords
regsub -all "\} \n" coords
regsub -all "\} \n" coords
regsub -all "\} \n" coords

#write the coordinates
puts $outf "$coords"
TdrFileClose $inp
close $outf

The next step consists of two parts. First is a tcl file that pre-processes a Matlab file to read the correct data files, and then the Matlab file itself. The pre-processing command is:

#setdep @node|tdr_cood@
#get node numbers
set cood @node|tdr_cood@
set node @node@
set FDTD @node|FDTD@

#process matlab file
set orig [open MapCarrierGenerationData.m "r"]
set text [read $orig]
close $orig
set text [string map "node|tdr_cood $cood node $node FDTD $FDTD" $text]
set out [open n@node@_import_gen.m "w"]
puts $out $text
close $out

#run matlab file
exec matlab -r n@node@_import_gen

The corresponding matlab code is:

set input data
xyz=dlmread('nnode|tdr_cood_OGO_COORDINATES.txt');
load('nFDTD.mat');

% 2Dify the generation profile
xnew=0:2.5e-9:max(x);
ynew=min(z):5e-9:max(z);
mx=length(xnew);
ny=length(ynew);
ts=0:0.1:2*pi-0.1;
nt=length(ts);
F = griddedInterpolant({x,y,z},G);
for i=1:length(xnew)
    for j=1:length(ynew)
        for k=1:nt
            g(k)=F(xnew(i)*cos(ts(k)),xnew(i)*sin(ts(k)),ynew(j));
        end
        Gnew(i,j)=mean(g);
    end
end

F2 = griddedInterpolant({xnew,ynew},Gnew);

% Interpolate onto the Sentaurus grid
parfor k =1:length(xyz)
    T(k)=F2(xyz(k,:).*1e-6)*1e-6
    if T(k)<0
        T(k)=1e10
    end
end

% Write absorption data to OGI.TXT file
fileID = fopen('nnode_OGI0.txt','w');
fprintf(fileID,'%1.0f\n',T);
fclose(fileID);
exit

If you do not have Matlab, this code is readily modified for Python, where the non-obvious bit is using `h5py` to open the generation data using:

import h5py
... 
filepath = 'n@node|FDTD@.mat'
arrays = {}
f = h5py.File(filepath, 'r')
for k, v in f.items():
    arrays[k] = np.array(v)

fdtd_x = f['x'][()][0]
fdtd_y = f['y'][()][0]
fdtd_z = f['z'][()][0]
fdtd_G = f['G'][()]

The final step is writing the list of generation rates generated by Matlab or Python to a tdr file. This is accomplished by the below TDX code:

#setdep @node|interpolation@
#manage files
puts "start"
set fdata [open n@node|interpolation@_interp_gen.csv "r"]
set file_data [read $fdata]
close $fdata
set data [split $file_data "\n"]
file copy -force "n@node|dark@_OGO.tdr" "n@node@_OGO.tdr"
set inp "n@node@_OGI.tdr"
puts "fin open and close files"

TdrFileOpen $inp

set ig 0
set is 0
set ir 1
set start 0
set end 0

#find the opt gen dataset
set nd [TdrRegionGetNumDataset $inp $ig $ir $is]
for {set id 0} {($id < $nd) {incr id}} {
    #find those that contain 'optical generation' quantity
if {[string match OpticalGeneration [TdrDatasetGetQuantity $inp $ig $ir $is $id]]} {
    set igen $id
    break
}

puts "found cat gen"
#count the number of data points
set nvc [TdrDatasetGetNumValue $inp $ig $ir $is $igen]
set end [expr $nvc-1]
#overwrite the data
puts "start=$start"
puts "end=$end"
set catdata [lrange $data $start [expr $end]]
set start [expr $end+1]
puts "made cat data"
for {set iv 0} {$iv<$nvc} {incr iv} {
    TdrDataSetComponent $inp $ig $ir
    $is $igen $iv 0 0 [lindex $catdata $iv]
}
puts "wrote cat data"

#now doing the other regions. notice how data is segmented into regions
TdrFileSave $inp
puts "tdr saved"
TdrFileClose $inp
puts "tdr closed"

#file delete n@node|FDTD@.mat

The region numbers and order need to be identical to the previous TDX step because the whole process depends on the order of the list of coordinates being preserved. This code is more verbose because it took me a while to sort all the pigeonhole errors that can occur. You are now ready to do the real device simulation. The command in sdevice for importing the optical generation is:

File {
...
    OpticalGenerationInput = "n@node|import_gen@_OGI.tdr"
...
}
Physics {
... 
Optics(
  OpticalGeneration(
    ReadFromFile( DatasetName = OpticalGeneration )
    ...
  )
)
)
}
Appendix B

SIMPLE THERMODYNAMICS OF GLOBAL WARMING

B.1 Introduction

My goal in writing this is to provide an accessible argument for why CO$_2$ and other greenhouse gases in the atmosphere can cause extreme global warming by building the two simplest models that capture the relevant physics. I started thinking about simple climate models because I figured that if I am trying to convince people to invest 10’s of trillions of dollars on fighting global warming, I should probably feel comfortable with the basic physics of the problem [212]. It turns out the simplest possible model puts a limit on global warming at $17.5^\circ$C. It takes a more complex model of the atmosphere to argue the effects of CO$_2$ on temperature can be large. Interestingly, these two models can be thought of as extremes on a spectrum of understanding how the atmosphere interacts with itself.

Making my argument will require three steps. First, I will outline the overall external sources of energy that heat up the Earth. Second, I will construct the simple model that shows saturation of warming because it includes a lot of the relevant physics and is easy to solve. Third, I will generalize that model to allow a more realistic picture of the atmosphere and use it to argue why global warming does not saturate with CO$_2$ concentration.

B.2 Energy Flows in the Earth-Sun-Space System

What we will end up doing is using conservation of energy between the Sun, Earth, and space to calculate the temperature of the Earth. Therefore, the first thing to do is figure out how much energy the Earth receives from other sources. The sources I will consider are power from the Sun as light or solar wind, power from Earth’s hot core, and power from the cosmic microwave background and other astronomical sources. The relevant power flows are illustrated in Fig. B.1.

Sunlight

Starting with sunlight, we can treat the Sun like a blackbody and use geometry to figure how much of its radiated energy hits the Earth. The power density
of radiation emitted at the surface of the sun follows the Stefan-Boltzmann Law:

$$P_{SL,S} = \sigma T_S^4,$$  \hspace{1cm} (B.1)

where $P_{SL,S}$ is the power density of sunlight at the surface of the Sun, $\sigma$ is the Stefan-Boltzmann Constant, and $T_S$ is the temperature of the Sun. We can then use inverse-square law scaling to figure out the intensity of sunlight incident on the Earth, $P_{SL,E}$:

$$P_{SL,E} = P_{SL,S} \left( \frac{r_S}{d_{E\rightarrow S}} \right)^2,$$  \hspace{1cm} (B.2)

where $d_{E\rightarrow S}$ is the distance from the Earth to the Sun, and $r_S$ is the radius of the Sun. Plugging in the relevant numbers, we get $1.33\text{ kW m}^{-2}$, which can be compared to the real value of $1.36\text{ kW m}^{-2}$ [260]. This number is the power density on the projected circle the Earth looks like from the Sun. It will be more useful later to consider the power density averaged over the whole surface area of the Earth, which is different by a factor of 4, so we will use $P_{SL,E} = 1.33\text{ kW m}^{-2}/4 = 340\text{ watt/m}^2$. 

Figure B.1: Power Sources for the Earth including solar light and wind, thermal energy from the Earth’s core, radio from the cosmic microwave background (CMB), and starlight.
The CMB is a blackbody as well, radiating at 2.73 K [36]. Although much colder than the Sun, it’s still worth considering because it covers the whole sky, while the Sun is only a small portion of the sky. Plugging 2.73 K into the Stefan-Boltzmann equation, we get a power of 3.2 μW m\(^{-2}\), much smaller than sunlight.

Starlight
To estimate the amount of energy from starlight, we can consider that the ground is much brighter on a night with a full moon, so we can use the full moon as an upper bound on the power from starlight. A full moon provides up to 0.3 lux of light, while the Sun is about 100,000 lux, so starlight is also a small contribution of at most 1 mW m\(^{-2}\).[174]

Solar Wind
The solar wind is a continuous stream of charged particles from the Sun’s corona. The pressure exerted by the solar wind is about 5 nPa, and the work done by it can be determine by simply multiplying the wind pressure and velocity, typically 500 km s\(^{-1}\), resulting in a power of 2.5 mW m\(^{-2}\).[48]

Geothermal Heating
The Earth’s core is an extremely hot iron-nickel ball due to radioactive decay and leftover heat from planetary accretion, so some of its heat makes it to the surface through conduction/convection of the mantle and things like volcanoes. The total geothermal heat flux is estimated to be 87 mW m\(^{-2}\).[198]

Overall Power Inflow
Taken together, the Earth’s energy inflow is dominated by blackbody radiation from the Sun, with other energy sources many orders of magnitude smaller. For subsequent calculations, we will use the average power inflow over the surface area of the Earth as \(P_{in} = 340\) W m\(^{-2}\).

B.3 A Simple Model
So we now know how much energy is flowing into the Earth. As the energy flows in, the Earth heats up and starts radiating energy as a blackbody, and keeps heating up until the power the Earth radiates is equal to the power it absorbs. To start out, let us imagine the Earth is a perfectly absorbing ball...
Figure B.2: Normalized blackbody spectra for 300 K and 5800 K. What’s important is that they peak at very different wavelengths, with the 5800 K spectra highest at 0.5 \( \mu \text{m} \), and the 300 K highest at 10 \( \mu \text{m} \).

with no atmosphere. Setting up the energy balance equation:

$$P_{in} = 340 \text{ W m}^{-2} = P_{out} = \sigma T_E^4,$$  \hspace{1cm} (B.3)

where \( P_{in} = P_{out} \) is the fundamental requirement of energy balance, and \( T_E \) is the Earth’s surface temperature, we get \( T_E = 273.8 \text{ K} = 5.1 ^\circ \text{C} \), which is close-ish to the real value of 14.4 \( ^\circ \text{C} \). One interesting thing to note is that the Earth is much colder than the Sun at 5780K. A consequence of this that the blackbody spectra of the Sun and Earth peak at very different wavelengths, as illustrated in Fig. B.2, with the Sun peaking in the visible spectrum at 0.5 \( \mu \text{m} \) (green), and the Earth peaking in the infrared (IR) at 10 \( \mu \text{m} \). This mismatch makes it possible for the atmosphere to behave very differently for visible light from the sun and IR light from the ground.

Let us now think about adding some more details to our model. First, we assumed that the Earth is black and perfectly absorbing, which it obviously is not because I can see a tree from my office that is basically green. It is generally accepted that Earth reflects 29% of the light incident on it, meaning \( P_{in} = (1 - 0.29) \times 340 \text{ W m}^{-2} = 241.4 \text{ W m}^{-2} \). Plugging this into the above
equation for the Earth’s temperature, we get $T_E = 255.4 \, K = -17.7 \, ^\circ C$, which is way too cold. To warm things up, we need to add an atmosphere. Let us think of the atmosphere as a slab on top of the ground that is characterized by some average absorption in the solar spectrum, $A_{Vis}$, and some absorption in the 300 K blackbody spectrum, $A_{IR}$. The power flows so far are visualized in Fig. B.3a with arrows indicating flow directions.

At this point, the visible optical properties are specified, so now we need to think about the IR optical properties of the system. Consider some amount of IR power emitted by the ground, $P_G$. The amount absorbed by the atmosphere is $P_G * A_{IR}$ and the amount that makes it to space is $P_G(1 - A_{IR})$. Now consider some IR power coming from the atmosphere, $P_A$. Because the atmosphere is radiating both up and down, half of the power will go to space, while the other half of the power will go down and be absorbed by the ground. We can now draw out the full energy flow diagram for the system, illustrated physically in Fig. B.3b, and as a graph in Fig. B.3c, where the weights specify how much of the power from the source is flowing in that direction and $P_{sp}$ is the power being radiated to space. From this, we can setup the following system

Figure B.3: Power flows in the Ground/Atmosphere/Space system. a. Visible light power flows. b. IR light power flows. c. Power flows as a graph.
of equations:

\[ P_G = (1 - A_{vis})P_{in} + 0.5P_A \]

\[ P_A = A_{vis}P_{in} + A_{IR}P_G \]  

(Equation B.4)

which specify how the sun pumps energy into the air and ground and how the air and ground couple to each other. Radiation to space is implicitly taken care of by the coupling between the air and ground being non-unitary. Solving these equations, we find:

\[ P_G = P_{in}\frac{A_{vis} - 2}{A_{IR} - 2} \]

\[ P_A = P_{in}\frac{2(A_{IR}(A_{vis} - 1) - A_{vis})}{A_{IR} - 2} \]  

(Equation B.5)

Qualitatively, it is interesting that temperature is not in these equations and that \( P_G \) can saturate with \( A_{vis} = 0 \) and \( A_{IR} = 1 \) producing \( P_G = 2P_{in} \). To calculate the temperature, we can use the fact that all the coupling in our model is through blackbody radiation, and so can use the Stefan-Boltzmann law to convert powers and temperatures. The emissivity of the ground is close to 1, and, from Kirchhoff’s law, emissivity is equal to absorption so the IR emissivity of the atmosphere is identical to \( A_{IR} \). Empirically, we also have the \( A_{vis} = 0.32 \), so that we now have enough information to determine how the Earth’s temperature varies with the atmosphere’s IR absorption, which is plotted in Fig. B.4a.

There are two interesting features. First, the atmosphere’s temperature dependence is non-monotonic in IR absorption. This is because absorption and emission are competing processes, where at low IR absorption the atmosphere needs a high temperature to dissipate its energy, but at higher IR absorption it becomes a better emitter and so its temperature decreases until it starts absorbing most of the IR coming from the ground. Secondly, the temperature of the ground saturates at 17.5°C. We can change the x-axis to be linear in greenhouse gas (GHG) concentration by assuming Beer Lambert absorption:

\[ A_{IR} = 1 - e^{-\beta C} \]  

(Equation B.6)

where \( \beta \) is proportional to the GHG’s IR absorption cross-section, and \( C \) is the GHG concentration. Arbitrarily choosing \( \beta = 1 \) and re-scaling the x-axis, we get Fig B.4b which again shows temperature saturating with GHG concentration. This is puzzling because Venus has extreme global warming
with a surface temperature of 450°C that is not possible in this model. It is also worthwhile to examine the effect of incident power on temperature. This is because some skeptics argue changes in solar output can account for climate change. A ground temperature of 14°C corresponds to a power flux at the ground of 385.5 kW m⁻². From Eqn. B.5, we know this scales linearly with the intensity of sunlight. Over the solar cycle, solar output varies by at most 0.1%, which supplies at most 0.07°C of warming, making it a very small effect[280].
B.4 More Complicated Model

To make the model a bit more sophisticated, we are going to allow the temperature of the atmosphere to vary with altitude by modeling the atmosphere as a stack of thin slices of air that are all coupled to each other through radiation. First, it will be convenient to switch to absorption coefficients that are related to overall absorption from the previous model as follows:

\[ A = 1 - e^{-\alpha h_0}, \]  

where \( A \) is the total absorption, \( \alpha \) is the absorption coefficient, and \( h_0 \) is the thickness of the atmosphere. Let us split the atmosphere into slices with thickness \( dz \) so that there are \( N = h_0/dz \) total slices, with the bottom of each slice at \( z_j = jdz \), where \( i \) is an index that runs from zero to \( N - 1 \). Some amount of light passing through a slice will be absorbed, with the total absorption given by \( A_s = 1 - e^{-\alpha dz} \). Finally, we will make the relevant variable in the atmosphere the power density of the slices, which is the total power coming out of the slice divided by its thickness. This will make converting everything to continuous coordinates and integrals later way simpler. A sketch of the relevant system is shown in Fig. B.5.

Writing out the power densities going into a slice from the Sun (\( P_{in,i} \)), the ground (\( P_{Gi,i} \)), and other slices (\( P_{ij,j} \), where \( i \) is the absorbing slice and \( j \) the
emitting): 

\[ P_{in,i} = P_{in}(1 - A_{s,Vis})^{N-1-i}A_{s,Vis}/dz \]

\[ P_{G,i} = P_{G}(1 - A_{s,IR})^{i}A_{s,IR}/dz \]

\[ P_{i,j} = 0.5P_{A,j}(1 - A_{s,IR})^{i-j|1}A_{s,IR} \]

By summing over the atmosphere, we can figure out the total power going into the ground and one slice, where \( P_{A,i} \) is the power density of slice \( i \):

\[ P_{G} = P_{in}e^{-\alpha_{Vis}h_0} + \sum_{i=0}^{N-1} 0.5P_{A,i}(1 - A_{s,IR})^{i}dz \]

\[ P_{A,i} = P_{in,i} + P_{G,i} + \sum_{j=0 \neq i}^{N-1} 0.5P_{A,j}(1 - A_{s,IR})^{i-j|1}A_{s,IR} \]

These coupled equations can be straightforwardly interpreted as a matrix inversion problem. It is also possible to take the limit as the slice thickness goes to zero to obtain an integral equation, which is done in a supplementary section.

Let us first consider ground temperature vs \( A_{IR} \) and GHG concentration, plotted in Fig. B.6. The most important feature is that the temperature does not saturate with IR absorption or GHG concentration, reaching temperatures over 150\(^\circ\)C in the regions considered.

To understand why this model does not exhibit temperature saturation, let us look at the power being radiated to space as a function of layer index for two different values of \( A_{IR} \), which is illustrated in Fig. B.7. As we increase \( A_{IR} \), the layers of the atmosphere that are most responsible for emitting radiation to space increase in altitude because the upper atmosphere does a better job of blocking IR radiation. Because the emitting layer’s altitude had increased, there’s now more atmosphere the can absorb upward radiation and send it back to the ground. The overall effect is that the ground has to radiate more to compensate which corresponds to increasing ground and lower atmosphere temperatures.

### B.5 Conclusion

At this point, I am personally convinced that extreme global warming is possible. What I would really like to do, but I do not think these models are
Figure B.6: Ground temperature vs IR absorption and GHG concentration of the atmosphere.

(a) Temperature vs IR absorption.

(b) Temperature vs GHG concentration, arbitrary x-axis scaling.
Figure B.7: Power being radiated to space as a function of altitude layer index.

capable of, is estimating the sensitivity of temperature to CO$_2$ concentration. The physics I would need to implement are altitude dependent atmospheric gas composition because in the real lower atmosphere, there’s water and CO$_2$, but in the upper atmosphere there is really only CO$_2$ because the water is frozen out, and take into account the actual absorption spectrum of water and CO$_2$.

Finally, it is interesting to note that these models represent two extremes of the degree of convective mixing in the atmosphere. In the first model, convective mixing is perfect because the whole atmosphere is sitting at the same temperature, and in the second model there is no convective mixing because all the coupling is through radiation. It would be interesting do devise a model with a tunable parameter representing convection. Unfortunately it would be a hassle in my current code because the convective power flow depends on the temperature gradient, so I would need to explicitly introduce temperature into the equations and break the symmetry between the up and down power flows from each layer of air. It would also make the problem non-linear because convective and radiative transfer scale differently with temperature, which results in the numerics being more difficult.

B.6 Dediscretizing Second Climate Model

The second climate model can be made into a single integral equation by making the slices thin. First, converting the atmosphere’s self-coupling into
an integral:

\[ \sum_{j=0}^{N-1} 0.5 P_{A,j} (1 - A_{s,IR})^{i-j-1} A_{s,IR}. \]

Subbing in z-coordinates for the middle factor which is the absorption loss between the two layers and the right factor which is the single layer loss:

\[ \sum_{j=0}^{N-1} 0.5 P_{A,j} e^{-\alpha_{IR}|z_i - z_j| - dz}(1 - e^{-\alpha_{IR}dz}), \]

\[ \sum_{j=0}^{N-1} 0.5 P_{A,j} e^{-\alpha_{IR}|z_i - z_j| + \alpha_{IR}dz} (1 - e^{-\alpha_{IR}dz}). \]

Multiply the right two factors and Taylor expanding because we can always make \( \alpha dz << 1 \):

\[ \sum_{j=0}^{N-1} 0.5 P_{A,j} e^{-\alpha_{IR}|z_i - z_j|} \alpha_{IR}dz. \]

We can now convert this to an integral and ignore the dropped term because its size goes to zero:

\[ \int_{z'=0}^{h_0} 0.5 \alpha_{IR} P_A(z') e^{-\alpha_{IR}|z - z'|} dz'. \]

For the power going into the ground:

\[ \sum_{i=0}^{N-1} 0.5 P_{A,i} (1 - A_{s,IR})^i dz \]

\[ \sum_{i=0}^{N-1} 0.5 P_{A,i} e^{-\alpha_{IR}dz_i} dz \]

\[ \sum_{i=0}^{N-1} 0.5 P_{A,i} e^{-\alpha_{IR}z} dz \]

\[ \int_{z'=0}^{h_0} 0.5 P_A(z') e^{-\alpha_{IR}z'} dz'. \]
To get rid of the \( A/dz \) factors in Eqn. 8, we can Taylor expand the \( A \)'s to turn them into \( \alpha \)'s. We can now write the whole coupled monstrosity:

\[
P_A(z) = P_n \alpha_{Vis} e^{-\alpha_{Vis}(h_0-z)} \quad \text{The Sun}
\]

\[
+ \left( P_n e^{-\alpha_{Vis}h_0} + \int_{z'=0}^{h_0} 0.5 P_A(z') e^{-\alpha_{IR}z'} dz' \right) e^{-\alpha_{IR}z} \alpha_{IR} \quad \text{The Ground}
\]

\[
+ \int_{z'=0}^{h_0} 0.5 \alpha_{IR} P_A(z') e^{-\alpha_{IR}|z-z'|} dz'. \quad \text{The Atmosphere}
\]

Unfortunately, I think the ‘kink’ introduced by the absolute value in the atmospheric coupling term keeps this from having an analytic solution, otherwise it would be a classic Fredholm equation of the second type with separable, integrable kernel.

### B.7 Directional vs Isotropic Optical Absorption

I realized there is a subtlety in that the visible optical absorption coefficient, \( \alpha_{Vis} \), is for regular plane wave light, but the IR absorption coefficient, \( \alpha_{IR} \) is implicitly for isotropic light. Taking \( \alpha_{IR,\perp} \) to be the regular, plane wave absorption coefficient, the two are related by:

\[
e^{-\alpha_{IR}d} = 2 \int_0^{\pi/2} \sin \theta \cos \theta e^{-\alpha_{IR,\perp}d/\cos \theta} d\theta, \quad (B.11)
\]

where \( d \) is the thickness of the slice of atmosphere and the integral is taking into account that the light is emitted in all directions.