Low Power, Integrated, Thermoelectric Micro-coolers for Microsystems Applications

By

Andrew John Gross

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering)
in The University of Michigan
2010

Doctoral Committee:

Professor Khalil Najafi, Chair
Professor Massoud Kaviany
Professor Ctirad Uher
Professor Kensall D. Wise
Assistant Research Scientist Rebecca L Peterson

Dedication

To Kelsey,

Thank you for all your love, help and support.

Table of Contents

Dedication ii

List of Figures viii

List of Tables xxi

Chapter 1 Introduction 1

1.1 Applications for micro cooling 3

1.1.1 Analog Electronics 4

1.1.2 Infrared (IR) Detectors 5

1.1.3 Resonant MEMS and gyroscopes 6

1.1.4 Other Applications 7

1.1.5 Summary of Applications 8

1.2 Comparison of Cooling Methods 8

1.2.1 Macro-scale coolers and refrigerators 9

1.2.2 Previous work in Micro-coolers 13

1.3 Scope of Research 16

1.4 Contributions 18

1.5 Thesis organization 18

Chapter 2 Thermoelectric Materials 20

2.1 Introduction 20

2.2 Thermal and Thermoelectric Material Properties 20

2.2.1 Thermal conductivity ... 20
2.2.2 Joule heating ... 22
2.2.3 The Peltier Effect ... 22
2.2.4 The Seebeck effect ... 23
2.2.5 Thomson ... 24
2.2.6 Thermoelectric Figure of Merit ... 25
2.3 Bulk Thermoelectric Materials ... 26
2.3.1 Metals ... 26
2.3.2 Semiconductors ... 27
2.4 Thin-film Materials ... 32
2.4.1 Poly-silicon ... 32
2.4.2 Bismuth Telluride and Antimony Telluride ... 33
2.4.3 Super-lattice Thermoelectrics ... 37
2.5 Materials Produced at the University of Michigan ... 37
2.5.1 Co-evaporation Process and System ... 38
2.5.2 Material Properties and characterization ... 43
2.6 Summary ... 49
Chapter 3 Modeling Thermoelectric Coolers ... 51
3.1 The Ideal Thermocouple ... 53
3.2 Non-Ideal Thermocouple ... 65
3.3 Multistage Cooling ... 69
3.3.1 Ideal 2-Stage Cooler ... 69
3.3.2 Sequential Current Optimization (Fixed Stage-1 Current) ... 72
3.3.3 Simultaneous Current Optimization (Variable Stage-1 Current) ... 75
3.3.4 Single Current Optimization ... 77

3.3.5 Parasitic Effects ... 80

3.3.6 Ultimate Cooling Limits ... 85

3.4 Prediction and Comparison to 2-D FEM ... 86

3.4.1 1-D Model Performance Predictions ... 87

3.4.2 2-D FEM Model ... 89

3.5 Summary ... 91

Chapter 4 Fabrication Process 1 – Designs and Results ... 93

4.1 Introduction ... 93

4.2 Structure ... 93

4.2.1 Planar vs. Vertical Design ... 93

4.2.2 Satisfying Thermal Requirements (cite other planar coolers fabricated on membranes) ... 97

4.2.3 Satisfying Structural Requirements ... 101

4.3 Fabrication Process ... 103

4.3.1 Processing Challenges ... 106

4.3.2 Detailed Fabrication Process ... 114

4.4 5-Stage Coolers ... 118

4.4.1 Specifications ... 118

4.4.2 Results and Discussion ... 121

4.5 6-Stage Coolers ... 126

4.5.1 Results and analysis ... 128

4.5.2 Modified 6-Stage Cooler (Version 2) ... 130

4.5.3 Test Results ... 134

4.6 1-Stage cooler ... 136

4.7 Process and Design Shortcomings ... 142

4.7.1 Parasitic Thermal Conduction ... 142
4.7.2 Contact Resistance ... 143
4.7.3 Die Attachment and support ... 144
4.8 Summary ... 144

Chapter 5 Single Wafer Process ... 145

5.1 Introduction ... 145
5.2 Device Overview ... 147
5.3 Process description ... 151
5.3.1 Process-related challenges and design decisions ... 153
5.3.2 Full Process Flow ... 154
5.4 The 4 stage device ... 160
5.4.1 Current drive mechanism ... 161
5.4.2 Structure and process simplification ... 165
5.4.3 Device parameters ... 166
5.4.4 Performance ... 168
5.5 1, 2, 3, and 5-Stage coolers ... 173
5.5.1 Parameters ... 173
5.5.2 Simulated Performance ... 175
5.5.3 Fabrication issues ... 179
5.5.4 Measured Performance and Material Analysis. ... 180
5.5.5 Revised performance predictions ... 181
5.6 Die attachment ... 182
5.6.1 TLP bonding ... 184
5.6.2 Target device alignment and transfer ... 186
5.6.3 Proof of Concept ... 186

5.7 Packaging **188**

5.8 Summary **189**

Chapter 6 Conclusion and Future Work **190**

6.1 Conclusion **190**

6.2 Future work **193**

6.2.1 Thermoelectric Materials 193

6.2.2 Structure and Process 194

6.3 Projected Long Term Results **195**

References **196**

List of Figures

Figure 2.1: The material being measured (red) generates voltage V1=α_1(T1-T2). However this cannot be directly measured because the test leads produce voltages $V_2 = \alpha_2 (T_3 - T_1)$ and $V_3 = \alpha_2 (T_2 - T_3)$. The total measured voltage is V=V1+V2+V3. Where α_1 is the Seebeck coefficient of the material being investigated, and α_2 is the the Seebeck coefficient of the test leads. α_1 can only be determined if α_2 is known. 24

Figure 2.2: A comparison TE material figures of merit between 0 C and 1200 C. It is important to note that most materials do not have a high enough Z to exceed ZT=1 over a wide temperature range, if at all. Below 200 C tertiary compounds of $(Bi,Sb)_2Te_3$ and $Bi_2(Te,Se)_3$ provide the highest figures of merit. Reproduced from [48] (annotations edited for readability). 28

Figure 2.3: Schematic of a two-source co-evaporation system. They system is housed in a high vacuum chamber and includes two crucibles, a crystal thickness monitor for each of the crucibles, and a third monitor near the wafer surface. The temperature of the substrate can be controlled and the substrate can be rotated (graphic reproduced from Huang et al [74]). 39

Figure 2.4: The original TE deposition system. 40

Figure 2.5: TE deposition system from Kurt J. Lesker Company. The main picture shows the interior of the deposition chamber. The inset shows the exterior of the tool..... 41

Figure 2.6: Schematic of a Seebeck measurement apparatus 43

Figure 2.7: Bismuth telluride films deposited at 453 K, 533 K, and 556 K, using a flux ratio of 2.4:1 along with measured thermoelectric properties. Images were reproduced from [74]. 45

Figure 2.8: Bismuth telluride films deposited at flux ratios 1.8:1, 2.4:1, and 3.0:1 at a temperature of 533 K, along with measured thermoelectric properties of the films. Images were reproduced from [74]. 45

Figure 2.9: Antimony telluride Films deposited at 453 K, 503 K, and 524 K, using a flux ratio of 2.4:1 along with measured thermoelectric properties. Images were reproduced from [74]. 47

Figure 2.10: Antimony telluride films deposited at flux ratios 2.1:1, 3.0:1, and 4.0:1 at a temperature of 503 K, along with measured thermoelectric properties of the films. Images were reproduced from [74]. 47

Figure 2.11: Atomic ratio of Te:Bi and Te:Sb vs. the flux ratio used during deposition. Image reproduced from [74]. 48

Figure 2.12: X-ray diffraction spectra for Bi_2Te_3 and Sb_2Te_3. The dashed lines are from the spectra of powdered, single-crystal Bi_2Te_3 and Sb_2Te_3. Image reproduced from [74]. 49

Figure 3.1: Illustration of a basic thermocouple arrangement for cooling. 53

Figure 3.2: Heat flow in an ideal thermocouple, when it is used as a cooler. The heat flows are shown as an electrical equivalent thermal circuit, over a diagram of the thermocouple 55

Figure 3.3: The generalized cooling curve of an ideal thermocouple with fixed T_c. The graph is plotted from Eqn. 3.10 and shows $\Delta T/\Delta T_{max}$ plotted versus I/I_{opt}. 57

Figure 3.4: The effect of ZT on cooling. The graph is plotted from Eqn. 3.14 and shows the normalized minimum temperature of a single, ideal thermocouple plotted versus ZT.. 59

Figure 3.5: The effect of current input on the cold side temperature of a single, ideal thermocouple. The graph is plotted from Eqn. 3.11 and shows normalized temperature of a single thermocouple plotted versus normalized current input for several values of ZT.. 60

Figure 3.6: The effect of aspect ratio on the optimum current of a single, ideal thermocouple. The optimum current, normalized with respect to material properties, is plotted from Eqn. 3.15 versus aspect ratio for several values of ZT.................... 61

Figure 3.7: The effect of aspect ratio on power consumption of single, ideal thermocouple. Power dissipation, normalized with respect to specific material properties, is plotted from Eqn. 3.17 versus aspect ratio for several different values of ZT. 62

Figure 3.8: The effect of a load on the cooling of a single, ideal thermocouple. The normalized minimum temperature is plotted versus load, normalized with respect to aspect ratio, specific material properties, and T_h (see Eqn. 3.19). Several different values of ZT are plotted. 64

Figure 3.9: The effect of parasitic thermal conductance on cooling for a single, ideal thermocouple. The normalized minimum temperature is plotted versus the K_{par}/K_{TE} or R_{par}/R_{TE} (see Eqn. 3.25) where $K_{TE}=\kappa A/l$ and $R_{TE}=\rho l/A$...................................... 68

Figure 3.10: Heat flow schematic for a 2-stage cooler. 70

Figure 3.11: The effect of the second-stage current on the temperature of a 2-stage cooler, when the currents are set by sequential optimization. Normalized temperature is plotted versus I_2, normalized with respect to $I_{2,opt}$ for several different values of N_1/N_2. All curves are for ZT=1. 73

Figure 3.12: The effect of N_1/N_2 on the temperature of a 2-stage cooler, when the currents are set by sequential optimization. Normalized minimum is plotted versus N_1/N_2 at several values of ZT... 73

Figure 3.13: The effect of N_1/N_2 on $I_{2,opt}$ for a 2-stage cooler, when the currents are set by sequential optimization. The ratio of $I_{2,opt}$:$I_{1,opt}$ is plotted versus the N_1/N_2, showing that $I_{2,opt}$ remains lower than $I_{1,opt}$, even at high values of N_1/N_2. 75

Figure 3.14: The effect of the second-stage current on the temperature across each stage of a 2-stage cooler, when the currents are set by sequential optimization. The normalized temperature difference is shown across the first-stage, the second stage, and both stages plotted versus I_2. ZT=1, and N_1/N_2=2. 75

Figure 3.15: The effect of the second-stage current on the temperature of a 2-stage cooler, when the currents are set by simultaneous optimization. Normalized temperature is plotted versus I_2, normalized with respect to $I_{2,opt}$ for several different values of N_1/N_2 shown. All curves are for ZT=1. This method shows almost no gain in performance compared to the sequential optimization process used for Figure 3.11. 76

Figure 3.16: The effect of N_1/N_2 on the temperature of a 2-stage cooler, when the currents are set by simultaneous optimization. Normalized minimum is plotted

versus N_1/N_2 at several values of ZT. This method shows almost no difference compared to the sequential optimization process used for Figure 3.12. 77

Figure 3.17: The effect of current on the temperature of a 2-stage cooler, when $I_1=I_2$. Normalized temperature is plotted versus I_2, normalized with respect to $I_{2,opt}$ for several different values of N_1/N_2. All curves are for ZT=1. 78

Figure 3.18: The effect of N_1/N_2 on the temperature of a 2-stage cooler, when $I_1=I_2$. Normalized minimum is plotted versus N_1/N_2 at several values of ZT. 79

Figure 3.19: The effect of N_1/N_2 on the I_{opt} of a 2-stage cooler, when $I_1=I_2$. Optimum input current, normalized with respect to geometry and material properties, is plotted versus the N_1/N_2 for several values of ZT. 79

Figure 3.20: The effect of N_1/N_2 on the power consumption of a 2-stage cooler, when $I_1=I_2$. Power consumption at the optimum input current, normalized with respect to geometry and material properties, is plotted versus the N_1/N_2 for several values of ZT. 80

Figure 3.21: The effect of parasitic thermal conductance on cooling for a 2-stage cooler where $I_1=I_2$. The normalized minimum temperature is plotted versus the ratio between the parasitic thermal conductance per thermocouple and the thermal conductance of one thermocouple. $N_1/N_2=2$ for all the curves. 81

Figure 3.22: Heat flows in a 2-stage cooler, including an intermediate region with a finite thermal conductance. 82

Figure 3.23: The effect of inter-stage thermal conductance on minimum temperature of a 2-stage cooler. Normalized minimum temperature is plotted versus K_{in} normalized

over temperature, material properties, geometry and the number of TCs in the second stage. N_1/N_2=2 and a single drive current is used for both stages............... 83

Figure 3.24: The effect of inter-stage thermal conductance on the optimum current of a 2-stage cooler for various values of ZT. Optimum current normalized with respect to temperature, geometry, and material properties versus K_{in} normalized over temperature, geometry, material properties, and the number of TCs in the second stage are shown. N_1/N_2=2 and a single drive current is used for both stages. 84

Figure 3.25: The effect of inter-stage thermal conductance on the optimum current of a 2-stage cooler for various values of N_1/N_2. Optimum current normalized with respect temperature, geometry, and material properties, versus K_{in} normalized over temperature, geometry, material properties, and the number of TCs in the second stage are shown. ZT=1 and a single drive current is used for both stages.............. 84

Figure 3.26: The effect of the number of stages on the normalized minimum temperature for several values of ZT. The ratio N_1/N_2 is set to infinity. 86

Figure 3.27: Simulated performance of 1-stage, 3-stage, and 6-stage coolers vs. current input. The coolers use the material properties for the Michigan TE thin films, and a contact resistivity of 1.45 μΩ-m. 89

Figure 3.28: (a) The FEM model of a 6-stage cooler showing the temperature profile at the optimum current input. (b) The same model, enlarged to make the TCs more visible.. 90

Figure 4.1: Pyramid to planar design transitions. (a) A traditonal pyramid structure. (b) A planar cooler with the stages arranged as concentric rings..................................... 96

Figure 4.2: A 3-D rendering of a planar multi-stage thermoelectric cooler. The coldest region is in the center of the device, with 6-stages arranged in concentric squares around the center. A glass tether supports the cold region to give some additional strength to the system.. 97

Figure 4.3: Cross sectional diagram of the thermal substrate for a 2-stage cooler built using the silicon-glass-silicon process. 97

Figure 4.4: 2-D axis symmetric FEM simulation of a 5-stage cooler structure used to calculate the thermal isolation of the Thermal Substrate. 101

Figure 4.5: Cross-section of the Structural Substrate used in the Silicon-Glass-Silicon process.. 101

Figure 4.6: 3-D Rendering FEM simulation of a glass tether, such as the one used in the silicon-glass-silicon process. The temperature difference was set at 10 °C 102

Figure 4.7: Illustration of a shadow mask cross sections (a) with a recess around the pattern, and (b) without a recess around the pattern. Using the recess allows flux impinging the wafer at shallower angles to reach the device wafer. This is important in systems where the source is not located along the axis perpendicular to the center of the wafer. 112

Figure 4.8: Images of TE material patterned by shadow masking. (a) The full test pattern. (b) The smallest element in the test pattern. It is 23 μm x 130 μm. (c) The largest element in the test pattern. It is 77 μm x 117 μm..................................... 113

Figure 4.9 (a) View of the top side of a shadow mask showing the recess and the pattern that will be deposited on the cooler. (b) Pattern of the backside of the wafer showing the pattern that will be deposited. Taken from [82]. 113

Figure 4.10: (a) A top view of a 5-stage cooler built using the Silicon-Glass-Silicon process. (b) The inner three stages of the same 5-stage cooler, showing the temperature sensor. 119

Figure 4.11: Pictures of (a) the T1 tether, (b) the T2 tether, and (c) the T3 tether designs. The pictures have been artificially colored to increase the contrast to the background and to illustrate different regions of the tether designs. Red areas indicate the regions that will be bonded to the Thermal Substrate. Blue regions are unbonded. 120

Figure 4.12: An SEM image of the T2 tether design used with a 5-Stage cooler. This shows the tether suspended over of 10 μm recess in a silicon wafer...................... 120

Figure 4.13: Top views of three metal designs, showing (a) M1, (b) M2, and (c) M3. The metallization used in M1 is also used in the M4 pattern (not shown) but the oxide bridges are longer in the M4 design... 121

Figure 4.14: A schematic representation of the thermal isolation measurement circuit. The four-point measurement system is used to avoid measuring changes in resistance due to parasitic elements that are not located at the center of the device. 123

Figure 4.15: Thermal resistance measurement of the 5-stage cooler in air and in vacuum. The thermal resistance in vacuum is 27 times higher than the thermal resistance in air. 123

Figure 4.16: A wafer map showing thermal resistance measurements across several variations of cooler designs. Die in green were testable. The top number indicates

thermal resistance in K/W, and V-x indicates a variation designator as explained in the text........ 124

Figure 4.17: Measured performance of a full 5-stage cooler plotted versus input current. Power consumption is plotted using the right hand axis, and modeled cooling performance is plotted using nominal........ 126

Figure 4.18: Top view of the 6-Stage cooler, Version 1. It shows the thermoelectric elements, the temperature sensor, and the silicon rings........ 127

Figure 4.19: An image of the tether used for Version 2 of the 6-stage cooler. The picture has been artificially colored to increase the contrast to the background. 128

Figure 4.20: Measured performance of Version 1 of a 6-stage cooler, plotted together with simulated performance that has been fit to measured data to extract the material properties of the TE. 130

Figure 4.21: Simulated TE-metal contact resistance for a variety of contact geometries. The assumed TE resistivity for these simulations is 20 μΩ-m, and the contact resistivity is 1E-8 $\Omega\text{-m}^2$. The current enters the contact in the TE layer and leaves in the metal layer........ 131

Figure 4.22: Comparison of contacts in Version 1 (right side) and Version 2 (left side). The images show contact dimensions and the current path through 3 stages of each cooler........ 132

Figure 4.23: Simulated resistance of a TE element deposited over gold. The TE resistivity is assumed to be 20 μΩ-m, and the contact resistivity is 1E-8 $\Omega\text{-m}^2$. The current enters and leaves the element through the TE material........ 133

Figure 4.24: Top view of the 6-stage cooler, Version 2. The inset highlights the thermoelectric elements at the inner 4 stages, and illustrates how a single element can span multiple stages 134

Figure 4.25: : Measured performance of several Version 2, 6-stage coolers plotted versus input current. All demonstrate cooling of greater than 19 K. 136

Figure 4.26: Measured performance of Version 2 of a 6-stage cooler, plotted together with simulated performance that has been fit to measured data to extract the material properties of the TE. 136

Figure 4.27: Top view of a single stage cooler with 16 TCs. 137

Figure 4.28: Measured performance of a 1-stage cooler, plotted together with simulated performance that has been fit to measured data to extract the material properties of the TE. 138

Figure 4.29: A 1-stage cooler tested at various thermal loads. The temperature change achieved and the optimum current both increase as the input load increases......... 140

Figure 4.30: Variation in ΔT and I_{opt} with respect to cold stage temperature for a 1-stage cooler. The ambient temperature is maintained at 300 K, and the temperature of the center region is increased by applying a thermal load with a resistive element..... 141

Figure 4.31: A Version 2 6-stage cooler tested at various thermal loads. The temperature change achieved and the optimum current both increase as the input load increases. 141

Figure 4.32: Variation in ΔT and I_{opt} with respect to cold stage temperature for a 6-stage cooler. The ambient temperature is maintained at 300 K, and the temperature of the center region is increased by applying a thermal load with a resistive element..... 142

Figure 5.1: Cross section of a 2-stage cooler implemented using the structure. The thermal properties of the cooler are implemented with an oxide-metal-oxide stack that is supported by oxide pillars. 147

Figure 5.2: Illustration of maximum dimension for inter-stage thermal conduction...... 149

Figure 5.3: Illustration of heat and current flow in (a) a conventional, serially-driven 2-stage cooler, and (b) a two stage cooler using a resistive network to distribute current. The electrical schematic of the device shown in (b) is provided as (c).... 162

Figure 5.4: (a) Illustration of a previously devised three stage cooler that distributes current through a resistive network. (b) A 3-stage cooler using the resistive network proposed in this section. (c) A modified version of the cooler shown in (b). It combines inter-stage regions of the same nominal voltage................................ 164

Figure 5.5: Cross section of a 2-stage cooler using the simplified vartion of the single wafer process. 165

Figure 5.6: A top-down SEM view of the 4-stage cooler with resistive network current distribution. The individual stages and the cold platform are highlighted. 167

Figure 5.7: An SEM showing the oxide pillars supporting the cold platform and stages 1 and 2 of the 4-stage cooler. This image also shows how the oxide has been removed, except directly under the TE materials. 167

Figure 5.8: An SEM image of the inter-stage zone between stages 2 and 3 of a 4-stage cooler. The zone is not supported by any oxide pillars, and curling of the inter-stage zone is clearly visible. This is in contrast to Figure 5.7, which shows very little deflection of the interstage zones where they are supported by the pillars. 168

Figure 5.9: Schematic representation of the 4 stage cooler. 170

Figure 5.10: FEM simulation of a cooler with resistive network current distribution, using the material properties reported in Chapter 2 for co-evaporated materials deposited at the University of Michigan. 170

Figure 5.11: Thermal resistance measurement of a 4-stage cooler fabricated using the simplified single wafer process. The thermal resistance is 42,539 K/W. 171

Figure 5.12: FEM simulation of a cooler with resistive network current distribution, using the material measured in section 5.4.4.2. 172

Figure 5.13: Top down SEM images of 5-stage (annotated), 1-stage (left), 2-stage (center), and 3-stage coolers (right), pictured released but without TE. Only a quarter of each cooler is shown because the devices are too large to visualize completely in the SEM 174

Figure 5.14: Simulated performance of 1, 2, 3, and 5-stage coolers with no parasitic thermal effects considered. 176

Figure 5.15: Simulated performance of 1, 2, 3, and 5-stage coolers with parasitic conduction due to supporting oxide membranes included in the model 177

Figure 5.16: Simulated performance of 1, 2, 3, and 5-stage coolers with parasitic conduction due to supporting oxide membranes and oxide pillars included in the model 179

Figure 5.17: An SEM image of an antimony telluride leg of a TC. The crack in the antimony telluride is clearly visible. 180

Figure 5.18: Performance predictions based on the material properties extrapolated from the measured performance of a 3-stage cooler. 182

Figure 5.19: Cross section illustration of the TLP bonding process showing the low melting temperature metal as purple and the high melting temperature metal in yellow. The illustration includes (a) The bond pads prior to bonding, (b) the arrangement at the initiation of the bond, and (c) the diffused intermetallic compound after completing the bonding process. 185

Figure 5.20: An SEM image of a die bonded to an unreleased cooler structure. 187

Figure 5.21: SEM images as different magnification levels of the bond region after the attached die has been pried away. Silicon has fractured, indicating a strong bond. 187

List of Tables

Table 1.1: Summary of applications benefitting from cooling. 4

Table 1.2: Summary of cooling techniques 9

Table 1.3 Comparison of thermoelectric microcoolers including structure, materials, temperature differentials and power consumption.. 17

Table 2.1: Thermoelectric Properties of Common Metals at ~300K[a] 27

Table 2.2: Comparison of Co-evaporated Bi_2Te_3 and Sb_2Te_3 Thin-films 36

Table 3.1: Material Properties and Dimensions used for simulation shown in Figure 3.27. 89

Table 4.1: Thermal Substrate fabrication process overview... 104

Table 4.2: Structural Substrate fabrication process overview. 105

Table 4.3: Cooler completion process overview.. 106

Table 4.4: Critical dimensions of the 5-Stage thermoelectric cooler............................ 119

Table 4.5: Critical dimensions of the Version-1 6-Stage thermoelectric cooler............ 127

Table 4.6: Critical dimensions of the Version-2 6-Stage thermoelectric cooler............ 133

Table 5.1: Design parameters of the 4-stage cooler fabricated using resistive network current distribution.. 166

Table 5.2: Design parameters for the 1, 2, 3 and 5-stage coolers fabricated using the single wafer process... 174

Chapter 1
Introduction

Advancements in microcooling technology are being driven by applications across a number of industries. The medical industry is interested in microcoolers to aid in cryosurgical techniques. Microcoolers that can be integrated with thermal imaging sensor would be useful for defense related applications. The communications industry would benefit from the decreased noise of cooled frequency references and amplifiers, but only if the solutions are small enough and low enough in power to integrated into mobile products. Digital electronics could operate with higher speed and stability if they could be paired with integrated cooling solutions, and some MEMS devices, such as resonant gyroscopes also demonstrate performance increases as they are cooled.

Some of these application areas are being addressed. Micro Joule-Thomson coolers have been developed with cryosurgery in mind, as well as for cooling microelectronics and MEMS. Thermoelectric cooling has been used to create integrated solutions for electronics with high power dissipations. However, there are applications where these solutions are not feasible. The Joule-Thomson coolers are too complex for many applications. Thermoelectric coolers offer compact solid-state operation, but the coolers developed to date have been designed for applications with high heat loads, where the total power consumption of the system is of little interest. Devices such as LNAs and gyroscopes, on the other hand are often used in mobile applications where the total power consumption of the system is very important. Applications like these, that

require a large temperature differential and low power consumption, have not yet been addressed, and they are the motivation behind the research in this thesis.

Designing a thermoelectric microcooler for low-power, mobile applications presents many challenges. First among them, is how to design for low power operation. Although several groups have previously developed thermoelectric micro-coolers with temperature differentials greater than 40 K, they require high power consumption. Reducing the power consumption of the thermoelectric cooler without reducing the temperature differential that can be achieved is fundamental to achieving the goal of a low-power integratable device.

The second challenge is integration of the thermoelectric materials with a micro-fabricated supporting structure. The deposition conditions of the thermoelectric material must be optimized, and the deposition process must to be integrated with the larger process flow used to make the device. These materials are incompatible with many clean room processes and have a low thermal budget, making the integration a significant challenge. The final set of challenges center on creating an entire system with the cooler. The cooler is of no use if it cannot be applied to the applications it is intended to address. It must be integrated with a MEMS or electronic device that requires cooling, and the entire system, comprising the cooler and the target device, must be appropriately packaged. The goal of this thesis is to present a thermoelectric microcooler that addresses the design challenges discussed above. The cooler will be capable of generating large temperature differentials with a power-consumption of less than 100 mW and will be integratable with arbitrary MEMS devices.

This chapter will present the motivation behind developing a micro thermoelectric cooler. It will begin with a discussion of the applications for micro cooling, and the cooling levels need for the targeted applications. It will then proceed to an overview of cooling technologies, including both macro and micro scale systems, and a discussion of why thermoelectric cooling is a good choice for many applications at the micro scale. It will provide a comparison of various thermoelectric micro coolers that have been developed to date, and discuss the contributions that this work makes to the state of the art. Finally this chapter will provide an outline of the remainder of this thesis.

1.1 Applications for micro cooling

A number of electronic and MEMS devices can exhibit improved performance characteristics when operated below room temperature. The desire to take advantage of these performance enhancements is one of the main driving forces for the development of micro scale integrated cooling systems. Different classes of devices, however, express such temperature-dependent performance increases in different ways. Similarly, different types of devices exhibit improved performance changes in different temperature ranges, and require different levels of cooling power. Table 1.1 summarizes several application areas that could potentially benefit from micro cooling. and the remainder of this section elaborates on those application areas.

Table 1.1: Summary of applications benefitting from cooling.

Application	Metric of Improved Performance	Temperature Range of Improved Performance
Low Noise Amplifiers	Reduced Johnson noise	19 K – 200 K
Infrared Detectors	Reduced Johnson and temperature noise	70 K – 250 K
Resonant MEMS	Increased quality factor	<250 K
Chromatography	Improved analyte adsorption; transit speed manipulation	<300 K
Temperature cycled applications.	Cooling speed	NA
Solid State Laser	Required for Continuous wave IR emmision	<264 K

1.1.1 Analog Electronics

One broad class of devices that can benefit from operation at decreased temperatures are analog integrated circuits; these circuits generally see a decrease in noise figure resulting from lower Johnson noise in transistors and resistors at lower temperature. Minimizing noise figure is particularly important in the first stage of cascaded amplifiers, because the noise introduced at this stage receives the most amplification, and ultimately dominates the noise output of the entire system amplifier. To address this, special low noise amplifiers (LNAs) have been developed, and are often employed as the first stage of amplification in the front end of radio receivers. Because LNAs dominate the noise performance of the overall system, further reduction in the noise figure of these devices is important. Several groups have demonstrated that reducing the operating temperature of the devices does, in fact, yield a significant performance increase.

In one case, an LNA designed for room temperature operation from high electron mobility GaAs transistors nonetheless demonstrated a decrease in noise temperature from

approximately 200 K when operate at 297 K to 20 K when operate at 19 K, demonstrating that improvement can be achieved even in devices that have been designed for room temperature operation [1]. In another case, a shift in noise figure from near 1 dB at 300K to 0.55 dB at 70 K was observed [2]. In both of these cases the low-temperature performance was investigated at truly cryogenic temperatures. However, for some devices, significant decreases in noise can be observed at more modest temperatures. In one instance the noise temperature was observed to change from 46 K at 293 K to 28 K at 200 K [3]. Although still significantly below room temperature, 200 K is within the temperature range that can be reached by small, commercially-available thermoelectric and Joule-Thomson coolers. An additional study addressed the use of thermoelectric cooling more directly. It showed a decrease in noise temperature from over 100 K to less than 80 K over a change in temperature from to 30 C to -15 C, using a commercially available thermoelectric cooler using less than 600 mW of input power [4].

1.1.2 Infrared (IR) Detectors

Infrared detectors constitute another large class of devices that can benefit from the application of micro cooling technology. There are two main types of IR detectors. One class of infrared sensors is based on the operation of photodiodes and photoconductors, in which electron-photon interactions increase the number of available carriers. However, unlike a standard photodiode, IR photon detectors are fabricated from semiconductors such as HgCdTe that have high absorption coefficients for light in the infrared region. For these detectors to be effective, the ratio of photon-generated carriers to thermally generated carriers should be greater than 1. For long and very long wavelength IR, it is necessary to cool the detectors to cryogenic temperatures in the range

of 70K or less in order to achieve this ratio. Short wave length IR detection, on the other hand, requires much less cooling, and detectors have been demonstrated that only require commercially available thermoelectric cooling units[5].

The second class of sensors are heated by impinging infrared light, and the temperature change is proportional to the intensity of the incident light. These devices consist of an IR absorber that is attached to a substrate through low thermal conductivity legs, and the temperature of the absorber can be measured either with a thermistor or thermoelectrically. In these devices, detectivity is primarily limited by thermal noise, which manifests itself both as Johnson noise in the electric components and as random temperature fluctuations of the absorber itself. The Johnson noise is proportional to T, while the temperature fluctuations are proportional to T^2[6]. Decreasing the temperature of such a detector from 300 K to 250 K would therefore lead to a reduction in Johnson noise of 16.6%, and a decrease in thermal fluctuations of 30%.

1.1.3 Resonant MEMS and gyroscopes

A third group of devices that can benefit from operation at decreased temperature are resonant MEMS. These devices are finding increasing applications as filters and potentially as a replacement for quartz resonators in clock and RF frequency generators. Additionally, resonant devices have been used to fabricate MEMS gyroscopes for several years now. It has been demonstrated that these devices experience an increase in the Quality Factor (Q) of their resonance at decreased temperatures. The change in Q can be quite significant, and it has been demonstrated to be a more sensitive measure of temperature than silicon RTDs for at least one resonator design [7]. However, the temperature range at which the change in Q occurs is highly dependent on the design of

the resonator. In the case of the resonant temperature sensor discussed above, which resonated at 1.3 MHz, the Q changed from less than 7 000 at 70 C to more than 12 000 at 0 C. However, in another case, resonators did not demonstrate a significant increase in Q unless they were cooled to below 110 K, at which point they showed an increase in Q from less than 100 000 to 362 768 at 5 K [8].

This variability in the temperature dependence of Q between different resonant devices means that care must be taken when pairing such devices with a cooling solution to ensure that low enough temperatures can be achieved. Gyroscopes, in particular, have demonstrated marked improvements in Q at relatively modest decreases in temperature. For example in one case, the sense mode Q was observed to change from 33,900 at 50 C to 51,700 at 5 C, while the drive mode Q changed from 54,400 to 72,500 [9]. In another design, the Q changed from approximately 10,000 at 50 C to more than 25,000 at 5 C, and this corresponded to a change in sensitivity of 57% [10].

1.1.4 Other Applications

One additional application that can benefit from micro cooling is IC lasers. They have exhibited improved performance when cooled thermoelectrically to temperature of 264 K [11]. Finally micro scale cooling has applications for a new generation of micro scale gas chromatographs that are under development. These devices can benefit from increased adsorption of analytes at various stages of the chromatography process. Increased adsorption in preconcentration beds leads to higher analyte collection, and 2D chromatography uses thermal modulation in the region between the two columns to separate peaks that would otherwise coelute [12].

1.1.5 Summary of Applications

It is clear from the descriptions above that a wide variety of micro scale devices can benefit from an integrated micro-scale cooling solution. In particular, a number of applications, including gyroscopes, thermal IR detectors, IC lasers, and certain LNAs have exhibited a significant increase in performance at temperatures above -50 C, potentially within the range of several types of cooling solutions. It is also important to note that several of these applications, particularly gyroscopes and passive IR sensors, dissipate very little power and are sometimes deployed in applications that require mobility and an overall low system power. This is an important consideration when evaluating potential cooling solutions; they should therefore be simple, robust and low power. Various types of cooling, their effectiveness, and their suitability to miniaturization and integration will be described in the following section.

1.2 Comparison of Cooling Methods

There are a wide variety of refrigeration and cryogenic cooling solutions at the macro scale. These can be based on several different mechanisms, and demonstrate different trade-offs between achievable minimum temperature, efficiency, overall power consumption, and simplicity. Some of these technologies have properties that allow them to be effectively miniaturized, while others remain primarily macro-scale. Table 1.2 summarizes the major cooling and refrigeration technologies used at the macro scale, and these are elaborated on in the remainder of this section.

Table 1.2: Summary of cooling techniques

Type of Cooler	Minimum temperature	Applications
Vapor Compression	~250 K	Consumer and commercial refrigeration
Absorption	~250 K	Consumer and commercial refrigeration
Stirling (including pulse-tube coolers)	1.78 K	Cooling for cryogenic processes; gas liquefaction possible
Joule Thomson	78 K (for nitrogen) <20 K (for helium)	Cooling for cryogenic process, gas liquefaction
Laser Cooling of Gasses	<1 mK	Specialized research
Laser Cooling of Solid	212 K	Mechanically and electrically isolated cooling
Demagnetization cooling	<1 K	Specialized research
Thermoelectric cooling	~ 230 K – single stage ~170 K – 6 stages	Consumer refrigeration, IR sensors, solid state lasers, other sensors and electronics

1.2.1 Macro-scale coolers and refrigerators

1.2.1.1 Fluid based systems

One broad class of macro-scale coolers is known as vapor-compression systems. The vapor-compression cycle relies on cooling through the evaporation of a working fluid. After the evaporation, the fluid undergoes adiabatic compression back to a liquid phase, during which it experiences an increase in temperature. The fluid then exchanges heat with the environment and undergoes cooling, after which it is fed back to the evaporator. This type of cooling cycle is widely used in commercial and consumer applications. Refrigerants include CO_2, ammonia, and engineered refrigerants such as R22. The minimum achievable temperature is related to the boiling point of the working fluid in the system, and is usually higher than -50 C [13],[14].

Absorption coolers are related to vapor compression coolers, but utilize heat energy instead of mechanical energy to power the cooling cycle. These systems remove heat from the cold region through evaporation of a liquid working fluid. After evaporation, the gas is absorbed into another fluid, which is then transported away from the cold region, and heated to desorb the refrigerant. The desorbed, high temperature gaseous refrigerant exchanges heat with the ambient and condenses back to a liquid state. Finally, the liquid refrigerant is fed back to the evaporator and the cooling process continues. The refrigerant is often ammonia, and the temperature differentials that can be achieved are similar to those of vapor compression systems[14].

A second class of fluid-based macro-scale coolers are based on the Stirling cycle. In this thermodynamic cycle, a working fluid is expanded adiabatically, after which it absorbs heat from the area being cooled. The fluid is then forced through a thermal regenerator where it is warmed and expands further. The fluid undergoes adiabatic compression and its temperature increases. It exhausts heat to the environment, and is moved back through the regenerator, where it is cooled and decreases in volume. Stirling cycle coolers are used in a wide variety of applications ranging from commercial refrigeration systems to cryogenic cooling systems. In traditional systems the gas flow, compression, and expansion are controlled by a system of multiple pistons, or pistons and valves in the case of Gifford-MacMahon coolers. A more recent variation of the Stirling cycle, the pulse-tube cooler, utilizes a single piston, combined with a reservoir connected to the main system through an orifice, and has achieved temperatures as low as 1.78 K [15].

A third type of fluid-based cooler utilizes the Joule-Thomson cycle. In this type of cooler, a gas under high pressure is forced through a small orifice where it undergoes isenthalpic expansion. Given an appropriate initial temperature, this results in cooling of the gas. The cooled low-pressure gas then travels away from the orifice, where it exchanges heat with the high pressure gas through a counter flow heat exchanger. Due to the use of counter-flow heat exchangers, the temperature of the gas after expansion will decrease over time until liquefaction begins to occur. These systems have been used to create liquid air, N2, O2 and He. The boiling point of the liquid will define the minimum achievable temperature, and the heat of vaporization, combined with the rate of liquid generation, will define the maximum heat load that can be cooled [16].

1.2.1.2 Solid State Systems

For systems where a compressed gas source is not available, there are solid-state solutions that can be applied. The most widely used are thermoelectric coolers. These devices make use of the Peltier effect, which will be discussed at length in the following chapters. At room temperature they typically use $(Bi,Sb)_2Te_3$ based materials, and have been used have been applied for a wide variety of applications. Consumer products include refrigerators and wine coolers capable of maintaining temperatures as low as 0 °C [17-19]. Although they cannot achieve the same minimum temperature that a conventional refrigerator can, these devices are appealing in applications where small size and quiet operation are important. They can also be useful when standard AC power is not available but DC power is. Thermoelectric modules are also available for a wide range of performance parameters for a variety of applications. Single stage devices are capable of reaching temperatures as low as 232 K, and are marketed for use with

bolometers and laser diodes. Two stage devices can reach temperatures of 203 K and are marketed toward use with CCD arrays and IR detectors. Four-stage devices can reach temperatures as low as 173 K, and 6 stage devices can achieve 167 K and are also targeted at CCD arrays and IR detectors. Integrated thermoelectric systems are also being applied to IC cooling [20].

A second solid-state technology that has been applied to certain specialized problems is laser cooling. Laser cooling of gasses has been available for several years, and works by shining properly tuned lasers into a magnetically confined low-density gas from multiple directions. Over time this arrangement gradually reduces the average kinetic energy of the atoms in the gas, thus cooling the system. This technique is used to cool gases to below 1 mK [21]. A potentially more useful technology involves laser cooling of crystalline solids. For this technique a laser is focused on special material such as ZBLANP (ZrF_4–BaF_2–LaF_3–AlF_3–NaF–PbF_2), and tuned so that electron transitions require the photon to interact with both an electron and a phonon. When the electron transitions back to its ground state, it emits all its excess energy as a photon. As a result there is a net loss in the energy of the system, equivalent to the energy of the phonon. The loss of energy causes the system cool, and has produced temperature differentials of 92 K at 300 K ambient temperature[22].

A final technique that has been used for cooling at the macro scale is demagnetization cooling. If a magnetic field is applied to certain materials under adiabatic conditions, they will increase in temperature due to the decrease in magnetic entropy. The material can then exchange heat with a thermal sink and cool back to its initial temperature. Next the sample is returned to adiabatic conditions and is

demagnetized. The increase in magnetic entropy results in a decrease in temperature. This technique can be used in refrigeration systems [23], but has also been used in cryogenic systems to achieve temperatures well below 1 K [24].

1.2.2 Previous work in Micro-coolers

Although there are a wide variety of cooling options at the macro scale, only a few technologies have been investigated at the micro scale. These include a significant investigation of both Joule-Thomson coolers and thermoelectric coolers, as well limited theoretical work with micro Stirling coolers. The remainder of this section will detail the state of the art of these three technologies, and compare their benefits and weaknesses.

1.2.2.1 Joule -Thomson coolers

Joule-Thomson coolers that utilize micro scale fabrication have been built by multiple groups, and with several design variations[25],[16],[26-31]. Single stage J-T coolers were fabricated at the University of Twente that reached a minimum temperature of 96 K and had a heat cooling power of at least 5 mW[25]. J-T coolers have also been fabricated at the University of Michigan [16] and the University of Colorado [31]. While effective at generating temperature differentials, such coolers have a number of drawbacks that make them an impractical solution for many integrated cooling applications. Although these devices use micro fabricated components, the total system size can still be relatively large. In all of the examples mentioned above, the counter-flow heat exchanger exceeded 1 cm in at least one dimension. In addition all J-T systems require a source of high-pressure gas. This can be supplied in an open-loop system with compressed bottled gas, or with a compressor to form a closed-loop system. In either case, the need for high pressure leads to an increase in the overall size complexity and

cost of the system. It also makes J-T systems impractical for applications where small size, simplicity and integration are more important than efficiency and the absolute minimum temperature.

1.2.2.2 Stirling

Limited work has been done on the development of micro Stirling coolers [32],[33]. These coolers are intriguing candidates for miniaturization because they require less compression of the gas than J-T systems, and implementation of the regenerator could potentially be less complex than the counter flow heat exchanger required for a J-T system. However, developing a micro-stirling cooler would still require fabrication of a complex, dynamic, micro mechanical system. Additionally the system must be sealed with the working fluid inside and be resistant to leaks. Although potentially smaller and more self contained than J-T coolers, this high level of complexity is still undesirable for many applications.

1.2.2.3 State of the art thermoelectric micro cooling

Unlike J-T and Stirling coolers, thermo-electric coolers offer an entirely solid state solution to cooling at the micro scale, and several groups have worked on developing small scale integrated thermoelectric coolers {insert citations}. Earlier work at the University of Michigan implemented a single stage design that used Bi_2Te_3 as the n-type material and Sb_2Te_3 as the p-type, material and achieved maximum cooling of 1.3 K while consuming 41 mW [34]. Several other investigators have also made use of a single stage vertical architecture, while using a variety of materials. Huang made use of polysilicon as the thermoelectric material and achieved a maximum temperature differential of 5.6 K, while consuming 844 mW of power [35]. Fan used a SiGe/Si

supperlattice as the p-type material and metal as the n-type material to create a cooler capable of cooling by 12 K at an ambient temperature of 200 C [36]. More successfully, Bottner has produced coolers utilizing Bi2Te3 as the n-type material and Sb2Te3 as the p-type material, which have demonstrated cooling of 48 K at an ambient temperature of 85 C, at an input current of 2.2 A [37]. Bulman has made use of more complex superlattice materials consisting of layers of quaternary alloys of $(Bi_x,Sb_{x-1})_2(Te_y,Se_{y-1})_3$. This research yielded a single stage cooler capable of generating 55 K temperature differentials at a total power draw of 362 mW at in input current of 3.9 A [38]. The same group has also produced a 3-stage vertical cooler capable of cooling by 102 K [39]. The performance of these coolers is compared in Table 1.3. It should be noted that the most effective vertical thin film thermoelectric coolers require a large amount of input power. This is largely due to the thinness of the films and will be discussed in greater depth in Chapter 3.

A different approach is to move the heat laterally across the die. This approach allows for greater control of the geometry of the thermocouple (TC) when using thin films, and can be used create coolers with lower power requirements. This approach was investigated by Yao, who achieved cooling of only 2 K utilizing a Si/SiGe superlattice as the thermoelectric material [40]. Goncalves has also investigated planar thermoelectric coolers and has achieved 5 K of cooling utilizing Bi2Te3 and Sb2Te3 on a polyimide substrate. This cooler achieved optimal operation at an input current of 4 mA with a total power consumption of 0.48 mW [41],[42].

1.3 Scope of Research

The opening of this chapter gave the major challenges associated with developing a low-power thermoelectric microcooler as:

1) Designing for low-power operation
2) Integration of the thermoelectric materials into the fabrication process
3) System level integration, including attachment of a device and appropriate packaging

While this thesis will seek to address all of the challenges above, its scope is nonetheless limited. The requirements of designing coolers for low power will be covered in depth, as will integration of the thermoelectric materials into the fabrication process. In fact, the materials have been integrated with two distinct fabrication processes and the challenges associated with each will be presented. Die attachment will be demonstrated, and it will be shown that the coolers are compatible with vacuum packaging techniques, both thermally and mechanically. The devices present will be tested under vacuum, however vacuum packaging itself will not be demonstrated.

Table 1.3 Comparison of thermoelectric microcoolers including structure, materials, temperature differentials and power consumption.

Reference	Bulman et al	Nextreme	Bottner et al	Micropelt	Bulman et al
Structure	1-Stage Vertical	1-Stage Vertical	1-Stage Vertical	1-Stage Vertical	3-stage Vertical
Materials	*n*-Bi_2Te_3/ $Bi_2Te_{2.83}Se_{0.17}$ *p*-Bi_2Te_3/Sb_2Te_3	*$(Bi,Sb)_2(Te,Se)_3$ based super lattice*	*n*-$Bi_2(Se,Te_3)$ *p*-$(Bi,Sb)_2Te_3$	*n*-$Bi_2(Se,Te)_3$ *p*-$(Bi,Sb)_2Te_3$	*n*-$Bi_2Te_3/Bi_2Te_{2.83}Se_{0.17}$ *p*-Bi_2Te_3/Sb_2Te_3
ZT	n-type: 1.4 p-type: 2.4	Unknown	n-type: 0.53 p-type: 0.71	Unknown	n-type: 1.4 p-type: 2.4
Cooling @ I_{opt}; T_h (1,2)	55 K @ 3.9 A	40 K @ 3.8 A; 25 °C 47 K @ 4.8 A; 85 °C	48 K @ 2.2 A; 85 °C	44.4 K @ 1.2 A; 23 °C 61.2 K @ 1.4 A; 85 °C	102 K @ current varies by stage
Estimated Power consumption	362 mW	760 mW @ 25 °C	Unknown	450 mW @ 23 °C	28.11 W
Estimated Extrinsic ZT	0.54	.31	.31	.35	NA

Reference	Snyder et al	Fan et al	da Silva et al	Goncalves et al	Huang et al
Structure	1- Stage Vertical	1-Stage Vertical	1-Stage Vertical	1-Stage Horizontal	1-Stage Vertical
Materials	*n*-Bi_2Te_3 *p*-Sb_2Te_3 (electroplated)	p-Si/SiGe superlattice	*n*-Bi_2Te_3 *p*-Sb_2Te_3	*n*-Bi_2Te_3 *p*-Sb_2Te_3	*n*-poly-Si *p*-poly-Si
ZT	unkown	unknown	n-type: 0.053 p-type: 0.027	n-type: 0.5 p-type: 0.84	n-type: 0.024 p-type: 0.016
Cooling @ I_{opt}; T_h (1,2)	2 K @ 100 mA; 82 °C	12 K @ 600 mA; 200 °C	1.3 K @ 23 mA; 62 °C	5 K @ 4 mA	5.6 K @ 80 mA
Estimated Power consumption	396 mW	Unknown	41 mW	.33 mW	844 mW
Estimated Extrinsic ZT	0.011	0.052	0.0078	0.034	0.038

[1] I_{opt} = Input current at maximum temperature differential
[2] T_h = Ambient temperature

1.4 Contributions

In solving the challenges associated with developing a thermoelectric microcooler this research makes the following contributions:

1) Analysis of the requirements for low-power thermoelectric cooling and application of those requirements to multiple processes and cooler designs.
2) Demonstration of the first planar, multistage, thermoelectric microcooler.
3) Integration of thin-film thermoelectric materials with a planar micro-fabrication process.
4) Development of a low power microcooler device that can be integrated with arbitrary MEMS and electronic devices

1.5 Thesis organization

The following chapters of this thesis will present the design, fabrication and performance of several thermoelectric microcoolers. Chapter 2 will discuss the thermoelectric materials used for the project. It will provide an overview of the thermoelectric effects and identify the properties necessary for a high quality thermoelectric material. It will compare several of the best materials used at both the macro and micro scales. It will present the co-evaporation deposition technique used to deposit the Bi_2Te_3 and Sb_2Te_3 thin films for this project. These films have demonstrated thermoelectric figures of merit, ZT, of 0.40 and 0.34 respectively. The test technique for the samples and the optimization of the films will also be discussed.

Chapter 3 will present the models used to design and optimize the coolers presented later in Chapter 4 and Chapter 5. It will provide an overview of the basic

cooler operation, and provide an analysis of the key general parameters that a cooler must possess for low power operation with a high temperature differential. Chapter 4 will present a fabrication process based on a multi-wafer Silicon-Glass-Silicon stack. A 5-stage cooler capable of $\Delta T=8$ K, a 1-stage capable of $\Delta T=17.9$ K and a 6-stage cooler capable of $\Delta T=22.3$ K and their performance will be analyzed based on the models developed in Chapter 3.

Chapter 5 will present a second fabrication process, which is based on a single wafer and a dry XeF_2 release. It has been designed to reduce fabrication complexity and allow for integration of MEMS devices with the cooler. 1-stage, 2-stage, 3-stage and 5-stage coolers are being fabricated with this process and have the potential to produce cooling of between $\Delta T=30$ K and $\Delta T=40$ K. A 4-stage cooler capable of $\Delta T=17.6$ K using unique current distribution mechanism will also be presented. This process allows for a successful die attachment to the cooler, and this aspect of the process will also be presented. All of these coolers will be analyzed based on the models that have been developed. Chapter 6 will look at future work and improvements to the technology that could be pursued, and provide a conclusion and summary of this thesis.

Chapter 2
Thermoelectric Materials

2.1 Introduction

This chapter will discuss the choice to use bismuth telluride and antimony telluride as the thermoelectric materials for this project. It will discuss the deposition technique used to create these materials in thin film form at the University of Michigan, and it will summarize the material properties that have been achieved so far. In order to clearly understand the reasons for this choice, the chapter will first present an overview of thermoelectric material properties, and the relationships those properties have to the quality of thermoelectric materials. It will then summarize the properties of several prominent bulk and thin-film thermoelectric materials. Finally the chapter will present the method of thin-film growth employed in this project, and discuss the material properties of the thermoelectric films that have been obtained.

2.2 Thermal and Thermoelectric Material Properties

2.2.1 Thermal conductivity

Thermal conduction is the transfer of heat through a specimen in the direction opposite an applied temperature gradient. For a discrete specimen, the ability to transport heat is given as thermal conductance, and is calculated as the amount of heat transferred from a one end of the specimen to the other end, per degree of temperature difference applied between the ends. In MKS units this is expressed with units of W/K. The ability

of a specific material to transport heat is given by the geometry-normalized property called thermal conductivity, which has units of W/m-K. In solids, the total thermal conductivity k is given by:

$$\kappa = \kappa_{ph} + \kappa_e \qquad \text{Eqn. 2.1}$$

where κ_{ph} is phonon thermal conductivity and κ_e is the thermal conductivity due to electrons.

The amount of phonon thermal conductivity is given by [43]:

$$\kappa_{ph} = \frac{1}{3}C\upsilon l \qquad \text{Eqn. 2.2}$$

when C is the phonon heat capacity per unit volume, ν is average phonon velocity, and l is the phonon mean free path between Umklapp scattering events. The two critical parameters when considering the phonon thermal conductivity are the heat capacity per volume and the mean free path. The mean free path is inversely proportional to the number of excited phonons. Since the number of phonons is proportional to T, l is inversely proportional to T, resulting in an increase in phonon thermal conductivity as temperature decreases. However, at very low temperatures C begins to vary as $1/T^3$ and drives the phonon thermal conductivity back toward 0 as the temperature approaches 0 K.

The electron contribution to thermal conductivity can be expressed by a modified version of Eqn. 2.2, instead using the electron heat capacity, velocity and mean free path. Substituting for these values, based on the free electron gas model, gives

$$\kappa_e = \pi^2 n k_b^2 T\tau / m = \pi^2 n k_b^2 T\mu / e \qquad \text{Eqn. 2.3}$$

where n is the density of free electrons, k_b is Boltzmann constant, τ is the average time between collisions and $\mu = e\tau/m$. As will be shown in later sections, this form of the

relationship for κ_e is useful for analyzing the relationship between electrical conductivity and thermal conductivity, and defining the parameters necessary for achieving a high quality thermoelectric material.

2.2.2 Joule heating

Joule heating is the conversion of electrical energy into thermal energy due to the electrical resistance of a conductor. The energy conversion occurs through scattering events, in which the kinetic energy of a carrier drifting in the applied electric field is transferred to the lattice. This results in an increase in the average energy and temperature of the lattice. The amount of joule heating in an element is given by the familiar equation:

$$P = VI = I^2R \qquad \text{Eqn. 2.4}$$

where P is the amount of joule heating, V is the voltage across a conductor, I is the current through the conductor and R is resistance of the conductor. The geometry-normalized measure of resistance in a material is resistivity, with units of Ohm-m, and for metals modeled as a free electron Fermi gas, it is given as:

$$\rho = \frac{m}{ne^2\tau} = \frac{1}{ne\mu} \qquad \text{Eqn. 2.5}$$

where m is the electron mass, τ is the collision time, e is the electron charge, n is the number of electrons per volume and μ is the electron mobility.

2.2.3 The Peltier Effect

The transport of heat through a sample due to a corresponding electrical current is called the Peltier effect. As has already been mentioned, electrical carriers such as ions, electrons and holes also have thermal and kinetic energy. As a result, a charge flux due

to a flow of electrical carriers will have a corresponding flux of energy. The amount of energy transported per coulomb of charge in a material is called the Peltier coefficient, and is given by:

$$\Pi = (E_c - \mu) + \frac{3}{2} k_b T / e \qquad \text{Eqn. 2.6}$$

where μ is the Fermi level of the material and *e* is the electron charge, and E_c is the conduction band energy.

2.2.4 The Seebeck effect

The Seebeck effect is related to the Peltier effect and results in the creation of an electric field in a sample due to applied temperature gradient. Like the Peltier effect, the Seebeck effect is a result of the fact that electrical carriers also play a role in thermal transport. In this case, the presence of a temperature gradient means there is also a gradient in the population of electrons at a given energy level. At the hot end, there is a high population of high-energy carriers, and a low population of low-energy carriers. The opposite is true at the cold end of the sample. This causes low-energy carriers to diffuse from the cold end to the hot end, and the high-energy carriers to diffuse from the hot end to the cold end. However, because of their greater kinetic energy, the high-energy carriers diffuse faster than the low-energy carriers. The result is a net transfer of electrical carriers, and therefore electrical charge, in the direction opposite the thermal gradient. Under open circuit conditions, the charge gradient due to the diffusion current will result in the build up of an electric field, until drift current in the opposite direction is equal to the diffusion current. The Seebeck coefficient of a material defines how strong of an electric field will be generated by a given temperature gradient in the material, but

is more often thought of as the voltage generated across a material due to an applied temperature difference:

$$V = \alpha_s \Delta T \quad \text{Eqn. 2.7}$$

The Seebeck coefficient is also directly related to the Peltier coefficient through temperature by:

$$\Pi = \alpha_s T \quad \text{Eqn. 2.8}$$

Because of this relationship, and the relative ease of measuring voltage and temperature, it is the Seebeck coefficient that is most often measured and reported for materials, and not the Peltier coefficient. It is important to note that the Seebeck coefficient cannot be directly measured, because any material that is used to connect the test sample to the voltage meter will interfere with the measurement due to its own Seebeck coefficient, as illustrated in Figure 2.1 below. As a result, Seebeck measurements must be performed using well characterized materials with a known Seebeck coefficient as the test leads.

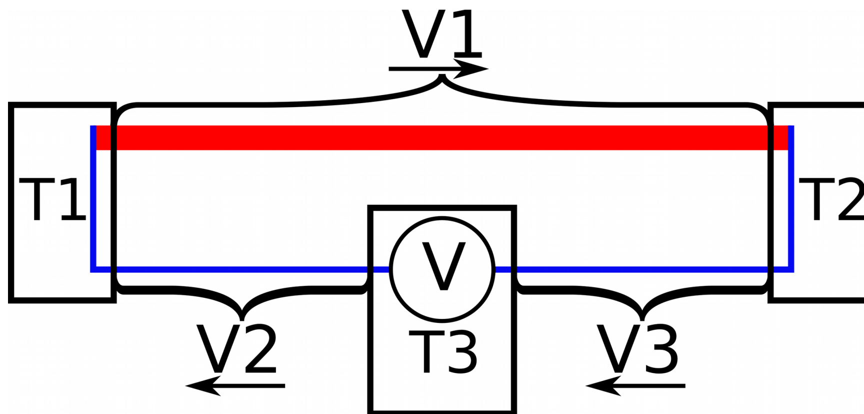


Figure 2.1: The material being measured (red) generates voltage V1=α_1(T1-T2). However this cannot be directly measured because the test leads produce voltages $V_2 = \alpha_2\ (T_3 - T_1)$ and $V_3 = \alpha_2\ (T_2 - T_3)$. The total measured voltage is V=V1+V2+V3. Where α_1 is the Seebeck coefficient of the material being investigated, and α_2 is the the Seebeck coefficient of the test leads. α_1 can only be determined if α_2 is known.

2.2.5 Thomson

The final thermoelectric effect is the Thomson effect and refers to the emission or absorption of heat by a current carrying conductor with a temperature gradient applied to it. The amount of heat emitted or absorbed is given by

$$\dot{S} = \mu J \frac{dT}{dx} \quad \text{Eqn. 2.9}$$

where μ is the Thomson coefficient and J is the current density. The Thomson coefficient is related to the Seebeck by:

$$\mu = T \frac{d\alpha_s}{dT} \quad \text{Eqn. 2.10}$$

2.2.6 Thermoelectric Figure of Merit

The key parameters that determine the quality of a thermoelectric material for cooling or power generation are the Seebeck coefficient, thermal conductivity, and electrical resistivity. The Peltier coefficient is not considered in the figure of merit because it is simply the Seebeck coefficient scaled by temperature. Similarly, the Thomson coefficient can be derived from the Seebeck coefficient if its value is known over a range of temperatures. The relationship between these properties can be quantitatively expressed by the thermoelectric figure of merit, Z, defined as [44]

$$Z = \frac{\alpha_s^2}{\kappa_t \rho_e} \quad \text{Eqn. 2.11}$$

revealing that high quality thermoelectric materials need a high Seebeck coefficient combined with low thermal conductivity and low electrical resistivity. However, optimization of Z is not simply a matter of optimizing the relevant individual material properties. This is because many of the properties are related, so optimizing one can have a negative effect on another. In particular, carrier thermal conductivity and

electrical resistivity are related through carrier concentration as well as carrier mobility. The Wiedemann–Franz relationship given in Eqn. 2.12 [43] states that for a free electron gas the ratio of electron thermal conductivity to electrical conductivity is a function of temperature and the L, the Lorenz number, where L=2.45e-8 WΩ/K^2.

$$\frac{k_e}{\sigma} = \frac{\pi^2 k_b^2 T n(\tau/3m)}{ne(\tau^2/m)} = \frac{T}{3}\left(\frac{k_b \pi}{e}\right) = LT \qquad \text{Eqn. 2.12}$$

In metals, where electron thermal conductivity dominates over phonon conductivity, this relationship means that thermal conductivity will always scale at the same rate. It has been shown that the Wiedemann–Franz relationship is also valid as a comparison between the electron thermal conductivity and electrical conductivity in semiconductor materials [44]. The lower carrier concentration in semiconductors means that thermal conduction has a strong relation to the phonon thermal conduction, and in some circumstances it is possible to reduce phonon thermal conduction without significantly effecting the thermal conduction or electrical resistivity [45]. However, semiconductors are not immune to the links between the various thermoelectric material properties. Both the Seebeck coefficient and resistivity are related to the separation between the conduction band (valence band for p-type) and the Fermi level of the material. As a results, attempts to raise the Seebeck coefficient by manipulating the Fermi level usually also result in an increase in resistivity. Because the Seebeck coefficient is more important than resistivity in determining the overall figure of merit, this type of optimization can lead to improvements, but has practical limits [46].

2.3 Bulk Thermoelectric Materials

2.3.1 Metals

In general, all metals exhibit a small (but non-zero) Seebeck coefficient. As shown in Table 1, these tend to be lower than 10 μV/K for pure metals, but can have magnitudes as high as 30 μV/K for some metals such as Cobalt. In addition to small Seebeck coefficients, metals have high thermal conductivities, leading to small Z values. This makes them a poor choice for thermoelectric cooling and power generation. However, they still find use as temperature sensors because their Seebeck coefficients are well-characterized and yield repeatable results. The low resistivity of metals means that temperature-sensing thermocouples can be made with low total resistance, helping to limit the Johnson noise of the sensor.

Table 2.1: Thermoelectric Properties of Common Metals at ~300K[a]

Material	Seebeck (μV/K)	Thermal Conductivity (W/m-K)
Gold	1.94	317
Silver	1.51	429
Aluminum	-1.66	237
Chrome	21.8	93.7
Nickel	-19.5	90.7
Tungsten	0.9	174
Platinum	-5.3	71.6
Copper	1.83	401

[a] Reference [47]

2.3.2 Semiconductors

A number of semiconductor materials have much better thermoelectric properties than the metals shown in the previous section. The presence of a bandgap means that these materials can have significantly higher Seebeck coefficients. By doping semiconductors, the carrier type can be manipulated, and a material can have either a positive or negative Seebeck coefficient, depending on whether it is doped n-type or p-type. Figure 2.2 shows the figure of merit for a number of common materials over a wide

temperature range. Between 0 °C and 200 °C, compounds based on Bi_2Te_2 and Sb_2Te_3 have the highest figures of merit among those shown. The following sections will summarize the thermoelectric properties of several commonly used bulk semiconductor thermoelectric materials including silicon, Bi_2Te_2 and Sb_2Te_3, BiSb, and PbTe. Silicon Germanium will not be summarized because it is only a useful option at temperatures well outside the scope of this research.

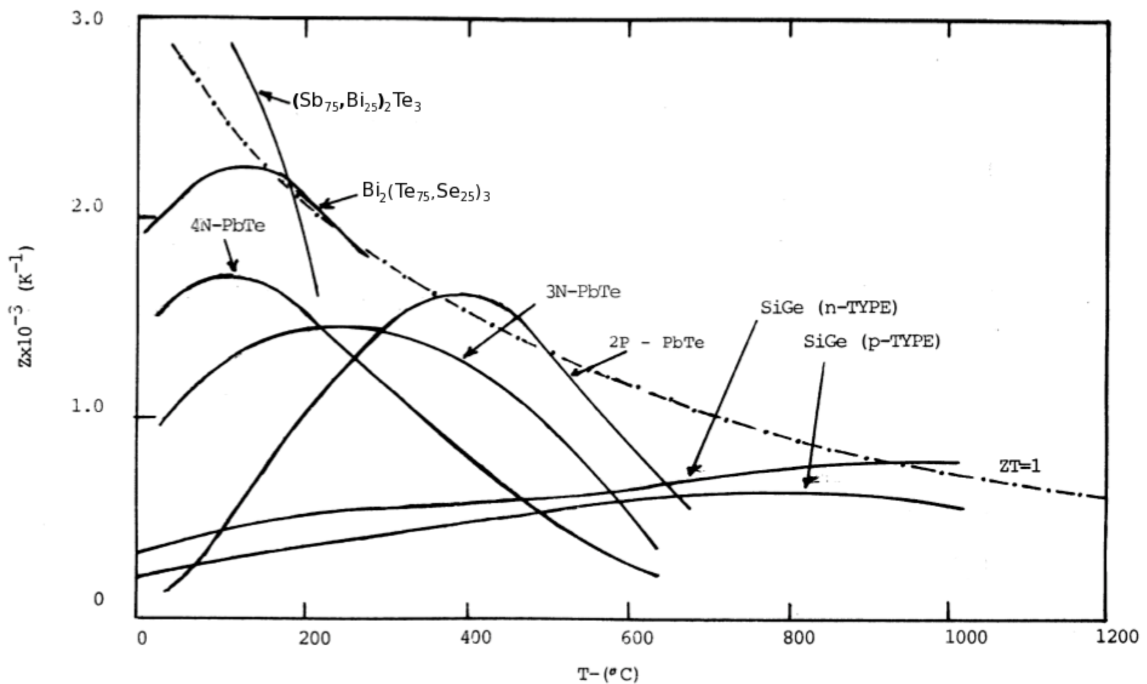


Figure 2.2: A comparison TE material figures of merit between 0 C and 1200 C. It is important to note that most materials do not have a high enough Z to exceed ZT=1 over a wide temperature range, if at all. Below 200 C tertiary compounds of $(Bi,Sb)_2Te_3$ and $Bi_2(Te,Se)_3$ provide the highest figures of merit. Reproduced from [48] (annotations edited for readability).

2.3.2.1 Silicon

Arguably the most widely-used semiconductor material, silicon is an intriguing possibility for thermoelectric applications. The wide variety of well-characterized processing technologies available means that TE technologies based on silicon could

potentially be easier to integrate with CMOS circuits and MEMS devices than other materials. Silicon has demonstrated high Seebeck coefficients that are logarithmically related to resistivity for both n-type and p-type materials. The magnitude of the Seebeck coefficient varies from about 450 μV/K at 35 μΩ-m to as high as 1600 μV/K for resistivities approaching 1 Ω-m. For sensing, the high response of high resistivity silicon may be useful, but the α^2/ρ dependence of Z means that the figure of merit for silicon is optimized in the range near 35 μΩ-m. However, even at its optimum point, silicon's high thermal conductivity of 140 W/m-k means that Z is still only 4E-5 K^{-1} (ZT=0.012 at 300K) [49]. Some researchers argue that silicon is still useful in applications that require a heat pump but not much temperature differential [50]. For applications requiring cooling below room temperature, such as those pursued in this thesis, silicon is an unacceptably poor thermoelectric material.

2.3.2.2 Lead Telluride

At temperatures above 550 K, lead telluride alloys are commonly used in thermoelectric applications. They have been shown to have a ZT of greater than 1 for portions of the temperature range between 550 K and 750 K, and can be used at temperatures as high as 900 K [48]. Because of its tolerance for high temperatures, lead telluride is often considered for use in applications such as power generation from waste heat. However, near 300 K, other materials such as bismuth telluride and antimony telluride surpass lead telluride and prove more suitable for this application.

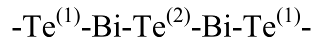
2.3.2.3 Bismuth Antimony Alloys

At the other end of the temperature spectrum, alloys of bismuth and antimony have been investigated. Although the optimum composition varies somewhat depending

on the exact temperature range in which the material will be used, n-type $Bi_{85}Sb_{15}$ has been shown to have a figure of merit of 6.5e-3 K^{-1} at 80 K (ZT=0.54), and this can be increased to as much as 11E-3 K^{-1} (ZT=.88) through the application of a large, properly aligned magnetic field [51]. Bi-Sb alloys is the best alloy for use at such low temperatures, and has been applied in coolers [52]. However, figures of merit diminish rapidly as the temperature increases. Thus, like lead telluride, it is not the best material for use in the temperature range near 300 K.

2.3.2.4 Bismuth Telluride

Bismuth telluride, Bi_2Te_3, has been widely studied as a thermoelectric material, particularly in the temperature range near 300 K. Its crystal structure is described as having a hexagonal cell, formed by alternating layers of bismuth and tellurium perpendicular to the c-axis with the following pattern:


$-Te^{(1)}-Bi-Te^{(2)}-Bi-Te^{(1)}-$

The $Bi-Te^{(2)}$ and $Bi-Te^{(1)}$ bonds are both ionic-covalent. When the cell repeats, the $Te^{(1)}-Te^{(1)}$ bonds that form are of the Van der Waals type, and are extremely weak [53]. Stoichiometric Bi_2Te_3 is p-type with α_s=227 μV/K, ρ=19.5 μΩ-m, and κ=1.73 W/m-K leading to Z=1.53e-3 (ZT=0.45). As a binary compound, however, bismuth telluride can be self-doped as either an n-type or a p-type material by creating either a tellurium rich composition or a bismuth rich composition respectively. Single crystal p-type bismuth telluride has a maximum Z perpendicular to the c-axis of 2.32e-3 (ZT=0.70), at 54.3 At.% TE. It has α_s=162 μV/K, ρ=5.5 μΩ-m, and κ=2.06 W/m-K. Single crystal n-type bismuth telluride achieves a maximum Z of 2.89e-3 (ZT=0.86) at 64.5 At.% TE, with α_s=-247 μV/K, ρ=14.3 μΩ-m, and κ=1.92 W/m-K [53].

2.3.2.5 Antimony Telluride

Antimony Telluride, Sb_2Te_3, has a similar structure to that of Bi_2Te_3, and is also widely used as a thermoelectric material near 300 K. Sb_2Te_3 is a p-type semiconductor and, like Bi_2Te_3, its composition can be manipulated to maximize the figure of merit. The optimum figure of merit occurs at 67 At.% Te. At this composition, the properties perpendicular to the c-axis are α_s=83 μV/K, ρ=1.89 μΩ-m, and κ=5.62 W/m-K, producing Z=6.54e-4 (ZT=0.20). Parallel to the c-axis the properties are better, and Z as high as 1.6e-3 (ZT=0.48) has been observed at 72 At.% Te [53]. However this is still significantly lower than the properties achieved with p-bismuth telluride.

2.3.2.6 Tertiary Alloys

As a result of having a similar crystal structure, bismuth telluride and antimony telluride can be combined in solid solutions. The creation of these tertiary alloys allows for the production of p-type materials with significantly higher figures of merit that those possible with either of the previously discussed binary compounds. Single crystal $Bi_9Sb_{31}Te_3$ has been measured to have α_s=206 μV/K, ρ=8.89 μΩ-m, and κ=1.5 W/m-K producing Z=3.2e-3 (ZT=0.96). This represents a 39% improvement over p-type Bi_2Te_3, and it is almost a factor of 5 better than Sb_2Te_3, measured perpendicular to the c-axis. On the n-type side, Bi_2Te_3 can form a solid solution with Bi_2Se_3. This is useful because achieving high quality n-type Bi_2Te_3 relies on achieving the correct composition within a very narrow tolerance. By adding selenium to the alloy, the thermoelectric properties are less dependent on the atomic ratio of the constituent elements $Bi_2(Te,Se)_3$ achieves its maximum Z of 2.9e-3 (ZT=0.87) at a composition of $Bi_{40}Te_{58.5}Se_{1.5}$ with α_s= -230 μV/K, ρ=11, μΩ-m, and κ=1.7 W/m-K [54].

2.4 Thin-film Materials

Bulk thermoelectric materials have been well optimized and have demonstrated good performance in macroscale applications such the TE modules from companies like Marlow industries that were discussed in Chapter 1. However these materials are not easily integrated with truly microscale devices. For bulk materials to be used, one of tow basic process must be used, and both have significant drawbacks. On option is to make all the elements that will form the TCs, then align them on the structure of the cooler and bond them into place. The difficulty with this process is that it requires aligning many individual parts that could be less than 20 μm on a side. The second possible process involves bonding larger pieces of TE material in place and etching (or dicing) the excess material away. Although this is potentially useful and has been shown to work in a TE generator application, it produces large amounts of wasted material. Additionally, because the available TE wafers are less than 2” in diameter, this process would be difficult to integrated with standard 4” silicon wafers. Instead, micro-scale TE devices rely on thin film thermoelectric materials that can be deposited directly onto the surface of the substrate.

Many of the same materials that are used in bulk form at the macro scale are also used in thin-film form for micro scale applications. However, because the films are not formed under equilibrium conditions, and are usually amorphous or polycrystalline, their properties can vary significantly from their single crystal versions. In some cases, the method of deposition is also of importance, as this will influence properties such as grain size and defect density in the resulting films.

2.4.1 Poly-silicon

Poly-silicon is one of the most readily available and easily process thermoelectric thin films[35],[55]. Thin film poly-silicon is similar to single crystal silicon and it has been investigated because it has reasonably good power factors and is compatible with semiconductor device manufacturing. Its thermal conductivity is around 24 W/m-K, which is much lower than single crystal silicon but still too high to be effective in most cooling applications. Like single crystal silicon, the absolute Seebeck coefficient is related to resistivity. Doping levels control both properties, but typical values of ZT at 300 K are 0.021 for n-type materials and 0.012 for p-type materials [55]. As section 2.4.2 will shown, this is over an order of magnitude lower than values that have been demonstrated by thin-film Bi_2Te_3 and Sb_2Te_3, making poly-Si an unattractive choice for applications requiring high temperature differentials.

2.4.2 Bismuth Telluride and Antimony Telluride

In bulk form, bismuth telluride, antimony telluride, and their solid solutions are the best-performing materials near room temperature. As a result, these materials have also received a significant amount of attention at the thin-film level. Several different techniques have been developed to deposit these films, and the remainder of this section will summarize these techniques and the major results from each.

2.4.2.1 Sputtering

Researchers at the University of Freiburg have used sputtered films of n-type $Bi_2(Te,Se)_3$ and p-type $(Bi,Sb)_2Te_3$. The technique relies on co-sputtering from three targets, for example Bi, Te, and Se, to create the TE thin-films. The films are deposited onto heated substrates and then further annealed at temperatures near 300°C [56]. This technique has produced n-type materials with a power factor (α_S^2/ρ) of 30 μW/cm-K, and

the p-type materials have demonstrated a power factor of 40 μW/cm-K [37]. If the thermal conductivity is assumed to be 1.7 W/m-K, such power factors correspond to a Z of 1.76e-3 (ZT=0.53) and 2.35e-3 (ZT=0.71) respectively. While the n-type material is somewhat lower than its bulk equivalent, the p-type material is approaching the performance of bulk $(Bi,Sb)_2Te_3$.

2.4.2.2 Electro-plating

Films of thermoelectric materials have also been produced by electrochemical deposition (ECD) [57-61]. The films produced have been binary and tertiary alloys of $(Bi,Sb)_2(Se,Te)_3$, and both n-type and p-type films have been grown. The constituent elements are dissolved in a nitric acid solution, which may also contain a chelating agent to prevent precipitation of insoluble oxides. The n-type Bi_2Te_3 has exhibited a Seebeck coefficient of 188.5 μV/K with a resistivity of approximately 20e-5 Ω-m for a power factor of 1.7e-3 (ZT=0.30) [61]. P-type $(Bi,Sb)_2(Te)_3$ has also been deposited, but with less success. Researchers measured Seebeck coefficients between 40 μV/K and 100 μV/K [60]. The resistivity of these films was not reported, but with such low values of α_s, the power factor and Z is almost certainly be low as well.

2.4.2.3 Metal Organic Chemical Vapor Deposition

Several groups have investigated metal organic chemical vapor deposition (MOCVD) as a means of growing thermoelectric materials. Deposited materials include binary films of Bi_2Te_3, Sb_2Te_3 [62],[63], and Bi_2Se_3 [64] as well as tertiary films of $(Bi,Sb)_2(Te)_3$ [65],[66]. Precursors for the reaction include trimethylbismuth, triethylantimony, diethyltellurium, and diethylsellenium for bismuth, antimony, tellurium, and selenium respectively. The growth temperature is in the range between

350 °C and 480 °C. N-type bismuth telluride has been produced with α_S= -216 μV/K and ρ=6.5e-6 Ω-m [63], resulting in a power factor of 7.2e-3 (ZT=1.27), more than the best value reported for bulk materials. High quality p-type materials have been produced by binary as well as tertiary films. Sb_2Te_3 films have demonstrated Seebeck coefficients of only 114 μV/K, but combined with resistivity of only 3.2e-6 Ω-m, these films nonetheless have a power factor of 4.0e-3 (ZT=0.71). $(Bi,Sb)_2(Te)_3$ films have shown increased Seebeck coefficients of α_S=240 μV/K, but also an increase in resistivity to $\rho \cong$17e-6 Ω-m [66], corresponding to a power factor of 3.3e-3 (ZT=0.58) [63]. Bi_2Se_3 trails the other materials discussed here, with α_S=-120 μV/K and ρ=18.2e-6 Ω-m, corresponding to a power factor of 0.7e-3 [64].

2.4.2.4 Molecular Beam Epitaxy

Molecular beam epitaxy (MBE) has also been use to deposit $(Bi,Sb)_2(Se,Te)_3$ compounds [67-70]. The systems are optimized to maintain very high purity, and are built around ultra-high-vacuum chambers that are capable of reaching pressures as low as 1e-10 Torr. Sources of the pure constituent elements are evaporated/sublimated from individual effusion cells that are separated from the substrate by shutters or valves. Cryogenic shields can also be used around the substrate and/or elsewhere in the chamber, to ensure that only pure beams of atoms from the effusion cells can reach the substrate. The substrate is heated so that there is enough energy at the substrate surface for the atoms in the incoming beams to form the proper bonding. Bi_2Te_3 has been deposited with a Seebeck coefficient of 180 μV/K [68], and $(Sb_{1-x},Bi_x)_2Te_3$ has been reported with a Seebeck coefficient of 184 μV/K, and a resistivity of ~2 μΩ-m [69].

2.4.2.5 Co-evaporation

Table 2.2: Comparison of Co-evaporated Bi_2Te_3 and Sb_2Te_3 Thin-films

Reference	Material	α_S (μV/K)	ρ_e (μΩ-m)	$(\alpha_S)^2/\rho_e$ (mW/K^2-m)	ZT[g]	T_{sub} (°C)
Zou et al[a]	p- Bi_2Te_3	81	3.2	2.05	0.36	314
	n- Bi_2Te_3	-228	13	3.99	0.70	260
Zou et al[b]	p- Sb_2Te_3	171	10.4	2.8	0.49	230
da Silva[c]	n- Bi_2Te_3	-228	28.3	1.84	0.32	260
	p- Sb_2Te_3	149	12.5	1.78	0.31	270
Goncalves et al[d]	n- Bi_2Te_3	-248	12.6	4.9	0.86	270
	p- Sb_2Te_3	188	12.6	2.8	0.49	220
Huang et al[e]	n- Bi_2Te_3	-208	18.8	2.3	0.41	260
	p- Sb_2Te_3	160	12.9	2.0	0.35	230

[a]Reference [71]
[b]Reference [73]
[c]Reference [46]
[d]Reference [72]
[e]Reference [74]
[g] at 300 K assuming a thermal conductivity of 1.7 W/m-K

Co-evaporation is similar to MBE in that the constituents are evaporated from individual sources and condense on the surface of the wafer, forming the final compound. However, these systems are simpler than MBE systems, and do not have the same level of control or purity as their MBE counterparts. As a result, they would be unlikely to produce single-crystal thin-film epitaxial layers. For applications where polycrystalline films are acceptable, co-evaporation can nonetheless produce films with high power factors. A number of researchers, including current work related to this thesis, have investigated co-evaporation of Bi_2Te_3 and Sb_2Te_3 thin-films, and the results are given in Table 2.2. The process and results [74] will be discussed at length in section 2.5.

2.4.3 Super-lattice Thermoelectrics

The thin-film materials discussed in the two previous sections have mostly exhibited ZT<1. Chapter 3 will show that major improvements in cooling capability are difficult without improvements in the figure of merit. To that end, super-lattice materials have been investigated. Super-lattices use alternating layers materials that are each 5 nm – 40 nm thick in order to create an artificial lattice constant that can influence the dispersion of both electrons and phonons. The most successful results of super-lattice research have produced p-type material with a ZT of 2.4, and n-type material with a ZT of 1.4 [45]. The p-type materials consist of alternating layers of MOCVD Bi_2Te_3 and Sb_2Te_3 with thicknesses of 10 Å and 50 Å respectively. The p-type material has layers of Bi_2Te_3 and $Bi_2Te_{2.83}Se_{0.17}$, again with thicknesses of 10 Å and 50 Å respectively. The materials achieve such high ZT by significantly decreasing the c-axis lattice thermal conductivity.

2.5 Materials Produced at the University of Michigan

Based on data presented in the previous sections of this chapter, it is clear that the best homogenous materials near 300 K are $(Bi,Sb)_2(Se,Te)_3$ alloys. While these materials are harder to integrate with MEMS and semiconductor process flows than silicon or poly-Si, their high figure of merit provides such a benefit to performance that it outweighs difficulties associated with processing the $(Bi,Sb)_2(Se,Te)_3$ thin films. Both n-type and p-type materials with high figures of merit can be grown using a number of different techniques.

Although tertiary thin films such as $(Bi,Sb)_2Te_3$ and $Bi_2(Se,Te)_3$, may ultimately provide the best performance, this work has focused primarily on binary films of Bi_2Te_3

and Sb_2Te_3. This has more to do with the limitation of the deposition system available, than any fundamental design choice. Super-lattice materials are another obvious choice because of the high figures of merit that have been demonstrated, but the available processing technology meant that super-lattice films could not be deposited in-house. Additionally, the SL materials have demonstrated their best figures of merit in the direction perpendicular to the layers of the super-lattice. However in Chapter 3, it will be shown that to minimize power consumption in thin film devices, the TE materials should be used in the planar orientation. This would mitigate some of the gains these materials provide.

2.5.1 Co-evaporation Process and System

For this project, we chose to use co-evaporation as the deposition technique. Co-evaporation has been previously demonstrated as an effective way to deposit both n-type and p-type materials. It has produced material with figures of merit comparable to other deposition techniques discussed above. It performs better than electroplating, and in-depth knowledge of electrochemistry is not required in order to optimize the process. It is also arguably the simplest of the vacuum deposition technologies that were presented; it requires no unusual chemical precursors or reaction effluents that need to be processed.

The co-evaporation process was described briefly in a previous section, and will be elaborated here. Co-evaporation must be used instead of evaporation from a single source because the thin-film TE materials being deposited are compounds of two elements. The constituent elements, in this case Bi, Sb, and Te, have different melting temperatures and vapor pressures. If an alloy or compound were used as the evaporation source, the rate at which each element evaporated from the source might not correspond

to the concentration of the element in the source material. The result would be an incorrect atomic ratio of elements in the thin film, despite a correct ratio in the source.

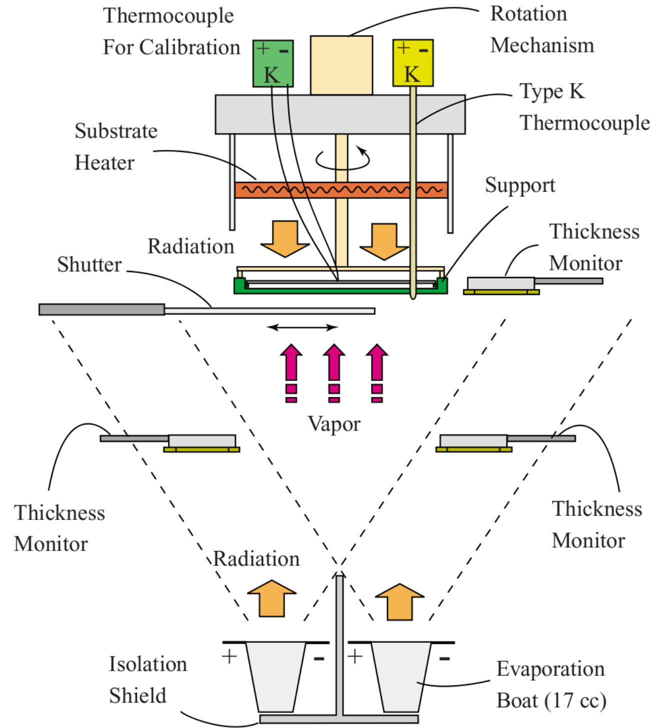


Figure 2.3: Schematic of a two-source co-evaporation system. They system is housed in a high vacuum chamber and includes two crucibles, a crystal thickness monitor for each of the crucibles, and a third monitor near the wafer surface. The temperature of the substrate can be controlled and the substrate can be rotated (graphic reproduced from Huang et al [74]).

To avoid this problem, the co-evaporation process provides a separate source for each element used to create the final compound. The deposition rate of each source can be measured independently, and the power supplied to heat each source is controlled independently. As a result, the relative flux rates of the different materials at the wafer surface can be controlled, allowing control over the atomic percentages of each material in the final film. However, film composition alone is not enough to ensure a high quality

TE material. The orientation, size and structure of the polycrystalline domains are also important. To manipulate the crystal growth, the temperature of the substrate should also be controlled. A basic schematic of the deposition system is shown in Figure 2.3.

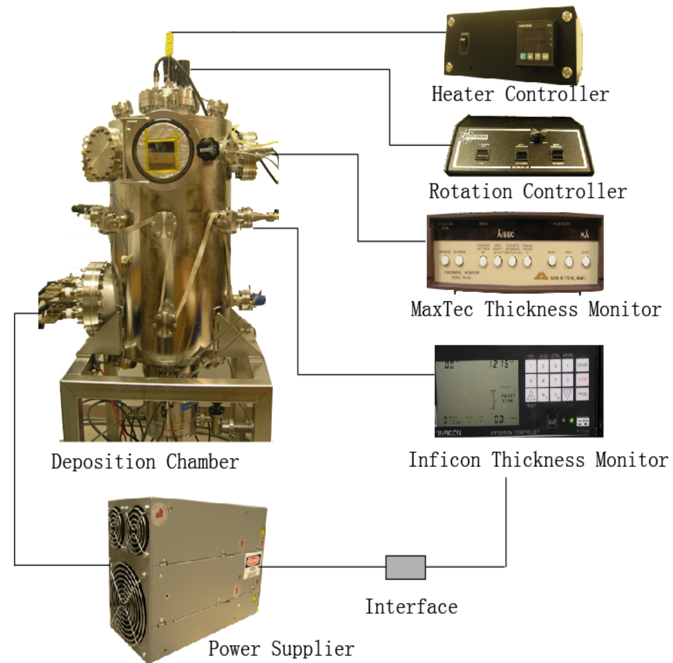


Figure 2.4: The original TE deposition system.

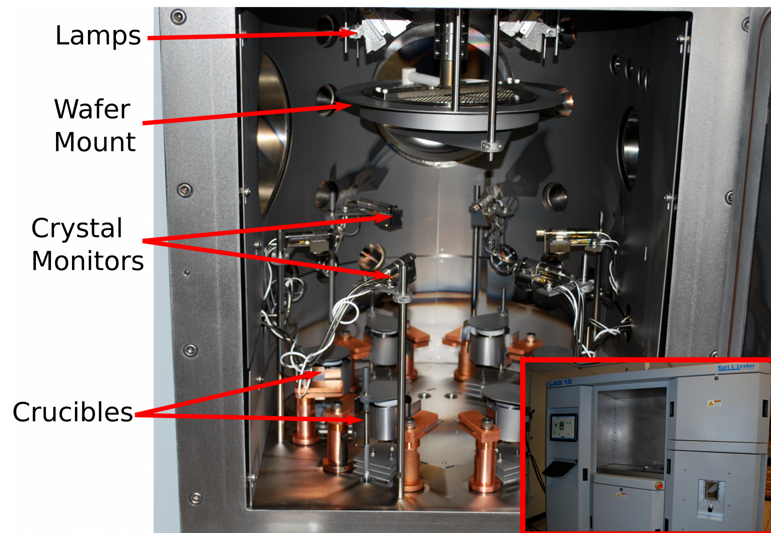


Figure 2.5: TE deposition system from Kurt J. Lesker Company. The main picture shows the interior of the deposition chamber. The inset shows the exterior of the tool.

Two different systems were used to deposit the thermoelectric materials used for this project. The first system was built in-house for a previous project and is shown in Figure 2.4. This system could accommodate one 100 mm wafer. It housed up to two source materials that were evaporated from resistively heated tungsten crucibles. These source materials could be controlled automatically by the thickness monitors or manually. Each source had its own crystal thickness monitor that was mechanically shielded from the opposite source to prevent cross talk. The wafer was manually loaded into the chamber. It could be heated by radiation from a resistive element above it, and rotated with a variable speed motor. Thickness of the final film was measured with a third crystal monitor located at the plane of the wafer, and wafer temperature was measured by a calibrated thermocouple located near the wafer. Vacuum is generated using a turbo-molecular pump backed by a rotary vane mechanical pump, and pressures as low as 2e-7 Torr can be achieved by backing the chamber.

The second system, supplied by Kurt J. Lesker company, is more modern, flexible and automated. It includes 6 sources, allowing tertiary alloys to be deposited. It also means that sources do not need to be changed between depositions. Each material source is held in an alumina crucible that is heated by a resistive element. Each source is individually shuttered, and there is also a wafer shutter. The system has a crystal thickness monitor for each source, and the crystals are shuttered to allow for discrete sampling of the deposition rate. The power to each source is automatically controlled based on the measured rate. This allows the lifetime of the crystals to be increased, and makes longer depositions possible. The wafer is loaded through a load-lock, eliminating the need to vent the chamber between consecutive runs. The main chamber uses a cryo-pump to achieve high vacuum as low as 1E-9 Torr and halogen lamps to heat the wafer.

Despite the differences, the actual deposition procedure is similar for both systems. The wafer is loaded into the process chamber face down. The process chamber is then pumped to its high-vacuum base pressure. Next the substrate heater is turned on, and the wafer temperature is allowed to stabilize. Next, the source crucibles are heated. During this step the wafer is isolated from the evaporated materials by shutters, however the rate monitors are exposed to the evaporated materials. Once the rates have stabilized at the appropriate levels, the shutters are opened, and the deposition proceeds until the correct thickness is reached, as measured by the crystal monitors. The shutters are closed, and both the source heaters and the substrate heaters are shut off. The substrate is allowed to cool slowly before the chamber is vented and the substrate is removed. This prevents oxidation that would be caused by immediate venting of the system.

2.5.2 Material Properties and characterization

The properties of the thermoelectric materials were characterized across a wide range of deposition conditions, and the best deposition conditions for each material were determined. The deposition conditions explored were the temperature of the substrate, and the ratio between the fluxes from each source. The thermoelectric thin films were deposited on silicon wafers that were coated with a thin film of CVD silicon dioxide. The measurements of the material properties were performed in a cryostat using the arrangement in Figure 2.6.

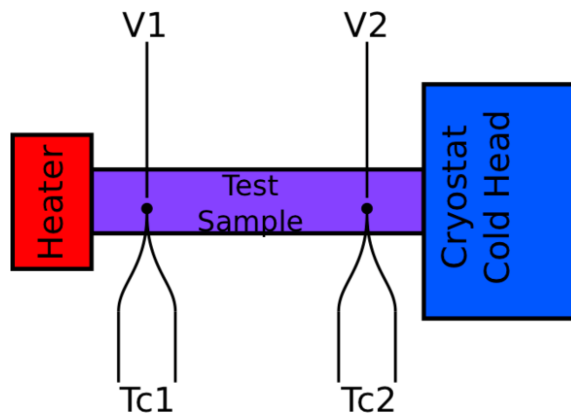


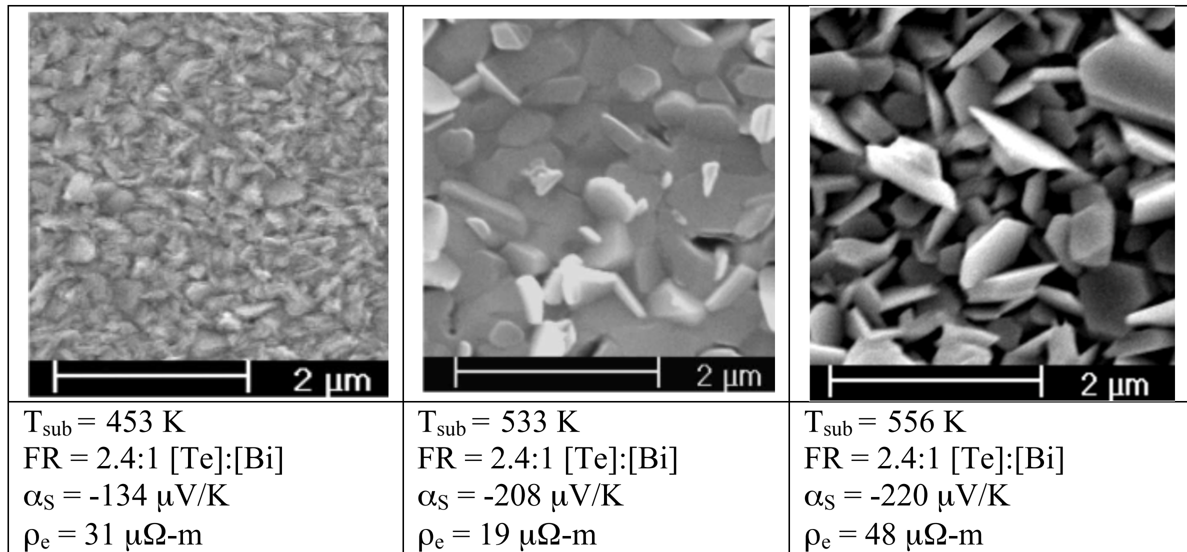
Figure 2.6: Schematic of a Seebeck measurement apparatus

To perform the tests samples were cleaved from the wafer, and a resistive heater was attached to one end of each sample using adhesive. K-type thermocouples were soldered to the thermoelectric films at two positions on the sample. An electrical connection was also made to the thermoelectric film at the position of each thermocouple. Finally, the samples were mounted to the temperature-controlled heat sink of the cryostat by mechanical clamping. While under vacuum, the heater was used to produce a temperature gradient across the sample. The gradient was measured by the thermocouples and the Seebeck voltage was measured using the attached electrical leads.

The Seebeck coefficients of the samples were calculated using the temperature and voltage data. The resistivity was measured using a four-point probe along the length of the sample. From these measurements thermo-power can be calculated, and Z can be estimated by assuming a value for the thermal conductivity of the thin films. Here, the thermal conductivity will be assumed to be 1.7 W/m-k.

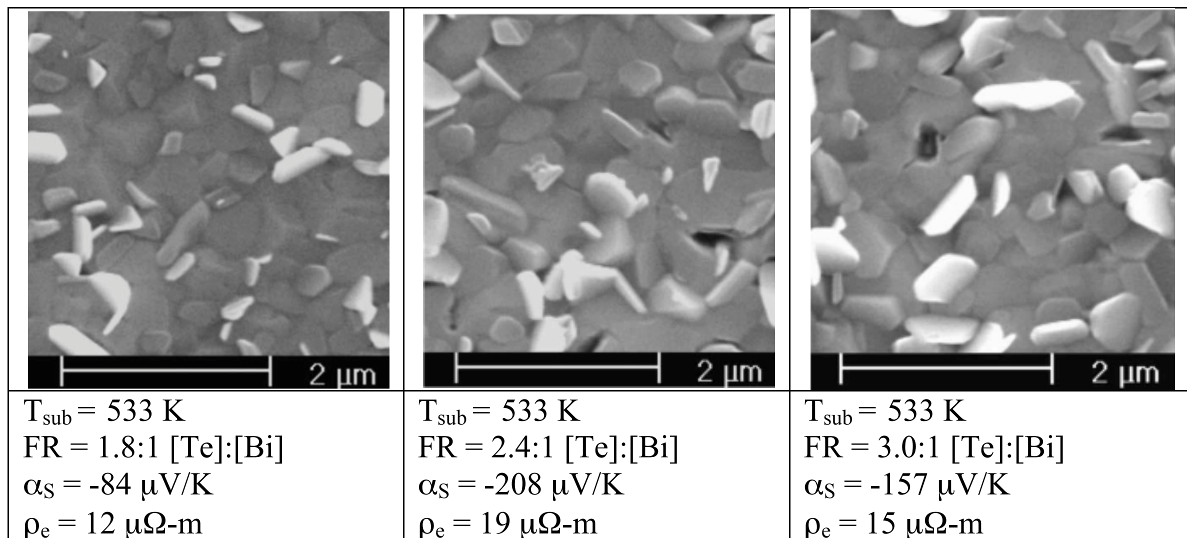
The material characterization summarized in this section was performed by Baoling Huang and was originally presented in [74]. Bismuth telluride was deposited at substrate temperatures of 453 K, 533 K and 556 K with a Te:Bi flux ratio (FR) of 2.4:1. The resulting films were observed using SEM imaging, and the resulting pictures are shown in Figure 2.7. It can be seen that between 453 K and 533 K, the grain size increased significantly. At the same time, the Seebeck coefficient increased and the resistivity dropped. Increasing the temperature further to 556 K resulted in still bigger grains, but they were not as well aligned, showing more out-of-plane protrusions. In this case, the Seebeck coefficient increased slightly, while the resistivity increased 2.5 times.

A second set of samples was deposited with the temperature held constant at 533 K, while the flux ratio was set to 1.8, 2.4 and 3.0. The surfaces resulting from these films are shown in Figure 2.8. In this case, the visual appearance of the films remained relatively constant while the resistivity and Seebeck coefficient varied. Based on these tests, the best films were produced at a substrate temperature of 533 K with a FR of 2.4, and produced a film with α_S = -208 mV/K and ρ_e = 19 mW-m. This corresponds to a power factor of 2.2E-3 K^{-1} and ZT= 0.40 at 300 K, assuming κ=1.7 W/m-K.



T_{sub} = 453 K FR = 2.4:1 [Te]:[Bi] α_S = -134 µV/K ρ_e = 31 µΩ-m	T_{sub} = 533 K FR = 2.4:1 [Te]:[Bi] α_S = -208 µV/K ρ_e = 19 µΩ-m	T_{sub} = 556 K FR = 2.4:1 [Te]:[Bi] α_S = -220 µV/K ρ_e = 48 µΩ-m

Figure 2.7: Bismuth telluride films deposited at 453 K, 533 K, and 556 K, using a flux ratio of 2.4:1 along with measured thermoelectric properties. Images were reproduced from [74].



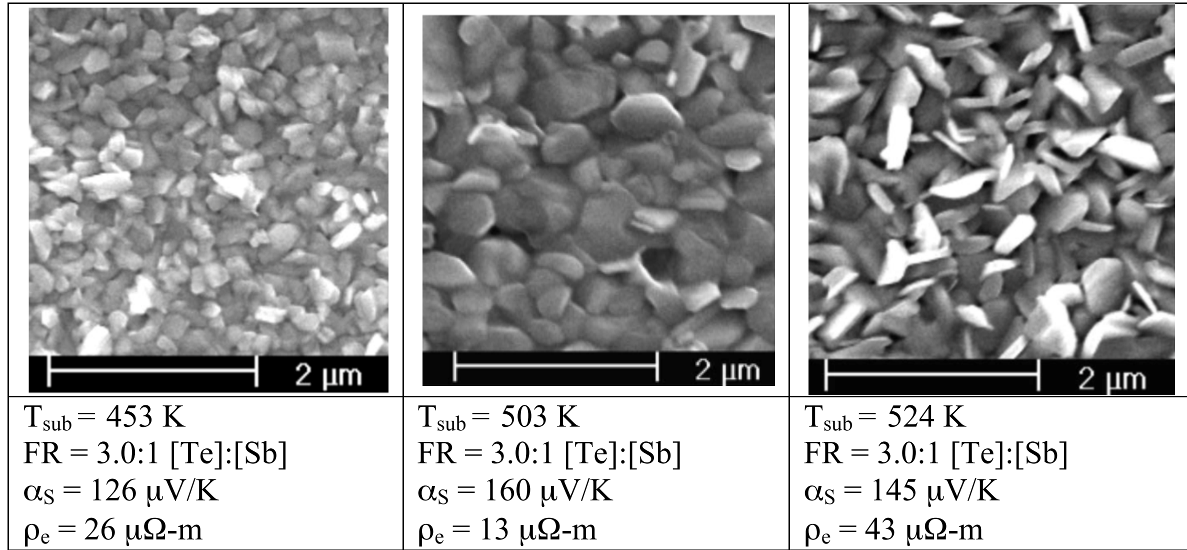
T_{sub} = 533 K FR = 1.8:1 [Te]:[Bi] α_S = -84 µV/K ρ_e = 12 µΩ-m	T_{sub} = 533 K FR = 2.4:1 [Te]:[Bi] α_S = -208 µV/K ρ_e = 19 µΩ-m	T_{sub} = 533 K FR = 3.0:1 [Te]:[Bi] α_S = -157 µV/K ρ_e = 15 µΩ-m

Figure 2.8: Bismuth telluride films deposited at flux ratios 1.8:1, 2.4:1, and 3.0:1 at a temperature of 533 K, along with measured thermoelectric properties of the films. Images were reproduced from [74].

A similar set of experiments was run using antimony telluride. It was tested at substrate temperatures of 453 K, 503 K and 524 K with a Te:Sb flux ratio of 3.0:1. The

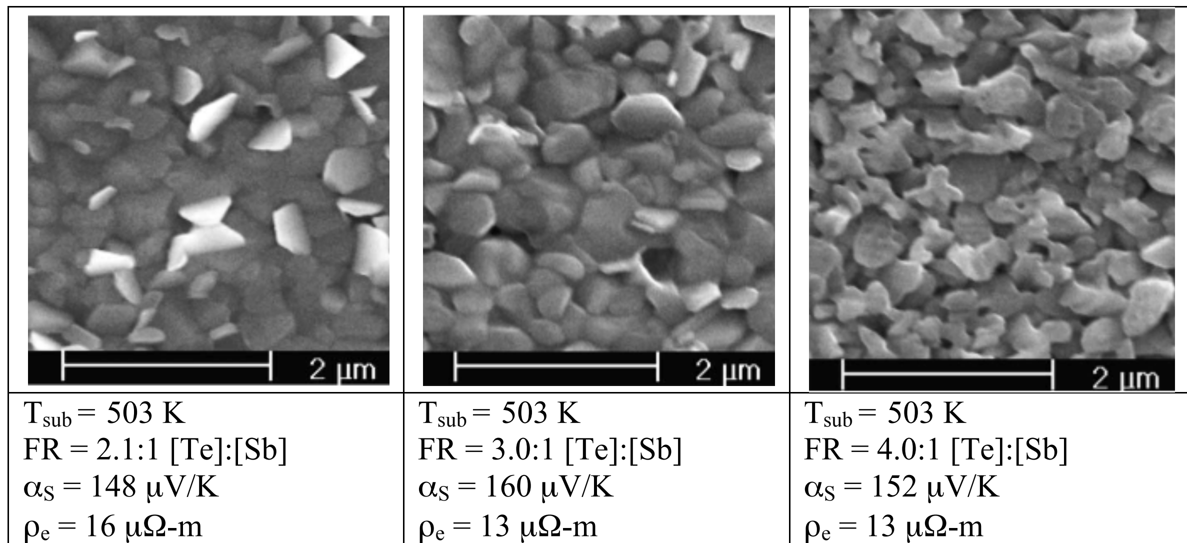
resulting films were observed using SEM imaging, and the resulting pictures are shown in Figure 2.9. Like the Bi_2Te_3 films, the Sb_2Te_3 films exhibited an increase in grain size between the lowest and middle deposition temperatures. Between the 503 K and 524 K, the grain size does not visibly change much, but the orientation of the grains is less uniform and the film is less dense, with visible dislocations at the grain boundaries.

A second set of samples was deposited with the temperature held constant at 503 K, while the flux ratio was set to 2.1, 3.0 and 4.0. The surfaces resulting from these films are shown in Figure 2.10. The two lower flux rates produce films that are visually similar, and vary only slightly in power factor. On the other hand, the tellurium-rich films grown at the highest FR have grains that are smaller and that appear more randomly shaped than the lower FR films. In this case, the visual appearance of the films remained relatively constant while the resistivity and Seebeck coefficient varied. Based on these tests, the best films were produced at a substrate temperature of 503 K with a FR of 3, and produced a film with α_S = 160 mV/K and ρ_e = 13 mW-m. This corresponds to a power factor of 1.9E-3 K^{-1} and ZT= 0.34 at 300 K, assuming κ=1.7 W/m-K.



T_{sub} = 453 K	T_{sub} = 503 K	T_{sub} = 524 K
FR = 3.0:1 [Te]:[Sb]	FR = 3.0:1 [Te]:[Sb]	FR = 3.0:1 [Te]:[Sb]
α_S = 126 μV/K	α_S = 160 μV/K	α_S = 145 μV/K
ρ_e = 26 μΩ-m	ρ_e = 13 μΩ-m	ρ_e = 43 μΩ-m

Figure 2.9: Antimony telluride Films deposited at 453 K, 503 K, and 524 K, using a flux ratio of 2.4:1 along with measured thermoelectric properties. Images were reproduced from [74].



T_{sub} = 503 K	T_{sub} = 503 K	T_{sub} = 503 K
FR = 2.1:1 [Te]:[Sb]	FR = 3.0:1 [Te]:[Sb]	FR = 4.0:1 [Te]:[Sb]
α_S = 148 μV/K	α_S = 160 μV/K	α_S = 152 μV/K
ρ_e = 16 μΩ-m	ρ_e = 13 μΩ-m	ρ_e = 13 μΩ-m

Figure 2.10: Antimony telluride films deposited at flux ratios 2.1:1, 3.0:1, and 4.0:1 at a temperature of 503 K, along with measured thermoelectric properties of the films. Images were reproduced from [74].

In addition to the thermoelectric properties, the atomic composition and the crystal structure of both films were analyzed. The influence of the flux ratio on the films

was investigated by depositing multiple samples with varying flux ratios while holding the substrate temperature constant. For bismuth telluride, the substrate temperature was held at 533 K while for antimony telluride the substrate temperature was held at 503 K. The flux ratio was then varied from 1.5 to 3.5 (4.0 for antimony telluride). The atomic percentage of each constituent in the films was measured by energy dispersive x-ray analysis (EDX). Pure Bi_2Te_3 and Sb_2Te_3 have an atomic ratio of Te:Sb/Bi equal to 1.5. In the Bi_2Te_3 films this is achieved with a flux ratio of approximately 2.25, and for Sb_2Te_3 it requires a flux ratio of approximately 2.5. The optimum thermoelectric properties for both films, however, are achieved at higher flux ratios, and both optimum films are tellurium rich. Figure 2.11 illustrates the data on flux ratio and composition for both materials.

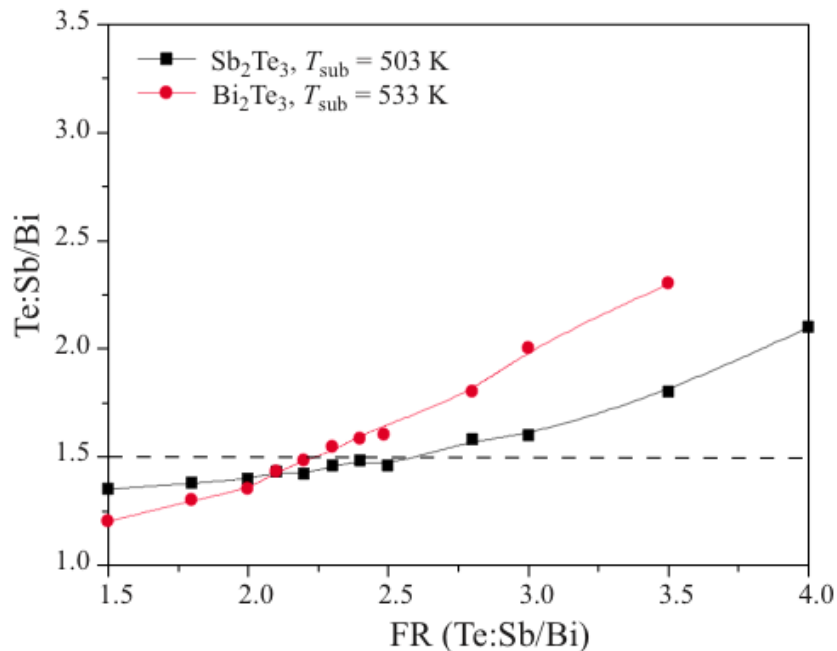


Figure 2.11: Atomic ratio of Te:Bi and Te:Sb vs. the flux ratio used during deposition. Image reproduced from [74].

Finally the films with the best thermoelectric properties were analyzed using x-ray diffraction to study their crystal properties. The spectrum of both films is shown in

Figure 2.12. The measured peaks match well with the spectra of powdered single crystal Bi_2Te_3 and Sb_2Te_3, shown as the dased lines in the figure. This indicates that in both films only a single crystal phase is present. Additionally, only (0,0,n) peaks are observed in the films, indicating a preferential c-axis orientation.

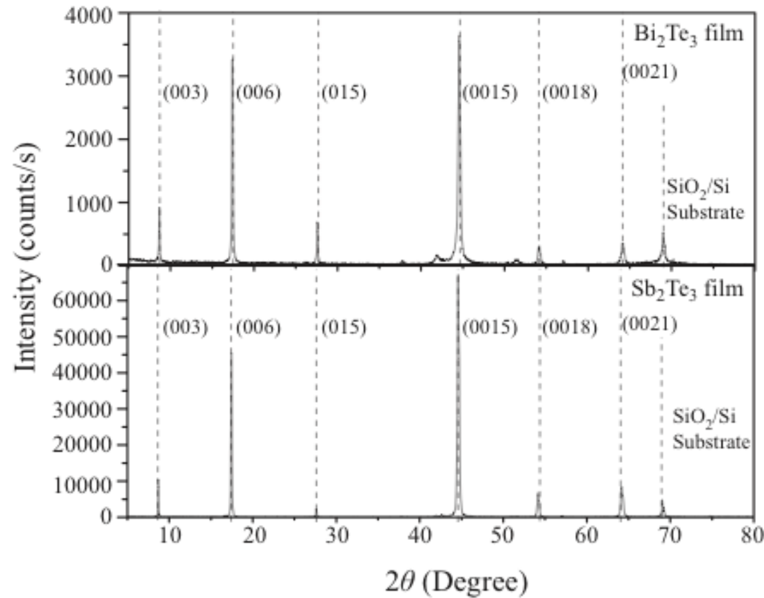


Figure 2.12: X-ray diffraction spectra for Bi_2Te_3 and Sb_2Te_3. The dashed lines are from the spectra of powdered, single-crystal Bi_2Te_3 and Sb_2Te_3. Image reproduced from [74].

2.6 Summary

For this project Bi_2Te_3 and Sb_2Te_3 were chosen because they provide the best thermoelectric properties at temperatures near 300 K. Co-evaporation was chosen because is offers good quality films at the wafer level. The films can be optimized by changing the flux ratio and the substrate temperature. Changing these parameters influences both film composition and grain structure in the thin films, as well as the Seebeck coefficient and resistivity. The results of the optimization process are presented above. The best Bi_2Te_3 and Sb_2Te_3 films have a ZT of 0.41 and 0.35 respectively. Although not the highest reported values of ZT for binary thin films, they exceed

previously realized films produced at the University of Michigan, and are within the range of values reported by others. Chapter 3 will discuss how thermoelectric thin films are used to create an effective cooler. It will also show how performance of the cooler is related to the thermoelectric properties of the material being used.

Chapter 3
Modeling Thermoelectric Coolers

Analytical modeling and optimization of thermoelectric coolers has been discussed by many groups [82],[81],[83-86],[40],[87],[75],[88],[76-80],[89],[39],[44],[90], and these investigations use a number of different performance criteria assessing performance. These include the minimum temperature that can be achieved, the cooling capacity of the device (i.e. the amount of heat removed from the cold stage), the power consumption of the device, and the coefficient of performance (COP) of the cooler (the cooling capacity divided by the power consumption of the device). The goal of this project is develop a low-power cooler capable of achieving large temperature differentials, in order to obtain improved performance from other devices. The key performance goals for this work are therefore achieving the minimum absolute temperature in tandem with minimizing the total power consumption of the device. As a result, the analysis for this project focuses primarily on achieving a high temperature differential, and secondarily on how to achieve that temperature differential with as little power as possible. Previous single and multistage coolers have been able to achieve temperatures below 260 K, from an ambient temperature of 300 K, but required more than 300 mW of power. The goal of this project has been to achieve similar temperature differentials while 100 mW of power. Therefore, the investigation modeling and optimization presented in this chapter will focus on maximizing the temperature differential and minimizing the power consumption. COP will not be part of the

optimization process because is has been shown that the optimum COP is achieved when temperature differentials are lower than the maximum level possible[44].

The previous chapter discussed the material properties that are important to the performance of a thermoelectric material, and Z, the figure of merit that is used as the relative measure of a material's thermoelectric quality. The goal of this chapter is to present how those materials can be used to construct a thermoelectric cooler, and to demonstrate how several key parameters affect the performance of the cooler. The chapter will begin with a description of how a basic, ideal thermocouple is used for single stage cooling. A 1-D analytical model of a thermocouple will be described and used to analyze the effect of thermocouple geometry on cooler performance and power consumption. Non-idealities such as parasitic thermal conduction and electrical resistance will be added to the model, and their effects on cooler performance will be analyzed. Next, multistage cooling will be investigated, primarily through the example of a two-stage cooler. Three different means for setting the current in the stages will be presented and compared. The effect of the relative number of thermocouples per stage will also be investigated, along with the influence of parasitic thermal conduction and electrical resistance.

3.1 The Ideal Thermocouple

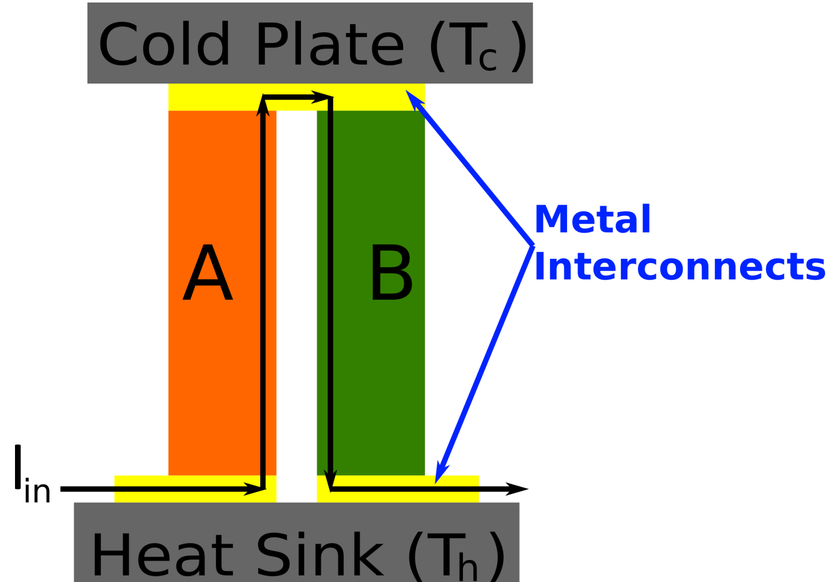


Figure 3.1: Illustration of a basic thermocouple arrangement for cooling.

A thermoelectric cooler is made by pairing two materials of differing Peltier co-efficients to form a thermocouple (TC) as illustrated in Figure 3.1. The electrical connection between the two materials is often referred to as a junction, but it is not necessarily a metallurgical junction (i.e., the junction of an n-type and a p-type material will not form a diode). If the Seebeck coefficients are such that $\alpha_A < \alpha_B$, and current flows from material A to material B, then material A is transporting heat into the junction equal to $Q_A=\alpha_A T_c I$. At the same time, material B is transporting heat out of the junction equal to $Q_B=\alpha_B T_c I$. In this case $Q_A<Q_B$. The result is a net removal of heat from the junction due to the Peltier effect, and causing the junction to decrease in temperature. The Peltier heat flowing out of the junction is maximized by maximizing the difference in the Seebeck coefficient between materials *A* and *B*. This can be accomplished by using an n-type material for *A* and a p-type material for *B*, making $\alpha_A<0$ and $\alpha_B>0$. In that

case, both materials remove heat from the junction and contribute to cooling. The total amount of heat removed by the Peltier effect is [44],[91]

$$Q_{\Pi} = T_c\left(\alpha_p - \alpha_n\right)I \qquad \text{Eqn. 3.1}$$

The Peltier effect alone does not determine how much cooling can be achieved. As discussed in the previous chapter, the electrical resistivity and thermal conductivity also determine the quality of a thermoelectric material. Thermal conduction through the TE material results in a heat flow from the hot side of the TC to the cold side. Assuming both the n-type and p-type material are the same size, the amount of heat transferred due to thermal conduction, Q_{Kth} is [44]

$$Q_{k_{th}} = \Delta T K_{th} = \left(T_h - T_c\right)\left(\kappa_n + \kappa_p\right)\frac{A}{l} \qquad \text{Eqn. 3.2}$$

Where K_{th} is the thermal conductance of the TE element, A is the cross sectional area of the TE element, and l is the TE elements length parallel to the current flow. Electrical resistivity adds heat to the system through Joule heating. The amount of Joule heat produced, Q_{Re}, in the thermocouple is [44]

$$Q_{R_e} = I^2 R_e = I^2\left(\rho_n + \rho_p\right)\frac{l}{A} \qquad \text{Eqn. 3.3}$$

Where R_e is the electrical resistance of the TE element. For a thermocouple element with a uniform cross-section, the Joule heat is produced uniformly across the length of the thermocouple, with half the Joule heat going to each end.

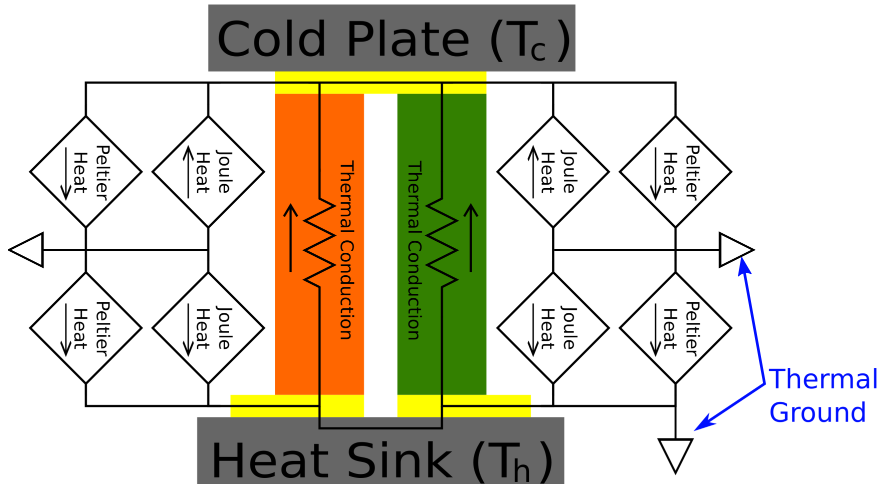


Figure 3.2: Heat flow in an ideal thermocouple, when it is used as a cooler. The heat flows are shown as an electrical equivalent thermal circuit, over a diagram of the thermocouple.

These heat flows are shown schematically in Figure 3.2. A steady state temperature is achieved when the heat flow into the junction (due to thermal conduction and joule heating) matches the heat flow out of the junction (due to Peltier effect). Summing the heat flows results in Eqn. 3.4 [44].

$$\begin{aligned} Q_{\Pi} &= \frac{1}{2} Q_{R_e} + Q_{K_{th}} = \\ T_c(\alpha_p - \alpha_n) I &= \frac{1}{2} I^2 R_e + \Delta T K_{th} = \\ T_c(\alpha_p - \alpha_n) I &= \frac{1}{2} I^2 (\rho_p + \rho_n) \frac{l}{A} + (T_h - T_c)(\kappa_p + \kappa_n) \frac{A}{l} \end{aligned} \qquad \text{Eqn. 3.4}$$

Which can be rearranged to express the temperature difference as:

$$\Delta T = \frac{T_c(\alpha_p - \alpha_n) I - \frac{1}{2} I^2 R_e}{K_{th}} = \frac{T_c(\alpha_p - \alpha_n) I - \frac{1}{2} I^2 (\rho_p + \rho_n) \frac{l}{A}}{(\kappa_p + \kappa_n) \frac{A}{l}} \qquad \text{Eqn. 3.5}$$

From this relationship, it is clear that the temperature difference has a quadratic relationship to the input current. It will therefore have a point of maximum temperature

differential with respect to current input. Solving Eqn. 3.6 for I gives the current that is needed to maximize ΔT, which is shown as I_{opt} in Eqn. 3.7. Unless otherwise noted, I_{opt} in this thesis refers to the current that optimizes the temperature differential, as opposed to the current that optimizes the COP, which is sometimes used in other sources.

$$\frac{d\Delta T}{dI} = \frac{T_c\left(\alpha_p - \alpha_n\right) - I\left(\rho_p + \rho_n\right)\frac{l}{A}}{\left(\kappa_p + \kappa_n\right)\frac{A}{l}} = 0 \qquad \text{Eqn. 3.6}$$

$$I_{opt} = \frac{T_c\left(\alpha_p - \alpha_n\right)}{R_e} = \frac{T_c\left(\alpha_p - \alpha_n\right)}{\left(\rho_p + \rho_n\right)\frac{l}{A}} \qquad \text{Eqn. 3.7}$$

By substituting I_{opt} back into Eqn. 3.5, the maximum achievable cooling is found to be:

$$\Delta T_{\max} = I_{opt} = \frac{T_c^2\left(\alpha_p - \alpha_n\right)^2}{R_e K_{th}} = \frac{T_c^2\left(\alpha_p - \alpha_n\right)^2}{2\left(\rho_p + \rho_n\right)\left(\kappa_p + \kappa_n\right)} \qquad \text{Eqn. 3.8}$$

In the ideal case, $-\alpha_n = \alpha_p = \alpha$, $\kappa_n = \kappa_p = \kappa$, $\rho_n = \rho_p = \rho$. When this is true, Eqn. 3.8 can be reduced to [44]:

$$\Delta T_{\max} = \frac{T_c^2 \alpha^2}{2\rho\kappa} = \frac{Z T_c^2}{2} \qquad \text{Eqn. 3.9}$$

where Z is the intrinsic thermoelectric figure of merit. The above relationship shows clearly how the thermoelectric figure of merit is related to the overall temperature performance of the thermoelectric cooler. This is an important result because it illustrates that for an ideal thermocouple with matched properties for the n-type and p-type materials, the maximum achievable ΔT is dependent only on the figure of merit of the materials used, and the temperature of the cold side. Based on Eqn. 3.5 through Eqn. 3.8, a normalized temperature differential can be defined as $\Delta T_{norm}=\Delta T/\Delta T_{max}$, and a

normalized input current can be defined as I_{norm}=I/I_{opt}. The equation for ΔT_{norm} with respect to I_{norm} is:

$$\Delta T_{norm} = 2I_{norm} - I_{norm}^2 \qquad \text{Eqn. 3.10}$$

This relationship is plotted in Figure 3.3 and is significant because it is independent of the specific material properties or geometric scale of the cooler being analyzed. It gives a sense of how all coolers will behave when the input current is swept.

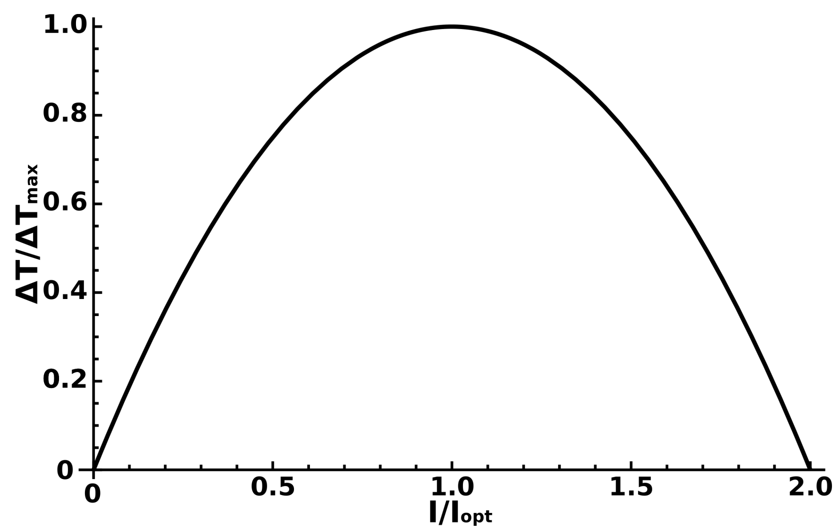


Figure 3.3: The generalized cooling curve of an ideal thermocouple with fixed T_c. The graph is plotted from Eqn. 3.10 and shows $\Delta T/\Delta T_{max}$ plotted versus I/I_{opt}.

However, the relationships described by Eqn. 3.4 through Eqn. 3.10 are problematic because they present ΔT as the dependent variable and T_c as an independent variable. In realistic coolers, T_h is the fixed temperature and T_c varies with input current. Solving Eqn. 3.4 for T_c instead of ΔT results in a slightly more complex relationship, as shown in equation Eqn. 3.11.

$$T_c = \frac{I^2\rho\left(\frac{l}{A}\right)^2 + 2\kappa T_h}{2\alpha I\left(\frac{l}{A}\right) + 2\kappa} \qquad \text{Eqn. 3.11}$$

In this case I_{opt} becomes

$$I_{opt} = \frac{-\kappa + \sqrt{\kappa^2 + \frac{2\alpha^2\kappa T_h}{\rho}}}{\alpha\left(\frac{l}{A}\right)} \qquad \text{Eqn. 3.12}$$

Subsitituting I_{opt}, back into Eqn. 3.11 produces the minimum temperature of the TC, given by

$$T_{c,\min} = \frac{\rho\kappa}{\alpha^2} + \frac{\rho}{\alpha^2}\sqrt{\kappa^2 + \frac{2\alpha^2\kappa T_h}{\rho}} = \frac{-1 + \sqrt{1 + 2ZT_h}}{Z} \qquad \text{Eqn. 3.13}$$

While slightly more complex than the formulation for ΔT_{max} that was expressed in Eqn. 3.9, this relationship for $T_{c,min}$ has many of the same properties. It is a function of only Z and T_h, once again indicating that the temperature performance of the TC is only dependent on the material figure of merit, not the geometry of the TC. Normalizing $T_{c,min}$ with respect to T_h gives [44]

$$\frac{T_{c,\min}}{T_h} = \frac{-1 + \sqrt{1 + 2ZT_h}}{ZT_h} \qquad \text{Eqn. 3.14}$$

Figure 3.4 shows this normalized cooling versus the ZT from 0 to 4. The most significant gains in the cooling are realized as ZT increases from 0 to 1, with diminishing returns thereafter. For example, at ZT = 0.25, $T_{c,min}/T_h$=0.90. For a hot side temperature of T_h=300 K, this corresponds to a cold side temperature of T_c= 270 K. At ZT = 0.5, $T_{c,min}/T_h$=0.83, corresponding to T_c= 249 K. At ZT = 1.0, $T_{c,min}/T_h$=0.73, corresponding to T_c= 219 K, and at ZT = 2.0, $T_{c,min}/T_h$=0.72, corresponding to T_c= 186 K. These values

represent a best-case analysis that could never be achieved in a real world scenario due to parasitic effects and temperature dependencies of the TE material properties. However, it is still useful information because it provides a bound on the potential temperature performance of a single stage cooler. While Figure 3.4 shows only how $T_{c,min}$ varies with ZT, Figure 3.5 shows how the normalized temperature changes with current for several values of ZT. The current in Figure 3.5 has been normalized with respect to I_{opt}.

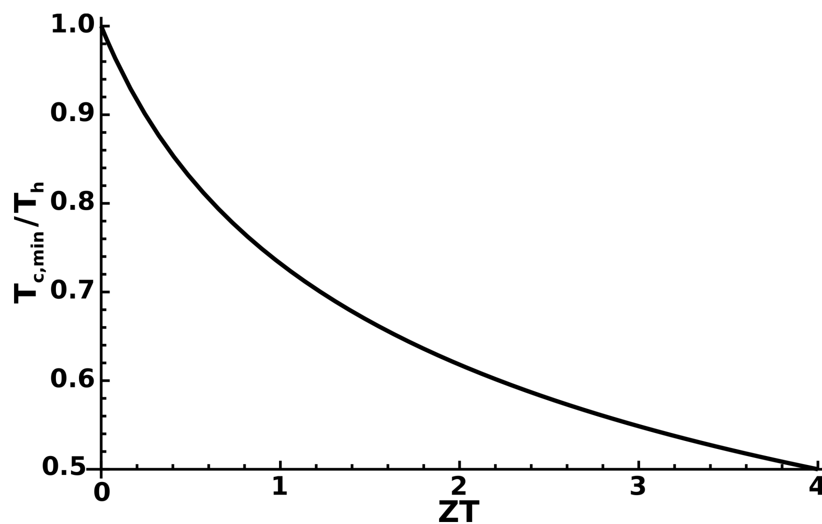


Figure 3.4: The effect of ZT on cooling. The graph is plotted from Eqn. 3.14 and shows the normalized minimum temperature of a single, ideal thermocouple plotted versus ZT.

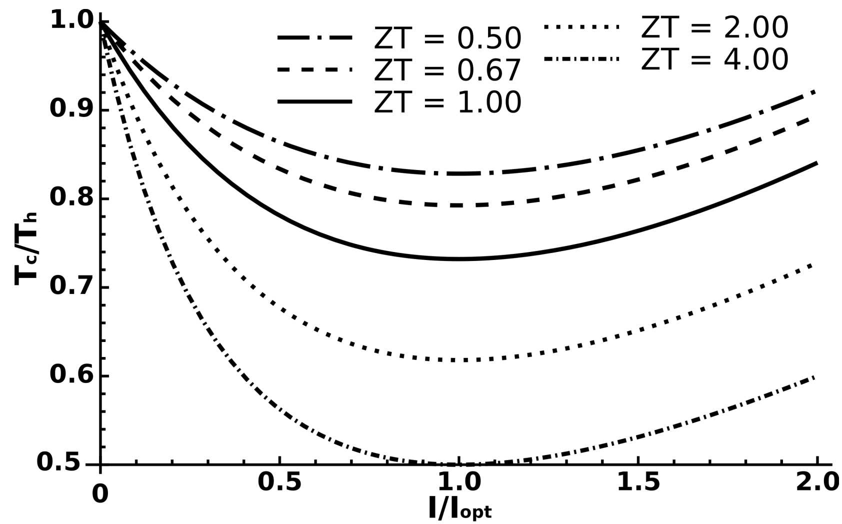


Figure 3.5: The effect of current input on the cold side temperature of a single, ideal thermocouple. The graph is plotted from Eqn. 3.11 and shows normalized temperature of a single thermocouple plotted versus normalized current input for several values of ZT.

The analysis has so far focused on the temperature performance that a single TC can achieve with different thermoelectric figures of merit. The results show that there is no relationship between the geometry of the cooler and minimum achievable temperature. Therefore, there is no fundamental limiting factor that should prevent the effective minimization of thermoelectric coolers into the micro domain. Indeed, as discussed in Chapter 1, several microscale thermoelectric coolers have been developed and commercialized that can compete with macroscale coolers in terms of temperature differential. However, these devices dissipate powers that are several times too large to be useful in integrated sensing applications. To overcome this limitation, it is important to understand how scaling effects the power consumption of the cooler, as well as the overall temperature performance. To understand the power scaling, it is useful to first look at how the optimum current scales with geometry. Looking back to Eqn. 3.12, I_{opt} is shown to be inversely proportional to the aspect ratio, (l/A), of the TC. This

relationship is dependent on α, ρ and κ, as well as T_h and the aspect ratio of the TC. Rearranging the expression gives a better indication of the effect each parameter has.

$$I_{opt} = \left(\frac{\kappa}{\alpha}\right)\frac{-1+\sqrt{1+2ZT_h}}{l/A} \qquad \text{Eqn. 3.15}$$

I_{opt} can be normalized with respect to κ/α in order to show how scaling will affect the current level in a generic material with a given ZT. Figure 3.6 illustrates that relationship, clearly showing an order of magnitude decrease in the quantity $I_{opt}(\alpha/\kappa)$ for every order of magnitude increase in the aspect ratio.

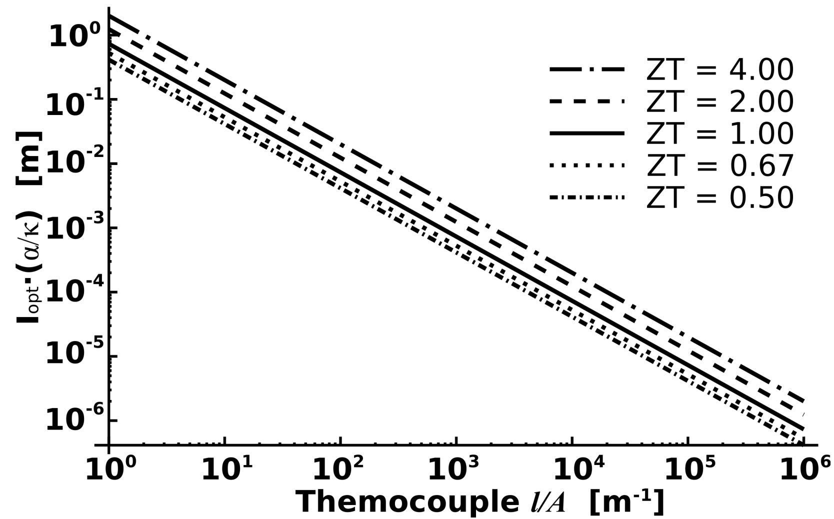


Figure 3.6: The effect of aspect ratio on the optimum current of a single, ideal thermocouple. The optimum current, normalized with respect to material properties, is plotted from Eqn. 3.15 versus aspect ratio for several values of ZT.

The power consumption of the thermocouple at I_{opt} is given by:

$$P_{\max} = 2I_{opt}^2R_e + 2\alpha\left(T_h - T_c\right)I_{opt} \qquad \text{Eqn. 3.16}$$

Substituting Eqn. 3.15 for I_{opt} and Eqn. 3.13 for $T_{c,min}$ results in an expression for power consumption that can be reduced to:

$$P_{\max} = \left(\frac{A}{l}\right)\left(\frac{\kappa}{Z}\right)\left[\frac{4(ZT_h)^2\sqrt{1+2ZT_h}}{(1+2ZT_h)} + \frac{2(ZT_h)\sqrt{1+2ZT_h}}{(1+2ZT_h)} - 2(ZT_h)\right] \qquad \text{Eqn. 3.17}$$

Figure 3.7 shows plots of the quantity QZ/κ. Like the optimum current, the power consumption at I_{opt} decreases 1 order of magnitude for every order of magnitude increase in aspect ratio of the TC.

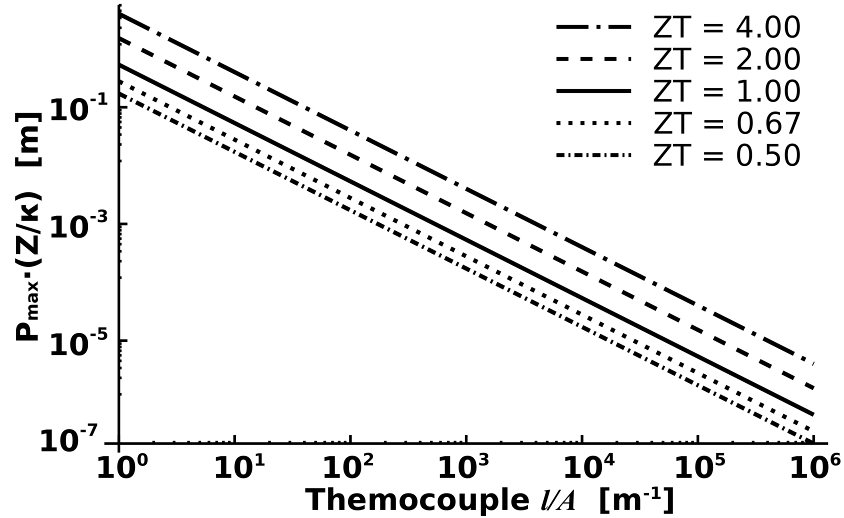


Figure 3.7: The effect of aspect ratio on power consumption of single, ideal thermocouple. Power dissipation, normalized with respect to specific material properties, is plotted from Eqn. 3.17 versus aspect ratio for several different values of ZT.

This clear relationship between power consumption and the aspect ratio of the thermocouple explains why existing micro thermoelectric coolers continue to have high power consumption levels, and also points the way to developing a thermoelectric micro cooler with low power consumption. To understand why, please consider the following examples. A small, macro-scale cubic TE element with 2 mm edges will have an aspect ratio of 500 m^{-1}. A cube of TE material with edges that are 20 μm on a side has an aspect ratio of 50,000 m^{-1}, and would represent a 100x decrease in power consumption from the macro TE cooler. This example is similar in dimension to existing thermoelectric micro

coolers [39],[37],[38], and it represents a significant reduction in power from the macro scale devices, however, as shown in Chapter 1. These micro-scale coolers still have power consumptions that are far to high to be useful in embedded applications. To further reduce power the aspect ratio must be increased even more. To achieve higher aspect ratios the cooler can be designed to carry heat in the plane of the thin film, instead of perpendicular to the plane of the film. By doing this, the thickness of the film no longer defines the length of the TC. Instead the length is defined litho-graphically, and the thickness of the film, combined with the TC width define the cross sectional area. If the thermocouple is constructed out of a thin film with a thickness of 2 μm, and has a width of 100 μm and a length of 60 μm, the aspect ratio would be 300,000 m^{-1}. This represents an additional 6x decrease in power from the cubic micro scale cooler and a 600x decrease in power for the macro scale example.

The decrease in total power consumption does not come without tradeoffs. By scaling the size, the optimum current and power consumption are reduced, but the thermal conductance of the TE segment is also scaled inversely to *l/A*. Any real device that requires cooling will dissipate some amount of power, and that power acts as a thermal load on the system. The thermal load is added to the analytical system as a heat source at the cold temperature node in Eqn. 3.4, as shown below.

$$
\begin{aligned}
Q_{\Pi} &= \frac{1}{2}Q_{R_e} + Q_{K_{th}} + Load \\
T_c\left(\alpha_p - \alpha_n\right)I &= \frac{1}{2}I^2R_e + \Delta T K_{th} + Load \\
T_c\left(\alpha_p - \alpha_n\right)I &= \frac{1}{2}I^2\left(\rho_p + \rho_n\right)\frac{l}{A} + \left(T_h - T_c\right)\left(\kappa_p + \kappa_n\right)\frac{A}{l} + Load
\end{aligned}
\qquad \text{Eqn. 3.18}
$$

$$\Delta T = \frac{T_c\left(\alpha_p - \alpha_n\right)I - \frac{1}{2}I^2 R_e - Load}{K_{th}}$$

$$= \frac{T_c\left(\alpha_p - \alpha_n\right)I - \frac{1}{2}I^2\left(\rho_p + \rho_n\right)\frac{l}{A} - Load}{\left(\kappa_p + \kappa_n\right)\frac{A}{l}} \qquad \text{Eqn. 3.19}$$

The result is that for given TE material, a cooler with a lower aspect ratio will experience a smaller increase in temperature when presented with a thermal load than a cooler with a large aspect ratio. Figure 3.8 illustrates this point. It shows how the normalized minimum temperature increases with aspect ratio when the TC is subject to a thermal load. If the load is equal to (1E-5)κT_h, then for ZT=1 $T_{c,min}/T_h$=0.73 when the aspect ratio is 500 m^{-1}. When the aspect ratio is 50,000 m^{-1}, $T_{c,min}/T_h$=0.87, and when the aspect ratio is 300,000 m^{-1}, $T_{c,min}/T_h$=1.45, indicating that TC can no longer cool below the temperature of the heat sink with a load of this power.

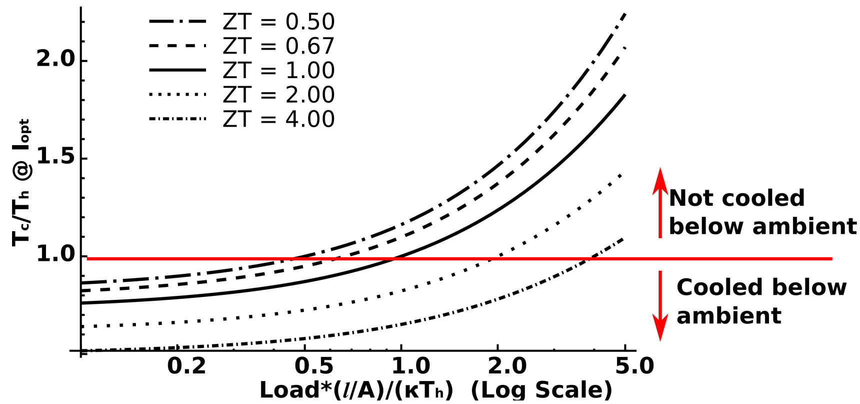


Figure 3.8: The effect of a load on the cooling of a single, ideal thermocouple. The normalized minimum temperature is plotted versus load, normalized with respect to aspect ratio, specific material properties, and T_h (see Eqn. 3.19). Several different values of ZT are plotted.

3.2 Non-Ideal Thermocouple

The previous analysis sheds a useful light on the effects of scaling TE coolers to the micro domain, but it is inadequate to fully anticipate the performance of a physical thermocouple. First, it assumes that the properties of both the n-type material and the p-type material are the same, and that the sizes of both of the thermocouple elements are the same. Second, it does not take into account any parasitic effects that may be present. These include contact resistances, resistances in any interconnecting metal, and additional thermal paths between the hot side and the cold side. All these factors will produce a reduction in the effectiveness of the TE cooler, and need to be accounted for in any complete model of a thermoelectric cooler. To see how these factors will impact the overall cooler, they must first be included in the heat balance equation initially expressed in Eqn. 3.4. Revising the equation to allow for materials with different properties and sizes yields the following relationship:

$$T_c\left(\alpha_p - \alpha_n\right)I = \frac{1}{2}I^2\left(\frac{\rho_p l_p}{A_p} + \frac{\rho_n l_n}{A_n}\right) + \left(T_h - T_c\right)\left(\frac{\kappa_p l_p}{A_p} + \frac{\kappa_n l_n}{A_n}\right) \qquad \text{Eqn. 3.20}$$

In this relationship, the heat removed from the junction is again shown on the left side of the equation, but now the aspect ratio of each material is different. Despite this change, the basic relationships in the previous section are still useful for estimating the optimal temperature performances of the TC. The intrinsic figure of merit, Z, can be replaced with an extrinsic figure of merit, Z_{ext}, that includes the geometric effects.

$$Z_{ext} = \frac{\left(\alpha_p - \alpha_n\right)^2}{K_{tot}R_{tot}} = \frac{\left(\alpha_p - \alpha_n\right)^2}{\left(\frac{\rho_p l_p}{A_p} + \frac{\rho_n l_n}{A_n}\right)\left(\frac{\kappa_p l_p}{A_p} + \frac{\kappa_n l_n}{A_n}\right)} \qquad \text{Eqn. 3.21}$$

It has been shown that the extrinsic figure of merit is maximized when the following condition is met [44]:

$$\left(\frac{A_p / l_p}{A_p / l_p} \right) = \sqrt{\frac{\kappa_n \rho_p}{\kappa_p \rho_n}} \qquad \text{Eqn. 3.22}$$

In that case, the geometric dependencies will cancel out, and the extrinsic figure of merit for the pair of materials becomes [44]:

$$Z_{ext,\max} = \frac{\left(\alpha_p - \alpha_n\right)^2}{\left[\left(\kappa_p \rho_p\right)^{1/2} + \left(\kappa_n \rho_n\right)^{1/2}\right]^2} \qquad \text{Eqn. 3.23}$$

This can be used to estimate the maximum temperature performance that can be expected from a pair of real materials, but still does not take into account additional parasitic effects. These include parasitic electrical resistance, R_{par}, and parasitic thermal conductance, K_{par}. R_{par} can be attributed to contact resistance and resistance in metal interconnects between TCs, and K_{par} is a result of any addition thermal paths between the cold side and hot side of the cooler. These thermal paths include conduction and convection through the air and any non TE material that may be present to provide structural support. To include parasitic resistance and thermal conduction, the following heat balance equation is used:

$$\left(\alpha_p - \alpha_n\right) T_c I = \left(T_h - T_c\right)\left(\frac{\kappa_n A_n}{l_n} + \frac{\kappa_p A_p}{l_p} + K_{par} \right) + \frac{1}{2} I^2 \left(\frac{\rho_n l_n}{A_n} + \frac{\rho_p l_p}{A_p} + C R_{par} \right) \qquad \text{Eqn. 3.24}$$

where C is the ratio of R_{par} on the cold side to R_{par} on the hot side. The equation for Z_{ext} then becomes.

$$Z_{ext} = \frac{(\alpha_p - \alpha_n)^2}{K_{tot}R_{tot}} = \frac{(\alpha_p - \alpha_n)^2}{\left(\frac{\rho_p l_p}{A_p} + \frac{\rho_n l_n}{A_n} + CR_{par}\right)\left(\frac{\kappa_p l_p}{A_p} + \frac{\kappa_n l_n}{A_n} + K_{par}\right)} \qquad \text{Eqn. 3.25}$$

Z_{ext} can replace Z in Eqn. 3.13 to calculate $T_{c,min}$, and Figure 3.9 show the effect of K_{par} and R_{par}. To isolate the effects of the parasitic values, this figure was made using the assumption that the magnitude of each material property is the same for both the n-type material and the p-type material, and that both legs of the thermocouple have the same size. The calculation also uses the assumption that the parasitic resistance is uniformly distributed between the two ends of the thermocouple so *C* in Eqn. 3.25 is 0. In the figure the parasitic effect being investigated (K_{par} or R_{par}) is varied in comparison to the intrinsic value for a single thermocouple, while the other parasitic effect is held at 0. So, for example, K_{par}/K_{TE}=1 is the point where the parasitic resistance per thermocouple is equal to the resistance of the TE material in one thermocouple, and the total resistance of the thermocouple is twice what it would be if K_{par}/K_{TE}=0.

The graph show that for parasitic resistance or conductance of less than 10% of the intrinsic value, there is little impact on the overall performance of the cooler. For example, a material with a ZT of 0.5 will exhibit an increase in normalized minimum temperature, $T_{c,min}/T_h$, of only 6.6% when the value of R_{par}/R_{TE} is increased from 0 to 0.1. Increases in R_{par}/R_{TE} of greater than 0.1, however, lead to significant performance degradation. Increasing R_{par}/R_{TE} to 1 is equivalent to reducing intrinsic Z by half, and results in the same loss of performance. This can be seen in the figures where the normalized temperature is the same for ZT=4 at R_{par}/R_{TE}=1 and for ZT=2 as R_{ar}/R_{TE} approaches 0. If the parasitic values are allowed to reach as high as 10 times the intrinsic

value, only a small percentage of the potential performance of the material can be realized.

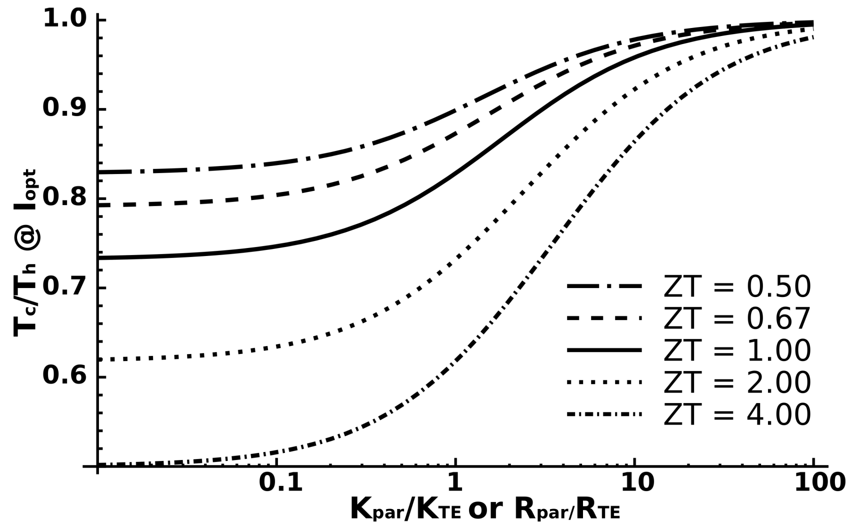


Figure 3.9: The effect of parasitic thermal conductance on cooling for a single, ideal thermocouple. The normalized minimum temperature is plotted versus the K_{par}/K_{TE} or R_{par}/R_{TE} (see Eqn. 3.25) where $K_{TE}=\kappa A/l$ and $R_{TE}=\rho l/A$.

The information taken from Figure 3.9 can be specifically applied to thermoelectric cooler design at the micro scale. As it was shown earlier, low power designs require high aspect ratio thermocouples that can only be practically realized using lateral planar coolers implemented at the micro-scale. The high aspect ratio also results in high electrical resistance in the thermocouple, and low thermal conductance in the thermocouple. This means that a micro-scale cooler can tolerate a level of absolute parasitic electrical resistance that is higher than its macro-scale counterparts. At the same time, the low thermal conductance means that the parasitic conductance must also remain low. This information helps to inform the design of the micro-cooler, and will be revisited in future chapters, when analyzing specific thermoelectric cooler designs.

3.3 Multistage Cooling

The analysis in the preceding sections focused on the temperature performance of a single thermocouple. However, based on Eqn. 3.13, there is a material dependent limit to the minimum temperature that can be achieved by single thermocouple. Because most available thermoelectric materials are limited to a ZT of less than 1, the minimum achievable temperature from a single thermocouple may not be low enough for many applications. To overcome the limitations of a single thermocouple, multistage cooling can be used [44],[39]. This arrangement creates a higher temperature differential by stacking multiple stages of thermocouples on top of one another, as illustrated in Figure 3.10. This section will build on the model for a single TC developed in the previous sections, and present a 1-D analytical model for two-stage thermoelectric cooler. It will also discuss the design principles that are necessary to effectively develop a multistage thermoelectric cooler, and the trade-offs associated with using a multistage configuration. The limits of multi-stage cooling will also be considered within the requirements of the low-power target applications.

3.3.1 Ideal 2-Stage Cooler

Analytical modeling of multistage systems has been explored by many groups since early in the development of thermoelectric cooling [75-80],[39]. However, as with single stage coolers, most of the investigation has focused on maximizing COP. The work here will start with the same basic model as presented elsewhere for two-stage and multistage coolers [44], but the analysis will focus only on how various parameters affect the overall temperature and power consumption. Constructing the model requires extending the basic heat balance expression in Eqn. 3.4 into a multi-equation system.

The one-stage system was modeled at steady state by equating the amount of heat flow into the cold side with the amount of heat flow out of the cold side. In a similar way, a multi-stage system is modeled by setting up such a relationship at each interface between consecutive stages. This interface is shown as the Inter-Stage Zone in Figure 3.10 below, and has a temperature T_1.

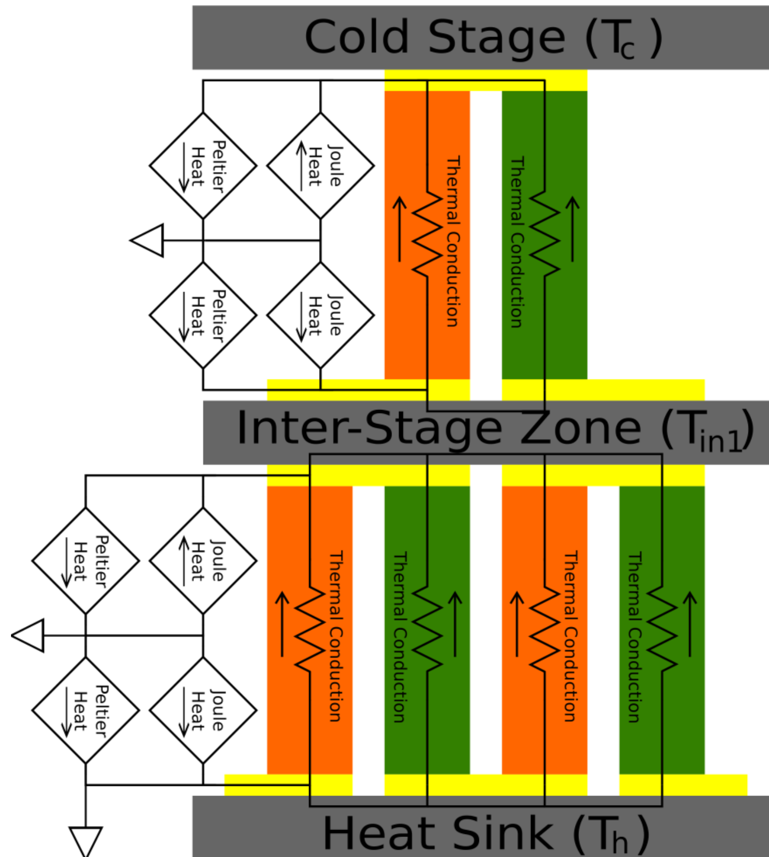


Figure 3.10: Heat flow schematic for a 2-stage cooler.

Equating the heat flows at the inter-stage zone results in the following relationship:

$$N_1(\alpha_p - \alpha_n)T_1I_1 + (T_1 - T_c)\left(N_2\left[\frac{\kappa_n A_{n2}}{l_{n2}} + \frac{\kappa_p A_{p2}}{l_{p2}}\right] + K_{par2}\right) = N_1(\alpha_p - \alpha_n)T_1I_2 + \ldots$$
$$(T_h - T_1)\left(N_1\left[\frac{\kappa_n A_{n1}}{l_{n1}} + \frac{\kappa_p A_{p1}}{l_{p1}}\right] + K_{par1}\right) + \frac{1}{2}I_1^2\left(N_1\left[\frac{\rho_n l_{n1}}{A_{n1}} + \frac{\rho_p l_{p1}}{A_{p1}}\right] + C_1R_{par1}\right) + \ldots \qquad \text{Eqn. 3.26}$$
$$\frac{1}{2}I_2^2\left(N_1\left[\frac{\rho_n l_{n2}}{A_{n2}} + \frac{\rho_p l_{p2}}{A_{p2}}\right] + C_2R_{par2}\right)$$

where, N_1 and N_2 are the number of thermocouples at their respective stages. Equating the heat flows at T_c, results in

$$N_2(\alpha_p - \alpha_n)T_cI_2 = (T_1 - T_c)\left(N_2\left[\frac{\kappa_n A_{n2}}{l_{n2}} + \frac{\kappa_p A_{p2}}{l_{p2}}\right] + K_{par2}\right) + \frac{1}{2}I_2^2\left(N_1\left[\frac{\rho_n l_{n2}}{A_{n2}} + \frac{\rho_p l_{p2}}{A_{p2}}\right] + C_2R_{par2}\right) \qquad \text{Eqn. 3.27}$$

Assuming the temperature of the hot side and the material properties are known, the above relationship represents a system of two equations with two unknowns, T_1 and T_c, and they can be solved simultaneously to determine analytical equations for these temperatures. The solutions can be quite complex, and are not shown here. The following analysis will assume that $\rho=\rho_n=\rho_p$, $\kappa=\kappa_n=\kappa_p$, $\alpha=\alpha_n=\alpha_p$, and that all of the thermocouples are sized equally. Also, unless otherwise noted, the analysis in this section will assume that the parasitic electrical resistance and thermal conductance are zero.

Adding a second stage to the cooler offers new degrees of freedom to the design. The new parameters include the ratio N_1 to N_2 (N_1/N_2), the current of the first stage, and the current in the second stage. Each of these parameters has an impact on the minimum achievable temperature and the power consumption of the cooler. The following sections

will explore the effect of different methods for setting the current in each stage, and the effect that N_1/N_2 has on temperature, current and power consumption.

3.3.2 Sequential Current Optimization (Fixed Stage-1 Current)

One method that can be used to set the current in each of the two stages is to optimize them sequentially. The first step is to find the optimum current of the first stage with zero current applied to the second stage. Because there is no thermal load, this is the same as finding I_{opt} for a 1-stage cooler, and Eqn. 3.12 can be used. The current in the first stage is fixed at this optimum current, and then the current in the second stage is varied to find the minimum value of $T_{c.}$ Figure 3.11 shows how the temperature varies with the I_2 for several different values of N_1/N_2. All the curves in the figure are for a material with ZT=1. The curves all have a Y-axis intercept of 0.73 because this is the amount of cooling that can be achieved by the first stage when the second stage completely off. The figure shows that even if $N_1/N_2 = 1$, there is a still a gain that can be realized by independently controlling the current through the second stage. To further look at how N_1/N_2 effects the temperature performance, Figure 3.12 plots $T_{c,min}/T_h$ vs. N_1/N_2, for several different values of ZT. When compared with the data for a one-stage cooler in the previous section, this analysis shows that there is a significant gain to be made by adding a second stage, regardless of ZT and even if $N_1/N_2 = 1$. Additional gains in performance are predicted by using ratios of 2 or more, but there is a limit to the amount that can be gained as the ratio increases past 10.

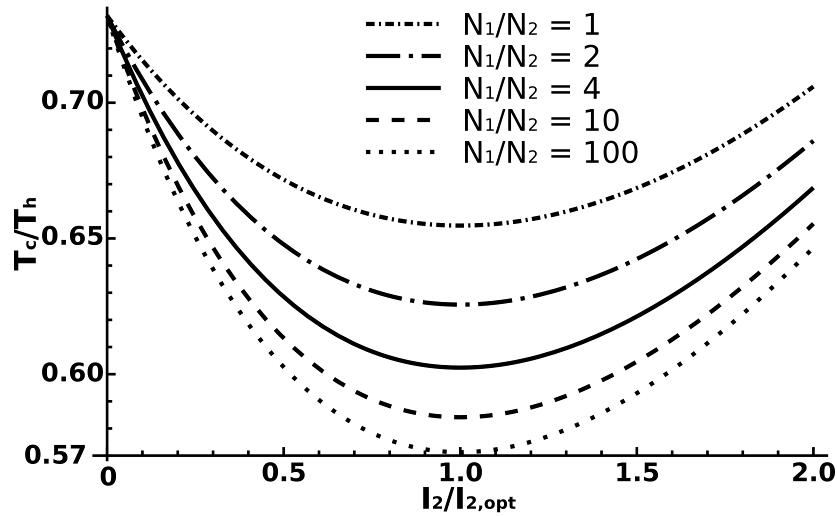


Figure 3.11: The effect of the second-stage current on the temperature of a 2-stage cooler, when the currents are set by sequential optimization. Normalized temperature is plotted versus I_2, normalized with respect to $I_{2,opt}$ for several different values of N_1/N_2. All curves are for ZT=1.

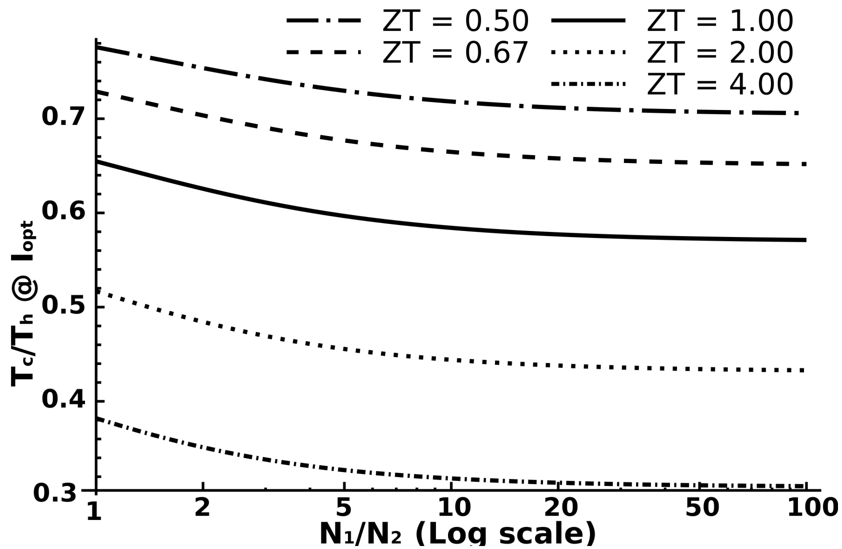


Figure 3.12: The effect of N_1/N_2 on the temperature of a 2-stage cooler, when the currents are set by sequential optimization. Normalized minimum is plotted versus N_1/N_2 at several values of ZT.

Both Figure 3.11 and Figure 3.12 show how performance at $I_{2,opt}$ varies in relationship to the ratio of thermocouples. But $I_{2,opt}$ varies with N_1/N_2 , and this variation is shown in Figure 3.13. As N_1/N_2 is increased, the current I_2 also increases, and at large

values of N_1/N_2 it approaches a stable value that is dependent on the ZT of the material. However, $I_{2,opt}$ never approaches $I_{1,opt}$, even for very large values of N_1/N_2. The reason is that I_{opt} of a given stage is temperature dependent and T_c is always less than T_1 at the optimum current points with no load.

To understand why the optimum current in the second stage varies with the N_1/N_2, as well as why the overall temperature varies with N_1/N_2, it is useful to look at how T_1 and T_c vary with respect to each other as the current I_2 is increased. Figure 3.14 shows an example of this variation. With $I_2=0$, and I_1 set to its optimum value, $T_c=T_{1,min}$. As I_2 is increased, $\Delta T=T_h-T_c$ also increases, but T_h-T_1 begins to increase due to the increased heat load produced by the second stage. ΔT continues to increase until $I_{2,opt,}$ at which point any further increase in I_2 will cause T_h-T_1 to decrease more than T_1-T_c will increase. Increasing N_1/N_2 means T_h-T_1 will decrease more slowly as I_2 is increased. This allows more power to be dissipated in the second stage, resulting in the increase in $I_{2,opt}$ and ΔT that has been shown. A similar effect would be observed by decreasing the aspect ratio of the first stage. The overall effect on cooling would be the same as increasing $N_1/N_{2,}$ but the current through the first stage would need to be adjusted to account for the revised size.

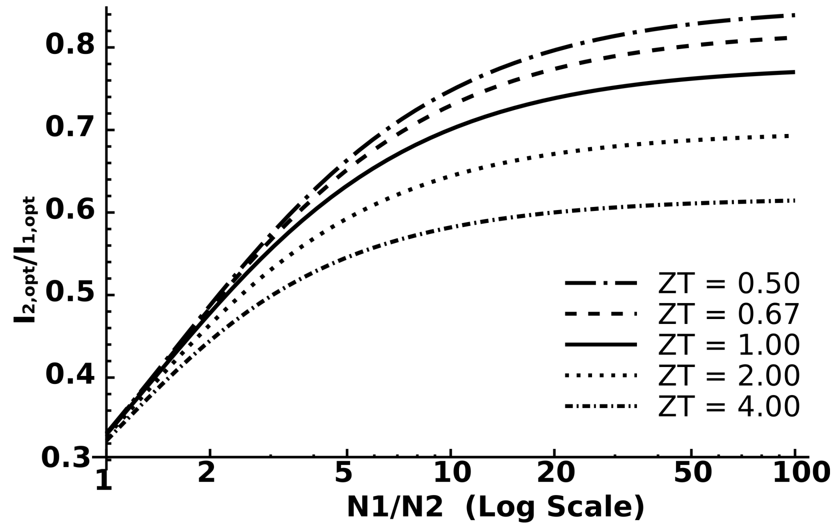


Figure 3.13: The effect of N_1/N_2 on $I_{2,opt}$ for a 2-stage cooler, when the currents are set by sequential optimization. The ratio of $I_{2,opt}:I_{1,opt}$ is plotted versus the N_1/N_2, showing that $I_{2,opt}$ remains lower than $I_{1,opt}$, even at high values of N_1/N_2.

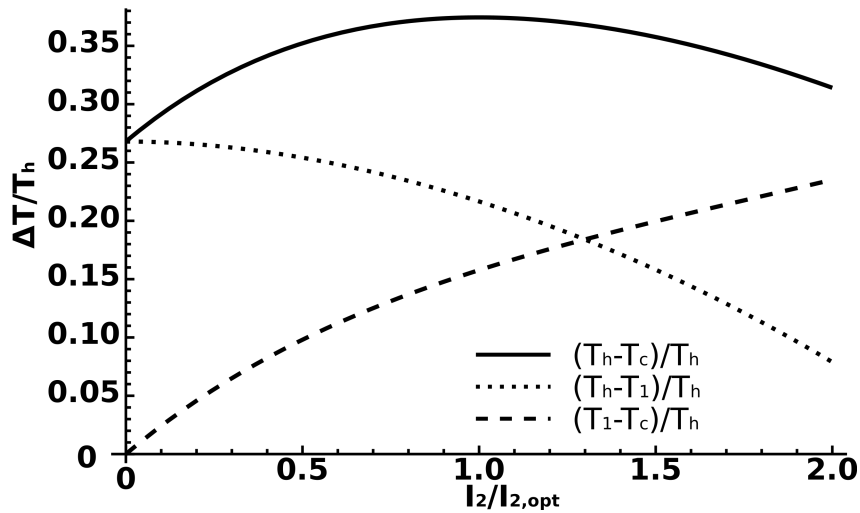


Figure 3.14: The effect of the second-stage current on the temperature across each stage of a 2-stage cooler, when the currents are set by sequential optimization. The normalized temperature difference is shown across the first-stage, the second stage, and both stages plotted versus I_2. ZT=1, and N_1/N_2=2.

3.3.3 Simultaneous Current Optimization (Variable Stage-1 Current)

The fact that the power dissipated by the second stage changes the temperature of the first stage means that the sequential current optimization in the previous analysis is

not the optimum means of setting the currents. The increase in T1 results in a slight increase in the Peltier coefficient (αT) at the inter-stage zone. This increases the optimum current for the first stage. In a practical application this would require an iterative approach to optimize the current, but in a modeled application, the entire system, including $I_{1,opt,}$ can be put in terms of I_2. Figure 3.15 shows how a device with ZT of 1 would behave versus I_2, while I_1 is set to the optimal current for any given value of I_2. Similarly, Figure 3.16 shows how the normalized minimum temperature varies with N_1/N_2 for a number of different ZT values. This scheme does provide an improvement in the normalized minimum temperature, but the gain is slight. For ZT=1 and $N_1/N_2 = 2$, it is 0.6250 compared to 0.6256 with the sequential current optimization scheme. With that in mind, the simplicity of the sequential current optimization would be appropriate for most practical applications.

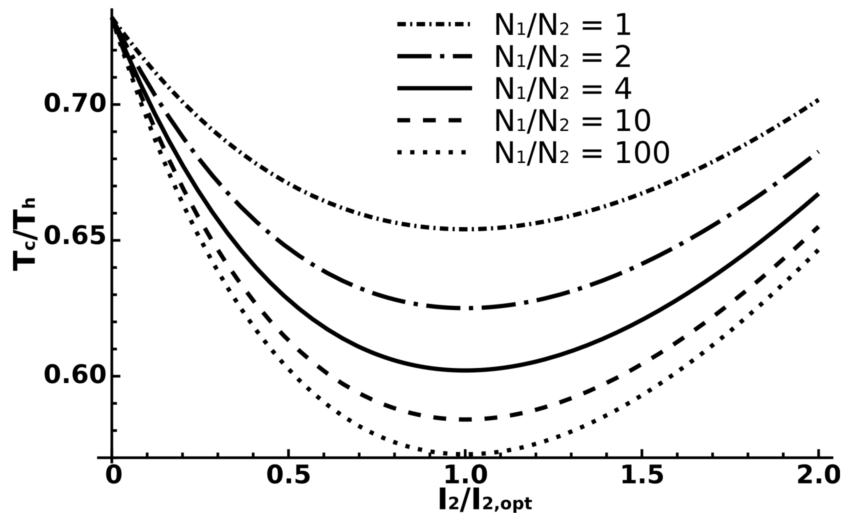


Figure 3.15:
normalized with respect to $I_{2,opt}$ for several different values of N_1/N_2 shown. All curves are for ZT=1. This method shows almost no gain in performance compared to the sequential optimization process used for Figure 3.11.

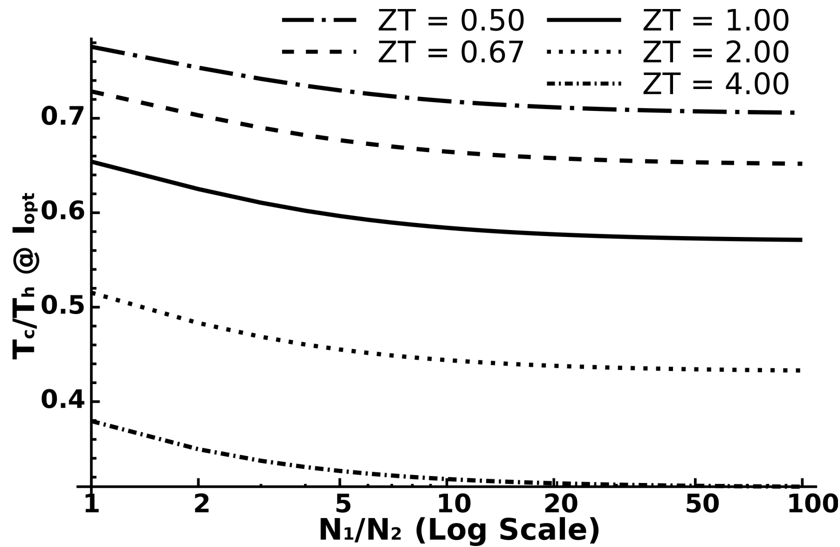


Figure 3.16: The effect of N_1/N_2 on the temperature of a 2-stage cooler, when the currents are set by simultaneous optimization. Normalized minimum is plotted versus N_1/N_2 at several values of ZT. This method shows almost no difference compared to the sequential optimization process used for Figure 3.12.

3.3.4 Single Current Optimization

A third method for setting the current in a multistage cooler is to supply the same current to both stages. This is the most straightforward method for practical applications because it requires only a single current supply. Figure 3.17 shows how the cooling varies versus input current at a number of different values of N_1/N_2. The calculated curves are for a device using materials with a ZT of 1. $N_1/N_2 = 1$ is not shown because it is equivalent to a single-stage cooler with an increased aspect ratio. Figure 3.18 shows how the normalized minimum temperature varies with the N_1/N_2 for several values of ZT. These curves are noticeably different from those in the previous two sections (Figure 3.11 and Figure 3.15) when N_1/N_2 is low, but converge to similar levels of normalized cooling as N_1/N_2 becomes large.

Figure 3.19 provides some explanation of this behavior. When N_1/N_2 is small, the optimum current of the 2-stage cooler is lower than the optimum current of a single stage cooler. This indicates that the temperature drop across the first stage is not as large as it would be in the simultaneously or sequentially optimized current input methods. As N_1/N_2 increases, the optimum current for the 2-stage system approaches that of the current in a single thermoelectric stage. Finally, it is important to note that power consumption does not scale directly with N_1/N_2. For large values of N_1/N_2 the power consumption scales linearly with N_1/N_2. However, for lower values, the power consumption is lower than would be expected with a linear relationship. This is because both the total number of thermocouples and the current through those thermocouples increase as N_1/N_2 increases. Figure 3.20 shows the power consumption of the cooler vs. the N_1/N_2.

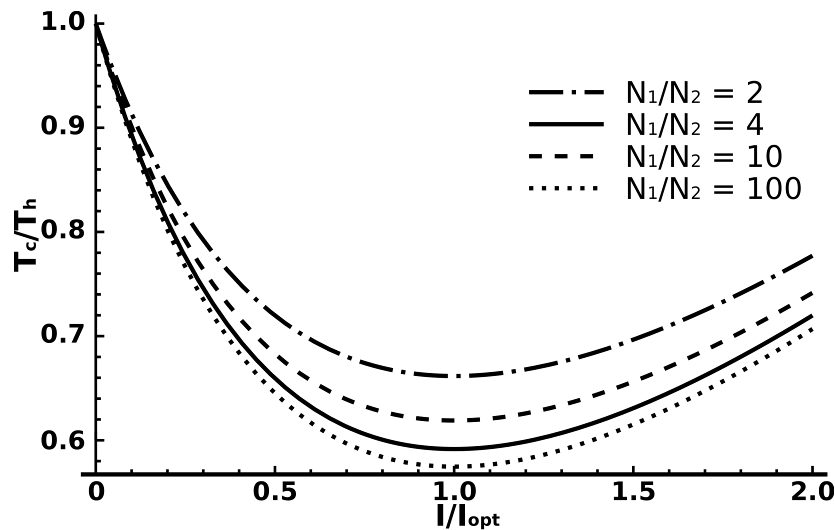


Figure 3.17: The effect of current on the temperature of a 2-stage cooler, when $I_1=I_2$. Normalized temperature is plotted versus I_2, normalized with respect to $I_{2,opt}$ for several different values of N_1/N_2. All curves are for ZT=1.

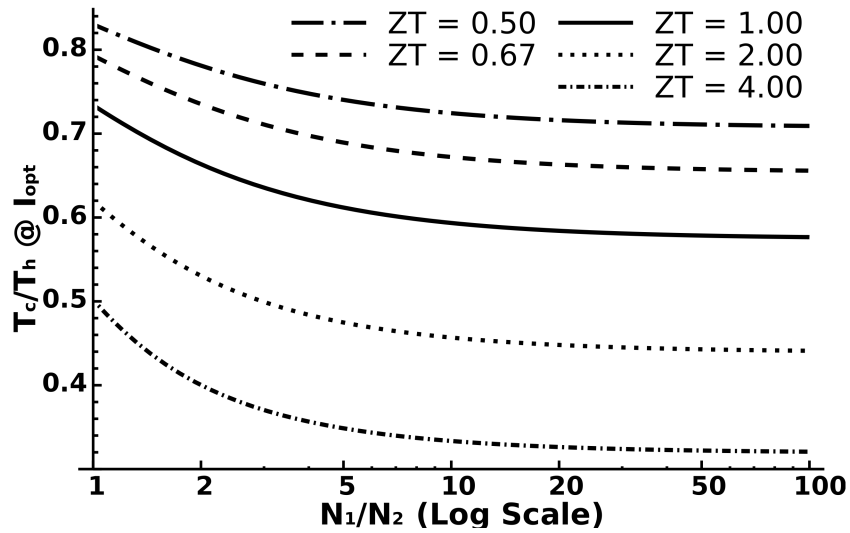


Figure 3.18: The effect of N_1/N_2 on the temperature of a 2-stage cooler, when $I_1=I_2$. Normalized minimum is plotted versus N_1/N_2 at several values of ZT.

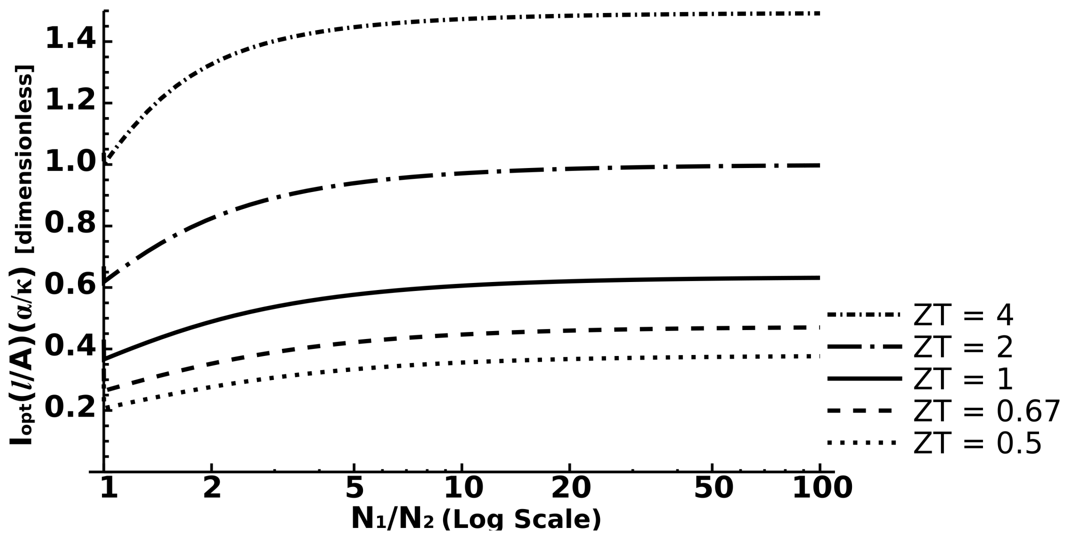


Figure 3.19: The effect of N_1/N_2 on the I_{opt} of a 2-stage cooler, when $I_1=I_2$. Optimum input current, normalized with respect to geometry and material properties, is plotted versus the N_1/N_2 for several values of ZT.

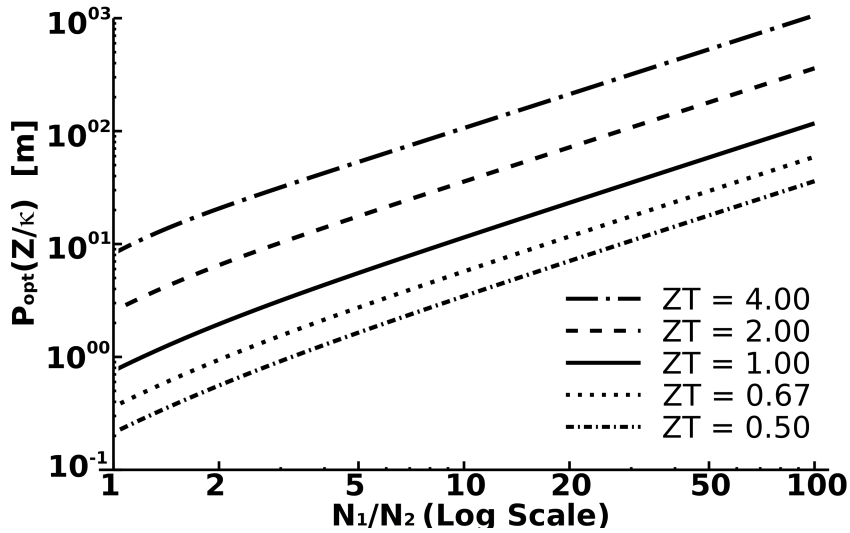


Figure 3.20: The effect of N_1/N_2 on the power consumption of a 2-stage cooler, when $I_1=I_2$. Power consumption at the optimum input current, normalized with respect to geometry and material properties, is plotted versus the N_1/N_2 for several values of ZT.

3.3.5 Parasitic Effects

Just as with the single stage devices discussed in the earlier part of this chapter, multi-stage devices are also affected by parasitic thermal conductance and parasitic electrical resistance. Figure 3.21 shows how K_{par} or R_{par} effect the performance of a 2-stage cooler with $N_1/N_2 = 2$, using the same current input for both stages. Both graphs show a strong similarity to their single-stage counterparts in Figure 3.9. Once again, little influence on performance is observed until either value is greater than 10% of the intrinsic value, and by the time the parasitic value per thermocouple is equal to the intrinsic value, the effective ZT has been reduced by a factor of 2.

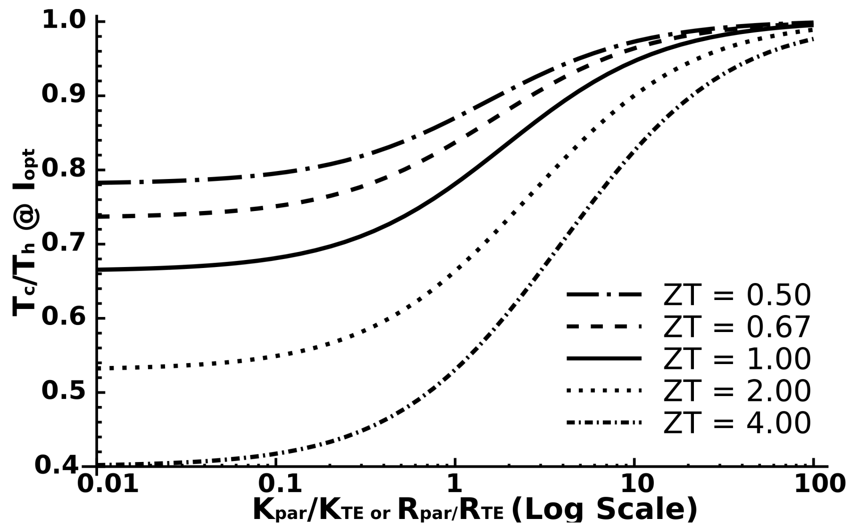


Figure 3.21: The effect of parasitic thermal conductance on cooling for a 2-stage cooler where $I_1=I_2$. The normalized minimum temperature is plotted versus the ratio between the parasitic thermal conductance per thermocouple and the thermal conductance of one thermocouple. $N_1/N_2=2$ for all the curves.

Figure 3.21 was presented to show that the same basic relationship to the parasitic thermal conductance and parasitic electrical resistance holds for both single stage and multistage coolers. But there is an additional parasitic effect that is unique to multistage coolers. In addition to the parasitic thermal conduction across individual stages, the thermal conduction between the two stages should also be considered. To investigate how inter-stage thermal conductance influences the performance of the cooler, a modification to the model is necessary. Instead of basing the model off of the schematic shown in Figure 3.10, the region between the stages must be split into two temperature nodes separated by a thermal conductance, as shown in Figure 3.22.

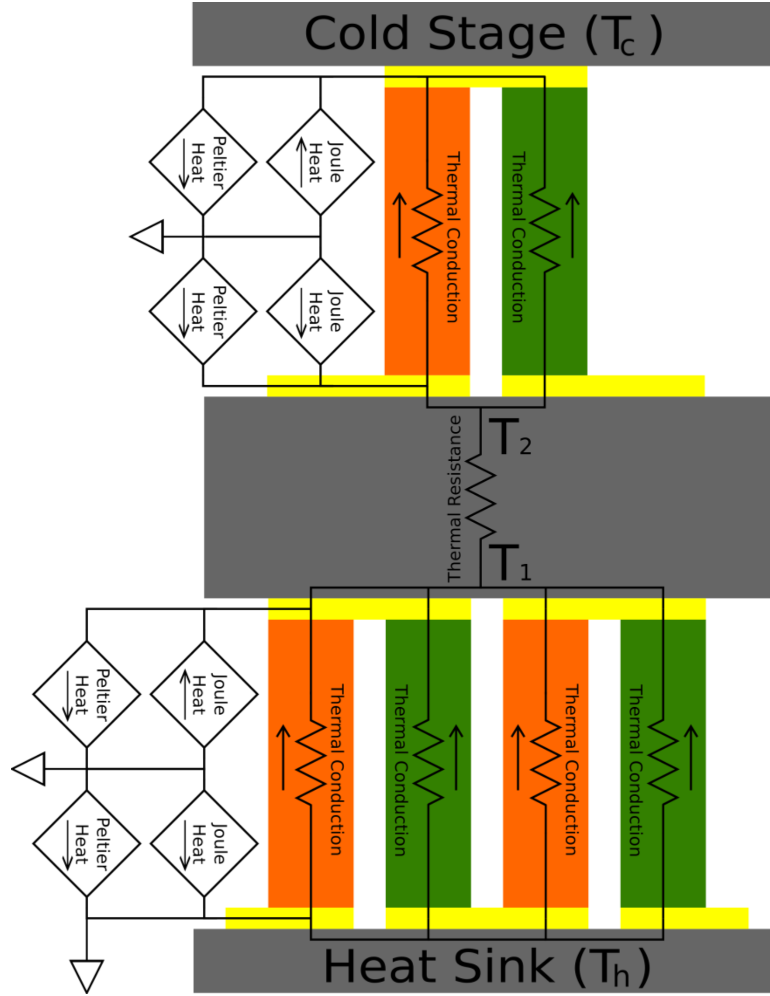


Figure 3.22: Heat flows in a 2-stage cooler, including an intermediate region with a finite thermal conductance.

Based on this schematic model, the equations that need to be solved to determine the temperature at each stage of a two stage cooler are:

$$N_1(\alpha_p - \alpha_n)T_1I_1 = (T_2 - T_1)K_{in} + (T_h - T_1)\left(N_1\left[\frac{\kappa_n A_{n1}}{l_{n1}} + \frac{\kappa_p A_{p1}}{l_{p1}}\right] + K_{par1}\right) + \frac{1}{2}I_2^2\left(N_1\left[\frac{\rho_n l_{n1}}{A_{n1}} + \frac{\rho_p l_{p1}}{A_{p1}}\right] + C_1R_{par1}\right) \qquad \text{Eqn. 3.28}$$

$$(T_2 - T_1)K_{in} + (T_2 - T_3)\left(N_2\left[\frac{\kappa_n A_{n2}}{l_{n2}} + \frac{\kappa_p A_{p2}}{l_{p2}}\right] + K_{par2}\right) = N_2(\alpha_p - \alpha_n)T_2I_2 + \frac{1}{2}I_2^2\left(N_2\left[\frac{\rho_n l_{n2}}{A_{n2}} + \frac{\rho_p l_{p2}}{A_{p2}}\right] + C_2R_{par2}\right) \qquad \text{Eqn. 3.29}$$

$$N_2(\alpha_p - \alpha_n)T_3I_2 = (T_2 - T_3)\left(N_2\left[\frac{\kappa_n A_{n2}}{l_{n2}} + \frac{\kappa_p A_{p2}}{l_{p2}}\right] + K_{par2}\right) + \frac{1}{2}I_2^2\left(N_2\left[\frac{\rho_n l_{n2}}{A_{n2}} + \frac{\rho_p l_{p2}}{A_{p2}}\right] + C_2R_{par2}\right) \qquad \text{Eqn. 3.30}$$

where the key difference is the inclusion of the K_{in} parameter, which represents the thermal conductance between the cold end of stage 1 and the hot end of stage 2. Figure 3.23 plots the normalized minimum temperature of a 2-stage cooler versus K_{in}. As expected, the plot indicates that the thermal conductance in the inter-stage zone should be as high as possible, and if it becomes too small, the performance of the cooler will be significantly impacted. In addition to the temperature, K_{in} also affects the optimum current. Figure 3.24 and Figure 3.25 illustrate how the optimum current changes with respect to K_{in} while maintaining fixed values for N_1/N_2 and ZT respectively. The figures show that as the K_{in} gets small, the optimum current point also decreases.

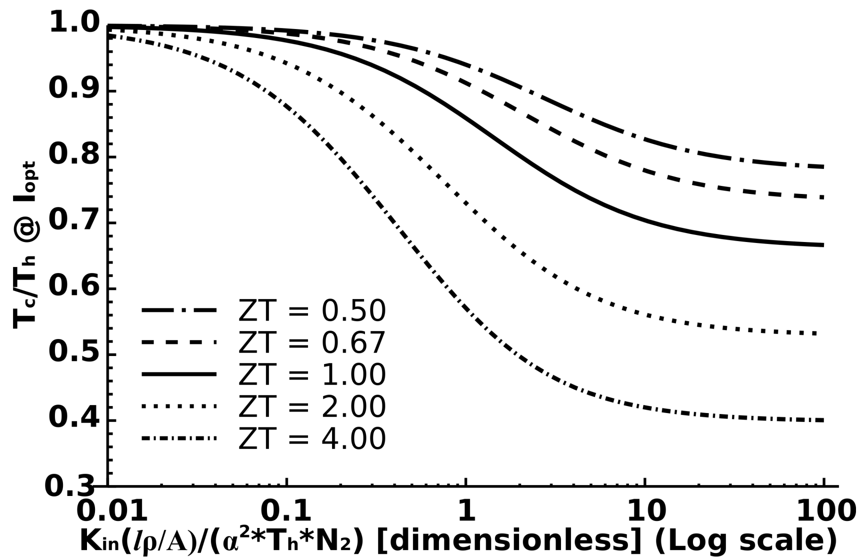


Figure 3.23: The effect of inter-stage thermal conductance on minimum temperature of a 2-stage cooler. Normalized minimum temperature is plotted versus K_{in} normalized over temperature, material properties, geometry and the number of TCs in the second stage. N_1/N_2=2 and a single drive current is used for both stages.

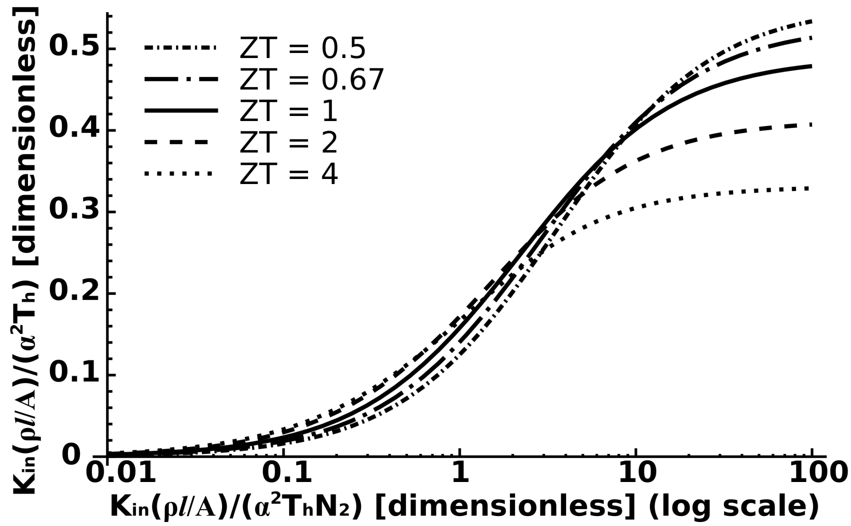


Figure 3.24: The effect of inter-stage thermal conductance on the optimum current of a 2-stage cooler for various values of ZT. Optimum current normalized with respect to temperature, geometry, and material properties versus K_{in} normalized over temperature, geometry, material properties, and the number of TCs in the second stage are shown. N_1/N_2=2 and a single drive current is used for both stages.

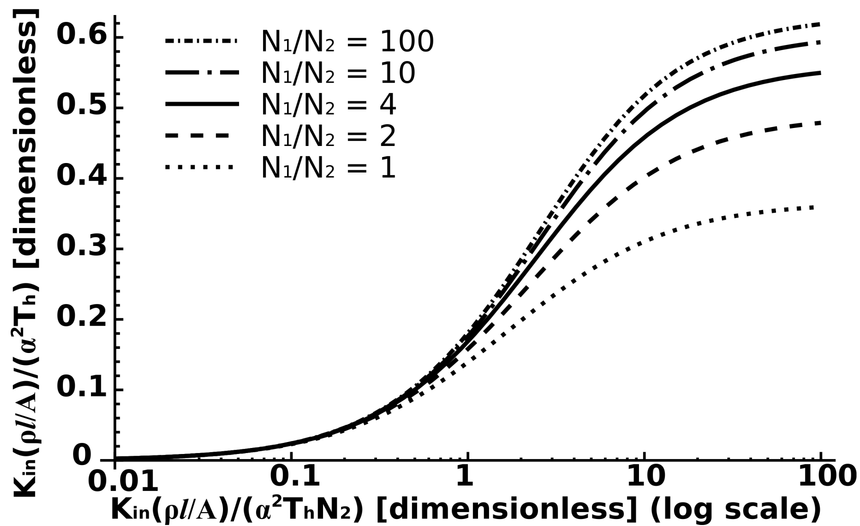


Figure 3.25: The effect of inter-stage thermal conductance on the optimum current of a 2-stage cooler for various values of N_1/N_2. Optimum current normalized with respect temperature, geometry, and material properties, versus K_{in} normalized over temperature, geometry, material properties, and the number of TCs in the second stage are shown. ZT=1 and a single drive current is used for both stages.

3.3.6 Ultimate Cooling Limits

The preceding sections have looked at the effect of several design parameters on the performance of a 2-stage cooler, but it is possible to build coolers with many more than 2 stages. The analysis in this section will examine the theoretical limits of multistage cooling. The resulting information is useful because it puts a bound on the performance of any potential multistage cooler, and shows how many stages might be useful in practical applications.

To begin the analysis, assume that all the dimensions of the thermocouples are held constant at all the stages, and the ratio of thermocouples at each consecutive stage is allowed to approach infinity such that

$$N_x / N_{x+1} \rightarrow \infty \qquad \text{Eqn. } 3.31$$

where *x* is the number of the stage. Under this circumstance, the temperature difference across each stage will approach its theoretical maximum value as given earlier in Eqn. 3.13. Figure 3.26 shows the minimum achievable temperature versus the number of stages for several values of ZT. The figure shows that the most significant gains to cooling are achieved in the first 4 stages, with modest gains through the 7th stage, and very limited added performance for each stage beyond that. Adding stages also adds power consumption. If $N_1/N_2 = 2$ each additional stage will approximately double the number of thermocouples necessary, leading to increased power dissipation through joule heating.

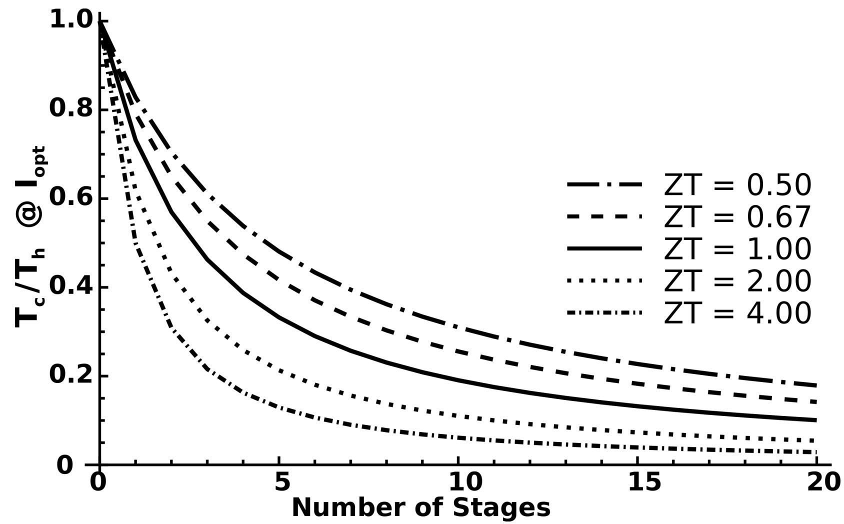


Figure 3.26: The effect of the number of stages on the normalized minimum temperature for several values of ZT. The ratio N_1/N_2 is set to infinity.

In summary, the one-dimensional model discussed above can be used to understand many of the general features that should be present in the design of any thermoelectric micro cooler. It should have up to 7 stages, and each higher stage should have fewer thermocouples that the stage below it. Parasitic thermal conductance should be limited to the same order of magnitude as the thermal conductance of the TE at each stage, and the thermal conductance between the stages should be as high as possible.

3.4 Prediction and Comparison to 2-D FEM

The previous sections have discussed how single and 2-stage coolers behave in a general way, but they have not made predictions about how a specific cooler will behave. This section will model the performance of simple single-stage and multi-stage coolers with specific material properties and geometries. It will then compare the results of these simple, 1-D models to predictions from a 2-D FEM model. The comparison is made to

show that the 1-D model is effective at predicting device performance, even though it does not take lateral heat transfer into account.

3.4.1 1-D Model Performance Predictions

Predicting device performance using a 1-D model is primarily a matter of solving the heat balance equations for T_c, while using real world parameters for the geometry, material properties and parasitic effects. However, the n-type and p-type materials used in a real cooler usually do not have the same material properties. Similarly, the size of the thermocouples may be different between the n-type and p-type legs of the thermocouple. Finally there may be additional sources of heat transfer present. These include radiation and parasitic thermal conductances that only affect a specific stage. When all these factors are taken into account, the resulting analytical solution becomes large, complex, and uninsightful. Instead it is helpful to use numerical tools that provide results on a case-by-case basis.

To perform the numerical analysis, two tools have been employed. A Matlab based program was developed by Gi-Suk Hwang [75]. This program represents the system of equations in matrix form, and it uses an iterative solver to find the solution. Using iteration instead of a direct solution of the matrix allows for inclusion of temperature dependent material properties and the radiation heat transfer, which is non-linear with respect to temperature. Additionally, a Mathematica based program was created. This program was based around an analytical representation of the heat balance equations, and a solution was found using the internal Mathematica solver. Both methods yielded numerically identical results, but had different strengths. The Matlab based model was more complete, with its inclusion of support for temperature dependent

material properties. The Mathematica based model was more flexible in its ability to quickly analyze the affect of an arbitrary independent variable on cooler performance, with little modification to the underlying code. Therefore, the 1-D simulation results presented in this chapter and in the following chapters are produced by the Mathematica-based model unless specifically indicated.

To get some indication of how a cooler made with thin films deposited at the University of Michigan will behave, the optimal material properties reported in Chapter 2 are used in the model. No parasitic thermal conductance is included at this time (the effects of parasitic thermal conduction on specific designs will be addressed on Chapter 4 and Chapter 5), but a parasitic electrical contact resistance is included, with a resistivity of 1.45E-8 $\Omega\text{-m}^2$. This value was chosen because it is high enough to affect the performance of the cooler, and will result in a difference to the FEM model if it is not simulated correctly. A 1-stage, 3-stage and 6-stage cooler have been modeled. A single current level is used to power all the stages of an individual cooler, and the geometry of the three different coolers is summarized in Table 3.1. Figure 3.27 shows the performance of the three coolers with respect to input current. The 1-stage cooler reaches a minimum temperature of 275.7 K at an input current of 19.1 mA. The 3-stage reaches 261.6 K at a current of 10.4 mA, and the 6-stage cooler reaches 243.5 K at a current of 7.9 mA.

Table 3.1: Material Properties and Dimensions used for simulation shown in Figure 3.27.

TE Material Properties				TE Dimensions (μm)		
	κ (W/m-K)	ρ (μΩ–m)	α (μV/K)	Length	Width	Thickness
Bi_2Te_3	1.5	19	208	30	100	4
Sb_2Te_3	1.5	13	160	30	100	4

# of TCs per Stage					
Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6
64	32	16	8	4	2

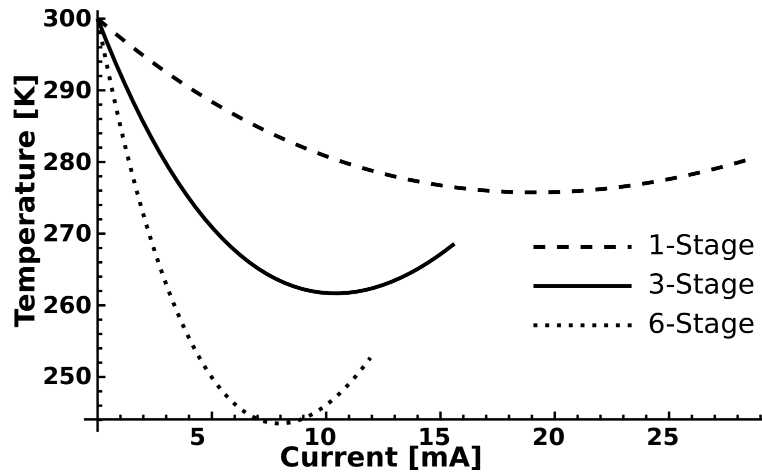


Figure 3.27: Simulated performance of 1-stage, 3-stage, and 6-stage coolers vs. current input. The coolers use the material properties for the Michigan TE thin films, and a contact resistivity of 1.45 μΩ-m.

3.4.2 2-D FEM Model

A 2-D FEM model has also been constructed. It was developed in Ansys using a thermal-electric coupled field element. The models were constructed using the Ansys scripting language. The parasitic electrical resistance was modeled through contact elements at the interface between the thermoelectric materials and the electrical conductors that connect them to each other. The electrical conductors were given conductivity of 30 nΩ-m, and the inter-stage material was set to a thermal conductance of 140 W/m-K to match the nominal values for gold and silicon respectively. The minimum

temperature and optimum current were found by repeatedly modeling the devices with slowly increasing current levels. The 1-stage cooler was modeled to reach a minimum temperature of 275.8 K at an input current of 19 mA, with the current incremented in 1mA steps. The 3-stage reached 262.2 K at a current of 10.5 mA, with the current incremented in 0.5 mA steps. The the 6-stage cooler reached 242.1 K at a current of 8.2 mA in 0.2 mA steps. Figure 3.28 shows the temperature profile of the 6-stage cooler.

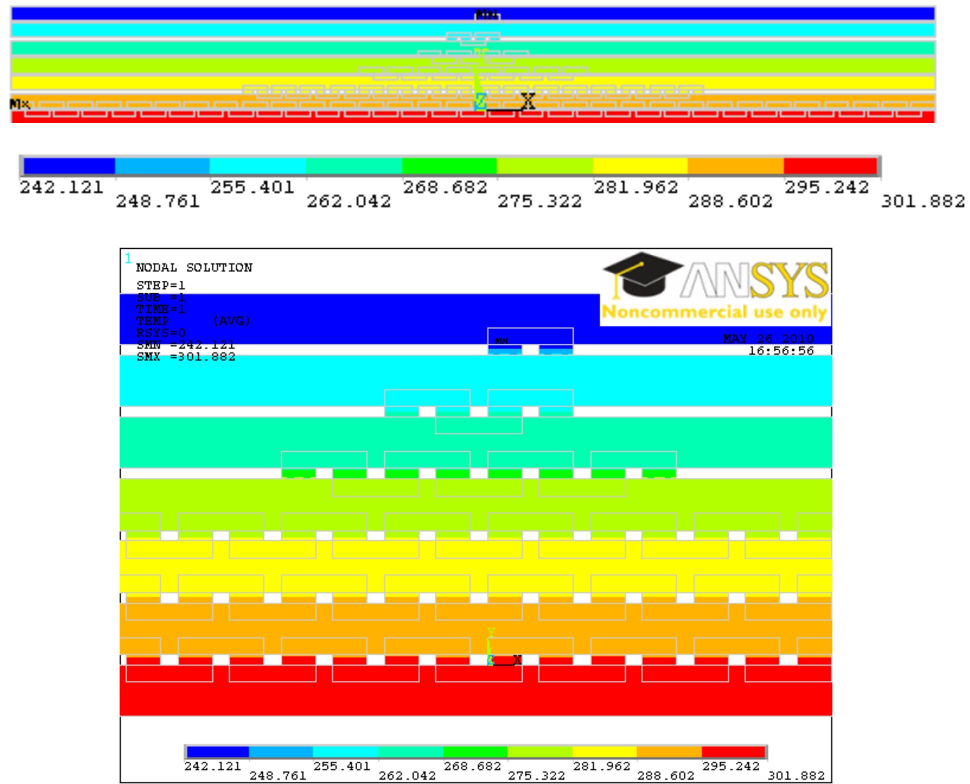


Figure 3.28: (a) The FEM model of a 6-stage cooler showing the temperature profile at the optimum current input. (b) The same model, enlarged to make the TCs more visible.

While the FEM model is useful in visualizing the temperature profile, it is also slow to execute and cumbersome to edit. The 1-D Mathematic model matches the 2-D FEM model to within 1.5 K, and is comparatively easy to modify and faster to execute.

This allows for quick tests of designs with widely varying parameters, and most of the analysis in the following chapters will be performed using the 1-D Mathematic model.

3.5 Summary

A 1-D analytical model has been implemented and found to compare well with results from 2-D FEM analysis. The model was used to understand the basic properties that need to be considered when designing a thermoelectric cooler to address the needs of low power MEMS and electronics applications. It was shown that using thermoelectric elements with large aspect ratios allows for low power operation, and fabrication at the micro-scale allows for very large aspect ratios to be achieved. Second, it was shown that adding stages to the cooler helps to achieve lower temperatures than are allowed by a single TC. However, there are diminishing returns, and using more than 7-stages would not be helpful. This is especially true for low power applications because power consumption increases quickly as additional stages are added, even though there is little gain in temperature performance. The model was also used to show how parasitic thermal conductances and electrical resistance can affect a generic cooler. On a per thermocouple basis, both should be kept below 10% of the values for TE materials in the thermocouple. They will significantly impact performance if they grow to the same order of magnitude as the thermocouple itself. Additionally, the thermal conductance of the region between the thermoelectric stages needs to be maximized, or the overall performance of the cooler will also suffer. Finally, the means of selecting the current in the various stages can be set in a number of ways. The simultaneous optimization concept provides the best performance, however the sequential optimization process is nearly as effective and simpler to implement. Using the same current level in all the stages of an ideal cooler

results in a device that does not perform as well as setting the current in each stage independently. However, in a real world design, it is important to consider whether the different drive schemes introduce additional parasitic effects. This will be explored further in later chapters as specific designs are analyzed.

Chapter 4
Fabrication Process 1 – Designs and Results

4.1 Introduction

Chapter 2 made the argument for the choice of Bismuth Telluride and Antimony Telluride thin films as the thermoelectric materials for this project, and Chapter 3 discussed the basic properties that a micro thermoelectric cooler must possess in order to be both effective and low power. This chapter will present a structure that implements the design criteria for a successful cooler presented in Chapter 3, and a process for fabricating that structure at the wafer level that is compatible with the thermoelectric materials. The chapter will then review several iterations of cooler designs based on this fabrication process and structure. Both simulated and measured data for each iteration will be presented.

4.2 Structure

4.2.1 Planar vs. Vertical Design

Many TE coolers at both the macro and micro scale use a vertical design as shown in Figure 4.1 (a), and there are two practical reasons for this. First, using a vertical design allows the TE materials to act as the structural support of the cold stage. This means that there will be no additional materials spanning between the interstage regions, making the parasitic thermal conductance very low, especially in vacuum. As discussed in Chapter 3, minimizing K_{par} is an important part of realizing the full potential

of the materials being used, and the ability to have low values of K_{par} is an important design advantage. The second advantage of the vertical design is that it can be used to implement multi-stage coolers in a spatially compact way. For example, if $N_1/N_2 = 2$, adding an additional stage to the hot side requires doubling the area occupied by TCs, but this only requires increasing the linear dimension by $\sqrt{2}$. However, vertical coolers do require a trade off in power for performance when implemented at the microscale. When using films with thickness less than 5 μm, the aspect ratio of the TCs can become very low, leading to high power consumption at the maximum ΔT. The example in the previous chapter was for a thermocouple with elements that were 20 μm on a side, giving it an aspect ratio of 50,000 m^{-1}. However, given the process capabilities of the deposition system at the University of Michigan, aspect ratios this high are not yet possible with a vertical process. Film thicknesses are limited to less than 7 μm, and reliable minimum dimensions are around 50 μm, giving a maximum possible aspect ratio of only 2,800 m^{-1}. While still much higher than the ratios of macro-scale coolers, this is still not high enough to realize power consumption levels below 100 mW.

This problem can be overcome by changing the orientation of the TC. Instead of moving heat vertically, the TC can be fabricated on its side, to move heat parallel to the plane of the cold platform, as shown in Figure 4.1 (b). Instead of the cold platform being located at the top of a pyramid, it is now at the center of a series of concentric rings, where each ring is a stage. This allows the aspect ratio to be lithographically defined over a very wide range of values. The final example of aspect ratio discussed in Chapter 3 applies to the case of a planar aligned TC:

> … if the thermocouple is constructed out of a thin film with a thickness of 2 μm, and has a width of 100 μm and a length of 60 μm, the aspect ratio would be 300,000 m^{-1}.

The planar TC, in this case, has an aspect ratio more than 100x larger than a vertical TC fabricated with the same materials, and would therefore have a power consumption that is 100x smaller than the vertical TC. Using a planar process also means that high aspect ratios can be implemented without pushing the limits of film thickness or the minimum dimensions of the pattern.

The planar design too has its weaknesses. The primary drawback is the TE thin films are fragile, and cannot withstand large amounts of shear or tensile stress. This means that they will not be able to provide adequate support to the cold stage on their own. Other materials must be present that span the stages and provide a structural support for both the TCs themselves, the inter-stage zones, and the cold platform. The result is that the parasitic thermal conductance, K_{par} will necessarily be higher in a planar thermoelectric micro cooler, and this will result in reduced temperature performance. The second drawback of the planar design is that the number of thermocouples is limited by the perimeter of a given stage, not the area. To add a stage to the hot side of a planar cooler requires doubling the perimeter of the device, resulting in a large increase in the overall size of the device.

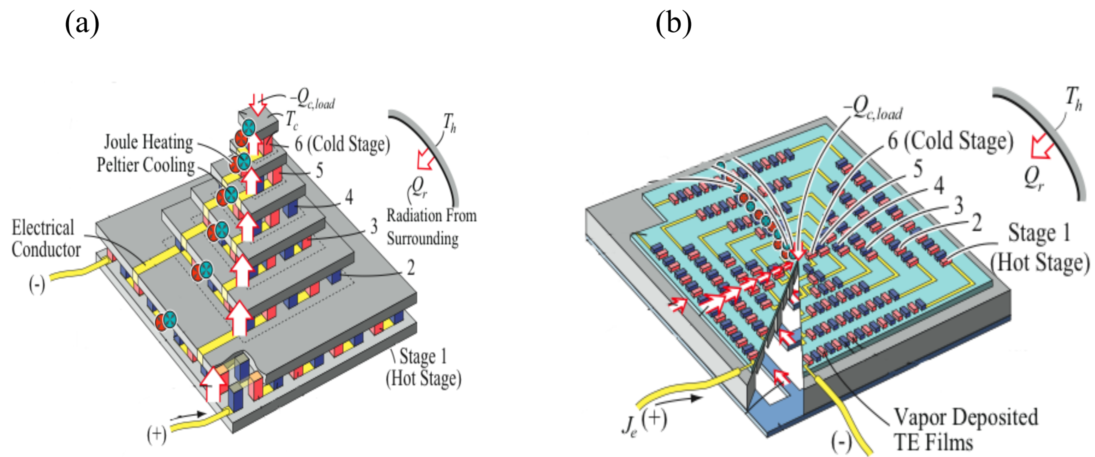


Figure 4.1: Pyramid to planar design transitions. (a) A traditonal pyramid structure. (b) A planar cooler with the stages arranged as concentric rings.

Despite these drawbacks, the planar design is the preferable choice because it is the only way to achieve low power cooling with thin film materials on the micro scale. However, the supporting material in any planar design must be carefully selected to minimize parasitic thermal conductance. In addition, limiting factor for the number of stages in some applications will be size restriction. The following sections will describe a planar MEMS structure that fulfills the requirements for multistage cooling, followed by a description of the process used to fabricate the device.

4.2.2 Satisfying Thermal Requirements

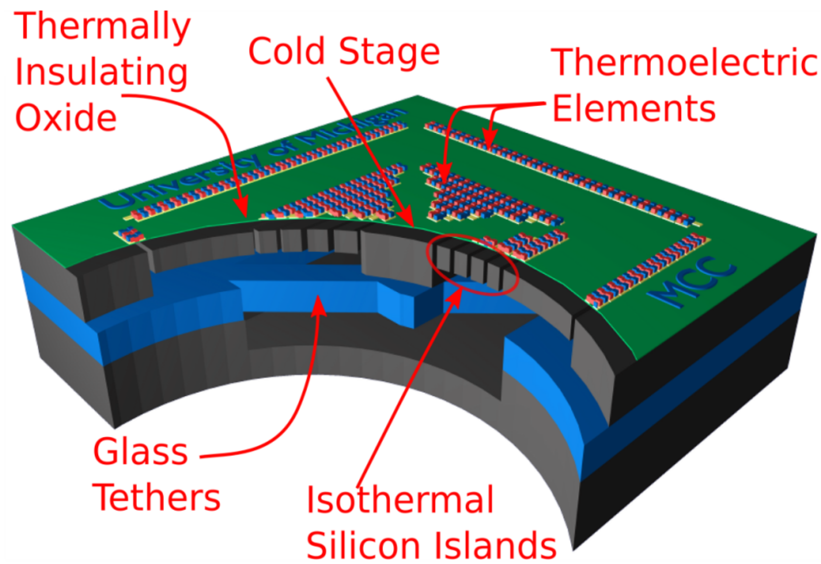


Figure 4.2: A 3-D rendering of a planar multi-stage thermoelectric cooler. The coldest region is in the center of the device, with 6-stages arranged in concentric squares around the center. A glass tether supports the cold region to give some additional strength to the system.

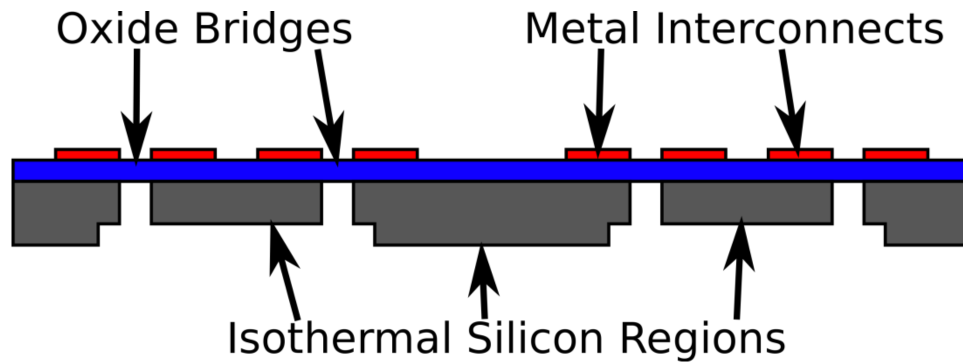


Figure 4.3: Cross sectional diagram of the thermal substrate for a 2-stage cooler built using the silicon-glass-silicon process.

Previous investigations of microcoolers have demonstrated in-plane designs [40],[72],[42],[41], by simply depositing the thermoelectric material on to a dielectric membrane. However these devices were all single stage. Creating a multistage cooler requires a more complex structure because of the need for thermally conducting, inter-

stage zones. The structure of any multistage thermoelectric cooler should satisfy the thermal requirements specified in the previous chapter:

1) Low parasitic thermal conductance across stages.
2) Higher thermal conductance between stages.
3) Increasing capacity with decreasing stage number.

To that end, the structure shown in Figure 4.2 was proposed. The top layer of the structure is shown in Figure 4.3 and implements that necessary thermal properties. The other layer of the device are for structural purposes and are described in a later section. In this case, the regions of high thermal conductance that are necessary between the stages are made from single crystal silicon, which has a thermal conductivity of 140 W/m-K. Each silicon region is a ring, with the coldest region in the center of the rings. The hottest region is assumed to be attached to a heat sink at the ambient temperature. Ideally, the only material to span between the silicon rings would be the TE materials; however, this is not possible because the TE materials have poor structural properties. Instead, the gap between adjacent silicon rings is spanned by silicon dioxide. Silicon dioxide was chosen for several reasons. First, it has a low thermal conductivity, approximately 1.1 W/m-K [92], which is close to the thermal conductivity of the TE materials being used for this project. Second, silicon dioxide is a standard material in microfabrication processes with well-characterized deposition techniques. Finally silicon dioxide has a very high thermal tolerance. It will not break down at the deposition temperature of the TE materials.

Other materials were considered, but rejected. Silicon nitride has been used as a structural material in MEMS devices in the past, but has a thermal conductivity several

times larger than that of silicon, in the range of 5-10 W/m-K. Polymers such as Parylene and polyimide were also considered because they have thermal conductivities as low as 0.12 W/m-K [93],[94]. However this was not a viable option because such polymers will not withstand the high temperatures necessary during deposition of the TE materials. Polymers are also generally not compatible with vacuum packaging, as they tend to outgas significantly.

FEM modeling of the structure, as shown in Figure 4.4, verifies that temperature gradients across the structure will be concentrated in the oxide bridges, and that the silicon rings will be effectively isothermal. The model was constructed as a 2-D cross-section, and a simulation was performed using axis symmetric boundary conditions. A 10 K temperature differential was applied across the model. Heat was only allowed to the leave the system though the central nodes of the structure, and this heat flux was set as output of the simulation. Using that information, the total thermal impedance of the structure was calculated, and that information will be used later to analyze specific designs.

A more precise measure of whether the second condition is met can be obtained by applying the generalized analysis of a two stage cooler in the previous chapter to the specific device being developed here. To see if the silicon is thermally conductive enough so as not to interfere with the cooler's ability to cool, the normalized relationships shown in Figure 3.25 can be used. It is clear from that chart that when the dimensionless quantity $K_{in}A_{ration}\rho/(\alpha^2 T_h N_2)$ is greater than 10, there is little impact on the temperature performance of the cooler. For the ring-shaped structure shown in Figure 4.4, with a first-stage radius of 6 mm, a second stage radius of 4.5 mm, a silicon thickness of 100um,

32 TCs at stage 1, 16 TCs on stage 2, ρ=12.9 mW-m, α=210 mV/K and a TC aspect ratio of 150000, the quantity $K_{in}A_{ration}\rho/(\alpha^2T_hN_2)$=1206. That is 2 orders of magnitude higher that the level necessary to avoid an degradation of performance.

The effect of oxide bridges on the potential performance of a cooler is more straight forward. The oxide creates a parasitic thermal conductance, K_{par}, which lowers the effective Z of the thermoelectric materials at each stage, as discussed in Chapter 3. For a stage where the oxide bridge is 30 μm with a radius of 6.2 mm, and an oxide thickness of 1.5 μm, K_{par} would be 8.84e-4 W/K. How much this affects the overall performance depends on the total thermal conductance of the TC that covers the oxide, and will be very design dependent. The extent to which the effectiveness is lowered is highly dependent on the design of the device and will vary by stage; it will be analyzed in greater depth when discussing individual cooler layouts.

To satisfy requirement three (increasing capacity), the perimeter of each stage must be large enough to accommodate more thermocouples than the adjacent colder stage. This brings to light one of the disadvantages of moving to a planar configuration for multi-stage coolers. With vertical coolers, doubling the number of TCs requires doubling the area of the stage. However, with a planar cooler, doubling the number of TCs requires doubling the perimeter of the stage, and results in a quadrupling of the area of the cooler for all outlines that only have convex corners (i.e. squares, circles, hexagons; star patterns can increase the perimeter with less impact on surface area).

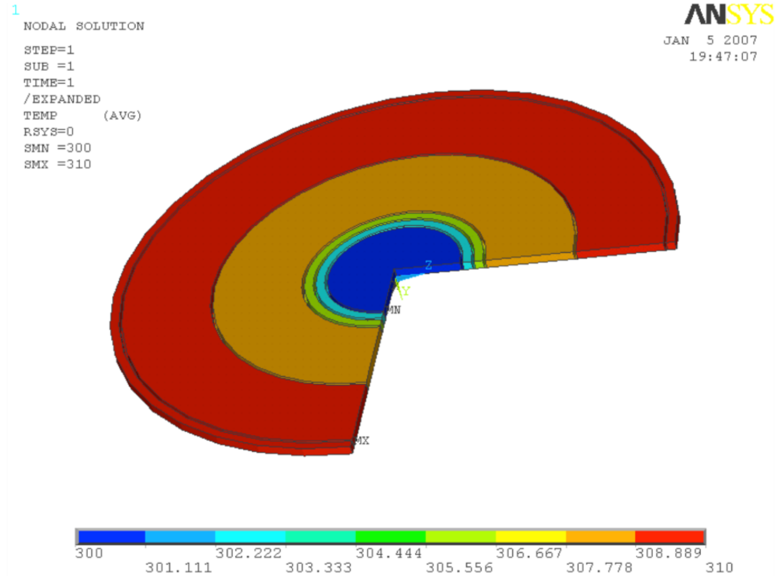


Figure 4.4: 2-D axis symmetric FEM simulation of a 5-stage cooler structure used to calculate the thermal isolation of the Thermal Substrate.

4.2.3 Satisfying Structural Requirements

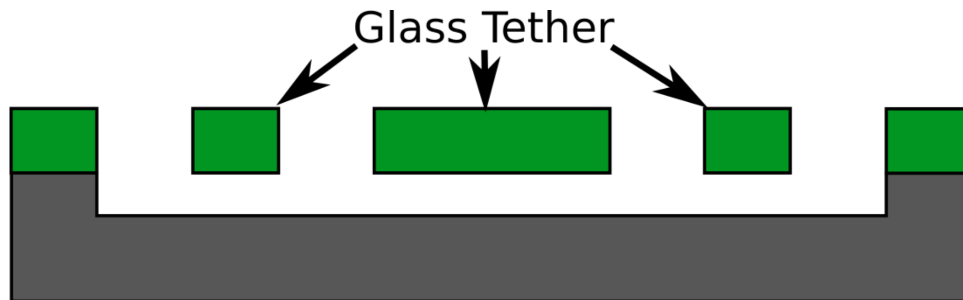


Figure 4.5: Cross-section of the Structural Substrate used in the Silicon-Glass-Silicon process.

The structure described in the previous section satisfies all the thermal requirements needed to implement a multistage cooler; however, it is also very fragile. The silicon rings are supported only by the thin SiO_2 bridges, and these bridges provide little mechanical stability or robustness. Glass tethers have previously been developed by