ABSTRACT

Title of Dissertation: On-Chip Thermoelectric Cooling
           of Semiconductor Hot Spot

           Peng Wang, Doctor of Philosophy, 2007

Directed By: Professor Avram Bar-Cohen
           Assistant Professor Bao Yang
           Department of Mechanical Engineering

   The Moore's Law progression in semiconductor technology, including shrinking feature size, increasing transistor density, and faster circuit speeds, is leading to increasing total power dissipations and heat fluxes on silicon chip. Moreover, in recent years, increasing performance has resulted in greater non-uniformity of on-chip power dissipation, creating microscale hot spots that can significantly degrade the processor performance and reliability. Application of conventional thermal packaging technology, developed to provide uniform chip cooling, to such chip designs results in lower allowable chip power dissipation or overcooling of large areas of the chip. Consequently, novel thermoelectric cooler (TEC) has been proposed recently for on-chip hot spot cooling because of its unique ability to selectively cool down the localized microscale hot spot.

In this dissertation the potential application of thermoelectric coolers to suppress on-chip hotspots is explored using analytical modeling, numerical simulation, and experimental techniques. Single-crystal silicon is proposed as a potential thermoelectric material due to its high Seebeck coefficient and its thermoelectric cooling performance is investigated using device-level analytical modeling. Integrated on silicon chip as an integral, on-chip thermoelectric cooler, silicon microcooler can effectively reduce the hotspot temperature and its effectiveness is investigated using analytical modeling and numerical simulation, and found to be dependent of doping concentration in silicon, electric contact resistance, hotspot size, hotspot heat flux, die thickness and microcooler size. The other novel on-chip hotspot cooling solution developed in this dissertation is to use a mini-contact enhanced TEC, where the mini-contact pad connects the silicon chip and the TEC to concentrate the thermoelectric cooling power onto a spot of top surface of the silicon chip and therefore significantly improve the hotspot cooling performance. Numerical simulation shows hotspot cooling is determined by thermal contact resistance, thermoelectric element thickness, chip thickness, etc. Package-level experiment demonstrates that spot cooling performance of such mini-contact enhanced TEC can be improved by about 100%.

ON-CHIP THERMOELECTRIC COOLING OF SEMICONDUCTOR HOT SPOT

By

Peng Wang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Committee:
Professor Avram Bar-Cohen, Chair
Assistant Professor Bao Yang, Co-Chair
Associate Professor Patrick
Professor Michael Ohadi
Professor Lourdes Salamanca-Riba

Dedication

To my wife, Yi Jin
and my daughter, Grace Wang
for their unconditional love and support

Acknowledgements

I would first like to thank my advisor Professor Avram Bar-Cohen for his mentoring and continuous support. He gave me the greatest opportunity to begin my Ph.D. study in thermal management of electronic packaging and introduced me to this very exciting on-chip hot spot cooling research. It was his wonderful supervision that made a profound impact on my life and future career development. I really enjoyed invaluable guidance, suggestion and inspiration he provided me. I would like to gratefully acknowledge my co-advisor Professor Bao Yang for his solid knowledge and advice in thermal physics and thermoelecctrics. Without their assistances for guiding the research directions, this dissertation could not be accomplished.

I would like to thank my dissertation committee members, Professor Patrick McCluskey, Professor Michael Ohadi and Professor Lourdes Salamanca-Riba for their valuable comments. I would also like to thank Professor Ali Shakouri and Dr. Yan Zhang at the University of California at Santa Cruz, Professor Gary Solbrekken at University of Missouri at Columbia and Dr. David Chau at Intel for their technical discussions, and Dae Whan Kim, KyongJoon Kim, Raj Bahadur, Chris Bachmann, Jeff Khan, Viatcheslav Litvinovitch, Emil Rahim and Dr. Ilai Sher for their friendship during my Ph.D. study. I am indebted to Intel for funding this three-year project.

Most importantly, I would like to thank my parents and sisters for their continuous support, understanding and patience during my graduate study in the U.S. I am especially grateful to my wife, Yi Jin, for her love, company and encouragements throughout these years and sharing all of the pressures and joys with me.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Overview and Motivation

The ongoing Moore's law progression in semiconductor technology continues to lead to shrinking feature size, increasing transistor density, faster circuit speeds, and higher chip performance. As shown in Figure 1.1, the International Technology Roadmap for Semiconductors (ITRS) [1], one of the major "roadmaps" for the semiconductor industry, predicts a continuous decrease in transistor size to 10 nm along with a rise in transistor density towards 10 billion transistors/cm^2 by 2018, while chip size remains almost constant at 260 mm^2 during the next two decades.
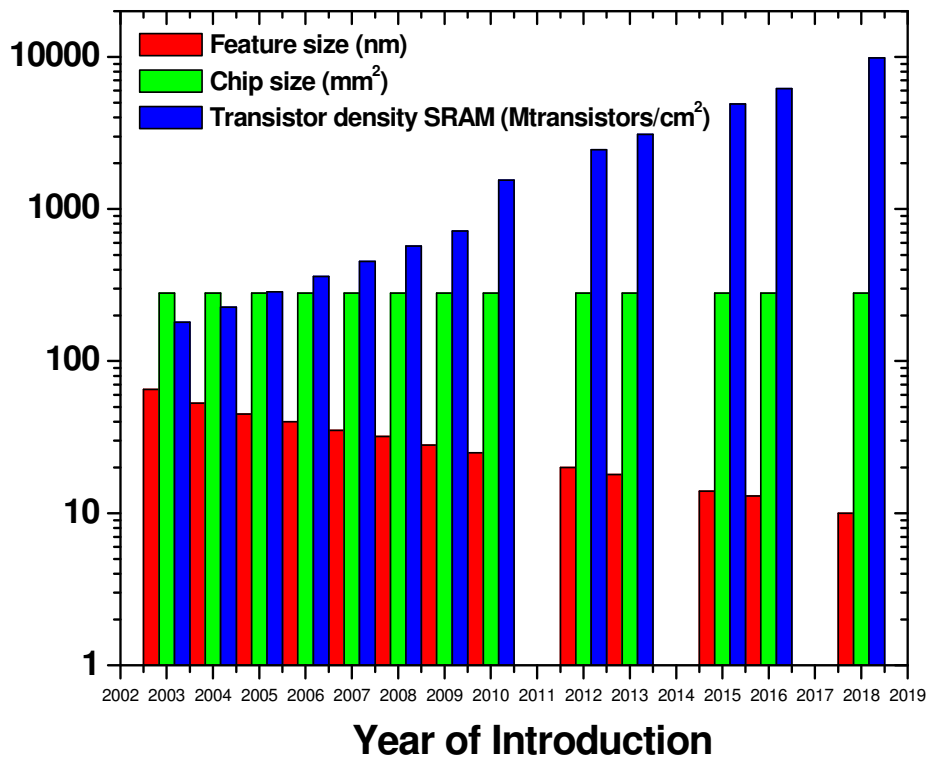
Figure 1.1: The 2005 ITRS predictions of feature size, chip size and transistor density for high performance microprocessor chips [1].

These changes in semiconductor technology can be expected to lead to ever faster and more computationally-complex chips. However, the consequence of increasing microprocessor performance is an increase in power dissipation because the chip power dissipation is governed by the following equation [2]:

$$P \approx NCV^2 f \qquad (1.1)$$

where N is the number of transistors per chip, C is the input capacity, V is the peak-to-peak voltage of the signal, and f is the operating frequency. During the last decade, extensive efforts have been made to reduce both the capacitance and the operating voltage on the chip. However, the chip power dissipation is still dramatically increasing every year due to increasing integration levels and increasing device speeds.

As is well known overheating of the chip is one of the major root causes of electronic device failures, due to both accelerated failure rates and reduced performance. For many semiconductor technologies, the reliability of individual transistors is exponentially dependent on the operating temperature, according to Black's equation [3]:

$$MTF = AJ^2 \exp(\frac{E_A}{k_B T}) \qquad (1.2)$$

where MTF is the mean time to failure, A is a constant, J is the current density, E_A is the active energy where the value for typical silicon failures is approximately 0.68eV, k_B is the Boltzmann constant and T is the absolute operating temperature. While small changes in the activation energy can lead to very large changes in the failure rate, at E_A values close to 0.7eV, a modest increase in operating temperature, in the range of

$10 \sim 15^{o}C$, could double the device failure rate [4]. In addition, increased operating temperature decreases transistor switching speed, due to increased gate delay, and thereby directly reduces microprocessor performance.

In recent years there is significant interest in cooling microprocessors [5-8] which is motivated by three respects [9]: (1) The drive to improve speed motivates circuit designers to compress the 'core' of the microprocessor, which contains the region of the most intense electrical activity, to ever smaller size. Along with the reduced "time-of-flight" between transistors, this spatial compression leads to high heat flux in the "core" areas of the silicon chip; (2) The temperature rise in the interconnects between transistors is growing fast in today's IC technology, owing primarily to the higher interconnect current densities and extreme interconnect aspect ratios. The problem is aggravated by the trend to replace the SiO_2 interconnect passivation layer with lower dielectric constant materials, such as novel organic and porous dielectrics, which also posses lower thermal conductivity and greatly impede the conduction of heat away from the interconnect and transistor; (3) The use of novel nanoscale electronics technologies further aggravates the local temperature rise around individual transistors. For example, decreasing channel dimensions result in higher power density and electron-phonon non-equilibrium within these nanoscale devices and offer greater thermal resistance between the transistors and the bulk silicon.

With the combined effects mentioned above, the technical challenges in today's thermal management of microprocessors arise from two respects. First, concomitantly with increasing performance, chip power dissipation and heat flux are

expected to increase further over the next decade. According to the International Electronics Manufacturing Initiative Technology Roadmap (iNEMI) [10], the maximum chip power dissipation is projected to be 360 W and the maximum chip heat flux to be more than 190 W/cm^2 for high performance CPUs by the end of the next decade, as indicated in Figure 1.2.
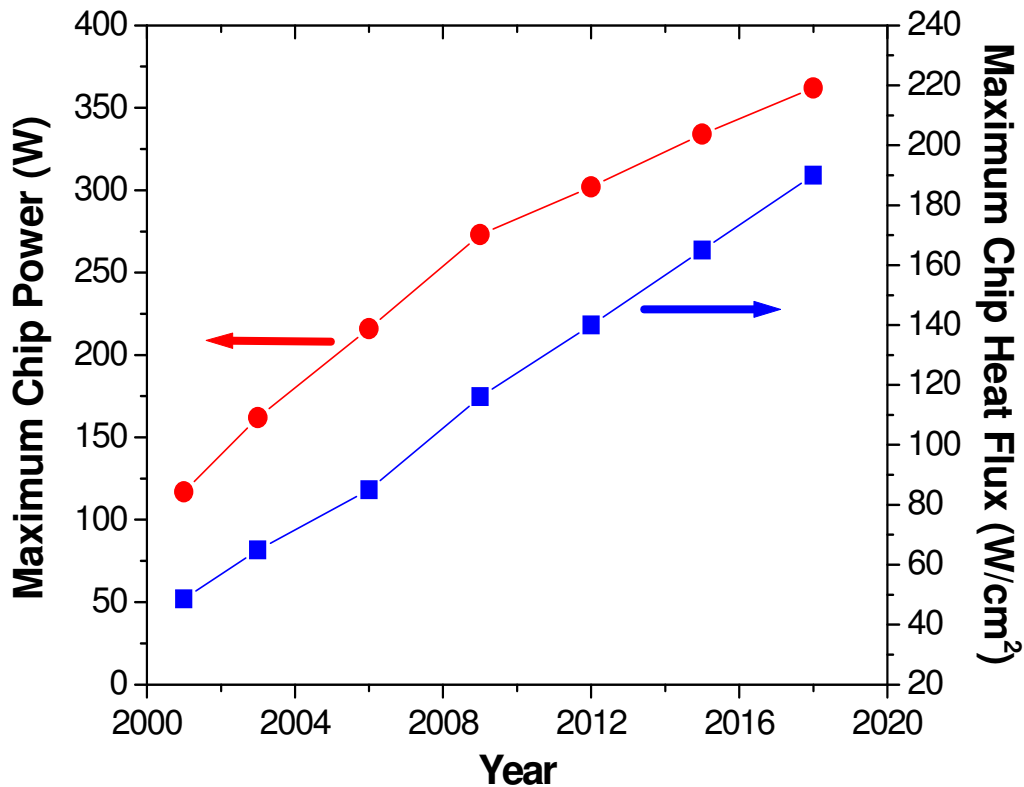
Figure 1.2: The 2004 iNEMI prediction of chip power dissipation and heat flux for high performance microprocessor chips [10].

Second, power dissipation on the chip is becoming highly non-uniform. Today's microprocessors have an average heat flux of about 10 ~ 50 W/cm^2 and a peak flux can reach six times that of the surrounding areas [11]. These large variations in heat flux arise from increasingly highly levels of device integration, specifically in the aforementioned "core" areas on the chip. Figure 1.3 shows a

4

typical power map and the resulting temperature distribution on an Intel silicon microprocessor chip, with a cell area of 1 mm × 1 mm [5,12]. With increasing performance the non-uniformity of power distribution increases, resulting in a large on-chip temperature gradient with localized, high heat flux "hot spots" that can be expected to degrade microprocessor performance and reduce reliability significantly.

(a) On-Chip Power Map   (b) On-Chip Temperature Map
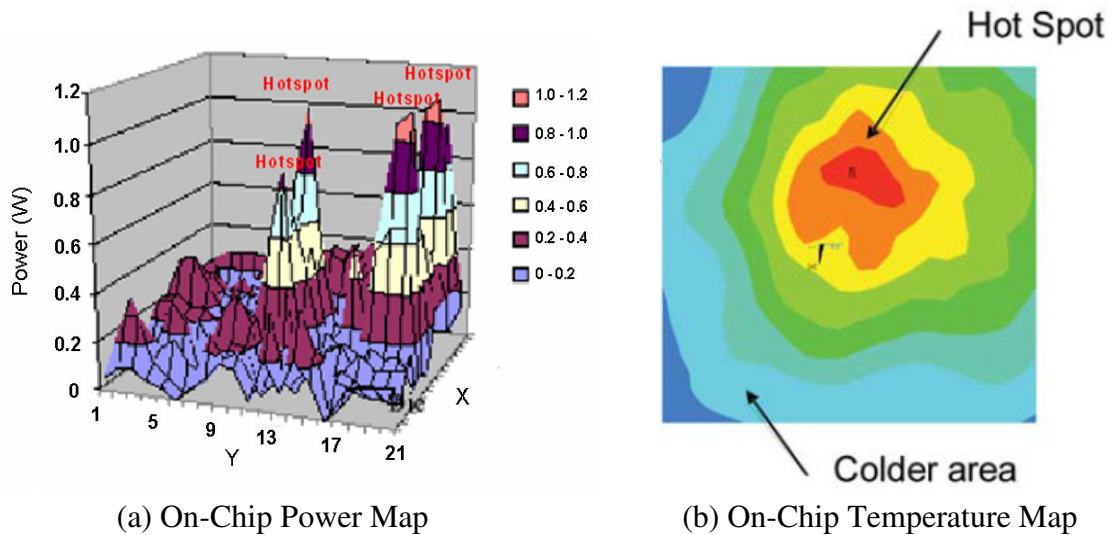
Figure 1.3: Schematic illustrating typical die power map (a) and the hot spots on the corresponding die temperature map (b). The red region represents the highest temperature spots [5].

Because chip thermal management must ensure that all junction temperatures in the microprocessor do not exceed an application-driven maximum temperature, typically in the range of 90 to 110ºC, it is often the hot spots, not the entire chip power dissipation that drives the thermal design. This leads to two undesirable consequences: (1) non-uniform heat generation limits the total heat dissipation that can be managed by a conventional thermal solution and, thus, a much more aggressive thermal solution than would be required for uniform heating, is required, and (2) the focus on controlling the temperature of the hotspot can lead to over-design

of the microprocessor cooling solution. Due to its complexity, on-chip hot spot cooling has become one of the most active and challenging research areas in thermal management of electronic devices and packages.

1.2 Potential Hot Spot Cooling Solutions

Thermal management for high flux electronic silicon chips can be classified into two strategies: passive cooling and active cooling, both of which have continued to be extensively studied during the past few years.

1.2.1 Passive Cooling Solution

Passive cooling solutions are those that do not have moving parts and generally require no external electric power to activate the cooling devices. These techniques mainly rely on heat spreading in high conductivity materials such as diamond coatings or spreaders bonded to the silicon chip, as well as on vapor transport along with evaporation and condensation in tubes and channels, to transport the heat from the high flux regions to the areas of lower heat flux.

Diamond is an attractive material for passive cooling of high flux regions on a chip because it has the highest thermal conductivity of any known materials and also has a very high electric resistance ($\sim 10^8$ Ω.m). The thermal conductivity is around 1500 ~ 2100 W/mK for single-crystal diamond fabricated by the high-pressure synthesis method and 500~1300 W/mK for polycrystalline diamond fabricated by CVD low pressure synthesis [13]. It might be the most promising spreading material because for integrated circuits the silicon is already a very good heat conductor with a thermal conductivity of about 150 W/mK, only modest spreading improvements can

be obtained from the other traditional spreading materials such as copper (390 W/mK), Beryllia (250 W/mK), Aluminum Nitride (220 W/mK) or other composites [14]. Deposition of diamond on substrates such as silicon is a reasonably mature technology and there are now multiple techniques that provide high-quality single-crystal or polycrystalline films. The fabrication of a diamond spreading layer on silicon's active region includes direct growth of diamond on the silicon substrate or bonding of a polished diamond film onto the silicon substrate. Figure 1.4 shows an example of diamond deposited directly on the aluminum metallization layers of a silicon chip. The contamination with impurities that may occur in the silicon wafer during diamond deposition, associated with the carbon, nitrogen, oxygen, hydrogen, and other elements diffusing into the device wafer from reactive gases, have kept diamond deposition from becoming the technique of choice. High temperature device processing like oxidation and annealing might enhance the diffusion of Aluminum, Nitrogen and Carbon into the device wafer. Moreover, diamond oxidizes above 600°C. To avoid these difficulties, diamond is often bonded to silicon devices using a gold–tin eutectic alloy solder film. However, the thermal contact resistance at the silicon substrate/diamond interface and metallization layers/diamond interface strongly impedes the cooling performance. Apart from the interfacial thermal resistance and cost which has limited this material to all but niche applications such as expensive laser diodes, poor adhesion of diamond to the metallization layer aggravated by the cooling and heating cycles which may occur in chips and in high frequency power devices, severely restricts the application of diamond as reliable heat spreaders.
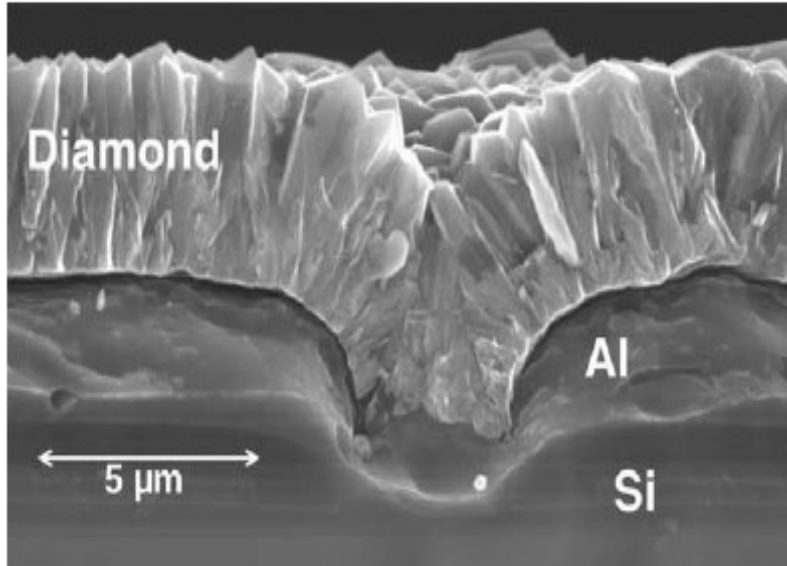
Figure 1.4: Diamond deposited on aluminum metallization in a silicon substrate using a microwave plasma technique (Courtesy of DaimlerChrysler, materials research group) [15].

Since the thermal conductivity of silicon is relatively high while the heat transfer resistance through the thickness is relatively small, a flat plate heat pipe or a high thermal conducting coating attached to the silicon chip surface is advantageous in spreading the heat in-plane at high heat fluxes. This allows them to spread heat, as well as reduce peak heat fluxes and associated temperature drops, for the largest possible fraction of the resistances in the system. The use of heat pipes in thermal management is increasing rapidly as power densities in electronics continue to rise and the electronics industry has embraced heat pipes as reliable, cost-effective solutions for high end cooling applications. Flat plate heat pipes have been found to be promising for situations where space available for the cooling system is a major constraint, such as in processor cooling in notebook computers. A heat pipe is essentially a passive heat transfer device with an extremely high effective thermal conductivity, which consists of an evacuated hermetically sealed enclosure, with three

distinct regions: an evaporator, a condenser, and an adiabatic region that separates these two regions. Figure 1.5 is a side view of a heat pipe showing the wick and the vapor/liquid flow characteristics. The enclosure contains a working fluid which absorbs heat by evaporation at the evaporator, travels as a vapor in the adiabatic region to the condenser, where it condenses and the heat is removed. The condensed fluid is then pumped back to the evaporator by the capillary forces developed in the wick structure. This continuous cycle can transfer large quantities of heat with very low thermal gradients in the internal volume. However, the contact resistance associated with the attachment of the heat pipe to the chip, conduction in the pipe wall, as well as thru the liquid saturated wick, and the attachment of the heat pipe to a cold-plate or heat sink, can lead to significant overall temperature difference from the chip to the local ambient.
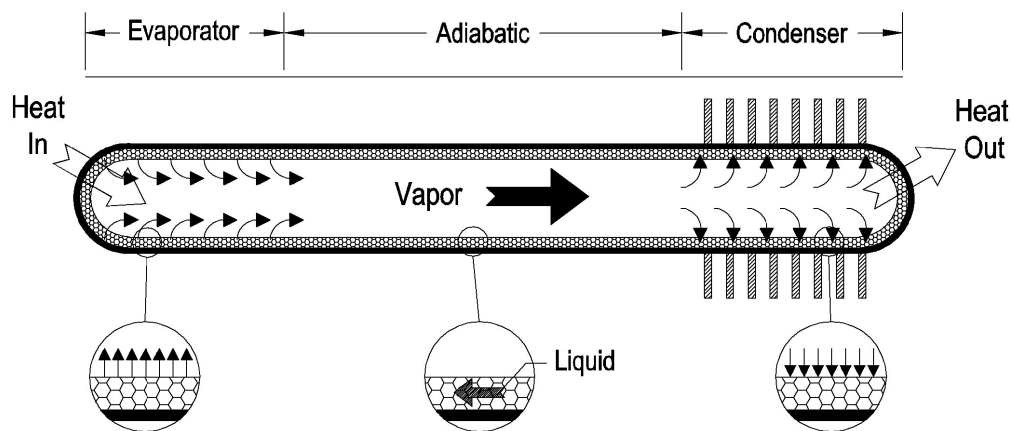
Figure 1.5: Structure of typical wicked heat pipe.

In most computer applications the operating temperatures are normally between 50 and 100$^{\mathrm{o}}$C. At this temperature range water is the best working fluid. This two-phase heat transfer mechanism lead to a dramatic increase in the effective

thermal conductivity for axial thermal transport in a heat pipe, yielding effective thermal conductivities of at least 800 W/m-K [16]. The wick pumping and local dry-out limits of the wick structure are the two most common constraints on heat pipe performance. Commercially available heat pipes can remove ~15W/cm^2 heat flux[17] and are currently used in cooling notebook computers. Research has been focused on increasing performance by using thermally driven pulsating two-phase flows [18,19], new capillary structures [20] and MEMS based heat pipes [21]. Plesch and Khrustalev reported that flat miniature heat pipes with axial grooves and using water as the working fluid were capable of sustaining heat fluxes on the order of 40 W/cm^2 [22,23]. Heat pipes with sintered powder wicks can tolerate concentrated heat fluxes of up to 80 W/cm^2 [24] without any sign of evaporator dry out. Gillot fabricated flat miniature heat pipes with micro capillary grooves inside a silicon substrate and the heat removal capability was reported to be 110 W/cm^2 [25]. Using miniaturized heat pipes Lin *et al.* reported that a cooling heat flux of 140 W/cm^2 was achieved using concentrated heating modes [20]. At higher heat fluxes boiling occurs inside the wick and the resulting liquid-vapor interactions could impede the returning liquid flow, leading to evaporator dry-out. However, using pulsating heat pipe techniques, the maximum local cooling capability can be expected theoretically to reach about 250 W/cm^2 without showing signs of evaporator dry-out [21]. For micro heat pipes integral to the chip, the transport of the dissipated heat away from the chip remains problematic. For attached silicon microcoolers, the contact resistance between the chip and the silicon microcooler can dominate the thermal performance. While heat pipes, thus posses certain inherent advantages, including the ability for self-pumping

10

and the working fluid is contained in the chamber with high reliability, heat removal rates and the overall thermal resistance of the heat pipe might still not be practical for high-flux cooling.

1.2.2 Active Cooling Solution

Active cooling solutions usually involve moving parts and require the input of electric energy for their operation.. The most common active cooling solution is air-cooled, forced convection heat sinks which have been long used for a wide range of electronic equipment, including office and desktop computers. However this conventional active cooling method has very limited capability for dealing with high flux zones on microelectronic chips. Compared to air cooling, the use of liquid coolants has many advantages such as high thermal conductivity, high specific heat, low viscosity and high latent heat of evaporation for two-phase application. As a result, different active liquid cooling technologies, such as microchannel heat sinks and direct jet impingement, have been developed due to the higher heat transfer coefficient and high cooling flux achieved as compared to air-cooling heat sink.

With microfabrication techniques developed by the electronics industry, it is possible to manufacture micro-scale three-dimensional structures. Microchannels may be machined in the chip itself or they may be machined in a substrate or heat sink to which a chip or array of chips is attached. As shown in Figure 1.6, microchannel heat sinks consist of closely-spaced parallel channels with rectangular, trapezoidal, or triangular cross sections with hydraulic diameters ranging from 100 to 1000 µm. Microchannel heat sinks can be used either with single-phase flow, where heat is transferred from the electronic chip via sensible heat gain to the coolant, or with two-

phase flow which also utilizes the latent heat of the coolant during liquid/vapor phase change. Single-phase liquid cooling systems utilize a pump to actively circulate the liquid to the micro-channels and have been studied for many years. In 1981 Tuckerman and Pease demonstrated the removal of 790 W/cm^2 heat flux by using single-phase liquid convection with water as coolant in silicon microchannels with a flow rate of 600 ml/min and the pressure drop below 50 psi [26]. Although the high heat flux capability is promising, and considerable pressure reductions and improvements in flow distribution have been achieved, the integration of such a microchannel cooler into a closed-loop system, with pipes, fittings, a pump, and a heat exchanger, is challenging due to the large pumping power requirements for the large flow rates and associated pressure drops. Apart from its reliability concern, the ability of pumps to provide both the pressure head and flow necessary for microprocessor cooling, while fitting within the chassis, is a challenge.

(a) Microchannel 1

(b) Microchannel 2

200 µm
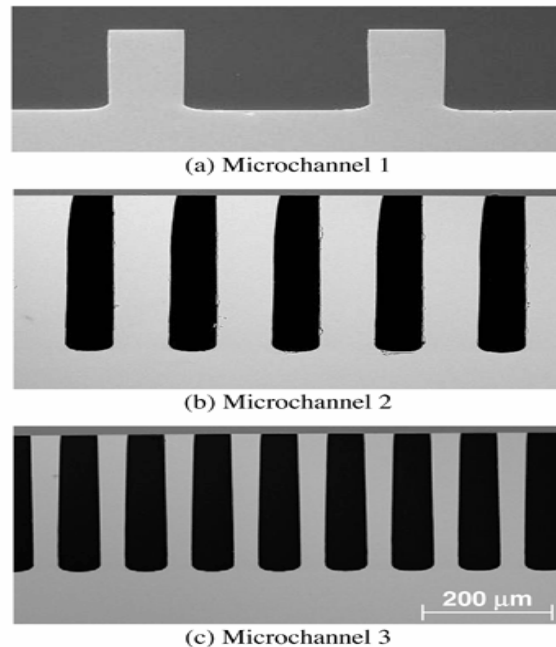
(c) Microchannel 3

Figure 1.6: SEM photos of cross-sections of three different microchannels [27].

Two-phase microchannel cooling using boiling is a more promising solution and recently has emerged as the focus of much research and practical applications because latent heat during phase-change process can be used to transfer and carry high heat fluxes. Furthermore, flow boiling requires far less coolant flow rate and often lower pumping power for removing a given amount of heat in comparison to single-phase convection. Bowers *et al.* showed two-phase microchannels could remove more than 200 W/cm^2 with flow rates of less than 65 ml/min and pressure drops of 5 psi [28]. Mudawar also demonstrated a heat removal rate of 361 W/cm^2 with two-phase forced convective cooling on an enhanced surface with FC-72 as the dielectric coolant [29]. However, the major challenge for two-phase microchannels is associated with the single and parallel channel instabilities during the liquid to vapor phase-change process. In the recent study of hot spot cooling using water-cooled two-phase microchannel, Prasher *et al.* found the biggest challenge is the huge temperature and pressure fluctuation and poor flow distribution. Poor flow distribution in two-phase microchannels might lead to less flow in the high flux regions, leading to localized dry out on the hot spot which will result in large and rapid increase in the hot spot temperature [27].

An alternative to microchannel heat sinks is jet impingement cooling, where high velocity liquid streams directly impinge onto the hot surface. This method offers several potential advantages, such as high heat transfer coefficients from the thin liquid boundary layer and uniform cooling with jet arrays. Typical jet impingement cooling might involve a single jet directed at a single component or an array of electronic components, multiple jets directed at a single component, arrays of jets

directed at an array of chips on a common substrate, or an array of jets directed at chip packages on a printed circuit board. The jets can be formed by circular slot-shaped orifices or nozzles of various cross sections. The space surrounding the jet may be filled with a gas, leading to a jet with a free surface. Alternately, liquid may occupy the space between the liquid distributor plate and the heated surface, leading to a submerged jet. As a final distinction, jet impingement cooling of electronic components may involve forced convection alone or localized flow boiling, with or without net vapor generation [6,30]. Jet impingement cooling provides very high heat transfer coefficients and could be used to meet high-flux cooling requirements. Although electronic cooling applications will require the use of dielectric liquids, Zhang demonstrated that two-phase jet impingement is capable of removing more than 100 W/cm^2 heat flux at water flow rates below 15 ml/min [31]. Boiling macrojets have demonstrated heat flux removal of over 400 W/cm^2 with water as coolant [32]. Kiper described a new method of cooling of VLSI circuits which allows one to obtain heat removal rate of more than 500 W/cm^2 using microscaled direct water impingement to IC chips from an orifice plate [33]. However, reliability, complexity, volume, weight and cost of this cooling device would be major barriers to successful commercial implementation of these approaches.

In recent years there has been increased interest in the application of solid-state thermoelectric coolers for high-flux thermal management because of their compact structure, fast response, high flux spot-cooling capability, and high reliability, due largely to the absence of moving parts [34-40]. Another major advantage of a thermoelectric cooler is that it can be miniaturized into microscale and

then integrated into the chip package. A thermoelectric cooler is based on the Peltier effect and consists of N- and P-type thermoelectric elements as shown in Figure 1.7, where the N- and P-type thermoelectric elements are joined by metallic connectors at the top and bottom. When a DC current goes through these thermoelectric element/metal contacts, heat is either released or absorbed in the contact region depending on the direction of the current. When the cold junction of the TEC is attached to the target to be cooled, the heat could be removed and the surface temperature will be reduced.
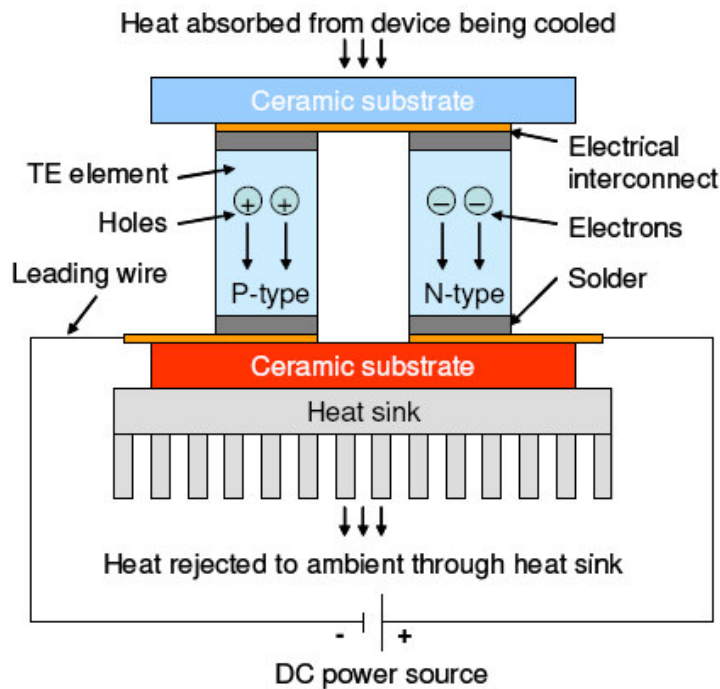
Figure 1.7: Schematic diagram of a thermoelectric cooler.

In the conventional thermoelectric cooler (TEC) design, the maximum achievable temperature reduction across the thermoelectric cooler can be estimated by:

$$\Delta T_{max} = \frac{S^2 T_c^2 \sigma}{2k} \tag{1.4}$$

and the maximum achievable cooling flux on the cold side of thermoelectric cooler can be estimated by Equation (1.5):

$$q''_{max} = \frac{S^2 T_c^2 \sigma}{2d} \tag{1.5}$$

where S is the Seebeck coefficient, k is the thermal conductivity, and ρ is the electrical resisitivity, T_c is the absolute temperature at the cold side and d is the thickness of thermoelectric elements [41-43]. Thus, the highest cooling temperature is attained for such thermoelectric materials when the Seebeck coefficient is large and the electrical resistivity and the thermal conductivity are as small as possible. Equation (1.5) indicates that the maximum cooling flux is inversely proportional to thermoelectric element thickness and thus the main advantage of going to thin-film thermoelectric coolers (TFTECs) is the dramatic increase in cooling heat flux. Fleurial estimated that the heat flux of several hundred W/cm^2 could be removed with thin film thermoelectric coolers when thermoelectric element thickness is on the order of 10 to 50 μm [44].

Recent attempts to improve the cooling performance of TEC's have focused on either material engineering to explore new materials and improve the TEC's figure of merit Z ($Z = \frac{S^2}{k\rho}$) using low-dimensional nanostructured superlattices to suppress its thermal conductivity. Venkatasubramanian reported thin-film superlattice Bi_2Te_3/Sb_2Te_3 coolers capable of providing up to 32^oC net cooling measured on the

16

cold side of the coolers and a maximum estimated cooling heat flux of 700 W/cm^2 at room temperature [35]. Harman demonstrated a thin film cooler based on quantum dot n-type PbSeTe/PbTe superlattice structure which provided 43.7^{o}C net cooling at room temperature [34]. Fan and Shakouri demonstrated the net cooling on the microcooler of up to 2.5^{o}C at room temperature and 7°C at 100°C ambient temperature for p-type thin film superlattice SiGeC/Si microcoolers and a maximum cooling heat flux as high as 680 W/cm^2 [36]. The other approach to TEC improvement is device miniaturization, to extract greater performance from existing bulk thermoelectric materials. Using ceramic thinning technology, Semenyuk developed and commercialized 200 µm-thick miniatured Bi$_2$Te$_3$ thermoelectric cooler which can provide approximately 100 W/cm^2 cooling heat flux at the cold side junction [45,46]. These reported results show solid-state thermoelectric coolers provide a powerful alternative to traditional high flux coolers and can offer great promise for reducing the severity of on-chip hot spots.

1.3 Objective and Scope of Work

Although there are extensive studies on the thermal performance of thermoelectric coolers, the ability of such devices to reduce the temperature of high-flux hot spots on the active side of packaged silicon chips and in the presence of considerable background heating, has not yet been studied. Therefore, the objective of this thesis is to explore the potential ability of thermoelectric coolers to suppress silicon chip hot spots using analytical modeling, numerical simulation, and experimental techniques. This Thesis describes innovations in basic thermoelectric

circuitry and in the enhancement of thin film TEC's for the thermal management of on-chip hot spots. The specific contributions of this thesis are: (1) proposing single-crystal silicon as a potential thermoelectric material and characterizing its thermoelectric cooling potential; (2) using the silicon chip itself to create an integral, on-chip thermoelectric cooler and reduce the hot spot temperature; (3) developing a modeling and design methodology for such on-chip thermoelectric coolers, (4) uncovering the thermal physics involved in these novel thermoelectric coolers and identifying the salient parametric effects on hot spot cooling, (5) utilizing a mini-contact pad, which connects the silicon chip and the thermoelectric cooler, to concentrate the thermoelectric cooling flux and thus significantly improve on-chip hot spot cooling performance; (6) experimentally demonstrating the cooling potential of such mini-contact assisted thin film thermoelectric coolers, and (7) defining the role played by thermal contact resistance in the efficacy of such mini-contact assisted TFTECs.

This dissertation consists of analytical modeling, numerical simulation and experiment demonstration to investigate on-chip hot spot cooling capability and is organized as follows:

Chapter 1 presents the thermal challenge of microprocessors, various hot spot cooling techniques, and the objectives of this research.

Chapter 2 reviews the principles of the thermoelectric effect and thermoelectric coolers, traditional thermoelectric materials and their cooling potential, recent progress in novel thermoelectric materials and coolers, such as superlattice structures and thin-film thermoelectric cooler.

Chapter 3 studies the thermal characteristics of a silicon microcooler using an analytical model which couples Peltier cooling with heat conduction and heat generation in the silicon substrate, and which includes heat conduction and heat generation in the metal lead. This work is used as a building block for the subsequent development of on-chip hot spot cooling using this silicon microcooler. The analytical modeling is validated by numerical simulation and experimental data. The effects of metal lead, electric contact resistance, silicon doping concentrations, and microcooler sizes on the cooling performance are investigated. The cooling potential of such thermoelectric devices, represented by the peak cooling and maximum cooling heat flux on the microcooler, is addressed.

Chapter 4 investigates on-chip hot spot cooling performance using a silicon thermoelectric microcooler through a three-dimensional analytical model. Allocation factors extracted from electro-thermal numerical simulations were combined with the analytical model to account for the impact of silicon Joule heating on the hot spot and the microcooler. It was determined that the system geometry and doping concentration in the chip can be optimized to achieve the maximum hotspot cooling performance.

Chapter 5 presents on-chip hot spot cooling performance of a silicon thermoelectric microcooler using a detailed three-dimensional numerical simulation while an approximate analytical model of on-chip hot spot cooling is developed in the last section for an idealized one-layer structure. In this section, a detailed three-dimensional package-level numerical model is developed and a simulation is performed to determine the hot spot temperature reductions associated with variations

in microcooler size, die thickness, and doping concentration along with the parasitic Joule heating effects from the electrical contact resistance and current flow through the silicon die.

Chapter 6 describes the novel use of miniaturized thermoelectric coolers for on-chip hot spot cooling enhanced with a copper mini-contact pad, which connects the thermoelectric cooler and the silicon chip to concentrate the thermoelectric cooling flux. A five-layer package-level numerical simulation is developed to predict the local on-chip hot spot cooling performance which can be achieved with such mini-contacts. Attention is focused on the hot spot temperature reduction associated with variations in mini-contact size and the thermoelectric element thickness, as well as the parasitic effect of the thermal contact resistance introduced by the mini-contact enhanced TEC. This numerical model and simulation results are validated by comparison to spot cooling experiments with a uniformly heated chip serving as the test vehicle.

Chapter 7 presents the experimental demonstration for mini-contact enhanced spot cooling performance. The experimental setup, thermal test procedure, and prototype fabrication are described in detail and the effects of mini-contact size and chip power on the spot cooling performance are investigated.

Chapter 8 summarizes the main findings of the original work presented in this thesis and provides recommendation for the future work.

Chapter 2

Thermoelectric Effects, Coolers, and Materials

Since the discovery of the Peltier effect in 1834, extensive research has been done to develop solid-state refrigeration devices and solid-state energy generators based on these thermoelectric effects. However, only limited applications of thermoelectric cooling existed until the middle 1950s, when it was discovered that doped semiconductors could achieve better thermoelectric properties than metallic materials. After several decades of research, the efficiency of thermoelectric cooling devices still reached only about 10% of Carnot efficiency in comparison to the 30% efficiency typical of vapor compression refrigerators. The progress in thermoelectrics research slowed down again until the early 1990s when theoretical predications indicated that low-dimensional materials, such as two-dimensional superlattices, could become excellent candidates as high performance thermoelectric materials. Since then extensive studies, theoretically as well as experimentally, have been conducted to explore new materials and new designs for fabrication of high-performance thermoelectric coolers. Since 2000 the demand for hot spot cooling in microprocessors provides another motivation to explore novel thermoelectric microcoolers which can be used to provide localized and high-flux cooling capability. In this chapter we review the thermoelectric effects, thermoelectric coolers, and new thermoelectric materials which can be used for thermoelectric cooling.

2.1 Thermoelectric Effect [47,48,49]

Thermoelectric coolers are energy conversion devices that use electrical energy to provide cooling capability based on the Peltier effect, while thermoelectric generators use the flow of heat across a temperature gradient to generate electrical energy based on the Seebeck effect. As shown in Figure 2.1, when two semiconductor materials A and B are joined together and electric current flows through the A/B junction, heat will be generated or absorbed at the junction at a constant rate. The heat generation rate is directly proportional to the current I and changes sign if the current changes sign:

$$q = \pi_{AB}I \tag{2.1}$$

This effect is called Peltier effect. The heat generation rate, q, is known as Peltier heat, and the coefficient π_{AB} is known as Peltier coefficient.
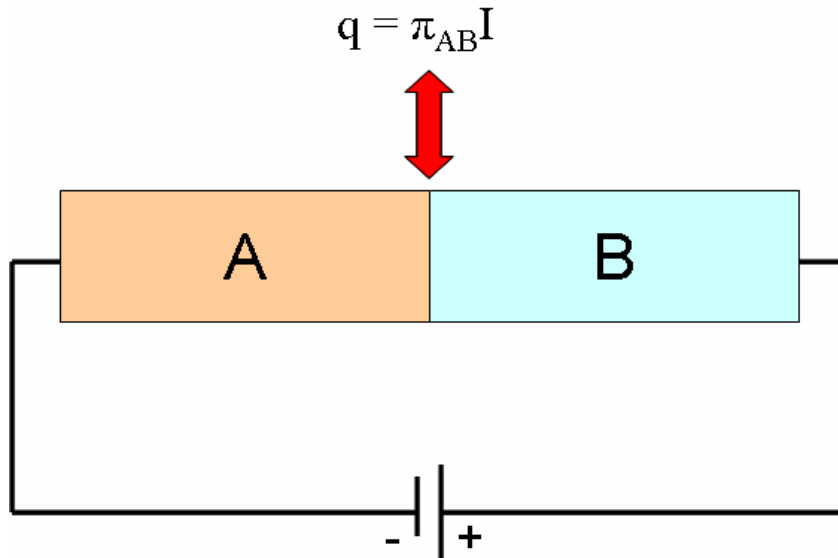
Figure 2.1: Peltier effect.

As shown in Figure 2.2, if two materials A and B are joined at two points 1 and 2 and the temperature difference ΔT is maintained between the two junctions, then an open-circuit potential difference ΔV will be developed. This effect is called Seebeck effect. The Seebeck coefficient S_{AB} is defined by:

$$S_{AB} = \lim_{\Delta T \to 0} \frac{\Delta V}{\Delta T} \qquad (2.2)$$
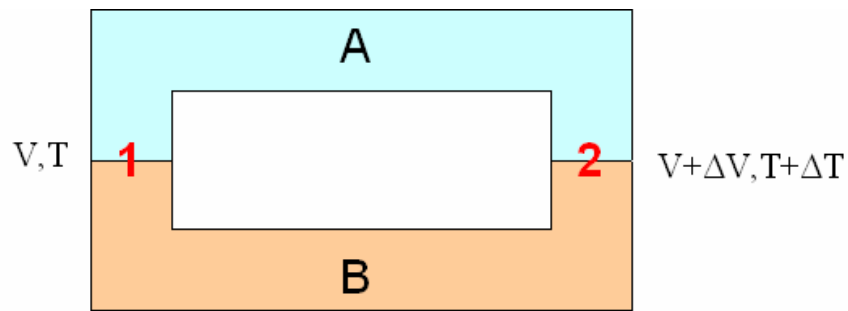
Figure 2.2: Seebeck effect.

If the Seebeck coefficient is known, the temperature difference can be measured through measuring Seebeck voltage. These two thermoelectric effects are thermodynamically related by means of the so-called Kelvin relations:

$$\pi_{AB} = S_{AB}T \qquad (2.3)$$

The third thermoelectric effect is Thomson effect in which heat is generated or absorbed by passage of a current I through a homogeneous conductor in the presence of a temperature gradient:

$$q = \mu I \frac{\partial T}{\partial x} \qquad (2.4)$$

where μ is called the Thomson coefficient. The above three effects represent

reversible interchange of electrical and thermal energies. Thomson effect is very small compared with Peltier effect and usually can be neglected.

2.2 Physical Origin of Thermoelectric Effects

In order to understand the Peltier effect from a point of view of solid state physics, let us first examine an n-type semiconductor in which conduction band electrons are the predominant charge carriers and which is connected to a voltage source by metallic conductors on both sides, as shown in Figure 2.3. Electric current flows through the semiconductor from left to right and, correspondingly, electrons flow from right to left. The electrons trying to flow into the semiconductor from the left metal conductors, face an energy barrier, which is the difference in energy between the conduction band edge (E_C) and the Fermi level (E_F), to enter the conduction band in the semiconductor. Only those electrons with energy greater than this barrier (E_C - E_F) can cross the metal-semiconductor junction and enter into the semiconductor. To gain the requisite energy level, these electrons will absorb energy from the surrounding metal lattice, creating Peltier cooling at the junction. These high-energy electrons then travel through the semiconductor but must shed the higher energy state when they cross the left metal-semiconductor junction. Consequently, the high-energy electrons will release energy to the metal lattice, creating Peltier heating. The overall result is that heat is transported from the right side to the left side of this material pair and the imposed electric current then generates a thermal "current" or heat flow. One side heats up while the other side cools down [50].
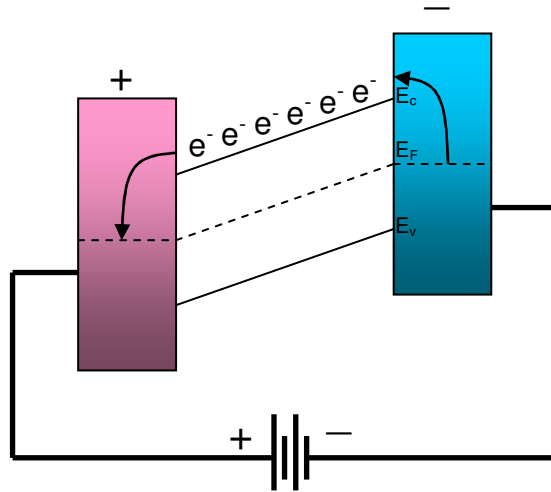
Figure 2.3: Illustration of the Peltier effect for n-type semiconductor between two pieces of metal.

(a) Open circuit

(b) Closed circuit
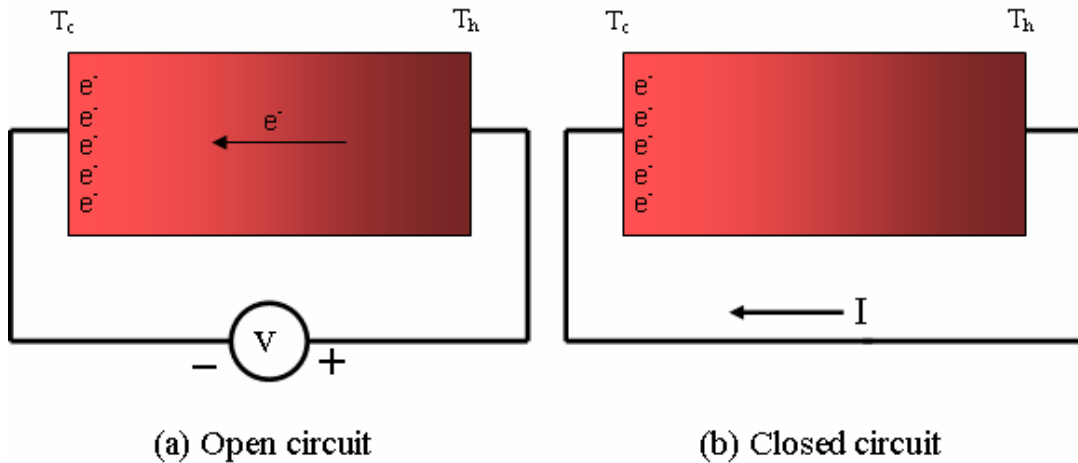
Figure 2.4: Illustration of Seebeck effect for n-type semiconductor. Note that the direction of current I is opposite the direction of electron flow.

The same type of reasoning applies to Seebeck effect. The mobile charge carriers, electrons or holes, tend to diffuse from the hot junction (T_h) to the cold junction (T_c), so that the cold junction acquires a potential of the same sign as the carriers. Figure 2.4 shows an n-type semiconductor where one side is kept at a higher temperature than the other. If the free electrons in the semiconductor are considered

to behave as a gas, the kinetic theory of gases predicts that the free electrons in the hot side of will on average have higher energy and will move at greater speeds than those in the cold side. As the faster moving electrons spread out, there will be a net flow of electrons from the hot side to the cold side of the semiconductor, creating an internal electric field to oppose further charge buildup (Figure 2.4(a)). In a closed circuit, as shown in Figure 2.4(b), electric current will flow to reduce the charge buildup and will continue to flow as long as the temperature gradient is maintained. The net result is that an imposed temperature gradient drives an electric current. For p-type semiconductor, the principle is the same except the sign of the effect is reversed.

2.3 Principle of Conventional Thermoelectric Cooler (TEC)

The principle of conventional thermoelectric coolers was developed 50 years ago by Ioffe and co-workers. A typical TEC consists of an array of n-type and p-type thermoelectric elements, two ceramic substrates that provide a mechanical integrity of a TEC, electric conductors that provide serial electric connection of thermoelectric elements and electric contacts to lead wires, solders that provide assembling of the thermoelectric elements, and lead wires that are connected to the ending conductors and deliver power from a DC electrical source. The array of p-type and n-type semiconductor elements is heavily doped and soldered to ceramic substrates so that it is connected electrically in series and thermally in parallel. As discussed in the previous section, electrons can move freely in the electric conductors but not so freely in the semiconductor. When the electrons leave the electric conductor and enter the p-

type semiconductor, they have to drop down to a lower energy level and release heat at the interface. However, as the electrons move from the p-type semiconductor into the electric conductor, the electrons are bumped back to a higher energy level and absorb heat when flowing across the interface. On the contrary, when the electrons move into the n-type semiconductor, they must bump up in energy level in order to move through the semiconductor, so that heat is absorbed. Finally, when the electrons leave the n-type semiconductor, they can move freely in the conductor and they have to drop down to a lower energy level and release heat during the process. In summary, heat is always absorbed when electrons enter n-type semiconductor or leave p-type semiconductor, and heat is always released when electrons enter p-type semiconductor or leave n-type semiconductor. The heat pumping capacity of a TEC is proportional to the current and is dependent on the element geometry, number of couples, and thermoelectric properties of the materials.
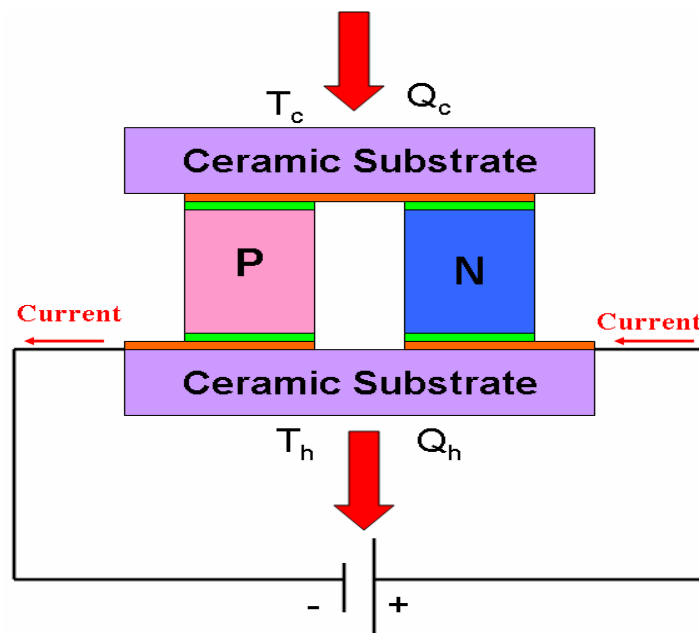
Figure 2.5: Sketch of a thermoelectric cooler.

Figure 2.5 shows the basic configuration of a TEC with one p-type element and one n-type element. To simplify analysis of a TEC, the following assumptions are made [41]:

(1) The thermal and electrical properties of the materials are independent of temperature.

(2) A linear temperature gradient exists between the hot and cold junctions.

(3) The Thomson effect can be neglected.

(4) The hot and cold junctions have no effect on the electric circuits.

(5) Convection and radiation heat transfer within the system are neglected (relative to conduction), and

(6) Geometrical similarity exists between the p-type and n-type materials.

For a single-stage TEC as shown in Figure 2.5, the amount of heat that can be pumped at the cold side of a TEC is the net of three contributions. If we assume to have perfect thermal interfaces at the cold side and the hot side, the net cooling power on the cold side of the TEC can be expressed by:

$$q_c = (S_p - S_n)IT_c - K(T_h - T_c) - \frac{1}{2}I^2 R_e \tag{2.5}$$

where R is the overall electrical resistance of the TEC:

$$R_e = \rho_p \frac{L_p}{A_p} + \rho_n \frac{L_n}{A_n} \tag{2.6}$$

K is the overall thermal conductance of the TEC

$$K = k_p \frac{L_p}{A_p} + k_n \frac{L_n}{A_n} \tag{2.7}$$

and S, k, ρ, A and L represent the Seebeck coefficient, thermal conductivity, electrical resistivity, the cross-section area, and the thickness of thermoelectric element, respectively. The p and n denote p-type and n-type thermoelectric materials. T_h and T_c are the temperature at the cold side and the hot side of the TEC, respectively. The overall cooling rate is driven by the Peltier cooling (the first term) and reduced by Joule heating in the element (the second term) and the heat flowing back from the hot side to the cold side of the TEC (the third term).

When a current is applied to the TEC, a voltage drop is generated, which includes the resistive voltages and the Seebeck voltages across the thermoelectric elements and is given by:

$$V = (S_p - S_n)(T_h - T_c) + IR \tag{2.8}$$

So, the electric power consumption of the TEC system is equal to:

$$W = (S_p - S_n)(T_h - T_c)I + I^2 R \tag{2.9}$$

The coefficient of performance (COP) is used to describe the performance of the TEC, which is the net cooling power at the cold side divided by the powder consumption to the system given by:

$$COP = \frac{q_c}{W} \frac{(S_p - S_n)IT_c - K(T_h - T_c) - \frac{1}{2}I^2 R_e}{(S_p - S_n)(T_h - T_c)I + I^2 R} \tag{2.10}$$

The maximum cooling rate that can be achieved by this TEC device can be determined by differentiating the cooling rate given in Equation (2.5) to find the optimal current I_{opt}:

$$\left(\frac{dq_c}{dI}\right)_{opt} = 0 \Rightarrow \quad I_{opt} = \frac{(S_p - S_n)T_c}{R_e} \tag{2.11}$$

When I = I_{opt}, the heat removal rate attains its maximum value given by:

$$q_{max} = \frac{1}{2} \frac{(S_p - S_n)^2 T_c^2}{R_e} - K(T_h - T_c)$$ (2.12)

To determine the largest temperature reduction that the TEC can achieve, i.e. the deepest cooling, the heat removed from the cold side is set equal to zero, $q_{max} = 0$, yielding the maximum temperature difference across the cold side and the hot side of the TEC:

$$\Delta T_{max} = \frac{(S_p - S_n)^2 T_c^2}{2KR_e} = \frac{ZT_c^2}{2}$$ (2.13)

Where Z is the figure of merit given by:

$$Z = \left[\frac{(S_p - S_n)}{(k_p \rho_p)^{0.5} + (k_n \rho_n)^{0.5}} \right]^2$$ (2.14)

For simplification, it is frequently assumed that Seebeck coefficient equal but opposite in sign, that is, $S_p = -S_n = S$, thermal conductivity, electrical resistivity and geometry are equal for the two elements, i.e., $\rho_p = \rho_n = \rho$, $k_p = k_n = k$, $A_p = A_n = A$, $L_p = L_n = L$, So Equation (2.14) can be simplified as:

$$Z = \frac{S^2}{k\rho}$$ (2.15)

Then the maximum cooling temperature at the optimized current, I = I_{opt}, can be calculated as:

$$\Delta T_{max} = \frac{S^2 T_c^2}{2k\rho} = \frac{1}{2} ZT_c^2$$ (2.16)

Similarly, the maximum cooling power at the optimized current $(I = I_{opt})$ and when the temperature difference across the cold side and the hot side is equal $(T_h = T_c)$ can be calculated as:

$$q_{max} = \frac{S^2 T_c^2}{2R_e} = \frac{S^2 T_c^2 A}{2d\rho} \qquad (2.17)$$

The corresponding cooling heat flux can be calculated as:

$$q_{max}^{''} = \frac{q_{max}}{2A} = \frac{S^2 T_c^2}{2d\rho} \qquad (2.18)$$

It is interesting to note that the maximum achievable cooling ΔT_{max} only depends on the figure of merit Z and the temperature of the cold side of the TEC, but it does not change with the geometry of the TEC including the cross-section area and thickness of the element. To attain the lowest temperature, a thermoelectric material with a high figure of merit Z is required. Alternatively, the power factor P, given by equation (2.19), is often used to characterize the thermoelectric properties of a given material:

$$P = \frac{S^2}{\rho} \qquad (2.19)$$

The maximum COP can be found by differentiating the COP given in Equation (2.11) to find the optimal current $I_{COP,opt}$:

$$\left(\frac{dCOP}{dI}\right)_{opt} = 0 \Rightarrow I_{COP,opt} = \frac{(S_p - S_n)(T_H - T_c)}{R_e(1 + ZT_{ave})^{0.5} - 1} \qquad (2.20)$$

When $I = I_{COP,opt}$, the maximum value of COP can be calculated:

$$COP_{max} = \frac{(1 + ZT_{ave})^{0.5} - T_h/T_c}{(1 + ZT_{ave})^{0.5} + 1} \qquad (2.21)$$

where $T_{ave} = (T_h + T_c)/2$ is the mean temperature of the TEC.

2.4 Thermoelectric Cooling Materials and Devices

As described in the last section, for conventional thermoelectric applications, the best thermoelectric materials should have Seebeck coefficients and electric conductivities as high as possible and thermal conductivity as low as possible. So far the best thermoelectric material properties are found in heavily doped semiconductors. Insulators have very low electrical conductivity while metals have relatively low Seebeck coefficients and high thermal conductivity. In addition, the thermal conductivity of a metallic material, which is dominated by electrons, is proportional to the electrical conductivity, as dictated by the Wiedmann–Franz law. Therefore it is impossible to improve electrical conductivity but at the same time suppress thermal conductivity for metallic materials. In semiconductors, the thermal conductivity is established by the flow of both electrons and phonons, but with much of the thermal transport ascribed to phonons. The phonon thermal conductivity can be reduced without causing too much reduction in the electrical conductivity. A common approach to reduce the phonon thermal conductivity is through alloying or doping because the mass difference scattering in alloys or doped semiconductors reduces the lattice thermal conductivity significantly without much degradation to the electrical conductivity [51].

Bismuth telluride-based compounds based on alloys of Bi_2Te_3 with Sb_2Te_3 (p-type) and Bi_2Te_3 with Bi_2Se_3 (n-type), are the best commercial state-of-the-art materials for thermoelectric cooling with the highest values of the figure of merit, *ZT,* and the highest power factor, P. In bulk materials a ZT of 0.75 for *p*-type $(BiSb)_2Te_3$ at room temperature was reported about 40 years ago. Since the 1960s much effort

has been made to raise the *ZT* of bulk materials based on bismuth telluride by doping or alloying other elements in various fabricating processes. Recently the highest *ZT* of 1.14 at 300 K has been reported for the *p*-type $(Bi_{0.25}Sb_{0.75})_2(Te_{0.97}Se_{0.03})_3$ alloy [52]. By annealing the ingots prepared by the Bridgman method, Yamashita *et al*. have most recently achieved a significant increase in the *ZT* value to 1.19 at 298 K for the *n*-type $Bi_2(Te_{0.94}Se_{0.06})_3$ and 1.41 at 308 K for the *p*-type $(Bi_{0.25}Sb_{0.75})_2Te_3$ alloy, so that both ZT values exceed 1 [53-55].

Besides alloying, several other approaches have been proposed to enhance ZT through either improving electrical conductivity or reducing thermal conductivity. In this respect, low-dimensional materials, such as quantum wells, superlattices, quantum wires, and quantum dots offer new ways to manipulate the electron and phonon properties of a given material. Some experiments have demonstrated that superlattice thermoelectric materials can achieve ZT more than 2.0. More recently theoretical predictions, based on nonconservation of lateral momentum which allows a higher number of electrons to participate in the thermionic emission process, show that metal-based superlattices with high energy barriers can achieve much large effective thermoelectric figure of merit (*ZT* > 5) at room temperature [56]. These inspiring results show the feasibility of applying thermoelectric materials to high-flux cooling applications.

Thermoelectric coolers (TEC) have traditionally been fabricated using bulk bismuth telluride materials and traditional processing techniques such as hot pressing sintering and extrusion. Commercial bulk thermoelectric coolers are made from thermoelectric elements, typically several mm in thickness and in the range of 1.8

mm × 3.4 mm (and 2.4 mm thick) to 62 mm × 62 mm (and 5.8 mm thick). For such commercial TE modules, the maximum cooling at room temperature is about 70°C, with relatively low cooling heat flux of 5 ~ 10 W/cm^{2}, which makes it impossible to use such TEC's for high-flux hot spot cooling application [57]. However, since the maximum cooling heat flux of a TEC is inversely proportional to the thickness of its elements, there have been extensive studies focusing on microscale thin film TECs, miniatured TECs and superlattice TECs in order to realize high flux cooling requirement for on-chip hot spot reduction.Significant progresses in recent years has been reported in making microscale thermoelectric coolers, as described in the three sections that follow.

2.4.1 Thin Film Thermoelectric and Coolers (TFTEC)

It is widely accepted that thin-film thermoelectric coolers (TFTEC) have great potential for high-flux cooling because of their main advantage – a dramatic increase in cooling heat flux with decreasing thermoelectric element thickness. Due to the excellent thermoelectric properties of bismuth telluride-based compounds, numerous state-of-the-art techniques have been applied to the deposition of thermoelectric thin films, including flash evaporation [58], molecular-beam epitaxy [59], metalorganic chemical vapor deposition (MOCVD) [60], sputtering [61], co-sputtering [62], co-evaporation [63], laser ablation [64], and electrodeposition [65].

Among these various deposition methods, electrochemical deposition is very attractive from an application perspective due to its ability to deposit thin films at high deposition rates of tens of microns per hour and at the much lower batch processing cost when compared to other state-of-the-art thin film fabrication

processes. Snyder *et al.* at the Jet Propulsion Laboratory used electrochemical-MEMS technique to fabricate thin film thermoelectric microcoolers which contains 63 n-type Bi_2Te_3 elements and 63 p-type $Bi_{2-x}Sb_xTe_3$ elements, each element being 20 μm in thickness and 60 μm in diameter with bridging metal interconnects, as shown in Figure 2.6. The cooler was fabricated on either glass or an oxidized silicon substrate (Si/SiO_2) and its area is close to 1700 μm × 1700 μm. The maximum cooling was found to be only 2^oC, under the optimized applied current of 110 mA at the ambient temperature of 80^oC, and the maximum cooling heat flux was calculated to be 7 W/cm^2. It was found that although the electrodeposited thermoelectric material has the expected elemental composition and crystal structure, the defect structure produced a high concentration of low mobility carriers, yielding a Seebeck coefficient of 60~100 μV/K in comparison to ~ 200 μV/K for corresponding bulk materials at room temperature. The resulting Z value was estimated to be 3.2×10^{-5} (1/K) and the effective ZT at room temperature only 0.01 [66-69], indicating the poor quality of the thin films.
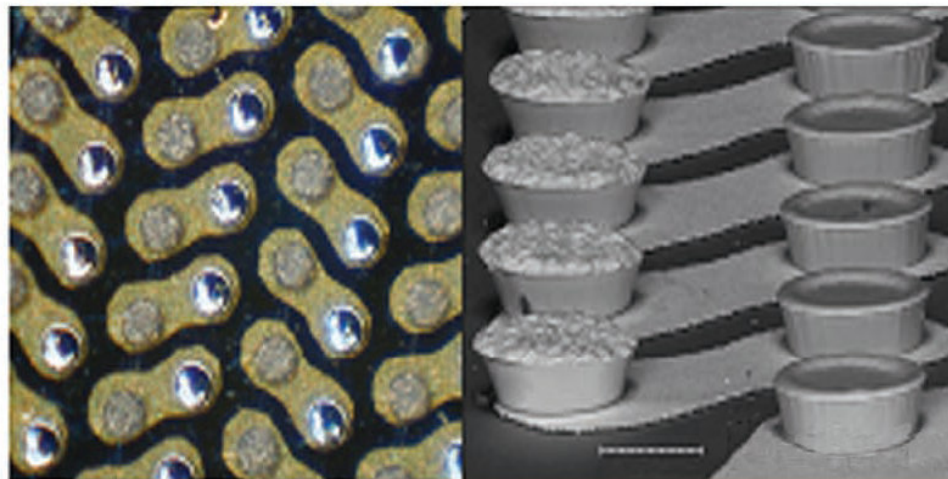
Figure 2.6: Thin film thermoelectric microcooler fabricated by MEMS [67].
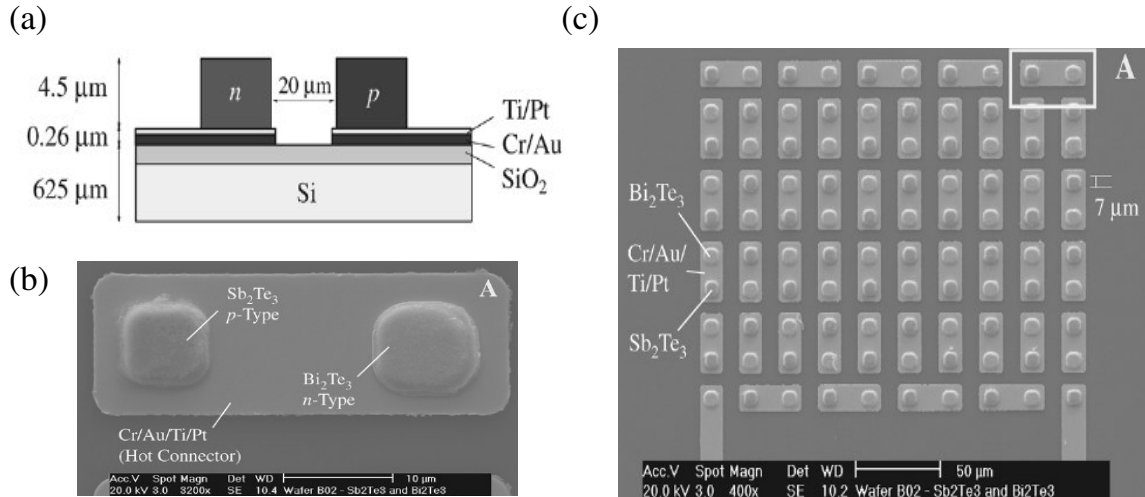
Figure 2.7: Bi$_2$Te$_3$ and Sb$_2$Te$_3$ films deposited on Cr/Au/Ti/Pt bottom connectors. (a) Sketch showing the approximate thickness of the films. (b) SEM micrograph showing a top view of the fabricated structures. This image is the enlarged section "a" indicated in the part (c), top right. (c) SEM micrograph of a device with 50 TE pairs [70].

Using co-evaporation as the deposition method, da Silva *et al.* fabricated a thin film thermoelectric microcooler which provided 60 n-type and p-type thermoelectric element pairs, with the thickness and width of the elements approximately 4.5 and 40 µm, respectively, as shown in Figure 2.7. The columns made of p-type Sb$_2$Te$_3$ and n–type Bi$_2$Te$_3$ are connected using Cr/Au/Ti/Pt layers at the hot junctions, and Cr/Au layers at the cold junctions. The measured Seebeck coefficient, electrical resistivity, and power factor of the thermoelectric films, which were deposited with a substrate temperature of 130°C, were 74µV/K, 3.6×10^{-5} Ω.m and 0.15 mW/K^{2}m for n-type Bi$_2$Te$_3$, respectively, and 97 µV/K, 3.1×10^{-5} Ω.m and 0.30 mW/K^{2}m for p-type Sb$_2$Te$_3$, respectively. The maximum cooling was reported about 1°C [70,71]. In her more recent work, the thermoelectric properties was improved to 228 µV/K, 2.83×10^{-5} Ω.m and 1.84 mW/K^{2}m for p-type Bi$_2$Te$_3$ thin films and 149 µV/K, 1.25×10^{-5} Ω.m and 1.78 mW/K^{2}m for p-type Sb$_2$Te$_3$ thin film at

36

the optimized deposition temperature [72]. However, the overall thermoelectric cooling performance achieved with such improved thin films has not been reported.

Böttner *et al.* developed a two-wafer process to fabricate thin film thermoelectric coolers (Micro-Peltier coolers) using chip-to-chip, chip-to-wafer, or wafer-to-wafer soldering. Figure 2.8 depicts a schematic drawing of the two wafer process in the left part and also schematic drawing of the resulting device in the right part. The polycrystalline n-type $Bi_2(Se,Te)_3$ and p-type $(Bi,Sb)_2Te_3$ materials were deposited by co-sputtering from 99.995% element targets (Bi, Sb, Te) onto pre-structured electrodes and the observed growth rate was in the range of five micrometers per hour. Regrettably, these alloys were not grown very well in thin film form due to delivery problems of the Se-target suppliers. For these coolers, the thickness of n-type Bi_2Te_3 elements and p-type $(Bi,Sb)_2Te_3$ elements is about 20 µm and the electric contact resistances is around $10^{-6}\Omega.m^2$ or better, as shown in Figure 2.9. A net cooling of $11^{o}C$ at the ambient temperature of $60°C$ was reported for current flow of 800 mA, with the highest Seebeck coefficient reaching 175 µV/K for n-type elements and 180 µV/K for p-type elements. This process yielded a maximum power factor of 1.57 mW/K^2m for n-$Bi_2(Se,Te)_3$ and 2.5 mW/K^2m for p-$(Bi,Sb)_2Te_3$ [73]. More recently, Böttner *et al.* reported the maximum temperature differences of around $48^{o}C$ could be achieved under vacuum conditions at an applied current of 2.1A and a maximum cooling flux of ~100 W/cm^2 for the complete device was measured [74].
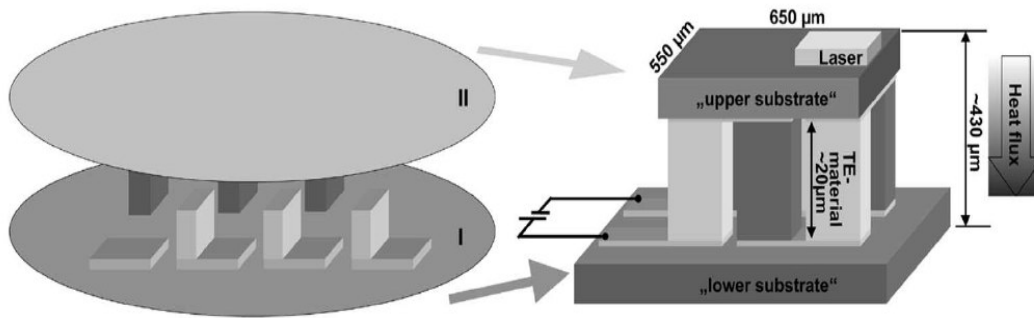
Figure 2.8: Micro-Peltier cooler. Left: Schematic drawing of the developed two wafer (I,II) concept. Right: Schematic drawing of the thermoelectric cooler used for telecommunication device [73].
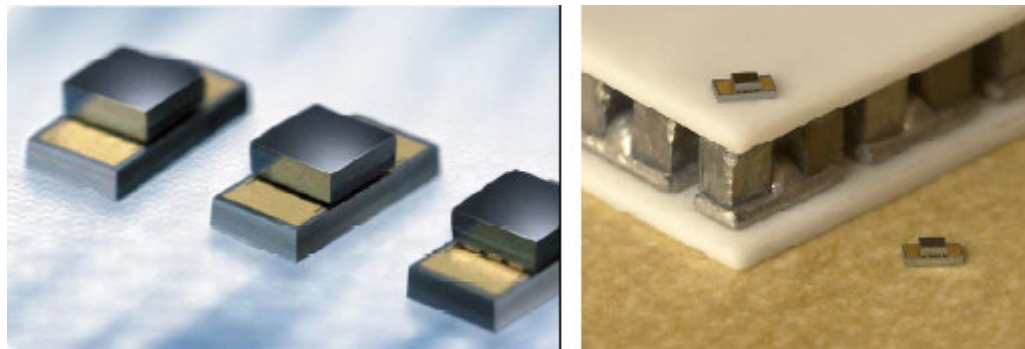
Figure 2.9: Micro-Peltier cooler. Left: individual miniaturized Micro-Peltier coolers. Right: comparison to a commercial bulk Peltier cooler [73].

Zou *et al.* found that direct vapor deposition of the bismuth telluride compounds is made difficult by the large difference in the vapor pressure between antimony, bismuth and tellurium, which could result in non-congruence and in a lack of stoichiometry. In his work, Sb_2Te_3 films were deposited by co-evaporation of antimony and tellurium and Bi_2Te_3 thin films by co-evaporation of bismuth and tellurium onto heated, clean glass substrates. High purity (99.999%) antimony, bismuth and tellurium were used as the evaporants. It is found that the best quality films can be obtained for p-type Sb_2Te_3 film (α=160 µV/K, ρ=3.12 ×10^{-5} Ω.m), and

for n-type Bi_2Te_3 (α=-200 μV/K, ρ=1.29 $\times10^{-5}$ Ω.m) at a substrate temperature of 190 ^{o}C and 230 ^{o}C, respectively. The figure of merit Z for the p-type Sb_2Te_3 film and n-type Bi_2Te_3 film was calculated and found to be approximately 1.04$\times10^{-3}$ at room temperature, corresponding to ZT of 0.32. The maximum value of temperature difference measured between the hot and cold end was 15.5^{o}C at a current of 55 mA, showing a promising procedure for fabricating thermoelectric microcooler [62].

Table 2.1: Summary of Cooling Performance of Bi_2Te_3 - based Thin Film TEC.

Author (Year)	Growth Method	ΔT_{max} (^{o}C)	q_{max} (W/cm^2)	TE properties
Snyder (2002)	electrochemical deposition	2.0@80°C	7	S = 60~100 μV/K ZT=0.01 (estimated)
Zuo (2002)	Co-sputtering	15.5@25ºC	N/A	p-type: S=160 μV/K ρ=3.12 $\times10^{-5}$ Ω.m P=0.82 mW/K^{2}m n-type: S=-200 μV/K ρ=1.29 $\times10^{-5}$ Ω.m P=3.10 mW/K^{2}m
da Silva (2005)	co-evaporation	1.0@25ºC	N/A	p-type: S= 228 μV/K ρ=2.83$\times$10-5 Ω.m P=1.84 mW/K^{2}m n-type: S=-149 μV/K, ρ =1.25$\times10^{-5}$ Ω.m P= 1.78 mW/K^{2}m
Böttner (2005)	Co-sputtering	48@25ºC	100@25ºC	p-type: S=180 μV/K ρ=1.30 $\times10^{-5}$ Ω.m P=2.5 mW/K^{2}m n-type: S=-175 μV/K ρ=1.95 $\times10^{-5}$ Ω.m P=1.57 mW/K^{2}m

In using the state-of-the-art deposition techniques to develop thermoelectric thin films, a primary difficulty is to maintain the stoichiometry of the bismuth telluride compounds. For example, the problem of resputtering during the film growth is present in sputter deposition while differences in volatility of the component elements pose difficulty in vacuum evaporation. A large deviation from stoichiometry arises in vapor deposition because the constituent elements in the target exhibit dissimilar sticking coefficients on the substrate. In addition, there is a tendency for re-evaporation of certain elements from the deposited thin films because of their higher vapor pressure. Therefore, the thermoelectric properties of these thin films reported in the above publications vary widely and the figure of merit (ZT) is always much small than ~1.0 for bulk bismuth telluride material. As shown in Table 2.1, to date thin film thermoelectric coolers are still not well developed and complete characterization of the materials properties of the various TEC thin-films is lacking.

2.4.2 Bulk Miniaturized Thermoelectric Coolers

Although thin film deposition technology has an advantage for mass production, currently it appears not to provide thermoelectric cooling performance comparable to that available in bulk coolers, due to the difficulty in controlling thin film growth conditions to obtain the desired stoichiometry, e.g., Bi/Te ratio and Se/Te ratio, and a defect-free microstructure. Alternatively, bulk miniatured thermoelectric coolers, based on thinning of bulk materials, seem to be more promising, because this technique can reduce thermoelectric element thickness down to microns and, at the same time, it can maintain the excellent thermoelectric properties of the bulk

materials. Table 2.2 shows the development progress in bulk miniaturized TEC since 1960s [75].

Table 2.2: Thermoelectric Cooling of Bulk Bi$_2$Te$_3$-based Miniaturized TEC.

Year	TEC Configuration	TE Element		ΔT_{max} (°C)	q''_{max} (W/cm^2)
		Thickness (μm)	TE Properties		
1967	Single TE couple	130	ZT=0.54	38 @30°C	95
1994	20 couple TECs	100	ZT=0.78	50 @30°C	100
1997	120 couple TECs	200	ZT=0.78	67 @30°C	65
2002	18 couple TECs	200	ZT=0.90	70.6 @30°C	80
				91.8 @85°C	98
2006	18 couple TECs	130	ZT=0.90	64.2 @30°C	110
				83.5 @85°C	132

In 1967 Semenyuk at Thermion Inc. began developing bulk bismuth telluride material technology so that thermoelectric elements could be shortened to several hundred microns. Starting from typically 1 to 2 mm thick thermoelectric elements, Semenyuk successfully fabricated a single-element TEC with 130 µm thick element in 1967. However, due to the poor quality of thermoelectric material at that time (Z = 1.8 ×10^{-3} K^{-1}) and also poor soldering processing with which electric contact resistance could be as high as 10^{-5} Ωcm^2, net cooling of only 38°C and cooling heat flux of only 95 W/cm^2 were demonstrated at an ambient temperature of 30°C [76]. Using improved extruded thermoelectric materials (Z = 2.6 × 10^{-3} K^{-1}) and Al$_2$O$_3$ as the substrates, in 1994 Semenyuk reported that bulk miniaturized TEC's could provide the maximum cooling of 50°C and the maximum cooling heat flux of 100

41

W/cm^2 for 100 µm thick TEC, and the maximum cooling of 62°C, the maximum cooling heat flux of 65 W/cm^2 for 200µm thick TEC [77,78].

In his later work in 1997, Semenyuk found that the thermal resistance of the Al$_2$O$_3$ substrates is the dominating factor controlling the cooling performance of such thin thermoelectric elements and that further improvement of thermoelectric efficiency could be achieved by either reducing ceramic thickness or using materials with higher thermal conductivity such as BeO, AlN or diamond. Experimental and detailed theoretical analysis of the performance of miniatured bulk coolers based on extruded thermoelectric materials (Z = 2.6 ×10^{-3} K^{-1}) and diamond substrates were reported. For a TEC with 200 µm thick thermoelectric elements, a maximum cooling of 67°C and a maximum cooling heat flux of 65 W/cm^2 was demonstrated and the result was comparable to those obtained in most commercial coolers [79]. Great progress in bulk thinning technology was achieved in 2002 when a new series of single stage TEC with the thermoelectric element as short as 200 µm was developed and commercialized [80]. AlN substrates were used and the extruded thermoelectric materials have excellent values of the thermoelectric figure of merit of 3.02×10^{-3} K^{-1}. The fabricated TEC's showed acceptable mechanical strength when sliced down to 200 µm. The maximum cooling of 70.6^{o}C and 91.8^{o}C, and the maximum cooling flux of 80 W/cm^2 and 98 W/cm^2 were demonstrated at 30°C and 85°C, respectively.
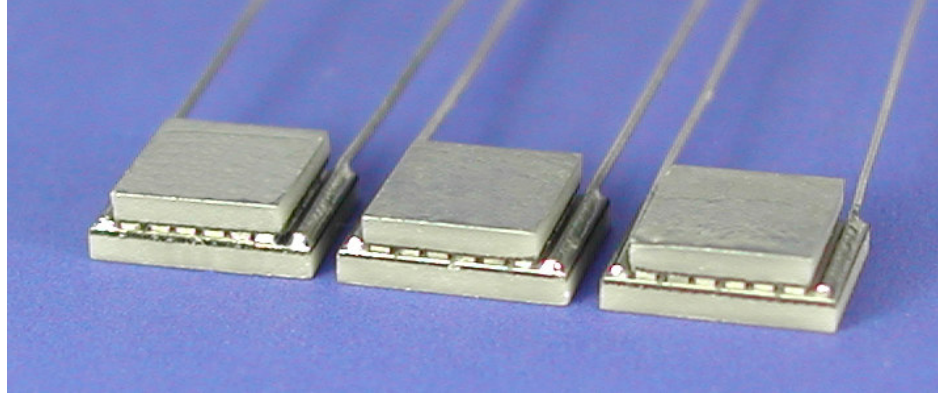
Figure 2.10: Thermion TECs with 130 µm thick TE elements [80].

In 2006 Semenyuk further thinned the thermoelectric element thickness down to 130 µm to achieve higher cooling performance as shown in Figure 2.10 [80]. The extruded p-type and n-type bismuth telluride thermoelectric materials were in a form of rods with Z values of 3.02×10^{-3} K^{-1} measured at 25ºC, corresponding to ZT value of 0.9. The 200 µm thick p-type and n-type slices were initially cut from the rods using electroerosion process. Then the slices were lapped to the final thickness of 130 µm, etched electrochemically, and nickel plated. The thermoelectric elements were produced in a lot by cutting slices to the dimensions of 370 µm $\times$ 370 µm and 635 µm thick AlN ceramic substrates were used with metal patterns obtained by standard microelectronics processing, including vacuum deposition of thin-film adhesive layers followed by electrochemical growth of thick copper films through the photoresist processing, nickel plating, and finally Tin solder electrodeposition. The modules were tested in vacuum and the maximum cooling of 64.2 and 83.5ºC and the maximum cooling flux is of 110 and 132 W/cm^2 were reported when measured at 30ºC and 85ºC, respectively. Compared with miniaturized TEC with 200 µm thick

elements, the present TEC did not improve maximum cooling as indicated in Figure

2.11. However, the maximum cooling heat flux did improve by 30% under both

conditions, showing great promise for on-chip hot spot cooling using this

miniaturized TEC.
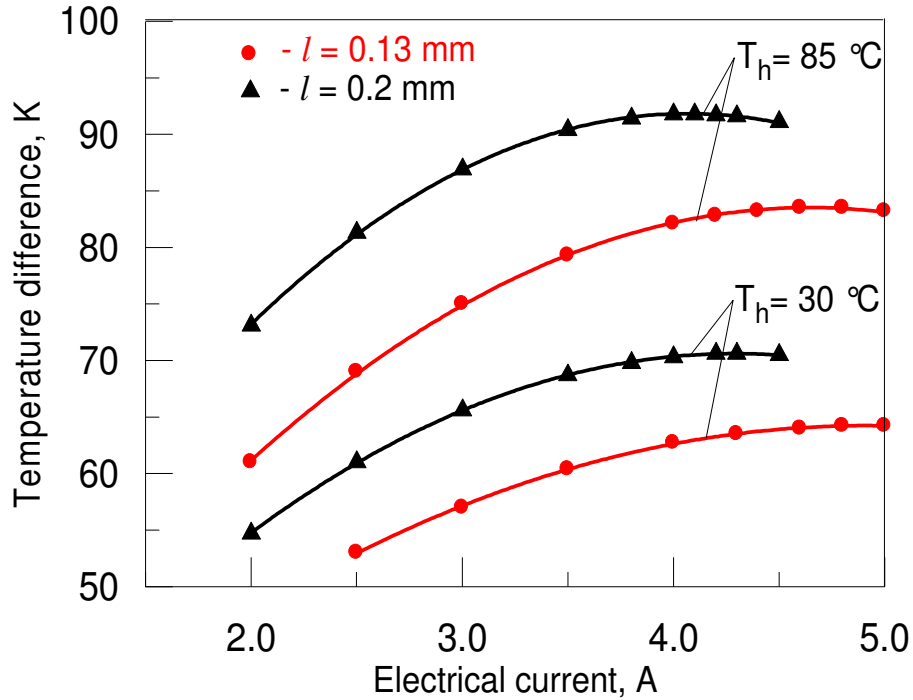
Figure 2.11: Variation of net cooling with applied current for Thermion TECs with 130 μm and 200 μm thick elements [80].

Similar to Semenyuk's work, Bierschenk *et al.* demonstrated TEC's with

thermoelectric element thickness of 25 µm, 84 µm, 100 µm and 137 µm developed

using Marlow's Micro Alloyed Material bulk Bi_2Te_3 thermoelectric materials [81].

However, no cooling performance was reported so far. It should be noted that due to

fragility of bismuth telluride thermoelectric materials and the lack of industrial

experience in microfabrication, the fabrication yield could be low. However,

compared with the cooling performance of TFTEC developed by state-of-the-art deposition technique, bulk materials-based techniques appear to be better suited to the manufacturing of highly effective microscale TEC with higher net cooling and high cooling heat flux. Bulk TECs with TE element length down to 200 μm are available at the market already and there are good prospects for further TE element miniaturization.

2.4.3 Superlattice Thermoelectric Cooler

As discussed in previous sections, the commercial thermoelectric cooling materials are based on alloys of Bi_2Te_3 with Sb_2Te_3 (such as $Bi_{0.5}Sb_{1.5}Te_3$, p-type) and Bi_2Te_3 with Bi_2Se_3 (such as $Bi_2Te_{2.7}Se_{0.3}$, n-type), each having a ZT approximately equal to one at room temperature. Low-dimensional materials, such as quantum wells, superlattices, quantum wires, and quantum dots offer new ways to manipulate the electron and phonon properties of a given material. In the regime where quantum effects are dominant, the energy spectra of electrons and phonons can be controlled through altering the size of the structures, leading to new ways to increase ZT. In this regime, the low-dimensional structures can be considered to be new materials, despite the fact that they are made of the same atomic structures as their parent materials. At each "scale" - resulting from a size reduction - these changed relationships can provide a "new" material whose properties must be examined, to a certain extent, both theoretically and experimentally, to define the resulting thermoelectric properties. Thus searching for high ZT systems in low-dimensional structures can be regarded as the equivalent of synthesizing many different bulk materials.

The benefits of low-dimensional structure can come from two respects. Firstly, it can improve electric properties and thus improve the power factor and ZT. For example, in a low-dimensional n-type material, the Fermi level is lower and the Seebeck coefficient is higher than that for corresponding bulk semiconductors with the same electron concentration, enhancing the value of the product of S^2/ρ, the power factor. Dresselhaus and co-workers theoretically predicted that the use of quantum wells could increase the power factor via quantum size effects, which improve the electron performance by taking advantage of sharp features in the electron density of states and ZT > 2~3 can be achieved [82]. Secondly, it can reduce thermal conductivity due to significantly modified phonon dispersion and enhanced phonon scattering mechanisms. In other words, it can reduce the thermal conductivity using a short period superlattice designed to impede phonon transport without excessively restricting the carrier flow [83,84]. There was extensive experimental evidence that superlattices could be made into superior thermal insulators, promising an effective approach to improving the figure-of-merit, ZT. The experimental studies have demonstrated significant thermal conductivity reduction in a wide variety of superlattices [85], and significant enhancements of the thermoelectric figure-of-merit were reported in Bi$_2$Te$_3$/Sb$_2$Se$_3$ superlattices [35] and PbTe/PbTeSe superlattices[34].

Table 2.3 compares the reported power factor, ZT and thermal conductivity of these structures with that of their corresponding bulk materials at room temperature. It is clear that thermal conductivity reduction plays a significant role in the reported ZT enhancement while there is only small improvement in the power factor. Thirdly, it can be used to achieve thermionic (TI) cooling, especially multilayer TI cooling

46

[86,87], which allows for low parasitic Joule heating since transport through the thin barriers is largely ballistic.

Table 2.3: Thermoelectric Properties of Superlattices with high ZT.

Thermoelectric Properties at 25°C	PbTe-PbSeTe Quantum Dot Superlattices	PbTe-PbSe Bulk Alloy	Bi_2Te_3-Sb_2Te_3 Superlattices	Bi_2Te_3-Sb_2Te_3 Bulk Alloy
Power Factor (mW/K^2m)	3.2	2.8	4.0	5.0
Thermal Conductivity (W/mK)	0.6	2.5	0.5	1.45
Figure of Merit ZT	1.6	0.35	2.4	1.0

In addition to the intensive theoretical and experimental research that has been performed on the thermoelectric properties of these nanostructured materials, the overall cooling performance of the devices based on such nanostructure has also been explored recently. LaBounty and Shakouri developed a InGaAs/InGaAsP thin film microcooler to demonstrate thermionic cooling effect. The tested cooler structure consisted of a 1 µm thick superlattice barrier (25 periods of 10 nm thick InGaAs and 30 nm thick InGaAsP) surrounded by InGaAs cathode and anode layers grown by metal-organic chemical vapor deposition (MOCVD). The microcooler size ranged from 20 µm × 40 µm to 100 µm × 200 µm. Experimental results demonstrated 1.2 °C and 2.3°C net cooling when tested at 25 and 90°C, respectively, and the maximum cooling heat flux was estimated at about several hundred W/cm^2 [88]. For more optimized structures and packaging, that is, removal of parasitic effect of the system,

47

simulations predicted a net cooling of 20~30°C at room temperature with a cooling heat flux of several 1000 W/cm^2 [89]. Similarly, Zhang *et al.* developed a AlGaAs/GaAs superlattice thermionic microcooler using metal-organic chemical vapor deposition (MOCVD) and a maximum cooling of 0.8 °C and 2.0 °C were demonstrated at 25°C and 100°C for 60 µm × 60 µm microcooler [90].However, due to the thermal properties of AlGaAs/GaAs superlattice not being available, the maximum cooling heat flux was not estimated.

Shakouri and co-workers fabricated thin-film SiGe/Si, and SiGeC/Si thermoelectric microcoolers based on superlattice structures using molecular beam epitaxy (MBE) [43,91-94]. As the SiGe/Si superlattice microcoolers can be monolithically integrated with microelectronic components to achieve localized cooling and temperature control, such devices provide great advantages for on-chip hot spot cooling. The microcooler structure is based on cross-plane electrical transport and the main part of the TEC is a 3 µm thick strain-compensated SiGe/Si superlattice, as shown in Figure 2.12. It consists of 200 periods of 12 nm $Si_{0.9}Ge_{0.1}$/3 nm Si, doped with boron to about 6×10^{19} cm^{-3}. A maximum cooling of 4.5°C at 25°C, 7°C at 100°C and 14°C at 250°C was demonstrated. The maximum cooling heat flux increases with decreasing microcooler size, increasing from 120 to 680 W/cm^2 when the microcooler sizes reduces from 100 µm × 100 µm to 60 µm × 60 µm. It is expected that cooling heat flux could be improved further at higher temperature and that the application of TEC's at higher temperature would be of more practical value. However, to date there is no report regarding the maximum cooling heat flux that such devices can provide at higher temperature.
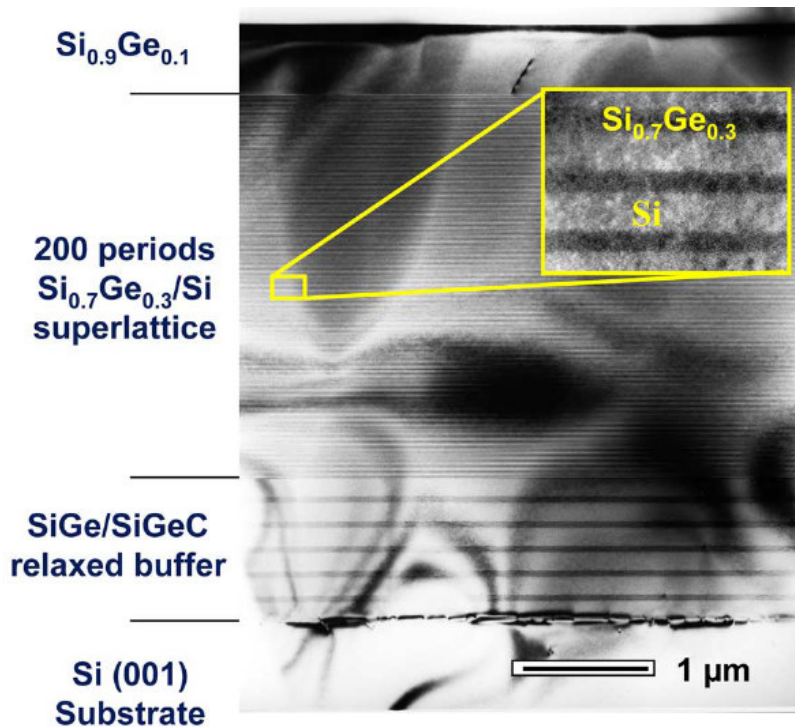
Figure 2.12: Transmission electron micrograph of 3 µm thick 200×(5 nm Si$_{0.7}$Ge$_{0.3}$ /10 nm Si) superlattice grown symmetrically strained on a buffer layer designed so that the in-plane lattice constant was approximately that of relaxed Si$_{0.9}$Ge$_{0.1}$. The n-type doping level (Sb) is 2 ×10^{19} cm^{-3}. The relaxed buffer layer has a ten-layer structure, alternating between 150-nm Si$_{0.9}$Ge$_{0.1}$ and 50-nm Si$_{0.845}$Ge$_{0.150}$C$_{0.005}$. 0.3 µm Si$_{0.9}$Ge$_{0.1}$ cap layer was grown with a high doping to get a good ohmic contact [43].

Venkatasubramanian *et al.* used metal–organic chemical vapor deposition (MOCVD) to epitaxially grow a 5 µm thick Bi$_2$Te$_3$/Sb$_2$Te$_3$ superlattice on GaAs substrates in 2001 [35,95]. These are phonon-blocking/electron-transmitting superlattices which are produced by alternately depositing thin (1~4 nm) films of Bi$_2$Te$_3$ and Sb$_2$Te$_3$. ZT is reported to be 2.4 for p-type nanostructured superlattices Bi$_2$Te$_3$/Sb$_2$Te$_3$ and 1.4 for n-type Bi$_2$Te$_3$/Bi$_2$Te$_{2.83}$Se$_{0.17}$ at room temperature. This high ZT was explained by a reduction of the lattice thermal conductivity due to scattering of the phonons at the superlattice interfaces. The maximum cooling of 32.2 and 40°C was measured using an infrared camera and the maximum cooling flux of

49

585 and 700 W/cm^2 was estimated for p-tyep Bi$_2$Te$_3$/Sb$_2$Te$_3$ superlattice at the temperatures of 25 and 80°C, respectively. Because of only 5 µm thickness, response time was only about 5 µsec.
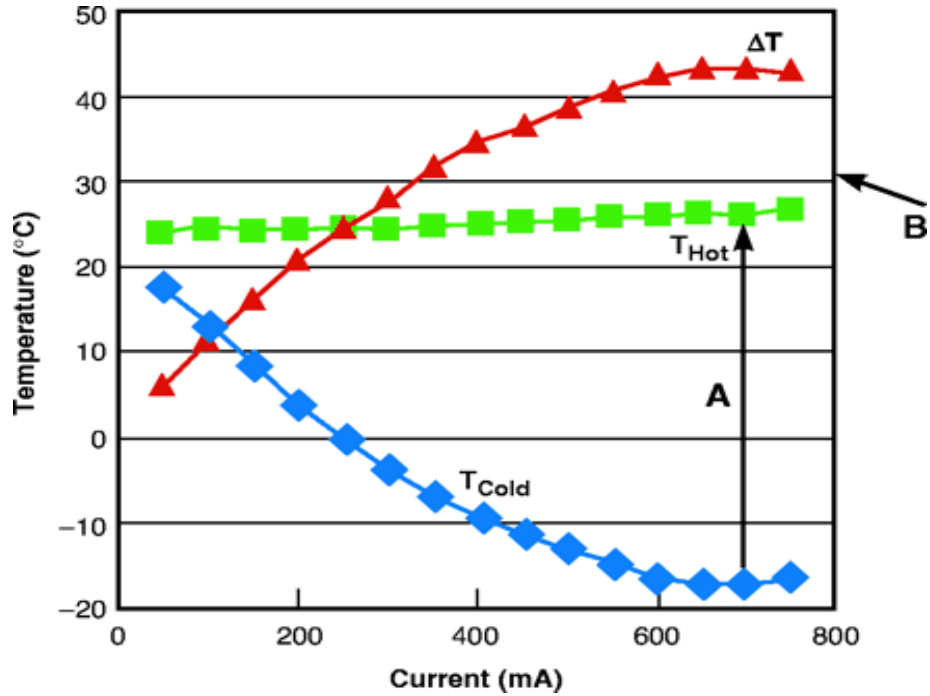
Figure 2.13: Thermoelectric cooling characteristics of one-leg device made from n-type PbSeTe/PbTe superlattice thermoelectric cooler. The red curve represents measured data points of temperature differential between the hot and cold junction temperatures versus the electrical current flowing through the device. (A) Maximum cooling of 43.7 ^{o}C measured for the superlattice cooler. (B) Maximum cooling of 30.8 ^{o}C measured for the conventional TE cooler made of n-type bulk (Bi,Sb)$_2$(Se,Te)$_3$ solid solution alloy and measured in the same test setup [35].

More recently PbSeTe-based quantum dot superlattice structures grown by molecular beam epitaxy (MBE) were reported by Herman's group for thermoelectric cooling applications [34]. The superlattice thin film with a thickness of approximately 100 µm is grown on BaF$_2$ substrates. The developed superlattice thin film n-type PbSeTe/PbTe has a ZT of 1.6~2.0 at the room temperature. Under vacuum

conditions, a maximum cooling of 43.7^{o}C was reported at 700 mA at 25°C ambient

temperature, as shown in Figure 2.14, in comparison to 30.8 ^{o}C for the conventional

n-type bulk $(Bi,Sb)_2(Se,Te)_3$ thermoelectric alloy measured in the same test setup.

Table 2.4 is the summary of cooling performance of thin film superlattice TEC

developed since 2000.

Table 2.4: Summary of cooling performance of superlattice TEC.

Year	Superlattice	Growth Method	ΔT_{max} (K)	q_{max} (W/cm^2)	ZT Properties
LaBounty (2000)	InGaAs/InGaAsP	MOCVD	1.2@25ºC 2.3@90ºC	Several 100's	N/A
Fan (2002)	SiGe/Si SiGeC/Si	MBE	4.5@25ºC 7.0@100ºC 14.0@250ºC	680@25ºC	S=200µV/K k=6.8~8.7 W/mK P=2.2 mW/K^{2}m ZT=0.085
Venkatasu bramanian (2001)	Bi_2Te_3/Sb_2Te_3	MOCVD	32.2@25ºC 40@80ºC	585@25ºC 700@80ºC	ZT=2.4 P=4.0 mW/K^{2}m
Herman (2002)	PbSeTe/PbTe	MBE	43.7 @25ºC	N/A	ZT=1.6 P=3.2 mW/K^{2}m
(Zhang) 2003	AlGaAs/GaAs	MOCVD	0.8@25ºC 2.0@100ºC	N/A	N/A

2.4.4 Silicon Thermoelectric Materials and Microcooler

While single-crystal silicon has been the key semiconductor material for much

of the microelectronics era, silicon's thermoelectric potential has been largely ignored

because of its high thermal conductivity and thus low value of figure of merit

(ZT≈0.017) [96,97]. However, silicon thermoelectric microcoolers, when formed on

the back of the silicon chip for hot spot cooling, provide unique advantages over TFTEC's. Despite its low ZT value, silicon constitutes a very viable candidate for high-flux cooling due to its high Seebeck coefficient and low electrical resistivity, which combine to yield a high power factor. Table 2.5 provides the thermal and electrical properties for three conventional thermoelectric materials, bulk Bi_2Te_3 alloy bulk SiGe alloy and single-crystal silicon, at room temperature. It can be seen that single-crystal silicon appears to offer the highest power factor of the materials shown, due to its high Seebeck coefficient and low electrical resistivity, and thus constitutes a very viable candidate for high-flux cooling.

Table 2.5: Typical values on the thermoelectric properties for Bi_2Te_3, SiGe and single-crystal silicon at room temperature.

Material	Seebeck Coefficient S (μV/K)	Electrical Resistivity ρ ($\mu\Omega$m)	Thermal Conductivity k (W/mK)	Figure of Merit	Power factor P (mW/K^{2}m)
Bi_2Te_3 (n-type)	-240	10	2.02	Z=2.85×10^{-3} ZT=0.86	5.76
Bi_2Te_3 (p-type)	162	5.5	2.06	Z=2.32×10^{-3} ZT=0.70	4.77
SiGe (n-type)	-136	10.1	4.45	Z=0.328×10^{-3} ZT=0.1	1.83
SiGe (p-type)	144	13.2	4.80	Z=0.413×10^{-3} ZT=0.12	1.57
Silicon (p-type)	450	35	150	Z=0.039×10^{-3} ZT=0.012	5.79

The thermoelectric properties of silicon depend on doping concentration which follows classic semiconductor theory. To attain the highest possible thermoelectric cooling flux, it is necessary to obtain as large a Seebeck coefficient and as low an electrical resistivity as possible and, in particular, to maximize the

power factor value of S^2/ρ. The electrical resistivity of semiconductors is known to decrease with increasing carrier concentration and carrier mobility and to be given by:

$$
\begin{aligned}
n-type \;\; Silicon: \;\; &\rho_{n-Si} = \frac{1}{en\mu_n} \\
p-type \;\; Silicon: \;\; &\rho_{p-Si} = \frac{1}{ep\mu_p}
\end{aligned}
\tag{2.22}
$$

where e is electron charge, n and p are the concentrations of electron and hole, respectively, and μ_n and μ_p are the mobility of the electron and hole, respectively. In silicon semiconductors, higher doping concentration leads to higher carrier concentration but lower carrier mobility and the combined effect is that the electrical resistivity decreases with increasing doping concentration [98].

The relationship between Seebeck coefficient and carrier concentration can be derived from solid state theory [99,100]. The value of the Seebeck coefficient is mainly determined by the difference in energy between the conduction (or valence) band edge and Fermi level and, when the weak dependence of the Seebeck coefficient on temperature is neglected, is approximately given by equation 2.23):

$$
\begin{aligned}
n-type \;\; Silicon: \;\; &S_{n-Si} \approx -\frac{k_B}{e}(\ln\frac{N_c}{n}+4) \\
p-type \;\; Silicon: \;\; &S_{p-Si} \approx \frac{k_B}{e}(\ln\frac{N_v}{p}+4)
\end{aligned}
\tag{2.23}
$$

where, N_c and N_v are the effective density of states at the conduction band and the valence band, respectively, n and p are the concentrations of electron and hole, respectively, and k_B is the Boltzmann constant. The Seebeck coefficient thus has a complex inverse relationship with doping concentration. Although both the electrical resistivity and the Seebeck coefficient are inversely dependent on doping in silicon,

due to their substantially different functional dependence on doping concentration, it is possible to find an optimum value that maximizes the thermoelectric cooling flux. The theoretical dependence of the silicon thermoelectric cooling "power factor", S^2/ρ, on the doping concentrations at $100^{\circ}C$, for boron-doped single-crystal silicon, is illustrated in Figure 2.14, which explicitly shows the optimum doping concentration to equal $2.3 \times 10^{19} cm^{-3}$.
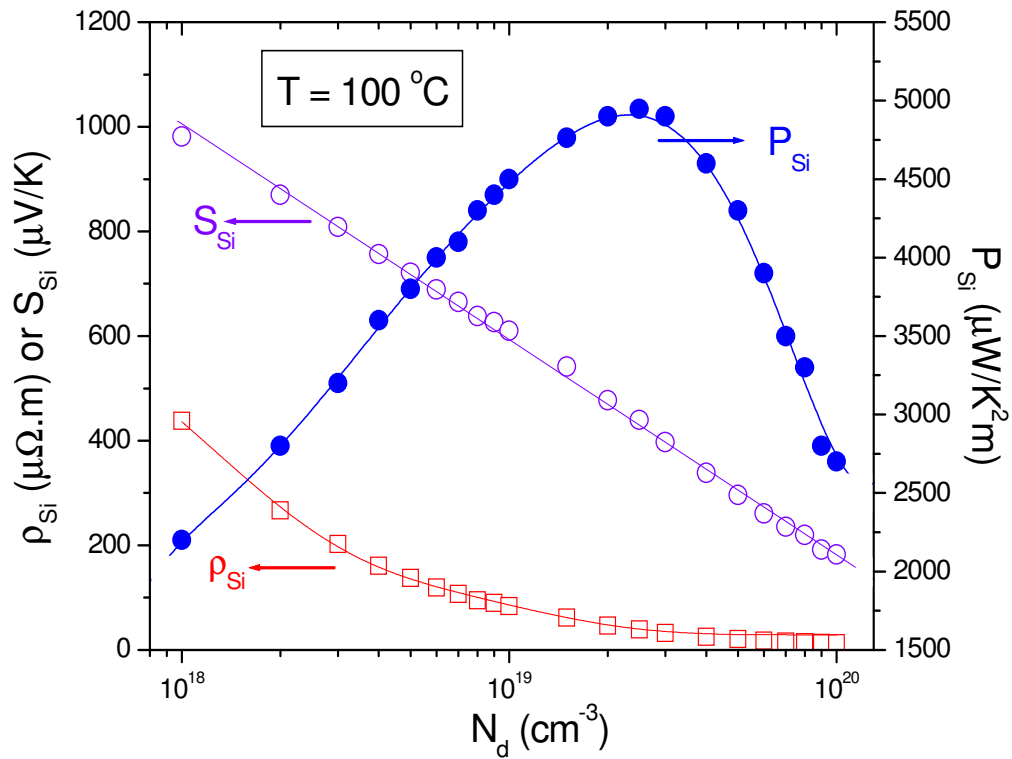
Figure 2.14: Dependence of thermoelectric properties on boron doping concentration for single-crystal silicon.
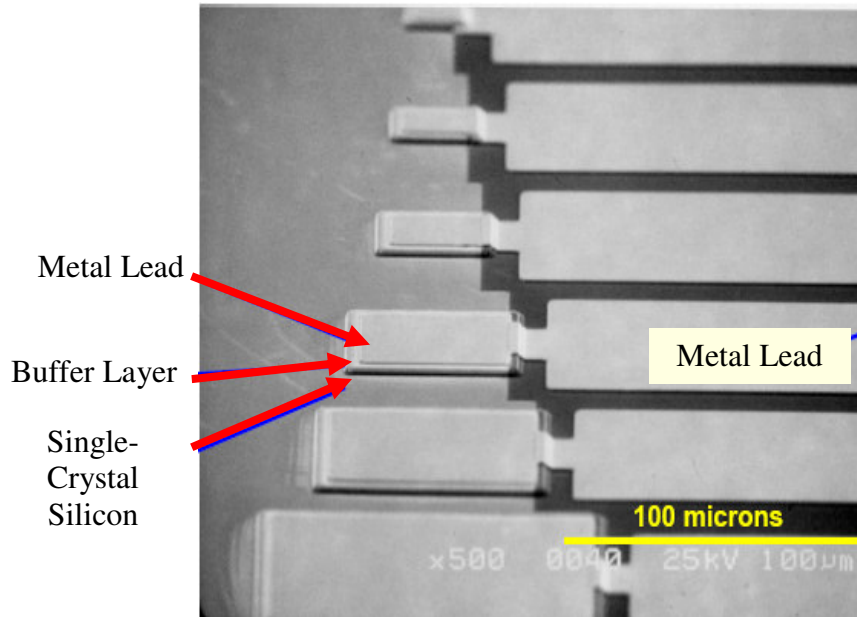
Figure 2.15: A SEM photo of silicon microcooler.

Zhang and Shakouri developed silicon thermoelectric microcooler using bulk silicon which is *p*-type boron doped at a doping concentration of around 10^{19} cm^{-3}. The device structure is illustrated in Figure 2.15, which was fabricated with standard microfabrication techniques: dry etch, lithography, metal evaporation, etc. They experimentally demonstrated the ability of silicon thermoelectric microcoolers to achieve a maximum cooling of 1.2°C for a 40 μm ×40 μm microcooler at the optimized current of 0.1 A and 0.88°C for a 75 μm × 75 μm microcooler at the optimized current of 0.32 A. The maximum cooling flux of 580 W/cm^2 and 250 W/cm^2 was estimated a 40 μm ×40 μm and 75 μm × 75 μm, respectively, showing that a silicon microcooler is a very attractive spot cooling solution for integrated circuits using the substrate's thermoelectric properties [101].

2.5 Conclusions

In this chapter the basic thermoelectric phenomena, including the Seebeck effect, the Peltier effect, and the Thomson effect are introduced and the operating principle of conventional TEC's is explained using simple configurations. Recent developments in thermoelectric cooler technology, such as thin film TEC's, miniaturized bulk TEC's and superlattice TEC's are reviewed and the thermoelectric properties of various thin film and bulk materials are discussed.

Chapter 3

Analytical Modeling of Silicon Mircrocooler

While single-crystal silicon has been the key semiconductor material for much of the microelectronics era, silicon's thermoelectric potential has been largely ignored because of its high thermal conductivity and thus low value of the figure of merit, ZT ($\approx$0.017). Recently Zhang and Shakouri demonstrated the concept of silicon microcoolers at room temperature with the maximum cooling heat flux (or cooling power density) of more than 500W/cm^2 and, more interestingly [101], it was predicted that 3D silicon microcooler structure could exceed a 1D device's cooling capabilities by more than doubling the maximum cooling estimated from Equation (1.4). Three-dimensional package-level FEA simulation also suggests that silicon thermoelectric microcoolers can be used to selectively cool on-chip hotspots [102,103]. However, careful thermal design and optimization will be needed to best exploit the Peltier cooling capability achievable in silicon microcoolers and to overcome the parasitic effects, such as electrical contact resistance and heat generation and conduction in the metal lead [104-109] inherent in the use of this technique. Consequently, in this section we develop an analytical model that can be used to predict the temperature reduction on the silicon microcoolers, reflecting the effects of the silicon doping concentrations, microcooler sizes, heat generation and conduction in the metal lead, and electrical contact resistance on the cooling performance. Results obtained from the analytical model will be compared with the available experimental data and with the three-dimensional thermal-electric numerical simulations. Please note that this work is device-level silicon microcooler modeling

without hot spot and thus not directly related with hot spot cooling. In hot spot cooling using silicon microcooler as shown in Chapter 4, there is an effective heat transfer coefficient applied on the silicon die and heat conduction path will be different from device-level model. However, this work is of great interest from a point of view of device physics and can be used to understand thermal physics involved in silicon microcooler and related thermal phenomena in this system.

3.1 Structure and Operating Principle

The structure of a silicon thermoelectric microcooler is illustrated in Figure 3.1. It is a single element silicon microcooler with cross-plane electrical transport through the silicon substrate. The metal lead, which is electrically isolated from the silicon substrate with a very thin SiN_x layer is employed to deliver electric current to the microcooler through the silicon cap layer. The current then flows into the silicon substrate and continues out through the ground electrode on the base of the silicon substrate, which is also maintained at a fixed temperature by an appropriate cooling system. In the reported studies of such microcoolers, the silicon substrate was boron-doped single-crystal silicon with a thickness of 500µm and the silicon cap layer, less than 1µm thick, was highly-doped silicon with a doping concentration larger than 1×10^{20} cm^{-3} (in order to improve ohmic contact between the metal contact and the silicon cap). The silicon microcooler sizes under current investigation ranged from 20µm×20µm to 100µm×100µm, while the width of the metal lead varied from 20µm to 200µm and the thickness varied from 1.0µm to 3.0µm. The passivation layer of SiN_x thin film was about 0.3 µm in thickness and of the same width and length as the metal lead layer.
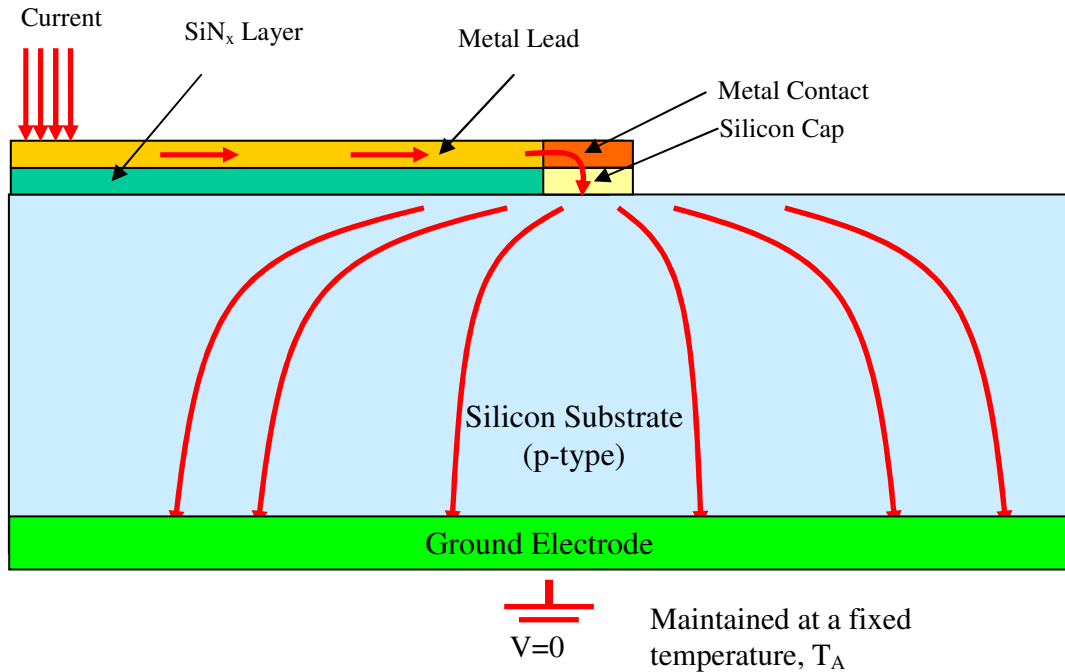
Figure 3.1: Structure of a silicon thermoelectric microcooler (The arrows indicate the direction for electric current flow).

A thermoelectric cooler uses an electric current to induce the Peltier effect, at the junction of two materials with different Seebeck coefficients, to provide localized cooling, and to transport the absorbed heat to the hot side of the thermoelectric circuit. Joule heating associated with the resistance to current flow in the thermoelectric circuit, and heat conduction from the hot to the cold side of the thermoelectric circuit, limits the thermoelectric cooling that can be achieved. Referring to the structure of the silicon microcooler depicted in Figure 3.1, it may be seen that, electric current flowing through the indicated circuit results in Peltier cooling at the junction of the metal contact/silicon cap and again at the silicon cap/silicon substrate interface, but causes Peltier heating at the silicon substrate/ground electrode interface, where energetic electrons must shed some of

their energy in entering the highly-conductive metal. The Peltier cooling rate at the metal contact/silicon cap interface is represented as:

$$q_{TE,1} = (S_{metal} - S_{cap})T_1 \, I \approx -S_{cap}T_1 I \qquad (3.1)$$

where S_{metal} and S_{cap} are the Seebeck coefficients of the metal contact and silicon cap layer, respectively. T_1 is the absolute temperature at the interface between the metal contact and the silicon cap layer and I is the applied current. It is to be noted that by comparison to the high Seebeck coefficient of silicon materials under consideration, the Seebeck coefficient of the metal contact, S_{metal}, is very low and can be neglected by comparison to S_{cap}.

The Peltier cooling rate at the silicon cap/silicon substrate interface is given by:

$$q_{TE,2} = (S_{cap} - S_{Si})T_2 I \qquad (3.2)$$

where S_{Si} is the Seebeck coefficient of the silicon substrate, which varies with the doping concentration, and T_2 is the absolute temperature at the interface between the silicon cap layer and silicon substrate.

Since the highly-doped silicon cap layer is very thin (<1µm) and the thermal conductivity is very large (100~150W/mK at operating temperatures), the temperature difference between these two interfaces can be neglected, i.e. $T_1 = T_2 = T_c$. So, to a very good approximation, the overall Peltier cooling rate of the silicon microcooler can be expressed as:

$$q_{TE} = q_{TE,1} + q_{TE,2} = -S_{cap}T_1I + (S_{cap} - S_{Si})T_2I \approx -S_{Si}T_cI \qquad (3.3)$$

where T_c is defined as the microcooler temperature. Therefore, in such a silicon thermoelectric microcooler configuration and to a very good first-order approximation, the overall Peltier cooling rate depends only on the Seebeck coefficient of the silicon substrate, the microcooler temperature, and the applied current.

In addition to volumetric Joule heating in the metal lead, in the silicon substrate, and in the silicon cap, such a parasitic effect will also arise at both the metal contact/silicon cap interface and silicon substrate/ground electrode interface. The interfacial Joule heating at the metal contact/silicon cap can be expressed as

$$q_{contact} = \frac{I^2\rho_c}{A_{contact}} \qquad (3.4)$$

where $A_{contact}$ is the cross-sectional area of the metal contact and ρ_c is the specific electric contact resistivity at the metal contact/silicon cap interface. The influence of interfacial Joule heating and Peltier heating at the silicon substrate/ground electrode interface on the cooling performance of the present microcooler configuration can be neglected as the temperature on the silicon substrate base is held constant.

3.2 Numerical Modeling of Silicon Microcooler and Its Limitations

Numerical modeling developed by Zhang can capture the microcooler cooling performance and temperature field of the silicon microcooler system very well [101].

We follow her approach to do numerical simulation of silicon mcirocooler using ANSYS finite element software. The modeled domains include the silicon substrate, SiN$_x$ layer, the metal lead, the silicon cap, and the metal contact, as shown in Figure 3.1. The thermal-electric elements, Solid 69, are used and densely located around the microcooler as shown in Figure 3.2 where the largest temperature gradient and electric potential gradient are expected to occur. To properly capture the large temperature and voltage gradients and the very small thickness of many of the geometric features, more than 100,000 elements were used to create the model.

(a)

(b)
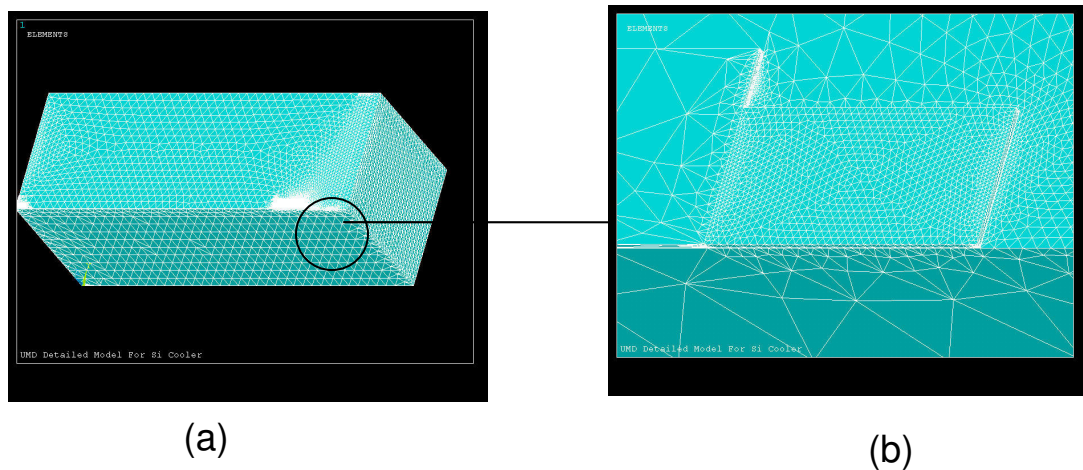
Figure 3.2: Mesh structure for silicon microcooler system: (a) low magnification, and (b) high magnification. Microcooler size is 75 µm × 75 µm, metal lead size is 3.0µm in thickness and 80 µm in width, and SiN$_x$ layer is 0.3 µm in thickness and 80 µm in width.

We found that numerical simulation can capture very detailed temperature and heat flux distributions of silicon microcooler system, however, due to very high aspect ratio for the metal lead, SiN$_x$ passivation layer and the metal contact, that is, for larger silicon microcooles, the metal lead and metal contact will expand to very large size in two directions on silicon substrate while the thickness of SiN$_x$ passivation layer has to keep at only 0.3 µm and the thickness of metal lead, metal

contact and silicon cap have to keep at only 1~3 µm to simulate the real devices developed in the laboratory [43,101]. We find it is extremely difficult to model larger silicon microcooler which correspond to large sizes in metal lead, SiN_x passivation layer, and metal contact. We find that the meshing capability of ANSYS model based on free meshing reported in [101] for detailed structures including metal contact, silicon cap, metal lead and SiN_x is restricted to around 80 µm × 80 µm microcoolers. If the ANSYS mode expands to larger microcooler size, either the computer is shun down automatically due to incompatible mesh structure or CUP running time seems becomes infinite. Therefore, our idea here is to use ANSYS to do some typical case studies for Shakouri's experimental configuration and then use these results to calibrate and validate an analytical modeling. Using these validated analytical modeling, we can expand metal lead, metal contact or SiN_x layer to any size as we want, which is indeed one of the major reasons why we develop analytical model. So in this section we demonstrate some typical temperature profiles and temperature contours using Zhang's ANSYS codes for her silicon microcooler experiment. More detailed parasitic studies such as doping effect, electrical contact resistance effect, and geometric configuration effects such as microcooler size and metal lead size, will be explored in details using our validated analytical model.
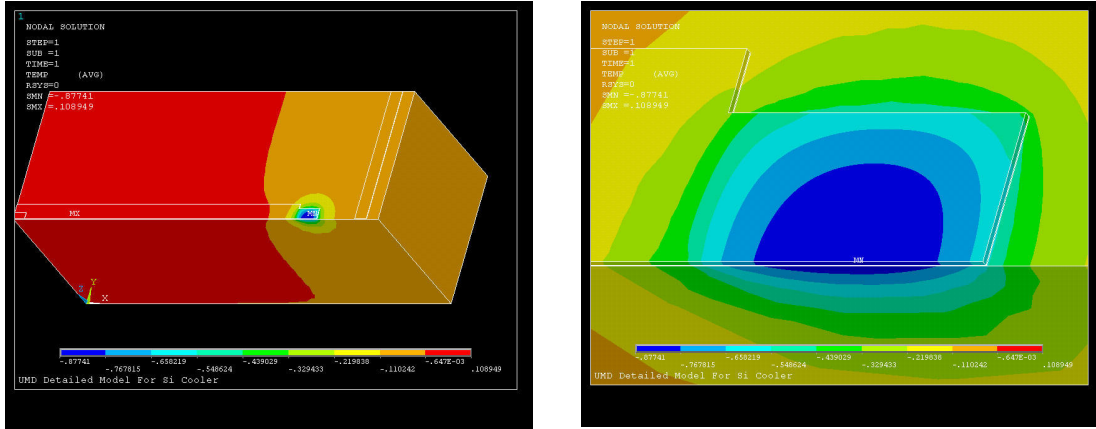
(a)   (b)

Figure 3.3: Temperature contour for silicon microcooler system for (a) low magnification and (b) high magnification. Microcooler size is 75 μm × 75 μm at 0.3A, the metal lead size is 3.0μm in thickness and 80 μm in width, and SiN$_x$ passivation layer is 0.3 μm in thickness and 80 μm in width.
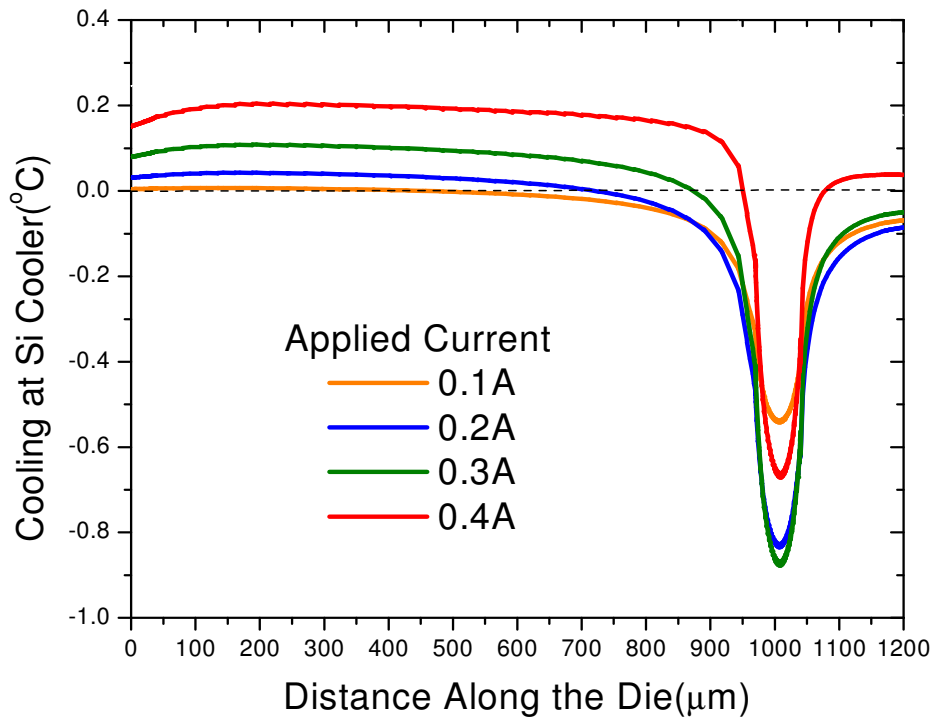
Figure 3.4: Temperature profile on the top of the silicon substrate with 75 μm × 75 μm microcooler at various applied currents.

64

Figure 3.3 shows the temperature contour of the silicon substrate for low magnification and high magnification. We can find there is a cold spot on the silicon substrate as shown in Figure 3.3 (a), on which the microcooler is located. On the left of the microcooler, the temperature is higher, which is due to metal lead Joule heating effect. Figure 3.3 (b) illustrates the temperature distribution around the microcooler, showing that the lowest temperature is around the center of the metal contact. Temperature becomes higher far away from the center of the microcooler.

(a)   (b)
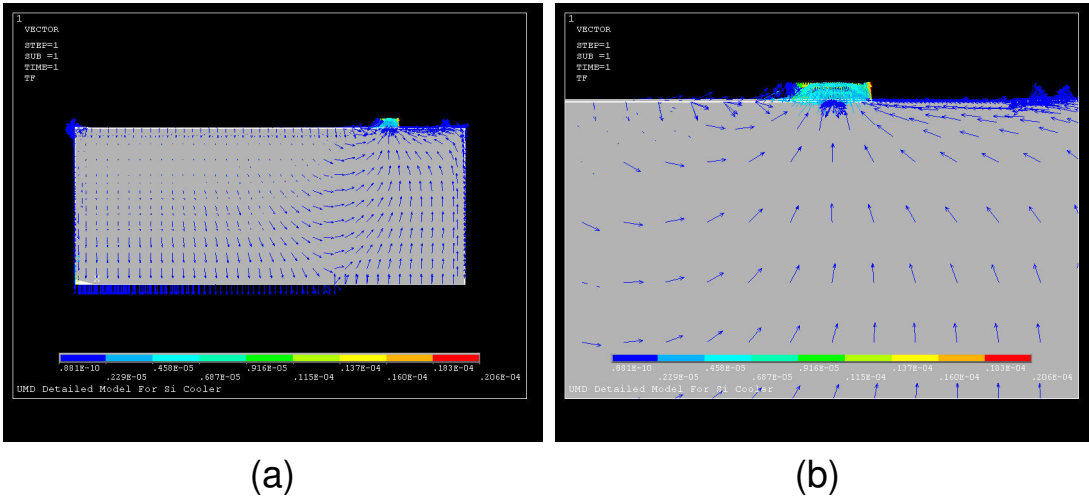
Figure 3.5: Heat flux contribution inside the silicon substrate for (a) low magnification, and (b) high magnification. Microcooler size is 75 μm × 75 μm at 0.3A, metal lead size is 3.0μm in thickness and 80 μm in width, and SiN$_x$ layer is 0.3 μm in thickness and 80 μm in width.

It might be clearer to check the temperature profile on the top of the silicon substrate where metal contact and metal mead are located. As shown in Figure 3.4 the temperature profile when silicon microcooler is activated with a current of 0.1A, 0.2A, 0.3A and 0.4A. We can find the temperature is reduced around the microcooler while the temperature on the metal lead is increased due to Joule heating in the metal

lead. Continuously increasing current from 0 to 0.3A leads to lower temperature on the microcooler but higher temperature on the metal lead.

Figure 3.5 show the heat flux inside the silicon substrate. It is interesting to find that Joule heat in metal lead has two dissipation paths: some of Joule heating in the metal lead will flow to the microcooler while the rest of Joule heating will flow to the bottom of silicon substrate where the heat sink is attached as shown in Figure 3.5(a). In the sequent sections, we will use analytical approach to model this Joule heating effect on cooling performance. Figure 3.5(b) shows the heat flux distribution around the microcooler, indicating that microcooler looks like a sink and absorbs the heat from the substrate to the microcooler.

3.3 Analytical Thermal Modeling

3.3.1 Modified Peltier Cooling Flux

The cooling performance of the silicon microcooler is determined by the balance between the Peltier cooling rate and the component of the parasitic heating rate due to the Joule heating at the metal contact/silicon interface, as well as the heat from the silicon substrate and the metal lead that diffuse to the microcooler zone. Therefore, the net cooling rate on the microcooler can be expressed as:

$$q_{net} = -STI + q_{J,contact} + q_{Si} + q_{lead} \tag{3.5}$$

where the terms on the right-side of Equation (3.5) represent, respectively, the Peltier cooling rate, Joule heating from the electric contact resistance at the metal contact/silicon interface, heat flow into the microcooler due to heat conduction/heat

generation inside the silicon substrate, and heat flow into the microcooler due to heat generation/conduction from the metal lead. Along with the effect of Joule heating in the substrate, the Joule heating at the metal contact/silicon interface and the heat generation and conduction in the metal lead are widely accepted as two major parasitic effects for thermoelectric microcoolers. In the present study, the thermal performance of a silicon microcooler is determined analytically by coupling the solution of the three-dimensional Laplace's equation for thermal diffusion in a silicon substrate subjected to a modified Peltier cooling boundary condition with a one-dimensional solution of heat generation /conduction in the metal lead.
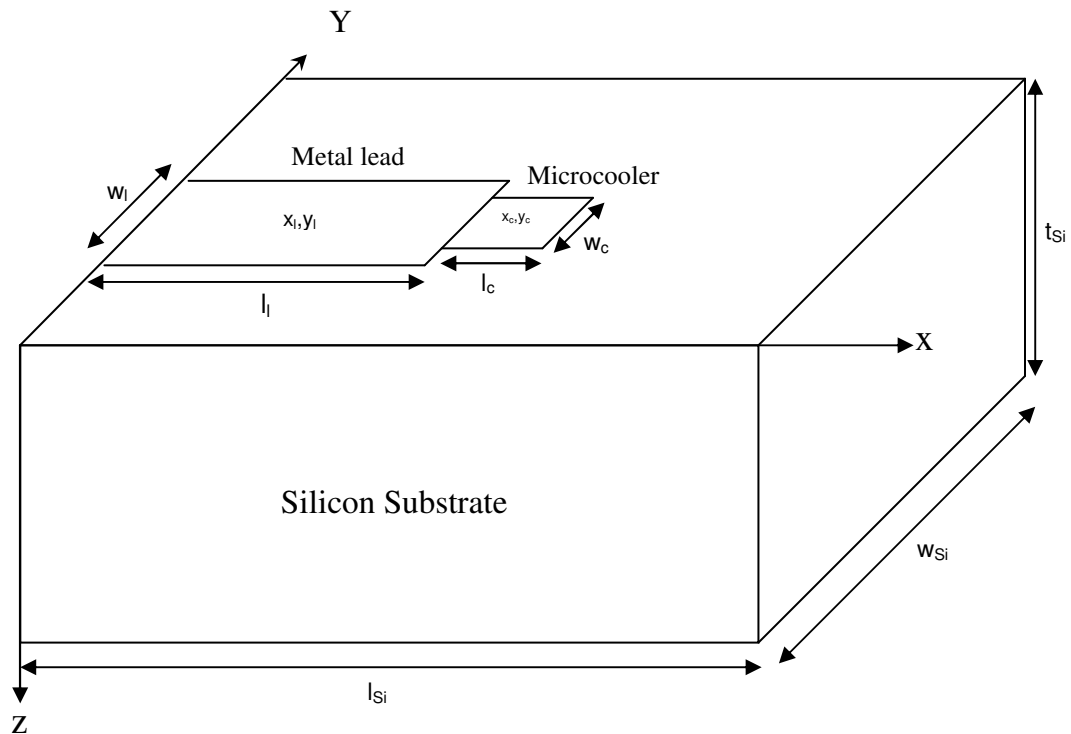
Figure 3.6: Schematic for modeling the silicon microcooler.

3.3.2 3D Analytical Thermal Model for Silicon Microcooler

Determination of the steady-state thermal performance of the silicon microcooler, described in Figure 3.6, requires the solution of the three-dimensional Poison's energy equation for the temperature distribution in a rectangular silicon slab subjected to the influence of Peltier cooling, Peltier heating, and Joule heating associated with electrical current flow through the silicon substrate and the metal lead, i.e.

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} + \frac{q_{Si}'''(x, y, z)}{k_{Si}} = 0 \tag{3.6}$$

where $q_{Si}'''(x,y,z)$ is the non-uniform volumetric heat generation due to Joule heating inside the silicon substrate.

Using the geometry depicted in Figure 3.2, the following boundary conditions can be applied to Eq. (3.6):

$$\frac{\partial T}{\partial z} = -\frac{q_{eff,cooler}''}{k} \quad x_c - 0.5l_c < x < x_c + 0.5l_c, \quad y_c - 0.5w_c < y < y_c + 0.5w_c \tag{3.7}$$

$$\frac{\partial T}{\partial z} = -\frac{q_{eff,lead}''}{k} \quad x_l - 0.5l_l < x < x_l + 0.5l_l, \quad y_l - 0.5w_l < y < y_l + 0.5w_l \tag{3.8}$$

where $q''_{eff,cooler}$ is the effective cooling heat flux over the footprint of the microcooler and $q''_{eff,lead}$ is the metal lead Joule heating that flows directly into the silicon substrate divided by the area of the metal lead footprint.

Unfortunately, solution of Equation (3.6) requires detailed knowledge of the internal heat generation function, $q_{Si}'''(x,y,z)$, resulting from the electric current flow in the silicon substrate and the associated three-dimensional Joule heating pattern. Determination of this function requires a parallel solution of the Poison's equation for

the voltage field. The resulting strongly non-uniform heat generation function can be expected to make Equation (3.6) nearly unsolvable analytically for all but the simplest approximations of q_{Si}'''. However, following conventional thermoelectric modeling procedures, it is possible to define an "allocation" factor, α, which defines the fraction of the Joule heating inside the silicon substrate flowing into the cold side of the thermoelectric circuit, which yields an acceptable approximation for the temperature on the microcooler. The appropriate allocation factor can be determined from an integrated numerical simulation of the thermal and electrical fields and is found to be approximately 0.36 if the largest temperature reduction (or peak cooling) on the microcooler is desired. With this approach, the internal Joule heating in the silicon substrate is replaced with a modified boundary condition on the surface of the micrcooler and the Poison's equation can be transformed into the Laplace's equation as:

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = 0 \qquad (3.9)$$

On the top surface of the silicon substrate, a uniform effective cooling heat flux, $q_{eff,cooler}''$, determined as the combined effect of the Peltier cooling (= STI), the fraction of silicon Joule heating flowing into the microcooler (= $\alpha I^2 R_{Si,e}$), Joule heating due to the electric contact resistance (= $I^2 R_{contact}$), and the heat diffusion directly from the metal lead into the microcooler (= $q_{lead,cooler}$), is assumed to prevail over the microcooler surface. The effective cooling heat flux on the microcooler surface can then be expressed as:

$$q_{eff,cooler}'' = \frac{-STI + I^2 R_{contact} + q_{lead,cooler} + \alpha I^2 R_{Si,e}}{w_c l_c} \qquad (3.10)$$

The effective metal lead heat flux onto the substrate accounts for the majority of the Joule heating in the metal lead and can be expressed as:

$$q''_{eff,lead} = \frac{q_{lead,substrate}}{w_l l_l} \tag{3.11}$$

where $q_{lead,substrate}$ is the metal lead Joule heating flowing into the substrate. The determination of the heat flow in the metal lead, $q_{lead,cooler}$ and $q_{lead,substrate}$, will be derived in a subsequent section.

The top surface of silicon substrate, outside the microcooler area and the metal lead area, is assume to be adiabatic and can be represented by:

$$\frac{\partial T}{\partial z} = 0 \qquad \text{(other areas at } z = 0) \tag{3.12}$$

Along the edges of the silicon substrate, an adiabatic boundary condition is assumed, i.e.

$$\frac{\partial T}{\partial y} = 0 \quad at \ y = 0, \ y = w_{Si}, \quad 0 < x < l_{Si}, 0 \le z \le t_{Si} \tag{3.13}$$

$$\frac{\partial T}{\partial x} = 0 \quad at \ x = 0, \ x = l_{Si}, \quad 0 < y < w_{Si}, 0 \le z \le t_{Si} \tag{3.14}$$

The base of the silicon substrate ($z = t_s$) is assumed to be isothermal, i.e.

$$T = T_A \qquad at \ 0 \le x \le l_{Si}, \ 0 \le y \le w_{Si}, \ z = t_{Si} \tag{3.15}$$

To quantify the internal heat generation in the domain of interest, it is necessary to determine the total electrical resistance of the silicon substrate, $R_{Si,e}$. Applying the electrical-thermal analogy for the current diffusion process in the cylindrical substrate[110], a closed-form equation for the electrical resistance in the rectangular substrate can be expressed as:

70

$$R_{Si,e} = \frac{\rho_{Si}t_{Si}}{\pi r_{Si}^2} + 0.9\left[\frac{\rho_{Si}}{4r_{Si}}(1 - \frac{r_c}{r_{Si}})^{3/2}\right] \tag{3.16}$$

The first term represents the one-dimensional electrical resistance and the second term the three-dimensional spreading resistance. The equivalent radius for the rectangular substrate and the microcooler, are r_{Si} and r_c, respectively. The application of this relation to the rectilinear geometry of the present microcooler requires the introduction of a 0.9 coefficient for the "current spreading" term and the electrical resistances predicted by Equation (3.16) are found to deviate no more than 2% from numerical simulation results when the silicon substrate thickness is larger than 200µm.
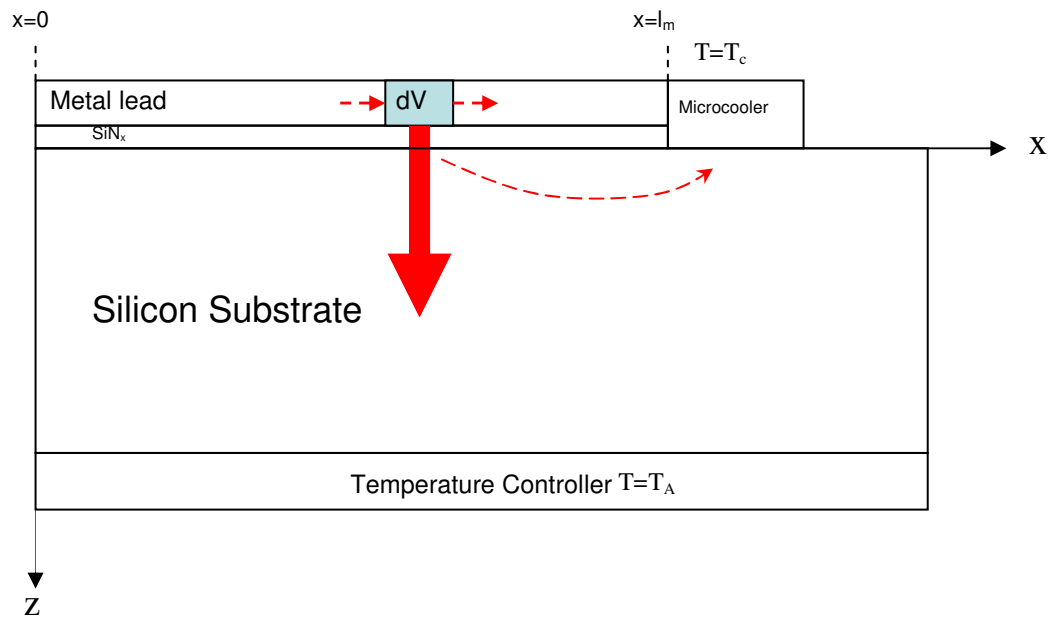
Figure 3.7: Schematic for modeling the heat generation/heat conduction in metal lead.

3.3.3 Analytical model for the metal lead

In the silicon microcooler, the metal lead layer transports electric current to the microcooler to induce the Peltier cooling effect at the interfaces. However, heat conduction through the lead to the region of the microcooler, as well as heat generation (Joule heating) inside the metal lead, could be a major parasitic source of cooling performance degradation. As illustrated in Figure 3.7, there are three heat flow paths for Joule heating generated in the metal lead. It can be anticipated that a majority of the Joule heat will flow into the silicon substrate and on to the temperature-controlled base, but some of this heat will flow laterally in the substrate into the microcooler zone, while some of the Joule heat will flow into the microcooler directly through the metal lead. In this section, we provide an analytical model to describe the parasitic effects of heat transfer related to heat generation/heat conduction inside the metal lead.

A schematic of the metal lead, SiN_x layer, and silicon substrate is shown in Figure 3.7. Consider a long rectangular metal lead of thickness t_m, width w_m, length l_m, electrical resistivity ρ_m and thermal conductivity k_m, separated from the underlying silicon substrate by a SiN_x layer of thickness t_{SiN}, width w_{SiN}, length l_{SiN}, electrical resistivity ρ_{SiN} and thermal conductivity k_{SiN}. With the length of the metal lead much larger than the width and the thickness, the temperature change is dominant in the longitudinal direction and, therefore, the metal lead can be viewed as a thermal fin. As illustrated in Figure 3.7, the heat conduction equation can be derived by examining a control volume of the metal lead of thickness t_m, width w_m, and length dx. Under steady state condition, the conservation of energy requires that the Joule

heating generated in the control volume, dV, is equal to the heat conduction and convection out of the volume as follows:

$$-k_m w_m t_m \frac{dT}{dx}\bigg|_x + J^2 \rho_m w_m t_m dx = -k_m w_m t_m \frac{dT}{dx}\bigg|_{x+dx} + h_{eff}(T-T_A)w_m dx \tag{3.17}$$

where T and T_A are the lead temperature and the silicon base temperature, respectively, J is the electric current density, and h_{eff} is the effective heat transfer coefficient. After taking the limit as $dx \rightarrow 0$, Equation (3.17) is simplified to

$$\frac{d^2 T}{dx^2} + \frac{J^2 \rho_m}{k_m} - \frac{h_{eff}}{k_m t_m}(T-T_A) = 0 \tag{3.18}$$

In order to solve Equation (3.18) two appropriate boundary conditions must be applied at both sides of the metal lead:

(a) At the one end of the metal lead (x = 0), the adiabatic boundary condition is applied:

$$\frac{dT}{dX}\bigg|_{x=0} = 0 \tag{3.19}$$

(b) At the other end of the metal lead (x = l_m), we assume the temperature to equal that of the microcooler:

$$T\bigg|_{x=L_m} = T_{cooler} \tag{3.20}$$

In this fin analysis an effective heat transfer coefficient, $h_{eff,}$ will be employed to represent the heat loss (albeit by conduction) from the bottom surface of the metal lead into the silicon substrate and given by:

$$h_{eff} = \frac{1}{(R_{SiN_x,t} + R_{Si,t})l_m w_m} \tag{3.21}$$

73

where $R_{SiNx,t}$ and $R_{Si,t}$ are the thermal resistance of SiN_x layer and silicon substrate, respectively.

Heat conduction inside the SiN_x passivation layer is assumed to be one dimensional (1-D) and perpendicular to the bottom surface of the SiN_x. The corresponding thermal resistance is given by:

$$R_{SiN_x,t} = \frac{t_{SiN_x}}{k_{SiN_x} l_{SiN_x} w_{SiN_x}} \tag{3.22}$$

As the length of the metal lead is much larger than its cross-section, the heat conduction downwards to the silicon substrate is assumed to take the form of two-dimensional "spreading," and the average value of this thermal resistance can be calculated as [111]:

$$R_{Si,t} = \frac{1}{2k_{Si}l_m}\{\alpha + \frac{2}{\pi^3 \varepsilon^2} \sum_{n=1}^{\infty} \frac{\sin^2(n\pi\varepsilon)}{n^3} \tanh(n\pi\alpha)\} \tag{3.23}$$

Where $\quad \alpha = \frac{2t_{Si}}{w_{Si}}, \quad \varepsilon = \frac{w_m}{w_{Si}}$

Therefore, the effective heat transfer coefficient, h_{eff}, can be expressed as:

$$h_{eff} = \left(\frac{w_m}{2k_{Si}}[\alpha + \frac{2}{\pi^3 \varepsilon^2} \sum_{n=1}^{\infty} \frac{\sin^2(n\pi\varepsilon)}{n^3} \tanh(n\pi\alpha)] + \frac{t_{SiN_x}}{k_{SiN_x}} \right)^{-1} \tag{3.24}$$

The temperature distribution along the metal lead can then be solved as:

$$T = T_A + \Delta T \frac{\cosh(\sqrt{\frac{h_{eff}}{k_m t_m}}x)}{\cosh(\sqrt{\frac{h_{eff}}{k_m t_m}}l_m)} + \frac{I^2 \rho_m}{w_m^2 t_m h_{eff}}(1 - \frac{\cosh(\sqrt{\frac{h_{eff}}{k_m t_m}}x)}{\cosh(\sqrt{\frac{h_{eff}}{k_m t_m}}l_m)}) \tag{3.25}$$

where ΔT is the temperature reduction at the silicon microcooler ($\Delta T = T_{cooler} - T_A$).

Applying Fourier's law together with Equation (3.25) at the lead tip ($x = l_m$), it is possible to determine the total heat transfer from the metal lead directly into the silicon microcooler, $q_{lead,\,cooler}$, as:

$$q_{lead,cooler} = -k_m w_m t_m \frac{dT}{dx}\bigg|_{x=l_m}$$

$$= I^2 \rho_m \sqrt{\frac{k_m}{w_m^2 t_m h_{eff}}} \tanh(\sqrt{\frac{h_{eff}}{k_m t_m}} l_m) - \Delta T \sqrt{w_m^2 t_m k_m h_{eff}} \tanh(\sqrt{\frac{h_{eff}}{k_m t_m}} l_m)$$

(3.26)

It is clear from Equation (3.26) that both the heat conduction due to the Joule heating inside the metal lead (the first term) and the heat conduction due to the temperature gradient (the second term) will reduce the effective cooling rate on the microcooler and thus degrade the cooling performance. Similarly, the heat transfer from the bottom surface of the metal lead into the silicon substrate, $q_{lead,substrate}$, can be calculated as:

$$q_{lead,substrate} = \int_{x=0}^{x=l_m} h_{eff} w_m l_m (T - T_A) dx$$

$$= I^2 \left(\rho_m \frac{l_m}{t_m w_m} - I^2 \rho_m \sqrt{\frac{k_m}{w_m^2 t_m h_{eff}}} \tanh(\sqrt{\frac{h_{eff}}{k_m t_m}} l_m) \right) + \Delta T \sqrt{w_m^2 t_m k_m h_{eff}} \tanh(\sqrt{\frac{h_{eff}}{k_m t_m}} l_m)$$

(3.27)

It should be noted that after flowing into the silicon substrate, some of $q_{lead,\,substrate}$ will flow laterally into the microcooler and the rest (the majority) will flow downwards to the silicon substrate base and become part of the overall heat diffusion in the substrate. In the next section, $q_{lead,cooler}$ and $q_{lead,substrate}$ will be coupled with 3D microcooler model so that the overall contribution of heat conduction/heat generation inside the metal lead to the cooling performance could be completely accounted.

3.3.4 Analytical Solution for Temperature Field

The separation of variables method was employed to find the solution to the temperature field of the microcooler. Using the given boundary conditions and expressing the heat flux distribution function as double Fourier series, the appropriate coefficients can be selected. Thus, the analytical solution for the temperature distribution in the silicon substrate can be obtained by solution of the governing equation, Eq. (3.9), as:

$$T(x,y,z) = T_A - [(\frac{q_{eff,cooler}''}{k_{Si}})(\frac{l_c w_c}{l_{Si} w_{Si}}) + (\frac{q_{eff,lead}''}{k_{Si}})(\frac{l_l w_l}{l_{Si} w_{Si}})](z - t_{Si}) \tag{3.28}$$

$$- \sum_{m=1}^{\infty} [(\frac{q_{eff,cooler}''}{k_{Si}}) \frac{2 l_{Si} w_c}{m^2 \pi^2 w_{Si}} \psi_{m0} + (\frac{q_{eff,lead}''}{k_{Si}}) \frac{2 l_{Si} w_l}{m^2 \pi^2 w_{Si}} \chi_{m0}] \cos(\alpha_m x) \frac{\sinh[\alpha_m (z - t_{Si})]}{\cosh(\alpha_m t_{Si})}$$

$$- \sum_{n=1}^{\infty} [(\frac{q_{eff,cooler}''}{k_{Si}}) \frac{2 w_{Si} l_c}{n^2 \pi^2 l_{Si}} \psi_{0n} + (\frac{q_{eff,lead}''}{k_{Si}}) \frac{2 w_{Si} l_l}{n^2 \pi^2 l_{Si}} \chi_{0n}] \cos(\beta_n y) \frac{\sinh[\beta_n (z - t_{Si})]}{\cosh(\beta_n t_{Si})}$$

$$- \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} [(\frac{q_{eff,cooler}''}{k_{Si}}) \psi_{m0} \psi_{0n} + (\frac{q_{eff,lead}''}{k_{Si}}) \chi_{m0} \chi_{0n}] \frac{4}{mn \pi^2 \gamma_{mn}} \cos(\alpha_m x) \cos(\beta_n y) \frac{\sinh[\gamma_{mn}(z - t_{Si})]}{\cosh(\gamma_{mn} t_{Si})}]$$

where

$$\alpha_m = \frac{m\pi}{l_{Si}}$$

$$\beta_n = \frac{n\pi}{w_{Si}}$$

$$\gamma_{mn} = \sqrt{\alpha_m^2 + \beta_n^2}$$

$$\psi_{m0}(m) = \sin \frac{m\pi(x_c + 0.5 l_c)}{l_{Si}} - \sin \frac{m\pi(x_c - 0.5 l_c)}{l_{Si}}$$

$$\psi_{0n}(n) = \sin \frac{n\pi(y_c + 0.5 w_c)}{w_{Si}} - \sin \frac{n\pi(y_c - 0.5 w_c)}{w_{Si}}$$

$$\chi_{m0}(m) = \sin \frac{m\pi(x_l + 0.5 l_l)}{l_{Si}} - \sin \frac{m\pi(x_l - 0.5 l_l)}{l_{Si}}$$

$$\chi_{0n}(n) = \sin \frac{n\pi(y_l + 0.5 w_l)}{w_{Si}} - \sin \frac{n\pi(y_l - 0.5 w_l)}{w_{Si}}$$

It is seen that the solution is in the form of an infinite double cosine series and, in actual calculation, it is apparent that we can sum only a finite number of terms. Consequently, the accuracy of the calculation is associated with the number of terms summed. It was found that the number of terms required for the solution to converge to within a desired degree of accuracy is related to the geometry of silicon substrate, the metal lead and the microcooler. For all the calculations, the infinite series are truncated at m = n = 300, beyond which a further increase in the number of terms has no influence on the results.

In the present analysis attention is focused on the determination of the temperature reduction achieved by the microcooler, with respect to the temperature of the silicon substrate base. However, in order to calculate the temperature reduction with consideration of the metal lead effect, Eqs. (3.26), (3.27) and (3.28) need to be integrated and the corresponding analytical solution is given:

$$\Delta T = \left[-STI + I^2 R_{cont} + \alpha I^2 R_{Si} + I^2 \sqrt{\frac{\rho_m^2 k_m}{w_m^2 t_m h_{eff}}} + \left(I^2 R_{lead} - I^2 \sqrt{\frac{\rho_m^2 k_m}{w_m^2 t_m h_{eff}}} \right) \frac{S_l}{S_c} \right] \tag{3.29}$$

$$\times \frac{S_c}{k_{si} + (S_c - S_l)\sqrt{w_m^2 t_m k_m h_{eff}}}$$

where α = 0.36 for peak cooling as previously noted and

$$S_{i=c,l} = \frac{t_{si}}{l_{si} w_{si}} + \sum_{m=1}^{\infty} \frac{2 l_{si}}{m^2 \pi^2 w_{si} l_i} \psi_{m0} \cos(\alpha_m x_c) \tanh(\alpha_m t_{si})$$

$$+ \sum_{n=1}^{\infty} \frac{2 w_{Si}}{n^2 \pi^2 w_i l_{si}} \psi_{0n} \cos(\beta_n y_c) \tanh(\beta_n t_{si})$$

$$+ \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{4}{mn \pi^2 \gamma_{mn} w_i l_i} \psi_{m0} \psi_{0n} \cos(\alpha_m x_c) \cos(\beta_n y_c) \tanh(\gamma_{mn} t_{si})$$

77

Equation (3.29) includes all cooling and heating effects in the silicon microcooler system: $-STI$ is the Peltier cooling rate, I^2R_{cont} the Joule heating from the electric contact resistance, $\alpha I^2 R_{Si,e}$ the silicon Joule heating flowing into the microcooler.

$I^2 \sqrt{\dfrac{\rho_m^2 k_m}{w_m^2 t_m h_{eff}}}$ is the metal lead Joule heating flowing into the microcooler through

the lead directly, and $\left(I^2 R_{lead} - I^2 \sqrt{\dfrac{\rho_m^2 k_m}{w_m^2 t_m h_{eff}}} \right) \dfrac{S_l}{S_c}$ is the metal lead Joule heating flowing into the microcooler laterally through the silicon substrate. The heat flow into the microcooler due to heat conduction (temperature gradient) between the metal lead and the microcooler can be quantitatively calculated and equal to $\Delta T \sqrt{w_m^2 t_m k_m h_{eff}}$, and its effect on the cooling performance is included in Equation (3.29) in the term of

$\left[\dfrac{S_c}{k_{Si}} - \dfrac{S_l}{k_{Si}} \right] \sqrt{w_m^2 t_m k_m h_{eff}}$. S_l and S_c are the shape factors determined by the geometry of silicon substrate, microcooler and metal lead.

3.3.5 Numerical model for microcooler system

A three-dimensional thermal-electric numerical simulation as described by Zhang [101], involving the determination of both the electric and thermal fields resulting from the application of an electric current to the silicon substrate was used to validate the analytical model. Joule heating inside the metal lead and the silicon substrate is accomplished automatically through the electric-thermal coupling in the finite element model, using solid 69 elements. The finite element simulator, ANSYS[TM], was used in this study with a total element number of more than 100,000. The thermo-electric elements are densely located around the microcooler where the largest temperature gradient is expected to occur. The structure of the modeled domains for the silicon microcooler system, which includes the silicon substrate, SiN_x layer, metal lead, metal contact, and silicon cap, is similar to that shown in Figure 3.1.

78

The geometry and the materials properties for silicon microcooler system are listed in Table 3.1.

Table 3.1: Geometry and material properties used for analytical modeling.

	Dimension (µm) (L×W×H)	Thermal Conductivity (W/mK)	Electrical resistivity (µΩ.m)
Silicon substrate	2000×2000×500	110	11~400[a]
Silicon cap	20×20×0.3 ~ 100×100×0.3	110	~11
Metal contact	20×20×3 ~ 100×100×3	300	0.0288
Metal lead (Au)	1000×20×3 ~ 1000×200×3	300	0.0288
SiN$_x$	1000×20×0.3 ~ 1000×200×0.3	30	>10^{18}

[a] The electrical resistivity of silicon substrate depends on doping concentration.

3.4 Results and Discussion

3.4.1 Experimental data and model validation

Preliminary experimental results of silicon microcooler cooling performance at room temperature are shown in Figure 3.8 for microcooler sizes ranging from 40µm×40µm to 75µm×75µm with 500µm thick silicon substrate. For each microcooler size, the measured temperature reduction below ambient temperature (hence below the temperature of the silicon substrate base) is seen to follow a characteristic dependence on applied current. Reflecting the competing contributions of Peltier cooling, with a linear dependence on electric current, and Joule heating, with a quadratic dependence on electric current, the microcooler temperature decreases with applied current until a minimum is reached and then rises back towards a zero cooldown. The lowest microcooler temperature, or the maximum temperature reduction, is achieved by the smallest microcooler size of 40µm×40µm, with approximately 1.1°C net cooling under the applied current of 0.2A. With

increasing microcooler sizes, the cooling performance degrades and the optimum current increases by a modest amount. As will be shown in the next sections, under the ideal case – no electric contact resistance and no metal lead effect, the maximum achievable temperature reduction on the microcooler is independent of microcooler sizes. The fact that the smallest microcooler demonstrates the largest temperature reduction suggests that there exists a large parasitic Joule heating effect from the electric contact resistance and/or the metal lead. Under such non-ideal conditions, because the smaller microcooler requires less current to achieve its maximum cooling, there is less Joule heating from the electric contact resistance and metal lead and thus the overall cooling is larger. Using data extraction technique we found the specific electric contact resistance for these fabricated silicon microcoolers varies from microcooler to microcooler ranging between 3×10^{-6} $\Omega.cm^2$ and 8×10^{-6} $\Omega.cm^2$, somewhat larger than the typical average value of 1×10^{-6} $\Omega.cm^2$ reported in [101], and the boron doping concentration is estimated to be around $2.5 \times 10^{19} cm^{-3}$ in the silicon substrate.
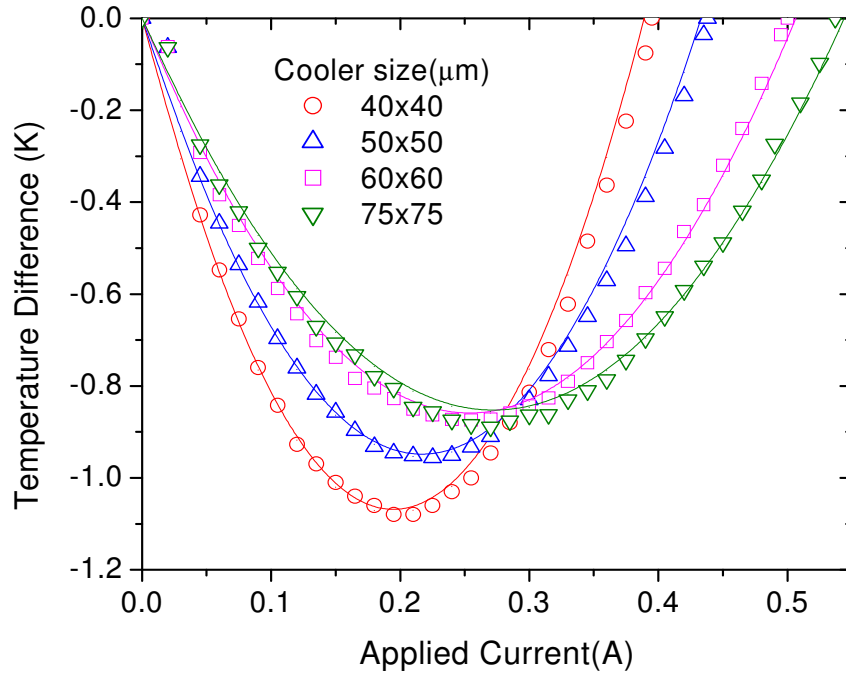
Figure 3.8: Comparison between analytical predictions with experimental data for cooling performance at 25°C. (symbols: experimental data, solid lines: analytical modeling results)
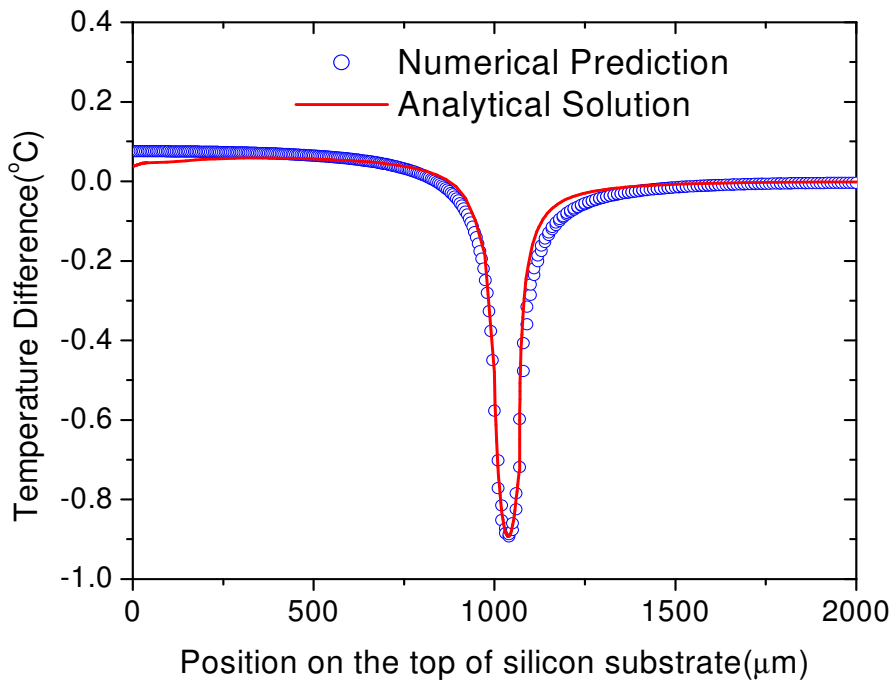
Figure 3.9: Comparison between analytical solutions and numerical predictions for temperature difference profile along the centerline of the top surface of silicon substrate at the applied current of 0.25A at 25°C.
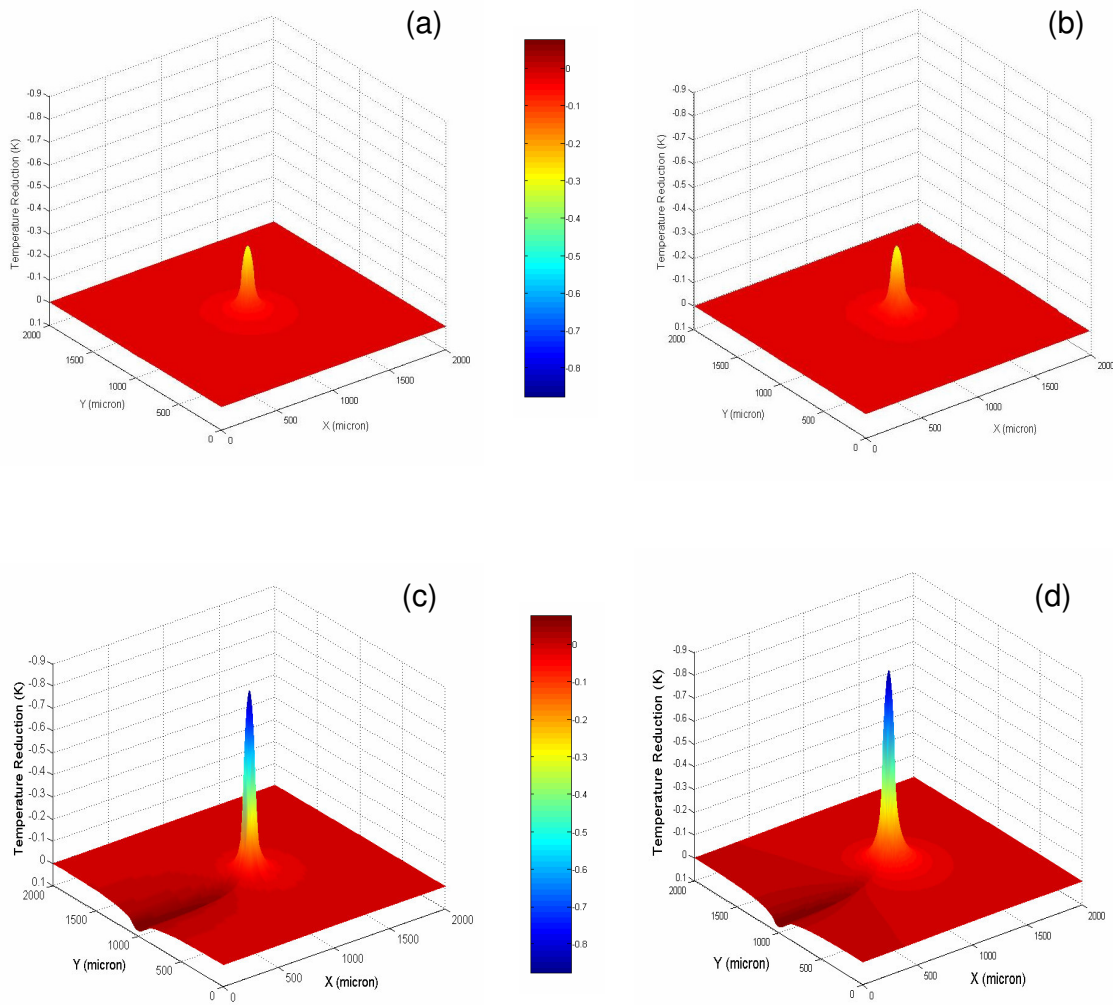
Figure 3.10: Comparison between analytical solutions and numerical predictions for temperature difference contour on the top surface of silicon substrate at 25$^{\mathrm{o}}$C. At the applied current of 0.05A: (a) analytical, (b) numerical; At the applied current of 0.25A: (c) analytical, (d) numerical.

The analytical model is first evaluated by comparing the calculated temperature reductions using Equation (3.29) with the experimental data for the peak cooldowns on the microcooler surfaces. As illustrated in Figure 3.8, the temperature reductions on the microcoolers calculated from analytical solutions are in good agreement with the experimental data across the four microcooler sizes and the range of applied currents. The slight discrepancy between measured data and calculated results might be due to uncertainty of thermal and electrical properties of metal lead thin film and/or additional cooling power loss mechanisms which are not included in our analytical model such as heat conduction through the thermocouple tip. In addition to the comparison with the experimental data, in Figure 3.9 the analytical solutions are also compared with the numerical results for 75µm×75µm microcooler under the applied current of 0.25A and with the specific contact resistance of 8×10^{-6} $\Omega.cm^2$. It can be seen that the analytical results agree very well with numerical results for the temperature difference profile. The temperature on the left side of silicon substrate is higher than that on the right side, which is due to Joule heating effect of the meat lead. Furthermore, under the same condition, the surface contours of analytical temperature difference on the top surface of the substrate, calculated with Equation (3.29) with the applied currents of 0.05A and 0.25A, are seen in Figure 3.10 to compare very well with the numerical simulation, with less than $0.05^{\circ}C$ difference at the microcooler center, and also to clearly display the highly-localized thermoelectric cooling around the microcooler and the self-heating effect of the metal lead at the higher current (Figure 3.10(c) and (d)). These comparisons thus provide confidence in the present analytical modeling methodology.

3.4.2 Silicon Microcooler at Elevated Temperature

In anticipation of the application of these microcoolers to the thermal management of microprocessor hot spots operating in the range of 100$^{\mathrm{o}}$C, the validated analytical model, with the embedded temperature dependence of electrical resistivity, thermal conductivity, and Seebeck coefficient based on values reported for single-crystal silicon [112,113], was used to predict the maximum achievable temperature reduction and parametric sensitivities of such thermoelectric microcoolers at 100$^{\mathrm{o}}$C. The 60µm×60µm microcooler, operating under four distinct 4 conditions: (1) an ideal case without any parasitic effects, (2) a non-ideal case with only Joule heating from electric contact resistance at metal contact/silicon cap interface, (3) a non-ideal case with only heat conduction and generation from the metal lead, and (4) a non-ideal case with both electric contact resistance effect and metal lead effects, was used as the test vehicle. The specific electric contact resistance for the microcooler is assumed to be the typical average value of 1×10^{-6} $\Omega.\mathrm{cm}^2$, the doping concentration is assumed to be $2.5\times10^{19}\mathrm{cm}^{-3}$ in the silicon substrate, and the metal lead is assumed to be a gold thin film with 3.0µm in thickness and 70µm in width.

The results displayed in Figure 3.11 reveal that, in the absence of parasitic effects, the silicon microcooler with the described configuration could achieve a maximum temperature reduction of 6.2$^{\mathrm{o}}$C on the microcooler at the optimum current of 0.9A. If Joule heating from the electric contact resistance is included, the maximum temperature reduction decreases to 4.6$^{\mathrm{o}}$C at the optimum current of 0.8A. If only the parasitic effects of the metal lead – heat generation and heat conduction –

are included, the maximum temperature reduction on the microcooler falls to about 4.7$^\text{o}$C at the optimum current of 0.70A. If both electric contact resistance effect and metal lead effects are included, there is a 3.6$^\text{o}$C maximum temperature reduction on the microcooler at the optimum current of 0.6A. In comparison with the ideal case, the parasitic effects from the electric contact resistance and metal lead result in 43% reduction in the maximum cooling temperature on the microcooler. Figure 3.12 displays the role played by the specific parasitic effects in the present microcooler configuration for Case 4. It is valuable to examine the magnitude of such parasitic effects at the optimum current of 0.6A with which the maximum temperature reduction is achieved: while it can be seen that Joule heating in the silicon is the largest parasitic heat source for the microcooler, accounting for 34% of the Peltier cooling rate, Joule heating due to the electric contact resistance (10%), and the parasitic contribution of the metal lead (13%), can also substantially degrade the net benefit of the Peltier cooling, leaving only 42% of the Peltier cooling rate as an effective net cooling power on the microcooler. To better define the role of doping concentration and the contribution of these parasitic phenomena to the performance of a silicon microcooler, subsequent sections will individually examine the effects of silicon doping concentration, electric contact resistance, and heat generation/ conduction in the metal lead.
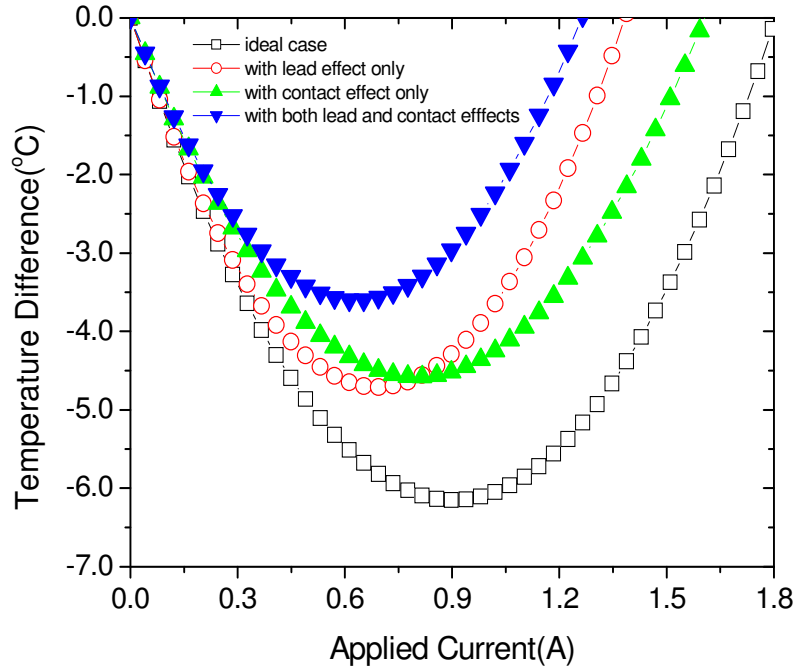
Figure 3.11: Variation of silicon microcooler performance with applied current at 100°C.
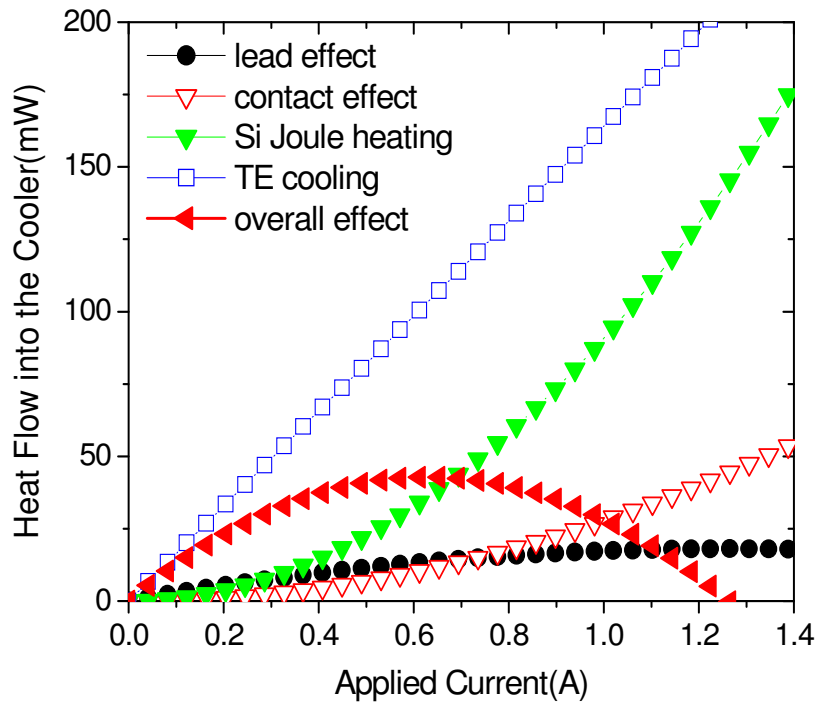
Figure 3.12: Comparison of various heat flows into the microcooler at 100°C.

3.4.3. Effect of Silicon Dopant Concentration

Dopant concentration in silicon can have a profound influence on silicon microcooler performance, strongly affecting the Seebeck coefficient and the electrical resistivity, but only modestly affecting the thermal conductivity if the operating temperature is at room temperature or above [114]. The relationship between Seebeck coefficient and doping concentration can be derived from solid state theory and shown to display a complex inverse relationship which has been experimentally corroborated. Chapman *et al* have shown that the electrical resistivity of silicon decreases with increasing dopant concentration, due to higher carrier concentration [113]. Combining these effects Figure 3.13 shows the dependence of the maximum temperature reduction at a 60µmx60µm microcooler on the boron doping concentration, for the specific contact resistance ranging from 1.0×10^{-9} $\Omega.cm^2$, representing a nearly-ideal interface, to 1.0×10^{-5} $\Omega.cm^2$, typical of a laboratory deteriorated interface. Figure 3.13 clearly demonstrates that cooling performance, as measured by the peak temperature reduction on the microcooler, strongly depends on the doping concentration and that for each specific contact resistance there is an optimum doping concentration. It is known that with increasing doping concentration the electrical resistivity decreases and, as a consequence, results in less Joule heating in the silicon substrate. Unfortunately, the Seebeck coefficient of silicon also decreases with increasing doping concentration, which leads to less Peltier cooling. The competition between these two factors results in an optimum doping concentration at which the maximum cooling performance could be obtained. It is interesting to note that progressive reductions in the specific contact resistance yield

87

greater cooldowns at larger doping concentrations, decreasing from $2.1^{\circ}C$ at the doping concentration of $4.0\times10^{18} cm^{-3}$ for the specific contact resistance of 1.0×10^{-5} Ωcm^2 to $4.6^{\circ}C$ at the doping concentration of $2.0\times10^{19} cm^{-3}$ for the nearly-ideal specific contact resistance of 1.0×10^{-9} Ωcm^2. It also suggests that selection of doping concentration to maximize thermoelectric cooling application has to consider the parasitic effects, and lower dopped silicon materials are preferable if larger parasitic effects exist in the microcooler. Figure 3.13 also shows that reductions in the specific contact resistance below 1×10^{-7} $\Omega.cm^2$ are unlikely to produce further cooldown improvements.
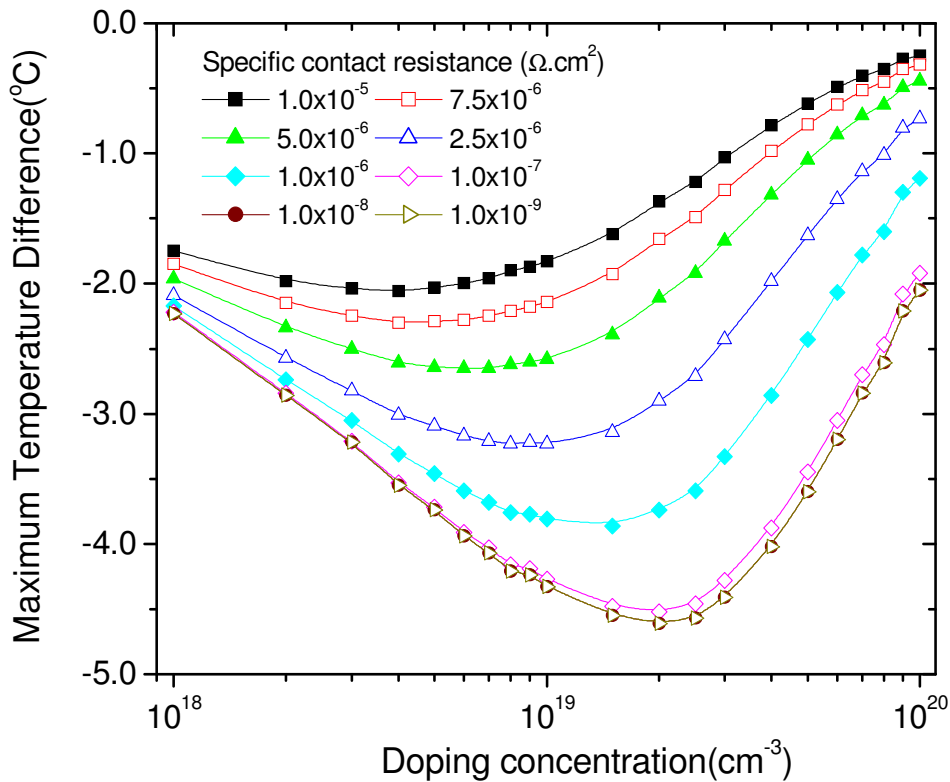
Figure 3.13: Dependence of maximum temperature difference on doping concentration for various specific contact resistances at $100^{\circ}C$.

3.4.4 Effect of Metal Lead

The metal lead is used to send the electric current into the microcooler but, at the same time, it can deteriorate the thermoelectric cooling performance through heat conduction and heat generation. The effect of the metal lead on thermoelectric cooling performance is determined by the geometry and the electrical/thermal properties of the metal lead. In this study a 60µm×60µm microcooler with a 3µm thick gold thin film is used as the vehicle to explore the sensitivity of cooling performance to the metal lead geometry, with an assumed specific contact resistance of 1.0×10^{-6} Ωcm^2. As shown in Figure 3.14 (a), increasing the lead width initially results in an improvement in cooling performance until a temperature reduction of $3.6^\circ C$ is reached at the lead width of 60µm~80µm. For larger widths, the cooling deteriorates. Please note that even with the optimized lead dimension heat flow into the micrcooler due to the metal lead effect still causes around $1.0^\circ C$ decrease in the cooldown on the micrcooler in comparison with a $4.6^\circ C$ temperature reduction if the metal lead effect is completely removed (Figure 3.11).

To understand the mechanism for this variation it is helpful to examine the magnitude of heat flow into the microcooler due to metal lead Joule heating and that due to the heat conduction (temperature gradient) between the microcooler and the lead. Figure 3.14 (b) shows the heat flow into the microcooler induced by Joule heating in the metal lead, and Figure 3.14 (c) illustrates the heat diffusion through the metal lead directly into the microcooler due to the temperature gradient. As can be seen in Figure 3.14 (c), with an increase in the lead width, from 20 µm to 200 µm, more heat will diffuse into the microcooler through the cold end of the lead. On the

other hand, as the lead width and thus cross-sectional area increase, the lead electrical resistance decreases, generating less Joule heating and inducing less associated heat flow to the microcooler, either through the metal lead directly or laterally through the silicon substrate, as illustrated in Figure 3.14 (c). Combining these two effects, as in Figure 3.14 (d), reveals that for the previously determined optimum current of 0.6A at the lead width of 60~80μm, net heat flow into the microcooler is minimized, yielding the largest temperature reduction achievable at the microcooler. Therefore, with variation of the metal lead geometry, the changes of heat generation and heat conduction in the metal lead move at an opposite direction and their parasitic effects on cooling performance can not be minimized at the same time. An optimum metal lead geometry could be found through the trade-off between the influences of heat generation and heat conduction on the cooling performance if the thermal and electrical properties of the metal lead have been fixed.
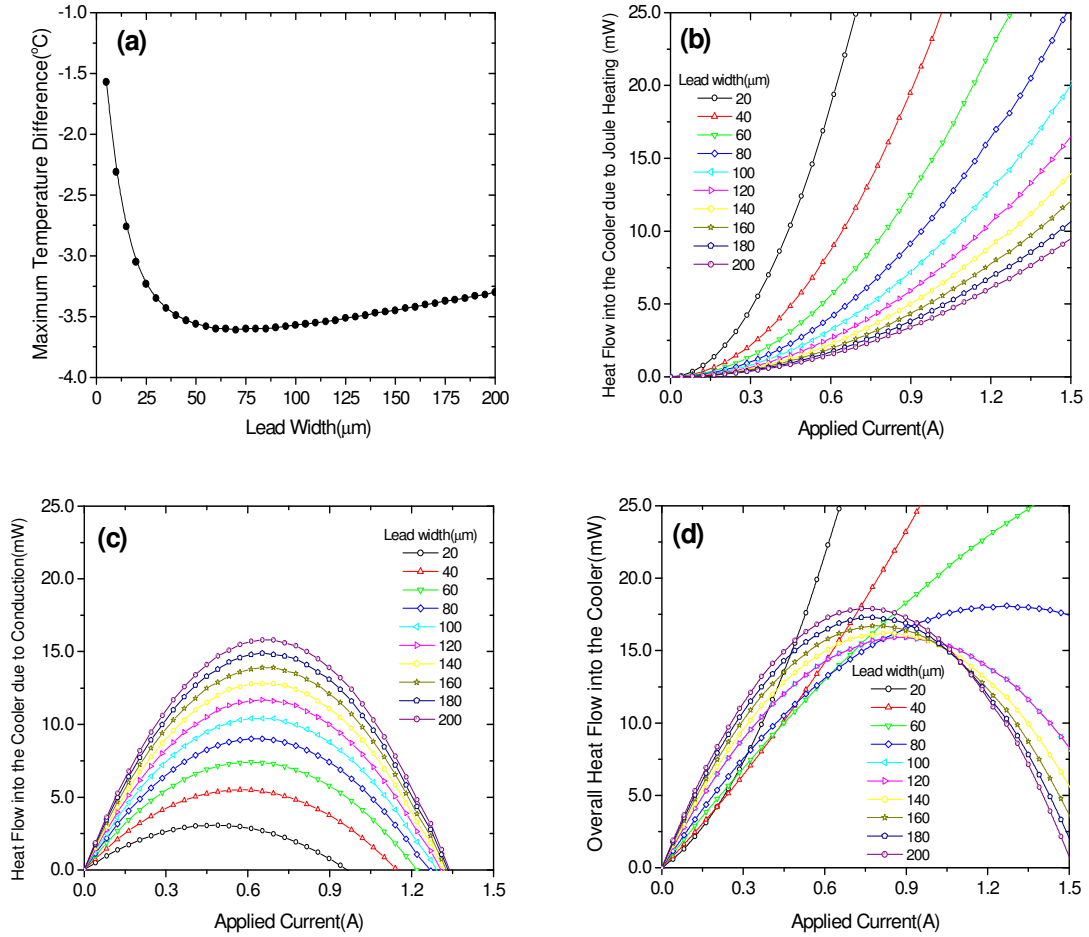
Figure 3.14: Influence of geometry on the metal lead effect for 60μm×60μm microcooler at 100°C: (a) maximum temperature difference on the microcooler, (b) heat flow into the micrcooler due to Joule heating, (c) heat flow into the microcooler due to conduction (temperature gradient), and (d) overall heat flow into the microcooler.

3.4.5 Cooling Potential of Silicon Microcooler

As is abundantly clear from the previous sections, with careful selection of doping concentration, Peltier cooling can be maximized and silicon Joule heating minimized in the silicon substrate. Moreover, the present state-of-the-art in thin film processing makes it possible to reduce the electric specific contact resistance to less than 1×10^{-7} Ω.cm^2 and thus almost eliminate Joule heating at the metal contact/silicon interface [115-117]. When such microcoolers are integrated with actual chip packages, power could be delivered directly to the microcoolers, rather than thru an attached lead, thus removing the deleterious effect of the metal lead. It is, therefore, interesting to evaluate the cooling potential of such silicon microcoolers when these parasitic effects are completely eliminated.

Figure 3.15 shows the maximum attainable temperature reduction on the microcooler, at an operating temperature of 100°C, for various doping concentrations with the microcooler size ranging from 20µm×20µm to 100µm×100µm. It can be seen in Figure 3.15(a) that, over the entire doping range, the maximum temperature reduction on the microcooler is nearly twice the temperature reduction predicted by a traditional one-dimensional thermo-electric analysis using Equation (1.2), as also reported for room-temperature silicon microcooler operation. The highest maximum temperature reduction of 6.2^{o}C is achieved at a doping concentration of 2.5×10^{19}cm^{-3}, and is independent of microcooler size. However, as shown in Figure 3.15(b), smaller microcoolers do achieve the optimal performance at lower currents. In Figure 3.15(a) the maximum average temperature reduction over the entire microcooler surface (average cooling), obtained by re-calibrating the analytical model with an allocation

factor, α, of 0.43 and modified shape factors of S_c and S_l based on surface integral, is also included for comparison. It is found the maximum average cooling is approximately 30% lower than the maximum peak cooling, but still 34% larger than that values predicted using the one-dimensional model. This thermal enhancement is related to the combined contribution of thermal spreading and electric current spreading from the discrete microcooler into the larger substrate.
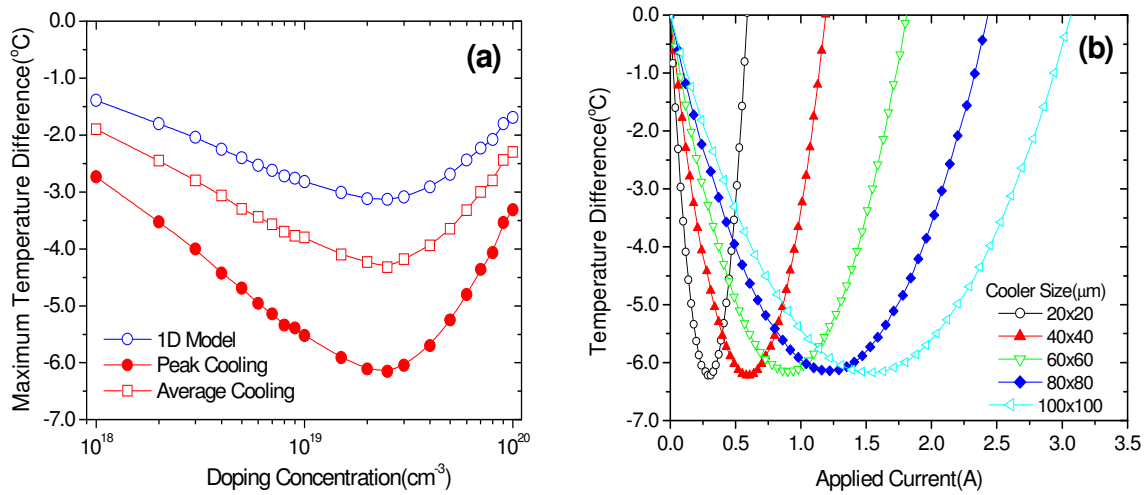
Figure 3.15: (a) Variation of maximum temperature difference with the doping concentration for the ideal case and (b) Dependence of temperature difference on the applied current for different microcooler sizes at 100ºC.

One of the main advantages of silicon microcoolers is the very high cooling heat flux made possible by the high power factor for silicon. As with any thermoelectric cooler, the maximum cooling flux is achieved at a negligibly small temperature reduction, while the greatest temperature reduction is achieved with negligibly small heat flux. For the present microcooler configuration, Figure 3.16 shows that the maximum cooling heat flux attains a predicted maximum value of 1k W/cm^2 for 100µm×100µm microcooler and 6k W/cm^2 for 20µm×20µm microcooler.

These results support the expectation that silicon microcoolers provide a very promising approach to high heat-flux spot cooling in silicon microprocessors.
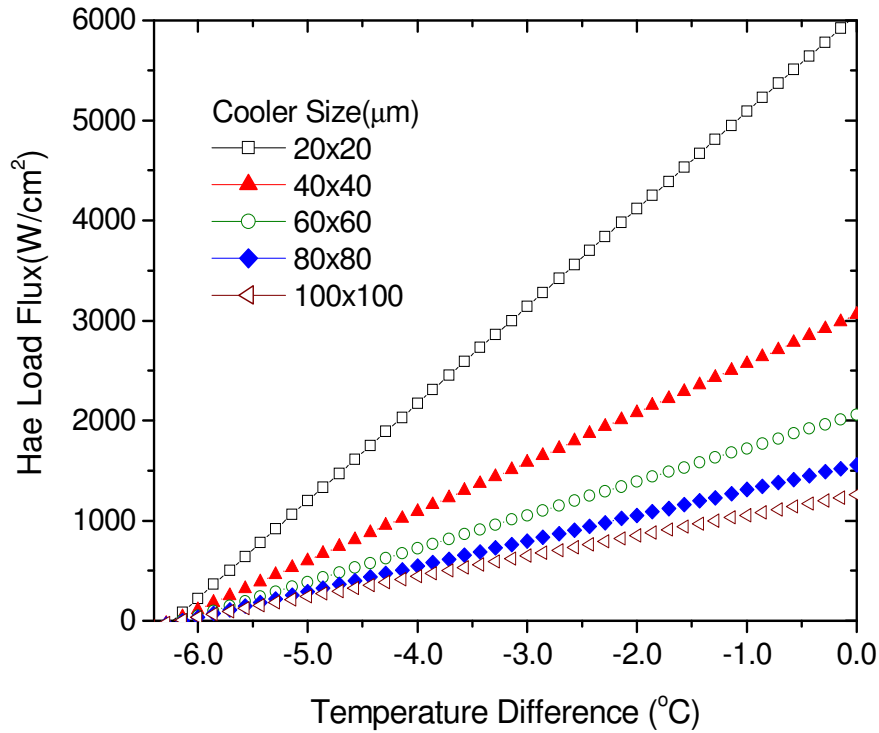
Figure 3.16: Variation of heat load flux with temperature difference on the microcooler at 100°C

3.5 Conclusion

An analytical thermal model for a silicon microcooler, which couples Peltier cooling with heat conduction and heat generation in the silicon substrate, and which includes heat conduction and heat generation in the metal lead, is derived and used to study the thermal characteristics of silicon thermoelectric microcoolers under various operating conditions.

It is found that the analytical modeling results are in excellent agreement with experimental data and detailed numerical finite-element simulations. The electric contact resistance and metal lead effect are found to degrade the cooling performance by as much as 43% for a 60µm×60µm microcooler under typical experimental condition. The doping concentration can be optimized to achieve the maximum cooling performance and it is found that larger electric contact resistances will push the optimum doping concentration to a lower level. Through optimizing geometry the metal lead effect due to heat generation and conduction could be minimized to enhance the cooling performance.

In the ideal case, it is found that the silicon microcooler could achieve a peak cooling of 6.2°C on the microcooler at the optimum doping concentration of 2.5×10^{19} cm^{-3} in silicon. At a negligibly small temperature reduction, the microcooler can extract a heat flux of 1k W/cm^2 for 100µm×100µm microcooler, and 6k W/cm^2 for 20µm×20µm microcooler. This successful thermal analytic characterization of silicon microcoolers is to be followed by subsequent studies on their effectiveness in reducing the hotspot temperature on advanced integrated circuits.

Chapter 4

Analytical Modeling of On-Chip Hot Spot Cooling Using Silicon Thermoelectric Microcooler

In Chapter 3, we developed analytical thermal model to predict thermoelectric cooling potential of silicon thermoelectric microcooler, showing a theoretical maximum cooling flux of several kW/cm^2 and could be a good candidate for on-chip hot spot cooling application. Moreover, silicon thermoelectric microcooler is based on metal-on-silicon fabrication techniques and can be monolithically grown on the silicon chip, yielding a very low thermal contact resistance. In addition, incorporation of the silicon chip into the thermoelectric circuit makes it possible to transfer the absorbed energy via the electric current to the edge of the chip, far from the location of the hotspot, thus substantially reducing the detrimental effect of thermoelectric heating on the temperature of the active circuitry. In this chapter a three-dimensional analytical thermal model of on-chip hotspot cooling is developed to investigate the effectiveness of such silicon thermoelectric microcoolers for a wide range of hotspot sizes and heat fluxes, microcooler sizes, silicon chip thicknesses, doping concentrations, and electric contact resistances. The analytical solution yields the temperature distribution in the silicon chip, under the influence of hotspot heating and background heating from related circuitry on the active surface, Peltier cooling, Peltier heating and conductive/convective cooling on the opposite surface, volumetric Joule heating inside the silicon chip, and interfacial Joule heating at the electric contact created by the silicon microcooler. The analytical solution employs numerically-derived allocation factors to redistribute the Joule heating inside the chip

to the hotspot and the microcooler. Results obtained from a three-dimensional electro-thermal finite-element numerical simulation are used to validate and calibrate the analytical model.
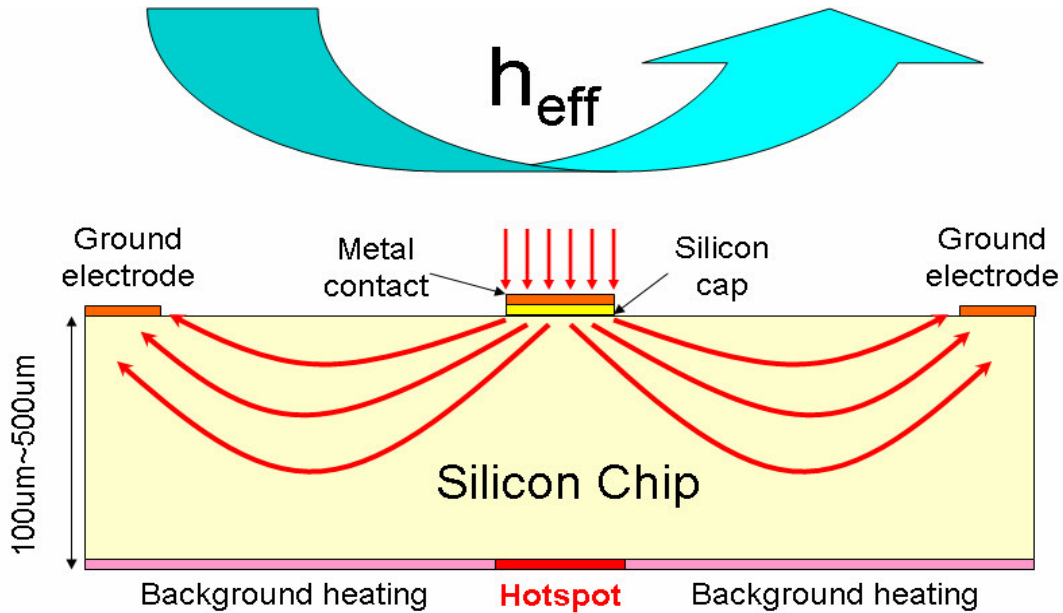
Figure 4.1: Silicon thermoelectric microcooler for on-chip hotspot cooling (The arrows indicate the direction for electric current).

4.1 Thermal Model

4.1.1 Silicon Microcooler for Hotspot Cooling

The structure of silicon thermoelectric microcooler for on-chip hotspot cooling, fabricated on the back of the silicon chip, is illustrated in Figure 4.1, which displays a single microcooler, activated by an electric current entering the silicon chip through the metal contact and the silicon cap, flowing laterally through the chip, and

exiting at the ground electrode located on the periphery of the chip. In a thermoelectric circuit, the flow of electrons across the interface between dissimilar materials, each with a distinct Seebeck coefficient, induces the Peltier effect, providing localized cooling when the direction of current flow is from the low Seebeck coefficient to the high Seebeck coefficient material. The flow of electric current also serves to transport the absorbed heat away from that junction and to deposit that heat at a secondary interface where the electric current flows from the high Seebeck coefficient to the low Seebeck coefficient material. Joule heating, associated with the resistance to electric current in the thermoelectric circuit, and heat conduction from the hot junction to the cold junction of the thermoelectric circuit limit the thermoelectric cooling that can be achieved.

Referring to the structure of the on-chip silicon microcooler depicted in Figure 4.1, it may be seen that Peltier cooling occurs at the junction between the metal contact and the silicon cap which is highly doped silicon with a doping concentration of more than $1 \times 10^{20} \text{cm}^{-3}$ and again at the silicon cap/silicon chip interface, and that Peltier heating is encountered at the silicon chip/ground electrode interface, located on the periphery of the chip, where the electrons must shed some of their energy in entering the highly-conductive metal. The overall Peltier cooling power of the silicon microcooler can be expressed as:

$$q_{TE.c} = -S_{Si}T_c I \tag{4.1}$$

where T_c is the absolute temperature at the microcooler, S_{Si} the Seebeck coefficient of the silicon chip, and I the applied current. Similarly, Peltier heating power at the silicon chip/ground electrode interface can be represented as:

$$q_{TE,h} = S_{Si} T_{ed} I \qquad (4.2)$$

where T_{ed} is the absolute temperature at the ground electrode. In addition to volumetric Joule heating inside the silicon chip, the silicon cap and the metal contact, these parasitic effects also arise at both the metal contact/silicon cap interface and the silicon chip/ground electrode interface. The interfacial Joule heating at the metal contact/silicon cap interface is given by:

$$q_{contact} = I^2 R_{cont} = I^2 \rho_c / A_{cont} \qquad (4.3)$$

where R_{cont} is the electric contact resistance, A_{cont} the cross-sectional area of metal contact, and ρ_c the specific electric contact resistance at this interface. Eq. (4.3) applies as well at the peripheral ground electrode/silicon chip interface, with the appropriately adjusted contact area and the specific electric contact resistance.

The target of this study is 12 mm × 12 mm silicon chip with 70 W/cm^2 background heat flux on the front of the chip and 70 µm × 70 µm hotspot with a heat flux of 680 W/cm^2 located on the center of the front of the chip but various hotspot sizes and hotspot heat fluxes are also investigated for comparison. The back of the chip experiences an effective heat transfer coefficient of 8700 W/m^{2}K in reference to an ambient temperature of 25°C, along with the local thermoelectric cooling flux and thermoelectric heating flux from the microcooler. The heat transfer coefficient,

applied on the exposed back surface of the chip, represents the effective cooling achieved by a typical air-cooled heat sink, heat spreader, and thermal interface materials used for electronic packages. The thermal conductivity of the silicon chip is assumed to be 110 W/mK, appropriate for 100°C operating temperature.
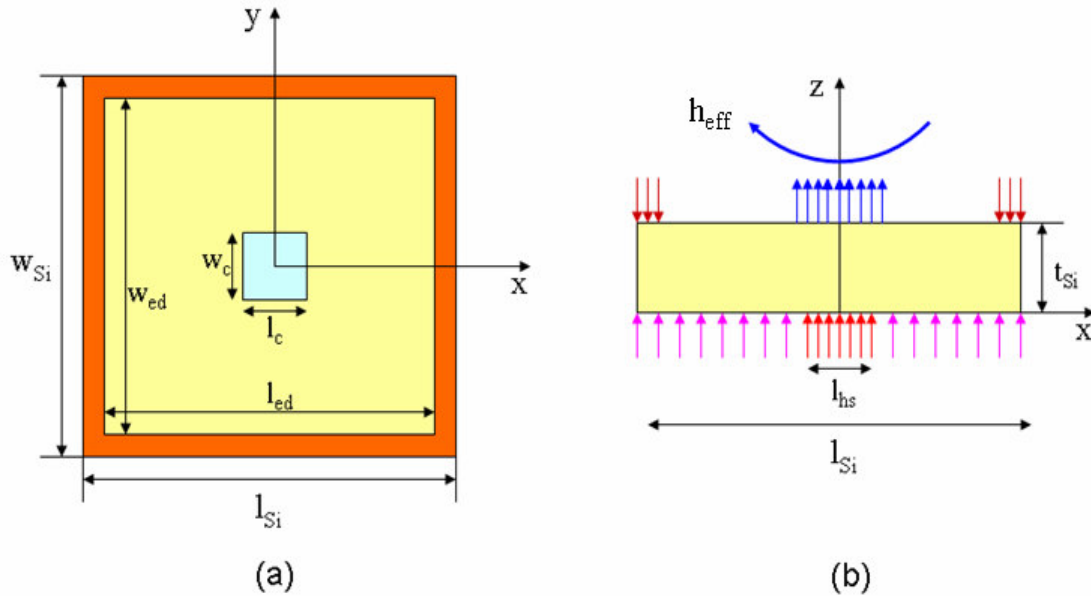
Figure 4.2: (a) Coordinate system and (b) boundary conditions in the analytical model for the silicon chip integrated with silicon thermoelectric microcooler.

4.1.2 Analytical Modeling

Prediction of hotspot cooling achievable with on-chip silicon microcooler , as described in Figure 4.1, requires the solution of the three-dimensional Poison's equation for the temperature distribution in a volume subjected to non-uniform heat generation, associated with the Joule heating in the silicon chip, heating and cooling boundary conditions, associated with Peltier cooling and Peltier heating on the back surface, along with the microprocessor heat generation on the front surface (active circuitry), i.e.,

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} + \frac{q_{Si}'''(x, y, z)}{k_{Si}} = 0 \qquad (4.4)$$

where $q_{Si}'''(x,y,z)$ is the non-uniform volumetric heat generation due to the silicon Joule heating and k_{Si} is the thermal conductivity of the silicon chip. Unfortunately, solution of Equation (4.4) requires detailed knowledge of the internal heat generation function, $q_{Si}'''(x,y,z)$. Determination of this function requires a parallel solution of the LaPlace's equation for the electric potential field which will vary significantly with the geometries of the silicon chip and the silicon microcooler and the placement of the ground electrode. The resulting highly non-uniform heat generation function can be expected to render Equation (4.4) analytically unsolvable for all but the simplest approximations of $q_{Si}'''(x,y,z)$. The simultaneous, conjugate solution of the temperature and electrical potential fields for the chip is, thus, beyond the scope of the present effort.

Alternatively, considering the common use of "allocation factors" in determining the performance of one-dimensional thermoelectric devices and the successful application of this approach to silicon thermoelectric microcoolers in an earlier publications [42], it is possible to re-formulate Equation (4.4) in the LaPlace's form by allocating an appropriate fraction of the Joule heating to the microcooler (α) and the hotspot (β), respectively. With this approach, the volumetric silicon Joule heating is replaced with modified boundary conditions at the microcooler and the hotspot, respectively, and the Poison's equation can then be transformed into the Laplace's equation for this same domain, i.e.,

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = 0 \qquad (4.5)$$

To obtain a general solution, we assume the dimension of the silicon chip is l_{Si} × w_{Si} × t_{Si}, the microcooler size is l_c × w_c and the hotspot size is l_{hs} × w_{hs}. For purposes of the present analysis, a centered microcooler and a peripheral ground electrode, along the outer edge of the chip, are assumed as well, as shown in Figure. 4.2 (a) and (b). The temperature gradient on the microcooler surface due to the effective cooling flux, $q_{c,eff}$'', is given by:

$$\frac{\partial T}{\partial z} = -\frac{q^{''}_{c,eff}}{k_{Si}} \qquad z = t_{Si}, \quad |x| \le 0.5l_c, \quad |y| \le 0.5w_c \qquad (4.6)$$

where q''$_{c,eff}$ is the combined effect of Peltier cooling, Joule heating from the electric contact resistance, and the fraction of silicon Joule heating allocated to the microcooler, given by:

$$q^{''}_{c,eff} = \frac{-S_{Si}TI + I^2(\frac{\rho_c}{l_c w_c}) + \alpha I^2 R_{Si}}{l_c w_c} \qquad (4.7)$$

where R_{Si} is the electrical resistance of the silicon microcooler and is determined by microcooler size, silicon chip thickness and doping concentration N_d. The temperature gradient, resulting from Peltier heating at the ground electrode, is given by:

$$\frac{\partial T}{\partial z} = -\frac{S_{Si}TI}{k_{Si}(l_{Si}w_{Si} - l_{ed}w_{ed})} \qquad z = t_{Si,} \quad 0.5l_{ed} \le |x| \le 0.5l_{Si}, \quad 0.5w_{ed} \le |y| \le 0.5w_{Si} \quad (4.8)$$

In addition, an effective heat transfer coefficient, representing heat transfer by conduction through the electronic package and convection through the heat sink, is applied on the exposed top surface of the silicon chip:

$$\frac{\partial T}{\partial z} = -\frac{h_{eff}}{k_{Si}}(T - T_a) \qquad z = t_{Si}, \quad |x| \leq 0.5 l_{Si,} \quad |y| \leq 0.5 w_{Si}, \tag{4.9}$$

where h_{eff} is the effective heat transfer coefficient, reflecting the entire resistance network from the top chip surface to the ambient, and T_a is the ambient temperature.

The edges of the silicon chip are assumed to be insulated, i.e.,

$$\frac{\partial T}{\partial x} = 0 \qquad x = \pm 0.5 l_{Si}, \quad |y| \leq 0.5 w_{Si}, \quad 0 \leq z \leq t_{Si} \tag{4.10}$$

$$\frac{\partial T}{\partial y} = 0 \qquad y = \pm 0.5 w_{Si}, \quad |x| \leq 0.5 l_{Si}, \quad 0 \leq z \leq t_{Si} \tag{4.11}$$

The thermal boundary conditions on the bottom of the silicon chip include both hotspot heating and background heating, produced by the active circuitry, i.e.,

$$\frac{\partial T}{\partial z} = -\frac{q''_{hs,eff}}{k_{Si}} \qquad z = 0, \ |x| \leq 0.5 l_{hs}, \ |y| \leq 0.5 w_{hs} \tag{4.12}$$

$$\frac{\partial T}{\partial z} = -\frac{q''_{bg}}{k_{Si}} \qquad z = 0, \ 0.5 l_{hs} \leq |x|, \ 0.5 w_{hs} \leq |y| \tag{4.13}$$

where q''_{bg} is the background heat flux and $q''_{hs,eff}$ is the effective heat flux over the hotspot, including the hotspot heat flux q_{hs}'' and the appropriate fraction of silicon Joule heating allocated to the hotspot, and given by:

$$q''_{hs,eff} = q''_{hs} + \frac{\beta I^2 R_{Si}}{l_{hs} w_{hs}} \tag{4.14}$$

The analytical solution for the temperature field in the thermoelectrically-cooled silicon chip can be derived by separation of variables with the assumption that the general solution has the form of T(x,y,z) = X(x)×Y(y)×Z(z) [119-121]. Applying

this method to the solution of Equation (4.5), with the thermal boundary conditions of

Equations (4.10) and (4.11), yields the following general solution:

$$T(x,y,z) = T_a + F_{00} + E_{00}z + \sum_{m=1}^{\infty} \cos(\lambda_m x)[E_{m0} \sinh(\lambda_m z) + F_{m0} \cosh(\lambda_m z)]$$

$$+ \sum_{n=1}^{\infty} \cos(\mu_n y)[E_{0n} \sinh(\mu_n z) + F_{0n} \cosh(\mu_n z)] \quad (4.15)$$

$$+ \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \cos(\lambda_m x)\cos(\mu_n y)[E_{mn} \sinh(\nu_{mn} z) + F_{mn} \cosh(\nu_{mn} z)]$$

where $\lambda_m = \dfrac{2m\pi}{l_{Si}}$, $\mu_n = \dfrac{2n\pi}{w_{Si}}$ and $\nu_{mn} = \sqrt{\lambda_m^2 + \mu_n^2}$ are the eigenvalues.

Application of the thermal boundary conditions at z = 0, i.e., the bottom surface of the

silicon chip (Equations (4.12) and (4.13)), yields the following results for the Fourier

coefficient of E_{00}, E_{m0}, E_{0n} and E_{mn}:

$$E_{m0} = -\frac{q_{hs}^{"} - q_{bg}^{"}}{k_{Si}} \frac{l_{Si} w_{hs}}{m^2 \pi^2 w_{Si}} \sin(\frac{\lambda_m l_{hs}}{2}) \quad (4.16)$$

$$E_{00} = -\frac{q_{bg}^{"}}{k_{Si}} - \frac{q_{hs}^{"} - q_{bg}^{"}}{k_{Si}} \frac{l_{hs} w_{hs}}{l_{Si} w_{Si}} \quad (4.17)$$

$$E_{0n} = -\frac{q_{hs}^{"} - q_{bg}^{"}}{k_{Si}} \frac{w_{Si} l_{hs}}{n^2 \pi^2 l_{Si}} \sin(\frac{\mu_n w_{hs}}{2}) \quad (4.18)$$

$$(4.19)$$

$$E_{mn} = -\frac{q_{hs}^{"} - q_{bg}^{"}}{k_{Si}} \frac{4}{mn\pi^2 \nu_{mn}} \sin(\frac{\lambda_m l_{hs}}{2}) \sin(\frac{\mu_n w_{hs}}{2})$$

Application of the thermal boundary conditions at z = t_{Si}, i.e., the top surface of the silicon chip (Equations (4.8) to (4.9)), yields the following results for the Fourier coefficient of F_{00}, F_{m0}, F_{0n} and F_{mn}:

$$F_{00} = -(t_{Si} + \frac{k_{Si}}{h_{eff}})E_{00} - \frac{q_{te,c}^{"}}{h_{eff}}\frac{l_c w_c}{l_{Si} w_{Si}} - \frac{q_{te,h}^{"}}{h_{eff}}\frac{l_{Si} w_{Si} - l_{ed} w_{ed}}{l_{Si} w_{Si}} \qquad (4.20)$$

$$F_{m0} = -\frac{\frac{h_{eff}}{k_{Si}}\sinh(\lambda_m t_{Si}) + (\lambda_m)\cosh(\lambda_m t_{Si})}{\frac{h_{eff}}{k_{Si}}\cosh(\lambda_m t_{Si}) + (\lambda_m)\sinh(\lambda_m t_{Si})}E_{mo} - \frac{\frac{q_{te,c}^{"}}{k_{Si}}\frac{2w_c}{m\pi w_{Si}}\sin(\frac{\lambda_m l_c}{2})}{\frac{h_{eff}}{k_{Si}}\cosh(\lambda_m t_{Si}) + (\lambda_m)\sinh(\lambda_m t_{Si})} \qquad (4.21)$$

$$-\frac{\frac{q_{te,h}^{"}}{k_{Si}}[\frac{2}{m\pi}\sin(\frac{\lambda_m l_{Si}}{2}) - \frac{2w_{ed}}{m\pi w_{Si}}\sin(\frac{\alpha_m l_{ed}}{2})]}{\frac{h_{eff}}{k_{Si}}\cosh(\lambda_m t_{Si}) + (\lambda_m)\sinh(\lambda_m t_{Si})}$$

$$F_{0n} = -\frac{\frac{h_{eff}}{k_{Si}}\sinh(\mu_n t_{Si}) + (\mu_n)\cosh(\mu_n t_{Si})}{\frac{h_{eff}}{k_{Si}}\cosh(\mu_n t_{Si}) + (\mu_n)\sinh(\mu_n t_{Si})}E_{on} - \frac{\frac{q_{te,c}^{"}}{k_{Si}}\frac{2l_c}{n\pi l_{Si}}\sin(\frac{\mu_n l_c}{2})}{\frac{h_{eff}}{k_{Si}}\cosh(\mu_n t_{Si}) + (\mu_n)\sinh(\mu_n t_{Si})}$$

$$-\frac{\frac{q_{te,h}^{"}}{k_{Si}}[\frac{2}{n\pi}\sin(\frac{\mu_n w_{Si}}{2}) - \frac{2l_{ed}}{n\pi l_{Si}}\sin(\frac{\mu_n w_{ed}}{2})]}{\frac{h_{eff}}{k_{Si}}\cosh(\mu_n t_{Si}) + (\mu_n)\sinh(\mu_n t_{Si})} \qquad (4.22)$$

$$F_{mn} = -\frac{\frac{h_{eff}}{k_{Si}}\sinh(\nu_{mn} t_{Si}) + \nu_{mn}\cosh(\nu_{mn} t_{Si})}{\frac{h_{eff}}{k_{Si}}\cosh(\nu_{mn} t_{Si}) + \nu_{mn}\sinh(\nu_{mn} t_{Si})}E_{mn} - \frac{\frac{q_{te,c}^{"}}{k_{Si}}\frac{4}{mn\pi^2}\sin(\frac{\lambda_m l_c}{2})\sin(\frac{\mu_n w_c}{2})}{\frac{h_{eff}}{k_{Si}}\cosh(\nu_{mn} t_{Si}) + \nu_{mn}\sinh(\nu_{mn} t_{Si})}$$

$$-\frac{\frac{q_{te,h}^{"}}{k_{Si}}\frac{4}{mn\pi^2}[\sin(\frac{\lambda_m l_{Si}}{2})\sin(\frac{\mu_n w_{Si}}{2}) - \sin(\frac{\lambda_m l_{ed}}{2})\sin(\frac{\mu_n w_{ed}}{2})]}{\frac{h_{eff}}{k_{Si}}\cosh(\nu_{mn} t_{Si}) + \nu_{mn}\sinh(\nu_{mn} t_{Si})} \qquad (4.23)$$

As shown in Equation (4.15), the analytical solution for the temperature field in the silicon chip is in the form of an infinite double cosine series. For the calculations reported in this study, all the series were truncated after 300 terms (m = n = 300),

yielding the hotspot and microcooler temperatures that change less than 0.1% between iterations.

The temperature at the center of the hotspot is the highest temperature on the chip and often governs the overall reliability of the circuit and has been widely used to characterize on-chip hotspot cooling performance. The temperature at the center of the microcooler represents a local minimum temperature on the chip and serves to characterize the cooling potential of the microcooler. The temperature at the hotspot center (T = T (0,0,0)) and that at the microcooler center (T = T(0,0,t_{Si})) calculated using Equations. (4.24) and (4.25), respectively, will be employed to characterize the silicon microcooler and its effectiveness for hotspot remediation, as follows:

$$T(0,0,0) = T_a + F_{00} + \sum_{m=1}^{\infty} F_{m0} + \sum_{n=1}^{\infty} F_{0n} + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} F_{mn} \tag{4.24}$$

$$T(0,0,t_{Si}) = T_a + F_{00} + E_{00}t_{Si} + \sum_{m=1}^{\infty}[E_{m0}\sinh(\lambda_m t_{Si}) + F_{m0}\cosh(\lambda_m t_{Si})] \tag{4.25}$$

$$+ \sum_{n=1}^{\infty}[E_{0n}\sinh(\mu_n t_{Si}) + F_{0n}\cosh(\mu_n t_{Si})] + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}[E_{mn}\sinh(\nu_{mn} t_{Si}) + F_{mn}\cosh(\nu_{mn} t_{Si})]$$

4.1.3 Numerical Modeling

A three-dimensional electro-thermal numerical simulation using the commercial finite-element software ANSYS[TM], involving the determination of both the electric field and the temperature field resulting from the application of electric current to the silicon chip was used in this study. The software provided detailed Joule heating and temperature distributions inside the chip, served to validate the analytical model and to define the allocation factors α and β needed for the analytical solution. Thermal-electrical coupling is achieved by solving the LaPlace's equation

for the voltage field, then determining the resulting Joule heating, and subsequently solving the Poison's energy equation for the resulting temperature field, with the defined material properties and boundary conditions. The accuracy of the numerical simulation has been previously verified with experimental data and analytical modeling results for the silicon microcooler.

In this study, the modeled domains include the silicon chip, the ground electrode, the silicon cap, and the metal contact, as shown in Figure 4.1. The thermal-electric elements, Solid 69, are used and densely located around the microcooler and the hotspot where the largest temperature gradient and electric potential gradient are expected to occur. Map meshing is applied for all configurations to assure that the electric current is uniformly distributed over the microcooler. Mesh refinement was conducted to a level that ensures the numerical solution is nearly independent of mesh size, i.e., ~0.02^{o}C change of hotspot temperature with further mesh refinement. To properly capture the large temperature and voltage gradients and the very small thickness of many of the geometric features, more than 150,000 elements were used to create the model. The CPU time, for each geometric configuration and a specific applied electric current, was in the range of 1 to 3 hours, depending on the system geometry. When it was desired to find the optimum current for a specified configuration, a total CPU time of 6 to 18 hours was required. On the other hand, the analytical solution time for the temperature at any specific location is approximately 20 to 30 seconds and less than 3 minutes are required to analytically determine the optimum current with which the maximum hotspot temperature reduction can be achieved.

4.1.4 Electrical Resistance and Allocation Factors

In order to analytically determine the microcooler and hotspot temperatures, using Equations (4.24) and (4.25), respectively, it is necessary to first determine the electrical resistance, R_{Si}, of the path linking the microcooler and the ground electrode through the silicon chip as indicated in Figure 4.1. Following Hewett's approache [122], current flow between a concentric metal contact and a ground electrode on the same surface of a semiconductor can be approximated as a circular spreading disk and the electrical resistance of the silicon microcooler under investigation can thus be shown to equal:

$$R_{Si} = 0.93 \left[\frac{4\rho_{Si}}{\pi r_c^2 r_{Si}^2} \sum_{n=1}^{\infty} \frac{1}{\tanh(\gamma_n t_{Si})} \frac{J_1^2(\gamma_n r_c)}{\lambda_n^3 J_1^2(\gamma_n r_{Si})} \right]$$ (4.26)

where ρ_{Si} is the electrical resistivity of the silicon chip, and r_{Si} and r_c are the equivalent radii of the annular ground electrode and the metal contact, respectively. For $l_c = w_c$ and $l_{Si} = w_{Si}$, these equivalent radii are given by:

$$r_c = \frac{l_c}{\sqrt{\pi}},$$ (4.27a)

$$r_{Si} = \frac{l_{Si}}{\sqrt{\pi}}$$ (4.27b)

The eignevalues, γ_n, in Equation (4.26) are the roots of the Bessel function relation $J_0(\gamma_n) = 0$ which can be computed by means of the following modified Stokes approximation [121]:

$$\gamma_n = \frac{\chi_0}{4}[1 + \frac{2}{\chi_0^2} - \frac{62}{3\chi_0^4} + \frac{15116}{30\chi_0^6}]$$ (4.28)

where $\chi_0 = \pi (4n-1)$ and $n = 1, 2, 3\ldots.$
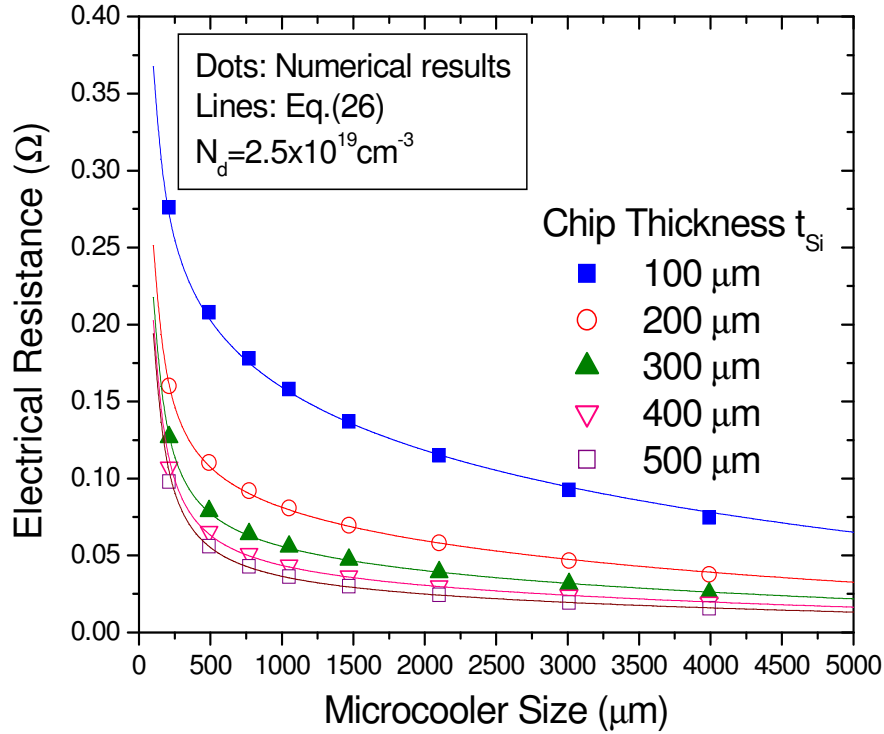
Figure 4.3: Dependence of electrical resistance of silicon thermoelectric microcooler on chip thickness and microcooler size.

Figure 4.3 reveals that the analytical solution, using Equation (4.26), yields the electrical resistance values that are nearly identical to the results from the numerical simulation, differing by less than 3% for the stated conditions. Thus, Equation (4.26) can be used as the basis for the analytical determination of the Joule heating in the silicon chip.
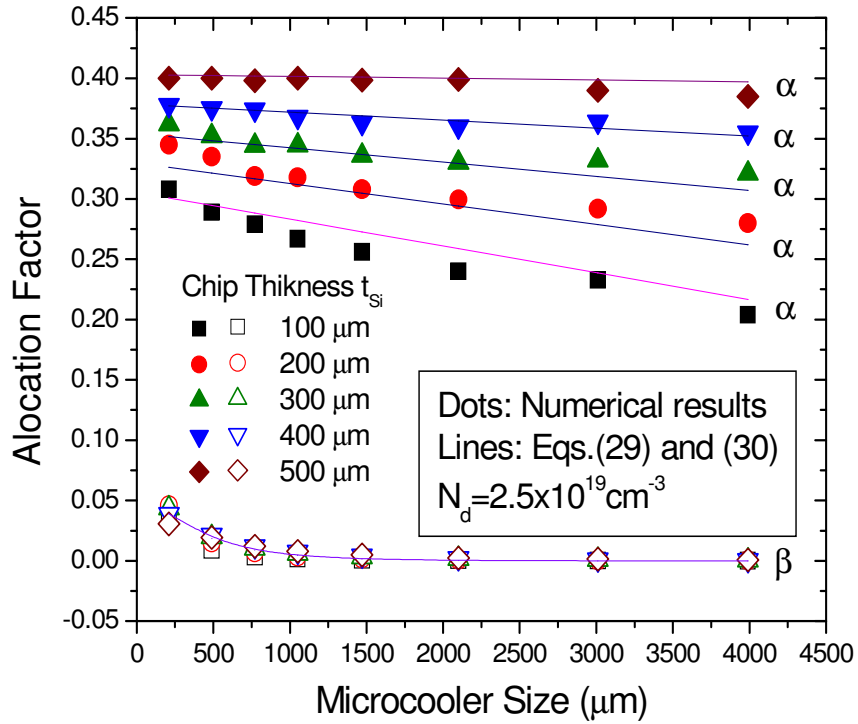
Figure 4.4: Dependence of Joule heating allocation factors on chip thickness and microcooler size.

The allocation factors, α - providing the fraction of silicon Joule heating ascribed to the microcooler, and β - the fraction ascribed to the hotspot can be extracted from the numerical simulation of the coupled thermal and electrical field equations. As shown in Figure 4.4, the microcooler allocation factor (α) is found to vary somewhat more strongly with chip thickness than microcooler size, ranging in value from 0.20 to 0.40 for our investigated configurations, while the hotspot allocation factor (β) is nearly independent of chip thickness at a value of 0.05 for very small microcoolers and asymptotically decays to zero for the microcoolers larger than 1500 μm × 1500 μm. The parametric dependencies of the microcooler allocation factor, α, can be correlated to a very good approximation in the polynomial form of Equation (4.29):

$$\alpha = 0.281 - 243.69\,t_{Si} + 27.40\,l_{cooler} + 5.19 \times 10^4\,l_{cooler}\,t_{Si}, \quad r^2 = 0.97 \quad (4.29)$$

The allocation factor for the hotspot, β, can be correlated well by the following exponential decay function:

$$\beta = 0.068\exp(-\frac{l_c}{377}) \qquad r^2 = 0.95 \qquad (4.30)$$

This observed variation of the allocation factors with geometry is mainly due to the redistribution of current flow and the resulting Joule heating inside the silicon chip as the microcooler size and chip thickness change.

4.1.5 Model Validation

Prior to using the derived analytical model for determining the parametric sensitivities of this novel thermal management approach, an effort was made to verify the accuracy of the analytical solution by comparison to the detailed finite-element simulations. Figure 4.5 and 4.6 display the analytically and numerically-derived temperature contours on the top and bottom surfaces (z = 0 and z = t_{Si}), respectively, of a thermoelectrically-cooled, 100 µm thick silicon chip with a 400 µm × 400 µm microcooler activated with an input power of 0.15 W. It is obvious from Figure 4.5 that the analytical solution properly captures the localized, deep "draw-down" in temperature around the microcooler, though with a slightly thicker "stem," and reaches a nearly identical temperature at the microcooler, within 0.05K of that obtained from the numerical simulation. Similar results were obtained for the temperature distribution on the bottom (active circuitry) side of the chip, again yielding the localized "draw-down" in temperature seen in Figure 4.6 but capturing, as well, the spike in temperature at the center of the hotspot that is anchored at the

111

bottom of the temperature well and then rises substantially above the temperature of the adjacent silicon.

A broader comparison between the analytical and numerical results for the typical configurations and materials properties that were examined in this study is shown in Figure 4.7 which displays the results of the maximum temperature reduction at the microcooler and at the hotspot for different microcooler sizes ranging from 100 µm × 100 µm to 5000 µm × 5000 µm and different chip thicknesses ranging from 100 µm to 500 µm. It is seen that there is less than 7% difference between the analytical and numerical modeling results for temperature reduction at the hotspot and less than 4% difference for the temperature reduction at the microcooler.

(a)

(b)
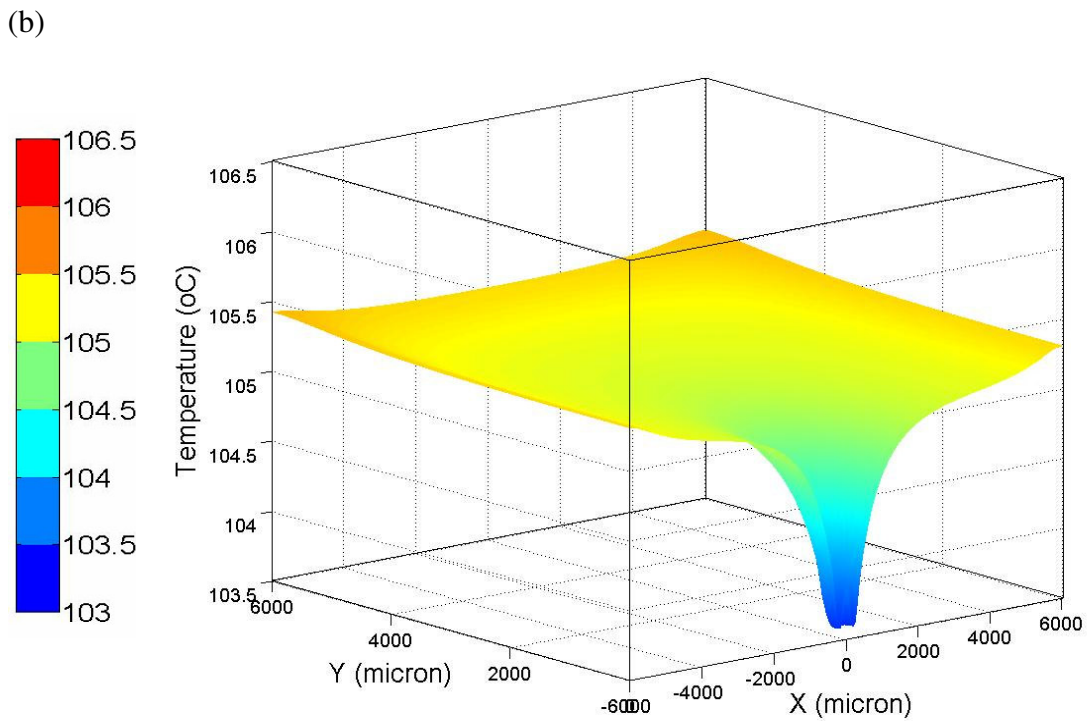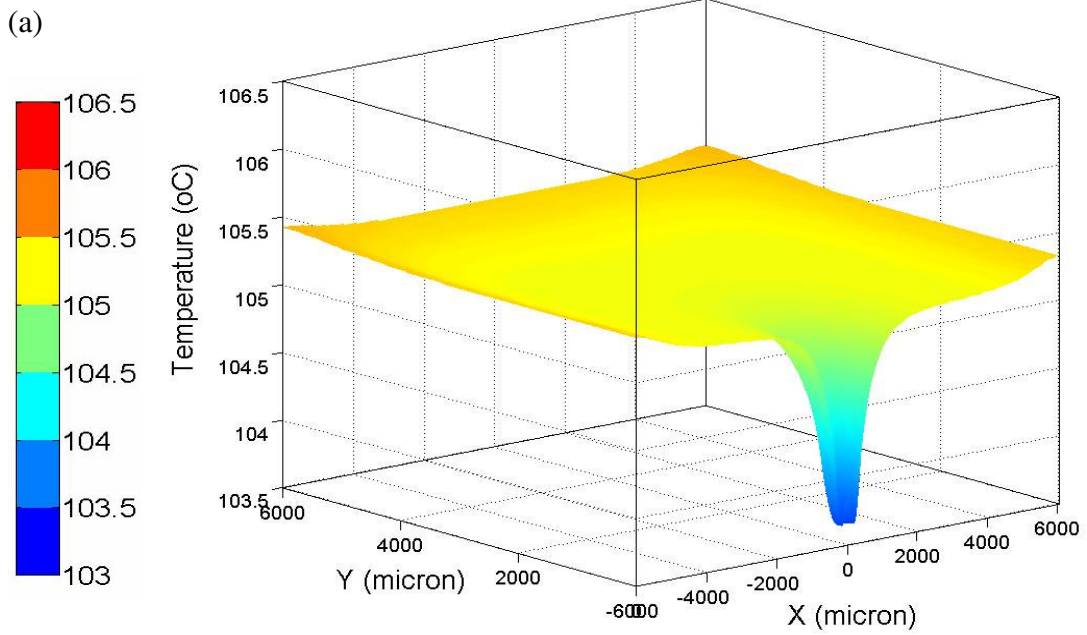
Figure 4.5: (a) Analytical and (b) numerical temperature fields on the top surface of 100 μm thick silicon chip thermoelectrically cooled by 400 μm × 400 μm silicon microcooler with an input power of 0.15 W. The hotspot is 70 μm × 70μm with a heat flux of 680 W/cm^2. (Only half plan is shown).

(a)
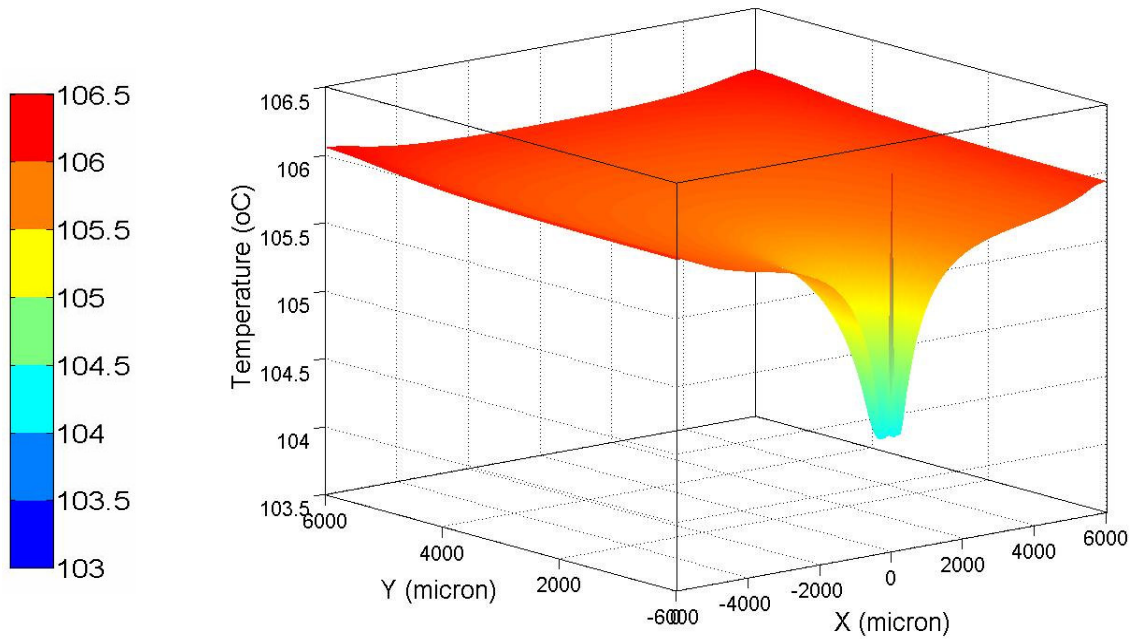
(b)
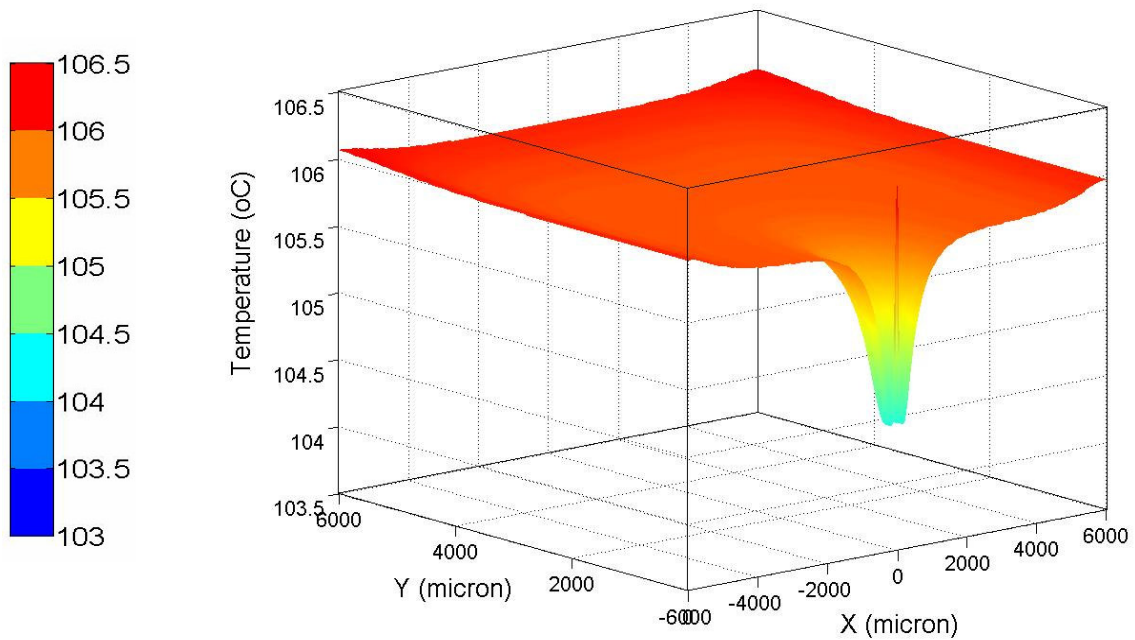
Figure 4.6: (a) Analytical and (b) numerical temperature fields on the bottom surface of 100 μm thick silicon chip thermoelectrically cooled by 400 μm × 400 μm thermoelectric microcooler with an input power of 0.15 W. The hotspot is 70μm × 70μm with a heat flux of 680 W/cm². (Only half plan is shown)
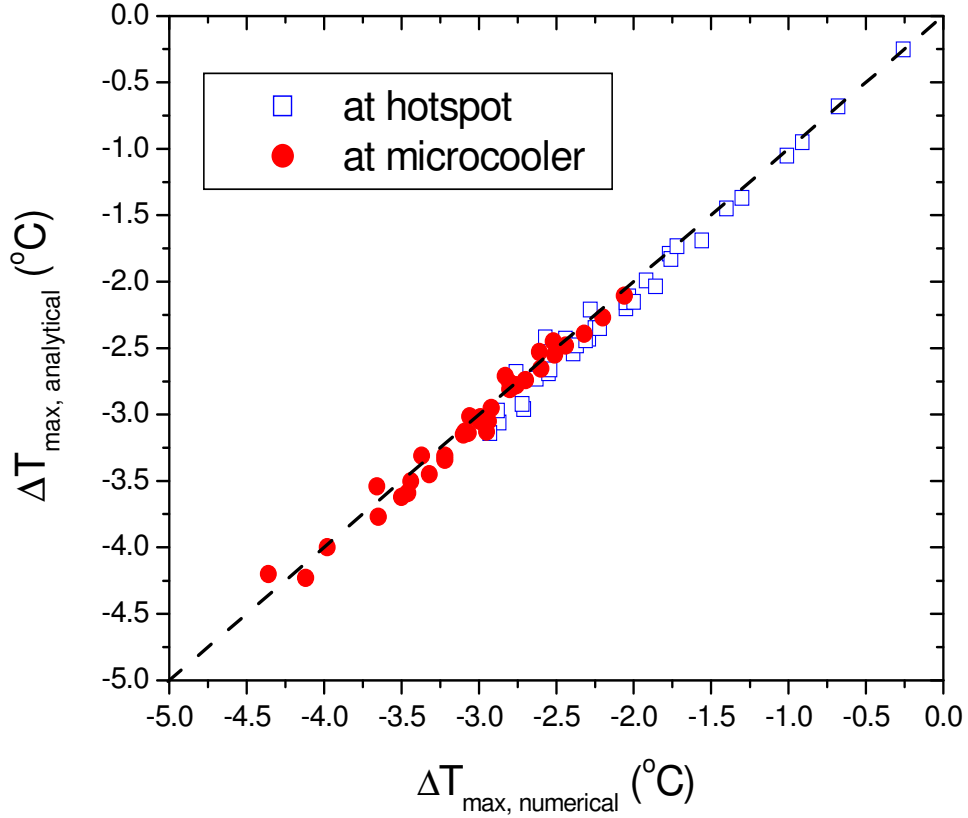
Figure 4.7: Comparison of analytical and numerical results for the maximum temperature reduction at the hotspot and at the microcooler for different microcooler sizes and chip thicknesses.

As a further verification on the accuracy of the analytical solution, it is instructive to compare the hotspot temperature, based on Equation (4.24), to the classical result for a specified hotspot on a semi-infinite slab, given by Equation (4.31) [123]:

$$\Delta T_{semi-\inf} = \frac{q^{''}_{hs,semi} w_{hs}}{\pi^{0.5} k_{Si}} \tag{4.31}$$

where $q_{hs,semi}$'' is the heat flux of the hotspot on the semi-infinite slab.

It is expected that for progressively smaller hotspots and thicker silicon chips, the present analytical solution for the uncooled condition would asymptotically approach the classical semi-infinite values. This behavior is clearly demonstrated in

Figure 4.8, showing the predicted temperature rise for the hotspot with a heat flux of 1000 W/cm^2. The hotspot size ranges from 20 µm × 20 µm to 400 µm × 400 µm, and the silicon chip thickness varies from 100 µm to 500 µm. Thus, Figures 4.5, 4.6, 4.7 and 4.8 provide confidence in the present analytical modeling methodology.
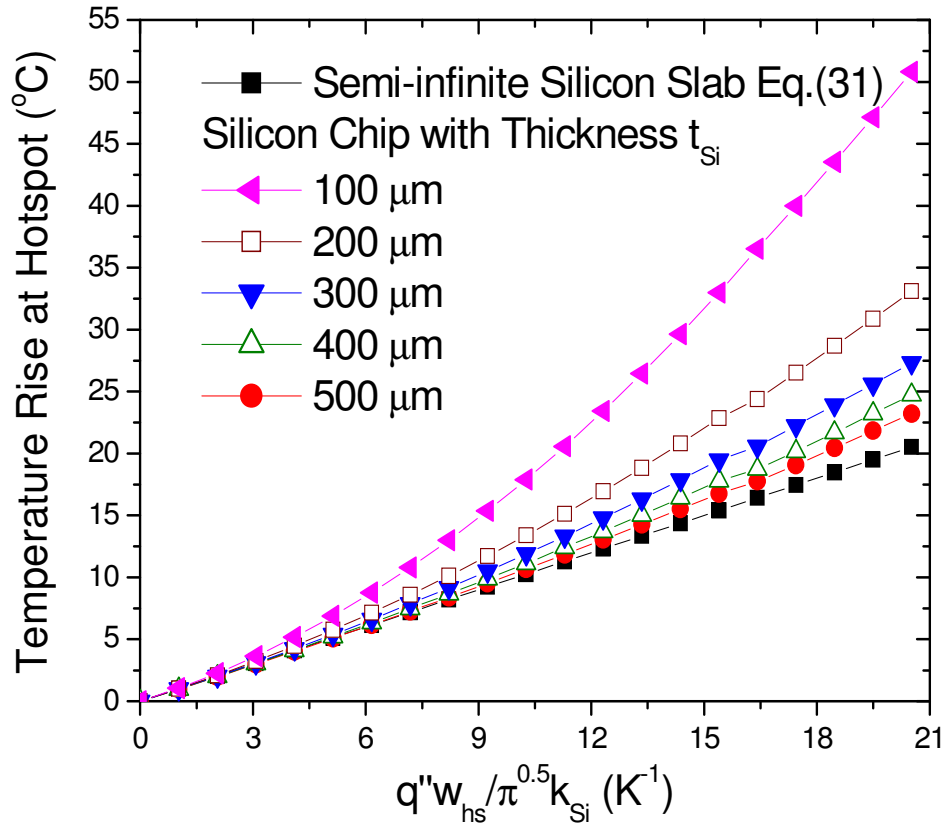
Figure 4.8: Comparison of temperature rise at the hotspot as a function of hotspot size for the silicon chip with finite thickness and semi-infinite silicon slab.

4.2 Results and Discussions

4.2.1 Cooling Metrics

In subsequent sections the hotspot remediation capability of silicon microcoolers will be characterized by three distinct metrics, including:

(1) ΔT - the temperature reduction anywhere in the studied domain that is achievable by activating the microcooler. This metric characterizes the intrinsic thermoelectric cooling capability of the silicon microcooler. It is generally applied to the hotspot or the microcooler in this study and given by:

$$\Delta T = T_{cooler,on} - T_{cooler,off} \qquad (4.32)$$

(2) $\Delta T_{hotspot}$* - the ratio of the temperature change at the hotspot due to activating the microcooler to the temperature rise engendered by the hotspot. This metric quantifies the hotspot cooling effectiveness of the silicon microcooler and is defined as:

$$\Delta T_{hotspot}^{*} = \frac{T_{hotspot\,on,cooler\,off} - T_{hotspot\,on,cooler\,on}}{T_{hotspot\,on,cooler\,off} - T_{hotspot\,off,cooler\,off}} \qquad (4.33)$$

For $\Delta T_{hotspot}$* = 1, the temperature rise engendered by the hotspot can be completely removed by the microcooler. For $\Delta T_{hotspot}$* = 0, the microcooler is totally ineffective and for 0 < $\Delta T_{hotspot}$* < 1, the microcooler can achieve partial success in reducing the hotspot temperature. For $\Delta T_{hotspot}$* > 1, the microcooler is capable of overcooling the hotspot relative to the base temperature of the silicon chip.

(3) π - the thermal impact factor which provides a measure of the power needed, P_{in}, to achieve a specified temperature reduction at the hotspot, $\Delta T_{hotspot}$. This dimensional metric (K/W_{elec}) can be expressed as:

$$\pi = \frac{-\Delta T_{hotspot}}{q_{in}} \qquad (4.34)$$

Clearly, as π increases less electric power is required in order to achieve a specific temperature reduction at the hotspot.

4.2.2 Doping Concentration Effect

The thermoelectric properties of semiconductors are strongly dependent on doping concentration but modestly on the doping type. As Figure 2.14 shows, the electrical resistivity of silicon decreases with increasing doping concentration, while the Seebeck coefficient also displays an inverse relationship with doping concentration. Thus, increasing doping concentration results in lower electrical resistivity and, as a consequence, less Joule heating in the silicon chip, but, the associated decrease in the Seebeck coefficient leads to reduced thermoelectric cooling power. The largest possible thermoelectric cooling power is attained by maximizing the thermoelectric power factor P $(=S^2/\rho)$, which for boron-doped single-crystal silicon at 100^oC occurs at about 2.5×10^{19}cm^{-3}. The variation of maximum hotspot cooling with doping concentration for various microcooler sizes is presented in Figure 4.9 for 100 µm thick chip and the specific electric contact resistance ranging from 1×10^{-7} Ωcm^2 to 1×10^{-4} Ωcm^2, revealing - as expected - that across the range of microcooler sizes studied, with increasing doping concentration the hotspot cooling increases until reaching a maximum value and then decreases with further increases in the doping concentration.

It is interesting to find that, despite the three-dimensional characteristic of heat spreading and electrical current spreading in the silicon chip surrounding the microcooler, for smaller electric contact resistance, e.g., $\rho_c < 1\times10^{-5}$ Ωcm^2, the optimum doping concentration is nearly equal to 2.5×10^{19} cm^{-3} which yields the maximum power factor shown in Figure 2.14. However, it is to be noted that the parasitic effect from larger electric contact resistance does have an influence on the

optimum doping concentration, yielding a lower optimized doping concentration of 1.5×10^{19} cm^{-3} for a 600 µm × 600 µm microcooler with the specific electric contact resistance of 1.0×10^{-4} Ωcm^2. It has been found that this trend becomes more pronounced as the microcooler size gets smaller.
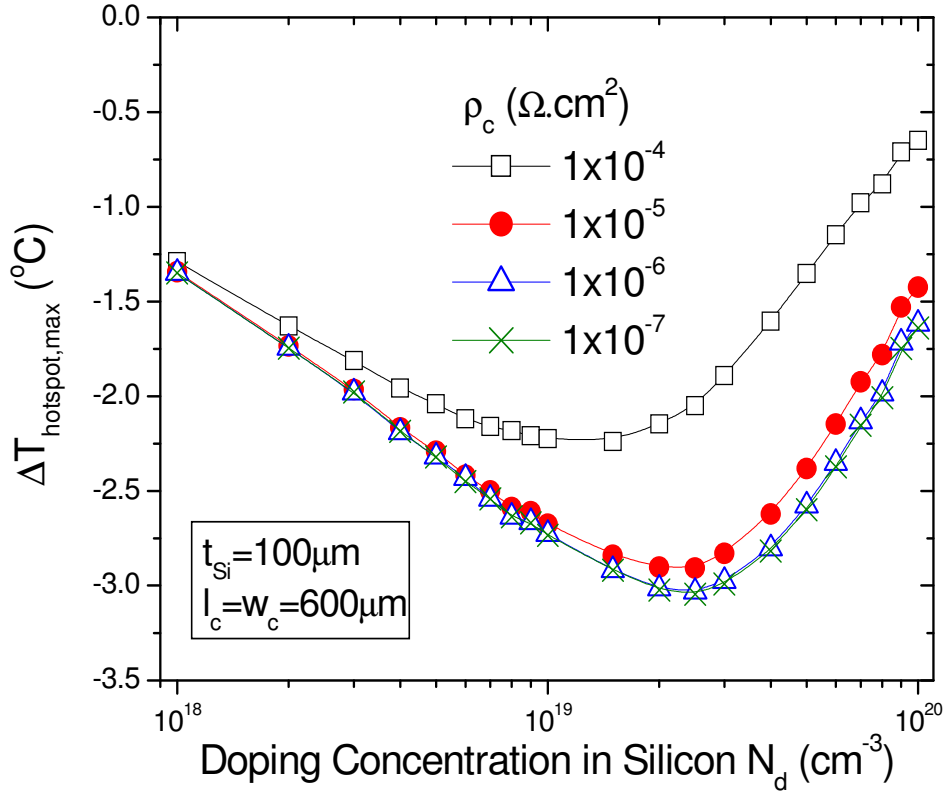
Figure 4.9: Hotspot cooling as a function of boron doping concentration for various specific electric contact resistances. The hotspot is 70 µm × 70 µm with a heat flux of 680 W/cm^2 and the microcooler size is 600 µm × 600 µm. The microcooler size is 600 µm × 600 µm.

It should be noted that the optimum doping level for silicon thermoelectric microcooler is likely to be substantially higher than commonly used in semiconductor silicon chips. However, as is almost always the case for chip thermal management, the present analysis assumes that the back of the chip is used for cooling while the

front is used for the active circuitry. Consequently, the doping concentration on the back side of the chip need not equal the more common doping concentration in the active semiconductor regions at the front of the chip, e.g., $1\times10^{16}\,cm^{-3}$. Due to its far higher electrical resistivity, the electric current that is used to activate thermoelectric cooling is not expected to penetrate into the active silicon layer.

4.2.3 Microcooler Size Effect

The effect of microcooler size on cooling performance involves the interplay of the thermoelectric cooling effect from the microcooler and the thermal diffusion from the hotspot to the microcooler. With decreasing microcooler size, the effective cooling flux and thus the temperature reduction at the microcooler increases, while the thermal resistance between the hotspot and the microcooler also increases. Consequently, this larger cooling flux at smaller microcoolers can not effectively translate into larger temperature reduction at the hotspot. On the other hand, with smaller thermal resistances between the hotspot and the microcooler, the more modest cooling flux on larger microcoolers can be projected effectively onto the hotspot, narrowing the temperature difference between the hotspot and the microcooler. However, the modest cooling flux achievable on the larger microcoolers reduces the beneficial temperature reduction at both the hotspot and the microcooler. The competition between these two effects results in an optimum microcooler size. Figure 4.10 displays this behavior and shows the temperature reductions at the hotspot and the microcooler for a wide range of microcooler sizes for 100 µm thick chip. For each microcooler, we carefully optimized the applied currents in order to achieve the maximum temperature reductions at the hotspot and the miccrocooler. It is seen that

120

the temperature reduction at the hotspot first increases with microcooler size and, after reaching the maximum value of 3.03°C for 600 µm × 600 µm microcooler, decreases with a further increase in microcooler size. Interestingly, the temperature reduction at the microcooler varies monotonically with microcooler size, yielding progressively larger temperature reductions, to as much as 3.9°C, as the microcooler dimension shrinks to 100 µm × 100 µm which, however, only provides 1.6°C temperature reduction at the hotspot.
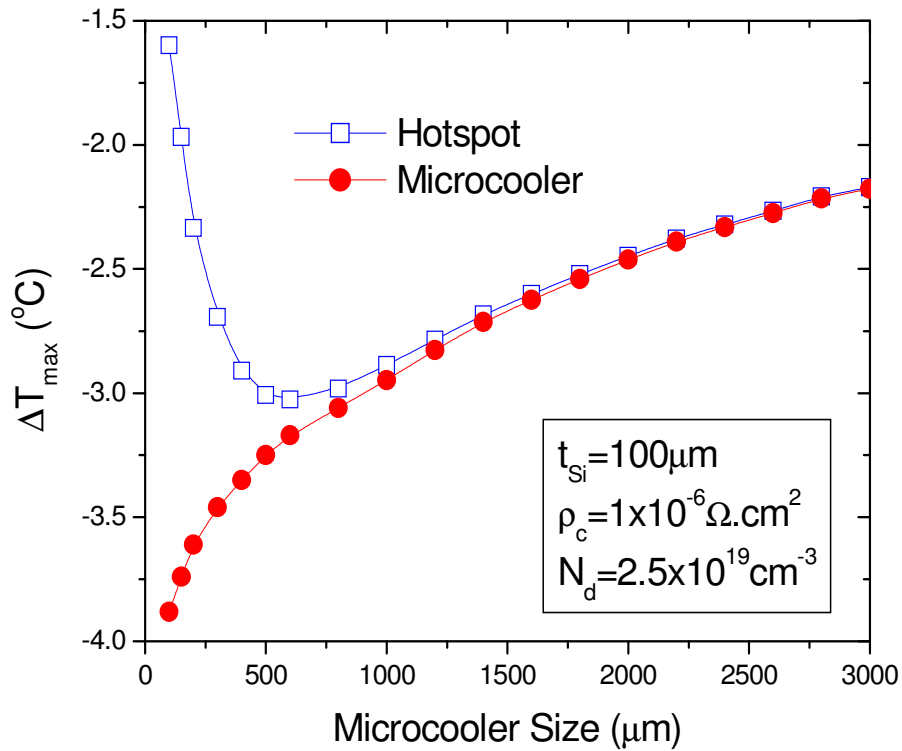
Figure 4.10: Variation of temperature reductions at the hotspot and the microcooler with microcooler size. The hotspot is 70 µm × 70 µm with a heat flux of 680 W/cm².

4.2.4 Chip Thickness Effect

In the application of silicon microcooler to hotspot remediation, the silicon chip plays multiple roles, functioning as a thermoelectric material, to provide on-chip cooling and, at the same time, as an electrical conductor to transfer electrons from the ground electrode to the microcooler, and as a thermal conductor to provide a diffusion path for the heat generated in the chip to the ambient. Therefore, the chip thickness influences Joule heating distribution inside the chip, heat spreading from the hotspot, heat diffusion from the hotspot to the microcooler, and heat diffusion from the ground electrode, where Peltier heating occurs, to the hotspot. As the chip becomes thinner, the thermal resistance between the microcooler and the hotspot decreases, allowing the microcooler to achieve greater hotspot temperature reductions, e.g., 2.05°C to 3.03°C as the chip thickness decreases from 500 µm to 100 µm, for the conditions of Figure 4.11. However, due to the smaller heat spreading effect in thinner chips, the temperature rise engendered by the hotspot is also higher for thinner chips and increases with decreasing chip thickness from 2.2°C for a 500 µm thick chip to 2.9°C for 100 µm thick chip. These two trends compete with each other, yielding the maximum hotspot cooling effectiveness at the chip thickness of 200 µm, with $\Delta T_{hotspot}^* = 1.2$ as shown in Figure 4.11. At this chip thickness, the silicon microcooler is, thus, capable of reducing the hotspot temperature below the baseline temperature of the chip by approximately 0.5°C. Moreover, for the present 70 µm × 70 µm hotspot with a heat flux of 680 W/cm^2, the silicon microcooler is capable of completely suppressing or over-cooling the hotspot, with $\Delta T^*_{hotspot} \geq 1$, for the chip thicknesses between 100 µm and 475 µm.
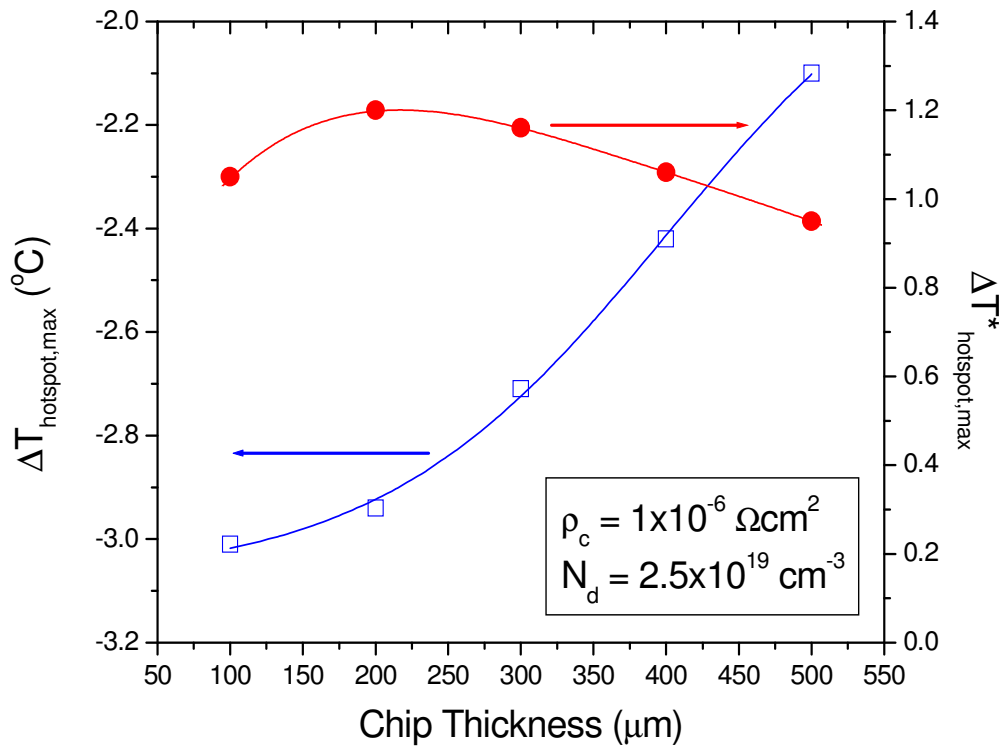
Figure 4.11: Hotspot cooling and hotspot cooling effectiveness as a function of chip thickness. The hotspot is 70 μm × 70 μm with a heat flux of 680 W/cm².

It is interesting to find that the chip thickness also influences the optimum microcooler size. As shown in Figures 4.12 and 4.13, the thicker the silicon chip the larger the optimized microcooler size. For 100 μm thick chip, the maximum hotspot cooling of 3.03°C is achieved with a 600 μm × 600 μm microcooler, while for 500 μm thick chip, 2500 μm × 2500 μm silicon microcooler is required in order to attain the maximum hotspot cooling of 2.1°C. It should be noted that the optimum ratio of microcooler size to chip thickness is, thus, approximately 5.5, with a modest sensitivity to chip thickness, reaching 6.0 for 100 μm, 200 μm and 300 μm thicknesses, dropping to 5.5 for the 400 μm thick chip, and to 5.0 for the 500 μm chip. This decreasing ratio can be related to the growing contributions of silicon Joule

heating and Peltier heating to the hotspot temperature as the optimized current - necessitated by the larger microcooler - increases. Figure 4.14 shows the dependence of the thermal impact factor, π, of the silicon microcooler on the microcooler size for different chip thicknesses, revealing that this factor and the relative benefit of input power decreases steeply with the microcooler size but more gently with the chip thickness. For example, 200 µm × 200 µm microcooler can achieve a π of 17.0 in comparison with 1.2 for 4000 µm × 4000 µm microcooler on 100 µm thick chip. With the increase of microcooler size, the effect of chip thickness on the thermal impact factor becomes less important. Consequently, the largest π values and the best returns on invested energy are attained when smaller microcoolers are used to remediate the hotspots on thinner chips.
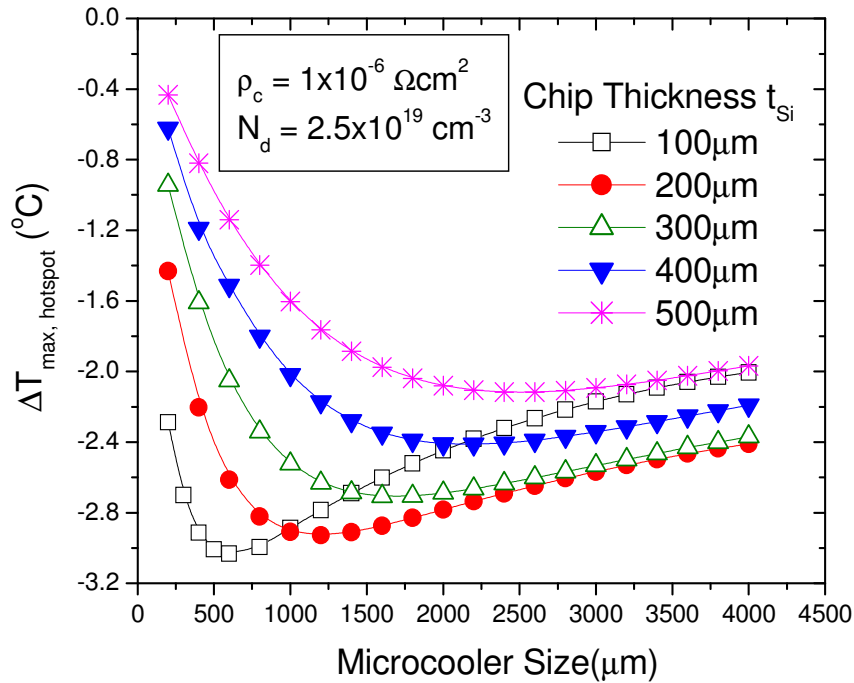
Figure 4.12: Hotspot cooling as a function of microcooler size for various chip thicknesses. The hotspot is 70 μm × 70 μm with a heat flux of 680 W/cm^2.
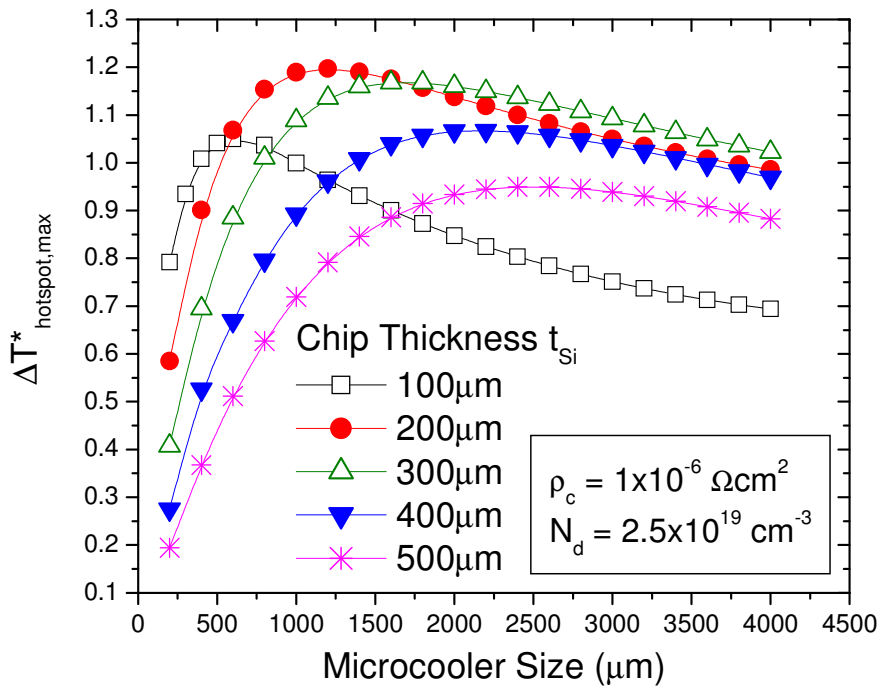
Figure 4.13: Hotspot cooling effectiveness as a function of microcooler size for various chip thicknesses. The hotspot is 70 μm × 70 μm with a heat flux of 680W/cm^2.
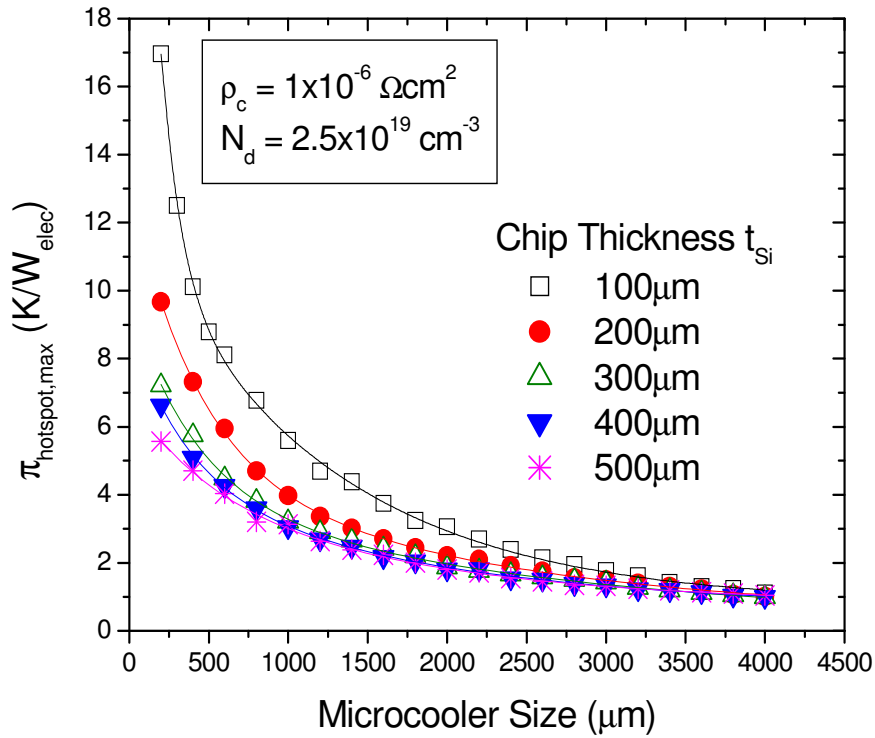
Figure 4.14: Thermal impact factor as a function of microcooler size for various chip thicknesses. The hotspot is 70 μm × 70 μm with a heat flux of 680 W/cm^2.

4.2.5 Electric Contact Resistance Effect

The miniaturization of thermoelectric coolers tends to exacerbate the deleterious effects of the electric contact resistance which is expected to occur at the interface between the metal contact and the silicon cap. The theoretical value of the specific electric contact resistance between highly-doped silicon and a metal contact is in the range of 1×10^{-9} Ωcm^2 at room temperature or above. However, due to process-related limitations, the typical specific electric contact resistance at such an interface usually ranges from 1×10^{-7} Ωcm^2 to 1×10^{-5} Ωcm^2, with significant batch to batch variations. Figure 4.15 shows the impact of the electric contact resistance on hotspot cooling for different microcooler sizes on 100 µm thick chip. In all cases, as

the specific electric contact resistance increases, hotspot cooling performance is degraded, but the electric contact resistance has a larger impact for smaller microcooler sizes because the contact resistance is inversely proportional to the microcooler area. For an increase in the specific electric contact resistance from 1×10^{-9} Ωcm^2 to 1×10^{-4} Ωcm^2, hotspot cooling will be degraded by a factor of 6.5 for 100 µm × 100 µm microcooler but only by 5% for 3000 µm × 3000 µm microcooler. It should be noted that for a typical state-of-the-art thin film process, which yields an average specific electric contact resistance of approximately 1×10^{-6} Ωcm^2 [43], the results displayed in Figure 4.15 reveal that the electric contact resistance induced degradation in hotspot cooling can be neglected.
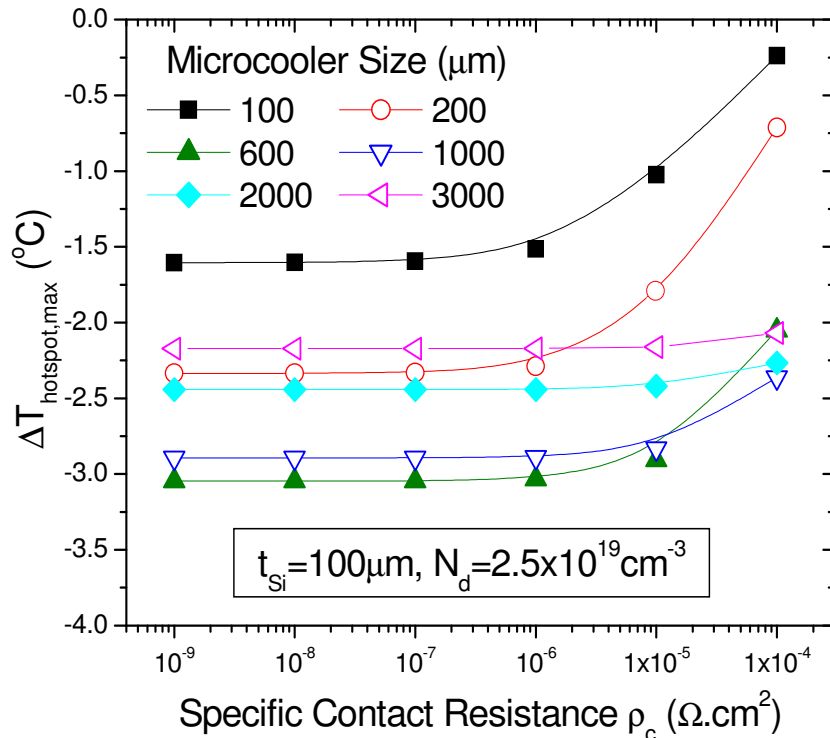
Figure 4.15: Hotspot cooling as a function of specific electric contact resistance for various microcooler sizes. The hotspot is 70 µm × 70 µm with a heat flux of 680 W/cm².

4.2.6 Hotspot Parameter Effect

Finally, attention is turned to the effects of the hotspot parameters - namely hotspot size and hotspot heat flux - on cooling performance, as evaluated by the three proposed metrics – ΔT, ΔT^*, and π. For each hotspot size and hotspot heat flux, the applied current, the microcooler size and the doping concentration have been optimized in order to achieve the maximum hotspot temperature reduction while the specific electric contact resistance is fixed at 1×10^{-6} Ωcm^2. It was found the optimized current and thus the optimized input power increase slightly with hotspot size and hotspot heat flux if the chip thickness and the doping concentration in silicon keep constant. As may be seen in Figures 4.16, 4.17 and 4.18 for 100µm thick chip, the efficacy of the silicon microcooler varies with these hotspot parameters in a complex manner. For example, the maximum temperature reduction at the hotspot, shown in Figure 4.16, increases from 3.03^{o}C for 70 µm × 70 µm hotspot with 680 W/cm^2 heat flux to 3.90^{o}C for 400 µm × 400 µm hotspot with 1000 W/cm^2 heat flux, primarily because of the effect of the higher chip temperature (105°C vs. 150°C) on Peltier cooling power. However, as seen in Figure 4.17, the maximum cooling effectiveness decreases steeply with hotspot size and hotspot heat flux, with the cooling effectiveness decreasing from 1.05 for 70 µm × 70 µm hotspot with 680 W/cm^2 heat flux to 0.08 for 400 µm × 400 µm hotspot with 1000 W/cm^2 heat flux. Interestingly, since as the hotspot size and the hotspot heat flux increases, the maximum hotspot temperature reduction increases while the optimized input power almost keep constant, π, the thermal impact factor, increases with the hotspot size and the heat

flux, as shown in Figure 4.18. It should, thus, be understood that the silicon microcoolers are most effective in remediating smaller hotspots.
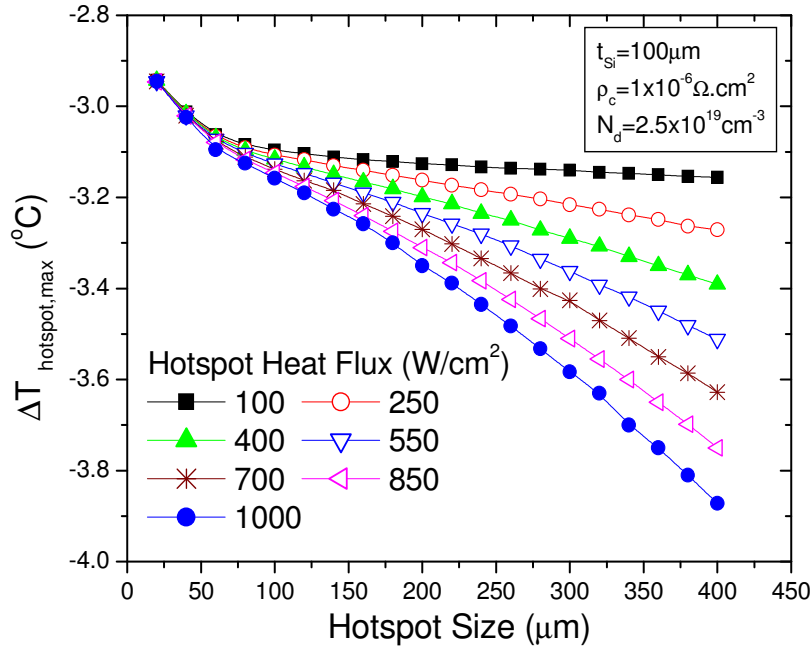
Figure 4.16: Hotspot temperature reduction as a function of hotspot size and hotspot heat flux.
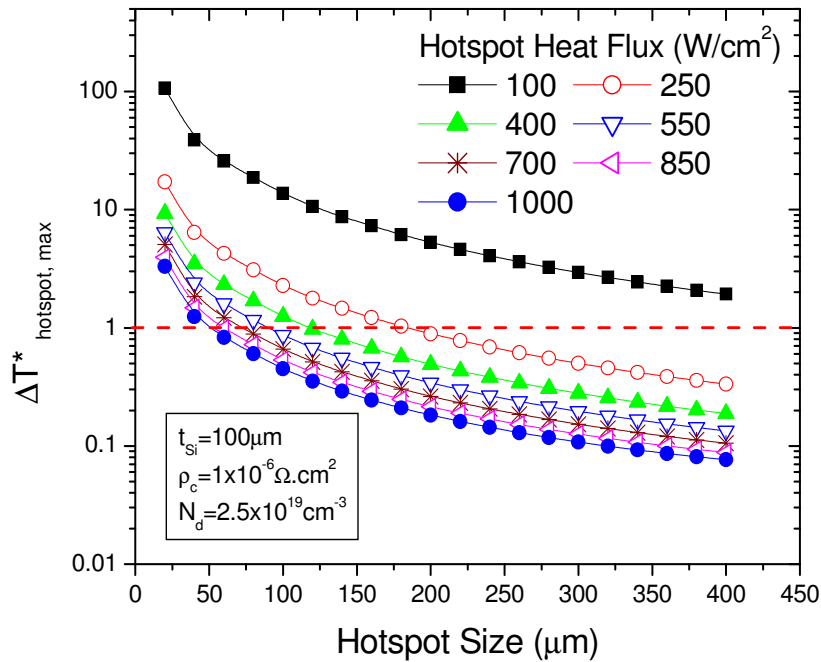
Figure 4.17: Hotspot cooling effectiveness as a function of hotspot size and hotspot heat flux.
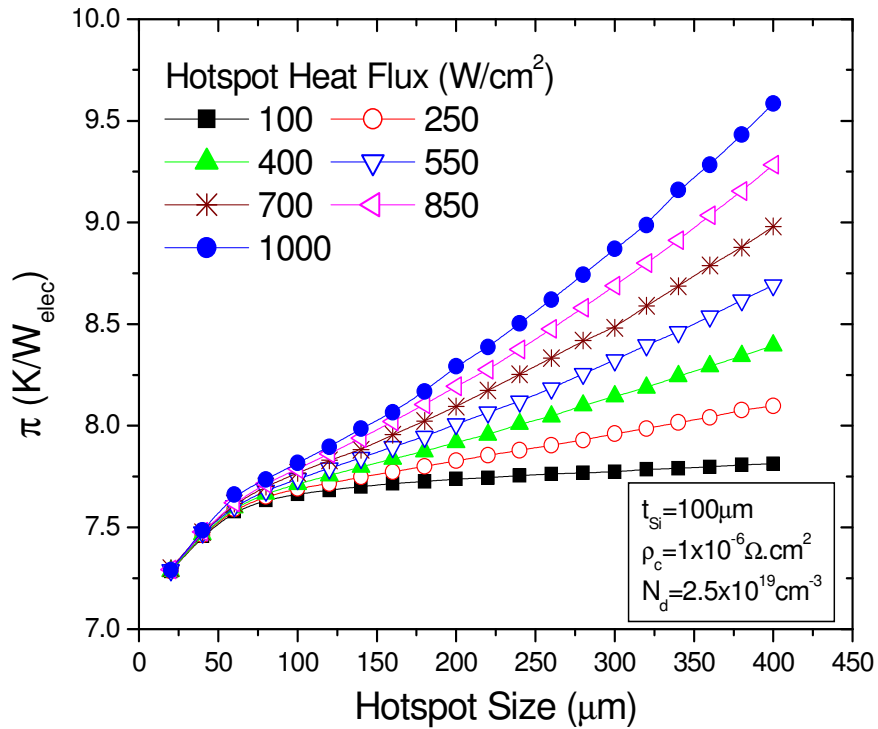
Figure 4.18: Thermal impact factor as a function of hotspot size and hotspot heat flux.

4.3. Simplified Thermal Modeling

4.3.1 Closed-Form Analytical Solutions

The hot spot temperature, with and without thermoelectric cooling, can be predicted very well using Equation (4.15) with an error of less than 7% in comparison with numerical simulation results. However, the exact analytical solution is too cumbersome to be used effectively in the development of an initial thermoelectric cooling design and the associated determination of the geometric values needed to achieve effective hot spot remediation. Consequently, the current effort focuses on the derivation and validation of simplified, closed-form equations, for the hot spot temperature, based on the superposition of three effects, background heating, hot spot heating, and thermoelectric cooling, as:

$$T_{hotspot} = T_{background\ heating} + \Delta T_{hotspot\ heating} + \Delta T_{hotspot\ cooling} \qquad (4.34)$$

The first term is due to the heat dissipation elsewhere on the chip, the second term the localized hot spot heating, and the third term the effective thermoelectric cooling effect at the hot spot resulting from the "projection" of the net cooling flux generated at the silicon microcooler. The contribution of silicon Joule heating on the hot spot temperature is neglected in this simplified closed-form solution because as, based on the allocation factors determined in previous detailed computations, less than 5% of Joule heating in the silicon penetrates to the hot spot in all cases and less than 1.5% when the microcooler size is more than 500 µm × 500 µm for all die thickness we investigated. As these three terms are interactive and the thermoelectric cooling power is approximately proportional to the chip temperature, the chip temperature has to be determined first and then hot spot temperature reduction can be calculated.

The temperature rise due to the background heating can be represented as the summation of one-dimensional heat conduction inside the silicon chip and heat convection from the top of silicon chip into the ambient air. The resulting analytical equation is very straightforward and given by:

$$T_{background\ heating} = q^{''}_{background}\left[\frac{t_{Si}}{k_{Si}} + \frac{1}{h_{eff}}\right] + T_a \qquad (4.35)$$

The temperature rise due to the hot spot is mainly due to heat spreading inside the silicon chip, which depends on the silicon chip thickness, hot spot size,

hot spot flux, and silicon thermal conductivity. The simplified closed-form equation to estimate such temperature rise is given as follows:

$$\Delta T_{hotspot\ heating} = \frac{q''_{hotspot}}{k_{Si}}\left(0.237\frac{w^2_{hotspot}}{t_{Si}} - 209.4w^2_{hotspot} + 0.56w_{hotspot}\right), \quad r^2 = 0.99 \tag{4.36}$$
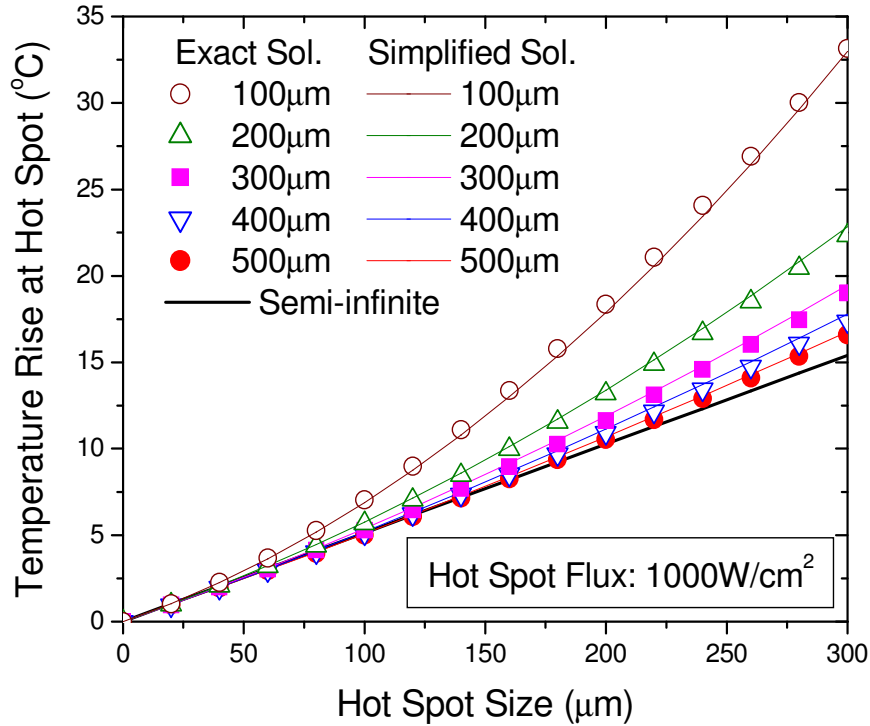
Figure 4.19: Hot spot temperature rise as a function of hot spot size for different silicon chip thickness. The hot spot heat flux is 1000 W/cm^2.

Figure 4.19 compares the uncooled hot spot temperature rise, predicted using Equation (4.15), with the values obtained from the simplified solution, using Equation (4.36). In this comparison the silicon chip thickness ranges from 100 µm to 500 µm, and hot spot size varies from 20 µm × 20 µm to 300 µm × 300 µm while its heat flux is kept at 1000 W/cm^2 as an example. As expected from the above

discussion, Figure 4.19 reveals that for progressively smaller hot spots and thicker silicon chips, the present closed-form equation for the uncooled condition asymptotically approaches the classical semi-infinite values. For larger hot spots and/or thinner chips, it is found that the difference between the exact solution and simplified solution is less than 1.8%. We also obtain similar results with various hot spot heat fluxes indicating that Equation (4.34) is a very convenient way to accurately predict hot spot temperature rise.

The thermoelectric cooling effect on the hot spot is very complicated and mainly determined by microcooler size, silicon chip thickness, silicon thermal conductivity and the effective cooling flux on the microcooler surface. We applied multiple regression method and found the following closed-form equation can be used to predict hot spot cooling performance when the cooler size/die thickness is less than 10 (e.g. $w_c/t_{Si} < 10$):

$$\Delta T_{hotspot,cooling} = \frac{w_c^2 q_c^{''}}{k_{Si}}[901 - 217(\frac{w_c}{t_{Si}})^{0.5} + 14003e^{\frac{-t_{Si}}{0.98 \times 10^{-4}}} \qquad (4.37)$$

$$+ 2670e^{\frac{-t_{Si}}{1.04 \times 10^{-4}}}(\frac{w_c}{t_{Si}})^{0.5}], \qquad r^2 = 0.98$$

4.3.2 Verification of Simplified Equation with Parametric Effects

The previous section established the ability of the simplified, closed form equations to accurately predict the distinct components of the uniform background heating effect, the hot spot temperature rise, namely self heating, and silicon thermoelectric cooling of the hot spot. In order to validate the ability of the simplified analytical equations to predict the behavior of a thermoelectrically cooled

hot spot subjected to the complex boundary conditions associated with an active chip, attention will now be turned to the exploring how the chip and cooler geometry, as well as the doping concentration and interfacial contact resistance affect the accuracy of closed-form simplified equations.
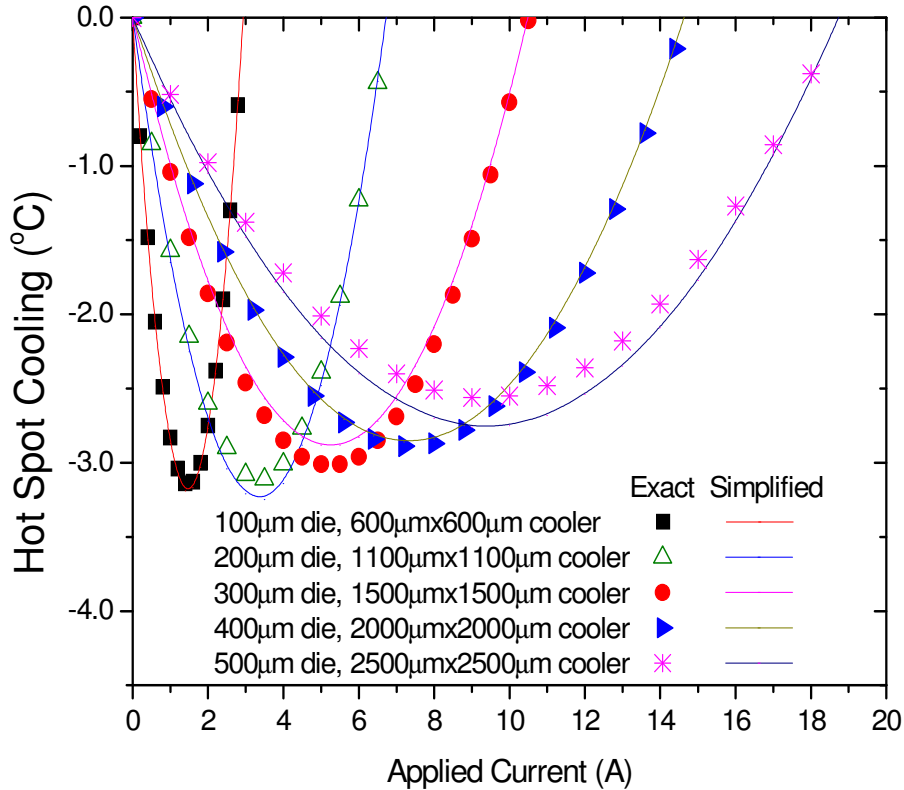
Figure 4.20: Comparison of hot spot cooling as a function of applied current for different silicon chip thickness under optimized microcooler size (boron doping concentration is 2.5×10^{19}cm^{-3} and electrical contact resistance is 1×10^{-6} Ω.cm^{2}). The symbols: exact analytical solutions; The solid lines: simplified closed-form solutions.

Figure 4.20 shows the hot spot temperature reduction as a function of applied current for the previously determined optimum ratios of cooler size to chip thickness for an assumed boron doping concentration of 2.5×10^{19} cm^{-3} and electrical contact resistance of 1×10^{-6} Ω.cm^{2}. As illustrated in Figure 4.20, these comparisons indicate

the results from exact solutions and those from closed-form simplified equations agree well. The largest discrepancies appear to occur at the optimum thermoelectric currents - yielding the deepest hot spot temperature reductions. The maximum discrepancies varies from 0.5% for 100µm and 300µm thick dies to 7% for 500µm thick die in comparison with hot spot cooling temperature derived by exact analytical solutions.
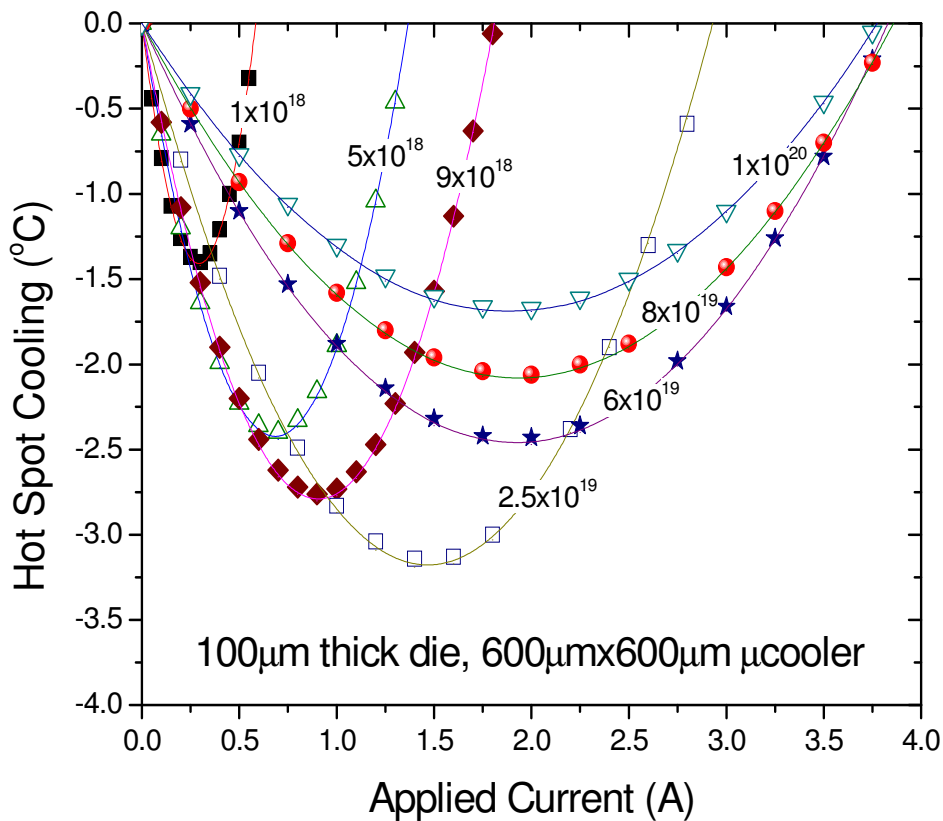
Figure 4.21: Comparison of hot spot cooling as a function of applied current for different doping concentration in silicon chip (chip thickness is 100 µm, microcooler size is 600 µm × 600 µm and electrical contact resistance is 1×10^{-6} $\Omega.cm^2$). The symbols: exact analytical solutions; The solid lines: simplified closed-form solutions. The number in the figure shows the boron doping concentration in silicon chip.

Next we examine the effect of boron doping concentration in the silicon chip on the accuracy of the simplified closed-form solutions. Here we use 100 µm thick silicon chip with a 600 µm × 600 µm microcooler as an example because the hot spot cooling performance of this configuration was well investigated in our previous research. Figure 4.21 illustrates the case where the electrical contact resistance is 1×10^{-6} $\Omega.cm^2$ and the boron doping concentration varies from 1×10^{18} cm^{-3} to 1×10^{20} cm^{-3}. The comparisons reveal that the simplified closed-form solution works extremely well for all the doping concentrations we investigated, providing excellent agreement to within 0.5% for the full parametric range we investigated.
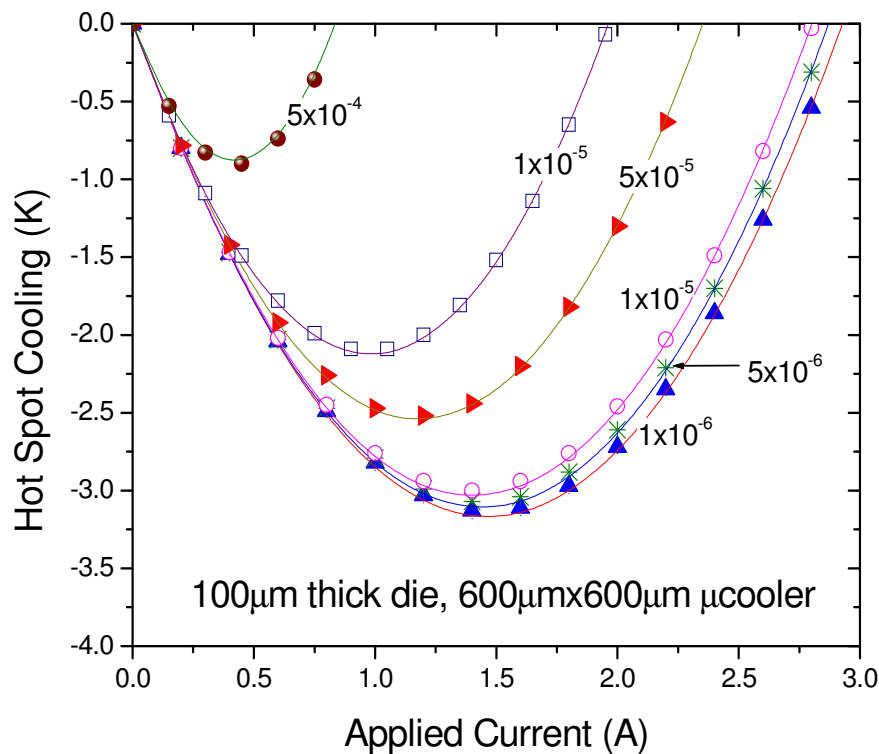
Figure 4.22: Comparison of hot spot cooling as a function of applied current for different electric contact resistances (chip thickness is 100 µm, microcooler size is 600 µm × 600 µm and boron doping concentration is 2.5×10^{19} cm^{-3}). The symbols: exact analytical solutions; The solid lines: simplified closed-form solutions. The number in the figure shows the electric contact resistance (unit: $\Omega.cm^2$) at silicon/metal interface.

Finally we examine the impact of electrical contact resistance on the accuracy of the simplified closed-form solutions for prediction of on-chip hot spot cooling performance, again assuming a 100 µm thick silicon chip with the microcooler size of 600 µm × 600 µm and boron doping concentration of 2.5×10^{19} cm^{-3}. As indicated in Figure 4.22, in all of the electrical contact resistance we investigated, the simplified closed-form solution is in excellent agreement with exact solutions, demonstrating that this simplified solution is a feasible approach to quickly examining and optimizing silicon microcoolers for maximum on-chip hot spot cooling.

4.4 Conclusions

A three-dimensional, numerically-validated analytical model was developed to investigate the remediation of microprocessor hotspots using the inherent thermoelectric properties of the silicon chip. Allocation factors extracted from electro-thermal numerical simulations were used in the analytical model to account for the impact of silicon Joule heating on the hotspot and the microcooler. The analytical model was used to study the parametric sensitivity of hotspot cooling and, typically, to determine the temperature reduction at 70 µm × 70 µm hotspot with a heat flux of 680 W/cm^2, under a variety of geometric and operating conditions. In the parametric range studied, the optimum microcooler size was found to vary from 5 to 6 times the chip thickness. The optimized doping concentration was found to be insensitive to system geometry but dependent on parasitic effects with high electric contact resistance pushing the optimized doping to lower levels. Larger hotspot size and higher hotspot heat flux results in larger temperature reductions at the hotspot but

lower hotspot cooling effectiveness. Under the optimized condition, the temperature rise engendered by the hotspot can be partially suppressed, completely removed, or even over-cooled depending on the hotspot size and hotspot heat flux, showing the promise of silicon thermoelectric microcoolers for on-chip hotspot cooling.

A simplified closed-form analytical thermal model for silicon thermoelectric cooling of on-chip hot spot is derived and compared to the exact three-dimensional analytical solution previously derived. It is found that the results from the simplified closed-form solutions are in very good agreement with those from the exact Fourier-series analytical solution, typically within 7% of the predicted hot spot temperature reduction, for the full parametric range investigated. It is expected that the closed-form solutions can be used effectively to reduce the complexity and required time for the design and optimization of silicon microcoolers for on-chip hot spot remediation.

Chapter 5

Numerical Simulation of On-Chip Hotspot Cooling Using Silicon Microcooler

Analytical modeling of on-chip hot spot cooling developed in the previous section is based on idealized one-layer structure where only silicon die is modeled while the thermal effects of thermal interface materials, heat spreader and heat sink are simplified using effective heat transfer coefficient applied on the top surface of the die, so that the three-dimensional LaPlace's heat conduction equation can be solved analytically. However, in real applications, the chip package is complicated with five-layer structure including silicon die, two layers of thermal interface materials (TIMs), heat spreader and heat sink, which makes a three-dimensional analytical solution impossible. In this section, the three-dimensional package-level numerical simulation is developed to investigate potential application of silicon microcooler for on-chip hot spot cooling. We anticipate that with detailed numerical model the predicted hot spot cooling performance will be more close to the real case. In addition, the detailed information such as temperature distribution and heat flux distribution inside the die can be obtained using numerical modeling. Thus, the objective of this study is to evaluate the potential application of silicon microcooler to remove an on-chip hot spot using three dimensional electro-thermal finite element modeling and to explore the parametric sensitivities of hot spot cooling performance, including the influence of boron doping concentration in silicon, microcooler size, silicon die thickness, hot spot size, and electrical contact resistivity between the

silicon and metal. These modeling results will help to define the optimum solid-state cooling configuration.

5.1 Numerical Modeling Methodology

To address the hot spot cooling needs of advanced microoelectronics, silicon thermoelectric microcooler can be fabricated on the top surface of the die, as shown in Figure 5.1(a), with a non-uniform heat flux distribution on the bottom of the die. A chip package cooled with such embedded thermoelectric microcooler includes the silicon die, two layers of thermal interface materials, an integrated heat spreader and a heat sink. The commercial finite element software, ANSYS^{TM}, was used in this study to simulate the thermal and electric behaviors of silicon microcooler using a half-symmetry three dimensional (3-D) thermal-electrical mode (Solid 69) with a total element number of 100,000 - 200,000 for the entire package. In order to facilitate thermal modeling of the IC package without the penalty of very large node counts and long computational runs, the detailed structures of the silicon microcooler, such as the silicon cap and the metal contact layer, are combined into a single "surface entity" attached to the top of the silicon chip. This "surface entity" is capable of generating a cooling heat flux equal to the combined Peltier cooling effect at the metal contact/silicon cap interface and at the silicon cap/silicon chip interface. The net Peltier cooling effect can then be expressed as an internal heat flux boundary condition on the surface of silicon microcooler given by Equation (5.1):

$$q_{TE,cooler}^{"} = -S_{Si}TI \, / \, A_{cooler} \tag{5.1}$$

where A_{cooler} is the surface area of silicon microcooler. This approximation simplifies significantly the numerical computations and provides a convenient way to determine

the extent to which the silicon microcooler will be useful in the overall system configuration. Joule heating due to electrical contact resistance, R_c, can again be represented as a heat flux boundary condition and directly added to silicon microcooler surface, and this additional parasitic term is expressible as:

$$q''_{Joule,contact} = I^2 R_c / A_{cooler} = I^2 \rho_c / A^2_{cooler} \tag{5.2}$$

Combining the Peltier cooling effect, Equation (5.1), and Joule heating from the electrical contact between the metal contact and silicon cap, Equation (5.2), the effective cooling effect of the silicon microcooler on the surface of the silicon chip can be expressed as the cooling heat flux on the microcooler:

$$q''_{eff,cooler} = (-S_{Si}TI + R_c I^2) / A_{cooler} \tag{5.3}$$

Peltier heating which occurs at the silicon/ring electrode interface is introduced as a heat flux boundary condition on the ring electrode surface:

$$q''_{electrode} = S_{Si}TI / A_{electrode} \tag{5.4}$$

where $A_{electrode}$ is the area of ring electrode.
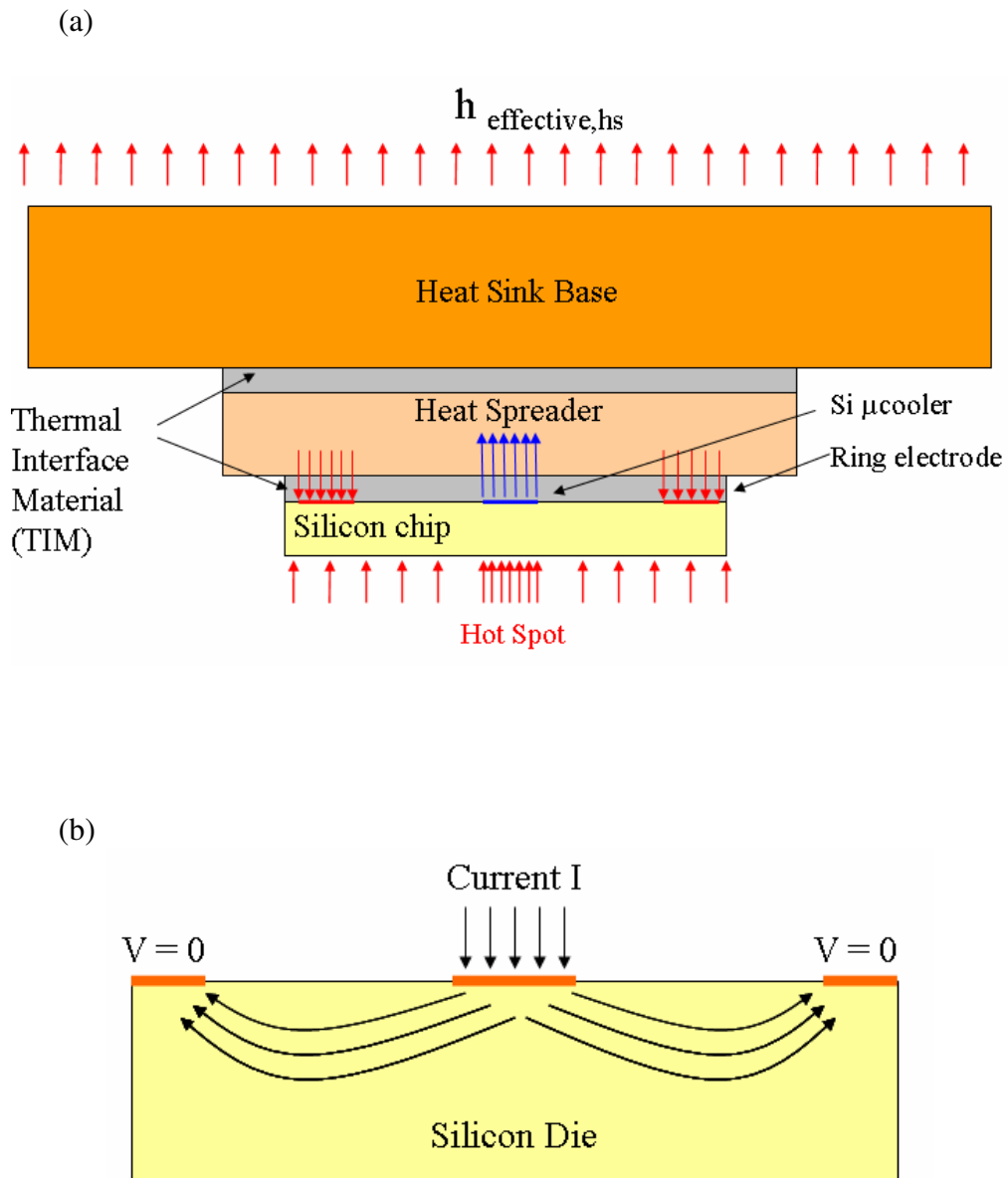
Figure 5.1: (a) Schematic of chip package with an embedded silicon microcooler, (b) structure and current flow of silicon microcooler developed on the top of the die.

The Joule heating effect inside the silicon chip is simulated using ANSYS' thermal-electrical element (Solid 69) which allows for both thermal and electric fields to be resolved through thermal-electrical coupling. As illustrated in Figure 5.1(b), a voltage of zero is applied as a boundary condition on the surface of the ring electrode, and the electric current is applied onto the microcooler surface, as a surface load. In this way, Joule heating from the silicon chip could be simulated directly using the thermo-electrical mode of ANSYS. The elements are densely located around the microcoolers and the hot spot where the largest temperature gradient is expected to occur. As shown in Figure 5.2, mesh density in the silicon chip near the microcoolers is also high in order to accurately calculate the thermal and electrical spreading effects and three dimensional distribution of Joule heating.
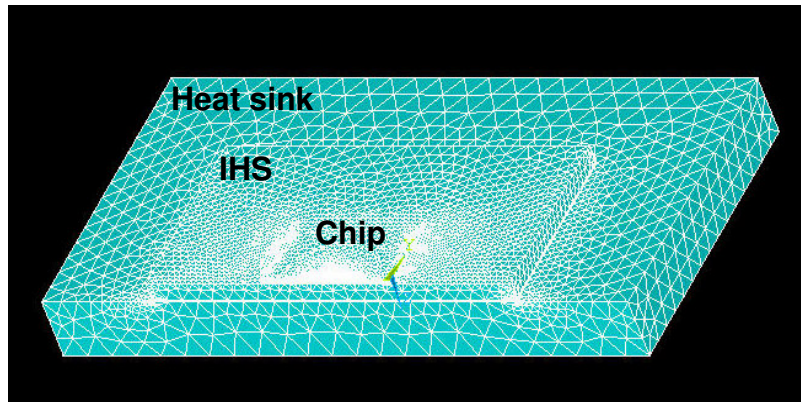
Figure 5.2: Mesh structure in chip package.

For purposes of the thermoelectric modeling study presented herein, the details of the solid-state circuitry in the chip, including individual transistors, gates, capacitors, etc., in the active regions of the chip are ignored and the heat generated from these components is represented, typically in most of this study, as a 680 W/cm^2 70 µm × 70 µm hot spot surrounded by a background heat flux of 70 W/cm^2 on the bottom surface of the silicon chip. A ring electrode at a distance of 4 mm from the

center and a width of 300 µm is also included. The major heat transfer path in this model is assumed to be from the active region at the bottom of the silicon chip to the top side, then through the heat spreader to the heat sink by conduction and from the heat sink to the ambient by forced convection. Two layers of solder-like thermal interface material (TIM) - on either side of the heat spreader - are included in the model. The thermal contact resistances at these two interfaces are included in the effective thermal conductivity value of 30 W/mK used for the 178 µm TIM's. To simplify the modeling geometry, the details of heat sink fins are not included in this model and, instead, an equivalent convective heat transfer coefficient of 730 W/m^2-K is applied as a boundary condition on the top surface of the heat sink base to achieve a commonly attained heat sink-to-ambient thermal resistance of about 0.55 K/W. Also, homogeneous material properties and uniform thicknesses are assumed for the silicon chip, thermal interface materials, heat spreader, and heat sink base. The geometric parameters and material properties for the packaging materials are listed in Table 5.1.

Table 5.1: Geometry and material properties used for the numerical model.

Materials	Dimension (L×W×H)	Thermal Conductivity (W/mK)
Silicon die	11mm×13mm× (25 to 200µm)	110
Thermal interface materials (1st layer)	11mm×13mm×178µm	30
Integrated heat spreader	31mm×31mm×1.5mm	150
Thermal interface materials (2nd layer)	31mm×31mm×178µm	30
Heat sink base	50mm×50mm×5mm	360

It should be noted that the optimum doping level for silicon thermoelectric microcooler is substantially higher than commonly used in semiconductor silicon chips. However, as is almost always the case for chip thermal management, the present analysis assumes that the back of the chip is used for cooling while the front is used for the active circuitry and that, therefore, the doping concentration on the back (for example 2×10^{19} cm^{-3}) need not equal the doping concentration on the front of the chip (for example 1×10^{16} cm^{-3}). Due to such great doping concentration difference, the electrical resistivity in the active silicon layer for circuitry is approximately two orders of magnitude higher than that suggested for the region of the silicon used for thermoelectric microcoolers. Therefore, the electric current that is used to provide thermoelectric cooling is not expected to penetrate deeply into the active silicon layer on the opposite side of the chip. Alternatively, this silicon microcooler configuration could be used for SOI (silicon-on-insulator) chip applications in which the active silicon layer is separated – i.e. electrically insulated - by a thin layer of SiO$_2$ (usually several hundred nanometer in thickness) from the bulk silicon.

5.2 Results and Discussions

5.2.1 Typical Cooling Performance

5.2.1.1 Typical Behavior

The typical temperature contour of an IC package containing a 50 µm thick 11 mm × 13 mm silicon die, with a 70 µm × 70 µm hot spot and a heat flux of 680 W/cm^2 and with the background heat flux of 70 W/cm^2, is shown in Figure 5.3. The corresponding temperature profile along the bottom (active surface) of the silicon die

is illustrated in Figure 5.4. It is observed that the presence of the background heat flux of 70 W/cm^2 on the bottom of the silicon die produces a parabolic temperature distribution, peaking at the center of the chip at around 104.5$^{\text{o}}$C. If a hot spot, 70 µm × 70 µm in size with the heat flux of 680 W/cm^2, is added at the center of the die, the peak temperature increases to around 107$^{\text{o}}$C, an increase of 2.5$^{\text{o}}$C. When a 150 µm × 150 µm microcooler is integrated onto the top of the silicon die to cool the hot spot, with an applied current of 0.6 A, the peak hot spot temperature is reduced to ~105$^{\text{o}}$C, and a more locally-complex temperature variation is created that includes a microcooler ring around the hot spot and a slightly elevated temperature in the second ring surrounding the hot spot. Therefore, by application of the silicon microcooler, 72% (~1.8$^{\text{o}}$C) of the 2.5$^{\text{o}}$C local temperature rise produced by the 680 W/cm^2 hot spot could be removed, at the expense of a small increase in the average temperature of the chip.
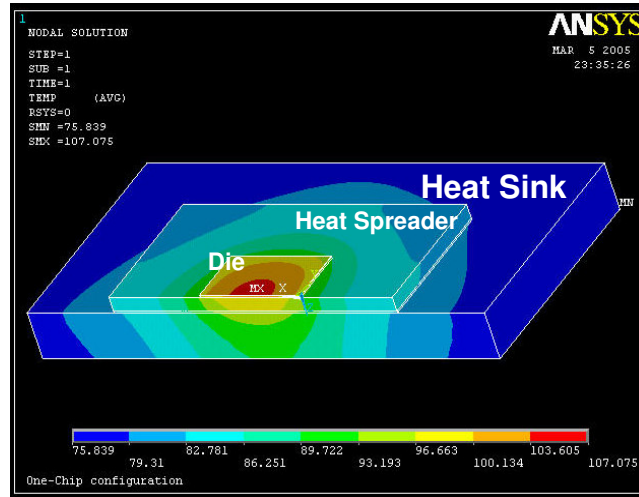
Figure 5.3: Temperature contours for chip package embedded with a silicon microcooler. The hot spot size is 70 µm × 70 µm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2, the die thickness is 50 µm, the microcooler size is 150 µm × 150 µm, the electric contact resistance is 1.0 × 10^{-6} Ω.cm^2, and the boron doping concentration is 2.5 × 10^{19} cm^{-3} in silicon.
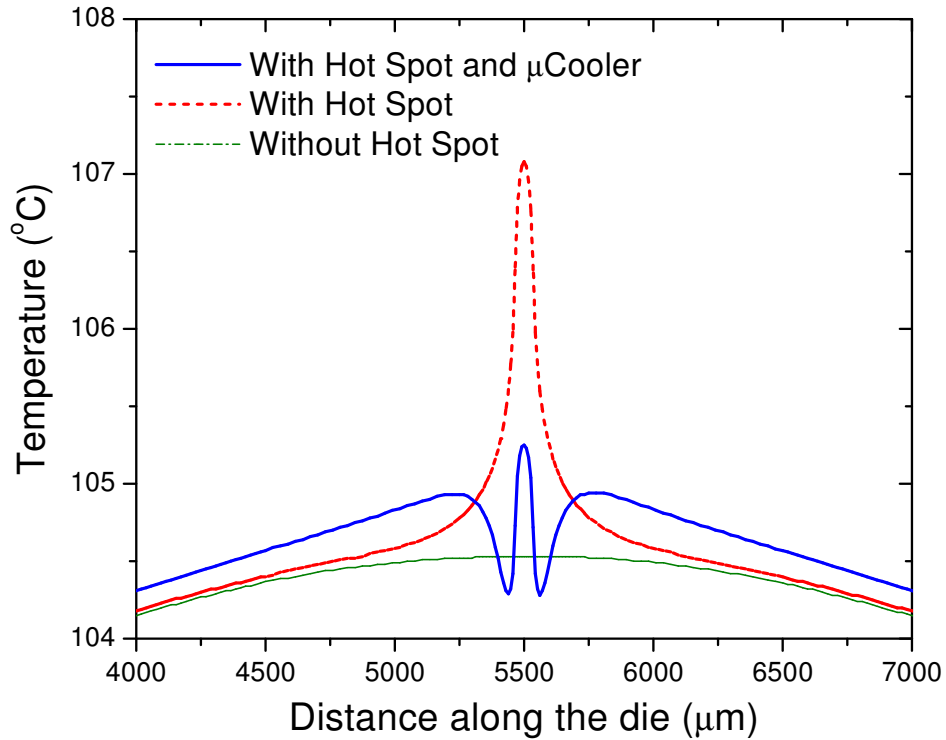
Figure 5.4: Temperature distribution on the bottom of the die. The hot spot size is 70 μm × 70 μm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2, the die thickness is 50 μm, the microcooler size is 150 μm × 150 μm, the electric contact resistance is 1.0 × 10^{-6} Ω.cm^2, and the boron doping concentration is 2.5 × 10^{19} cm^{-3} in silicon.

Figure 5.5 shows the typical variation of hot spot cooling and hot spot cooling effectiveness with an applied current for a 150 μm × 150 μm microcooler on a 50 μm thick silicon die. With an increase in the applied current, both the hot spot cooling and the cooling effectiveness improve monotonically and reach their maximum values at the optimum current of 0.6 A, at which a hot spot temperature reduction of 1.8^{o}C is achieved, representing 72% of the hot spot temperature rise on the silicon die. With further increases in the applied current, the hot spot cooling performance

and cooling effectiveness deteriorate, due to the growing dominance of the Joule heating in the silicon die.
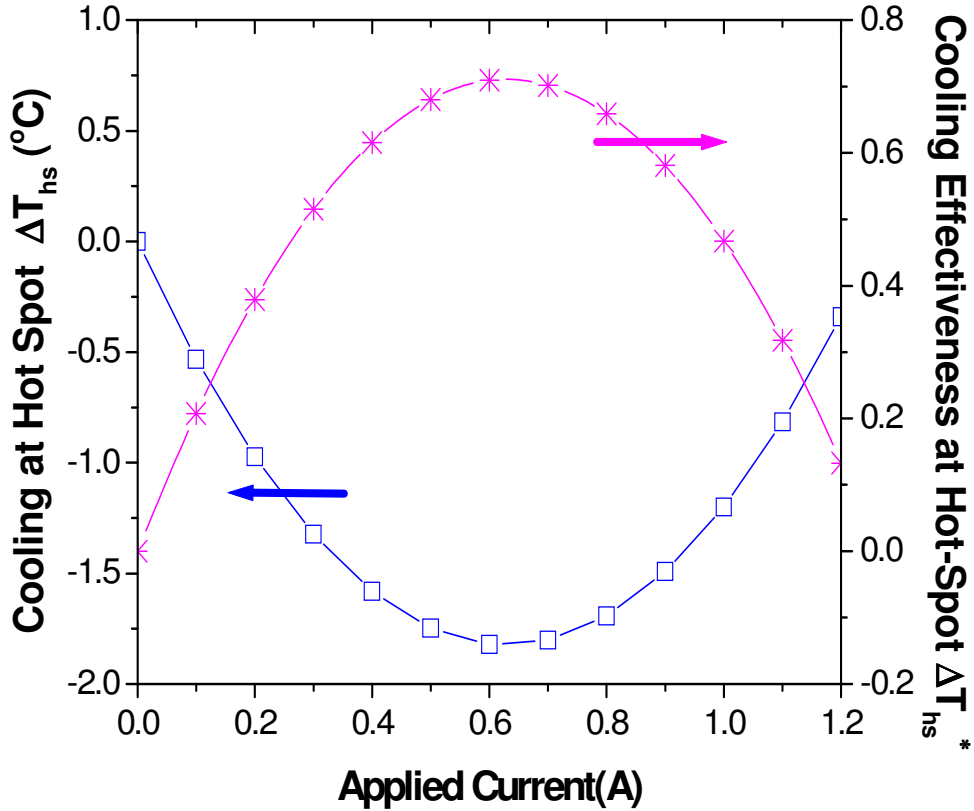
Figure 5.5: Typical hot spot cooling and cooling effectiveness for 150 µm × 150 µm microcooler. The hot spot size is 70 µm × 70 µm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2, the die thickness is 50 µm, the microcooler size is 150 µm × 150 µm, the electric contact resistance is 1.0 × 10^{-6} Ω.cm^2, and the boron doping concentration is 2.5 × 10^{19} cm^{-3} in silicon.

5.2.1.2 Heat Flux Profiles

Figure 5.6 shows the heat flux distribution around the hot spot and the microcooler inside the die, highlighting the physical advantage of using a silicon microcooler to locally cool down the hot spot. In the absence of a microcooler, Figure 5.6(a) shows the heat from the hot spot spreading in a nearly radially symmetric

manner in the 50 µm thick silicon die. When a 150 µm × 150 µm microcooler is

positioned right above the hot spot and activated with a current of 0.6 A, the heat

leaving the hot spot is drawn towards the microcooler and the two-dimensional heat

flow pattern seen in Figure 5.6 (b) approximates classic source-sink diffusion in a

slab. Due to the use of an oversized microcooler of 150 µm × 150 µm in size to

remove a small hot spot of 70 µm × 70 µm in size, the silicon microcooler not only

cools the hot spot but also cools the surrounding silicon.
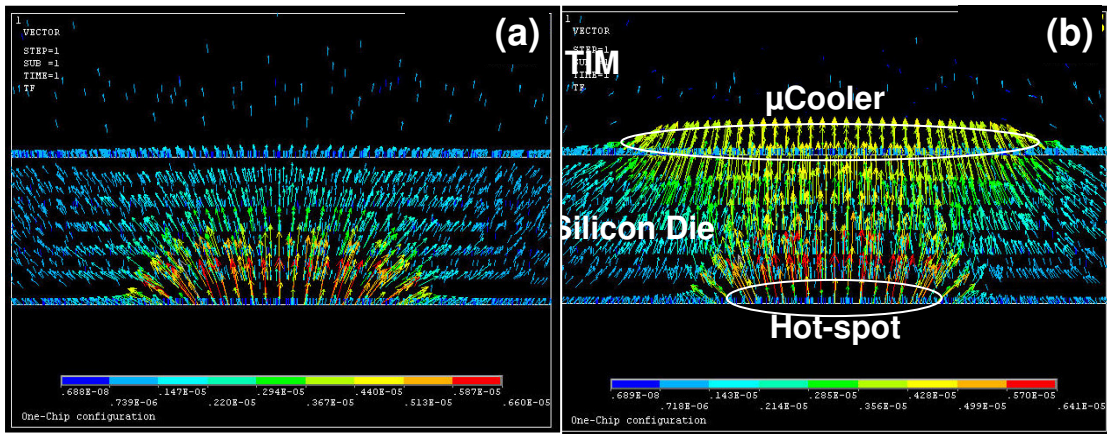
Figure 5.6: Heat flux distribution inside the silicon die: (a) without silicon microcooler, and (b) with 150 µm × 150 µm microcooler activated with a current of 0.6 A. The hot spot size is 70 µm × 70 µm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2, the die thickness is 50 µm, the electric contact resistance is 1.0×10^{-6} Ω.cm^2, and the boron doping concentration is 2.5×10^{19} cm^{-3} in silicon.

5.2.1 Doping Concentration Effect

To attain the highest possible thermoelectric cooling, it is necessary to obtain

as large a Seebeck coefficient and as low an electrical resistivity as possible and, in

particular, to maximize the value of power factor S^2/ρ. Here we use 150 µm × 150 µm

microcooler and 50 µm thick die as an example to explore doping effect on hot spot cooling and to search for optimum doping concentration under various conditions. Once the optimized doping concentration is found, it can be applied to different microcooler sizes or different die thicknesses and thereby optimize geometric configurations to maximize hot spot cooling performance. Figure 5.7 shows the hot spot cooling versus applied currents for 150 µm × 150 µm microcooler under different boron doping concentrations in silicon. The hot spot size is 70 µm × 70 µm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2 and the electric contact resistance is 1.0×10^{-6} Ω.cm^2 in this case study. It can be seen from Figure 5.7 that with the electric contact resistance of 1×10^{-6} Ω.cm^2, as the doping concentration increases, the optimized current extends to a higher value due to lower electrical resistivity of the silicon, while the maximum temperature reduction at the hot spot initially increases and, after reaching a maximum value of 1.8^{o}C at the doping concentration of 2.5×10^{19} cm^{-3}, it decreases with further changes. The reason is that increasing doping concentration results in lower electrical resistivity and, as a consequence, less Joule heating in the silicon substrate. Unfortunately, the Seebeck coefficient of silicon also decreases with increasing doping concentration, which leads to less Peltier cooling. The competition between these two factors results in an optimum doping concentration at which the maximum cooling performance could be obtained.

Figure 5.8 gives the variation of the maximum hot spot cooling with boron doping concentration for different electric contact resistances. It shows the combined effect of the parasitic electrical contact Joule heating and doping for five different

electric contact resistance ranging from 1.0×10^{-8} $\Omega.\text{cm}^2$, a theoretical electric contact resistance achievable at silicon/metal interface, to 1.0×10^{-5} $\Omega.\text{cm}^2$, a laboratory deteriorated interface value, on the cooling performance. In all cases, the dependence of maximum temperature reduction on the doping concentration follows a similar trend, but the optimized doping concentration seems to depend on the electric contact resistance. As shown in Figure 5.8, higher electrical contact resistance pushes the optimized doping to a lower level, i.e. for a specific contact resistance of 1.0×10^{-5} Ωcm^2, the optimized doping concentration occurs at around $\sim 1.5 \times 10^{19}$ cm^{-3}, while for a specific contact resistance of 1.0×10^{-7} Ωcm^2 or better, the maximum temperature reduction occurs at the doping concentration of 2.5×10^{19} cm^{-3}. This result is in excellent agreement with the theoretical optimum "power factor" value shown in Figure 2.14. It can be expected that if the microcooler size is very small and the specific contact resistance is very large, the optimized doping concentration would be pushed to values below 1×10^{19} cm^{-3}. Therefore, care is required in the selection of doping concentration for this silicon microcooler application and, in the presence of higher parasitic losses, best cooling results may be obtained with lower doping concentrations. In addition it is found that the optimized doping concentration is independent of geometric configurations, like microcooler size and die thickness. In all of the geometric ranges we investigated, if the electric contact resistance is less than 1.0×10^{-6} Ωcm^2 or better, the optimized doping concentration always occurs at 2.5×10^{19} cm^{-3}.
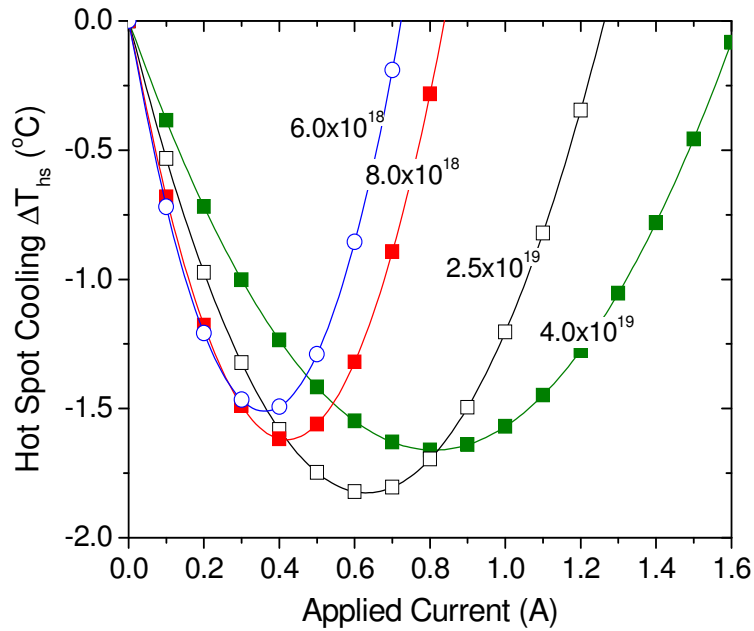
Figure 5.7: Variation of hot spot cooling with applied current for 150 μm × 150 μm microcooler on 50 μm thick silicon die. The hot spot size is 70 μm × 70 μm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2 and the electric contact resistance is 1.0×10^{-6} Ω.cm^2.
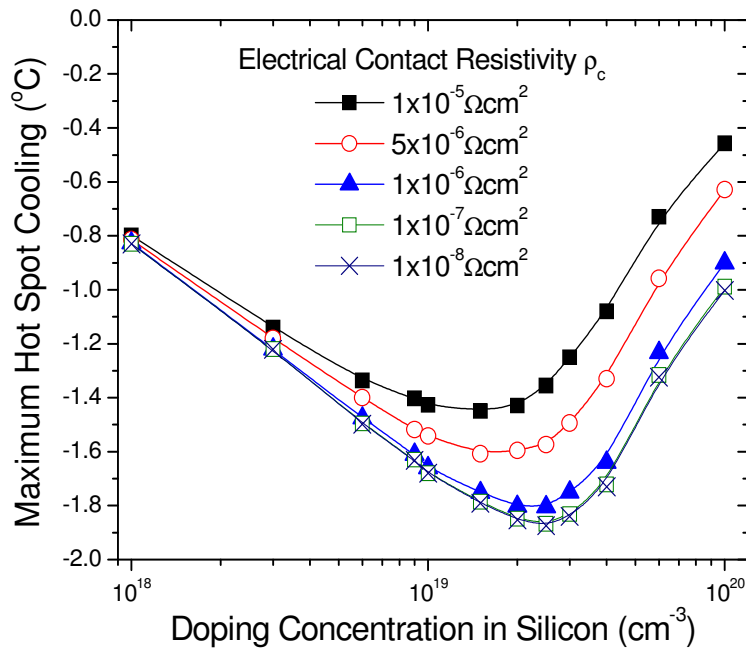
Figure 5.8: Dependence of maximum hot spot cooling on boron doping concentration in silicon for 150 μm × 150 μm microcooler on 50 μm thick silicon die. The hot spot size is 70 μm × 70 μm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2.

5.2.2 Silicon Die Thickness Effect

The influence of the silicon die thickness on thermoelectric cooling performance is illustrated in Figure 5.9 and 5.10. Increasing the die thickness from 10 µm to 500 µm is seen to lower the cooling at the hot spot but raise the temperature reduction at the surface of the microcooler. It is also found that with the decrease of die thickness the maximum cooling at the hot spot approaches that of the microcooler. Clearly, a higher cooling capability at the microcooler does not guarantee a higher cooling at the hotspot. Under the stated conditions, at 500 µm die thickness the microcooler temperature has decreased by more than 4°C while the hot spot has been cooled by just 0.1°C. Moreover, the optimum current for cooling at the hotspot - the true target of this thermal management approach - is much smaller than that required to maximize cooling of the microcooler itself if the die thickness is as large as 500 µm, as indicated in Figure 5.10. However, when the die thickness is very thin, like 10 µm, the cooling vs. current curves for the hot spot and that for the microcooler become very similar.

Since the silicon die works as both a thermoelectric material to provide thermoelectric cooling power and a thermal conductor to provide a dissipation path for the heat generated in the chip, the silicon die thickness will have an influence on both the Joule heating in the silicon die and the thermal resistance between the microcooler and the hotspot. With the decrease of the die thickness, the thermal resistance between the micro-cooler and the hot spot decreases while the electrical resistance and thus Joule heating in the silicon die becomes larger. These two factors compete with each other, leading to the results shown in Figure 5.9.
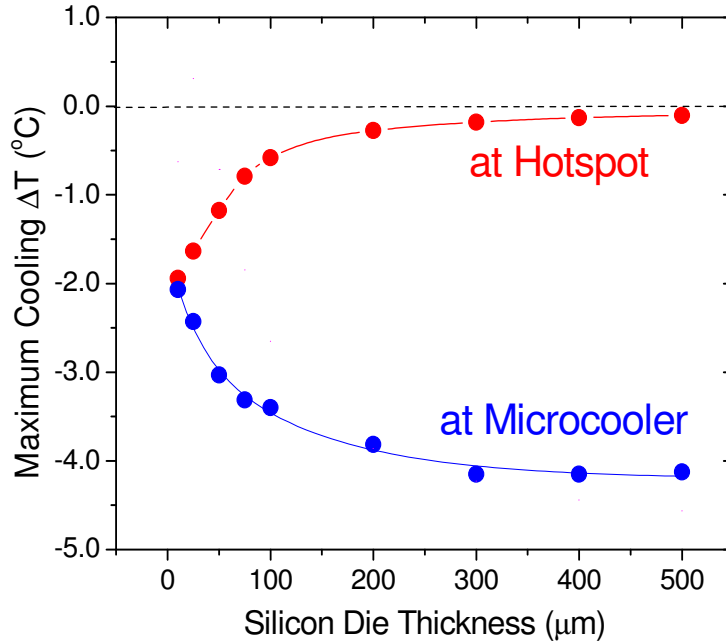
Figure 5.9: Dependence of maximum hot spot cooling on die thickness. The microcooler size is 70 μm × 70 μm, the hot spot size is 70 μm × 70 μm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2 and the electric contact resistance is 1.0×10^{-6} Ω.cm^2.
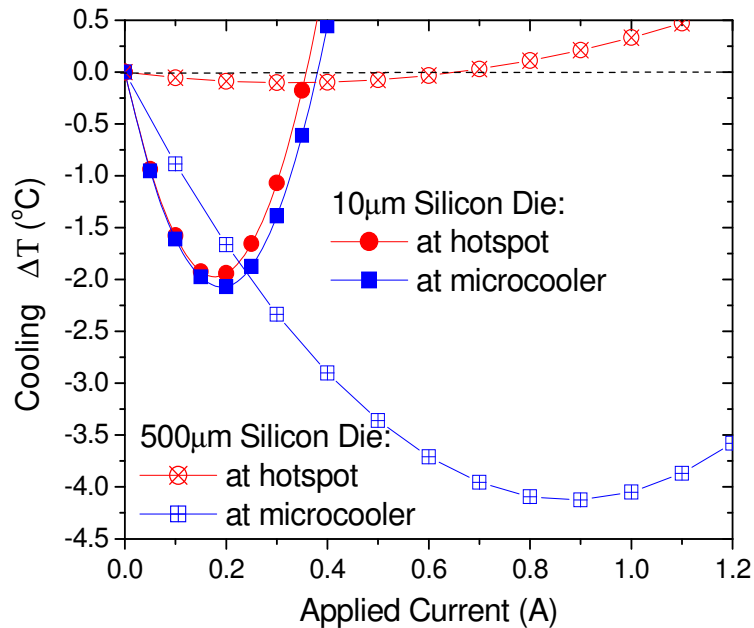
Figure 5.10: Variation of cooling with applied current for two different die thicknesses. The microcooler size is 70 μm × 70 μm, the hot spot size is 70 μm × 70 μm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2 and the electric contact resistance is 1.0×10^{-6} Ω.cm^2.

5.2.3 Microcooler Size Effect

Figure 5.11 shows the variation of the maximum cooling with microcooler size for a die thickness of 50 µm and doping concentration of $2.5 \times 10^{19} cm^{-3}$. As the microcooler size increases from 35 µm to 600 µm, cooling of the hot spot first improves, reaching its best value at 250 µm and then decreases continuously. The microcooler temperature follows a similar trend, though it reaches its lowest value at a far smaller size. Figure 5.12 displays the typical relationship of cooling performance versus applied current for the micro-cooler sizes of interest, showing an inverse parabolic temperature distribution with a distinct, yet different, optimum current, which minimizes the hot spot temperature, for each microcooler size. Similar to the case of thicker die, if the microcooler size is larger, like 300 µm × 300 µm, the cooling vs. current curves for the hot spot and that for the microcooler become very similar.

It is to be noted that this optimum current is associated with a balance between the beneficial effect of the thermoelectric energy conversion at the microcooler/silicon interface and Joule (resistive) heating due to the current flow in the silicon. The effect of microcooler size on cooling performance involves the interplay of three factors: thermoelectric energy conversion at the interface, Joule heating in the nearby silicon, and thermal diffusion from the hot spot to the microcooler. For small microcoolers, the high electrical current flux results in more intense local cooling, but also higher Joule heating in the silicon adjacent to the microcooler and in higher thermal resistance (or poorer "form factor" in the vernacular of multidimensional thermal conduction) from the hot spot to the

microcooler. Consequently, the choice of small microcooler results in a low microcooler temperature that doesn't translate into effective hot spot cooling. Alternatively, for very large microcoolers, the resistance between the hot-spot and micro-cooler is reduced - drawing the hot spot temperature very close to the microcooler temperature - but Joule heating in the substrate limits the total current flow to values that yield lower current flux at the microcooler than achievable for small microcoolers, thus resulting in less cooling and higher micro-cooler temperatures.
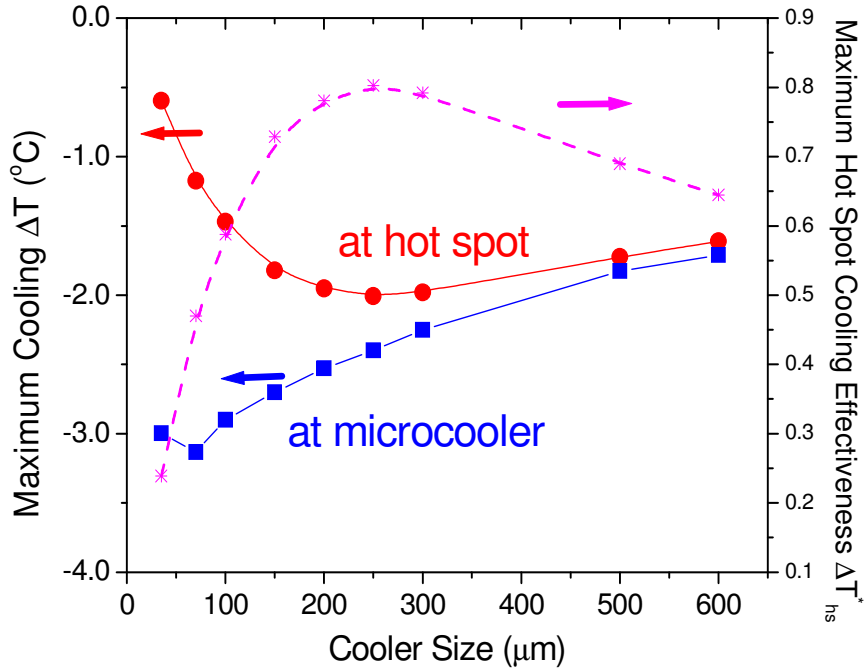
Figure 5.11: Variation of maximum cooing on microcooler sizes. The hot spot size is 70 μm × 70 μm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2 and the electric contact resistance is 1.0×10^{-6} Ω.cm^2.
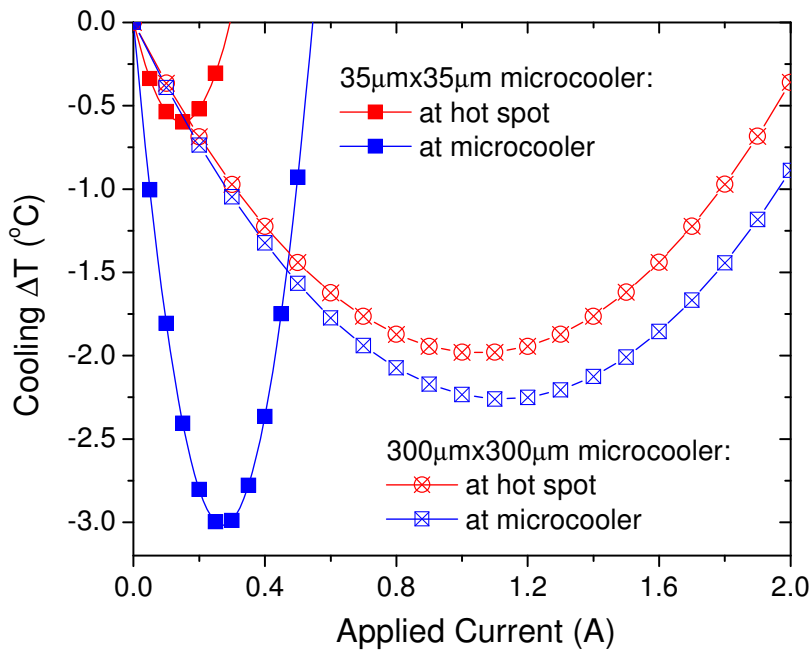
Figure 5.12: Variation of cooling with applied current for two different microcooler sizes. The hot spot size is 70 μm × 70 μm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2 and the electric contact resistance is 1.0×10^{-6} Ω.cm^2.

It was found the optimized microcooler size depends on die thickness, and the thicker the silicon die the larger the optimized microcooler size with which silicon microcooler can achieve the highest hot spot cooling. Figure 5.13 shows the variation of maximum cooling at the hot spot and at the microcooler for various microcooler sizes on 25 µm, 50 µm, 100 µm and 200 µm thick silicon die. The hot spot size is 70 µm × 70 µm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2, the electric contact resistance is 1.0×10^{-6} Ω.cm^2 and the doping concentration is 2.5 × 10^{19} cm^{-3}. We can find that for 25 µm thick die - the thinnest die we investigated, if the microcooler size is 150 µm × 150 µm, the optimized microcooler size for hot spot cooling, silicon microcooler can achieve a maximum cooling of 2.06^{o}C at the hot spot and 2.63^{o}C at the microcooler. However, if the die thickness is increased to 200 µm, the maximum cooling at the hot spot will reduce to 1.63^{o}C while the maximum cooling at the microcooler will increase to 4.07^{o}C. Apparently, the thinner silicon die leads to larger cooling at the hot spot but less cooling at the microcooler itself. Table 5.2 summarizes the maximum cooling at the hot spot and at the microcooler as well as the optimized geometric configuration to maximize hot spot cooling temperature, indicating that the optimum ratio of micrcooler size to die thickness is around 6 for 25 µm thick die and reduces to 4 for 200 µm thick die.

It is a little bit complicated to compare hot spot cooling effectiveness as it is also dependent of temperature rise due to the self-heating of the hot spot. Figure 5.14 (a) shows the dependence of the hot spot temperature rise and the hot spot cooling on the die thickness. It is found that with increasing die thickness, the hot spot temperature rise reduce due to better heat spreading effect, decreasing from 3.08^{o}C

for 25 µm thick die to 2.19^{o}C for 200 µm thick die. On the other hand, the hot spot

cooling also decreases with increase of the die thickness, decreasing from 2.06^{o}C for

25 µm thick die to 1.63^{o}C for 200 µm thick die. The combined effect is that there is

an optimized die thickness with which silicon microcooler can achieve the highest hot

spot cooling effectiveness, as shown in Figure 5.14(b). Under the condition we

investigated, this optimized thickness occurs at 100 µm die thickness, with which the

hot spot cooling effectiveness of 0.85 is achieved, indicating that 85% of temperature

rise at the hot spot can be suppressed. Although silicon microcooler on 25 µm thick

die provides the highest cooling temperature of 2.07 ^{o}C among all the die thickness

we investigated, it can only reduce about 67% of hot spot temperature rise. On the

other hand, silicon microcooler on 200 µm thick die provides the hot spot cooling

temperature of 1.67^{o}C, the lowest among all the die thickness we investigated and,

however, it can still reduce about 74% hot spot temperature rise.

Table 5.2: Summary of cooling performance for various die thicknesses and microcooler sizes under the electric contact resistance of 1×10^{-6} Ω.cm^2 and the boron doping concentration of 2.5 × 10^{19} cm^{-3}.

Die Thickness (µm)	25	50	100	200
Optimum Cooler Size (µm)	150	250	500	800
Optimum ratio	6:1	5:1	5:1	4:1
Maximum Hot Spot Cooling (^{o}C)	2.07	2.03	1.97	1.63
Maximum Cooling at Cooler (^{o}C)	2.67	3.08	3.71	4.07
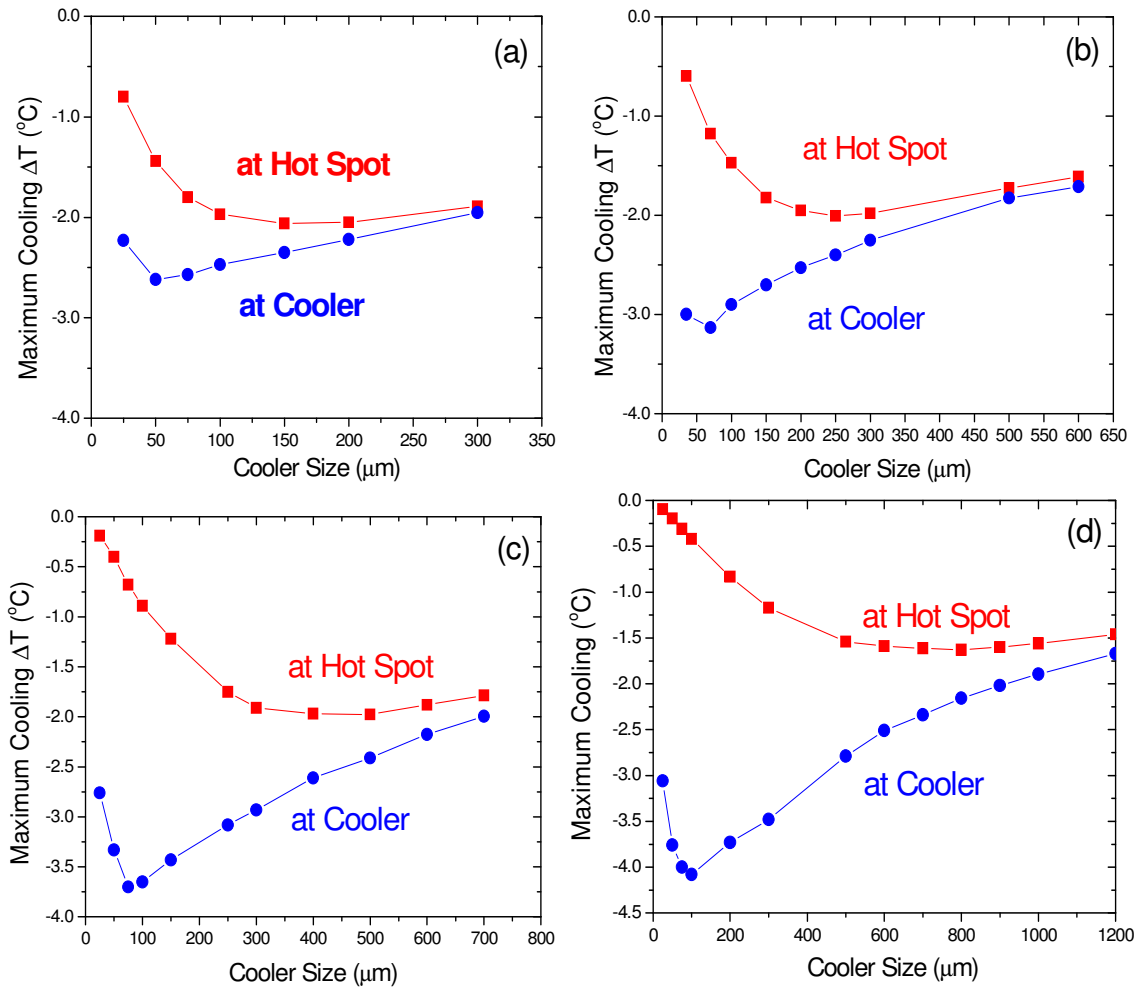
Figure 5.13: Variation of cooling at silicon microcooler and at hot spot with applied current for different die thickness and microcooler sizes: (a) 25 µm thick die, (b) 50 µm thick die, (c) 100 µm thick die and (d) 200 µm thick die. The electric contact resistance is 1×10^{-6} Ω.cm^2 and the boron doping concentration is 2.5 × 10^{19} cm^{-3}.
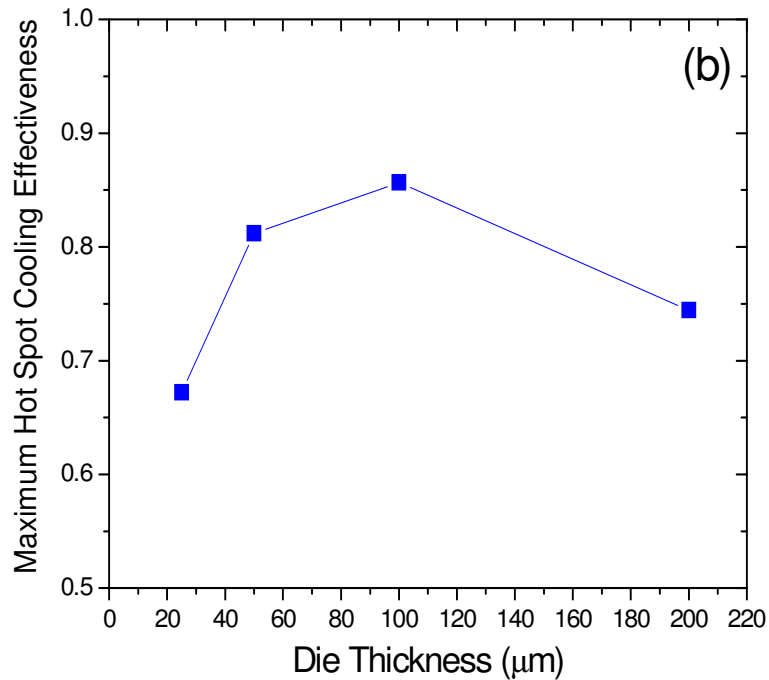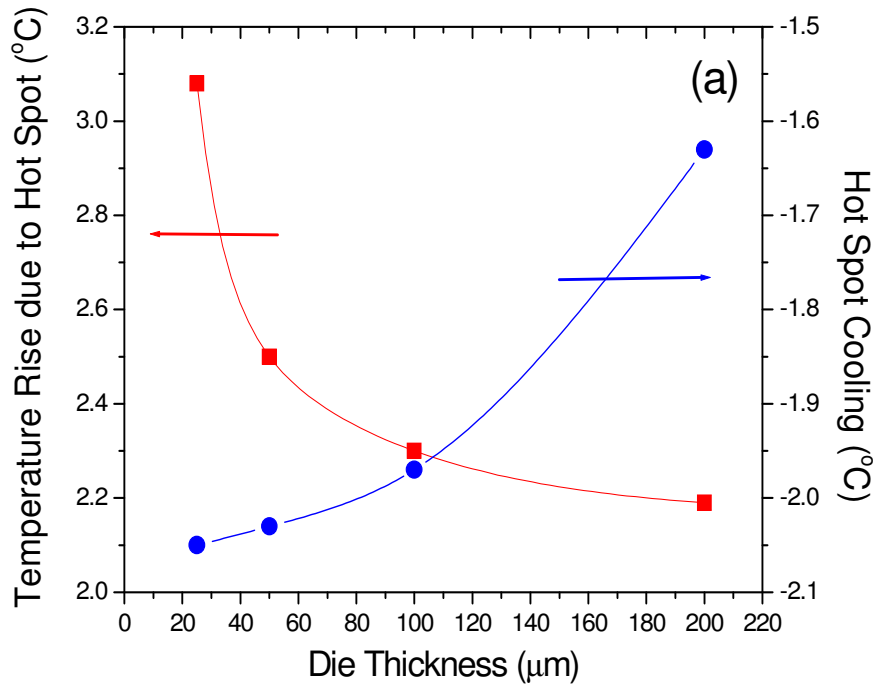
Figure 5.14: Effect of silicon die thickness on (a) maximum hot spot cooling and (b) maximum hot spot cooling effectiveness. The electric contact resistance is 1×10^{-6} $\Omega.cm^2$ and the boron doping concentration is 2.5×10^{19} cm^{-3}. The microcooler size is optimized at 150 μm × 150 μm, 250 μm × 250 μm, 500 μm × 500 μm, 800 μm × 800 μm, for 25 μm, 50 μm, 100 μm and 200 μm thick die, respectively.

5.2.4 Hot Spot Size Effect

In this study, a 150 µm × 150 µm silicon microcooler integrated on a 50 µm thick die is used as an example to study the effect of hot spot size on hot spot cooling performance. The hot spot size varies from 35 µm × 35 µm to 300 µm × 300 µm with a heat flux of 680 W/cm^2, corresponding to the power dissipation on the hot spot from 8.3 mW to 613 mW. The electric contact resistance is assumed to be 1×10^{-6} Ω.cm^2 at the metal/silicon interface and the boron doping concentration is assumed to be 2.5 × 10^{19} cm^{-3}. As shown in Figure 5.15, the hot spot cooling temperature changes very little with the hot spot size in our investigated range, increasing from around 1.79^{o}C cooling at 35 µm × 35 µm hot spot to 1.81^{o}C cooling at 300 µm × 300 µm hot spot. However, because small hot spot produces small temperature rise, about 1°C temperature rise at 35 µm × 35 µm hot spot versus about 13°C temperature rise at 300 µm × 30 µm hot spot, if 150 µm × 150 µm microcooler is turned on with an optimized current of 0.6 A to suppress 35 µm × 35 µm hot spot, it can overcool this hot spot by 0.8^{o}C. However, if such a microcoole is applied to suppress larger hot spot, it becomes ineffective. Figure 5.15 shows that 150 µm × 150 µm microcooler can remove only 14% of temperature rise at 300 µm × 300 µm hot spot.
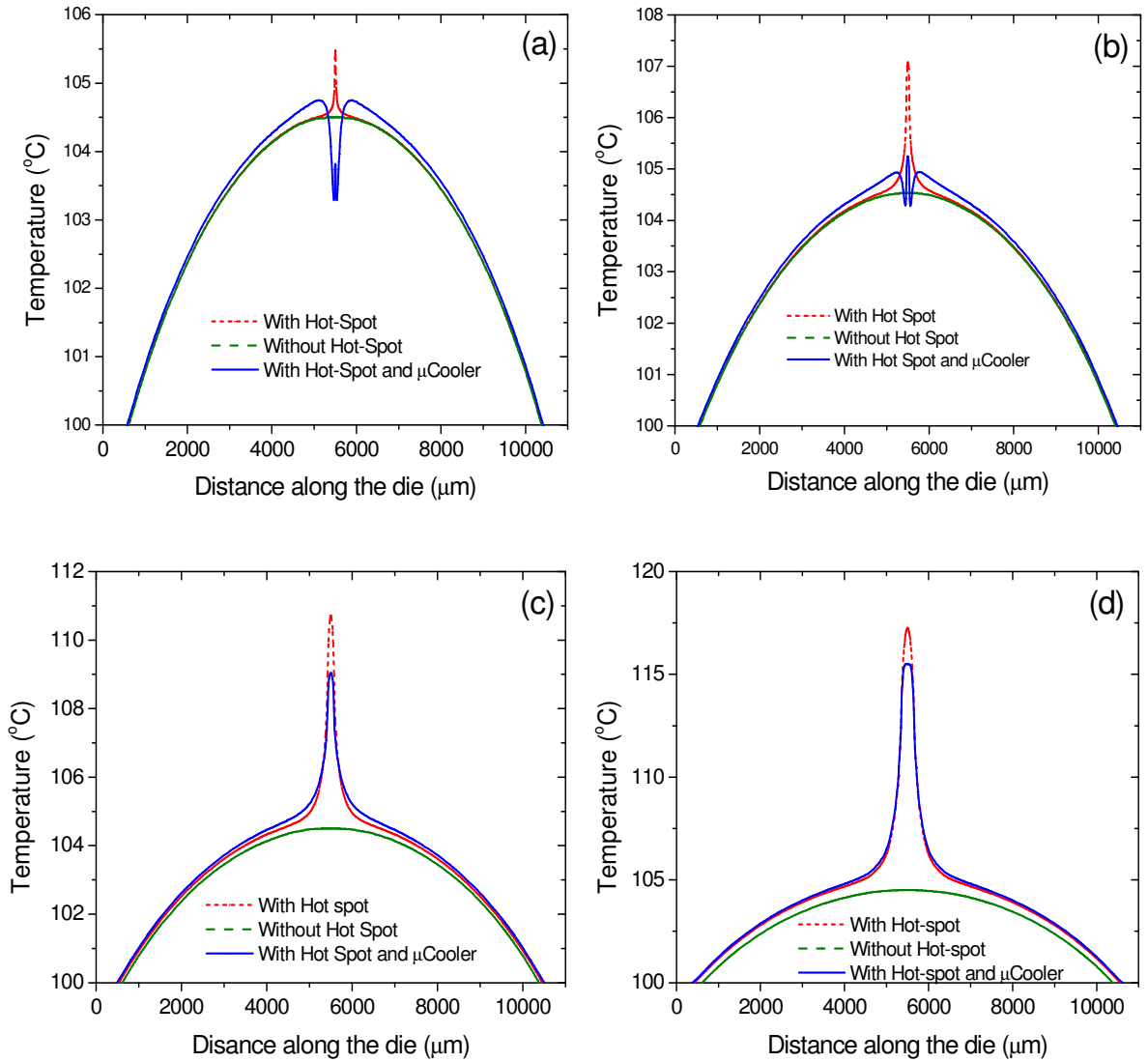
Figure 5.15: Effect of hot spot size hot spot cooling: (a) 35 μm × 35 μm hot spot, (a) 70 μm × 70 μm hot spot, (c) 150 μm × 150 μm hot spot, and (d) 300 μm × 300 μm hot spot. The heat flux of the hot spot is 680 W/cm². The electric contact resistance is 1×10⁻⁶ Ω.cm², the boron doping concentration is 2.5 × 10¹⁹ cm⁻³, and the microcooler size is 150 μm × 150 μm.

163

5.2.5 Electrical Contact Resistivity Effect

The miniaturization of the thermoelectric microcooler tends to exacerbate the deleterious effects of the electrical contact resistance expected to occur at the interface between the micro-cooler and the metal contact. If the electrical contact resistivity is constant, the smaller cooler size will result in larger electrical contact resistance and thus larger Joule heating at the metal/silicon interface and increase the parasitic heat load. In general, for a semiconductor/metal contact, the carrier transport mechanism can be separated into three regions: field emission, thermionic-field emission and thermionic emission, depending on the doping concentration of the semiconductor. Under high doping concentration ($N_d > 10^{20} cm^{-3}$), the field emission is dominant at the semiconductor/metal interface and the theoretical electrical contact resistivity is given by [115]:

$$\rho_c = \frac{1}{E_{00}} \exp(\frac{\varphi_b}{E_{00}})$$ (5.5)

where φ_b is the barrier height and E_{00} is a reference energy which depends on the doping concentration N_d, the semiconductor permittivity ε_s and the effective mass m*:

$$E_{00} = \frac{qh}{4\pi}(\frac{N_d}{\varepsilon_s m^*})^{1/2}$$ (5.6)

The theoretical value of the electrical contact resistivity between high-doped silicon and the metal contact, such as Al, PtSi and CoSi, is in the range of 1×10^{-8} $\Omega.cm^2$ to 1×10^{-9} $\Omega.cm^2$ at room temperature or above [116]. However, due to process-related

limitations, the typical electrical contact resistivity between silicon and Ti/Al/Au is much higher and may vary from batch to batch, usually ranging from 1×10^{-5} to 1×10^{-7} $\Omega.cm^2$.

Figure 5.16 gives the maximum cooling and cooling effectiveness at hot spot as a function of electrical contact resistivity for different micro-cooler sizes. In all cases, as the electrical contact resistivity increases, the cooling effect is degraded continuously and electrical contact resistivity has larger impact for smaller cooler sizes: if the electrical contact resistivity decreases from 1×10^{-5} $\Omega.cm^2$ to 1×10^{-7} $\Omega.cm^2$, the maximum cooling and cooling effectiveness at hot-spot could be improved by as large as 300 % for 35µm × 35µm microcooler while only 20% for 150 µm ×150 µm microcooler. Therefore, for smaller microcoolers, reducing the electrical contact resistivity to less than 1×10^{-6} $\Omega.cm^2$ is very critical.
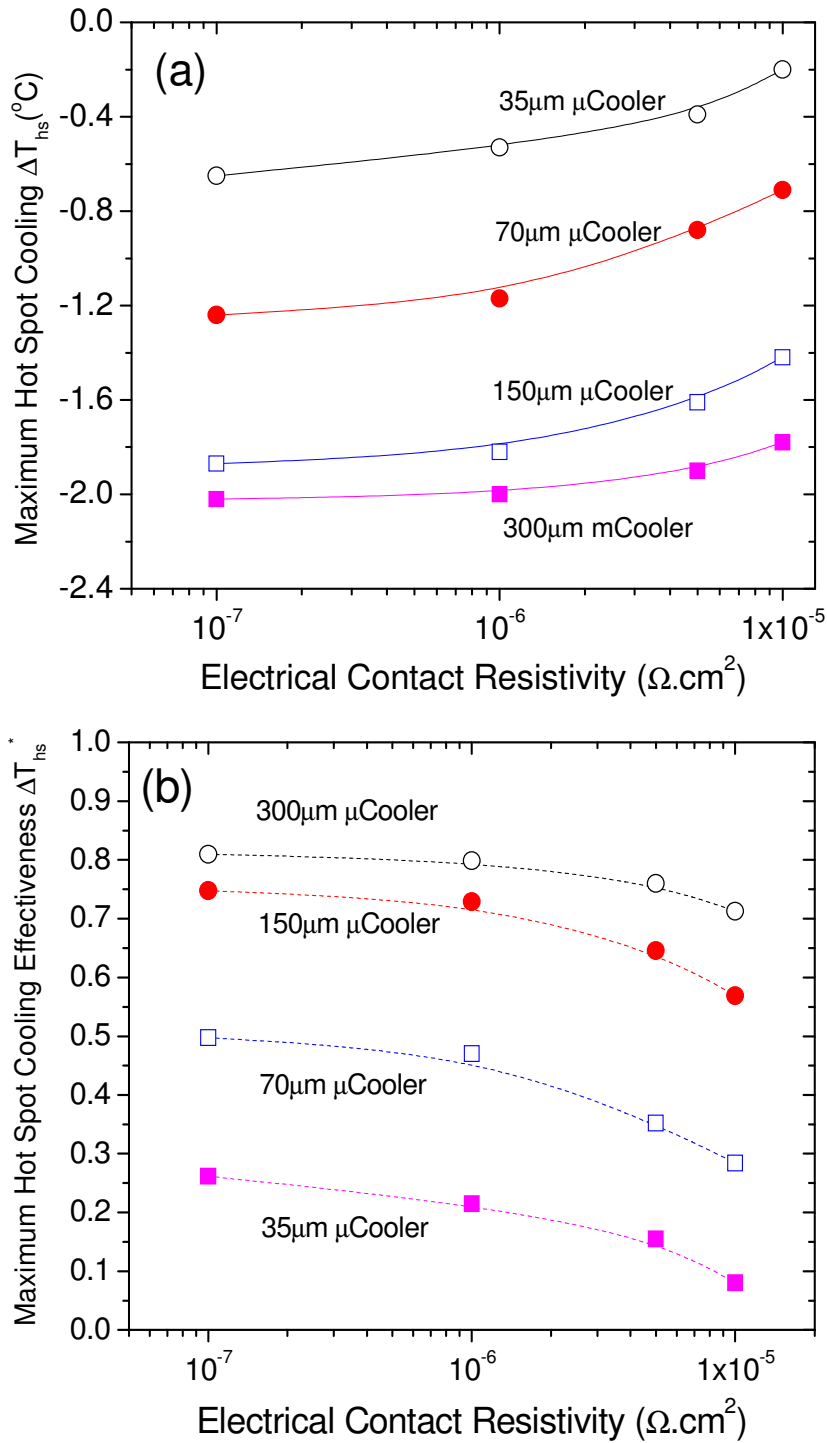
Figure 5.16: Dependence of maximum hot spot cooling on electrical contact resistivity: (a) maximum hot spot cooling, and (b) maximum hot spot cooling effectiveness. The hot spot size is 70 µm × 70 µm with the heat flux of 680 W/cm^2, the background heat flux is 70 W/cm^2 and the boron doping concentration in silicon is 2.5 × 10^{19} cm^{-3}.

5.2.6 Multiple Microcooler Effect

The foregoing discussion has focused on the use of just one microcooler, but attention will now turn briefly to the potential benefits of applying multiple microcoolers to thermal control of the hotspot. In the absence of silicon Joule heating effects, use of multiple microcoolers would be expected to always lower the hotspot temperature. However, the additional electric current needed to operate multiple microcoolers results in a significant increase in Joule heating within the silicon and necessitates careful thermal-electrical optimization to achieve the desired improvement in cooling effectiveness.

The temperature profile developed along the centerline on the bottom (active) surface of a 50µm thick silicon chip is shown in Figure 5.17. It is observed that the presence of the background heat flux of 70W/cm^2 on the bottom surface of the silicon chip produces a parabolic temperature distribution, peaking at the center of the chip at 104.51$^\text{o}$C. When the 70µm×70µm hotspot with a heat flux of 680W/cm^2 is activated at the center of the silicon chip, the peak temperature increases to 107.05$^\text{o}$C, an increase of 2.5$^\text{o}$C. When a single 70µm×70µm microcooler is monolithically integrated onto the top surface of the silicon chip to cool the hotspot, at an optimized current of 0.3A, the predicted peak hotspot temperature is reduced to 105.95$^\text{o}$C, leading to removal of 44% of the temperature rise due to the hotspot. Moreover, if nine microcoolers (each microcooler is 70µm×70µm in size) are integrated onto the silicon chip, at the optimized current of 0.1A for each microcooler, the peak hotspot temperature is reduced to 105.05$^\text{o}$C, nearly doubling the cooling effectiveness to approximately 80%.
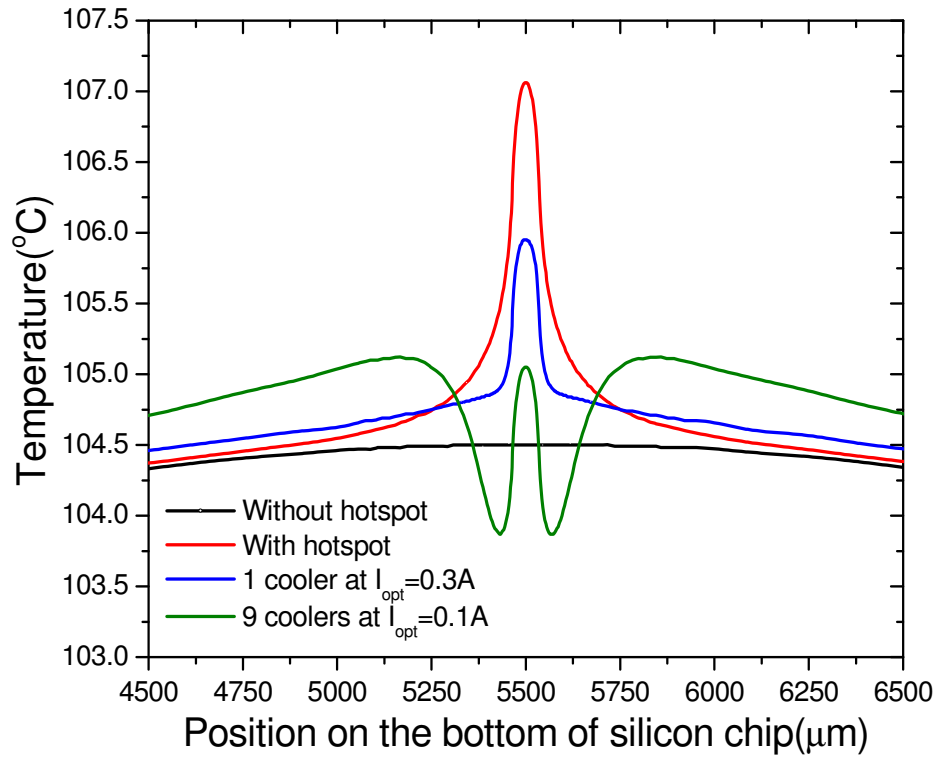
Figure 5.17: Temperature profile on the bottom of silicon chip (Chip thickness is 50µm, each microcooler size is 70µm×70µm, cooler-to-cooler gap is 5µm, boron doping concentration is 2.5×10^{19} cm^{-3}, and specific contact resistance is $1 \times 10^{-6} \Omega.cm^2$).

Figure 5.18 illustrates the results for the temperature reduction and cooling effectiveness achieved at the hotspot when cooled with multiple 70µm×70µm microcoolers on 50µm thick silicon chip. The maximum temperature reduction at the hotspot is seen to increases from 1.1°C (0.44 cooling effectiveness) for a single microcooler to 1.6°C (0.62 cooling effectiveness) for three microcoolers, 1.9°C (0.75 cooling effectiveness) for five microcoolers, and a modest increase to 2.0°C (0.80 cooling effectiveness) for nine microcoolers.
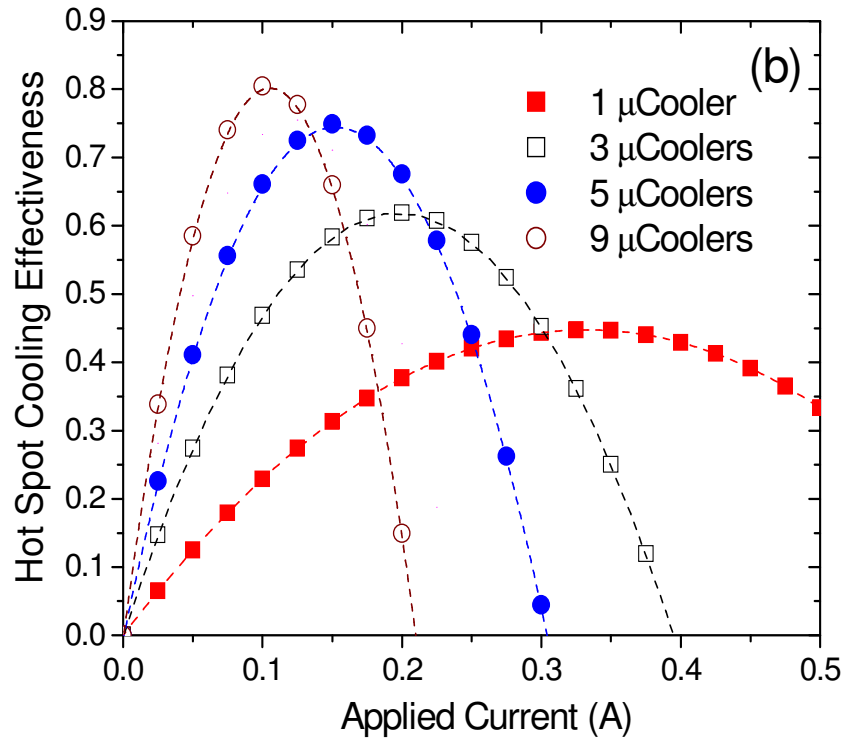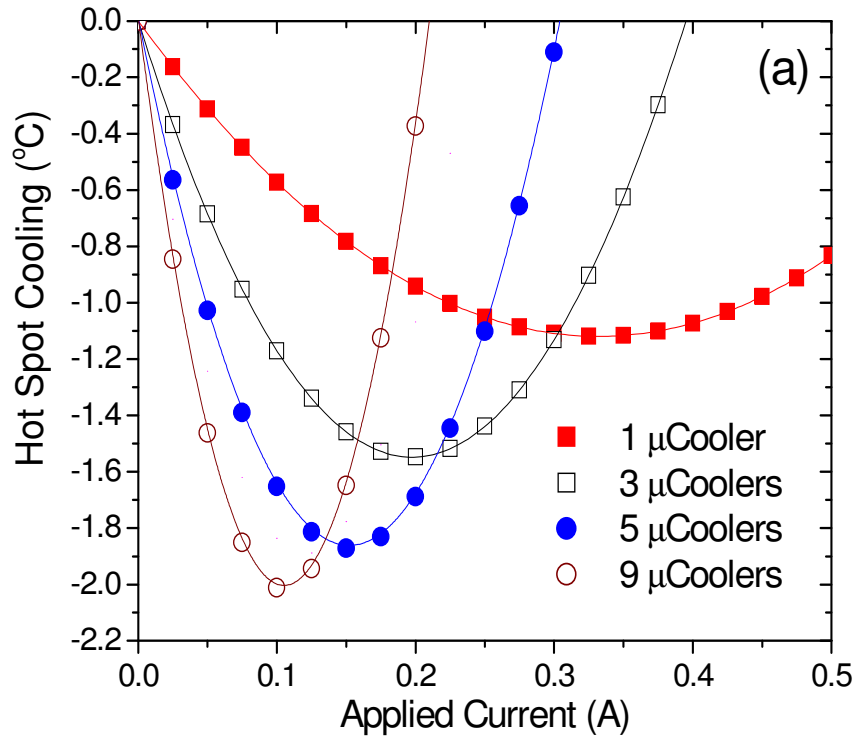
Figure 5.18: Variation of (a) hotspot cooling and (b) cooling effectiveness with electric current applied on each microcooler (Chip thickness is 50µm, each microcooler size is 70µm×70µm, cooler-to-cooler gap is 5µm, boron doping concentration is $2.5\times10^{19}cm^{-3}$, and specific contact resistance is $1\times10^{-6}\Omega.cm^{2}$).
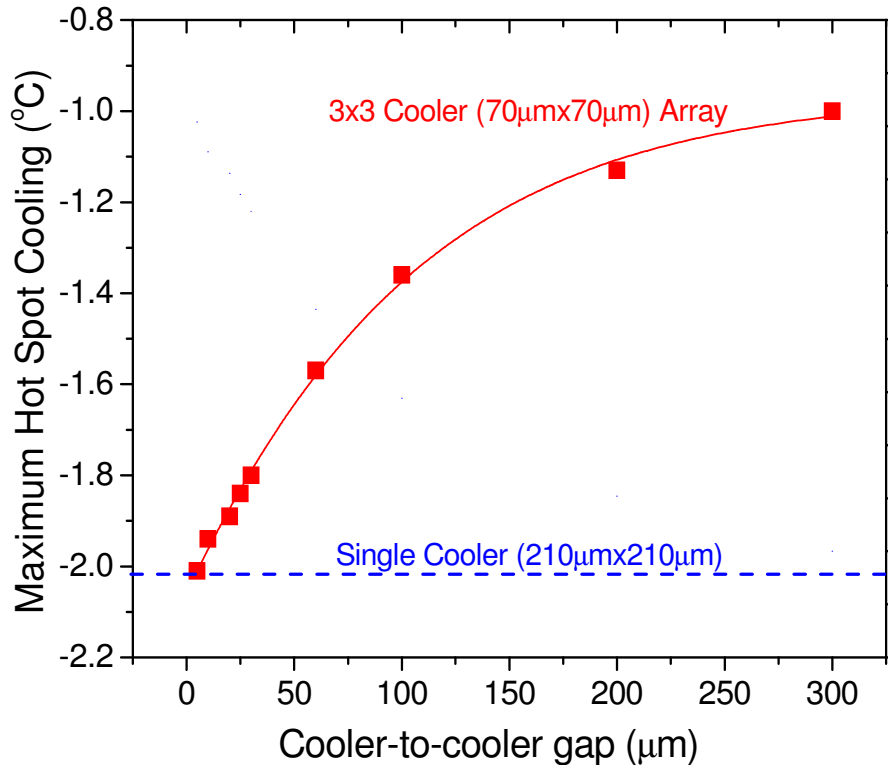
Figure 5.19: Effect of cooler-to-cooler gap on maximum hotspot cooling (Each microcooler size is 70µm×70µm, silicon chip thickness is 50µm, boron doping concentration is 2.5×10^{19}cm^{-3}, and specific contact resistance is $1\times10^{-6}\Omega.$cm^2).

When the distance among the microcoolers in a multi-cooler array shrinks, the interaction of multiple electric currents, each injected from a single microcooler and flowing into the silicon chip, will intensify the Joule heating in the silicon adjacent to the hotspot. On the other hand, more closely spaced microcoolers are advantageously positioned to absorb the hotspot heat. For the investigated configuration, a decrease of the spacing of the microcoolers, or cooler-to-cooler gap, is found to always improve the hotspot cooling performance, and the closer the microcoolers, the larger temperature reduction at the hotspot. As shown in Figure 5.19, for the present configuration, best results are obtained with the closest microcooler spacing. As the

gap grows, the hotspot temperature reduction deteriorates. It should be noted that when the multiple microcooler gets closer and closer, its hotspot cooling effect will be equivalent to a single larger cooler which has the same cooling surface area as multiple coolers. As illustrated in Figure 5.19, when the 3×3 nine coolers (70µm×70µm for each microcooler size) get very close, its cooling performance will be very close to that of a single larger cooler (210µm×210µm for microcooler size). Therefore, the maximum hotspot cooling performance for multi-cooler configuration should be same as its equivalent large cooler. However, the use of such microcooler arrays can be expected to enhance the electrical design flexibility, fabricability, and reliability of this cooling strategy and to facilitate cooling of multiple or spatially variable hotspots on a single chip.

5.3. Conclusions

This chapter provides a methodology for the three-dimensional thermal-electrical finite-element simulation of silicon thermoelectric microcooler used to remove hot spot on packaged silicon chips. The effects of doping concentration in the silicon, die thickness, microcooler size, electrical contact resistance, and hot spot size are examined and the most promising configurations are examined and discussed. Hot spot cooling is seen to improve by decreasing the chip thickness and electric contact resistance. It is found that to achieve the best cooling performance the microcooler size needs to be optimized, with larger microcooler size preferable for thicker silicon chips and that the optimum silicon doping concentration decreases with higher specific electrical contact resistance. Under optimized conditions, it was found that

around 2°C cooling can be achieved on the hot spot on 25µm, 50µm and 100µm silicon die, and around 80% of the temperature rise at the hot spot, 70 µm × 70 µm in size with 680 W/cm^2 heat flux, can be removed.

Chapter 6

Mini-Contact Enhanced TEC for On-Chip Hot Spot Cooling

Solid state thermoelectric coolers (TECs) have been recently studied for hot spot thermal management because these solid state devices offer high reliability, can be locally applied for spot cooling, provide high cooling heat flux, and can be integrated with IC processing. However, the relatively low cooling flux, 5 ~ 10 W/cm^2, of conventional Bi$_2$Te$_3$-based TEC modules severely limits the application of these devices to hot spot remediation. In this chapter we propose the novel use of a copper mini-contact pad, which connects the thermoelectric cooler and the silicon chip and therefore concentrates the thermoelectric cooling power on the silicon chip, to significantly improve hot spot cooling performance [124,125]. A package-level numerical simulation (ANSYS) is developed to predict and optimize the on-chip hot spot cooling performance using such mini-contacts. The targeted hot spot is 400 µm × 400 µm with a heat flux of 1250 W/cm^2 and the dimensions of the silicon chip are 11 mm × 13 mm × 500 µm, with a background heat flux of 70 W/cm^2. Our attention is focused on the hot spot temperature reduction associated with variations in copper mini-contact size, thermoelectric element thickness and thermal contact resistance to explore the parametric sensitivities and establish the optimum cooling configuration.

6.1. Simulation of Mini-Contact Enhanced TEC for Hot Spot Cooling

The mini-contact pad, which is made of a high thermal conductivity material and is inserted between the TEC and the silicon chip to concentrate the thermoelectric

cooling power, is proposed as a novel approach to increasing the cooling flux on a local area on the top surface of the silicon chip. Figure 6.1 compares hot spot cooling mechanisms using a TEC without a mini-contact pad and a TEC enhanced with a mini-contact pad, indicating that the mini-contact pad can effectively increase the spot cooling capability through concentration of the thermoelectric cooling power at the reduced cross-sectional area of the mini-contact tip. It can be expected that - to a first approximation - the smaller the mini-contact tip, the larger the cooling flux on the top of the silicon chip. The other unique benefit of the mini-contact is that it allows the TEC to act mostly on the hot spot heat load. Since the TEC does not have to pump the additional background heat load, either the TEC input power requirement is reduced or the hot spot can be cooled further for the same input power on the TEC.
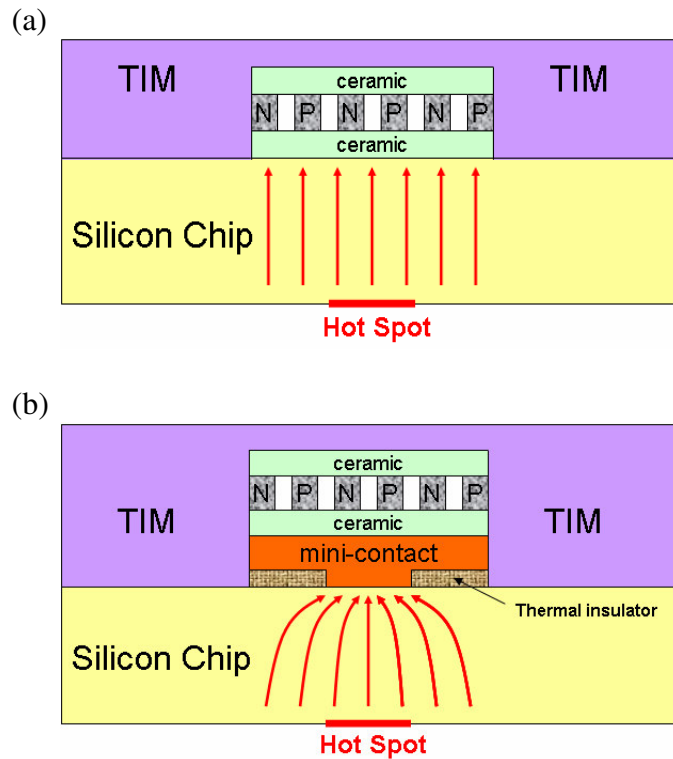
Figure 6.1: Schematic of a thermoelectric cooler attached on silicon chip and embedded inside thermal interface material (TIM) of chip package: (a) TEC without a mini-contact and (b) TEC with an integrated mini-contact.
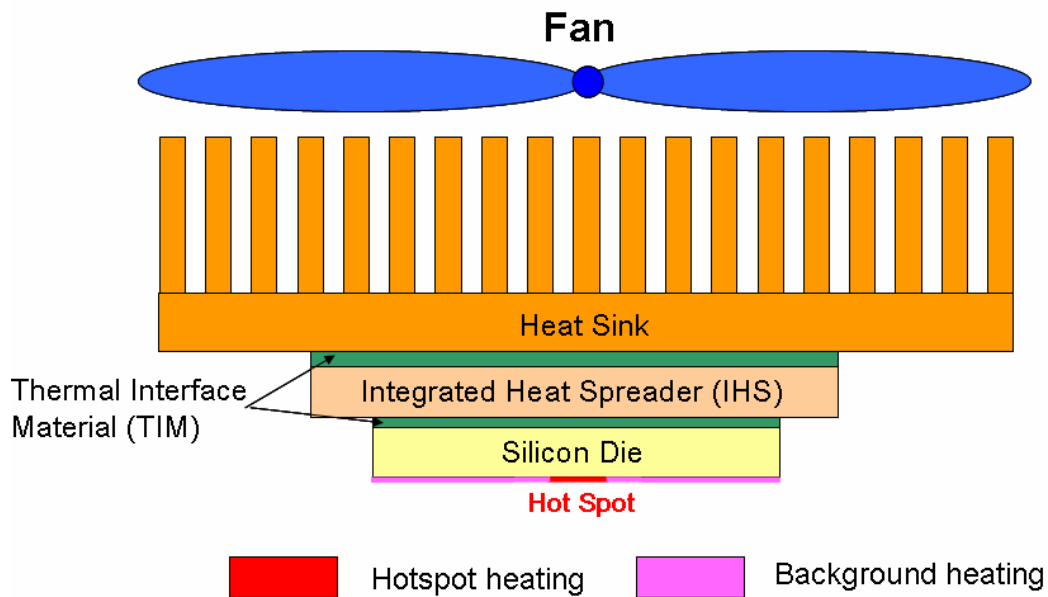
Figure 6.2: Typical chip package without mini-contact TEC.
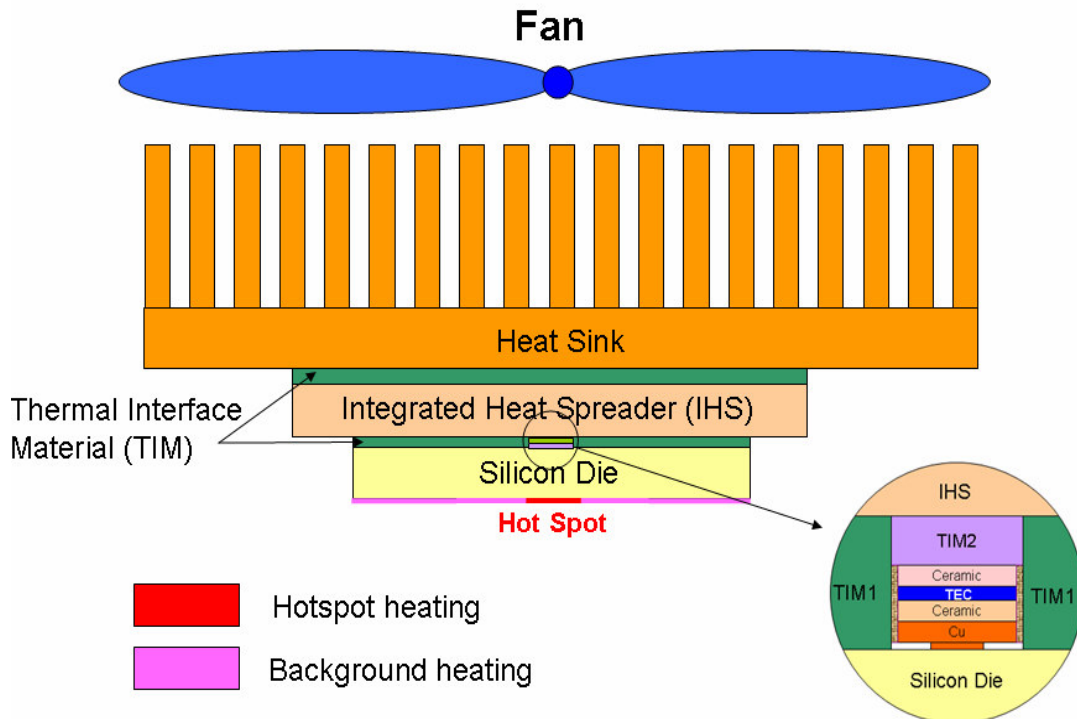
Figure 6.3: Typical chip package with embedded mini-contact TEC.

A three-dimensional (3D) package-level numerical model was developed using commercial finite element software ANSYS^{TM} to analyze and optimize on-chip hot spot cooling performance. Figure 6.2 and 6.3 show the typical chip packages for our numerical simulation without integrated TEC (Figure 6.2) and with a mini-contact enhanced TEC (Figure 6.3). The typical package consists of a silicon chip, thermal interface materials (TIM), a copper integrated heat spreader (IHS), and an aluminum heat sink with an integrated fan. As illustrated in Figure 6.3, a 2.4 mm × 2.4 mm miniaturized TEC with the thermoelectric element's packing density of 50%, which has the similar dimensions as reported by Intel [126] is integrated with a mini-contact pad, attached on the top of the silicon chip and then embedded inside the thermal interface materials, TIM1 and TIM2. The TEC consists of two 50 µm thick ceramic substrates and a 20 µm thick Bi_2Te_3 thermoelectric element, making it 120 µm in overall thickness. The mini-contact pad consists of a 100 µm thick copper mini-contact base, which is used to facilitate heat spreading, and a 50 µm thick mini-contact tip, which is employed to concentrate the effective cooling power of the TEC. For purposes of this study, the thickness of the TIM1 was fixed at 300 µm and the thickness of the mini-contact base and tip were held constant. A quarter-symmetry three dimensional (3-D) thermal model with a total element number of 150,000 is created and the elements are densely located around the TEC and the hot spot where the largest temperature gradients are expected to occur. The details of the solid-state circuitry in the chip, including individual transistors, gates, capacitors, etc., in the active regions of the die are ignored in this model but the heat generated from these components is modeled as a heat flux of 1250 W/cm^2 on a 400 µm × 400 µm hot spot

and a background heat flux of 70 W/cm^2 on the bottom surface of an 11 mm × 13 mm silicon chip. The major heat transfer path in this model is assumed to be from the active region at the bottom of the silicon die to the top side, then through a heat spreader to the heat sink by conduction and from the heat sink to the ambient by forced convection. Adiabatic boundary conditions are applied to all the side surfaces of the die, the spreader and the heat sink. To simplify the modeling geometry, the details of heat sink fins and fan are not included in this model and, instead, an equivalent convective heat transfer coefficient of 730 W/m^{2}K, representative of a high-performance, air-cooled heat sink, is applied as a boundary condition on the top surface of heat sink base to achieve an overall heat sink-to-ambient thermal resistance of 0.55 K/W, with an assumed ambient temperature of 25°C. Also, homogeneous material properties and uniform thicknesses are assumed for the silicon chip, thermal interface materials, heat spreader and heat sink base. The geometry and material properties used in this numerical model are listed in Table 6.1. Figure 6.4 shows the mesh structure and temperature contour of chip package with embedded TEC. Please note in order to save simulation run, quarter-symmetry model is used.

(a)

(b)
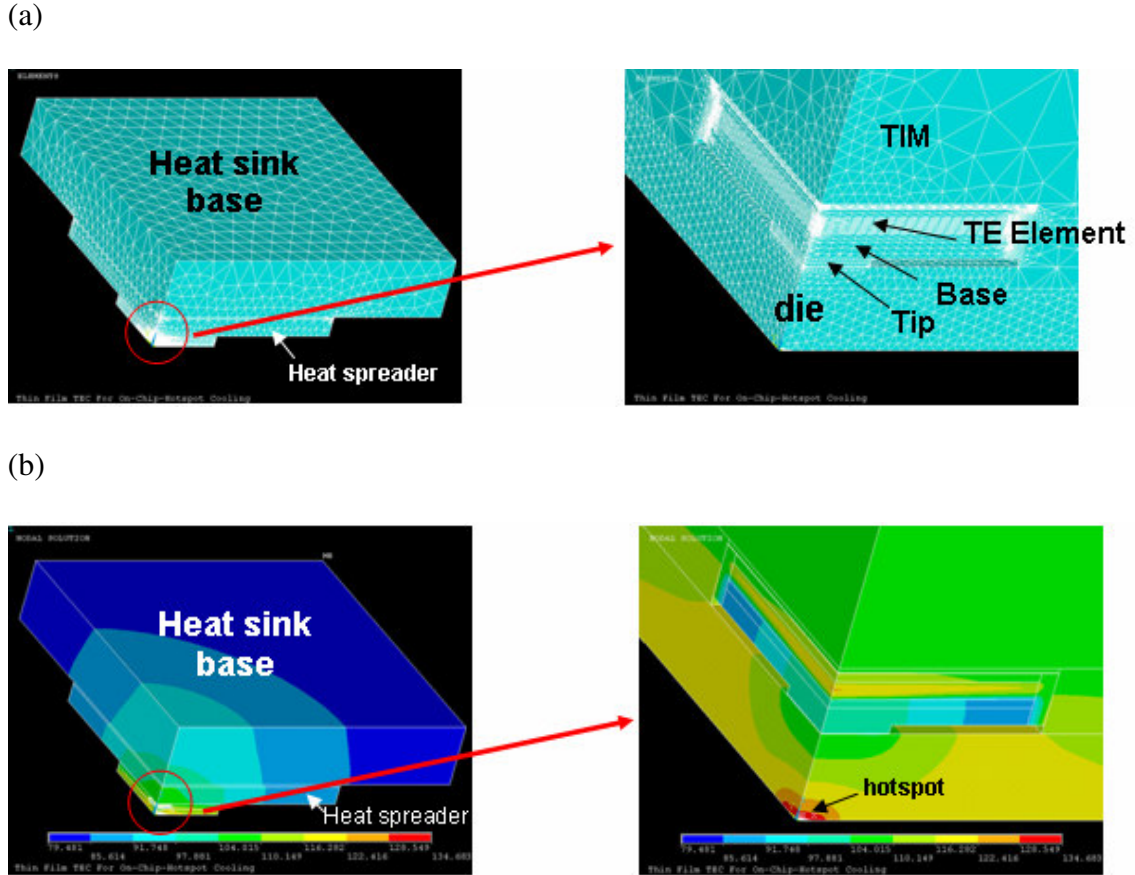
Figure 6.4: Meshing (a) and temperature contour (b) of chip package with embedded mini-contact enhanced TEC.

When a mini-contact enhanced TEC is integrated into the chip package, it introduces several thermal contact interfaces. In consideration of possible assembly procedures for such an enhanced TEC, it is assumed that the two most important thermal contact interfaces occur at the interface between the top ceramic substrate/TIM2 interface and at the interface between the mini-contact tip and the silicon die, as indicated in Figure 6.2. We use R_{c1} and R_{c2} to stand for these two thermal contact resistances with each varying from 1×10^{-7} to 1×10^{-4} m^2K/W which is the typical range reported for electronic package application [127]. The specific electrical contact resistance between the Bi_2Te_3 element and the metallization layer is

assumed to be 1×10^{-7} Ωcm^2 as reported previously. Thermoelectric cooling power is give by $q_{TE\ cooling} = ST_cI$, where S is the Seebeck coefficient of the thermoelectric material, T_c the cold-side junction temperature at the TEC, I the applied current on the TEC. Similarly, thermoelectric heating power is give by $q_{TE\ heating} = ST_hI$, where T_h is the hot-side junction temperature at the TEC. Thermoelectric cooling and heating powers are represented as heat flux boundary conditions in our numerical model and directly added to the cold and hot sides of the TEC, respectively, while Joule heating is modeled as uniform volumetric heat generation inside the bismuth telluride elements. Joule heating from electrical contact resistance is modeled as the surface boundary condition at the two TEC junctions. As the hot spot cooling performance is strongly dependent on input power supplied to the TEC, in the course of this simulation, various electric currents are applied to the TEC until the lowest hot spot temperature is achieved.

Table 6.1: Geometry and material properties used for numerical package-level model.

	Geometry (l × w × h)	Materials	k (W/mK)
Heat sink base	50mm×50mm×5mm	Al	180
HS	31mm×31mm×1.5m	Cu	360
TIM3	31mm×31mm×175μm	/	30
TIM2	31mm×31mm×30μm	/	30
Top ceramic substrate	2.4mm×2.4mm×50μm	AlN	180
TEC	2.4mm×2.4mm×20μm	Bi_2Te_3	1.4 - 1.5*
Bottom ceramic substrate	2.4mm×2.4mm×50μm	AlN	180
Mini-contact base	2.4mm×2.4mm×100μm	Cu	360
mini-contact tip	h = 50μm **	Cu	360
Die	11mm×13mm×500μm	Si	90 - 150*
TIM1	11mm×13mm×300μm	/	30
hotspot	400μm×400μm	/	/

* Temperature dependent thermal conductivity is used.
** The cross-sectional area of the mini-contact tip will change from 600 µm × 600 µm to 2400 µm × 2400 µm.

6.2 Results and Discussions

6.2.1 Effect of Input Power on TEC

Thermoelectric cooling performance is dependent on applied current or input power to the TEC in a non-linear manner, as Peltier cooling has a favorable linear dependence on electric current while the parasitic Joule heating effect has a quadratic dependence on electric current. The competition of these two opposite contributions will lead to an optimum current or input power at which the maximum hot spot cooling can be achieved. Figure 6.5 demonstrates the variation of hot spot temperature as a function of the input power to the TEC with the mini-contact tip size of 1250 µm × 1250 µm and the assumed thermal contact resistance of 1×10^{-7} Km^2/W at both the mini-contact tip/silicon chip interface and the ceramic/TIM2 interfaces. If there is no TEC, the peak hot spot temperature is found to reach 137.0^{o}C. However, if we activate the TEC, the hot spot temperature decrease steeply as the power increases, reaching a minimum of 120°C at approximately 10 W, providing a temperature reduction of 17.0^{o}C at the hot spot compared to the temperatures encountered without the TEC, and then rises more gently as the power increases further. It is worth noting that if the TEC is present but not activated, the hot spot temperature will increase by 7°C due to the additional thermal resistance to heat flow from the hot spot created by the presence of the TEC. Figure 6.6 shows the temperature contour and the heat flux inside the silicon die without TEC and with TEC. From the temperature contour, we can see if the TEC is integrated on the silicon die and turned on, there is a cold zone on the top of the die. Also, we can find the temperature distribution inside the die is changed compared with the case without

181

TEC. From the heat flux profile, we can see that if there is no TEC integrated on the die, the heat from the hotspot will be full spreading in all directions. However, if we turn on the TEC, most of the heat from the hotspot will flow to the TEC. Also, TEC not only cools down the hotspot but also cools the background of the die. We can see the heat from the background also flows to the TEC.
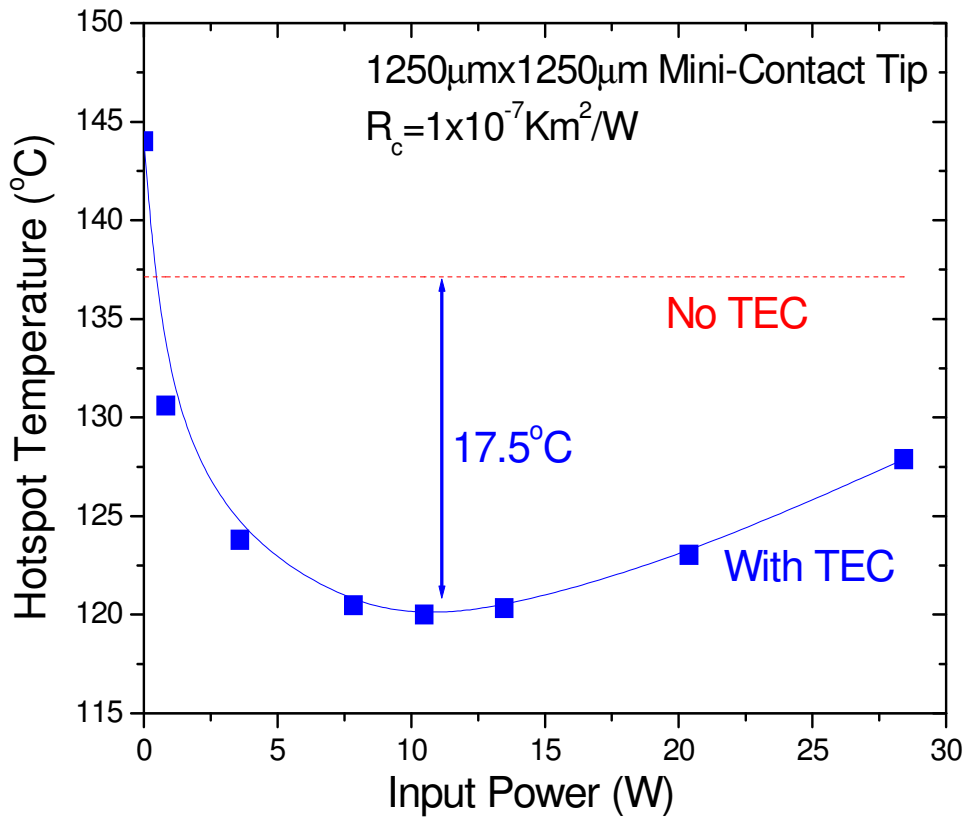
Figure 6.5: Hot Spot Temperature Variation with TEC Input Power (Bi$_2$Te$_3$ TEC, Thermoelectric Element Thickness = 20 µm, Copper Mini-Contact Tip Size = 1250 µm × 1250 µm).

<center>(a) Without TEC   (b) With TEC</center>
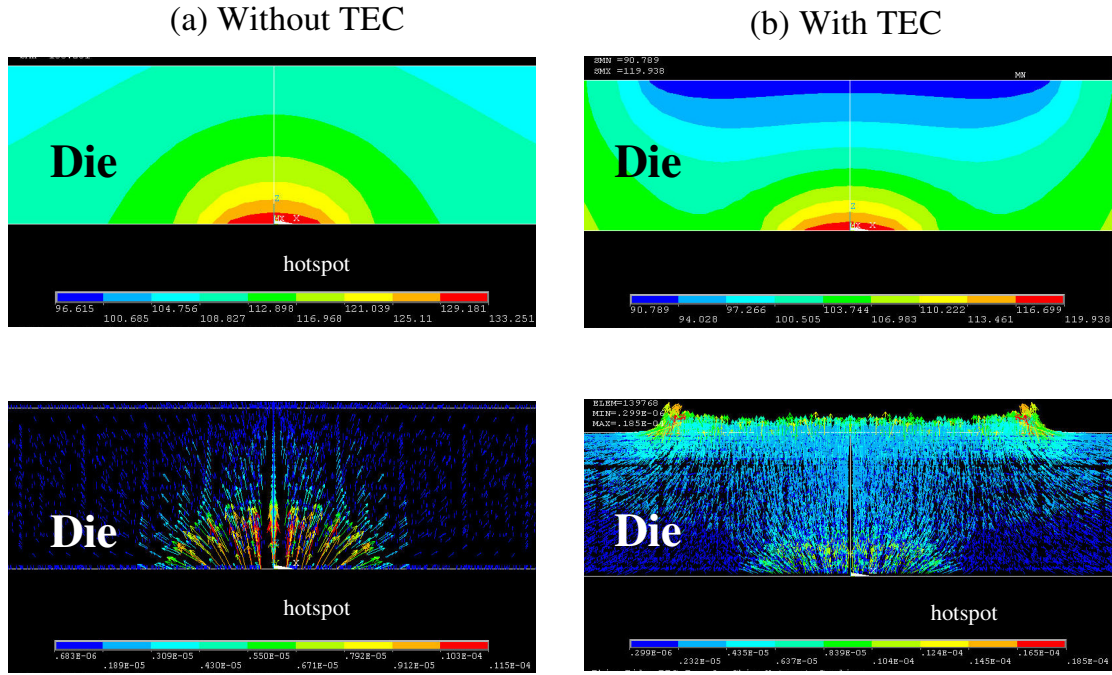
Figure 6.6: Temperature contour and heat flux inside the die. Bi$_2$Te$_3$ thermoelectric element thickness is 20 μm, copper mini-contact tip size is 1250 μm × 1250 μm, input power is 10W, thermal contact resistance is 1×10^{-7} m^{2}K/W, die thickness is 500 μm.

.

6.2.2 Effect of Mini-Contact Size

The mini-contact pad sandwiched between the TEC and the silicon chip is used to concentrate the thermoelectric cooling power and its beneficial effect on hot spot cooling is limited by the heat spreading effect inside the mini-contact pad as well as inside the silicon chip. Consequently, care should be taken to optimize the geometric configuration to achieve the maximum hot spot cooling performance. Figure 6.7 shows the temperature profile on the bottom of the silicon chip with different mini-contact sizes. First we can find if there is no hot spot and no TEC on the chip, the peak chip temperature is about 109°C. However, if there is a 400 μm × 400 μm hot spot with a heat flux of 1250 W/cm^2, the peak chip temperature will

<center>183</center>

increase to 137°C. So, the hot spot leads to about 28°C peak temperature rise on the chip. If the TEC is activated with 10W input power and enhanced with a 600 µm × 600 µm copper mini-contact pad, the hot spot temperature is reduced to 128°C, leading to 9°C maximum hot spot cooling. If the mini-contact tip grows to 1250 µm × 1250 µm, the hot spot temperature will reduce further, down to 120°C, resulting in 17°C maximum cooling at the hot spot. However, if we expand the mini-contact size to as large as 2400 µm × 2400 µm, the hot spot cooling is limited to just 12°C. Obviously there exists an optimized mini-contact size for each configuration.
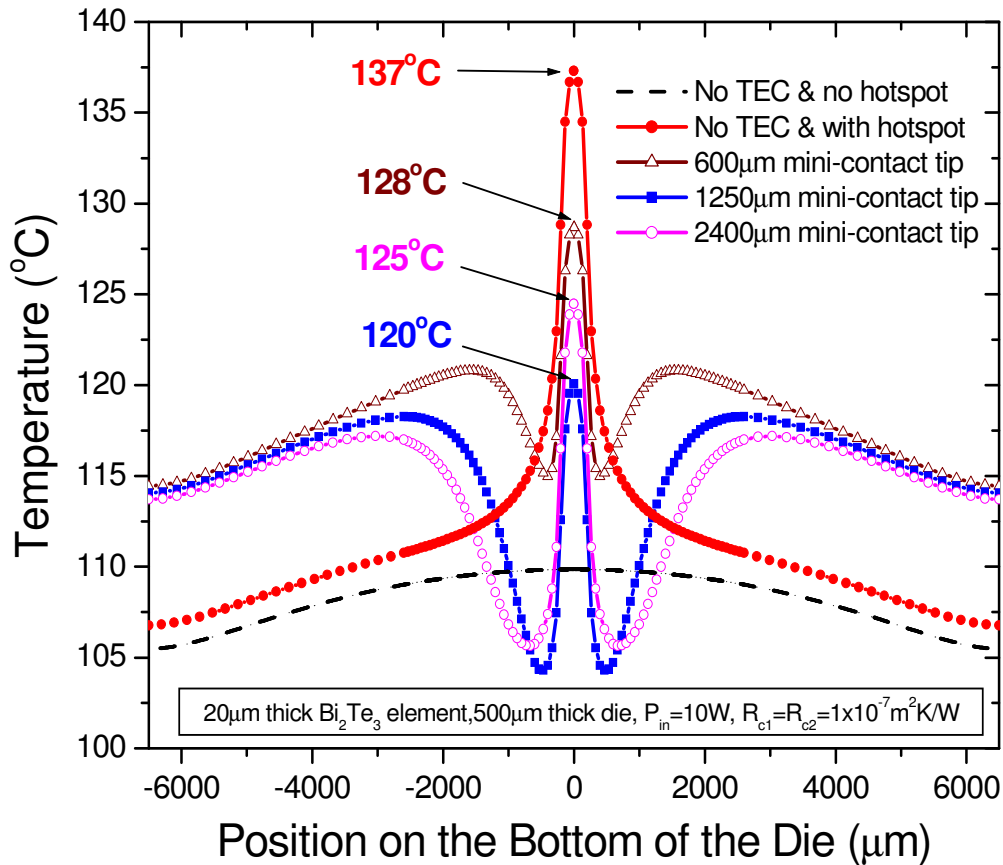
Figure 6.7: Effect of Mini-Contact Size on TEC-Induced Temperature Profile (Thermoelectric Element Thickness = 20 µm, Input Power =10 W).

6.2.3 Effect of Thermoelectric Element Thickness

Thermoelectric element thickness is a key parameter for improving the hot spot cooling performance as the maximum achievable cooling flux of the TEC is inversely proportional to the thermoelectric element thickness. Figure 6.8 illustrates the variation of hot spot cooling with the mini-contact size for three different thermoelectric element thicknesses under optimized input power on the TEC. As is expected, thinner thermoelectric elements allow the TEC to achieve better hot spot temperature reductions, e.g. 6°C to 11.2^{o}C and to 17.0^{o}C as the thermoelectric element decreases from 100 µm to 50 µm and to 20 µm in thickness, using the optimum mini-contact tip size. Even though the mini-contact tip size is kept constant, thinner thermoelectric elements always yield better hot spot cooling than thick elements. However, it is interesting to find that the optimum mini-contact tip size increases with decreasing element thickness, from 800 µm × 800 µm for a 100 µm thick element to 1000 µm × 1000 µm for a 50 µm thick element, and to 1250 µm × 1250 µm for a 20 µm thick element. It should be emphasized that, as may be seen in Figure 6.8, the use of the mini-contact pad results in much better hot spot cooling improvement when combined with thicker rather than thinner thermoelectric elements. While for 20 µm thick TE elements, the addition of an optimally-sized mini-contact pad improves the cooling by 4.3^{o}C (from the hot spot cooling of 12.7^{o}C with no mini-contact to 17°C with a 1250 µm × 1250 µm mini-contact), for 100 µm thick TE elements, the addition of an optimally-sized mini-contact pad (with a 800 µm × 800 µm tip), reduces the hot spot temperature by an additional 6°C.
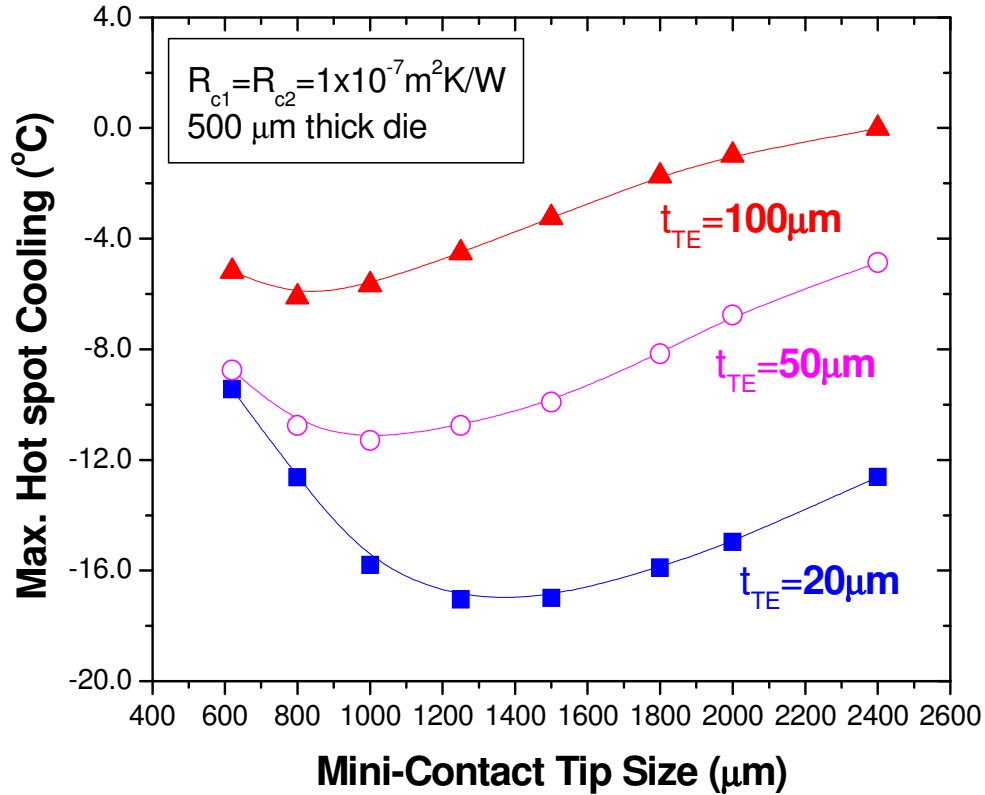
Figure 6.8: Effect of TE Element Thickness on the Optimized Mini-Contact Tip Size. The optimized input power of 10W, 7.5W and 6.1W is applied on 20μm, 50μm and 100μm thick TEC, respectively.

6.2.4 Effect of Thermal Contact Resistance

Low thermal resistance interfaces are critical to mini-contact enhanced hot spot cooling, since a high thermal resistance at the mini-contact/die interface - where the heat flux is highest - will significantly reduce the effectiveness of the cold mini-contact tip and a bad thermal interface between the TEC and the TIM2 will impede heat transfer into the heat spreader and then into the heat sink and the ambient. Figure 6.9 displays the interplay between the thermal contact resistance and the achievable hot spot cooling with the assumption of equal thermal contact resistance at the two interfaces (e.g. $R_{c1} = R_{c1} = R_c$) and reveals that with increasing thermal contact

186

resistance at the both interfaces, the net cooling achievable on the hot spot diminishes. It may be seen that if the thermal contact resistance is 1×10^{-5} Km^2/W or higher, the hot spot temperature will increase to 140°C and the embedded TEC will actually raise the hot spot temperature rather than make it cooler. The thermal contact resistance also has an impact on optimized mini-contact size. As shown in Figure 6.9, with increasing thermal contact resistance, the optimized mini-contact increases from 1250 µm × 1250 µm for the thermal contact resistance of 1×10^{-7} Km^2/W to 2000 µm × 2000 µm for the thermal contact resistance of 1×10^{-5} Km^2/W. It should be noted that over the full thermal contact resistance range investigated, use of a reduced cross-section mini-contact tip always provides a lower hot spot temperature than achieved without the mini-contact or with the mini-contact of the same size as the TEC base. This is especially true at the low contact resistances. However, the mini-contact is seen to provide diminishing returns as the contact resistances increase.

Thermal contact resistance at different interfaces might not necessarily be the same depending on surface roughness, surface chemistry and applied stress, etc [128]. Next we assume the thermal contact resistance at the two interfaces is not the same (e.g. $R_{c1} \neq R_{c1}$), with one being nearly "perfect" thermal contact ($R_c = 1\times10^{-7}$ Km^2/W) while the other being "bad" thermal contact ($R_c = 1\times10^{-5}$ Km^2/W). Figure 6.10 demonstrates the dependence of the maximum hot spot temperature as a function of mini-contact size for these two different cases. First it can be found that if the thermal contact between TEC/TIM is nearly "perfect" but that between mini-contact/chip is "bad", at the mini-contact size of 2000 µm × 2000 µm, a maximum cooling of 6.4^{o}C can be achieved at the hot spot. On the other hand, if the thermal contact between

TEC/TIM is "bad" while that between mini-contact/die is nearly "perfect", a maximum cooling of 4.7°C can be obtained at the hot spot with the mini-contact size of 1250 µm × 1250 µm. Therefore, if one interface is good but the other interface is bad, the mini-contact enhanced TEC can still achieve $5 \sim 6^{\circ}$C hot spot cooling under optimized mini-contact size and optimized input power on the TEC. Also, in comparison of these two extreme cases, it seems the thermal contact at TEC/TIM interface is more critical than that at mini-contact/chip interface.
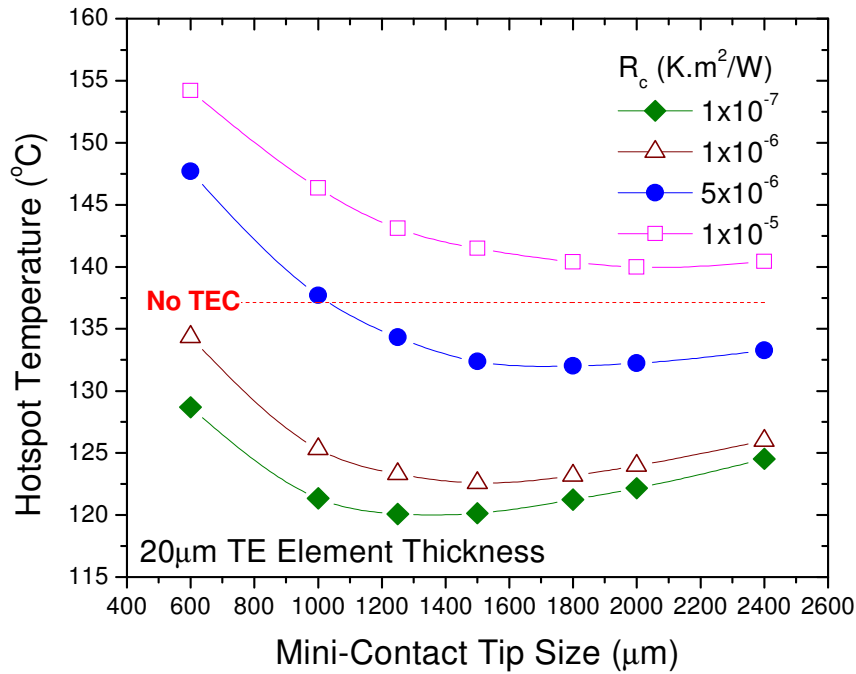
Figure 6.9: Influence of Thermal Contact Resistance on Mini-Contact Enhanced TEC Hot Spot Cooling (The contact resistance is assumed equal at TIM/TEC and mini-contact/Silicon chip interfaces).
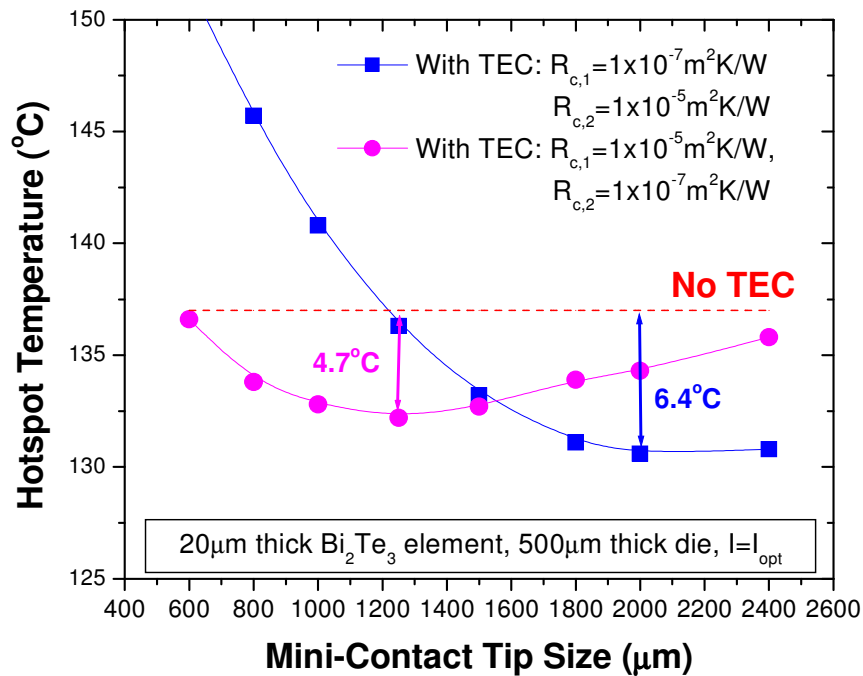
Figure 6.10: Influence of Thermal Contact Resistance on Mini-Contact Enhanced TEC Hot Spot Cooling (The contact resistance is assumed not equal at TIM/TEC and mini-contact/Silicon chip interfaces).
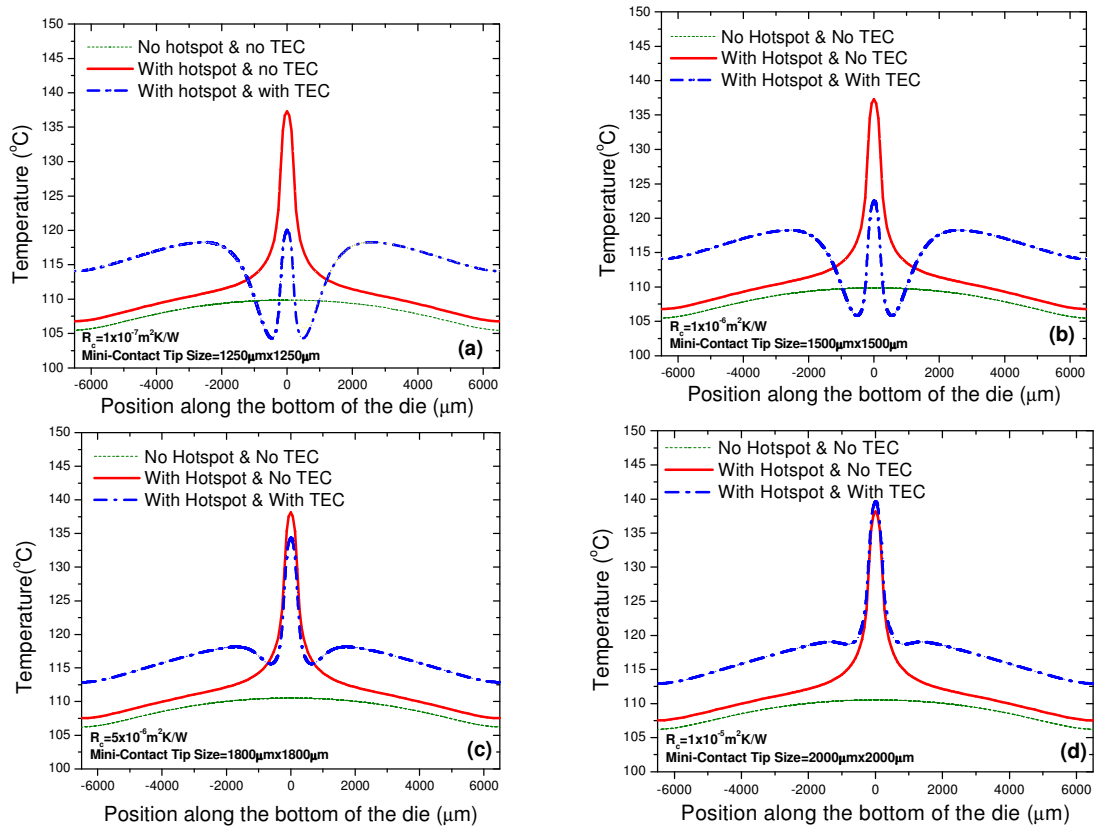
Figure 6.11: Temperature distribution on the silicon die for mini-contact enhanced 20μm thick TEC-cooled chip at the optimized input power of 10W: (a) $R_c = 1×10^{-7}$ Km²/W, (b) $R_c = 1×10^{-6}$ Km²/W, (c) $R_c = 5×10^{-6}$ Km²/W, (d) $R_c = 1×10^{-5}$ Km²/W.

In order to better understand the beneficial effect of mini-contact enhanced, miniaturized TEC for hotspot cooling, the temperature profile on the silicon chip for different thermal contact resistances, under the optimized input power and optimized mini-contact tip size, are shown in Figure 6.11(a) - (d). In each of the figures, the chip temperature profile that would occur without the hotspot, the temperature with the hotspot, and the temperature resulting from the activation of the TEC are displayed. In Figure 6.11(a) for the case of excellent bonding ($R_c = 1×10^{-7}$ Km²/W) at the two

interfaces, the peak temperature - created by the 1250W/cm^2 hotspot with 70W/cm^2 background heating - is seen to reach 137.5^{o}C, suggesting that the hotspot brings about a 27.5^{o}C temperature rise. With the mini-contact enhanced miniaturized TEC activated, the peak temperature drops to 120°C. The targeted cooling of the TEC is clearly reflected in the characteristic "W" temperature profile, with the 17.5^{o}C temperature reduction at the hotspot accompanied by a 7°C temperature increase in base temperature. This increase in the "side lobes" of the profile is due to the additional heat dissipation created by operating the TEC, but produces an effect similar to what would have been achieved by transferring heat from the hotspot to the adjacent area on the chip. Figure 6.11 (b)-(d) reflects the previously observed deterioration in cooling capability with increasing thermal contact resistance of 1×10$^-$^{6}, 5×10^{-6} and 1×10^{-5} K.m^2/W, respectively, at the two interfaces. Clearly, as the thermal contact resistance increases, the beneficial effect of the TEC for on-chip hotspot cooling is diminishing.

6.3 Conclusions

Mini-contact enhanced TEC has great potential application for on-chip hot spot cooling, providing significant cooling improvement relative to directly attached TEC cooling, by concentrating the thermoelectric cooling power on small regions of the silicon chip. Numerical simulation shows the peak temperature on the silicon chip with a 400 µm × 400 µm hot spot of a 1250 W/cm^2 heat flux can be reduced by as much as 17°C, 10°C and 6°C with an optimized mini-contact enhanced, 20 µm, 50 µm and 100 µm thick TEC, respectively, under optimized TEC input power and with nearly perfect thermal contact at the related interfaces. However, the thermal contact

resistance is critical to this novel cooling solution and must be maintained well below 1×10^{-5} Km^2/W in order to obtain significant hot spot cooling capability. The optimized mini-contact size is determined by thermoelectric element thickness and thermal contact resistance with larger thermal contact resistance and thinner thermoelectric element leading to larger optimized mini-contact size.

Chapter 7
Proof of Concept for Spot Cooling Improvement Using Mini-Contact Enhanced TEC

In the previous chapter, we use 3D finite element modeling approach to predict that the copper mini-contact, which connects the TEC and the silicon die, can enhance hot spot cooling performance significantly by concentrating thermoelectric cooling power to a spot on the back surface of the die. In this chapter we design some experiments to demonstrate spot cooling improvement using mini-contact pad. Due to the restraint of state-of-the-art microfabrication equipment to fabricate microscale thin film hot spot on the silicon die and the related thin film temperature sensing systems, we do not fabricate the hot spot on the silicon die to measure the cooling performance at the hot spot. Instead, we measure the temperature of the spot which is just below the mini-contact tip but separated from it by the silicon die, to characterize the spot cooling performance. Therefore, this experimental design is for "Proof of Concept", to demonstrate that mini-contact pad can improve spot cooling performance. From the principle of superposition, it is expected that cooling down at a spot on the silicon die will have the similar cooling effect at the hot spot if it is built at that spot. In order to demonstrate hot spot cooling using this concept, some sophisticated thin film deposition process and microelectronics packaging/assembling process, such as soldering and alignment, etc. are required. It will be the future work for this project.
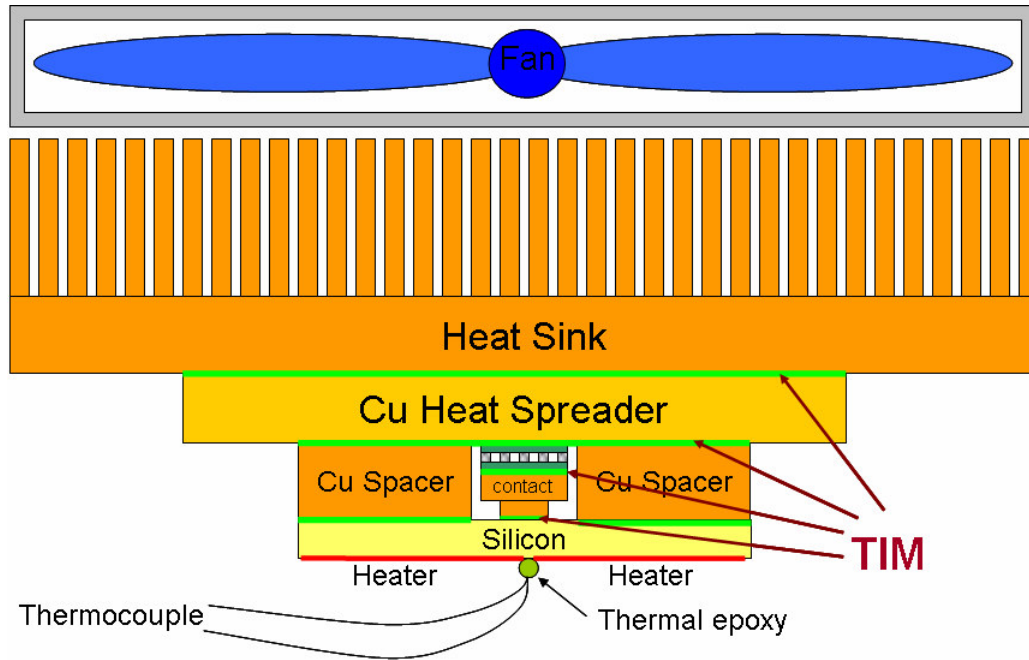
Figure 7.1: Schematic of test vehicle for mini-contact enhanced TEC for spot cooling.

7.1 Design of Experiment for Spot Cooling Testing

Figure 7.1 illustrates the schematic of the vest vehicle designed for our experiment to demonstrate spot cooling improvement with a copper mini-contact. At the center of the top surface of the silicon die is bonded with a copper mini-contact enhanced TEC by soldering. A copper spacer, which is used to accommodate the TEC and the copper mini-contact, is also bonded to the top surface of the die. Above the copper spacer and mini-contact TEC we attach a copper heat spreader. On the copper heat spreader there is a heat sink with an integrated fan. On the bottom of the silicon die are attached some heaters to simulate the power dissipation of the chip and, in addition, a thermocouple is bonded to the center of the die using thermal epoxy and well aligned with the copper mini-contact to characterize spot cooling performance. There are several thermal interfaces at heater/die, mini-contact/TEC,

194

die/spacer, spacer/spreader, TEC/spreader and spreader/heat sink. Thermal grease is applied to fill these interfaces to facilitate heat conduction.

7.1.1 Miniaturized Thermoelectric Cooler.

As introduced in Chapter 2, today's most powerful and commercially available thermoelectric cooler is Thermion miniaturized TEC. As shown in Figure 7.2, Thermion miniaturized TEC used in these experiments has a dimension of 3.6 mm × 3.6 mm × 1.6 mm with a total of 36 diced pieces of p- and n-type 200 µm thick bismuth telluride thermoelectric elements. The thermal conductivity of bismuth telluride thermoelectric material is 1.3 W/mK, the Seebeck coefficient 200 µV/K and the electrical resistivity 10 µΩm, with a figure merit value Z of 3×10^{-3} K^{-1} and ZT of 0.9 at room temperature. The two ceramic substrates are made of AlN each with a thickness of 635 µm and indium-tin solder was pre-tinned to the end faces to facilitate solder connections. 36 pieces of bismuth telluride thermoelectric elements are soldered onto the AlN substrates at around $183^{o}C$. The maximum cooling temperature is reported to be $68^{o}C$ and $91^{o}C$ according to the vendor when the heat sink temperature is maintained constant at $25^{o}C$ and $85^{o}C$, respectively. The maximum cooling power is around 5 W and 6 W when the heat sink temperature is maintained constant at $25^{o}C$ and $85^{o}C$, respectively.
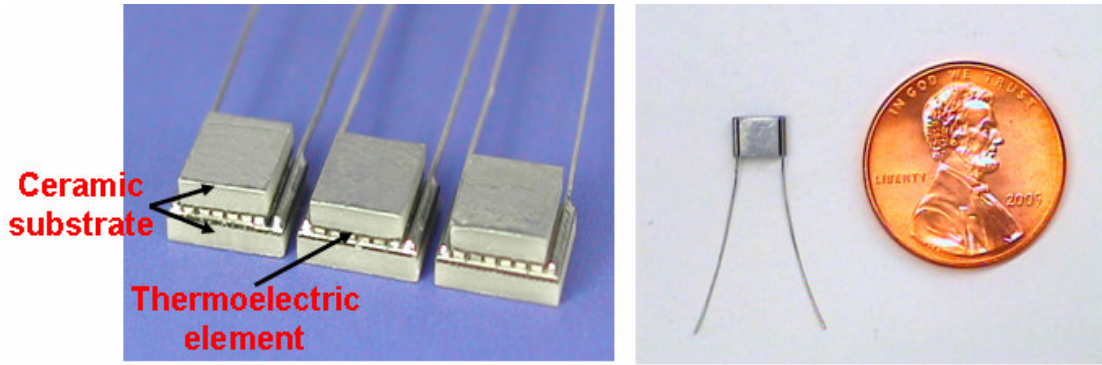
Figure 7.2: Optical micrograph of the Thermion TEC (Model: 1MC04-018-02-2200D).

7.1.2 Fabrication of Copper Mini-Contact

Due to its high thermal conductivity and low price, copper is probably the best material to for mini-contact pad. Based on our numerical simulation for design of the experiment, we find the optimized copper mini-contact tip size should be more than 500 µm × 500 µm. So we fabricate five different mini-contact sizes with the mini-contact tip' cross-section size ranging from 800 µm × 800 µm to 3.6 mm × 3.6 mm. Theoretically, the shorter the mini-contact tip, the smaller the thermal resistance and the better the spot cooling performance. However, considering the difficulty in soldering such a short mini-contact tip to the silicon die, we choose 0.5 mm as mini-contact tip thickness. As the Thermion TEC' base area is 3.6 mm × 3.6 mm, we choose the mini-contact base size to be 3.6 mm × 3.6 mm × 2.5 mm for all of the mini-contacts. In order to have very smooth surface to improve soldering process and also to reduce thermal contact resistance at silicon die/mini-contact interface, we ground the top and bottom surfaces of the copper mini-contact at CALCE Materials Characterization Laboratory. We start from coarse silicon carbide ground paper of #400, to #800, #1000, and end with #1200 (Applied High Tech Products, Inc.). After

grounding we polished the two surfaces using the alumina slurry (alumina particle size is ~ 1 µm) on variable speed grind-polisher (Bechler, Inc.) to make the surfaces very smooth, with a toughness of around 1 µm. Figure 7.3 shows the copper mini-contacts with five different tip sizes.
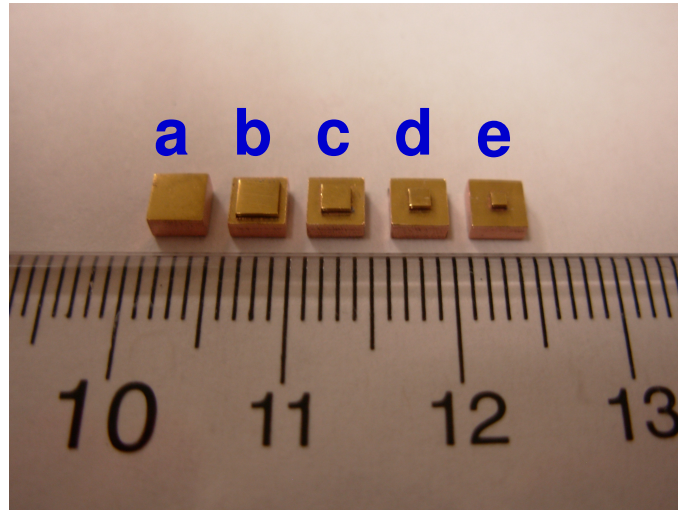
Figure 7.3: Copper mini-contacts with mini-contact tip size as (a) 3.6 mm × 3.6 mm, (b) 2.5 mm × 2.5 mm, (c) 1.8 mm × 1.8 mm, (d) 1.3 mm × 1.3 mm and (e) 0.8 mm × 0.8 mm. The mini-contact base size is 3.6 mm × 3.6 mm, the same size as Thermion TEC base.

7.1.3 Cr/Au Deposition on Silicon Wafers and Copper Mini-Contacts

To achieve good heat conduction through mini-contact tip/silicon interface, particularly for smaller mini-contact tips which could be as small as 800 µm × 800 µm in cross section, soldering of the copper mini-contact onto the silicon die to achieve as good thermal interface as possible is required. Thermion TEC modules are assembled using the solder with the melting temperature of 183°C during the manufacturing process and can withstand heating to 175°C. Therefore, for our experiment indium solder is used due to its relatively low melting temperature of

156.6 °C and good ductility. Alternatively, indium-tin solder (In52-Sn48), the indium–tin binary system with a eutectic temperature of 118 °C at 52% indium and 48% tin, is also used for developing a bonding process. However, the thermal conductivity of indium-tin solder is lower (34 W/mK) compared with indium (86 W/mK).

We found that copper mini-contact could not solder onto the silicon directly using any kind of solders. Therefore, the silicon and copper mini-contact has to be coated with some metal thin film to facilitate soldering process. Similar to the soldering approach reported by Lee [129,130], we develop the similar procedure to deposit a thin layer of Au thin film onto the silicon wafer and copper mini-contact. As indium solder will react with gold thin film above $160^{o}C$ to form an intermatallic compound $AuIn_2$, a good bonding between gold and copper can be achieved by such indium-gold reaction. However, gold can not deposit and adhere on the silicon wafer and thus Cr thin film has to be deposited onto the silicon first as an adhesive layer. After that a layer of gold thin film is deposited onto the Cr thin film. This two-layer metallization structure for soldering is illustrated in Figure 7.4.

(a)   (b)
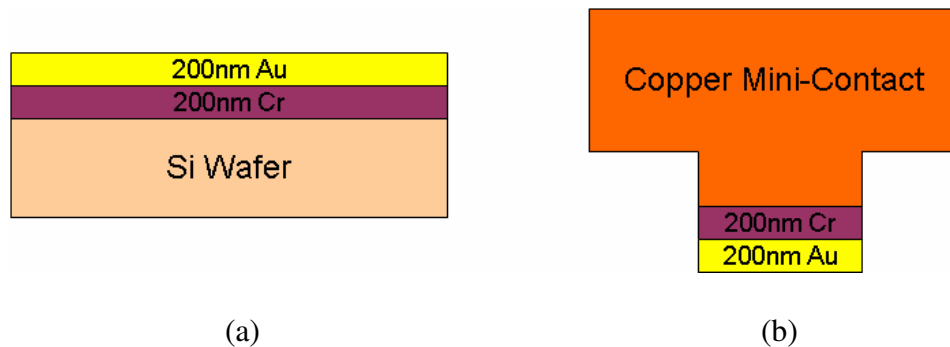
Figure 7.4: Schematic shows (a) silicon wafer and (b) copper mini-contact coated with 200 nm thick Cr and 200 nm thick Au thin films

We use the following processes at the clean room in the NanoCenter at the University of Maryland to clean the copper mini-contact:

1. Soak the copper mini-contact in acetone for 2-5 minutes;

2. Rinse in methanol for 2-5 minutes;

3. Rinse in by DI water for 2-5 minutes;

4. Do a wet etch of 10% nitric water solution to remove any possible copper oxide on the copper surface. It is found if there is copper oxide on the surface of mini-contact Cr can not adhere to the surface very well;

5. Do nitrogen blow to dry the copper mini-contact and store it in the sample box. Using the similar procedure, we cleaned the silicon wafer with acetone, methanol and DI water for 2 minutes.

We employ e-beam evaporation approach to fabricate Cr and Au coating on the silicon wafer. In this thin film process, a block of the material (source) to be deposited is heated to the point where it starts to boil and evaporate. Then it is allowed to condense on the targeted substrate, the material that we want to coat. This process takes place inside a vacuum chamber, enabling the molecules to evaporate freely in the chamber, where they then condense on all surfaces. For e-beam evaporation, an electron beam is used to heat the source material and cause evaporation. The electron beam evaporation process typically involves such components as electron beam evaporation gun, a system controller, power supply, crucibles for the evaporation material, materials for evaporation, material to be coated. The entire process takes place inside of a vacuum chamber and multi-layer coatings can be deposited in one duty cycle.

After we clean the copper mini-contact and silicon wafer and removed any oxides and greases, we mounted all the samples onto the coating stage inside the e-beam evaporator (Model: Airco FC-1800 e-beam evaporator). We first deposited the 200 nm thick Cr thin film at 25~35 Armstrong per second and then 200 nm thick Au thin film at 15~25 Armstrong per second under vacuum condition with a pressure of approximately 8×10^{-7} torr. In order to avoid oxidation of Cr coating, Au coating should be done immediately after Cr deposition without opening the chamber. Therefore, Care should be taken that deposition of Cr and Au should be two continuous processes in the vacuum. If Cr thin film is oxidized after the first deposition, for example due to opening of the chamber, we find Au thin film can not adhere on it. After Cr/Au thin films are successfully deposited on the silicon wafer, we cut the silicon wafer into 27 mm × 27 mm square pieces using a diamond cutter and use them as "silicon die" for our experiment.

7.1.4 Fabrication of Copper Spacer and Copper Heat Spreader

Five copper spacers with a dimension of 27 mm × 27 mm × 3.6 mm and five copper heat spreaders with a dimension of 40 mm × 40 mm × 10 mm are fabricated in the Mechanical Engineering Machine Shop. In each copper spacer, 5 mm × 5 mm square hole is drilled on the center to accommodate the TEC and the copper mini-contact. Several grooves are also made to accommodate the leads of the TEC module. Similar to copper mini-contact, in order to have very smooth surface to reduce thermal contact resistance, we ground the top and bottom surfaces of the copper spacers and heat spreaders at CALCE Materials Characterization Laboratory in the

University of Maryland. We started from a coarse silicon carbide ground paper of #400, to #800, #1000, and finally ended with #1200 (Applied High Tech Products, Inc.). After that we polished the two surfaces of each copper spacers and spreaders using variable speed grinder-polisher (Buehler, Inc) with the alumina slurry (alumina particle size is ~ 1 µm) to make the surfaces smooth with toughness of around 1 µm. Figure 7.5 shows the typical copper spacer after grounding and polishing, indicating that the surface is very smooth and shining.
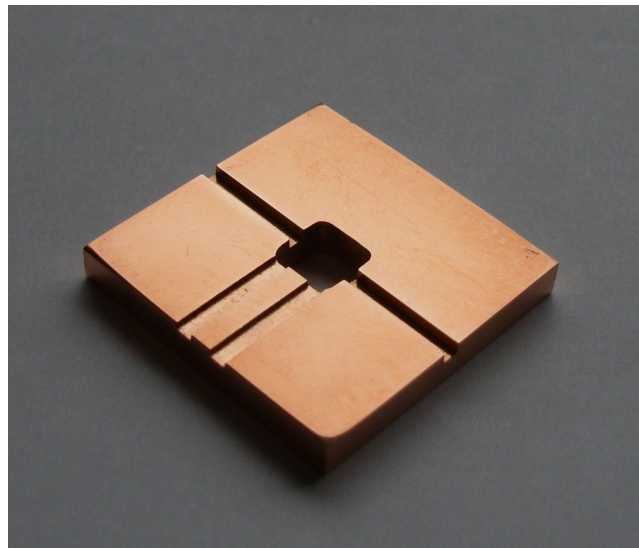
Figure 7.5: Photograph of copper spacer used to accommodate mini-contact TEC and leads.
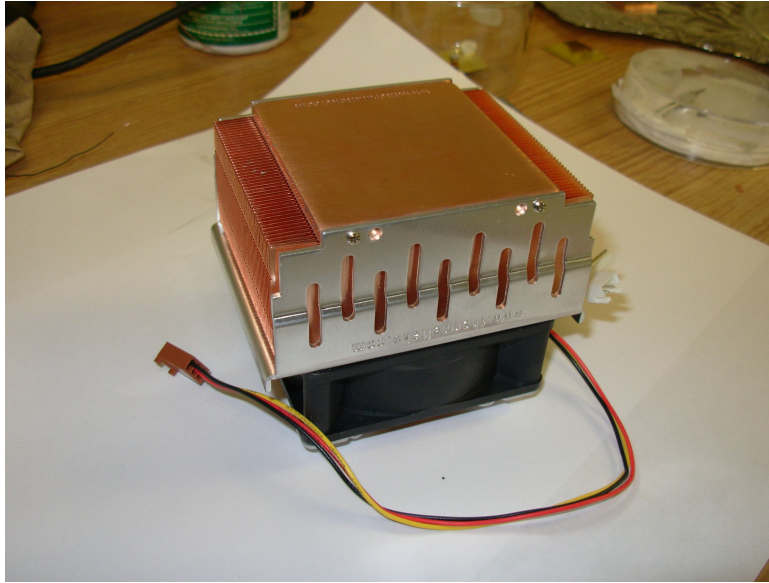
Figure 7.6: Photograph of copper heat sink with an integrated fan.

7.1.5 Copper Heat Sink

The heat sink we use is Thermaltake P4 Spark 7 A1715 highest performance CPU cooler. It is a HSF unit that features an auto-adjusting dual ball bearing fan. The fan operates at 1300 ~ 6000 RPM depending on the temperature and settings, while producing 17 – 43 dbA and generating 10.42 - 49.17 CFM. The dimension of the fan is 70 × 25 × 70 mm while the heat sink is made of copper with a dimension of 63 × 35.5 × 82 mm. The lowest thermal resistance achievable is reported to be around 0.26 K/W. Figure 7.6 shows the photo of the copper heat sink with an integrated fan.

7.1.6 Assembly of Testing Package

When all components necessary for thermoelectric spot cooling testing are available, we assemble it following the procedure shown in Figure 7.7 to 7.14. First we clean Cr/Au coated silicon die and copper mini-contact using acetone and

methanol to remove any grease. Then we put the silicon die onto the hot stage (Model: Corning Scholar 170) and turn on the power to increase to temperature of the silicon die to 160°C. Wait until the temperature becomes stable. Cut 1 mm × 1 mm × 0.25 mm indium solder piece from a 50 mm × 50 mm indium foil (Alfa Aesar, 0.25 mm thick, 99.99% pure) onto the silicon die and it is melt immediately. Then we attach the copper mini-contact onto the melt indium solder. Apply some pressure by hand on the top of the mini-contact to achieve better contact between the silicon die and the copper mini-contact as illustrated in Figure 7.7. Hold for 2 minutes with pressure and then turn off the power of the hot plate. After the temperature drops to below 100°C, copper mini-contact can be bonded to silicon die well as shown in Figure 7.8 and we can release the pressure. After the hot stage cooled down to room temperature, we put small amount of thermal grease onto the top surface of the copper mini-contact and spread it uniformly by hand. Then we attached the TEC on the top surface of the mini-contact and apply some pressure by hand on the top of the TEC as indicated in Figure 7.9. We find the TEC can attach onto the mini-contact well as illustrated in Figure 7.10

Next we put the thermal grease on the top surface of the silicon die and assemble the copper spacer onto the silicon die with some pressure as shown in Figure 7.11. Care should be taken that the pressure can not be too high. Otherwise, the silicon die will break. Then we put some thermal grease onto the top surface of the copper spacer and the TEC, and assemble the copper heat spreader onto them using pressure, as illustrated in Figure 7.12. Next we put thermal grease onto the top surface of the heat spreader and attach the heat sink onto it with pressure as shown in

Figure 7.13. Finally, we flip up the whole package and bond a thermocouple onto the center of the silicon die using thermal epoxy (Arctic Thermal Adhesive). In addition, four Minco heaters are bonded on the silicon die, as indicated in Figure 7.14, using thermal grease. In our experiment the thermal grease we use is Silver Arctic Thermal Grease. To our knowledge this kind of thermal grease has the best thermal performance available on the market and the reported thermal resistance of thermal grease per unit area is about 1.8×10^{-6} m^{2}K/W and the thermal conductivity is more than 7.5 W/mK [131]. However, depending on application condition, this data could be dramatically different from case to case [132].
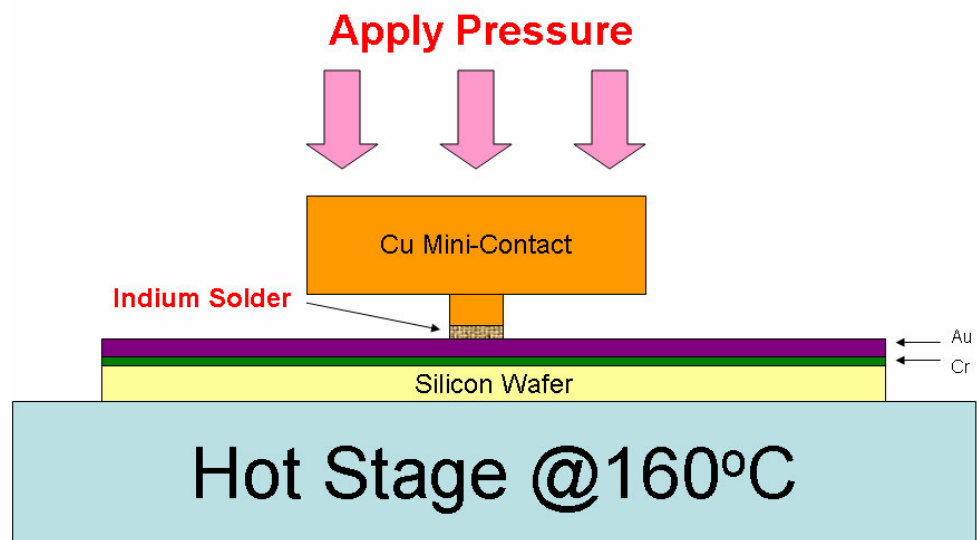
Figure 7.7: Schematic of how to solder mini-contact onto the silicon.
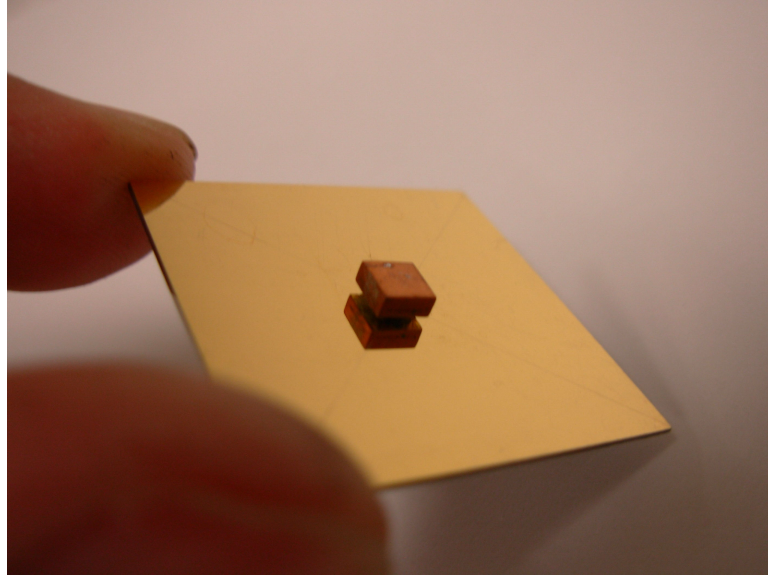
Figure 7.8: Copper mini-contact is soldered onto Au/Cr coated silicon die
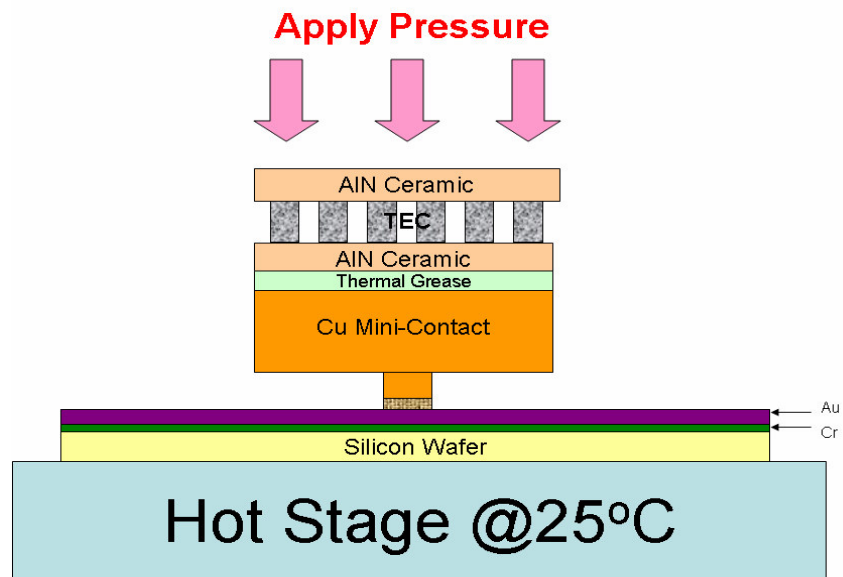
Figure 7.9: Schematic of how to attach TEC onto mini-contact.
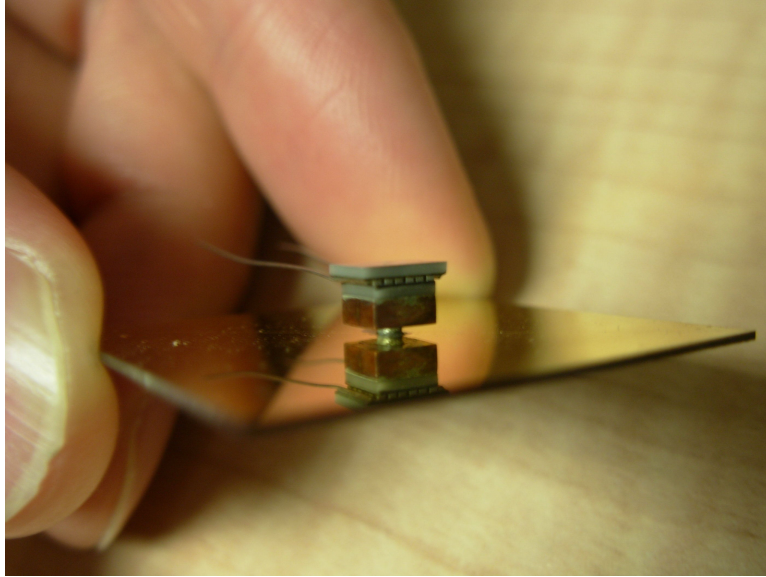
Figure 7.10: TEC is bonded onto copper mini-contact.
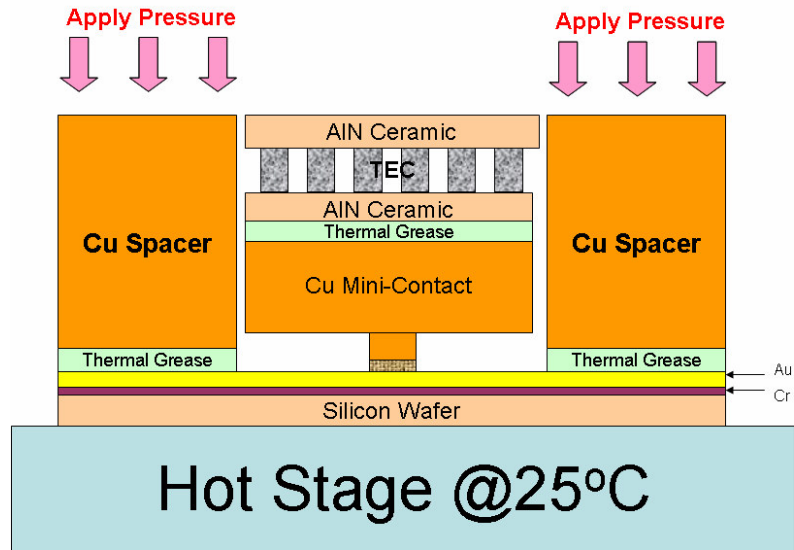
Figure 7.11: Schematic of how to attach TEC onto mini-contact.
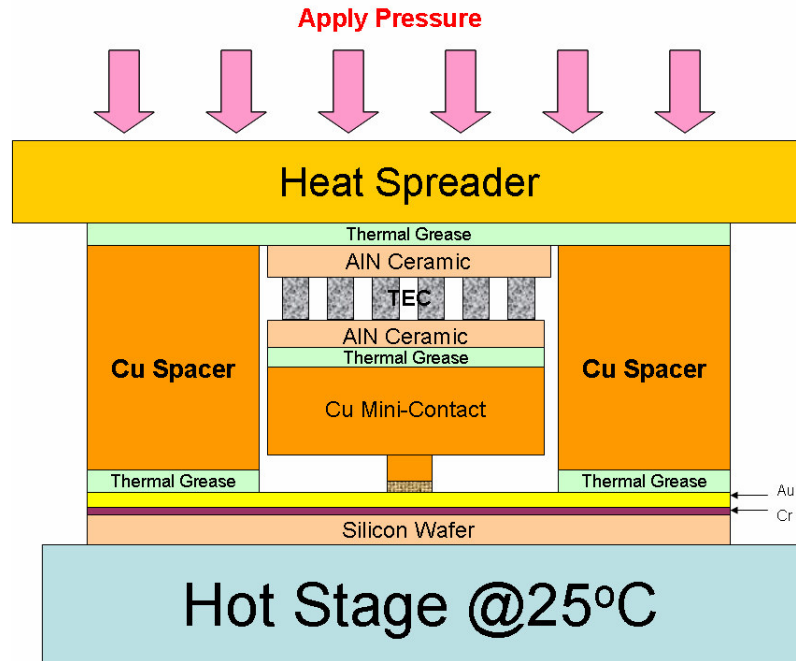
Figure 7.12: Schematic of how to attach heat spreader onto copper spacer and TEC.
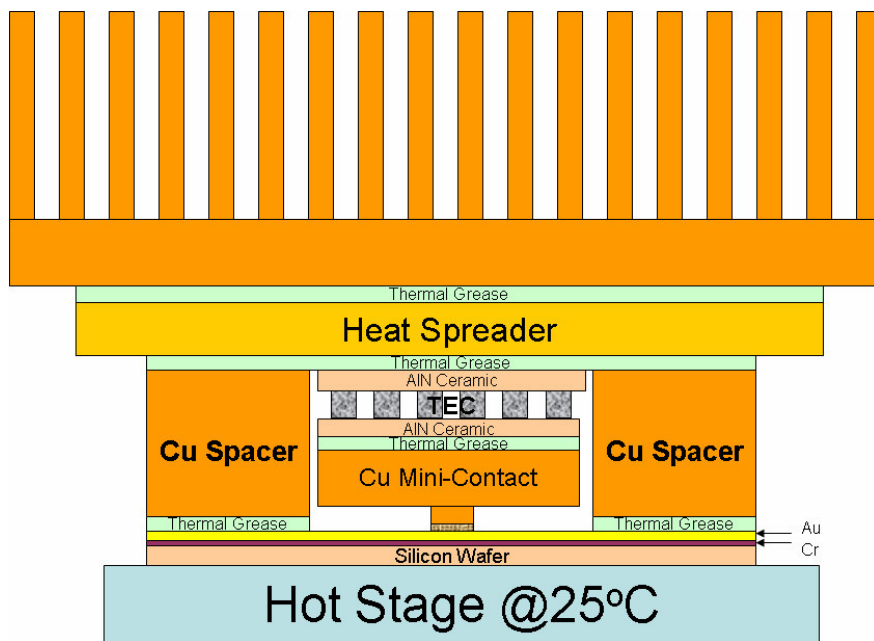
Figure 7.13: Schematic of how to attach heat sink to heat spreader.
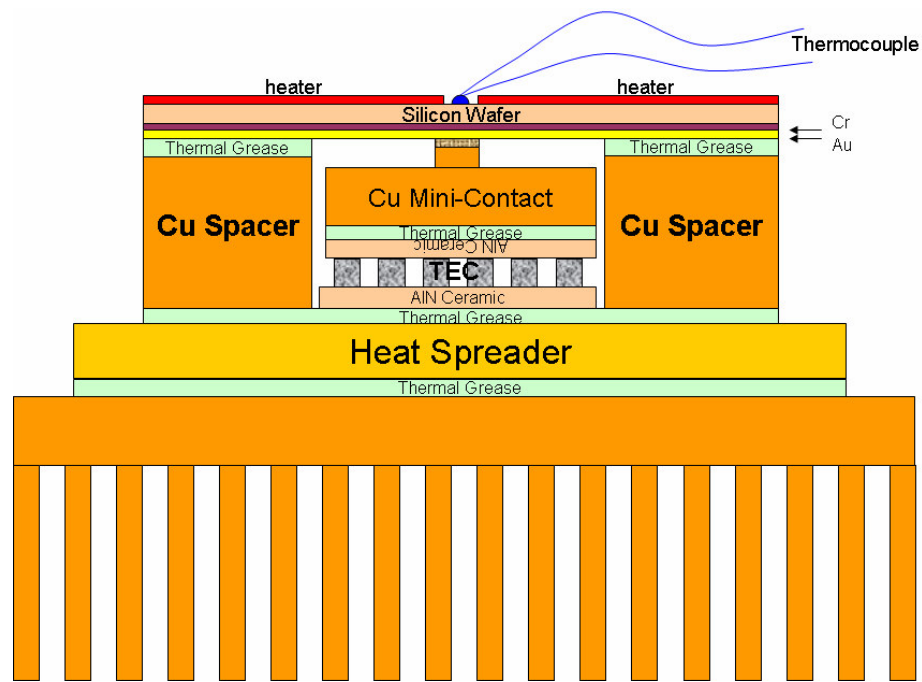
Figure 7.14: Schematic of how to attach thermocouple and heaters to the silicon die.

7.1.7 Experimental Setup and Thermal Test

A package-level experiment has been designed to demonstrate the spot cooling improvement using a copper mini-contact pad. Due to fabrication constraints, in this "proof of concept" experiment there are no micro-scaled hot spot and related temperature-sensing systems on the chip and, instead, the heat flux on the bottom of the chip is assumed to be uniform by attaching four Minco thin-film heaters in our experiment. With this approach the feasibility of spot cooling improvement using mini-contact enhanced TEC can be established. Figure 7.15 shows the chip package where four Minco heaters are bonded to the silicon die using thermal grease, and please note that in order to make the heaters have good contact with the silicon die, a glass fixture is used to apply pressure onto the silicon die. The glass fibrous insulator

208

is inserted between the silicon die and the glass fixture as shown in Figure 7.15 to eliminate thermal leakage.
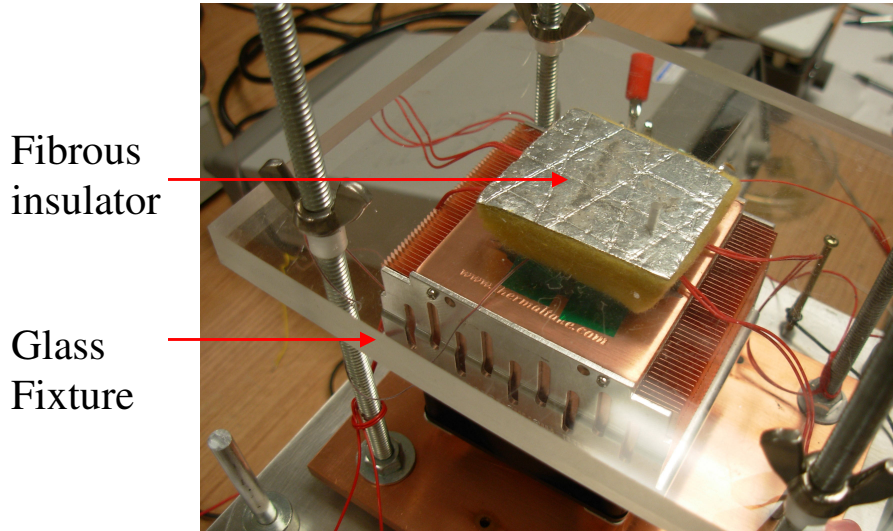
Figure 7.15: Testing Assembling with heaters attached on the silicon chip. As the heaters are covered with fibrous insulator, we can not see the heaters.

The experimental testing structure is shown in Figure 7.16 and 16A power of 0, 30 W, and 67 W was supplied to the bottom of the silicon wafer using four thin-film Minco heaters activated by HP E3611A System DC Power Supply to simulate different power dissipations on the silicon chip. The Omega E-type thermocouple with a diameter of 76 µm is bonded to the bottom of the silicon wafer using thermal epoxy to measure the spot cooling performance. Based on thermocouple calibration at 0°C and 100°C, respectively, the measurement uncertainty is estimated to $\pm 0.2^{o}$C. The tip of the thermocouple is well aligned with and separated by the silicon wafer with the copper mini-contact pad and the TEC. With specific power dissipation on the silicon wafer, the electric current was applied to the TEC by Agilent 6038A System

Power Supply and the temperature, current and voltage will be recorded automatically by Agilent 34970A Data Acquisition/Data Logger System. The thermal data reported herein relate to steady-state conditions, usually obtained some 10~30 minutes after the test vehicle was powered.
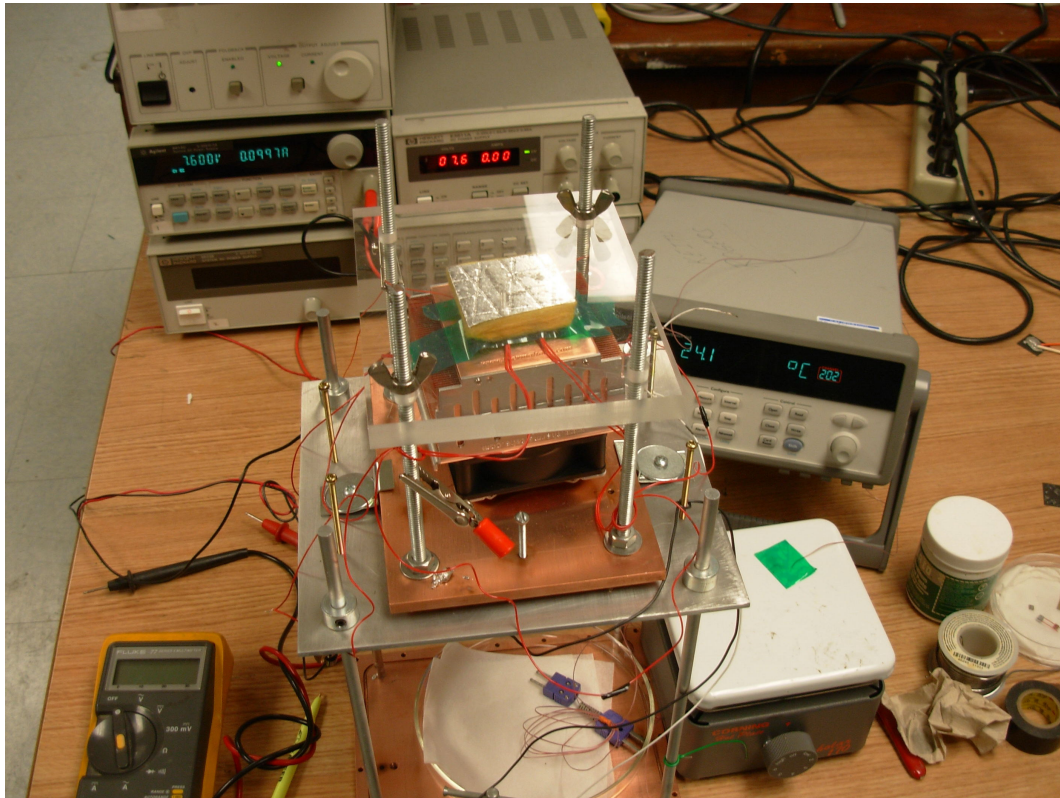
Figure 7.16: Experimental setup for mini-contact enhanced TEC for spot cooling.

7.2 Results and Discussions

In this work, thermal measurements were performed on the chip package test vehicle shown in Figure 7.16 to provide a validation of the numerical modeling approach discussed in the opening sections, to demonstrate and quantify the spot cooling improvement provided by the mini-contact pad, and to determine its relationship to the TEC input power and power dissipation on the silicon chip. Figure 7.17 shows the experimentally-determined dependence of targeted spot temperature and spot cooling for this 500 µm thick silicon die on the TEC input power, with the mini-contact tip size kept at 1.8 mm × 1.8 mm and the power dissipation on the silicon chip varying from 0 to 67 W. It is found that when the power dissipation on the silicon die increases from 0 to 30W and then to 67 W, the temperature at the targeted spot, e.g., the center of the silicon chip, increases from 25°C to 52.5^{o}C and then to 82.5^{o}C as illustrated in Figure 7.17 (a). The temperature reduction at the spot varies parabolicaly with the TEC input power, reflecting the competing mechanisms of rapidly improving Peltier cooling at lower input powers and progressively more damaging Joule heating, as well as reverse heat conduction, at the higher input powers [133]. Figure 7.17 (b) shows that in this test vehicle the silicon chip, thus, experiences its largest value of spot cooling at an optimum input power of 4 ~ 6 W.
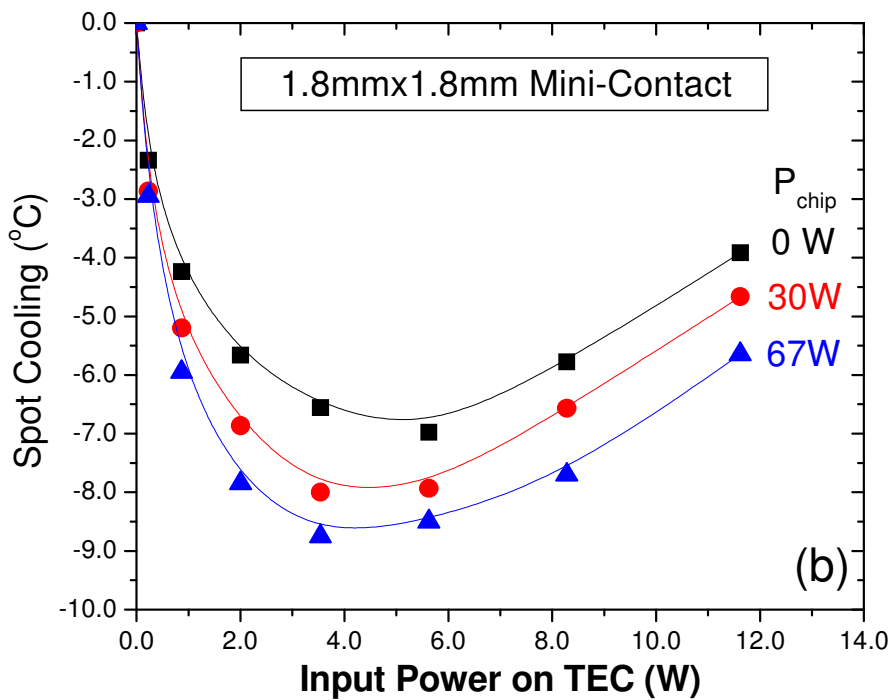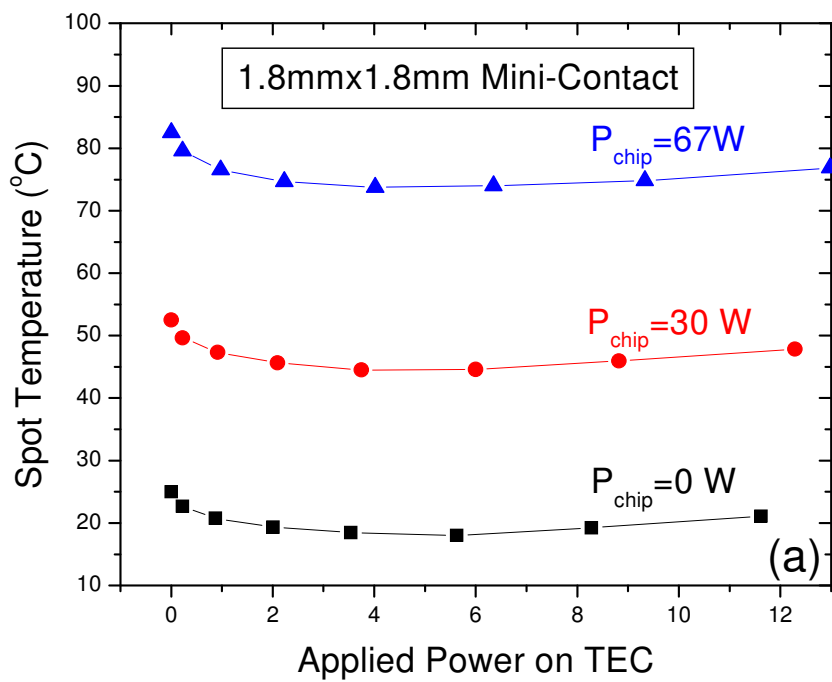
Figure 7.17: Variation of (a) measured spot temperature and (b) measured spot cooling with TEC input power.

It is interesting to note that the power dissipation on the silicon chip has a great effect on spot cooling and the larger the power dissipation, the greater the temperature reduction on the chip. For example, if there is no power dissipation on the silicon wafer, a maximum spot cooling of about 7.0°C can be achieved. However, if the power dissipation on the chip is increased to 67 W, the maximum spot cooling increases to 8.8°C with a 26% improvement. Interestingly, the experimental results suggest that increasing power dissipation on the silicon chip leads to lower values of the optimum TEC input power. As illustrated in Fig. 11, with an increase of the power dissipation on the silicon wafer from 0 to 67W, the optimum TEC input power decreases from 5.7 W to 3.6 W.
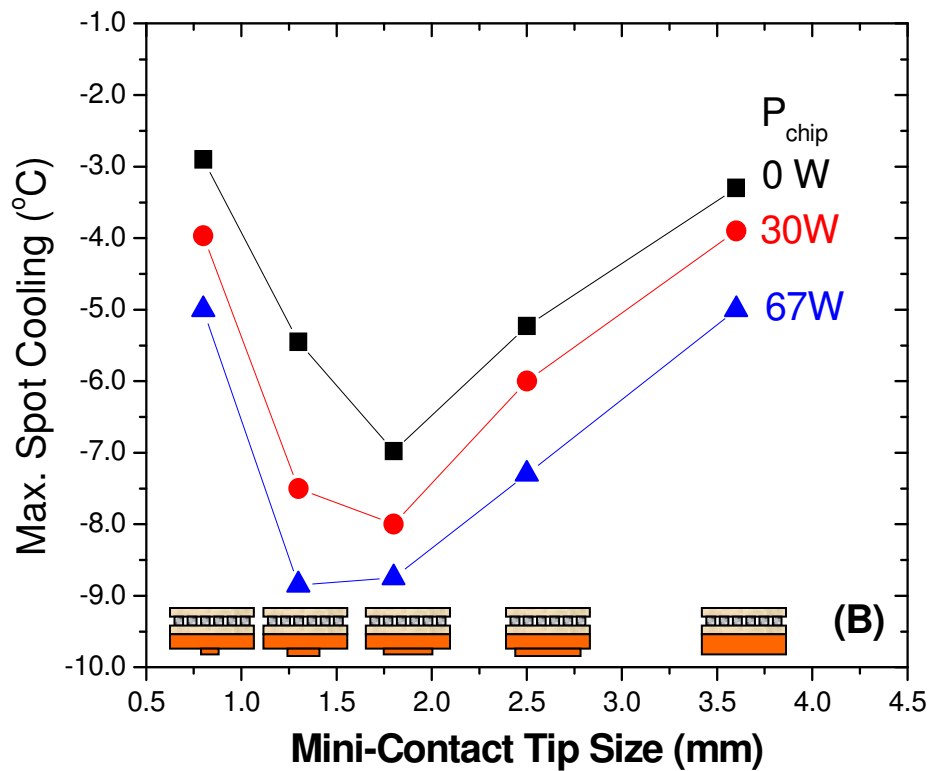
Figure 7.18: Experimental dependence of maximum spot cooling on copper mini-contact tip size.

Figure 7.18 displays the experimentally observed effect of the mini-contact tip size on the temperature reduction at the targeted spot. For the three different power dissipations and a 500 µm thick chip, the maximum spot cooling is seen to display a parabolic dependence on the tip size, showing very favorable improvements as the mini-contact tip size decreases in area from the "full coverage" limit, but ultimately reversing direction as the tip size shrinks below an optimum value and approaches point contact. The presence of an optimum tip size reflects the competing effects of the favorable cooling flux concentration and the parasitic spreading resistance in the mini-contact tip. As shown in Figure 7.18 for the case of no power dissipation on the silicon chip, if the mini-contact is of the same size as the TEC base, the measured maximum spot cooling is about 3.3°C. However, if a 1.8 mm × 1.8 mm copper mini-contact is integrated onto the TEC, 7.1°C maximum spot cooling can be obtained which results in 115% improvement on spot cooling performance. Similarly, spot cooling performance can be improved by 100% and 80% if the power dissipation of the silicon chip is 30 W and 67 W, respectively. It is interesting to note that the power dissipation on the silicon chip has an impact on the optimized mini-contact size and the larger the power dissipation on the silicon chip, the smaller the optimized mini-contact size. As clearly illustrated in Figure 7.18, as the power dissipation on the silicon chip increases from 0 to 67 W the optimized mini-contact size reduces from 1.8 mm × 1.8 mm to 1.3 mm × 1.3 mm.
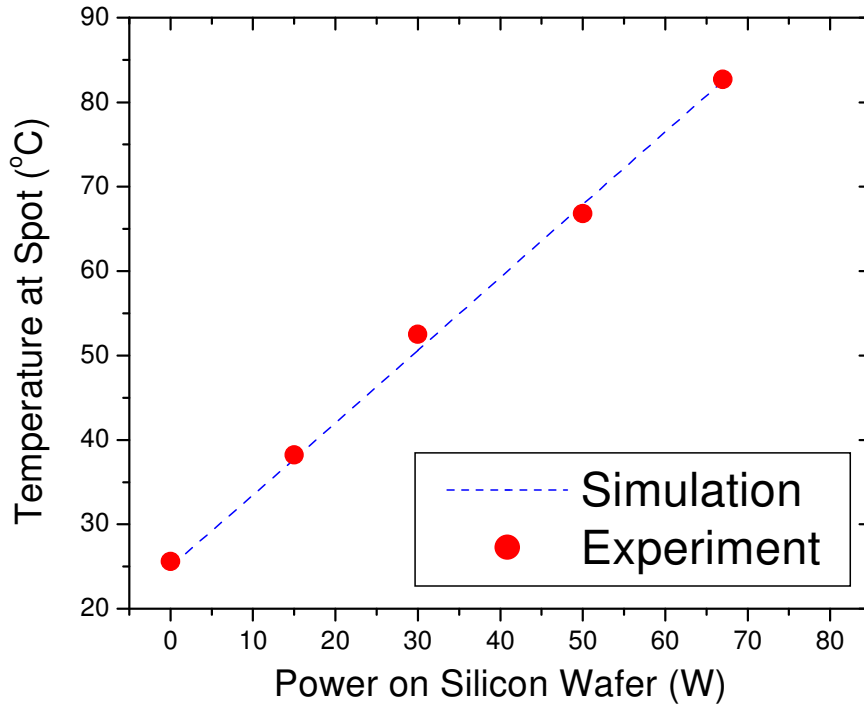
Figure 7.19: Temperature at the targeted spot as a function of chip power dissipation when TEC is turned off.
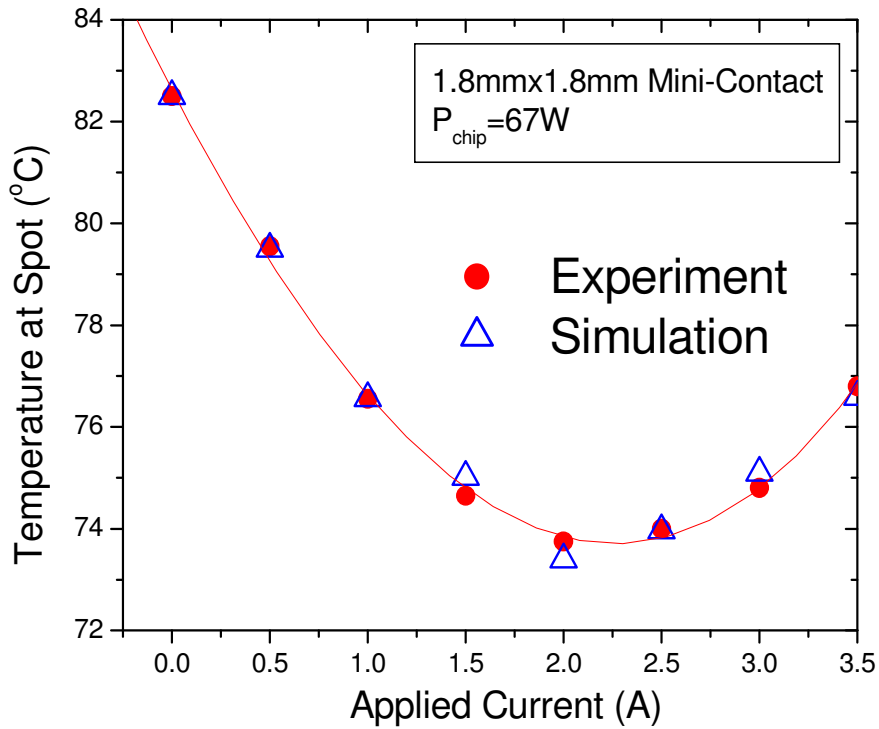
Figure 7.20: Temperature at the targeted spot as a function of applied current on TEC with chip power dissipation of 67W.

Figure 7.19 compares simulation results with experimental data of the targeted spot temperature when the TEC is off and the chip power dissipation varies from 0 to 67 W. Figure 7.20 shows such comparison when the TEC is activated with the mini-contact tip size of 1.8 mm × 1.8 mm and the chip power dissipation of 67 W. The numerical simulations were performed with assumed values of thermal contact resistance equal to 2×10^{-6} m^2 K/W and 8×10^{-5} m^{2}K/W, at the mini-contact/chip interface (e.g. the solder interface) and TEC/TIM interface (e.g. the thermal grease interface), respectively. These values were obtained using data extraction techniques [134] and were found to agree well with the reported thermal resistance values in the literature for solder and thermal grease interfaces, respectively. With these contact resistance values the experimental data are found to be in good agreement with the simulation results for the uniformly heated chip. This validation of the numerical model for the mini-contact cooled spot and also provides strong support for the earlier discussed, hot spot cooling results obtained via numerical simulation.

7.3 Conclusions

Mini-contact enhanced TEC has great potential application for on-chip hot spot cooling, providing significant cooling improvement relative to directly attached TEC cooling, by concentrating the thermoelectric cooling power on small regions of the silicon chip. In this chapter, the package-level experimental results demonstrates that use of the optimized copper mini-contact pad can improve spot cooling by 80~115%, depending on the power dissipation on the silicon chip showing the great promise of this novel technique for on-chip hot spot cooling. Close agreement between the experimental data and the numerical simulation, with inversely-

determined thermal contact resistances at the solder interface and the thermal grease interface, respectively, validates the TEC modeling approach used in this study and provides strong support for the hot spot cooling results obtained via numerical simulation.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

This doctoral dissertation addresses advanced applications of thermoelectric coolers to on-chip hot spot temperature reductions, including the use of a silicon microoooler and a mini-contact enhanced thermoelectric cooler. A novel silicon microcooler, based on the unique thermoelectric properties of single-crystal silicon, is proposed and applied to on-chip hot spot cooling by using the silicon die itself as the thermoelectric material and thus dramatically simplifying the fabrication process and eliminating the parasitic thermal contact resistance effect. The other novel application developed in this dissertation involves the use of a mini-contact enhanced TEC for on-chip hot spot cooling, where the mini-contact, connecting the silicon die and the TEC, is used to concentrate the thermoelectric cooling flux on the top surface of the silicon die and therefore significantly improve the hot spot cooling performance. The objective of this dissertation is to explore the thermal physics involved in the application of thermoelectric coolers to on-chip hot spot cooling, to develop suitable predictive models for the hot spot temperature, and to design and optimize the respective thermoelectric coolers, as well as the chip package, to maximize hot spot cooling performance. The conclusions reached in the course of this dissertation are summarized below:

Silicon Thermoelectric Microcooler: To understand the thermal physics and develop a building block for subsequent study of on-chip hot spot cooling using silicon microcooler, a device-level analytical thermal model for such a microcooler,

which couples Peltier cooling with heat conduction and heat generation in the silicon substrate, and which includes heat conduction and heat generation in the metal lead, is derived and used to study the thermal characteristics of silicon thermoelectric microcoolers under various operating conditions. It is found that the analytical modeling results are in excellent agreement with experimental data and detailed numerical simulations. The analytical model, thus, provides a very convenient approach to the design and optimization of silicon thermoelectric cooling devices, and can serve as well to investigate the complex thermal physics involved in silicon thermoelectric microcooler systems. Using the analytical model - in combination with the numerical simulations - it has been found that the optimum doping concentration lies in the range of $2.5 \times 10^{19} cm^{-3}$, some three order of magnitude higher than the commonly used semiconductor doping, and that larger electric contact resistances will push the optimum doping concentration to a lower level. In the ideal case, it is found that the silicon microcooler - placed in an air, natural convection environment - could achieve a maximum microcooler temperature reduction of about $6^{o}C$ at near zero heat flux and, alternatively, extract a heat flux of several thousand's W/cm^2 at a negligibly small temperature reduction.

Silicon Microcooler for On-Chip Hot Spot Cooling: On-chip hot spot cooling using a silicon microcooler is investigated using a three-dimensional chip-level analytical model as well as a three-dimensional package-level electro-thermal numerical simulation. The goal is to investigate the hot spot cooling potential using the inherent thermoelectric properties of the silicon chip. For analytical modeling, allocation factors extracted from the electro-thermal numerical simulations are used

219

in the analytical model to account for the impact of silicon Joule heating on the hot spot and the microcooler temperatures. The resulting analytical model can be used to quickly study the parametric sensitivity of hotspot cooling under a variety of geometric and operating conditions, while the numerical simulation, which models the five-layer package structure in details, can provide more complete temperature and heat flux distributions. In the parametric range studied, i.e., doping concentration in silicon varying from 1.0×10^{18} to 1.0×10^{20} cm^{-3}, silicon die thickness varying from 100 to 500 µm, microcooler size varying from 100 µm × 100 µm to 5000 µm × 5000 µm, and electric contact resistance varying from 1.0×10^{-9} to 1.0×10^{-4} Ω.cm^2, the optimum microcooler size is found to vary from 5 to 6 times the die thickness. The optimized doping concentration of approximately 2.5×10^{19} cm^{-3} is found to be insensitive to system geometry but dependent on parasitic effects, with high electric contact resistance pushing the optimized doping to lower levels. For large hot spots operating at high heat flux, temperature reductions of approximately 4°C can be achieved. For optimum microcooler designs, it is found that the temperature rise engendered by the hot spot could be partially suppressed, completely removed, or even over-cooled, showing the promise of silicon thermoelectric microcoolers for on-chip hotspot removal.

Mini-Contact Enhanced TEC for On-Chip Hot Spot Cooling: A mini-contact enhanced TEC is found to have great potential for on-chip hot spot cooling, providing significant cooling improvement relative to directly attached TEC device, by concentrating the thermoelectric cooling on small regions of the silicon die. A three dimensional numerical model is developed to predict on-chip hot spot cooling

performance using this novel technique. The numerical simulations show that the peak temperature on the silicon die with a large, high flux hot spot could be reduced by as much as 17°C, with an optimized mini-contact pad. However, maintaining low thermal contact resistance along the mini-contact surfaces is critical to the performance of this cooling technique. A chip package-level experiment was designed and performed, demonstrating that use of an optimized copper mini-contact pad with an advanced miniaturized thermoelectric cooler can improve spot cooling by as much as 115% relative to the un-enhanced configuration on a 500 µm thick silicon die. Close agreement between the experimental data and the numerical simulation, with inversely-determined thermal contact resistances at the solder interface and the thermal grease interface, respectively, served to validate the TEC modeling approach used in this study and provided strong support for the hot spot cooling with a mini-contact enhanced TEC.

8.2 Future Work

In this dissertation extensive modeling effort is devoted to the prediction of device-level silicon microcooler cooling performance and package-level hot spot cooling performance. However, only limited experimental work, based on microfabrication techniques, is performed. Therefore, most of our future work should focus on the fabrication of microcooler devices and experimental characterization of their thermoelectric cooling performance as follows:

Fabrication/Testing of Silicon Microcoolers: Some future effort should focus on forming and assembling silicon microcoolers using microfabrication techniques

and testing these microcoolers at 100$^\text{o}$C or higher. Then such silicon microcoolers should be assembled into chip packages to study their hot spot cooling capability. Since the cooling performance depends on the doping concentration, silicon microcoolers with different doping concentrations should be fabricated and tested. This work would also have scientific importance in the thermal physics community.

Low Resistance Thermal Interfaces for Mini-Contact TEC's: The mini-contact enhanced TEC devices are another research area which has great potential application. Future work could focus on applying advanced microelectronic packaging techniques, commonly used in the electronic industry, to develop low thermal resistance solder bonds at the mini-contact/silicon die interface, mini-contact/TEC interface, and TEC/heat spreader interface. Alternatively, copper mini-contact pads could be patterned onto the silicon die using thin-film techniques to significantly reduce the thermal contact resistance at the mini-contact/silicon die interface. Alternatively, the TEC ceramic substrate could be replaced by a copper mini-contact pad so as to eliminate the original mini-contact/TEC interface completely.

Liquid/Phase Change Cooling of TEC: The current thermoelectric cooling capability is constrained by the thermoelectric properties of the materials and by the effective heat transfer coefficient of the heat removal component. It could be our future work that other cooling techniques can be combined to current TEC cooling technique to achieve better hot spot cooling performance. As is known, currently thermoelectric heating is removed through heat sink and fan which is not powerful as expected. Greater effectiveness than explored in this dissertation could be achieved

by assembling the TEC onto a heat pipe or liquid-cooled microchannel cooler to more effectively cool the thermoelectric device and improve hot spot cooling significantly.

Thermal Modeling: To eliminate the need for allocation factors in the analytical solution of the temperature field produced by the silicon thermoelectric microcooler, it would be desirable to derive the three-dimensional analytical solution for non-uniform Joule heating inside the silicon die (3D Poison's equation). While it is unlikely that an exact analytical solution can be found for an arbitrary distribution of internal heat generation, the identification and derivation of solutions for distinct Joule heating distributions could be most helpful in the design and analysis of silicon-based on-chip thermoelectric microcoolers.

References

[1] ITRS, The International Technology Roadmap for Semiconductors, 2004, Semiconductor Industry Association, http://www.ITRSnemi.org.

[2] Pedram, M., and Nazarian, S., "Thermal modeling, analysis, and management in VLSI circuits: principles and methods," *Proceedings of IEEE*, Vol. 94, No.8, pp.1487-1518, 2006.

[3] Black, J. R., "Electromigration - A brief survey and some recent results," *IEEE Transactions on Electron Devices*, Vol. ED-16, pp. 338-347, 1969.

[4] Viswannath, R., Wakharkar, V., Watwe, A., and Lebonheur, V., "Thermal performance challenges from silicon to systems," *Intel Journal of Technology*, Vol. 5, No. 3, pp. 1-16, 2000.

[5] Mahajan, R., Chiu, C., and Chrysler, G., "Cooling a microprocessor chip," *Proceedings of IEEE*, Vol. 94, No.8, pp.1476-1486, 2006.

[6] Bar-Cohen, A., Arik, M., Ohadi, M., "Direct Liquid Cooling of High Flux Micro and Nano Electronic Components," *Proceedings of the IEEE*, Vol. 94, No. 8, pp.1549-1570, 2006.

[7] Mudawar, I., "Assessment of high-heat-flux thermal management schemes," *IEEE Transactions on Components and Packaging Technologies, Part A: Packaging Technologies,* Vol. 24, No.2, pp. 122-141, 2001.

[8] Garimella, S. V., "Advances in mesoscale thermal management technologies for microelectronics," *Microelectronics Journal*, Vol. 37, pp.1165–1185, 2006.

[9] Shelling, P., Li, S., and Goodson, K. E., "Managing Heat for Electronics," *Materials Today*, Vol.8, No.6, pp. 30-35, 2005.

[10] iNEMI, Electronics Manufacturing Initiative Technology Roadmap, 2004, http://www.nemi.org.

[11] Shakouri, A., and Zhang, Y., "On-chip solid-state cooling for integrated circuits using thin-film microrefrigerators," *IEEE Transactions on Components and Packaging Technologies, Part A: Packaging Technologies*, Vol. 28, No.1, pp. 65 - 69, 2005.

[12] Watwe, A., and Viswanath, R., "Thermal Implications of Non-Uniform Die Power Map and CPU Performance," *Proceedings of InterPACK'03*, Maui, Hawaii, USA. Paper No. 35044, 2003.

[13] Dahlgren, S., "High-pressure polycrystalline diamond as a cost effective heat spreader," *Proceedings of Thermal and Thermomechanical Phenomena in Electronic Systems, (ITHERM 2000)*, Vol.1, pp.23-26, 2000.

[14] Jannadham, K., Watkins, T. R., and Dinwiddie, R. B., "Novel heat spreader coatings for high power electronic devices," *Journal Materials Sciences*, Vo.37, 1363 – 1376, 2002.

[15] Goodson, K. E., and Ju, Y. S., "Heat conduction in novel electronic films," *Annual Review of Materials Science*, Vol. 29, pp. 261-293, 1999.

[16] Benson, D., A., Mitchell, R., T., Tuck, M., R., Palmer, D., W., and Peterson, G., P., "Ultrahigh-capacity micromachined heat spreaders," *Microscale Thermophysics Enginggering*, Vol. 2, pp. 21-30, 1998.

[17] Thermacore Inc, http:// www.thermacore.com

[18] Akachi, H., Structure of a Heat Pipe, U.S. Patent #4,921,041, 1990.

[19] Zuo, Z. J., North, M. T., and Wert, K. L., "High heat flux heat pipe mechanism for cooling of electronics," *IEEE Transactions on Components and Packaging Technologies*, Vo. 24, No.2, pp.220-225, 2001.

[20] Lin C., Ponnappan, R., and Leland, J., "High performance miniature heat pipe," *International Journal of Heat and Mass Transfer*, Vol.45, No. 15, pp.3131-3142, 2002.

[21] Lee, M., Wong, M., and Zohar, Y., "Characterization of an integrated micro heat pipe," *Journal of Micromechanics and Microengineering*, Vol. 13, No.1, pp.58-64, 2003.

[22] Plesch, D., Bier, W., Siedel, D., and Schubert, K., "Miniature heat pipes for heat removal from microelectronic circuits," *Micromechanical Sensors, Actuators and Systems. New York*: ASME, Vol. DCS-32, 1991.

[23] Khrustalev, D., and Faghri, A., "Thermal characteristics of conventional and flat miniature axially-grooved heat pipes," *Journal of Heat Transfer*, Vol.117, No.3, pp.740-7,1995.

[24] Therma-Base™ Heat Sink Application Notes, Thermacore, Inc., Lancaster, PA, 1999.

[25] Gillot, G., Avenas, Y., Cezac, N., Poupon, G., Schaeffer, C., and Fournier, E., "Silicon heat pipes used as thermal spreaders," *IEEE Transactions on Components and Packaging Technologies*, Vol. 26, No.2, pp. 332-339, 2003.

[26] Tuckerman, D. B., and Pease, R. F. W., "High-performance heat sinking for VLSI," *IEEE Electron Device Letters*, Vol. EDL-2, pp.143-150, 1981.

[27] Prasher, R., Chang, J., Sauciuc, I., Narasimhan, S., Chau, D., Chrysler, G., Myers, A., Prstic, A., and Hu, C., "Nano and micro technology-based next-generation package level cooling solutions," *Intel Journal of Technology*, Vol.9, No.4, pp.285-296, 2005.

[28] Bowers, M., and Mudawar, I., "High-flux boiling in low-flow rate, low-pressure drop mini-channel and microchannel heat sinks," *International Journal of Heat and Mass Transfer*, Vo. 37, No.2, pp.321-332, 1994.

[29] Mudawar, I. Maddox, D. E., "Enhancement of critical heat flux from high power microelectronic heat sources in a flow channel", *Journal of Electronic Packaging*, Vol.112, No.3, pp.241-248., 1990.

[30] Garimella, S. V., Singhal, V., and Liu, D., "On-chip thermal management with microchannel heat sinks and integrated micropumps," *Proceedings of the IEEE*, Vol. 94, No. 8, pp.1534-1548, 2006.

[31] Zhang, L., Wang, E., N., Koo, J. M. Goodson, K. E., Santiago, J. G., Kenny, T. W., "Microscale liquid jet impingement", *Proceedings of AMSE IMECE*, Nov. 2001, Vol.2, MEME-23820.

[32] Wolf, D. H., Incropera, F. P., Viskanta, R., "Local jet impingement boiling heat transfer", *International Journal of Heat and Mass Transfer*, Vol.39, No.7, pp.1395-1406, 1996.

[33] Kiper, A. M., "Impinging water jet cooling of VLST circuits", *International Communications in Heat and Mass Transfer*, Vol. 11, pp.126-129, 1984.

[34] Harman, T. C., Taylor, P. J., Walsh, M. P. and LaForge, B. E., "Quantum dot superlattice thermoelectric materials and devices," *Science*, Vol. 297, No. 5590, pp.2229-2232, 2002.

[35] Venkatasubramanian, R., Siivola, E., Colpitts, T., and O'Quinn, B., "Thin-film thermoelectric devices with high room-temperature figures of merit," *Nature* (London), Vol. 413, No. 6856, pp.597-602, 2001.

[36] Fan, X., Zeng, G., LaBounty, C., Bowers, J., Croke, E., Ahn, C., Huxtable, S., and Majumdar, A., "SiGeC/Si superlattice micro-coolers," *Applied Physics Letters*, Vol. 78, No.11, pp.1580-1600,2001.

[37] Chen, C., Yang, B., and Liu, W. L., "Engineering nanostructures for energy conversion," in *Heat Transfer and Fluid Flow in Microscale and Nanoscale Structures*, M. Faghri and B. Sunden, Eds.: WIT Press, pp. 45-91,2004.

[38] Yang, B., Liu, W. L., Wang, K. L., and Chen, G., 2002, "Simultaneous measurements of Seebeck coefficient and thermal conductivity across superlattice," *Applied Physics Letters*, Vol. 80, pp. 1758-1760.

[39] Zhang, Y., Zeng, G. H., Piprek, J., Bar-Cohen, A., and Shakouri, A., "Superlattice microrefrigerators fusion bonded with optoelectronic devices," *IEEE Transactions on Components and Packaging Technologies*, Vol. 28, No.4, pp. 658-666, 2005.

[40] Simons, R. E., Ellsworth, M. J., and Chu, R., C., "An assessment of module cooling enhancement with thermoelectric coolers," *Journal of Heat Transfer*, Vol. 127, No.1, pp. 76-84, 2005.

[41] Yeh, L., and Chu, C., *Thermal Management of Microelectronic Equipment*, ASME Press, New York, 2002.

[42] Kraus, A. D., and Bar-Cohen, A., *Thermal Analysis and Control of Electronic Equipment*, Hemisphere Publishing Corporation, New York, USA, 1983.

[43] Fan, X., Ph.D. Thesis, University of California at Santa Barbara, March 2002.

[44] Fleurial, J.-P., Borshchevsky, A., Ryan, M. A., Phillips, W., Kolawa, E., Kacisch K., and Ewell, R., "Thermoelectric microcoolers for thermal management applications," *Proceedings of 16th International Conference on Thermoelectrics*, pp. 641-645, 1997.

[45] Semenyuk, V., "Thermoelectric micro modules for spot cooling of high density heat sources," *Proceedings of the 20th International Conference on Thermoelectrics,* pp.391-396, 2001.

[46] www.thermion–company.com.

[47] Nolas, G. S., Sharp, J. and Goldsmid, H. J., *Thermoelectrics Basic Principles and New Materials Developments*, Springer, 2001.

[48] Rowe, D. M., *CRC Handbook of Thermoelectrics*, Boca Raton, FL: CRC Press, 1995.

[49] Hannay N. B., *Semiconductors*, Reinhold Publishing Corp, 1959.

[50] Decher, R., *Direct Energy Conversion*, Oxford University Press, Inc., 1977.

[51] Ioffe, A. F., *Semiconductor Thermoelements and Thermoelectric Ccooling.*, London, UK: Infosearch Ltd., 1957.

[52] Ettenberg, M. H., Jesser, M. A., and Rosi, E., D., "A new n-type and improved p-type pseudo-ternary $(Bi_2Te_3)(Sb_2Te_3)(Sb_2Se_3)$ alloy for Peltier cooling," *Proceedings of the 15th International Conference on Thermoelectrics*, pp. 52-56, 1996.

[53] Yamashita, O., and Tomiyoshi, S., "Effect of Annealing on Thermoelectric Properties of Bismuth Telluride Compounds," *Japan Journal of Applied Physics*, Part.1, Vol. 42, No. 2A, pp.492-500, 2003.

[54] Yamashita, O., Tomiyoshi, S., and Makita, K., "Bismuth telluride compounds with high thermoelectric figures of merit," *Journal of Applied Physics*, Vol.93, No.1, pp.368-374, 2003.

[55] Yamashita, O., and Tomiyoshi, S., "High performance n-type bismuth telluride with highly stable thermoelectric figure of merit", *Journal of Applied Physics*, vol. 95, no. 11, pp. 6277–6283, 2004.

[56] Vashaee, D, and Shakouri, A., "Improved thermoelectric power factor in metal-based Superlattices," *Physical Review Letters*, Vol.92, No.10, pp.106103-1-4, 2004.

[57] Volklein, F., Baier, V., Dillner, U., and Kessler, E., "Transport properties of flash-evaporated $(Bi_{1-x}Sb_x)_2Te_3$ films I: Optimization of film properties," *Thin Solid Films*, Vol. 187, pp. 253-262, 1990.

[58] Mzerd, A., Sayah, D., Tedenac, J.C., and Boyer, A., "Effect of substrate temperature on crystal growth of Bi_2Te_3 on single crystal Sb_2Te_3," *Journal of Material Science Letter*, Vol. 13, No.5, pp. 301-304, 1994.

[59] Giani, A., Pascal-Delannoy, F., Boyer, A., Foucaran, A., Gschwind, M., and Ancey, P., "Elaboration of Bi_2Te_3 by metal organic chemical vapor deposition," *Thin Solid Films*, Vol.303, No.1-2, pp. 1-3, 1997.

[60] Dauscher, A., Thomy, A., and Scherrer, H., "Pulsed laser deposition Bi_2Te_3 thin film," *Thin Solid Films*, Vol.280, No.1-2, pp. 61-6, 1996.

[61] Stolzer, M., Bechstein, V., and Meusel, J., "Deposition Bi_2Te_3 films by magnetron co-sputtering from two targets-first results," *Proceedings of 16th International Conference on Thermoelectrics*, pp. 422-424, 1996.

[62] Zhou, H., Rowe, D. M., and Williams, S.,G.,W., "Peltier effect in a co-evaporated Sb_2Te_3(P)-Bi_2Te_3 (N) thin film thermocouple," *Thin Solid Films*, Vol. 408, No.1, pp. 270–274, 2002.

[63] Raghuveer, S., Makala, K., Jagannadham, A., and Sales, B.,C., "Pulsed laser deposition of Bi_2Te_3-based thermoelectric thin films", *Journal of Applied Physics*, Vol,94, No.6, pp.3907-3918, 2003.

[64] Nedelcu, M., Sima, M., Visan, T., Pascu, T., Franga, I., and Craciuniou, F., "Bi_2Te_3 thick thermoelectric films obtained by electrodeposition from hydrochloric acid solutions," in *Proceedings of 20th International Conference on Thermoelectrics*, pp. 322-326, 2001.

[65] Pandey, R.K., Sahu, S.N., and Chandra, S., *Handbook of Semiconductor Deposition*, Ed. M. Dekkar, New York, 1996.

[66] Lim, J. M., Snyder, G. J., Huang, C. K., Herman, J. A., Ryan, M. A., and Fleurial, J-P., "Thermoelectric microdevice fabrication process and evaluation at the Jet Propulsion Laboratory (JPL)," *Proceedings of 21rd International Conference on Thermoelectrics*, pp. 535-539, 2002.

[67] Snyder, G. J., Lim, J. R., Huang, C., and Fleurial, J.-P., "Thermoelectric microdevice fabricated by a MEMS-like electrochemical process," *Nature Materials*, Vol.2, No.8, pp.528-531, 2003.

[68] Nedelcu, M., Sima, M., Visan, T., Pascu, T., Franga, I., and Craciuniou, F., "Bi_2Te_3 thick thermoelectric films obtained by electrodeposition from hydrochloric acid solutions," in *Proceedings of the 20th International Conference on Thermoelectrics*, pp. 322-326, 2001.

[69] Fleurial, J. P. Snyder, G. J., Herman, J. A., Giauque, P. H., Philips, W., M., Ryan, M., A., Shakkotai, P., Kolawa, E., A., and Nicolet, M. A., "Thick-film thermoelectric microdevices," *Proceedings of the18th International Conference on Thermoelectrics,* pp. 294-302, 1999.

[70] da Silva, L., W., Kaviany, M., "Fabrication and measured performance of a first-generation microthermoelectric cooler," *Journal of Microelectromechanical Systems*, Vol.14, No.5, pp.1110-1117, 2005.

[71] da Silva, L.W. et al., "Micro thermoelectric cooler fabrication: Growth and characterization of patterned Sb_2Te_3 and Bi_2Te_3 films", *Proceedings of the 22nd International Conference on Thermoelectrics*, pp. 665-668, 2003.

[72] da Silva, L., W., Kaviany, M., and Uher, C., "Thermoelectric performance of films in the bismuth-tellurium and antimony-tellurium systems," *Journal of Applied Physics*, Vol.97, pp.114903-1-10, 2005.

[73] Böttner, H., Nurnus, J., Gavrikov, A., Kühner, G., Jägle, M., Künzel, C., Eberhard, D., Plescher, G., Schubert, A., and Schlereth, K., "New Thermoelectric

Components Using Microsystem Technologies," *Journal of Microelectromechanical Systems*, Vol.13, No.3, pp.414-420, 2004.

[74] Bottner, H.," Micropelt miniaturized thermoelectric devices: small size, high cooling power densities, short response time," *Proceedings of the 24th International Conference on Thermoelectrics*, pp.1-8, 2005.

[75] Semenyuk, V., "Miniature thermoelectric modules with increased cooling power", *Proceedings of the 25th International Conference on Thermoelectrics*, pp.322-326, 2006.

[76] Semenyuk, V, Ph.D. Dissertation, Odessa Technological Institute of Food and Refrigerating Engineering, Odessa, USSR, 1967 (in Russian).

[77] Semenyuk, V., "Miniature thermoelectric coolers for semiconductor lasers", *Proceedings of the 13th International Conference on Thermoelectrics* , pp. 150-153, 1994.

[78] Semenyuk, V.et al, "Miniature Thermoelectric Coolers for Semiconductor Lasers", *Proceedings of the 13th International Conference on Thermoelectrics*, pp. 150-153, 1994.

[79] Semenyuk, V. et al, "Novel High Performance Thermoelectric Microcoolers with Diamond Substrates", *Proceedings of the 16th International Conference on Thermoelectrics*, pp. 683-686, 1997.

[80] Semenyuk, V., "Thermoelectric cooling of electro-optic components," *Thermoelectrics Handbook: Macro to Nano*, edited by D.M. Rowe, CRC Press, (Boka Raton, Fl), Chapter 58, pp. 1-21, 2006.

[81] Bierschenk, J., and Johnson, D., "Extending the limits of air cooling with thermoelectrically enhanced heat sinks," *Proceedings of ITHERM'04*, PP.679-684, 2004.

[82] Hicks, L. D., and Dresselhaus, M.S., **"**Thermoelectric figure of merit of a one-dimensional conductor," *Physics Review* B Vol.47, No.24, 16631-16634,1993.

[83] Balandin, A., and Wang, K. L. "Effect of phonon confinement on the thermoelectric figure of merit of quantum wells," *Journal of Applied Physics*, Vol.84, No.11, pp. 6149-6153,1998.

[84] Balandin, A., and Lazarenkova, O. L., "Mechanism for thermoelectric figure-of-merit enhancement in regimented quantum dot superlattices," *Applied Physics Letters*, Vol.82, No.3, pp. 415-417,2003.

[85] Yang, B., and Chen, G., *Thermal Conductivity: Theory, Properties and Application*, Edied by T.M. Tritt, Kluwar Press, pp. 167-186, 2005.

[86] Mahan, G. D., and Woods, L. M., "Multilayer thermionic refrigeration," *Physical Review Letter*, Vol. 80, No. 18, 4016-4019, 1998.

[87] Shakouri, A., LaBounty, C., Piprek, J., Abraham, P., and Bowers, J. E., "Thermionic emission cooling in single barrier heterostructures," *Applied Physics Letters*, Vol.74, No.1, pp. 88-89, 1999.

[88] LaBounty, C., Shakouri, A., Abraham, P., Bowers, J. E., "Monolithic integration of thin film coolers with optoelectronic devices," *Optical Engineering*, Vol.39, No.11, 2847–2852, 2000.

[89] LaBounty, C., Shakouri, A., Abraham, P., Bowers, J. E., "Design and characterization of thin film micro-coolers", *Journal of Applied Physics*, Vol. 89, No.7, pp.4059-4064, 2001.

[90] Zhang, J., Anderson, N., Lau, K., "AlGaAs superlattice microcoolers," *Applied Physics Letters*, Vol.83, No.2, pp.374-376, 2003.

[91] Zeng, G., Shakouri, A., LaBounty, C., Robinson, G., Croke, E., Abraham, P., Fan, X., Reese, H., and Bowers, J. E., "SiGe micro-cooler," *Electronics Letter*, Vol. 35, No. 24, pp. 2146–2147, 1999.

[92] Zeng, G., Fan, X., LaBounty, C., Croke, E., Zhang, Y., Christofferson, J., Vashaee, D., Shakouri, A., and Bowers, J. E., "Cooling power density of SiGe/Si superlattice micro refrigerators," *Proceedings of Thermoelectric Materials Research and Applications,* Vol. 793, pp. 43–49, 2003.

[93] Fan, X., Zeng, G., LaBounty, C., Vashaee, D., Christofferson, J., Shakouri, A., and Bowers, J. E., "Integrated cooling for Si-based microelectronics,: *Proc. 20th Int. Conf. on Thermoelectrics,* pp.405-408, 2001.

[94] Fan, X., Zeng, G., LaBounty, C., Vashaee, D., Shakouri, A., and Bowers, J. E., B, "High cooling power density SiGe/Si micro coolers,[Electron. Lett., vol. 37, no. 2, pp. 126–127, Jan. 18, 2001.

[95] Venkatasubramanian, R., Colpitts, T., Liu, S., El-Masry,N., and Lamvik, M.,"Low-temperature organometallic epitaxy and its application to superlattice structures in thermoelectrics," Applied Physics Letter, Vol.75, No. 8, pp. 1104-1106, 1999.

[96] Herwaarden, A. W., and Sarro, P. M., "Thermal Sensors Based on the Seebeck Effect," *Sensors and Actuators*, Vol.10, pp. 321-346, 1986.

[97] Geballe, T. H., and Hull, G. W., "Seebeck Effect in Silicon", *Physical Review*, Vol. 98, No.4, pp.940-70, 1955.

[98] Horn, F. H., "Densitometric and Electrical Investigation of Boron in Silicon", *Physical Review*, Vol. 97, No.6, pp. 1521–1525, 1955.

[99] Cuttle, M., and Mott, M. F., "Observation of Anderson localization in an electron gas", *Physical Review*, Vol.181, No.3, pp.1336-40, 1969.

[100] Fritzsche, H., "A general expression for the thermoelectric power", *Solid State Communication*, Vol.9, No. 21, pp.1813-15, 1971.

[101] Zhang, Y., Shakouri, A., and Zeng, G., "High-power-density spot cooling using bulk thermoelectrics," Applied Physics Letters, Vol. 85, No.14, 2977-2979,2004.

[102] Wang, P., Bar-Cohen, A., Yang, B., and Shakouri, A., "Thermoelectric Microcooler for Hot Spot Thermal Management", *Proceedings of the Pacific Rim/ASME International Electronic Packaging Technical Conference and Exhibition (InterPack'05)*, Paper No: IPACK2005‑73244.

[103] Solbrekken, G. L., Zhang, Y., Bar-Cohen, A., Shakouri, A., "Hot spot reduction in convectively-cooled chip", *Proceedings of 19th Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, Vol. 2, pp.610-616, 2004.

[104] Daryoosh, V., Christofferson, J., Zhang, Y., Shakouri, A., Zeng, G., LaBounty, C., Fan, X., Bowers, J. E., and Croke, E., "Modeling and optimization of single-element bulk SiGe Thin-film coolers," *Microscale Thermophysical Engineering*, Vol.9, No. 1, pp.99- 118, 2005.

[105] Wijngaards, D.D.L., de Graaf, G., Wolffenbuttel, R. F, "Single-chip micro-thermostat applying both active heating and active cooling," *Sensors and Actuators A*, Vol.110, No.1-3, pp. 187-95, 2004.

[106] LaBounty, C., Shakouri, A., and Bowers, J., E., "Monolithic integration of thin-film coolers with optoelectronic devices," *Optical Engineering*, Vol.39, No.11, 2847-2852, 2000.

[107] Wijngaards, D.D.L.; Wolffenbuttel, R.F., "Study on temperature stability improvement of on-chip reference elements using integrated Peltier coolers," *IEEE Trans. Instrumentation and Measurement*, Vol.52, No.2, pp. 478-82, 2003.

[108] Kong, H., Wijngaards, D.D.L., Wolffenbuttel, R.F , "Study of IC compatible on-chip thermoelectric coolers," *Japanese Journal of Applied Physics*, Part 1 Vol.44, No.7B, pp. 5736-9, 2005.

[109] Wijngaards, D.D.L.; Kong, S.H.; Bartek, M.; Wolffenbuttel, R.F., "Design and fabrication of on-chip integrated polySiGe and polySi Peltier devices," *Sensors and Actuators* A, Vol.A85, No.1-3, pp. 316-23, 2000.

[110] Nakwaski, W., and Osinski, M., "Thermal resistance of top-surface-emitting vertical-cavity semiconductor lasers and monolithic two-dimensional arrays," *Electronics Letters,* Vol.28, No.6, 572 - 574 (1992).

[111] Negus, K. J., Yovanovich, M. M., and Roulston, D. J., "An Introduction to thermal electrical coupling in bipolar transistors," *Proceedings of ASME/JSME Thermal Engineering Joint Conference*, Vol.3, p. 395-401, 1987.

[112] Nieveld, G. D., "Thermopiles Fabricated Using Silicon Planar Technology," *Sensors and. Actuators*, Vol.3,pp.179-183, 1983.

[113] Chapman, P.W., Tfte, O. N., Zook, J., D., and Long, D., "Electrical properties of heavily doped silicon," *Journal of Applied Physics*, Vol.34, No.11, pp.3291-3295, 1963.

[114] Asheghi, M., Kurabayashi, K., Kasnavi R., and Goodson, K.E., "Thermal conduction in doped single-crystal silicon films," *Journal of Applied Physics*, Vol. 91,No.8, 5079-5088, 2002.

[115] Chang, C. Y., Fang, Y. K., and Sze, S., M., "Specific contact resistance of metal-semiconductor barriers," *Solid-State Electronics*, Vol.14, No.7, pp.541-550, 1971.

[116] Akiya, M., and Aihara, M., "Specific contact resistance at extremely low temperatures," *Journal of Physics D*, Vol.30, No. 2, 271-273, 1997.

[117] Oh, K., Chun.J., Banerjee, K., Duvvury,C., and Dutton, R. W., "Modeling of temperature dependent contact resistance for analysis of ESD reliability," *Proceedings of 41st Annual IEEE International Reliability Physics Symposium*, pp. 249-255, 2003.

[118] *CRC Handbook of Chemistry and Physics*, 75th edition, edited by D R Lide (CRC, Boca Raton, FL,1994).

[119] Muzychka, Y.S., Culham, J.R. and Yovanovich, M.M., "Thermal Spreading Resistance of Eccentric Heat Sources on Rectangular Flux Channels," *Journal of Electronic Packaging*, Vol. 125, No. 2, pp. 178-185, 2003.

[120] Muzychka, Y.S., Yovanovich, M.M. and Culham, J.R.,Influence of Geometry and Edge Cooling on Thermal Spreading Resistance," *Journal of Thermophysics and Heat Transfer*, Vol. 20, No. 2, pp. 247-255, 2006.

[121] Yovanovich, M.M., Muzychka, Y.S. and Culham, J.R., "Spreading Resistance of Isoflux Rectangles and Strips on Compound Flux Channels," *Journal of Thermophysics and Heat Transfer,* Vol. 13, No. 4, pp. 495-500, 1999.

[122] Hewett, C. A., Taylor, M. J., Zeidler, J. R. and Geis, M. W., "Specific contact resistance measurements of ohmic contacts to semiconducting diamond," *Journal of Applied Physics*, 77(2), pp.755-760, 1996.

[123] Madhusudana, C. V., *Thermal contact conductance*, New York, Springer-Verlag, 1996.

[124] Yang, B., Wang, P., Bar-Cohen, A., "Mini-Contact Enhanced Thermoelectric Cooling of Hot Spot in High Power Devices", *IEEE Transactions on Components and Packaging Technologies*, 2007 (in press).

[125] Yang, B., Wang, P., Bar-Cohen, A., "Thermoelectric Mini-Contact Cooler for Hot Spot Removal in High Power Devices", *Proceedings of the 56th Electronic Components and Technology Conference (ECTC-56)*, San Francisco, California, May 30 - June 2, 2006, pp. 997-1002.

[126] Narasimhan, S., Lofgreen, K., Chau, D., and Chrysler, G., "Thin film thermoelectric cooler thermal validation and product thermal performance estimation," *Proceedings of 10th Intersociety Conference on Thermal and Thermo-mechanical Phenomena in Electronics Systems*, pp.470-475, 2006.

[127] Mahajan, R., Chiu, C., and Prasher, R., "Thermal interface materials: A brief review of design characteristics and materials," *Electronics Cooling*, Vol.10, No.1, 2004.

[128] Singhal, V., Siegmund V., and Garimella, S., "Optimization of thermal interface materials for electronic cooling applications," *IEEE Transactions on Components and Packaging Technologies, Part A:* Vol. 27, No.2, pp. 244-252, 2006.

[129] Lee, C., C., Cho, S., "Fluxless In-Sn bonding process at 140 °C," *Materials Science and Engineering A,* Vol.33, No.1-2, pp. 45-50, 2002.

[130] Lee, C., C., Wang, Y., C., and Matijasevic, G., "A new bonding technology using gold and tin multilayer composite structures," *IEEE Trans. Components Hybrids Manufacture Technologies,* Vol.14, No.2, p. 407-412, 1991.

[131] www.arcticsilver.com.

[132] Gwinn, J. P., and Webb, R. L., "Performance and testing of thermal interface materials," *Microelectronics Journal*, Vol.34, pp.215-222, 2003.

[133] Ramanathan, S., and Chrysler, G. M., "Solid-state refrigeration for cooling microprocessors," *IEEE Transactions on Components and Packaging Technologies, Part A:* Vol. 29, No.1, pp. 179-183, 2006.

[134] Wang, P., Bar-Cohen, A., Yang, B., Solbrekken, G. L., and Shakouri, A., "Analytical modeling of silicon thermoelectric microcooler," *Journal of Applied Physics*, Vol.100, No.1, pp.14501-1-13, 2006.