Thermal Management of Electronics and Optoelectronics: From Heat Source Characterization to Heat Mitigation at the Device and Package Levels

by

Chen Li

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)
in the University of Michigan
2019

Doctoral Committee:

Professor Kevin P. Pipe, Chair
Professor Ellen M. Arruda
Professor Jay Guo
Professor Jinsang Kim
Dr. Paul O. Leisher, Lawrence Livermore National Laboratory

Chen Li
chenlium@umich.edu
ORCID iD: 0000-0002-9592-0277

DEDICATION

To my parents and grandparents.

To Yao.

ACKNOWLEDGEMENTS

Much of the work presented in this thesis is the result of team work by many talented individuals, whom I am deeply indebted to. In particular, I want to thank my advisor, Prof. Kevin Pipe, for his sharp scientific insights and priceless guidance that made my time here at U of M enjoyable and productive. Kevin's ability to distill useful information from complex experiments and his ability to explain intricate things in a way that's both clear and scientifically sound never cease to amaze me. I thank Prof. Jinsang Kim for his keen insights on polymer science, which I learned a lot as a mechanical engineer. His passion on scientific research has served as continued inspiration for me. Dr. Paul Leisher is my co-advisor for the laser back-irradiance project, and it has been my privilege to work under his guidance. I remember I was quite nervous when I broke several diode lasers by mistake, and Paul kindly comforted me by telling me when he was his doctoral study he has broken much more expensive wafers and that it was truly OK.

My fellow labmates at U of M have made this journey truly enjoyable. I am grateful to Prof. Gunho Kim and Dr. Kejia Zhang who mentored me closely in my early years. Working with Apoorv Shankar has been rewarding and enjoyable. We spent countless late nights in lab, and Apoorv has made those intense periods filled with joy. I want to thank Aman Kumar for his support on several projects. I truly enjoyed working with Aman. Hui Si, Paddy Chan, Vahid Rashidi, Shantonio Birch, Connie Lee, Da Seul Yang, Yingying Zeng, Joonkoo Kang, Lixin Zang, Wenhao Shao, Byongseop Song, Kyeongwoon Chung, Ricardo Vazquez - I could not ask for a better set of colleagues and friends, it was an honor to work alongside all of you.

I want to express my thanks to the dear friends that I made at U of M. Longji Cui, Ruijiao Miao, Daning Huang for their warm friendships. It was enjoyable to hang out with them. Lingchen Sun, Aibo Gong, Boyu Tian, Yinghan Xu, Xinyue Zhang, Yining Lu, Wenbo Sun and all the other board game enthusiasts that I met, you have made my time outside of research so much more colorful and joyful.

Finally, I owe the deepest thanks to my family. My parents, Jian Li and Donghong Jin, and my grandma, Shengyu Gao, have always been a continuous source of support. Their love and care have become the solid backbone of my various pursuits in life. I am in gratitude to my twin brother, Hao Li, who is a lot of fun to hang out with. Finally, my wife, Yao (Iris) Zhang, who has been my closest friend and family for the longest time. She has made my life infinitely better. I want to thank her for her understanding when I was truly working hard and ask for her forgiveness when I pretend to work hard from time to time. I can't wait to continue the adventure ahead with her!

Chen Li
Ann Arbor, Michigan
December 26th, 2018

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 4

Chapter 5

Chapter 6

Chapter 7

LIST OF TABLES

ABSTRACT

Thermal management of electronic and optoelectronic devices has become increasingly challenging. For electronic devices, the challenge arises primarily from the drive for miniaturized, high-performance devices, leading to escalating power density. For optoelectronics, the recent widespread use of organic light emitting diode (OLED) displays in mobile platforms and flexible electronics presents new challenges for heat dissipation. Furthermore, the performance and reliability of increasingly high-power semiconductor lasers used for telecommunications and other applications hinge on proper thermal management. For example, small, concentrated hotspots may trigger thermal runaway and premature device destruction.

Emerging challenges in thermal management of devices require innovative methods to characterize and mitigate heat generation and temperature rise at the device level as well as the package level. The first part of this dissertation discusses device-level thermal management. A thermal imaging microscope with high spatial resolution (~450nm) is created for hotspot detection in the context of diode lasers under back-irradiance (BI). Laser facet temperature maps reveal the existence of a critical BI spot location that increases the laser's active region temperature by nearly a factor of 3. An active solid-state cooling strategy that could scale down to the size of hotspots in modern devices is then explored, utilizing energy filtering at carbon nanotube (CNT) junctions as a means to provide thermionic cooling at nanometer spatial scales. The CNT cooler exhibits a large effective Seebeck coefficient of 386μV/K and a relatively moderate thermal conductivity, together giving rise to a high cooling capacity (2.3×10^6 W/cm^2).

Thermal management at the package level is then considered. Heat transfer in polymers is first studied, owing to their prevalence in thermal interface materials as well as organic devices (e.g., OLEDs). Employing molecular design principles developed to engineer the thermal properties of polymers, molecular-scale electrostatic repulsive forces are utilized to modify chain morphologies

in amorphous polymers, leading to spin-cast films that are free of ceramic or metallic fillers yet have thermal conductivities as high as 1.17 $Wm^{-1}K^{-1}$, which is approximately 6 times that of typical amorphous polymers. Electronics packaging designs incorporating phase change materials (PCMs) are then considered as a means to mitigate bursty heat sources; PCM incorporation in a packaged accelerator chip intended for large-scale object identification is found to suppress the peak die temperature by 17%.

Chapter 1
Introduction

1.1 Thermal challenges in electronic and optoelectronic (E&O) devices

A typical computer chip today has an area of only 500-1000 mm^2 yet incorporates ~10 billion transistors with a characteristic length scale of ~10 nm[1,2]. The pursuit of faster and cheaper computing has driven the dimensions of transistors to shrink from micrometer to nanometer spatial scales in less than a decade[2,3]. In addition, the widespread usage of mobile devices, touchscreens and emerging technologies such as flexible displays have motivated intense research and development efforts to come up with new materials and designs compatible with additional requirements such as transparency and flexibility[4].

Thermal challenges often accompany these technologies evolving at a fast pace. First, power densities have been rapidly escalating in both electronic (e.g., the power scaling of transistors in processors) and optoelectronic (e.g., the optical power of semiconductor diode lasers used for pumping solid-state lasers or laser machining[5-8]) devices. The scaling of complementary metal-oxide-semiconductor (CMOS) transistor which is the fundamental building block of modern processors has been predicted by the Moore's law[9], which states that CMOS device density doubles and the dimension halves every 18 months. Even though the switching power per transistor drops with scaling (Fig. 1.1A), the power density of CMOS chips has been predicted to exhibit an increasing trend due to escalating transistor density[10] (Fig. 1.1B). Since the allowable power density is limited by heat removal capability of the chip package, further scaling of CMOS requires innovative thermal management strategies with augmented cooling capacities[3,10].

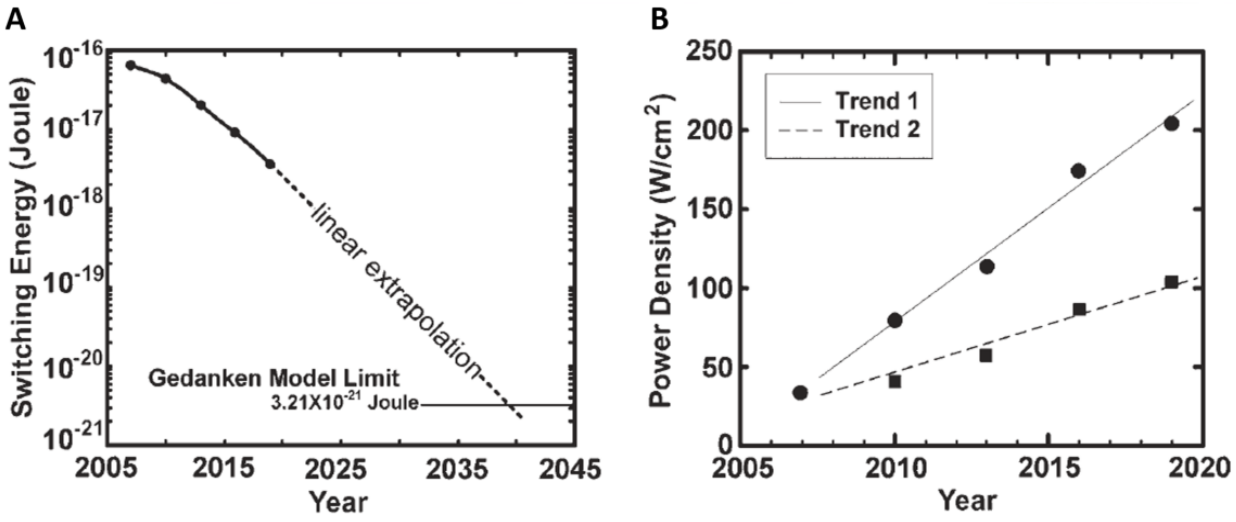

Figure 1.1 | Transistor scaling. (A) The scaling of switching energy for each transistor as predicted by the International Technology Roadmap for Semiconductors (ITRS). **(B)** Trend of integrated circuit power density (Trend 1) as per ITRS projection. The power density can be scaled down (Trend 2) if the switching quantity (i.e. the number of transistors that are turned on simultaneously) is halved. However, this merely alleviates the rapidly escalating power density to a limited extent. Image courtesy of ref. [10].

In addition to a rising power density, spatially non-uniform and/or time-varying heat load represents another key challenge in device thermal management. The spatial nonuniformity of heat generation at transistor junctions[3] or laser facets[7,8] leads to an increasingly prominent local "hotspot" with a temperature significantly greater than the average temperature, making it a potential trigger site for catastrophic thermal runaways. Hotspots are especially prevalent at the facets of edge-emitting diode lasers where their characteristic dimension can be as small as ~1 μm, a length scale at which conventional thermal characterization techniques such as infrared (IR) imaging (which offers a spatial resolution > 1 μm) are inadequate. Therefore, novel thermal characterization techniques with high spatial resolution are invaluable for hotspot detection.

Besides spatial nonuniformity, time-varying heat loads in devices give rise to transient thermal responses which are intimately tied to device heat capacitance. In certain mobile applications such as speech recognition and web browsing, a short burst of computation that lasts only a few seconds could significantly heat up the chip due to its small thermal mass. Since adding mass is undesirable especially in volume-sensitive devices such as cellphones and wearable electronics, incorporating fast-response phase change materials (PCMs) as a means to boost heat capacitance can be very effective in attenuating transient overheating. In this dissertation, an on-chip metallic-PCM

heatsink is investigated, which shows promising initial results for managing thermal transients of an accelerator chip intended for large-scale visualization application.

Lastly, from a heat transfer perspective, the lowered intrinsic thermal conductivity of nanoscale semiconductor building blocks also exacerbates the thermal challenge. Thermal conductivity (κ) characterizes a substance's ability to conduct heat, which directly links to the effectiveness of heat-spreading components. As the device characteristic dimension decreases down to nanometer scale, the classical treatment of heat transfer described by Fourier's law is replaced by a quantum mechanical description in which heat is carried by quantized atomic lattice vibrations (phonons). For insulators and weakly conducting semiconductors in which heat conduction by charge carriers is small, thermal conductivity is dominated by phonon contributions. As the characteristic length scale of the device (e.g., film thickness, channel width, grain boundary size) becomes comparable to or smaller than the phonon mean free path, intense phonon scattering leads to a reduced thermal conductivity compared to the bulk value[11-14] (Fig. 1.2). This holds true for almost all semiconductor materials that are critical for the microelectronics industry, making thermal management even more challenging.

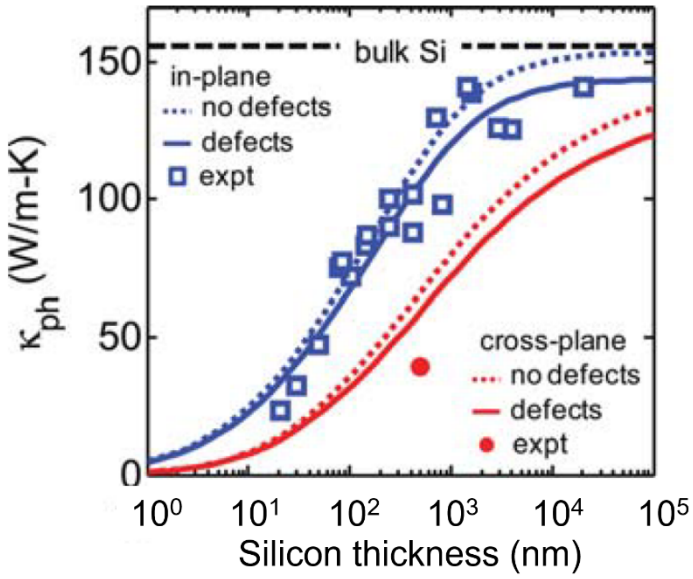

Figure 1.2 | Reduction in silicon thermal conductivity as thickness decreases. The thermal conductivity of silicon as a function of film thickness. A significant reduction in κ is found as length scale decreases down to below 100 nm, with a 4-5 times reduction of the bulk value at 10 nm. The x-axis has been recast in unit of nanometer. Image courtesy to ref. [14].

1.2 Issues caused by device overheating

Device overheating adversely impacts device performance, reliability, and lifetime. Device performance as a function of temperature is rooted in the temperature dependences of various material and transport properties[15], among which the intrinsic carrier density, generation lifetimes, and leakage currents increase exponentially with temperature[14.] A rising leakage current coupled with a decreasing thermal conductivity (κ of crystalline semiconductors scales negatively with temperature) will deteriorate the performance of integrated circuits.

In electronic packages, high temperature typically generates a large thermomechanical stress that triggers several failure modes. Excessive thermomechanical stresses can cause electrical connections (such as wire bonds or solder balls) to break off[16]. Other failure modes include cracking of the silicon die or substrate, resulting from a localized high thermomechanical stress in conjunction with a coefficient of thermal expansion (CTE) mismatch between silicon and other more compliant materials (metals or epoxy resin), and accelerated electro-migration in bonding wires in which high temperature plays a critical role[16].

In semiconductor diode lasers, device degradation originates from dislocations and other defects in the active region, metal intermixing and alloy reactions which deteriorate the electrode, and solder migration that affects the bonding parts[17,18]. All these degradation pathways can be accelerated by a rise in device temperature. Overheating in the active region generates light-absorbing defects that facilitate thermal runaway at the laser facet, causing the semiconductor material to melt and form dark-line defects which are symptoms of catastrophic optical damage (COD). Fig. 1.3 depicts the results of an accelerated aging test for a diode laser bar subject to two different heatsink temperatures, 20 ^{o}C and 50 ^{o}C. For the device tested, it was found that under continuous-wave operation, increasing the heatsink temperature from 20 ^{o}C to 50 ^{o}C decreases the characteristic lifetime (defined as the operation time corresponding to a 50% failure probability) from 11,200 hours to 7,100 hours[19].

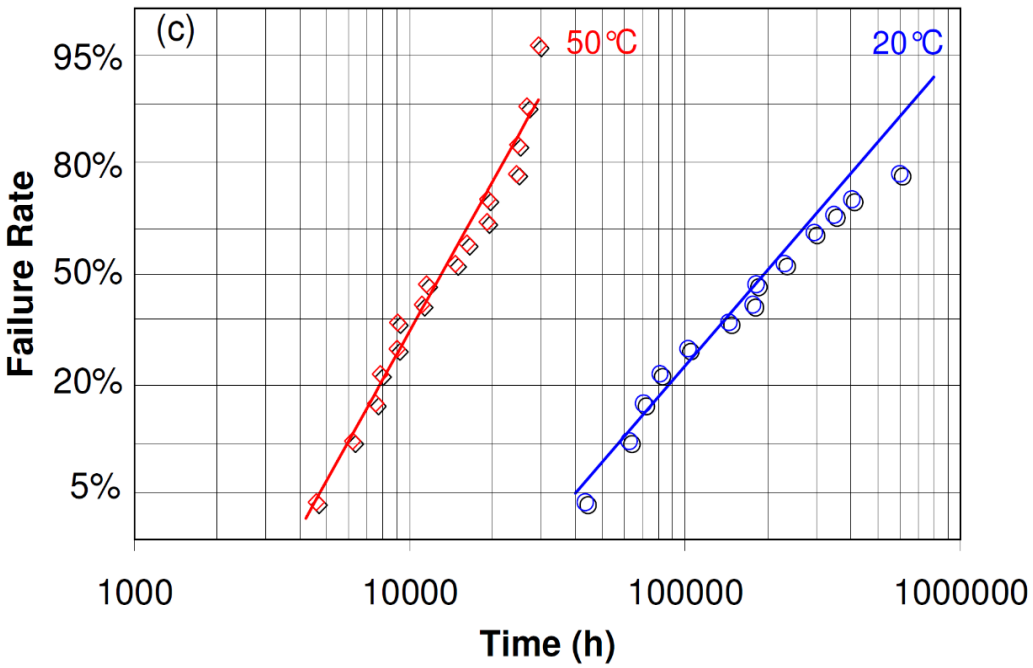

Figure 1.3 | Accelerated aging test of a laser bar subject to two heatsink temperatures. The laser bar compromises 25 high-power, edge-emitting diode lasers that operate in continuous-wave (CW) mode during the aging test. A sizable reduction in lifetime is observed when the heatsink temperature increases. Image courtesy of ref. [19].

1.3 Characterizing and mitigating thermal challenges at the device level

As discussed above, thermal challenges span multiple spatial scales (from device hotspot to packaging) and time scales (short transient to steady-state), and they can cause catastrophic damage to a variety of electronic and optoelectronic devices. As such, this dissertation sets out to explore prototypical examples of thermal issues and ways to mitigate them in several contexts.

This dissertation starts off by introducing several useful tools for microscale thermal measurements. As mentioned in the previous section, highly non-uniform heat dissipation in electronic and optoelectronic (E&O) devices gives rise to localized hotspots that experience much higher temperature than the device average. While the characteristic dimension of hotspots in microprocessors can range anywhere between ~50 μm to several millimeters at the chip level[3], in edge-emitting diode lasers, hotspots typically occur in the active region which is only one to a few microns thick, hampering the usage of an IR camera due to insufficient spatial resolution. As such, high spatial resolution (sub-micron) thermal characterization techniques geared towards hotspot

detection are valuable tools in providing design guidance at the device level; Chapter 2 of this dissertation discusses two such techniques.

Chapter 3 discusses the application of a charge-coupled-device-based thermoreflectance (CCD-TR) technique along with destructive testing to investigate high-power diode lasers subject to back-irradiance (BI). It is well known that high power semiconductor lasers used in real laser systems are subject to different conditions than those in isolated lifetime tests. One such difference is the presence of BI, which can originate from back-scatter by micro-optics, residual specular reflection from optical flats, amplified spontaneous emission from the laser gain medium, etc. BI is known to be a critical factor that accelerates device degradation and COD. Despite the detrimental effect of BI and its prevalence in diode-pumped laser systems, a systematic framework correlating the conditions of back-irradiance (magnitude, BI spot size, BI spot location, etc.) and the temperature rise in the active (light emitting) region to device lifetime has not been established. This study bridges the knowledge gap by controlling the BI spot magnitude and location, and quantitatively measures the temperature at the active region, which is a key parameter that dictates device reliability.

Techniques to remove heat at the device level can be broadly separated into two categories: passive cooling methods that ultimately rely on natural convection for heat removal, or active cooling methods that utilize forced convection or thermoelectric/thermionic effects, etc. Chapter 4 of the dissertation investigates the usage of one-dimensional (1D) carbon nanotubes (CNTs) as active cooling materials via the solid-state thermionic (TI) effect. The compact CNT TI cooler can be placed close to where the micron-size hotspot is, affording a high cooling capacity from a Carnot engine perspective (i.e. large temperature difference between the hot and cold sides). In a TI cooler, interfacial energy filtering is used to block carriers with low energies (cold carriers) and allow high-energy carriers (hot carriers) to pass through. For example, a conventional solid-state heterostructure TI cooler consists of three layers (cathode, barrier, and anode) in which the cathode-barrier junction presents a step potential that selectively transmits hot carriers. Cold carriers accumulate at this junction and reach equilibrium by removing heat from the lattice, thus realizing solid-state cooling. Hot carriers transmitted across the barrier come into equilibrium over the energy relaxation length (ERL), heating the lattice by phonon emission. For a group of carriers with a prescribed out-of-equilibrium energy distribution, the ERL measures how far they travel on

average before the quasi-equilibrium energy distribution is restored. For larger energy relaxation lengths, the barrier can be made thicker, increasing the thermal resistance through which this heat is transmitted back to the cold side and improving device efficiency.

Since ERL is strongly correlated with the carrier mean free path (MFP), the long MFP (~1.6 μm) of electrons in CNT as measured in ref. [20] indicates that it can be a promising material for TI refrigeration. A Monte Carlo simulation framework is developed to track the evolution of the energy exchange between electron and lattice (phonons) in real time for a CNT-based TI cooler and calculate the cooling capacity of the device.

1.4 Characterizing and mitigating thermal challenges at the package level

The rising thermal challenge in high-power electronic and optoelectronic devices provides a strong incentive to develop novel E&O packages and heatsinks, from the fundamental material level to packaging architectures. Chapter 5 presents a study that develops amorphous polymeric packaging materials with high thermal conductivities. Such materials form an integral part of thermal interface materials (TIMs) which are ubiquitous in VLSI packages, and hold promise as flexible substrates desirable for emerging technologies such as flexible displays. Specifically, this dissertation discusses a novel molecular design approach that differs from the traditional blending or mechanical stretching methods to enhance the thermal conductivity of an amorphous polymer to 1.17 $Wm^{-1}K^{-1}$, which is approximately 6 times that of typical amorphous polymers.

The polymer of choice is poly(acrylate acid) (PAA), which upon dissolving in basic aqueous solution ionizes to carry charges on the side group. The charge-charge repulsive forces between the side groups translate into the expansion and un-entanglement of the polymer backbone, reducing phonon scattering and improving thermal conductivity. In addition to a change in morphology, the positively-charged sodium ions bond with the negatively-charged polymer side groups, forming ionic interactions that are much stronger than Van der Waals interactions. Together they lead to a polymer thin film with a thermal conductivity exceeding 1.17 $Wm^{-1}K^{-1}$, which is one of the highest values reported for pure amorphous polymers. This chain-extension approach can potentially advance thermal management at the packaging level when it is geared toward increasing the thermal conductivity of the polymer matrix in TIMs. Additionally, it can be adapted to facilitate heat dissipation at the device level, if the insights of how κ and elastic modulus

correlate with chain morphology are applied to the molecular design of OLED materials for improved heat-spreading capabilities and better thermal stability (higher glass transition temperature T_g).

While thermally conductive polymers can enhance the performance of TIMs which are critical in the thermal management of devices with a steady-state heat source, electronics such as mobile phones are sometimes subjected to a transient heat source. Specifically, a large fraction of existing and emerging mobile applications, including speech recognition, visual recognition and web browsing, demands short bursts of intense computations separated by extended (at least tens of seconds) idle states waiting on user input. Therefore, the heat source is intense but lasts only a short amount of time (i.e. it is "bursty" in nature). To suppress the thermal transient caused by the bursty packets of heat, phase change material (PCM) with large thermal capacitance and fast thermal-response time is a promising addition to existing heatsink solutions.

Prior research by our group examined the thermal performance of a silicon thermal test chip (TTC) integrated with on-chip PCMs as a proxy for mobile chips. It was found that incorporating PCMs can help suppress peak and average chip temperatures when the TTC is fully activated for a short period of time. However, due to the use of a proxy chip that was unable to perform real computation tasks, the practical applicability of the measurement results was limited.

In Chapter 6, the effectiveness of a PCM as heat storage material is evaluated experimentally in a realistic state-of-the-art neural network accelerator test-chip (ATC) designed for deep learning. Embedding the PCM into the package substrate ensures zero volume added for this heat-sink solution, meeting the volume constraint of the wearable device that the ATC is designed for. By choosing a metallic phase change material comprising low melting point alloy having relatively high thermal conductivity and latent heat, it is found that the die peak temperature during sprint can be suppressed by ~17% for the chip with PCM compared to the one without.

Chapter 7 includes a summary of the presented thesis and a proposal of future work.

An overview of the presented work is summarized in Fig. 1.4.

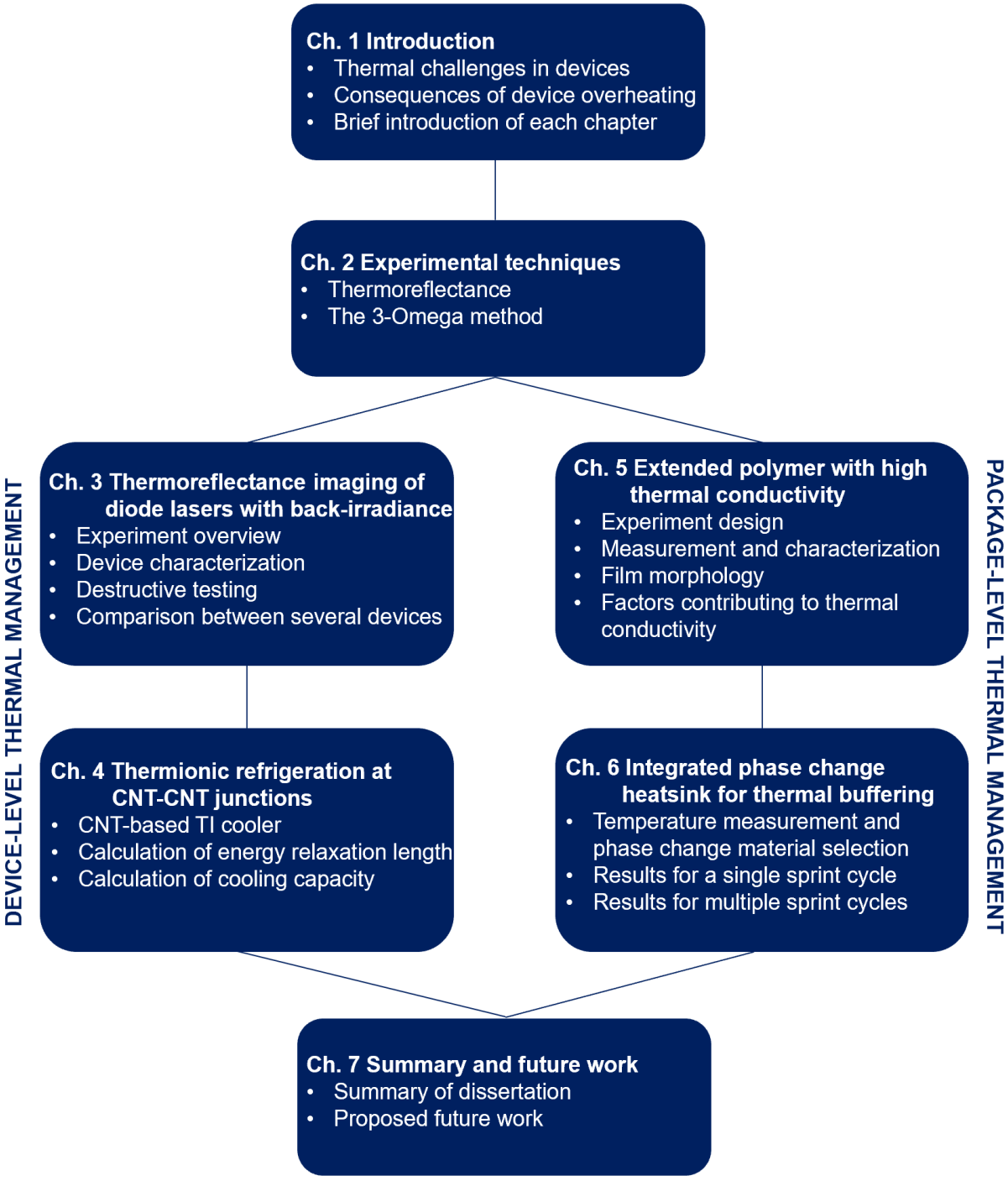

Figure 1.4 | An overview of the presented thesis.

References

1. J. D. Plumer, M. Deal, and P. D. Griffin, Silicon VLSI Technology: Fundamentals, Practice, and Modeling. Pearson, Prentice Hall (2000).

2. M. M. Waldrop. More than Moore. *Nature News* **530**, 145-147 (2016).

3. A. L. Moore and L. Shi. Emerging challenges and materials for thermal management of electronics. *Materials Today* **17**, 163-174 (2014).

4. A. Shanker, C. Li, G.-H. Kim, D. Gidley, K. P. Pipe, and J. Kim. High thermal conductivity in electrostatically engineered amorphous polymers. *Sci. Adv.* **3**, e1700342 (2017).

5. Z. Yao *et al.* High power and high reliability InGaAs broad area lasers emitting between 910 and 980 nm. *Lasers and Electro-Optics Society 2000 Annual Meeting. LEOS 2000. 13th Annual Meeting. IEEE* **2**, 510-511 (2000).

6. H. U. Pfeiffer, S. Arlt, M. Jacob, C. S. Harder, and I. D. Jung. Reliability of 980 nm pump lasers for submarine, long-haul terrestrial, and low-cost metro applications. *Optical Fiber Communication Conference* (2002).

7. V. Rossin, T. Strite, E. Zucker, and T. H. Ju. AlGaInAs broad-area lasers exhibit outstanding reliability. *Compound Semiconductor*, 43-46 (2003).

8. P. A. Crump *et al*. High-efficiency high-power 808-nm laser array and stacked arrays optimized for elevated temperature operation. *Proc. SPIE* **5336** (2004).

9. G.E. Moore. Cramming more components onto integrated circuits. *Electronics* **38**, 114-117 (1965).

10. S.-C. Lin and K. Banerjee. Cool chips: opportunities and implications for power and thermal management. *IEEE Transactions on Electron Devices* **55**, 245-255 (2008).

11. G. Chen. Size and interface effects on thermal conductivity of superlattices and periodic thin-film structures. *J. Heat Transfer* **119**, 220-229 (1997).

12. G. Chen. Thermal conductivity and ballistic-phonon transport in the cross-plane direction of superlattices. *Phys. Rev. B* **57**, 14958 (1998).

13. B. Poudel, Q. Hao, *et al*. High-thermoelectric performance of nanostructured bismuth antimony telluride bulk alloys. *Science* **320**, 634-638 (2008).

14. C. Jeong, S. Datta, and M. S. Lundstrom. Thermal conductivity of bulk and thin-film silicon: A Landauer approach. *J. Appl. Phys.* **111**, 093708 (2012).

15. W. Wondrak. Physical limits and lifetime limitations of semiconductor devices at high temperatures. *Microelectronics Reliability* **39**, 1113-1120 (1999).

16. H. Ye, M. Lin, and C. Nasaran. Failure modes and FEM analysis of power electronic packaging. *Finite Elements in Analysis and Design* **38**, 601-612 (2002).

17. J. P. Power, An introduction to fiber optics systems. Akesen Associates Inc. (1993).

18. M. Fukada. Laser and LED Reliability Update. *Journal of Lightwave Technology* **6**, 1488-1495 (1988).

19. H. Kissel, *et al*. A comprehensive reliability study of high-power 808 nm laser diodes mounted with AuSn and indium. *Proc. SPIE* **6876** (2008).

20. J. Y. Park, S. Rosenblatt, Y. Yaish, V. Sazonova, H. Üstünel, S. Braig, T. A. Arias, P. W. Brouwer, and P. L. McEuen. Electron-phonon scattering in metallic single-walled carbon nanotubes. *Nano Lett.* **4**, 517-520 (2004).

Chapter 2
Experimental Techniques for Microscale Temperature Measurements and Thermal Property Measurements

2.1 Thermoreflectance

2.1.1 Introduction and background

Thermoreflectance is a non-invasive, optical temperature measurement technique that uses the linear relation between a change in the normalized reflection coefficient ($\Delta R/R$) and a change in temperature[1-5] (ΔT). The dependence of reflectivity on temperature can be understood in terms of the temperature-dependent bandgap. Bandgap can be considered as the energy required to excite an electron from the valence band into the conduction band, or equivalently the energy needed to break a bonded electron free. In a semiconductor material, electrons are bounded by the potential well formed by nearby atoms. As temperature rises, lattice undergoes thermal expansion so that the atoms are further away from the electron, weakening the binding energy and therefore lowering the bandgap. The reduction in the electronic bandgap would then alter the refractive index (n) of the material, and hence the reflection coefficient.

For light at normal incidence with zero absorption, reflectivity is related to refractive index by

$$R = \frac{(n - 1)^2}{(n + 1)^2} \tag{2.1}$$

The relation between a change of the refractive index and the associating change in the bandgap is described by the Moss rule[6]

$$E_g n^4 = constant \tag{2.2}$$

Thermoreflectance techniques, when applied to study thermal phenomena in electronic devices and various materials, can usually be classified into two categories: a) Steady-state (DC) or quasi-steady-state thermoreflectance, and b) transient thermoreflectance. In the case of DC thermoreflectance, a charge-coupled device (CCD) is usually used as the light detector, and hence the name "CCD-thermoreflectance (CCD-TR)". In order to achieve a large bit depth that is sufficient to distinguish thermoreflectance signal from the noise in the acquired image (typically require a SNR of ~10^{-5} – 10^{-4}), the frame rate of the CCD camera is relatively low (<100 fps), limiting the time resolution of the CCD-TR technique to >0.01s. However, the CCD pixel array can image a large field of view in parallel (as opposed to scanning a single point across a surface), facilitating the acquisition of large-area 2D temperature maps[4.] CCD-TR has been applied to study numerous thermal phenomena in electronic and photonic devices, including profiling the hot-spots in integrated circuits[7], quantifying the thermal lensing effect at high bias in high power diode lasers[8], and measuring the thermal relaxation time constant of the active region in a quantum dot laser[9] (Fig. 2.1).

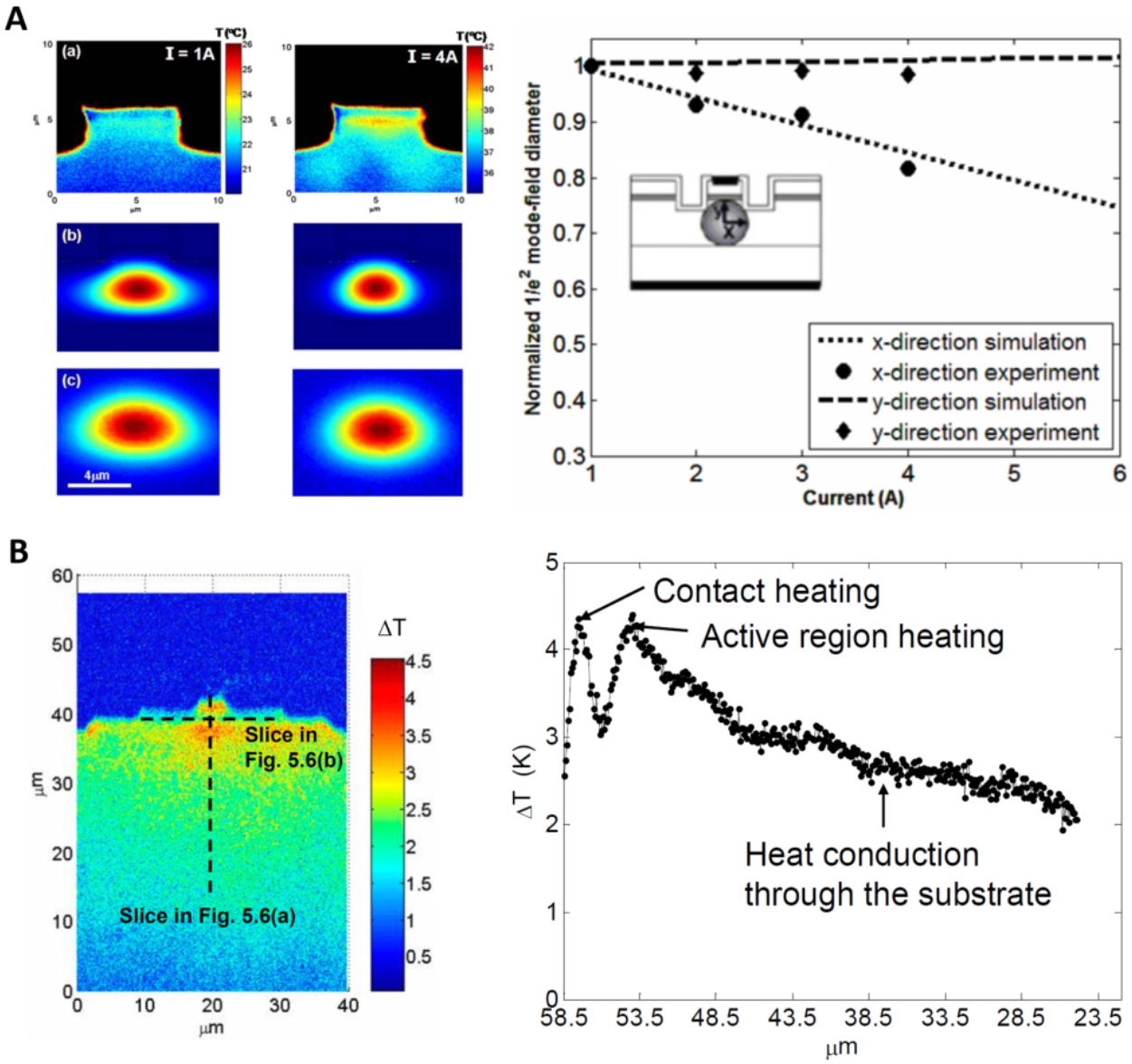

Figure 2.1 | Previous application of CCD thermoreflectance on thermal effects in electronic and photonic devices. (A) The 2D temperature profiles and optical modes for a slab-coupled optical waveguide laser (left panel) and the normalized diameter of the optical mode as a function of current (right panel). Image courtesy of ref. [8] **(B)** Imaged temperature profile of a quantum dot laser (left panel) and vertical line-out of the 2D temperature map, averaged in the horizontal direction (right panel). Image courtesy of ref. [9]

Transient thermoreflectance, on the other hand, is usually implemented as a pump-probe technique where the pump light induces a temperature excursion that leads to a change in the reflection coefficient, which is measured by a change in the probe light intensity after it's reflected off of the sample and collected by a photodetector[10]. By adopting a mechanical delay stage or an asynchronous technique, the thermal transient of the material can be probed with nanosecond resolution (Fig. 2.2A). The disadvantage of this technique is that it is typically a single point measurement, acquiring a large-area 2D temperature map requires scanning the pump-probe spot across a material surface which can be time-consuming. Two of the widely adopted applications of this technique are time-domain thermoreflectance (TDTR) and frequency-domain thermoreflectance (FDTR) which can be used to measure the thermal properties (including but not limiting to thermal conductivity, thermal diffusivity, specific heat, etc.) of nanometer thin-film samples[11], as well as the phonon mean free path spectra of materials[12,13] (Fig. 2.2B).

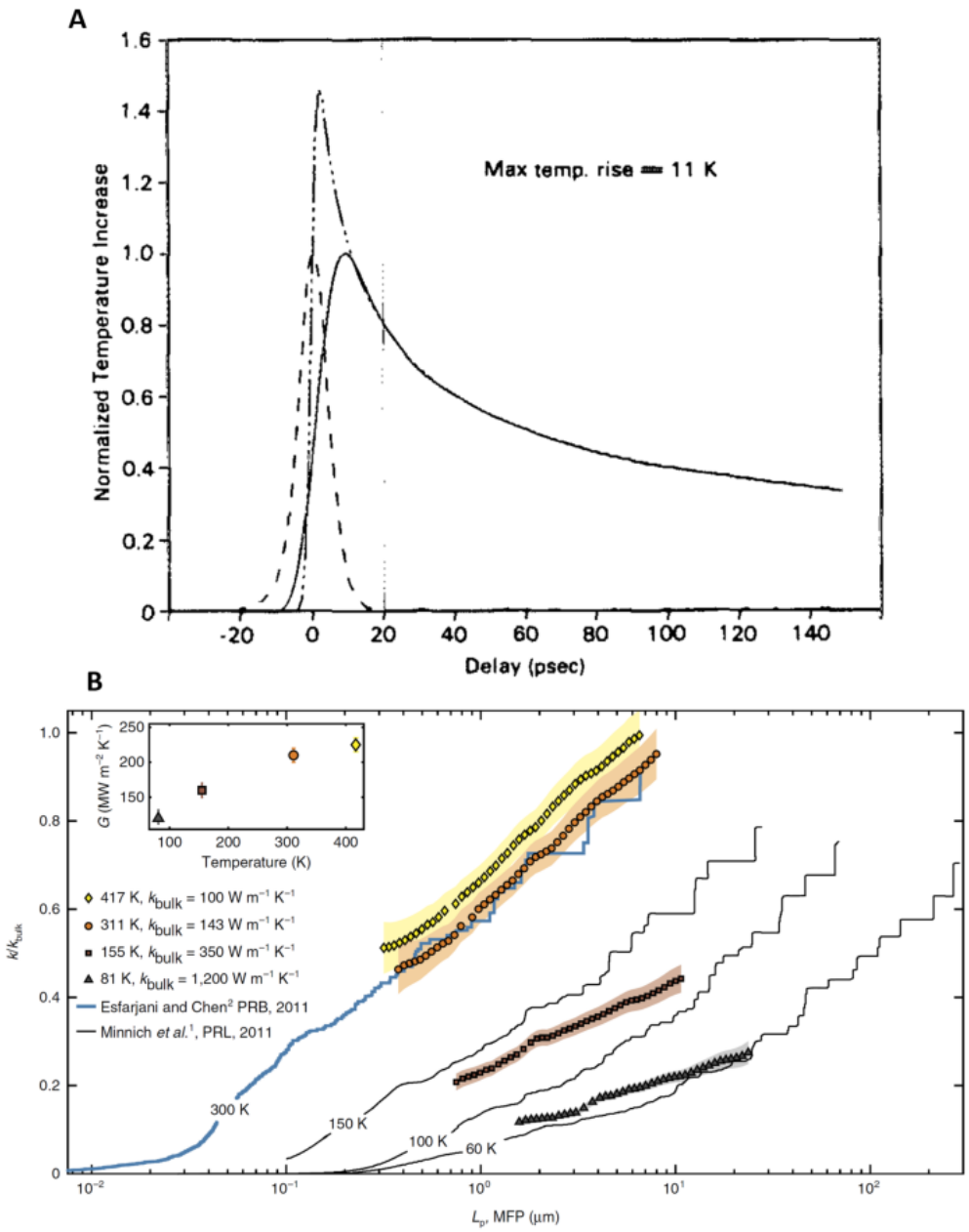

Figure 2.2 | Application of the transient pump-probe technique in studying micro-to-nanoscale thermal phenomena. (A) TDTR measurement and fitted transient profile for single crystal nickel. The solid line is the normalized change in reflectivity measured experimentally. By fitting this curve with a thermal model, thermal properties of the nickel film can be extracted. Image courtesy of ref. [10] **(B)** Temperature-dependent phonon mean free path of crystal silicon, measured with a FDTR technique. The y-axis represents the fraction of bulk thermal conductivity that is contributed by phonons with MFP up to a certain threshold value, which is represented by what is on the x-axis Image courtesy of ref. [13].

2.1.2 Thermoreflectance equation and thermoreflectance coefficient

Within a small temperature range, the change in reflectivity for semiconductor materials is linearly correlated with a change in temperature ΔT:

$$\Delta T = k^{-1} \frac{\Delta R}{R} \tag{2.3}$$

Where $\Delta R/R$ is the normalized variation in reflectivity and k is a constant known as the thermoreflectance coefficient. Since the temperature-induced change in band structure varies with materials and the refractive index is wavelength-dependent, the value of k is typically material- and wavelength-dependent. It is also reported to be dependent on the optics used in the thermoreflectance setup (e.g., the numerical aperture of the objective used for magnification). For these reasons, k is usually calibrated independently in a separate *in-situ* experiment.

2.1.3 Working principles of CCD-based thermoreflectance

This section discusses a "four-bucket" CCD-TR imaging scheme used to calculate $\Delta R/R$ based on the TR images acquired during experiments. Firstly, a current oscillating at a frequency $f_{\text{Dev.}}$ is sent to power the device (Fig. 2.3). A signal with a frequency f_{CCD} that is four times $f_{\text{Dev.}}$ is used to trigger the CCD such that each pixel starts integrating at the rising edge of the trigger signal, giving rise to four CCD images taken per device on-off cycle (two each for "on" and "off"). The light-emitting diode (LED) used to illuminate the device surface operates at constant power, keeping the optical flux (Φ) constant.

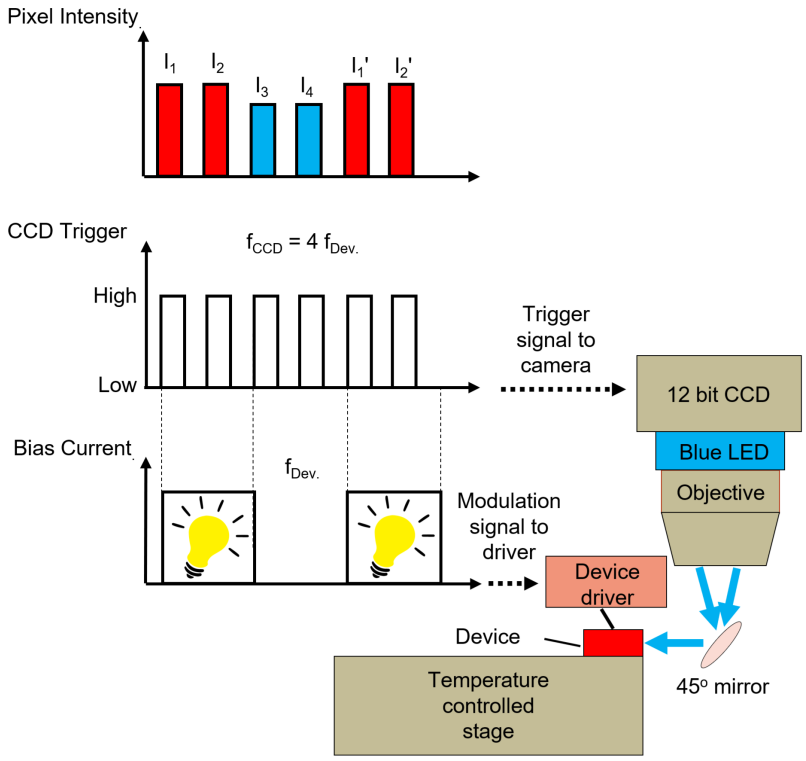

Figure 2.3 | Schematic of the CCD-thermoreflectance setup. During experiment, an oscillating bias current is sent to power the device, and a CCD trigger signal phase-locked to the bias signal is used to control the timing of pixel integration. The device under test is placed on a feedback-controlled temperature stage to make sure its back-side temperature remains stable throughout the experiment.

The bias current is usually sinusoidal or square-shaped. In the following derivation, a square-wave signal will be assumed since it is used for most TR experiment described in this dissertation. The variation in surface reflectivity due to device heating that is captured by the CCD camera follows a generic shape as shown in Fig. 2.4A. Different thermal time constants exist for the initial temperature ramp-up (when the device is turned on) and the cooling-off when bias is set to zero. As the device tested has a thermal time constant of <50μs in the critical region (i.e., the active region of the diode laser), which is much smaller than the integration time of the CCD camera (24.5ms), we can safety approximate the temperature variation as a square wave (Fig. 2.4B).

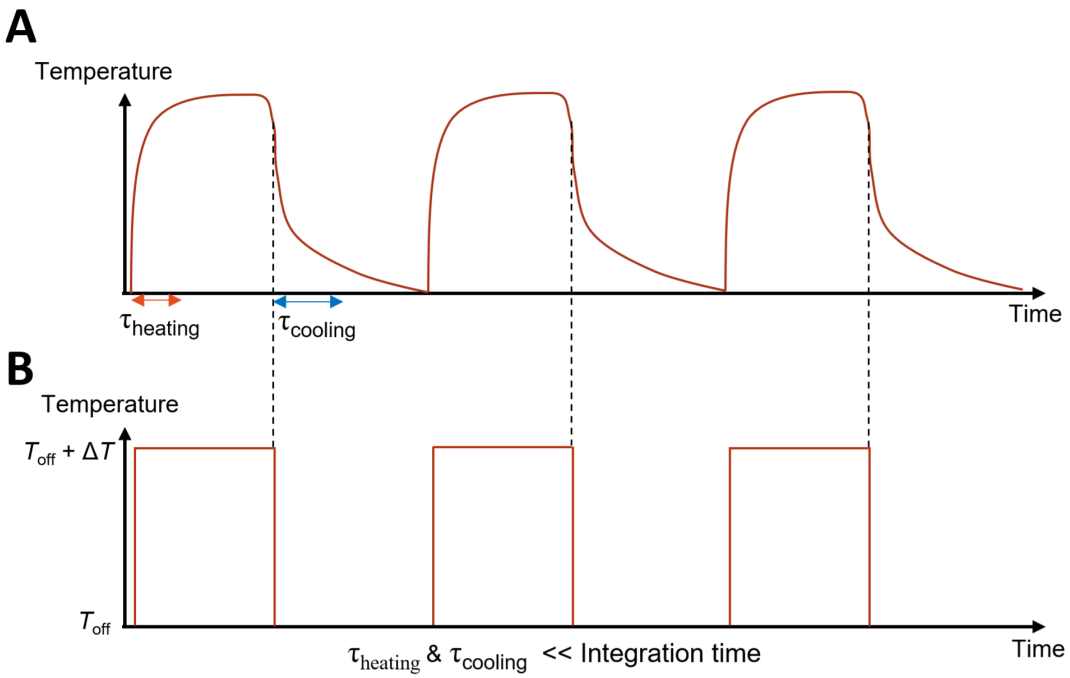

Figure 2.4 | Temperature variation as a function of time for device subject to a square-wave bias. (A) The general shape of the temperature variation, comprising a heating and a cooling transient. **(B)** This is the temperature evolution for the case where the device has a very short transient compared to the integration time such that the profile can be approximated as square wave.

The reflectivity of the device surface, under the assumption of a square-wave bias current and a square-wave temperature variation, can be written as

$$R = R_0\,(x, y, t) \;+\; \Delta R\,(x, y, t)\ \mathrm{Sqr}\,(t) \tag{2.4}$$

Where

$$\mathrm{Sqr}\,(t) = \begin{cases} 1 \;\; \text{where } t \;\in \left(\dfrac{T_{\mathrm{Dev.}}}{2}n,\ \dfrac{T_{\mathrm{Dev.}}}{2}(n{+}1)\right),\ n = 0, 2, 4, 6... \\ 0 \;\; \text{where } t \;\in \left(\dfrac{T_{\mathrm{Dev.}}}{2}m,\ \dfrac{T_{\mathrm{Dev.}}}{2}(m{+}1)\right),\ m = 1, 3, 5, 7... \end{cases} \tag{2.5}$$

This is a square wave representing the temperature variation where the heating cycle (device on) precedes the cooling cycle (device off). $T_{\mathrm{Dev.}}$ is the device operating period ($T_{\mathrm{Dev.}} = 1/f_{\mathrm{Dev.}}$). The reflectivity is equal to $R + \Delta R$ and R when the device is on and off, respectively. In general, there is a time delay between the device bias signal and the pixel detecting the temperature variation due to thermal relaxation time of the material, which varies based material thermal conductivity, heat capacity, and density. In our experiment, it is very small compared to the integration time, so it's set to zero.

Multiplying the optical flux afforded by the LED with R gives the signal (S) seen by the CCD camera

$$S\,(x, y, t) \;= \Phi R_0\,(x, y, t) \;+\; \Phi\Delta R\,(x, y, t)\ \mathrm{Sqr}\,(t) \tag{2.6}$$

The intensity of the 4 images captured by the CCD can then be calculated as the reflected light see by each pixel integrated over approximately a quarter of the device bias cycle, given as

$$I_1 = I_2 = C\frac{T_{\text{Dev.}}}{4}\ [\Phi R_0\ (x, y, t)\ +\ \Phi\Delta R\ (x, y, t)] \tag{2.7a}$$

$$I_3 = I_4 = C\frac{T_{\text{Dev.}}}{4}\ \Phi R_0\ (x, y, t) \tag{2.7b}$$

where C is a constant that accounts for the transmission losses (e.g., due to optics in the imaging pathway) before the light is collected by the detector pixel, and the quantum efficiency of the CCD detector.

Combining Eqn. 2.3 to 2.6, the temperature variation between the device on and off states can be expressed as

$$\Delta T = k^{-1}\frac{\Delta R}{R + \Delta R}\ \approx k^{-1}\frac{\Delta R}{R} = k^{-1}\ \frac{I_1 + I_2 - I_3 - I_4}{I_3 + I_4} \tag{2.8}$$

During the experiment, many heating/cooling cycles (>4000) are recorded and averaged in order to mitigate the error caused by long-term fluctuations in the LED intensity, as well as read-out noise in the CCD camera.

2.1.4 Calibration of the thermoreflectance coefficient

As mentioned in Section 2.1.2, calibration of the thermoreflectance coefficient needs to be done *in situ* due to its dependence on material, wavelength, and optical setup. In general, there are two approaches for carrying out the calibration experiment[4.] In the first approach, a micro-thermocouple (~25μm in diameter at the tip) is used to measure the temperature variation (ΔT) of

the device during a normal thermoreflectance experiment. k can be derived from the collected thermoreflectance signal $\Delta R/R$ and the measured ΔT. While this method is convenient, it works only when the device of interest is much bigger than the micro-thermocouple, which is not the case when it comes to thin epitaxial layers in a lot of semiconductor devices (LEDs, FETs, diodes lasers, etc.). Therefore, in most cases a second approach is used where the device is placed on a broad-area thermoelectric cooler (TEC), forming a good thermal contact with the help of thermal grease and/or thermal adhesive. The TEC is biased by a square-wave current having low frequency (~0.25 Hz), which modulates the temperature of the entire device while acquiring thermoreflectance images. Since the temperature of the device is modulated uniformly, as verified by thermocouple measurements at various locations on the device, the thin-layers of interest would assume a similar ΔT. By selecting a high-power TEC, a peak-to-peak ΔT of ~5K can be achieved. In addition, in general k is temperature-dependent, which should be taken into account when converting $\Delta R/R$ to ΔT because the temperature of the device also oscillates during experiments. Hence, it is critical to measure k at a range of temperatures. The diode laser characterized using CCD-TR have several thin epitaxial layers near the hotspot in the device (i.e. quantum well near the facet): the substrate (undoped GaAs), waveguide, and cladding. Fig. 2.5 depicts the calibrated k for these materials up to 65 ^{o}C, which is approximately the maximum temperature the device reaches during experiment. No significant temperature dependence has been found in this temperature range.

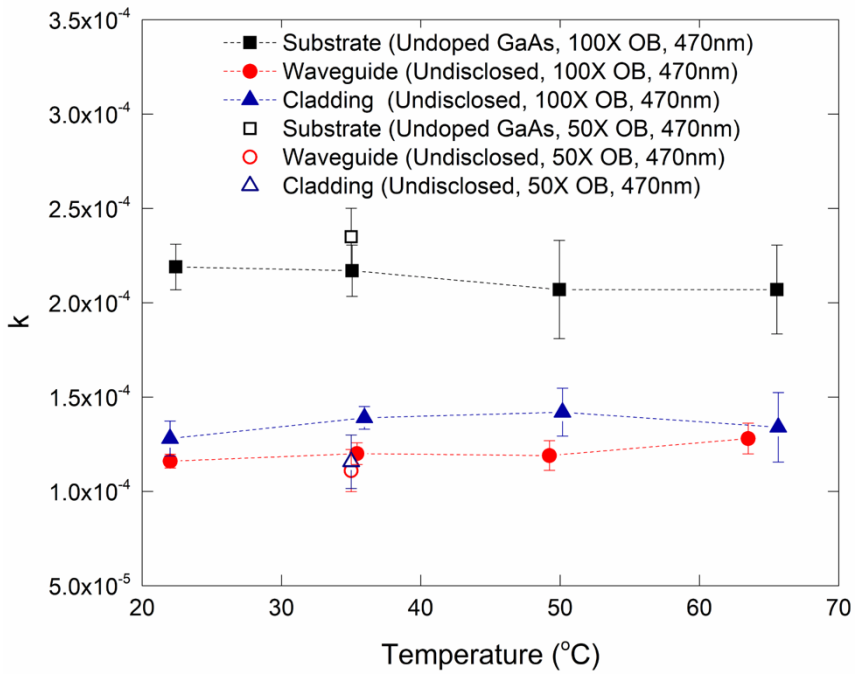

Figure 2.5 | Calibration of the thermoreflectance coefficients. The TRC of the substrate, waveguide and cladding layers of the diode laser tested in this dissertation have been measured from 22°C to 65°C. No significant temperature dependence has been found beyond measurement error**.** The legend entails (in order from left to right) material composition, magnification of the objective ("OB") used in the calibration, and illumination wavelength.

2.2 Differential 3ω technique for thermal conductivity measurements

The 3ω method has been a useful technique for determining the thermal properties of bulk and small structures. It is particularly suitable for measuring micro-to-nanoscale samples where the dimensions of thermistors and thermocouples could severely perturb the temperature field of the sample. The first application of the 3ω method is reported by Corbino in 1910 [14], and the first practical application of the 3ω method in measuring material thermal conductivities (κ) was pioneered by Cahill and Pohl in 1987 [15].

In a standard implementation of the 3ω technique[15], a sinusoidal current with frequency ω (I_ω) is supplied to a thin metal line deposited on the material surface (Fig. 2.6), leading to a Joule heating with a frequency of 2ω ($Q_{2\omega} = I_\omega{}^2 r$, where r is the resistance of the metal line) and hence and a temperature rise in the heater line (ΔT). Since r is linearly proportional to temperature within a small temperature range, there would be a small 2ω AC signal in r that is caused by the 2ω Joule

heating (i.e., $r = r_{DC} + \Delta r_{2\omega}$). Hence, a voltage signal that oscillates with a frequency of 3ω arises ($V_{3\omega} = I_{\omega}\Delta r_{2\omega}$).

The 3rd harmonic voltage is related to a change in metal line temperature through

$$\Delta T = 2T_{CR}^{-1}\frac{V_{3\omega}}{V_{\omega}} \tag{2.9}$$

Where T_{CR} is the heater line's temperature coefficient of resistance.

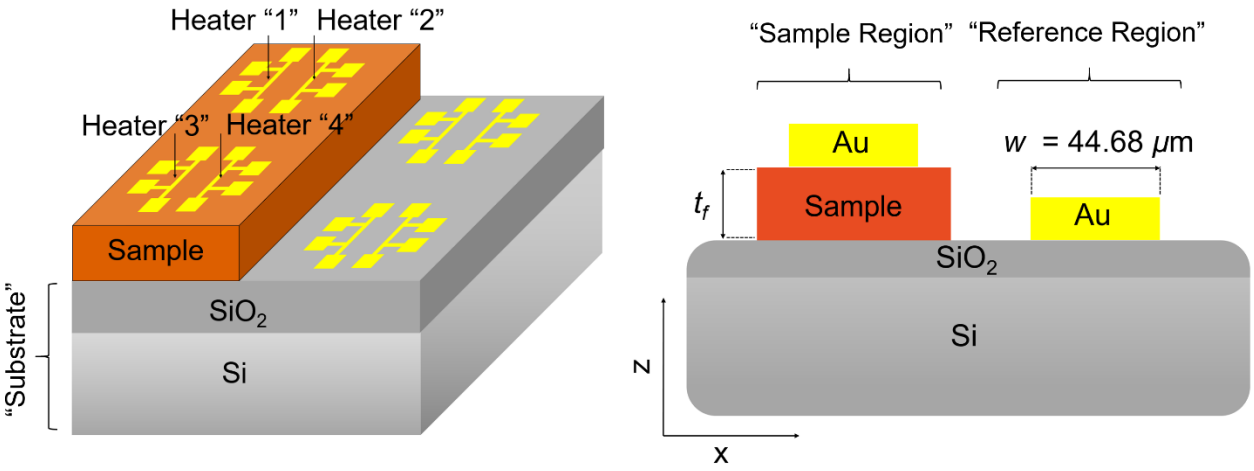

Figure 2.6 | Configuration for the differential 3ω measurement used to measure cross-plane thermal conductivity (z-direction) of the sample. Wide gold heater lines (~45 μm) are deposited onto the "sample region" with the thin film of interest on top of the substrate, and the "reference region" with a bare substrate. Multiple heater lines are deposited to facilitate cross-checking and comparing measurement consistency. The substrate used is a doped Si covered by ~100 nm SiO_2 for electrically insulating the conductive Si from the metallic heater line. Doped silicon substrate rather than undoped is used because it is more cost-effective.

In this dissertation a variation of the 3ω technique, called "the differential 3ω technique", is used to measure the cross-plane thermal conductivity of thin film samples[16]. In this method, the substrate usually has a region covered by the sample of interest ("sample region") and a region without the sample[16] ("reference region"). Identical thin metallic heater lines with a known width (w) are deposited onto both regions.

By supplying the same power input (Q) to the identical heater lines in the sample and reference regions, the additional temperature rise in the "sample region" due to the presence of the thin film can be calculated by

$$\Delta T_f = \Delta T_s - \Delta T_r = 2T_{CR}^{-1}\left(\frac{V_{3\omega,s}}{V_{\omega,s}} - \frac{V_{3\omega,r}}{V_{\omega,r}}\right) \tag{2.10}$$

The subscript "s" and "r" represents "sample region" and "reference region", respectively.

For sample with thickness (t_f) much smaller than w, heat transfer through the sample is quasi 1-dimensional, and the temperature rise across it can be calculated using 1D Fourier's law.

$$\Delta T_f = \frac{Qt_f}{wl\kappa} \tag{2.11}$$

Where l is the length of the heater line. In our experiment l = 2.6 mm.

Combing Eqn. 2.6 and 2.7, we arrive at an expression for calculating the cross-plane thermal conductivity of the isotropic thin film from the measured 1ω and 3ω voltages in the sample region and the reference region:

$$\kappa = [2T_{\mathrm{CR}}^{-1}(\frac{V_{3\omega,s}}{V_{\omega,s}} - \frac{V_{3\omega,r}}{V_{\omega,r}})]^{-1}\ \frac{Q\,t_f}{l\,w} \tag{2.12}$$

The differential 3ω method is relatively insensitive to the substrate thermal conductivity and contributions from other layers as well as interfacial thermal resistances, as long as they are present both in the sample region and the reference region.

References

1. D. Guidotti and H. M. Driel. Spatially resolved defect mapping in semiconductors using laser-modulated thermoreflectance. *Applied Physics Letters* **47**, 1336-1338 (1985).

2. P. W. Epperlein. Micro-temperature measurements on semiconductor laser mirrors by reflectance modulation: A newly developed technique for laser characterization. *Japanese Journal of Applied Physics* **32**, 5514-5522 (1993).

3. S. Dilhaire, S. Jorez, L.-D. Patiño-Lopez, W. Claeys, and E. Schaub. Laser diode light efficiency determination by thermoreflectance microscopy. *Microelectronics Journal* **32**, 899-901 (2001).

4. J. Christofferson and A. Shakouri. Thermoreflectance based thermal microscope. *Review of Scientific Instruments* **76** (2005).

5. M. Farzaneh, *et al*. CCD-based thermoreflectance microscopy: principles and applications. *Journal of Physics D: Applied Physics* **42**, 143001 (2009).

6. J. I. Pankove, Optical Processes in Semiconductor. Dover, New York (1975).

7. G. Tessier, *et al*. Thermoreflectance temperature imaging of integrated circuits: calibration technique and quantitative comparison with integrated sensors and simulations. *Journal of Physics D: Applied Physics* **39**, 4159 (2009).

8. P. K. L. Chan, *et al*. Temperature mapping and thermal lensing in large-mode, high-power laser diodes. *Appl. Phys. Lett.* **89**, 201110 (2006).

9. P. K. L. Chan, *et al*. Thermal relaxation time and heat distribution in pulsed InGaAs quantum dot lasers. *Appl. Phys. Lett.* **89**, 01110 (2006).

10. C. A. Paddock and G. L. Eesley. Transient thermoreflectance from thin metal films. *Journal of Applied Physics* **60**, 285 (1986).

11. D. G. Cahill. Analysis of heat flow in layered structures for time-domain thermoreflectance. *Review of Scientific Instruments* **75**, 55119 (2004).

12. A. J. Minnich, *et al*. Thermal conductivity spectroscopy technique to measure phonon mean free path. *Phys. Rev. Lett.* **107**, 095901 (2011).

13. K. T. Regner, *et al*. Broadband phonon mean free path contributions to thermal conductivity measured using frequency domain thermoreflectance. *Nat. Comm.* **4**, 1640 (2013).

14. O. Corbino. Thermal oscillations in lamps of thin fibers with alternating current flowing through them and the resulting effect on the rectifier as a result of presence of even-numbered harmonics. *Physikalische Zeitschrift* **11**, 413417 (1910).

15. D. G. Cahill and R. O. Pohl. Thermal conductivity of amorphous solids above the plateau. *Phys. Rev. B* **35**, 4067 (1987)

16. T. Borca-Tasciuc, A. R. Kumar, and G. Chen. Data reduction in 3ω method for thin-film thermal conductivity determination. *Review of Scientific Instruments* **72**, 2139 (2001)

Chapter 3

Thermal Characterization at the Device Level – Thermoreflectance Imaging and Failure Testing of Diode Lasers with Back-Irradiance

IEEE Journal of Quantum Electronics **54**, 2001213 (2018)

Applied Physics Letter **112**, 101101 (2018)

3.1 Introduction

As mentioned in Chapter 1, the active region of semiconductor lasers represents one of the smallest hotspots in optoelectronic devices, with a characteristic dimension on the order of ~1μm. The active region is simultaneously the most critical region in a laser for light generation and typically the hottest. Overheating in the active region leads to defects, which are absorbing centers and generates more heat due to non-radiative recombination. This localized excess heat then causes more defects to form, completing the positive feedback loop. Eventually this gives rise to a thermal run-away process that leads to catastrophic optical damage (COD) where the semiconductor would experience extremely high local temperature and begin to melt, at which point the device is irreversibly destroyed. While advancements in epitaxial designs, diode chip architecture, novel cooling approaches, and facet passivation technologies have led to rapid increase in the maximum output power which can be produced by diode pumps[1-14], the maximum output power of a diode laser is still primarily limited by catastrophic optical damage (COD), a topic which has been studied in tremendous detail over the past twenty years[15-21]. Commercial diode suppliers have developed numerous COD mitigation strategies which have enabled continuous improvement in the rated operating power while maintaining excellent reliability in isolated testing[3,7,18]. However, there have been few reports on the reliability of high power diode lasers deployed in real diode-pumped laser systems.

High power semiconductor lasers deployed in real laser systems are subject to different conditions than those present during isolated testing[2]. One such difference is the presence of back-irradiance

of optical emission. There are numerous sources of back-irradiance in diode pumped laser systems[22], including: 1) back-scatter and reflection from micro-optics, 2) residual specular reflection from antireflection-coated optical flats, 3) amplified spontaneous emission from the laser gain medium, 4) imperfect optical isolation leading to pulsed or continuous wave laser emission leak-through, and 5) unabsorbed pump light in double-pass and double side- or end-pumped laser systems. Back-irradiance is known to alter diode behavior in several ways including a broadening of the spectral linewidth, greater fluctuations in the outcoupled optical power, a reduction in the threshold current, a shift in the emission wavelength[22-25] and so on.

An important consequence of optical feedback is accelerated device degradation leading to premature failure of high-power diode lasers[22,25-28]. This reliability hazard of laser back-irradiance is of great concern in the design of virtually all diode-pumped laser systems. Despite this, no framework has yet been established to quantify the maximum permissible back-irradiance for reliable laser operation.

3.2 Experiment overview

In this work, the effect of optical feedback on laser diodes is investigated. The study is limited to devices with an emission wavelength around 800 nm (where the GaAs substrate is highly absorbing) that operate in continuous-wave (CW) mode, however the approach and framework established is expected to applicable to diode lasers at all wavelengths and operating conditions. Nondestructive thermoreflectance testing and thermal modelling are employed to investigate how optical feedback would impact the temperature profile of the laser facet. Destructive testing is carried out to bound the activation energy of back-irradiance-induced device failure, shedding light on whether COD in laser diode with back-irradiance is primarily grounded in thermal, similar to the case for free-running devices (i.e. without back-irradiance). Finally, a framework is established to describe the impact of optical feedback on the reliability and lifetime of diode lasers. It is expected that the development of such a tool would be of great value to the diode-pumped laser community, enabling diode manufactures to better understand how their products will behave in integrated laser systems and allow solid state laser developers to improve their designs to mitigate such failures.

The particular diode laser structure studied herein is a quantum well broad-area diode laser

emitting TM-polarized light around 800 nm and is epitaxially grown by molecular beam epitaxy (MBE) on a GaAs substrate. The design is based on the AlInGaAsP material system and with tensile-strained quantum wells. The quantum well is surrounded by an optical waveguide approximately 1 μm thick which itself is surrounded by doped cladding layers of appropriate thickness. In the fast axis (growth) direction, the design permits laser oscillation in a stable fundamental mode while in the slow axis, the ~200 μm wide broad area stripe operates highly multimode. The diode laser follows wafer fabrication processes: current injection apertures are lithographically defined by via openings in a deposited dielectric layer. Ohmic contacts are deposited as appropriate. Laser bars are cleaved to a cavity length of 1500 μm prior to facet coating. Bars and single emitters are both bonded junctions down using AuSn (hard) solder to expansion-matched heatsinks.

The experimental testing approach for the work herein proceeds as follows. The output from the laser facet is collimated in both the fast and slow axis. This collimated beam is subsequently relayed to a feedback mirror of varying reflectance. This mirror redirects the beam back to the facet of the diode laser while allowing its position with respect to the emitted beam to be controlled. This permits the general study of both the overall intensity and location of back-irradiance on the operation and reliability of the diode pump source. A schematic illustration is shown in Fig. 3.1.

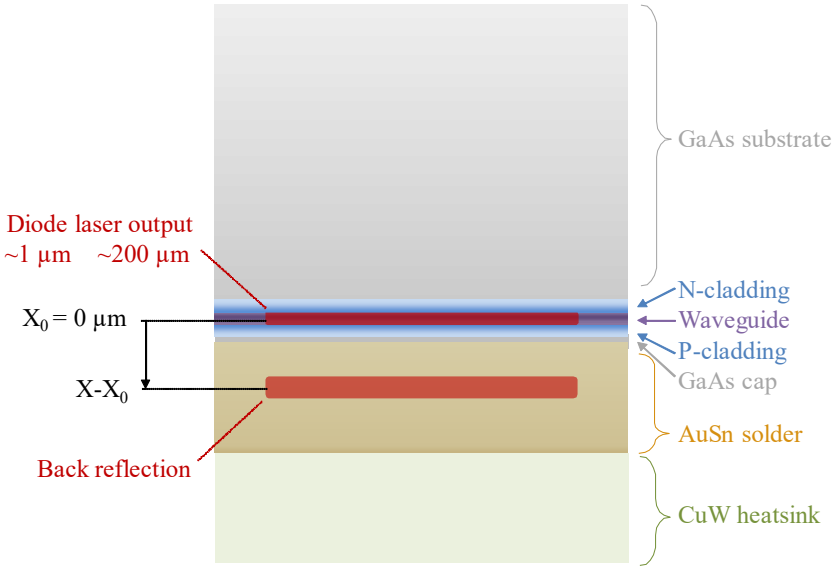

Figure 3.1 | Cross-section (front view) schematic of the laser diode facet. External feedback is applied to reflect a controllable portion of the laser diode output back at the facet. The return spot position in the fast axis is controlled through rotation of feedback mirror. The offset of the return beam from perfect boresight alignment is denoted $X–X_0$, with a negative value indicating pointing of the return beam toward the solder/heatsink and a positive value indicating pointing toward the substrate.

3.3 Device characterization and thermoreflectance imaging

Charge-coupled device (CCD)-based thermoreflectance microscopy technique was employed to obtain thermal images of the laser diode facet with high spatial (~ 550 nm) and temperature (~400 mK) resolution[29-30]. The high spatial resolution of this microscopy technique[31-37] enabled precise quantification of temperature profile across the laser diode facet, with the ability to accurately discern the back-irradiance spot location. A detailed schematic illustration of the setup is presented in Figure 3.2.

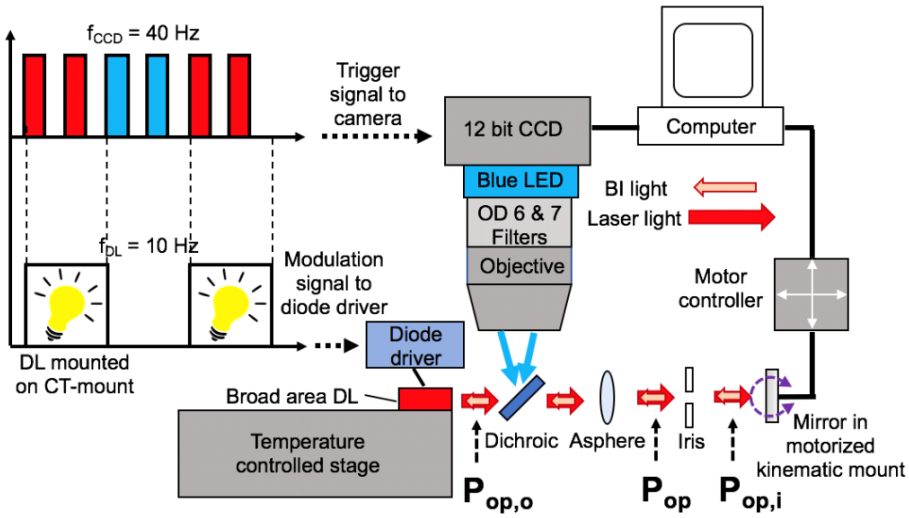

Figure 3.2 | Schematic illustration of the thermoreflectance setup for nondestructive assessment of the temperature profile arising at the laser diode facet due to back-irradiance.

To obtain a controllable back-irradiance spot on the facet, the emitted beam was split into two paths using a dichroic beamsplitter. In the back-irradiance pathway, the beam was reflected by a dielectric mirror after undergoing collimation through an aspheric lens. The mirror was affixed to a motorized kinematic mount which enabled high precision dual-axis tilt control. This allowed for a positioning accuracy of ~1 μm for the back-irradiance spot onto the laser diode facet. The back-irradiance reflectance level (BI R_{eff}) was controlled by adjusting the aperture of an iris inserted between the aspheric lens and mirror, as well as using dielectric mirrors with various reflection coefficients. A back-irradiance spot with a full width at half maximum (FWHM) equal to 9 μm was achieved; this relatively large value is caused by aberrations introduced in the back-irradiance pathway due to the dichroic beamsplitter and the aspheric lens. The imaging pathway consisted of standard thermoreflectance microscope with a 50X objective lens and a 470nm LED for illumination of the laser diode facet. The LED light is first reflected by the dichroic, which is placed at 45˚ with respect to the microscope axis, and then reflected off of the laser diode facet. The reflected light is focused by the 50X objective onto the CCD camera after filtering by two optical filters (with optical density of 6 and 7, respectively) which prevent the 800nm laser light from reaching the camera sensor and contaminating the measurement.

The thermoreflectance setup described above was used to characterize the effect of back-irradiance on the 2D facet temperature profile of the 800 nm diode laser. The diode laser was fabricated on a

GaAs substrate and soldered onto a CT-mount heatsink. The basic device characteristics such as out-coupling optical power ($P_{op,o}$), electrical-to-optical conversion efficiency (E/O efficiency) and voltage (V) as a function of drive current (I) have been plotted in Figure 3.3, with the inset showing

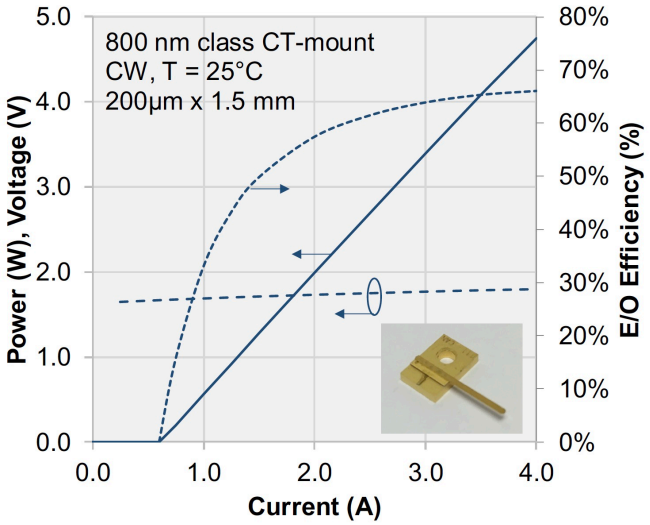

Fig. 3.3 | Power, efficiency, and voltage as a function of drive current for the high-power single emitter device used in the thermoreflectance testing. Inset depicts a photograph of the device, attached to a CT-mount style heatsink.

a photograph of the device tested. As can be seen from this plot, the threshold current for the laser is ~700 mA, and the laser operates >4.5 W output power and >65% E/O efficiency at a drive current of 4.0 Amps. In order to determine the thermal resistance of the diode, the relationship between the waste heat generated in the system and the shift in emission wavelength was analyzed. The waste heat (calculated as the difference between the input power VI and output power P_{op}) has been plotted against the rise in average junction temperature (inferred from the linear shift in centroid wavelength shift with temperature, 0.26 nm/K) in Figure 3.4, showing a well-behaved linear relationship with a slope of 11.1 K/W.

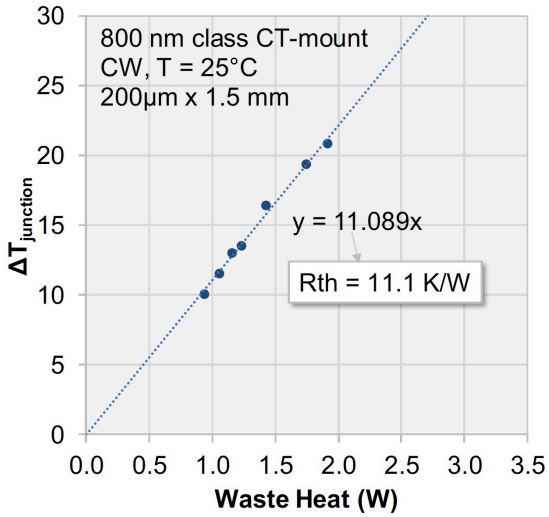

Figure 3.4 | Average junction temperature rise (calculated from the emission wavelength shift) as a function of waste heat (calculated from the net difference between the input electrical power and output optical power) for the CT-mount 800 nm class laser diode used in the thermoreflectance testing.

In order to quantify the facet temperature profile with respect to back-irradiance spot position, the spot was swept in the fast axis direction, starting from deep within the solder, through the active region to deep within the substrate. Figure 3.5A depicts the variation in emission spectrum of the laser diode for a range of back-irradiance spot positions. It is evident that a significant red shift in

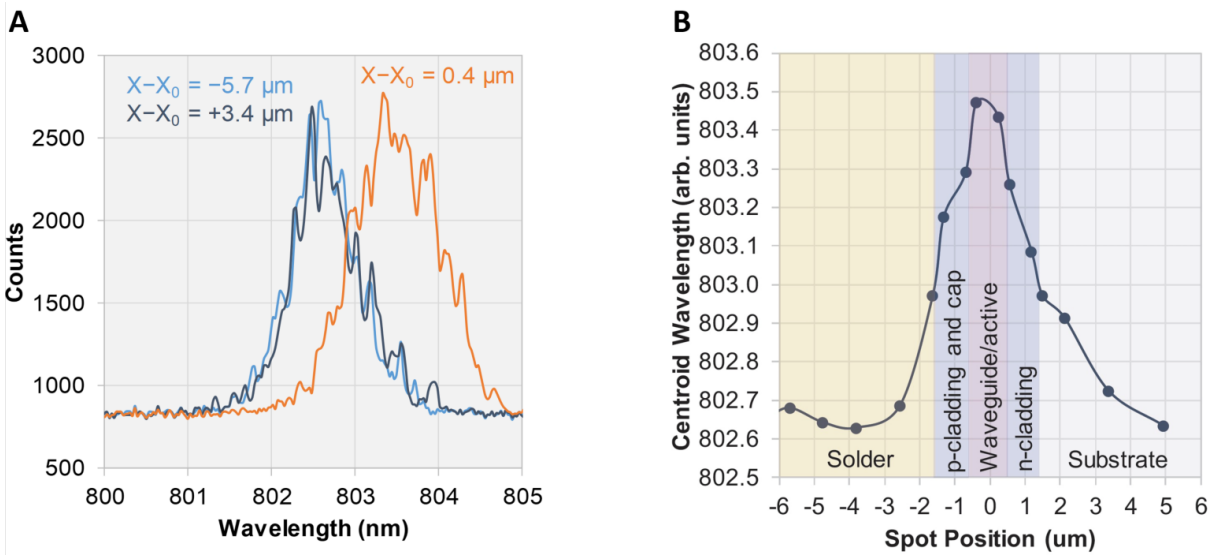

Figure 3.5 | Effect of back-irradiance on emission spectrum. (A) Laser emission spectrum obtained for several reflected beam spot offset values. A significant red shift in the emission spectrum is obtained when the reflected beam is aligned to the laser diode active region. **(B)** Calculated centroid emission wavelength as a function of reflected beam offset. A maximum red-shift in the emission wavelength centroid is used to identify the location of boresight alignment of the reflected beam with the diode laser active region.

emission spectrum is observed when the back-irradiance spot is in the vicinity of the active region. The physical origin of the red shift at boresight alignment is attributed to threshold reduction caused by resonant feedback to the diode cavity. The reduction in the threshold current and hence threshold carrier density results in pinning of the quasi Fermi levels at a lower energy. This causes the peak of the carrier distribution to occur at a lower energy level, resulting in laser emission at a longer wavelength. Figure 3.5B depicts the variation of centroid emission wavelength as a function of back-irradiance spot position. For reference, the plot is overlaid with shaded regions representing the solder, p-side cladding and cap layers, waveguide, n-side cladding, and substrate regions. The spot position which leads to a maximum in the plot corresponds to boresight alignment of back-irradiance spot with the active region and is set as the $X_0 = 0$ position[22].

Characterization of temperature profiles using the thermoreflectance technique involves recording the normalized linear change in reflectance of a material $\left(\frac{\Delta R}{R}\right)$, caused by a temperature change (ΔT) which are related by Eq. 3.1.

$$\Delta T = k^{-1} \frac{\Delta R}{R} \tag{3.1}$$

Here, k is the thermoreflectance coefficient whose value is dictated by the material composition, the wavelength of the illumination light and the numerical aperture of the objective used in the setup[35-37]. In our thermoreflectance system, the laser diode was driven with a square-wave bias current at a frequency of 10 Hz. For every on/off cycle of the laser diode, the CCD camera took four images (two for on-state and two for off-state), each integrated over 25% of the complete cycle time. Hence, the camera was triggered by a square wave of frequency 40Hz. The images were averaged over many periods (>6000) to obtain a good signal-to-noise ratio[34.]

In order to extract temperature from the reflectance data, it is essential to quantify the values of κ for all the epitaxial layers as well as the substrate. This is achieved by placing the diode laser on a broad-area thermoelectric unit to modulate its temperature uniformly while taking thermoreflectance images to acquire $\frac{\Delta R}{R}$. ΔT is then measured in a separate experiment using a calibrated micro thermocouple in good thermal contact with the facet, and finally using Eq. 3 to quantify κ. The thermoreflectance coefficients were measured to be 2.23×10^{-4} K^{-1} for GaAs

substrate (in good agreement with published values[31]) and 1.18×10^{-4} K^{-1} for the laser epitaxial layers (cladding and waveguide). Using these values of k, we can convert the obtained thermoreflectance map to a 2D temperature profile of the laser facet within the field of view.

Figure 3.6 shows the temperature profiles of the laser facet (spatially averaged in the slow-axis direction) measured for three values of operating current, in the absence of back-irradiance. The peak in temperature occurs at the position of the active region (which includes the QW and ~550 nm of nearby layers due to the spatial resolution of the thermoreflectance setup), caused by absorption of generated light by the facet. The secondary peaks close to the active region are artifacts caused by thermal expansion. The periodic temperature variation leads to an infinitesimal cyclic motion of the layers which leads to falsely high $\frac{\Delta R}{R}$ values at the boundaries between two different layers.

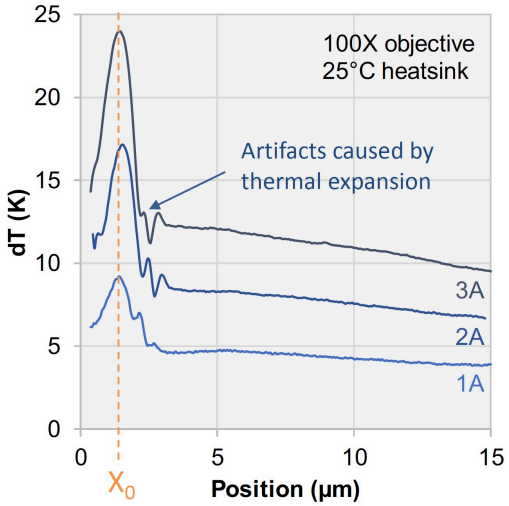

Figure 3.6 | Thermoreflectance data depicting temperature rise as a function of position along the fast axis direction (averaged along the slow axis direction) in the absence of back-irradiance. The active region location is denoted X_0.

Next, the temperature profile of the laser diode facet in the presence of back irradiance, for a bias current $I = 3$ Amps, was measured. The upper panel in Figure 3.7 shows the two-dimensional map of $\frac{\Delta R}{R}$ measured over the entire laser facet with the back-irradiance spot centered at three distinct locations: within the solder ($X–X_0 = -10$ μm), in the substrate close to the active region ($X–X_0 = +5$ μm), and deep within the substrate ($X–X_0 = +28$ μm). The corresponding temperature

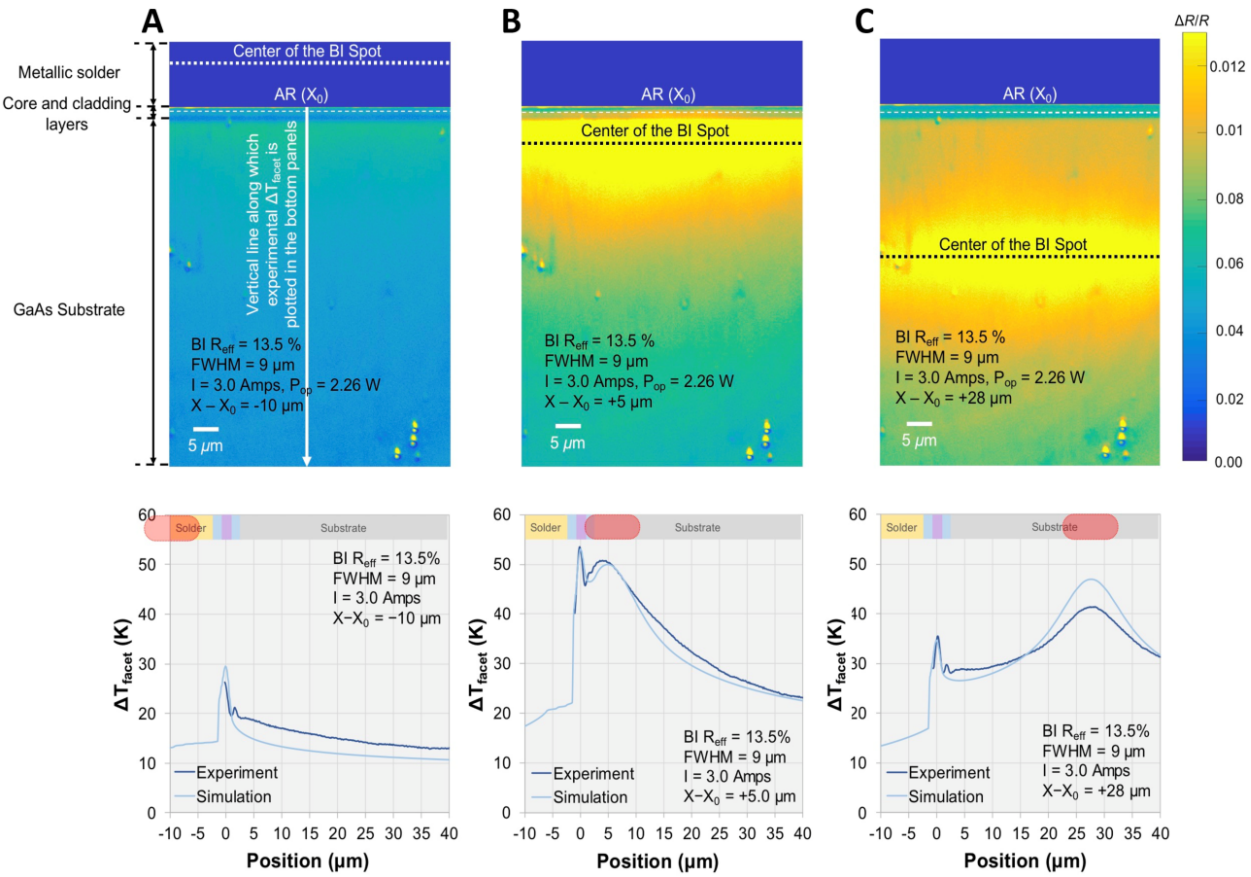

Figure 3.7 | 2D thermoreflectance (ΔR/R) images of the laser facet (upper panel) and horizontally averaged vertical temperature profiles (bottom panel) corresponding to three distinct back-irradiance spot locations. (A) within the solder, 10 μm above the active region, **(B)** within the substrate, 5 μm below the active region, and **(C)** deep within the substrate, 28 μm below the active region. The back-irradiance reflectance level and injection current are specified in the plots, and the position of the active region is set as the reference point (X_0 = 0 μm). A white arrow is drawn on the upper left panel to indicate the vertical slice along which experimentally measured ΔT_{facet} is plotted in the bottom panels. The thermoreflectance signal for the metallic solder (upper panel), which is very noisy due to the solder's large roughness, has been manually removed in post-processing.

profiles spatially averaged in the lateral direction, are depicted in the bottom panel. The temperature spike in the active region is due to self-heating of the diode laser, absorption of the output beam at the facet, and absorption of back-irradiance. In the absence of back-irradiance, the facet temperature rise was found to be ~20 K; thus, the contribution of back-irradiance absorption to the facet temperature at the active region is measured to be 6.3 K, 34.7 K and 14.4 K for back-irradiance spot positions of $X–X_0$ = –10 μm, +5 μm and +28 μm, respectively. It is quite evident from these results that there exists a critical region in the substrate close to the active region of the laser facet, such that when the back-irradiance spot is focused in this region, the temperature rise

in the active region is maximized[22,29,30].

Figures 3.8A and B depict the dependence of back-irradiance component of the active region temperature rise, on the BI spot position, the optical power and the effective back-irradiance level. As shown, the temperature rise in the active region reaches a maximum value when the back-irradiance spot is located nearly 5 µm away from the active region in the substrate. There exist two conflicting phenomena that together give rise to this critical location. Moving the spot closer to the active region leads to an increase in the temperature spike owing to the heat source being closer. However, if the return spot overlaps significantly with the active region, the temperature spike would reduce due to the transparency of the core region at the emission wavelength. As a compromise of these two conflicting phenomena, the critical position is located into the substrate below the active region at a distance roughly equal to half the FWHM of back-irradiance spot. Figures 3.8A and B also show that the temperature rise scales approximately linearly with the optical power of the laser diode as well as the back-irradiance reflectance levels.

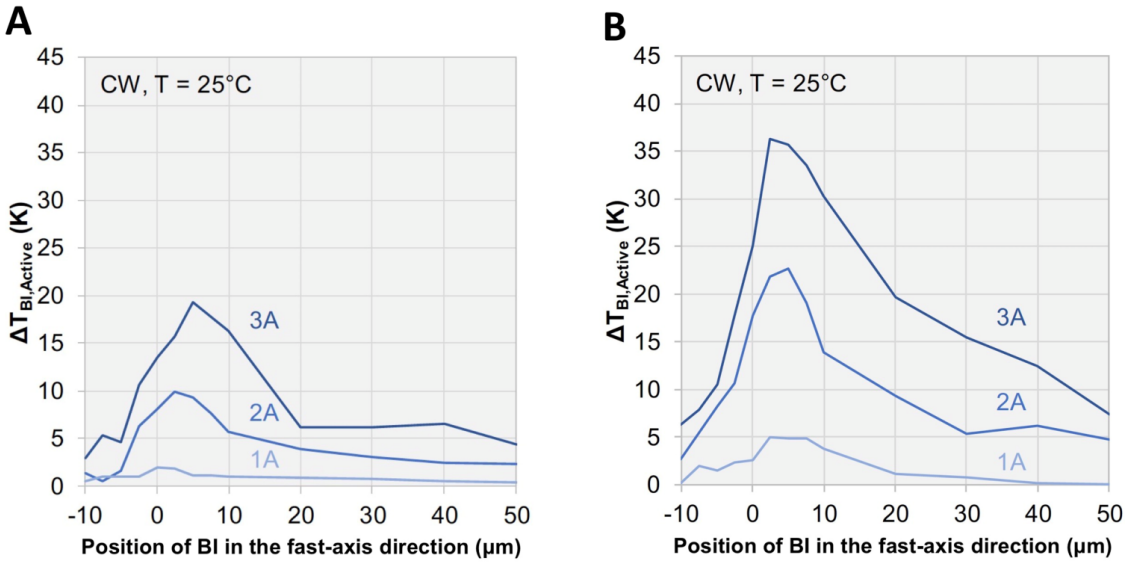

Figure 3.8 | Facet temperature rise at the position of the active region due to back irradiance as a function of the beam offset position and obtained at 1A, 2A, and 3A injection current. Results are depicted for the case of **(A)** 7% back reflection and **(B)** 13.5% back reflection. In all cases, the active region temperature at the facet reaches a maximum when the return beam is pointed several microns toward the substrate.

A two-dimensional (2D) microscopic finite element thermal model that simulates the distributed thermal load presented by the absorption of back irradiance in the various layers of the structure was developed to aid in analysis of the experimental campaign[22,29,30]. The 2D model simulates a 1600 μm x 50 μm domain taken from the center of the bonded laser emitter and comprises the solder, non-pumped window insulator, *p*-side metallization layers, all epitaxial layers to the substrate, the facet coating, and a portion of the submount. Figure 3.9 depicts a cross-sectional illustration of the simulation domain including the mesh definition and key features included in the simulation. The longitudinal thermal loading due to self-heating is included in the model. Thermal loading due to the back-irradiance is treated as follows. The overlap of back-irradiance profile with each layer is first calculated, and the fraction of power absorbed in each layer is estimated based on the reflectance coefficient of each layer in the structure. For the transparent layers comprising the cladding, waveguide, and active region, reabsorption of light near the facet is negligible and therefore not considered. For the partially absorbing substrate and cap layers, the volumetric heating is applied over the characteristic absorption length (Beer-Lambert law). For the metallic layers (including the solder and submount heatsink), the surface heating is assumed. Appropriate boundary conditions are used to represent heat transfer due to conduction from the chip *p*-side into the mount, conduction from the chip *n*-side through the wirebonds to the mount, and convection from the chip facets. Other model features include forward power facet loading, heat spreading through the facet mirror coating, thermal boundary resistance for the non-pumped window insulator, the submount corner radius, and chip/submount facet overhang.

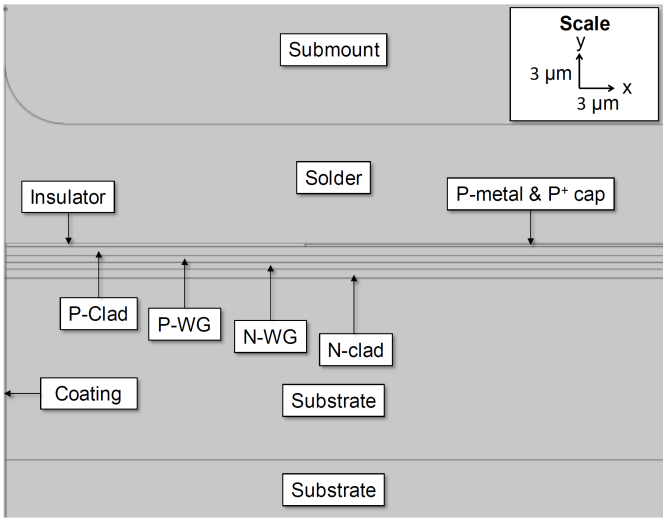

Figure 3.9 | Illustration of the simulation domain and mesh definition for the finite element model.

A series of simulations were carried out to corroborate the results of the nondestructive thermoreflectance testing effort. Figure 3.10 depicts 2D temperature profiles taken along the center of the 200 µm laser stripe over the domain near the laser facet. The simulations depict the results for the laser operating at 3.0 W output power with 10% of the output power being reflected back to the laser facet. Results are shown for three different positions of the return beam. In the first case, the return beam is positioned 3 µm above the active region (striking the solder). In this case, most of the light is reflected and the peak of the temperature distribution occurring at the active region is caused primarily by self-heating. As the return beam is moved to 3 µm below the active region (approximately at the *n*-clad / substrate interface), a large increase in temperature is observed in the vicinity of the active region and significant cross-heating causes the temperature of the active region to increase. As the beam is shifted further into the substrate, the peak in the temperature profile follows, and the facet temperature at the position of the active region decreases.

These simulation results were included in the plots in the bottom panel of Figure 3.7 and show excellent agreement with the experimental results. Figure 3.11 compares the back-irradiance induced temperature rise at the active region versus the return spot location obtained from the simulation and the thermoreflectance experiment. This also indicates excellent agreement between measurement and simulation and conclusively shows that the critical point is real.

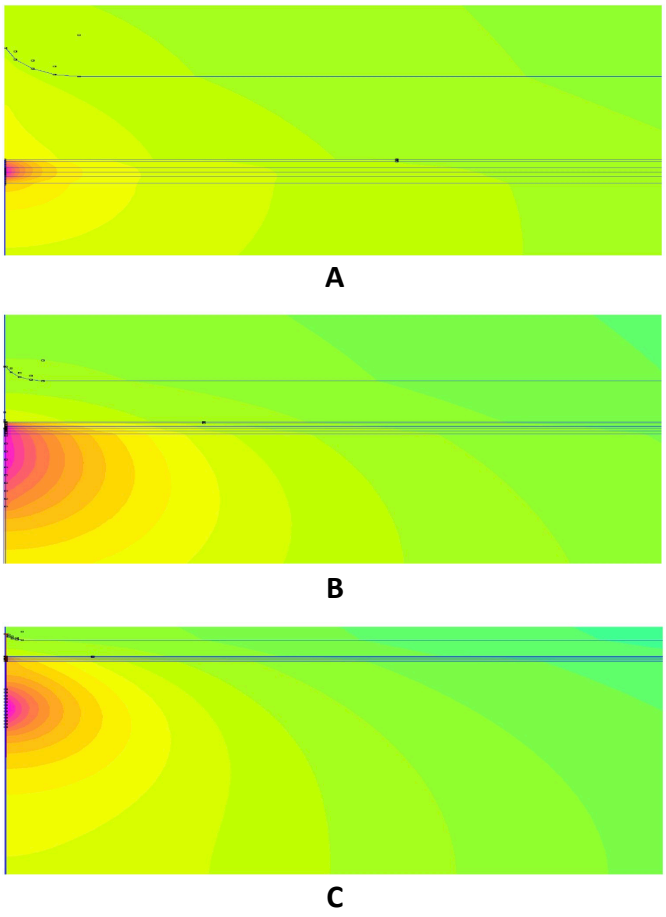

Figure 3.10 | Example images of the 2D temperature profile taken at the center of the 200 µm stripe. In this simulation, the 800 nm class emitter is operating at 3.0 W and 10% of the output power is being reflected back to facet. The offset of the beam with respect to perfect boresight alignment is **(A)** $X-X_0$ = −3 µm, **(B)** $X-X_0$ = +3 µm and **(C)** $X-X_0$ = +10 µm, with X_0 being the position of the quantum well (QW) in the fast axis direction of the facet, a negative "$X-X_0$" means positions above the QW from the *p*-side to the metallic solder and a positive "$X-X_0$" corresponds to positions from the *n*-side to the substrate.

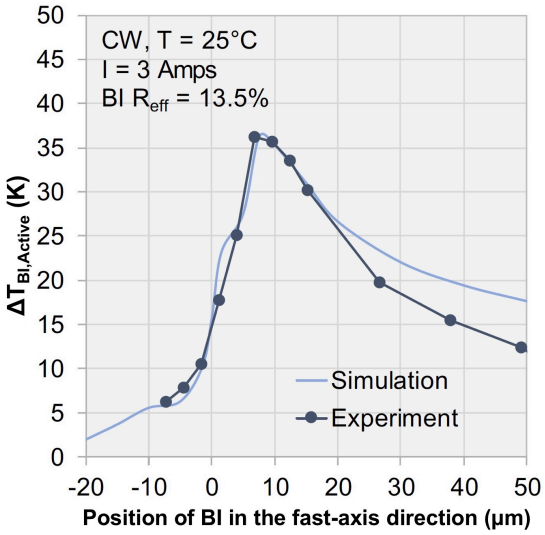

Figure 3.11 | Comparison of the experiment and simulation results of the active region temperature as a function of the position of the back-irradiance spot in the fast-axis direction at I = 3A and a back-irradiance reflectance level of 13.5 %.

3.4 Destructive testing

In order to further assess the impact of back-irradiance, a series of destructive tests were performed on devices subject to varying levels of optical feedback. The overall goal of the effort was to establish and validate a framework which could be used to quantify the effect of back-irradiance on the reliability of diode pump lasers. The experiment was designed to facilitate the introduction of back-irradiance in a manner similar to that used in the non-destructive thermoreflectance testing work, but to do so to a bar array of laser diodes in order to collect failure statistics over time.

A schematic of the experimental setup is shown in Figure 3.12. The back-irradiance system comprises a 4F image relay (with 3X magnification) with the collimated diode laser output plane at one end and a feedback mirror at the other. The 4F image relay allows for precision rotation the feedback mirror to control the position of the return beam in the facet axis at the diode facet to within 140 nm without vignetting. Pickoff diagnostics enable in-situ monitoring of diode wavelength, bar power, and the near-field intensity profile. The optical system operates nearly diffraction-limited, and the size of the return spot is estimated to be ~2 µm FWHM. We note that the emitters comprising the laser bars in this experiment are identical to those used in the nondestructive thermoreflectance testing.

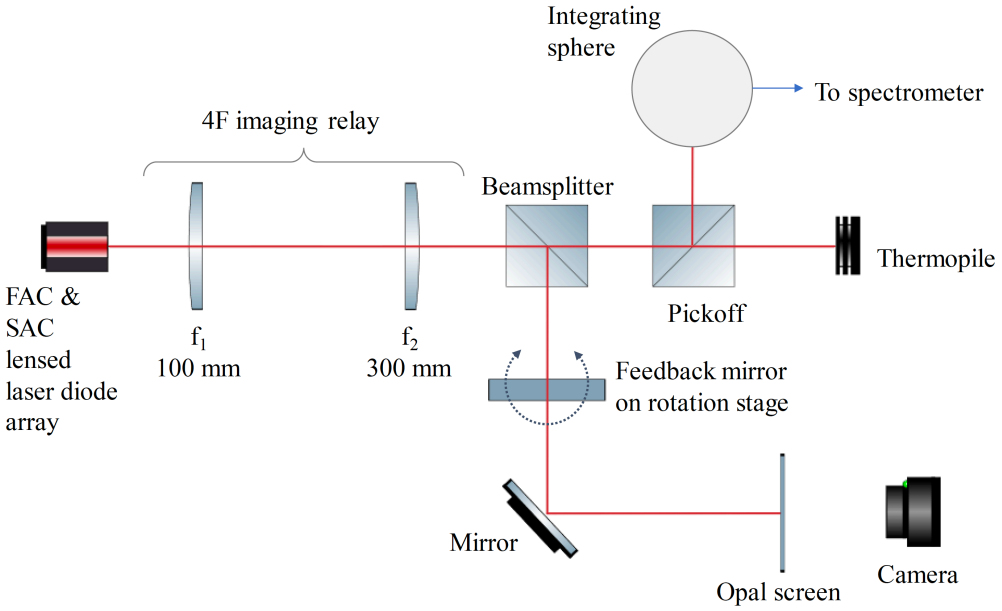

Figure 3.12 | Schematic layout of the destructive testing setup

Calibration of the effective back-irradiance reflectance was carried out as follows. A single-mode collimated probe beam operating at the same emission wavelength as the laser diode array (~800 nm) was directed into the optical setup from a position behind the front focal point of the first lens in the 4F imaging relay at a small angle with respect to the optical axis. In this manner, the return spot was able to be spatially separated from the source laser and captured by a power meter. Various combinations of feedback mirrors and beamsplitters were introduced and the effective reflectance of the setup was calculated by the ratio of the incident beam power to reflected beam power.

In an initial test, a 24-emitter high-power diode laser bar was subjected to back-irradiance to assess its effect on catastrophic optical damage (Fig. 3.13). The bar was operated at 2.9 A per emitter with 12 °C coolant flow. Using a reference $X_0 = 0$ corresponding to the location of the quantum well, the position of the 2 µm return spot was swept in the fast axis from deep in the laser heatsink to deep in the GaAs substrate. Fig.3.13A and Fig.3.13B depict the results of testing at back-irradiance reflectance levels of $R_{eff} = 17\%$ and $R_{eff} = 24\%$, respectively. As shown in both figures, as the region, the position of the return beam is swept from the solder and *p*-side through the waveguide

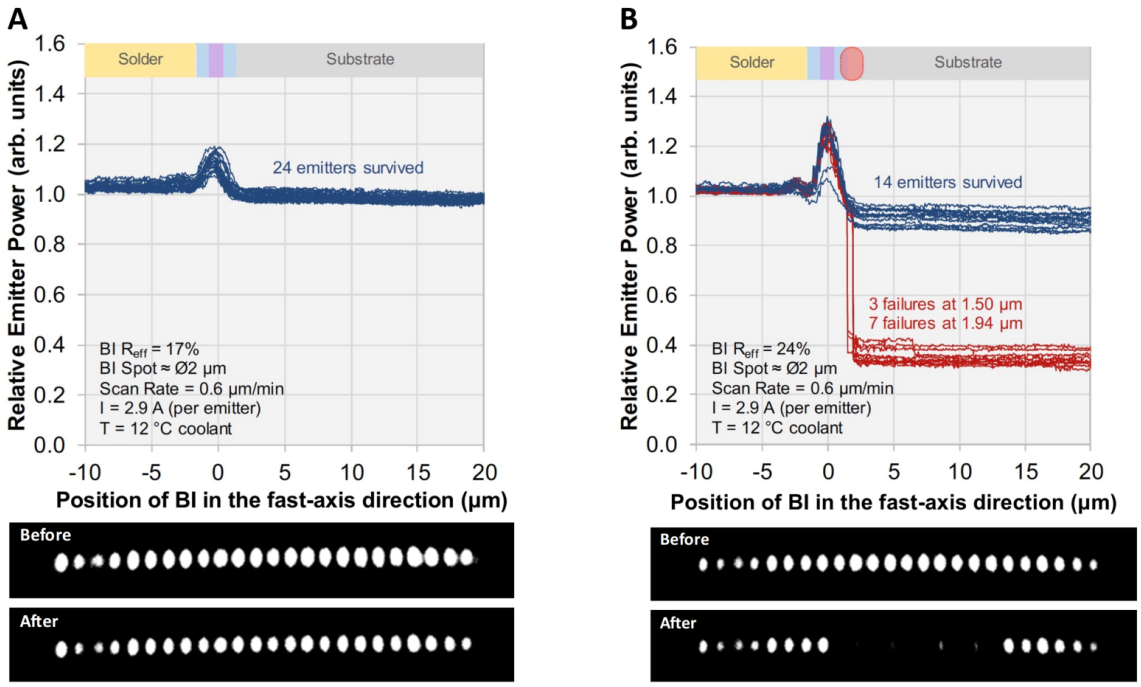

Figure 3.13 |Scanning failure test. (Top) Relative emitter power vs. back-irradiance spot position as it is swept from deep in the heatsink (with the position of BI in the fast-axis direction $X\text{-}X_0 < 0$), through the active region ($X\text{-}X_0 = 0$) and into the substrate ($X\text{-}X_0 > 0$). No emitters undergo catastrophic damage when the spot passes through $X\text{-}X_0 \approx 1.75$ µm at **(A)** 17 % back-irradiance reflectance level (BI R_{eff}) and **(B)** 24% back-irradiance reflectance level. Parameters of the test are displayed in the bottom left corner. (Bottom) Before and after near-field intensity photographs showing the 10 emitter failures (contrast enhanced for publication). These emitters undergo catastrophic damage as the spot passes through X-X0 ≈ 1.75 µm at 24 % back-irradiance reflectance level.

and active relative emitter power increases significantly and subsequently decreases as the beam is further swept into the n-side and substrate. The origin of this power increase is the threshold reduction brought about by the overall reduction in mirror loss (at this operating point, the associated reduction in slope efficiency due to reduced mirror loss is not sufficient to overcome the improvement brought about by the threshold reduction). For the case of 17% effective feedback, the process did not lead to permanent change in the diode characteristics and was repeatable. However, when the effective reflectance was increased to 24%, failure of 10 out the 24 emitters was observed. Interestingly, these failures did not occur at the point of boresight alignment, but rather at a point positioned ~1.75 µm into the substrate (Fig. 3.13B). This critical point is in line with the position corresponding to the peak in $\Delta T_{BI,Active}$ obtained from the non-destructive thermoreflectance testing as well as the 2D finite element model (though the absolute distance from the active region is lower due to the smaller FWHM of the back-irradiance beam),

suggesting a thermal basis for back-irradiance induced COD at this critical location.

Temperature-induced acceleration of failures can be described using the Arrhenius relation (Eq. 3.2) with an activation energy E_a quantifying the strength of the interaction, Boltzmann's constant k_b, and T describing the associated temperature driving the effect[38].

$$Relative\ lifetime \propto \exp\left(\frac{E_a}{k_b T}\right) \tag{3.2}$$

This formalism is often already used to describe the temperature dependence of catastrophic optical mirror damage in high power diode lasers[3,9,10,20], and so represents a reasonable approach for describing the thermal acceleration of COD due to back irradiance.

A series of laser bar failure tests were carried out to bound the activation energy associated with back-irradiance-induced COD. The back-irradiance spot was scanned from a position $X–X_0 = –25$ μm to $X–X_0 = +50$ μm and back again over the course of 3 hours. If no device failure was observed, the back-irradiance reflectance level was increased by changing the feedback mirror and the process repeated until at least one emitter was observed to fail. After this, a new laser bar was tested at a higher injection current and the process repeated to assess the maximum permissible back-irradiance level as a function of operating current (laser power). Figure 3.14A plots the results of this experiment.

Next, a power law dependence (Eq. 3.3) was used to translate the relative power at which COD occurred into a relative device lifetime[3,7,9], normalized to an output power of 3.5 W.

$$Relative\ lifetime = (Emitter\ COD\ power/3.5)^m \tag{3.3}$$

The power factor m is highly empirical and varies according to characteristics of the diode laser. Prior studies at this wavelength have shown values of m ranging from 3 to 6.[5,9,20] The effective reflectance values were translated into temperatures as shown in Eq. 3.4, where T_{Active} is the temperature at the facet at the position of the active region, T_{HS} is the heatsink temperature (directly measured), ΔT_J is the junction

temperature rise due to self-heating (measured from the emission spectrum and corroborated by finite element thermal modelling), and $\Delta T_{BI,\,Active}$ is the temperature rise at the active region due to back-irradiance (based on the 2D microscopic finite element model which is validated by the thermoreflectance testing).

$$T_{Active} = T_{HS} + \Delta T_J + \Delta T_{BI,\,Active} \tag{3.4}$$

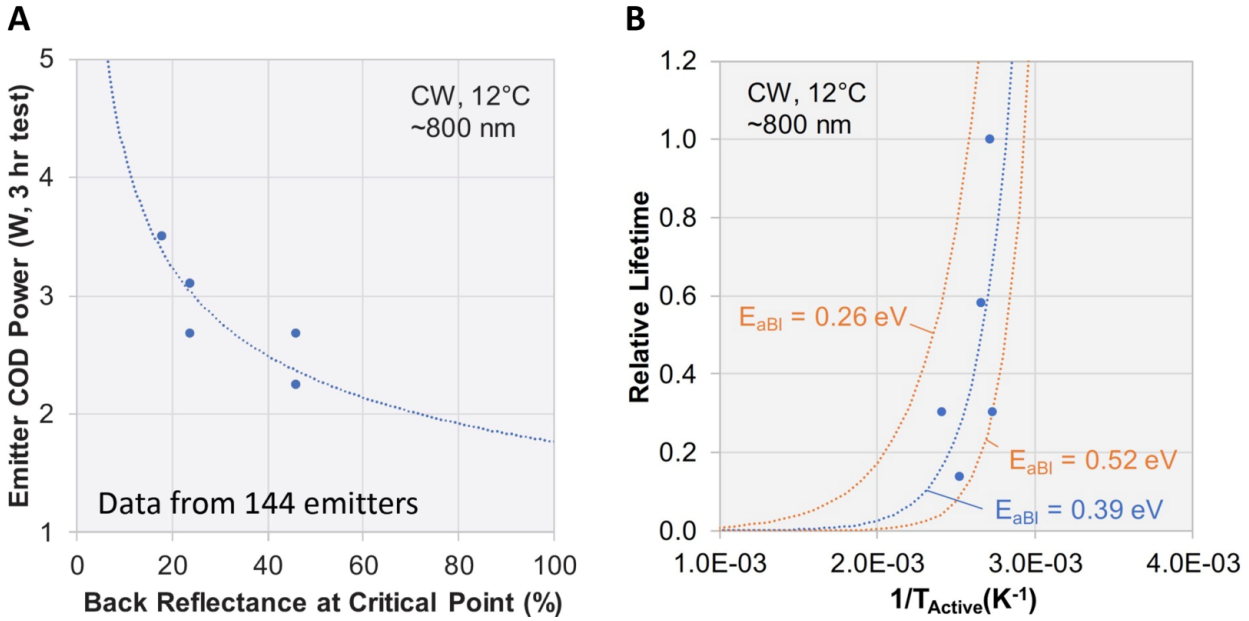

Figure 3.14 | (A) Maximum permissible back reflectance levels at the critical location for several emitted powers. **(B)** The relative lifetime calculated from equation 2 as a function of inverse active region temperature used to extrapolate the activation energy of laser diode COD subject to back-irradiance. The extrapolated E_{aBI} of 0.26 eV, 0.39 eV and 0.52 eV are based on m equal to 6, 4.5 and 3, respectively. The data points are calculated based on m = 4.4.

With these parameters converted, Figure 3.14B plots the relative lifetime of the devices versus the inverse facet temperature at the position of the active region. As shown, for the case of m = 4.5, the data is well fit by an exponential function (Arrhenius equation) having an activation energy 0.39 eV. Curves are also shown for the limiting cases of m = 6 and m = 3, bounding the activation energy between 0.26 eV and 0.52 eV. Interestingly, these two bounds fit well within the range of activation energies (0.2 eV – 0.7 eV) previously reported for catastrophic optical damage[5,9,20]. This hints at the possibility that back-irradiance does not actually introduce a new failure mode – it simply accelerates the standard catastrophic optical damage process through thermal loading.

Having established a preliminary framework to describe the reliability of diode lasers subject to back-irradiance, a reliability lifetest was carried out to demonstrate the validity of the Arrhenius model and corroborate the activation energies estimated in the previous experiment. In this lifetest, a microchannel-cooled stacked laser diode array comprising ten 24-emitter diode bars was subjected to long-term life-testing. A slit aperture was introduced into the image relay of the experimental setup (Fig. 3.12) in order to prevent optical feedback to all bars in the array except one. Optical feedback (17.4%) was applied to one bar at a time (two bars total) with the mirror pointing adjusted to direct the back-irradiance to the worst-case critical point (in the substrate near the n-cladding). This reflectance value was selected in order to ensure failures would not occur too quickly. The stack was run continuously for 433 hours total (150 hours while feedback was applied to Bar 9 and then an additional 283 hours while the feedback was applied to Bar 6). Performed in this manner, the experiment ensured that all of the functioning emitters in the laser array which did not see any back-reflection from the feedback mirror would serve as experimental controls (0% optical feedback).

Near-field images were collected at one-minute intervals for the laser bar under test so that failure times of the feedback bars could be accurately recorded. Full near-field images of the stack array could only be collected before and after execution of the experiment. The array was operated at a constant current of 3.75A per emitter. The facet temperature at the position of the quantum well active region was calculated based on the prior thermoreflectance testing and finite element thermal simulation work. The lifetest data is shown in Table 3.1. A total of 76,798 raw device hours were collected in the study. Note that several bars do not show 24 emitters total – these were bars which exhibited a number of dead emitters prior to the execution of the lifetest (for example, Bar 1 had 12 dead emitters at the start of the lifetest).

The failure data of Table 3.1 was analyzed using the maximum likelihood estimate approach as implemented within a commercial statistical data analysis software package[39]. The data was best-fit by a Weibull distribution and an Arrhenius law model was used to calculate the

TABLE 3.1
BACK-IRRADIANCE LIFETEST DATA

# in State	Last Inspected (hours)	Failed / Suspended	Time (hours)	Facet Temperature at QW (K)	Current per Emitter (A)	Effective Back-Reflectance (%)	Bar ID / Comment
12	433	Suspended	433	315	3.75	0%	Bar 1 – Control
23	433	Suspended	433	315	3.75	0%	Bar 2 – Control
1	1	Failed	433	315	3.75	0%	Bar 2 – Control
20	433	Suspended	433	315	3.75	0%	Bar 3 – Control
4	1	Failed	433	315	3.75	0%	Bar 3 – Control
11	433	Suspended	433	315	3.75	0%	Bar 4 – Control
2	1	Failed	433	315	3.75	0%	Bar 4 – Control
23	433	Suspended	433	315	3.75	0%	Bar 5 – Control
1	1	Failed	433	315	3.75	0%	Bar 5 – Control
16	433	Suspended	433	379	3.75	17.4%	Bar 6
1	148.5	Failed	148.5	379	3.75	17.4%	Bar 6
1	149	Failed	149	379	3.75	17.4%	Bar 6
1	164	Failed	164	379	3.75	17.4%	Bar 6
1	251	Failed	251	379	3.75	17.4%	Bar 6
1	309	Failed	309	379	3.75	17.4%	Bar 6
1	320.5	Failed	320.5	379	3.75	17.4%	Bar 6
1	370.5	Failed	370.5	379	3.75	17.4%	Bar 6
9	433	Suspended	433	315	3.75	0%	Bar 7 – Control
24	433	Suspended	433	315	3.75	0%	Bar 8 – Control
20	150	Suspended	150	379	3.75	17.4%	Bar 9
1	23	Failed	23	379	3.75	17.4%	Bar 9
1	44	Failed	44	379	3.75	17.4%	Bar 9
1	140.5	Failed	140.5	379	3.75	17.4%	Bar 9
20	433	Suspended	433	315	3.75	0%	Bar 10 – Control

acceleration due to the facet temperature at the quantum well. Statistical confidence bounds were directly calculated from the likelihood ratio, and the 90% single-sided upper and lower bounds are reported herein. The shape parameter of the Weibull fit was $0.9 < \beta < 1.8$ with a nominal value 1.3. This observed value is consistent with the assumption that the observed failures were random and independent (i.e. failed emitters were not causing a significant acceleration in failure of neighboring emitters in the array). The activation energy of the Arrhenius equation was found to be $0.15 < E_a < 0.38$ eV, with a nominal value 0.25 eV. This range is consistent with both the bounded estimate determined in prior testing (Fig. 3.14B) and with previously published values of activation energy for catastrophic optical damage. At a nominal use condition with no back-irradiance (corresponding to a peak facet temperature of 315K) the characteristic mean time to failure (MTTF) of an emitter was found to be $1664 < \text{MTTF} < 8729$ hours with a nominal value of 3811 hours. For an effective back-reflection of 17.4% incident at the critical location, the mean emitter lifetime was reduced by $518 < \text{MTTF} < 1271$ hours with a nominal value of 811 hours.

Figure 3.15 depicts the use-level probability Weibull plot of the analyzed data at a back-irradiance level of 0%. The 90% single-sided confidence bounds are also shown on the plot in grey. Interval censored data from the control units which failed at an unknown time in the test is indicated by large horizontal lines in the distribution.

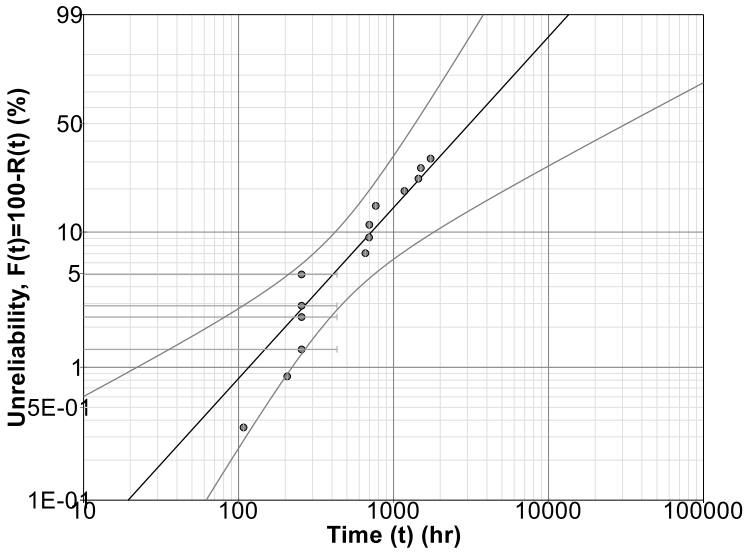

Figure 3.15 | Unreliability Weibull probably plot of the lifetest data at for a use condition of 0% effective back-reflectance.

Figure 3.16 depicts the dependence of the failure acceleration factor for both the 90% upper confidence and 50% median (nominal) values of activation energy (0.38 and 0.25 eV, respectively) versus the effective back-reflection on the diode laser facet at the critical location. It is important to note that these tests were performed for the case of the back-reflected beam being directed to the worst-case critical location in the substrate adjacent to the n-cladding. Because the failures appear to be thermally-accelerated, back-irradiance to other locations are expected to increase the facet temperature at the position of the quantum well active region by a lesser amount. Therefore, the plot in Fig.3.16 can be interpreted as an expected worst-case failure acceleration model for back-irradiance induced failure of 800 nm class diode lasers operating continuous wave.

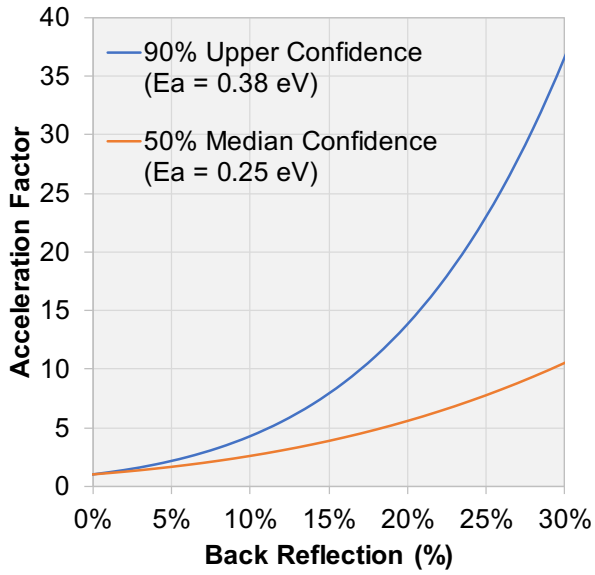

Figure 3.16 | Acceleration factor vs. effective back-reflection condition for the continuous wave 800 nm diode lasers. Plots are shown for the 90% upper single-sided confidence and 50% median confidence values of activation energy and correspond to the case of the effective back-reflection being directed to the critical location in the device structure (in the substrate adjacent to the *n*-cladding layer).

3.5 Comparison between 800 nm, 900 nm, and 1 μm device fabricated on GaAs substrate

Since we've identified substrate absorption as one of the main causes for BI-induced heating, devices fabricated on undoped GaAs (having an absorption cut-off wavelength at ~870 nm, see Fig. 3.17) with different wavelengths are expected to react differently to BI. For this reason, we tested three sets of devices having emission wavelengths of ~800 nm, ~900 nm, and 1μm, at which the GaAs substrate is strongly absorbing, weakly absorbing, and non-absorbing, respectively.

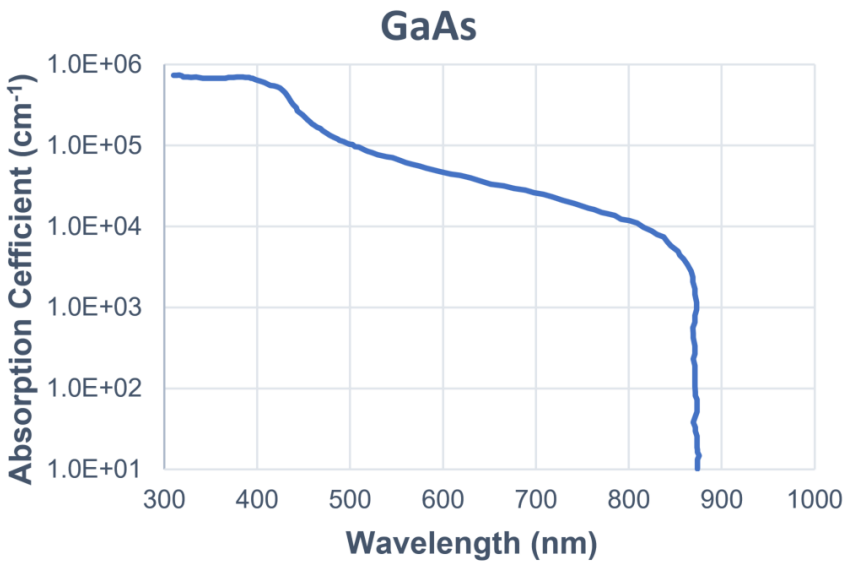

Figure 3.17 | Room temperature optical absorption coefficient of undoped GaAs. The absorption values were extracted from Ref. 40 and recast in terms of wavelength.

In order to test the "most-dangerous" scenario where the BI spot is tightly focused such that its power is concentrated, we redesigned the TR setup to reduce aberrations in the BI beam path (Fig. 3.18), obtaining a sharp BI spot having a 1.3 μm FWHM (Fig. 3.19A). In addition, a 4F system is placed in the BI pathway to minimize the change in the angle content of the reflected beam as it's steered by the dielectric mirror affixed to a motorized mount. This leads to a consistent BI power across different BI spot locations (Fig. 3.19B).

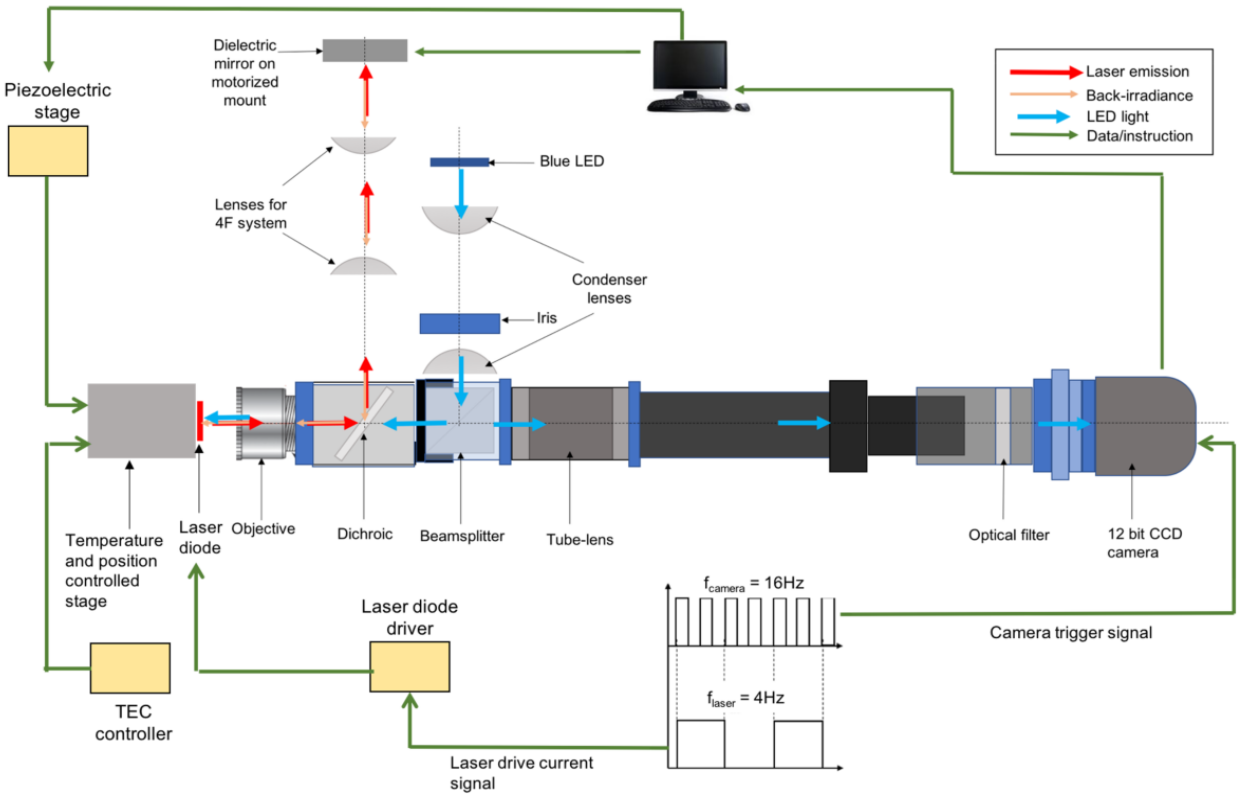

Figure 3.18 | Improved thermoreflectance imaging setup with diffraction-limited BI spot. The outcoupled laser light is first being collimated by the objective before reaching the angled dichroic, reducing aberrations. A 4F image relay system is placed in the BI beam path (between the 45-degree dichroic and the dielectronic mirror) to conserve the angle content as the BI is steered to enter the back focal-plane of the objective at various angles. Image courtesy of Aman Kumar.

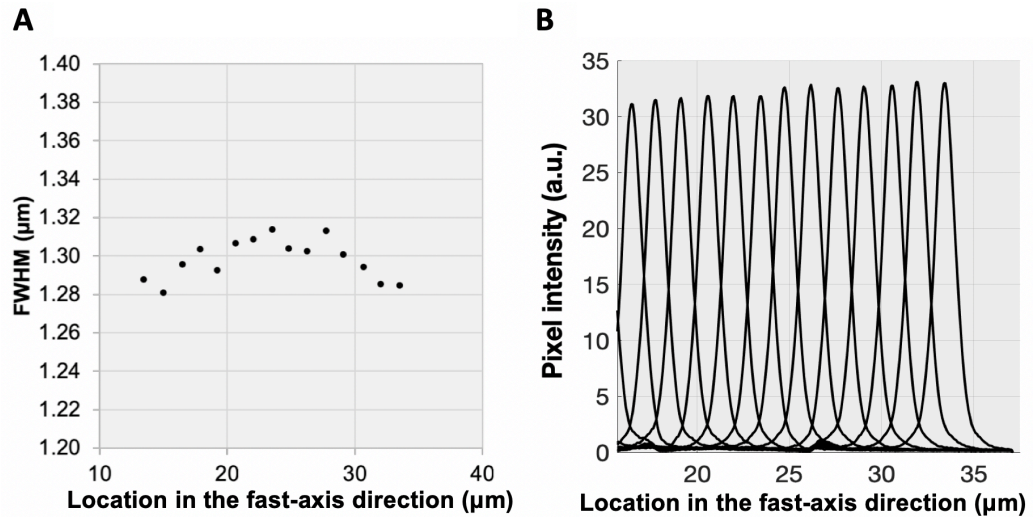

Figure 3.19 | (A) Focused BI spot with a FWHM of 1.29 μm when it's located at various locations in the fast-axis direction (i.e. the direction perpendicular to the epitaxial layers). **(B)** Pixel intensity of the BI spot at different locations are very consistent due to the presence of the 4F image relay system. Image courtesy of Aman Kumar.

~800 nm Diode laser (substrate is strongly absorbing)

Fig. 3.20 shows the 2D temperature maps of the facet for the ~800nm diode laser as the BI spot is located deep into the solder, 2µm below the quantum well (QW) into the substrate, and deep into the substrate.

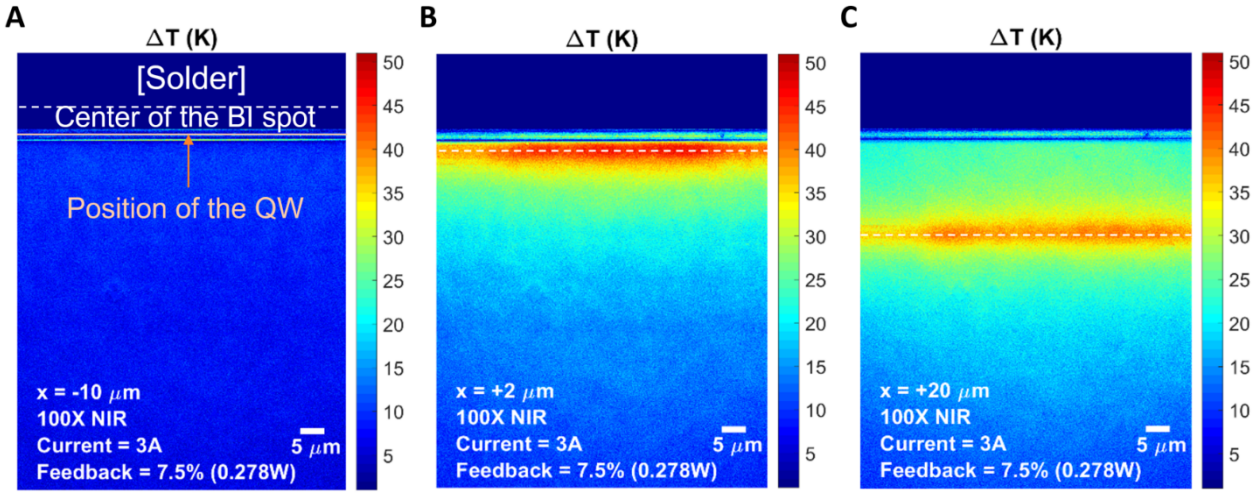

Figure 3.20 | 2D temperature maps for the facet of the 800nm device when the return spot is located **(A)** deep into the solder, **(B)** close to the QW into the substrate, and **(C)** deep into the substrate. The text at the bottom left corresponds to the position of the BI spot, the magnification of the objective, the driving current as well as R_{eff} followed by BI power reaching the facet (shown in brackets)

Fig. 3.21 depicts how ΔT_{QW} evolves as the BI spot is swept in the epi-layer stacking direction, with the position of the active region set as reference (X_0 = 0µm). For the 800nm DL, a sharp temperature rise is observed when the BI is centered ~2.5 µm from the QW into the absorbing substrate, a distance approximately equal to the FWHM (~2 µm) of the return spot. This critical location maximizes ΔT_{QW} since at this location a large fraction of the spot is within the substrate and the heat generated is also close to the QW. Moving the BI spot close to the QW reduces ΔT_{QW} due to reduced absorption, while moving the BI spot further into the substrate also leads to a drop in ΔT_{QW} due to the absorption hot-spot located further away from the QW. It is observed that when the C-mount laser is run at 3A with 7.5 % effective back-reflectance level, ΔT_{QW} with BI (29.28K) nearly triples that of the free-running (i.e. without back-irradiance) case (11.36K). Interestingly, a secondary peak in ΔT_{QW} also appears when the BI spot is ~2 µm above the QW into the metallic solder; one possible explanation for this peak is the weak absorption of 800 nm light by the solder material.

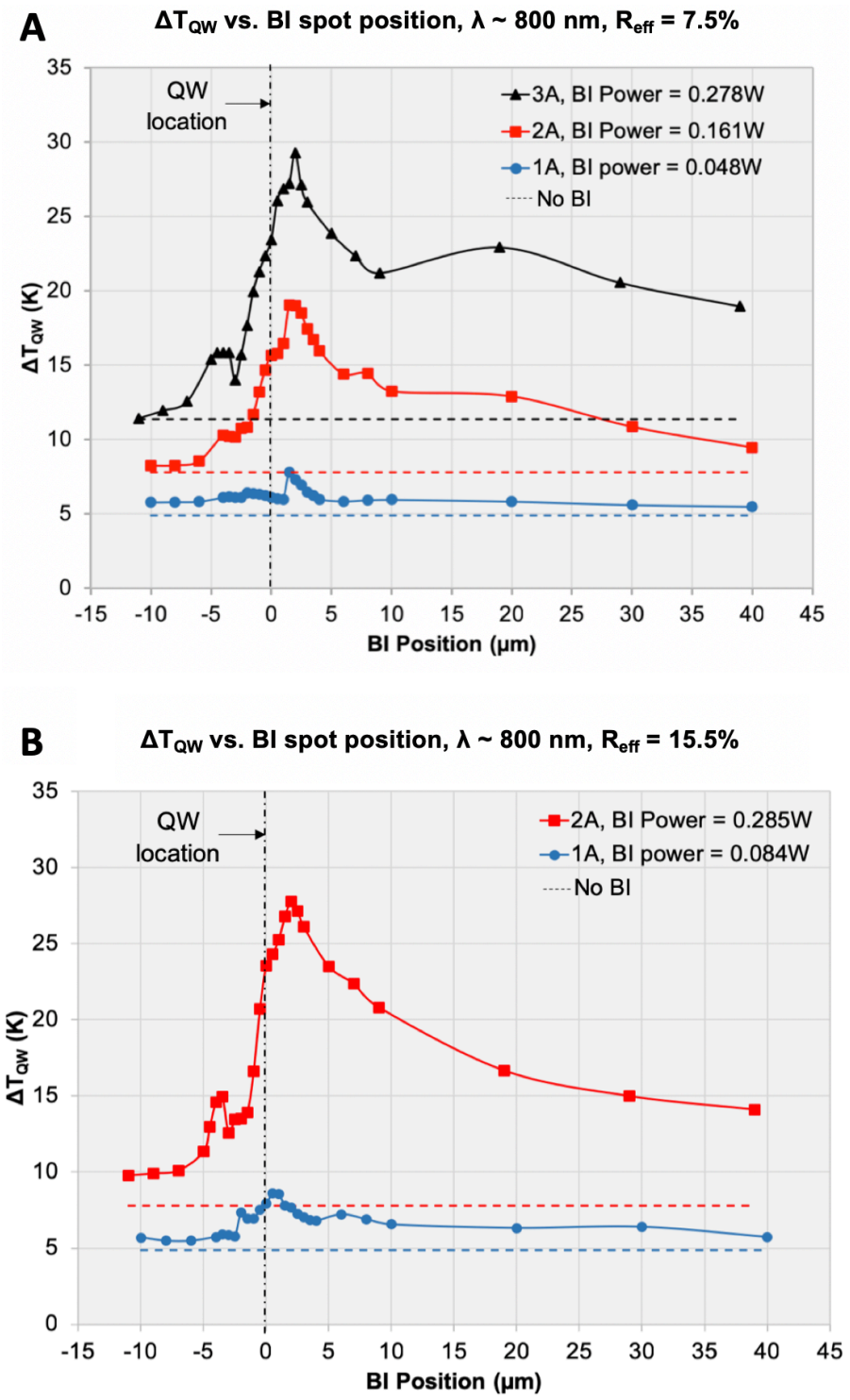

Figure 3.21 | ΔT_{QW} versus the BI spot position with respect to the quantum well location (X_0 = 0μm) for the ~800 nm device with the direction going towards the substrate being positive, at feedback levels of **(A)** R_{eff} = 7.5% and **(B)** R_{eff} = 15.5%. For all currents and feedback levels tested, a surge in ΔT_{QW} can be observed at a BI position (X) of ~2.5μm, corresponding to a critical location at which the device is potentially most susceptible to COD. The dotted lines correspond to the ΔT_{QW} for the case without back-irradiance. Image courtesy to Aman Kumar.

~900 nm Diode laser (substrate is weakly absorbing)

Fig. 3.22 shows the 2D ΔT map of the facet for three distinct BI spot locations. The substrate heating is minimal due to low absorptivity of GaAs at 900nm.

Similar to Fig. 3.21, Fig. 3.23 depicts how ΔT_{QW} varies as the BI spot is swept in the direction perpendicular to the epi-layer stacking direction for the ~900nm DL. As mentioned above, a different mirror (50%R) is used to yield an R_{eff} = 10.98 % (Fig. 3.23A), which compensates the 900nm DL's lower P_{op} at 3A compared to that of the 800nm diode laser (Fig. 3.21), resulting in a P_{BI} of 0.293W which is close to the 0.276W BI power for the 800nm DL. For the 900nm device, the rise in ΔT_{QW} when BI is directed towards the critical location (~2.5 μm away from the QW into the substrate) becomes much less pronounced. This can be seen by its comparable magnitude with the secondary shoulder at the solder/epi-layer interface, whereas in the case of the 780nm device the sharp rise at the epi-layer/substrate interface overwhelmingly exceeds the secondary shoulder. This is attributed to the much smaller absorption coefficient of GaAs at 900 nm. We note that the same values of TRCs with those of the 800nm device are used to convert raw dR/R to ΔT. Hence, the ΔT_{QW} shown in the figures can only be compared in relative magnitudes; an independent calibration, will be carried out in the near future.

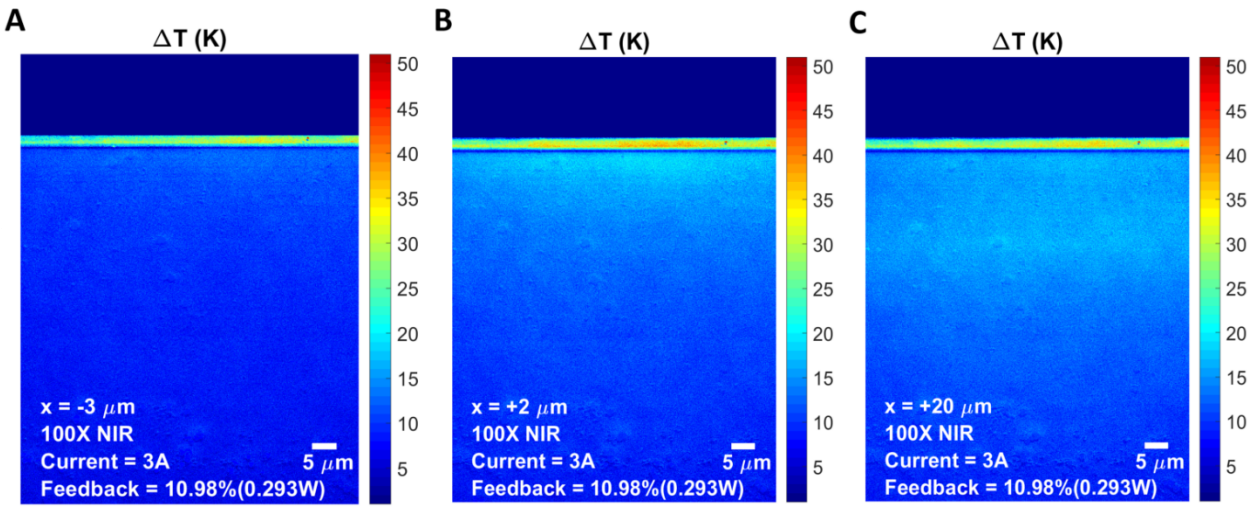

Figure 3.22 | 2D temperature maps for the facet of the 900nm device when the return spot is located **(A)** deep into the solder, **(B)** close to the QW into the substrate, and **(C)** deep into the substrate. The text at the bottom left corresponds to the position of the BI spot, the magnification of the objective, the driving current as well as R_{eff} followed by BI power reaching the facet (shown in brackets). The TRC of the core and cladding for this laser has not been calibrated.

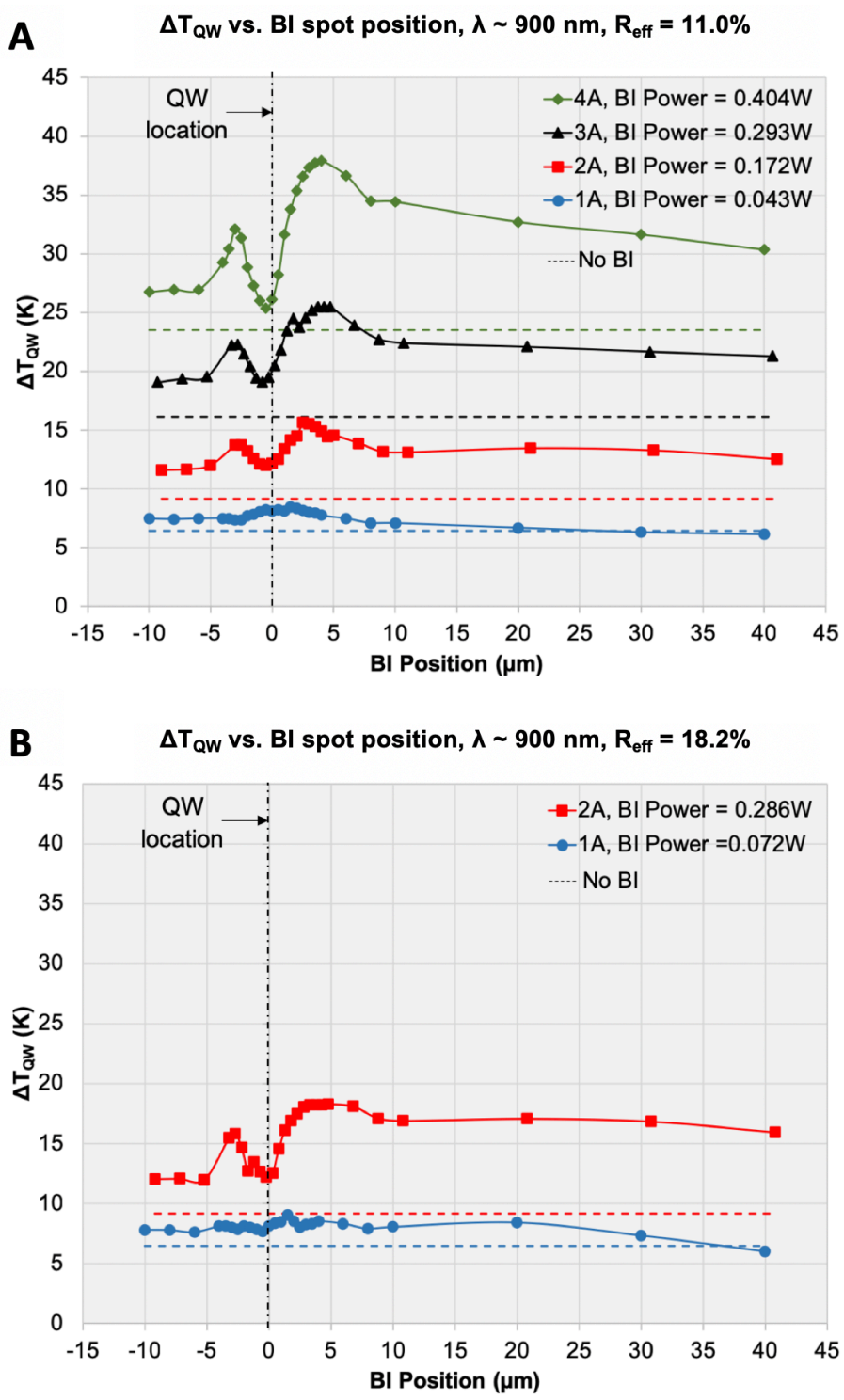

Fig. 3.23 | ΔT_{QW} versus the BI spot position with respect to the quantum well location (X_0 = 0μm) for the 900nm device, with the direction going towards the substrate being positive, for TRBI experiment with **(A)** R_{eff} = 10.98% and **(B)** R_{eff} = 18.19%. The surge in ΔT_{QW} observed at X ~ 2.5μm becomes much less pronounced due to weak absorption of 900 nm light by the substrate. The secondary shoulder at the solder/epi-layer interface becomes almost comparable (although still lower) to the primary peak caused by substrate absorption. Image courtesy to Aman Kumar.

~1 μm Diode laser (substrate is non-absorbing)

At ~1μm GaAs has an extremely small absorption coefficient, so we anticipate the additional temperature rise in the active region due to substrate absorption to be negligible. It can be clearly seen from Fig. 3.24 that the active region temperature reaches a maximum when BI is located at the solder/epi-layer interface (left panel) rather than at the epi-layer/substrate interface (middle panel).

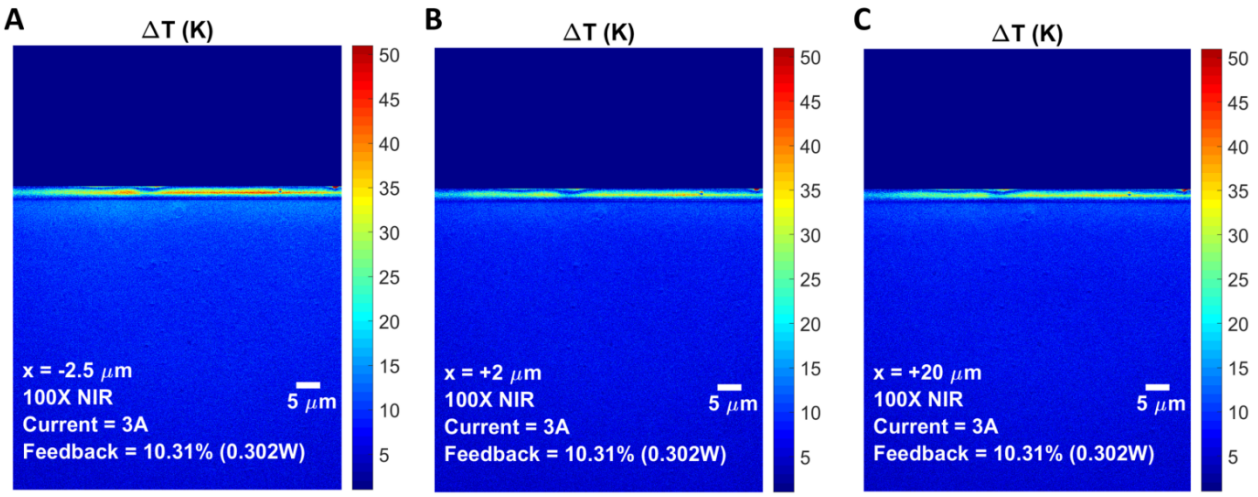

Figure 3.24 | 2D temperature maps for the facet of the ~1μm device when the return spot is located **(A)** deep into the solder, **(B)** close to the QW into the substrate, and **(C)** deep into the substrate. The text at the bottom left corresponds to the position of the BI spot, the magnification of the objective, the driving current as well as R_{eff} followed by BI power reaching the facet (shown in brackets). The TRC of the core and cladding for this laser has not been calibrated.

In Fig. 3.25 shown below we can clearly see that the shoulder due to substrate absorption becomes extremely weak and much less pronounced than the secondary shoulder possibly due to absorption by metallic solder. A recent publication on a 950nm high-power diode laser with GaAs substrate showed similar trend where the maximal ΔT_{QW} occurs as the BI is positioned close to the solder/epi-layer interface[41]. The implies that devices having non-absorbing substrates would be most susceptible to BI-induced COD when the BI spot hits the solder close to the solder/epi-layer interface.

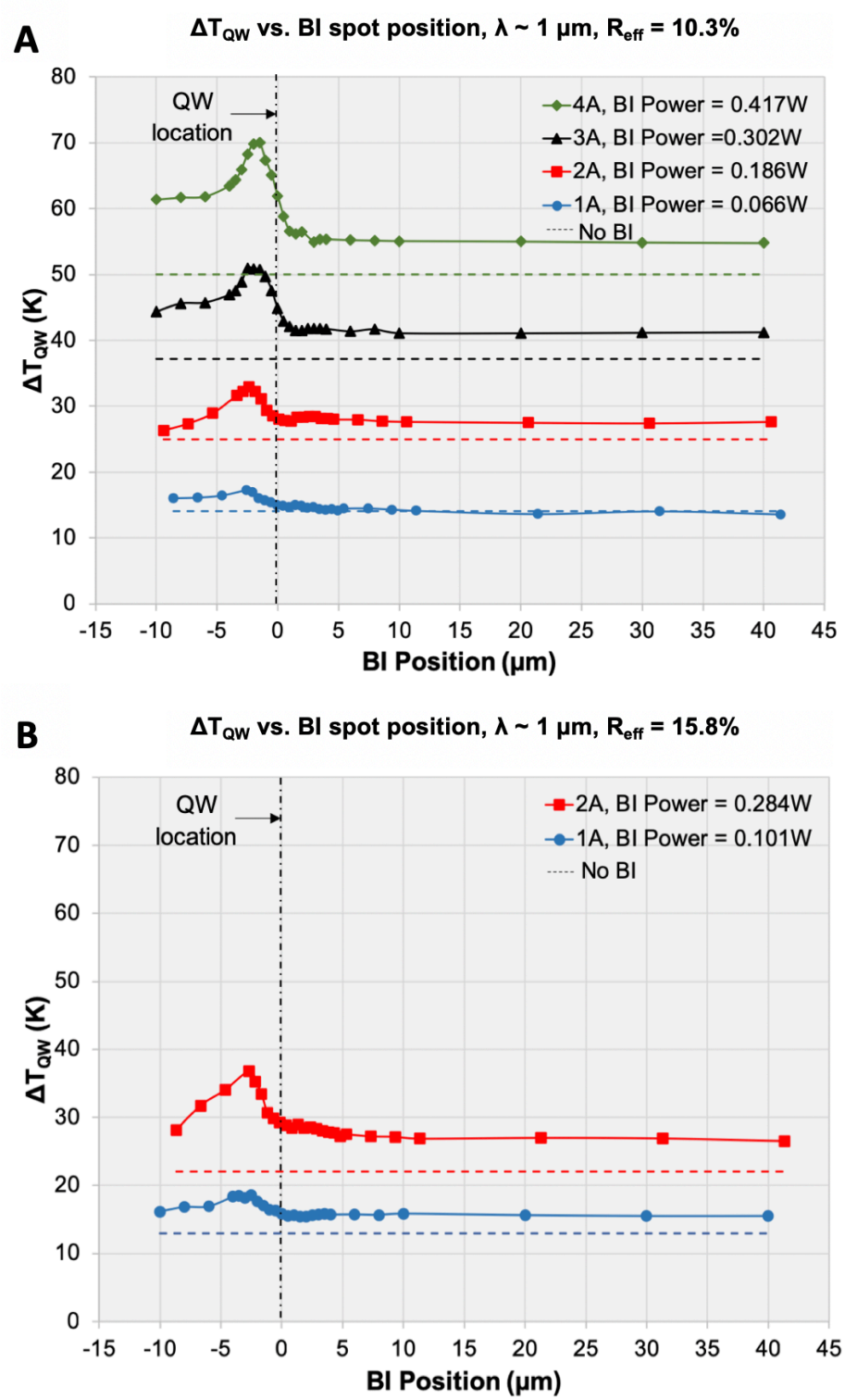

Figure 3.25 | ΔT_{QW} versus the position of the BI spot with respect to the quantum well location (X_0 = 0μm) for the 980nm diode laser, with **(A)** R_{eff} = 10.31% and **(B)** R_{eff} = 14.80% (using a metallic feedback mirror).The shoulder in ΔT_{QW} due to BI absorption at the substrate (at X ~ 2.5μm) becomes very weak due to negligible absorption of 1 μm light by the substrate material. The secondary shoulder at the solder/epi-layer interface becomes pronounced in comparison. Image courtesy to Aman Kumar.

3.6 Discussion and summary

Understanding the impact of back-irradiance on the reliability of high-power diode lasers is critical for the robust design and deployment of such devices in modern high energy diode-pumped laser systems. Nondestructive high-resolution thermoreflectance imaging has been applied to measure the temperature profile of high-power diode lasers subject to back-irradiance. For devices fabricated on absorbing substrates and operated continuous wave, this work demonstrates that the position of the back-irradiance spot matters greatly, with a critical location (corresponding to maximum facet temperature rise at the position of the quantum well active region) being in the substrate adjacent to the n-cladding layers of the device. This experimental result is further corroborated through microscopic finite element thermal simulations. The temperature rise in the active region caused by absorption of back-irradiated light at the critical location is considerably greater than that when the return spot is at the solder, deep into the substrate, or even directly incident on the active region of the device. Destructive testing further confirms that this critical location is the most "dangerous" spot for the back-irradiance to cause catastrophic optical mirror damage

Based on the assessed thermal nature of back-irradiance induced failure, a framework is established following quantitative accelerated lifetesting methodology to describe the impact of back-irradiance on laser reliability. The acceleration model is found to follow an Arrhenius law which uses as its input the back-irradiance induced facet temperature rise at the position of the quantum well active region. Preliminary short-term lifetime assessment for laser diodes with back-irradiance initially bounded the activation energy of back-irradiance induced COD between 0.26 eV and 0.52 eV, and subsequent (independent) long-term lifetesting shows that the activation energy is between 0.15 eV and 0.38 eV. Both of these results lie within the range of activation energies associated with thermal-induced catastrophic optical damage in high power diode lasers operating without back-irradiance, further corroborating a thermal basis for back-irradiance-induced failure.

It is also important to reiterate that the failure tests were carried out only for diode lasers operating around 800 nm, where the substrate is highly absorbing. We have confirmed that devices operating at longer wavelengths will be more resilient to this effect, especially for the case where the GaAs substrate is transparent to the laser wavelength. Nevertheless, the framework developed herein is

expected to be applicable to these circumstances, but will likely require additional validation (for example, the activation energy will likely be somewhat different).

Finally, this study has demonstrated that generic rules-of-thumb for maximum permissible back-irradiance (e.g. back reflections below about 5% are “safe”) in a laser system designs is likely inadequate because details such as the location of the return spot on the diode facet would impact device reliability significantly. It is expected that these results will enable improved diode-pumped laser systems.

References

1. Z. Yao *et al.* High power and high reliability InGaAs broad area lasers emitting between 910 and 980 nm. *Lasers and Electro-Optics Society 2000 Annual Meeting. LEOS 2000. 13th Annual Meeting. IEEE* **2**, 510-511 (2000).

2. H. U. Pfeiffer, S. Arlt, M. Jacob, C. S. Harder, and I. D. Jung. Reliability of 980 nm pump lasers for submarine, long-haul terrestrial, and low-cost metro applications. *Optical Fiber Communication Conference* (2002).

3. V. Rossin, T. Strite, E. Zucker, and T. H. Ju. AlGaInAs broad-area lasers exhibit outstanding reliability. *Compound Semiconductor*, 43-46 (2003).

4. P. A. Crump *et al*. High-efficiency high-power 808-nm laser array and stacked arrays optimized for elevated temperature operation. *Proc. SPIE* **5336** (2004).

5. J. Van de Casteele, M. Bettiati, F. Laruelle, V. Cargemel, P. Pagnod-Rossiaux, P. Garabedian, L. Raymond, D. Laffitte, S. Fromy, D. Chambonnet and J. P. Hirtz. High reliability level on single-mode 980 nm-1060 nm diode lasers for telecommunication and industrial applications. *Proc. of SPIE* **6876** (2008).

6. P. Leisher *et al.* Highly reliable high-efficiency wavelength-stabilized 885 nm diode laser bars. *Proc. SPIE* **7325** (2009).

7. J. Wang *et al.* Reliability and performance of 808-nm single emitter multi-mode laser diodes. *Proc. SPIE* **7583** (2010).

8. R. Feeler, J. Junghans, J. Remley, D. Schnurbusch, and E. Stephens. Reliability of high-power QCW arrays. *Proc. SPIE* **7583** (2010).

9. L. Bao *et al*. High reliability and high performance of 9xx-nm single emitter laser diodes. *Proc. SPIE* **7918** (2013).

10. L. Bao *et al.* Reliability of high power/brightness diode lasers emitting from 790 to 980 nm. *Proc. SPIE* **8605** (2013).

11. P. Crump *et al*. Efficient High-Power Laser Diodes. *IEEE Journal of Selected Topics in Quantum Electronics* **19**, 1501211-1501211 (2013).

12. J. Campbell, T. Semenic, P. Leisher, A. Bhunia, M. Mashanovitch, and D. Renner. 980nm diode laser pump modules operating at high temperature. *Proc. SPIE* **9730** (2016).

13. J. Rieprich, M. Winterfeldt, J. Tomm, R. Kernke, and P. Crump. Assessment of factors regulating the thermal lens profile and lateral brightness in high power diode lasers. *Proc. SPIE* **10085** (2017).

14. S. Heinemann, S. D. McDougall, G. Ryu, L. Zhao, X. Liu, C. Holy, C.-L. Jiang, P. Modak, Y. Xiong, T. Vethake, S. G. Strohmaier, B. Schmidt, H. Zimer. Advanced chip designs and novel cooling techniques for brightness scaling of industrial, high power diode laser bars. *Proc. SPIE* **10514** (2018).

15. H. Kressel and H. Mierop. Catastrophic Degradation in GaAs Injection Lasers. *Journal of Applied Physics* **38**, 5419-5421 (1967).

16. A. Moser, E. E. Latta, and D. J. Webb. Thermodynamics approach to catastrophic optical mirror damage of AlGaAs single quantum well lasers. *Applied Physics Letters* **55**, 1152-1154 (1989).

17. P. W. Epperlein, P. Buchmann, and A. Jakubowicz. Lattice disorder, facet heating and catastrophic optical mirror damage of AlGaAs quantum well lasers. *Applied Physics Letters* **62**, 455-457 (1993).

18. D. A. Livshits, *et al*. Improved catastrophic optical mirror damage level in InGaAs/AlGaAs laser diodes. *Electronics Letters* **36**, 1848-1849 (2000).

19. M. B. Sanayeh *et al.* Temperature-power dependence of catastrophic optical damage in AlGaInP laser diodes. *Applied Physics Letters* **91** (2007).

20. K. Häusler, U. Zeimer, B. Sumpf, et al. Degradation model analysis of laser diodes. *Journal of Material Science: Materials in Electronics* **19**, 160-164 (2008).

21. Q. Zhang, Y. Xiong, H. An, K. Boucke, and G. Treusch. Unveiling laser diode "fossil" and the dynamic analysis for heliotropic growth of catastrophic optical damage in high power laser diodes. *Scientific Reports* **6**, 1-12 (2016).

22. A. Takeda, R. Shogenji, and J. Ohtsubo. Dynamics and pulse-package oscillations in broad-area semiconductor lasers with short optical feedback. *Applied Physics Letters* **101** (2012).

23. T. Tachikawa, R. Shogenji, and J. Ohtsubo. Observation of multi-path interference in broad-area semiconductor lasers with optical feedback. *Optical Review* **16**, 533-539 (2009).

24. B. Leonhäuser, H. Kissel, J. W. Tomm, M. Hempel, A. Unger, and J. Biesenbach. High-power diode lasers under external optical feedback. *Proc. SPIE* **9348** (2015).

25. Y. Takiguchi, T. Asatsuma, and S. Hirata. Effect of the threshold reduction on a catastrophic optical mirror damage in broad-area semiconductor lasers with optical feedback. *Proc. SPIE* **6104** (2006).

26. B. Leonhäuser, H. Kissel, A. Unger, B. Köhler, and J. Biesenbach. Feedback-induced catastrophic optical mirror damage (COMD) on 976nm broad area single emitters with different AR reflectivity. *Proc. SPIE* **8965** (2014).

27. M. Hempel, M. Chi, P. M. Petersen, U. Zeimer, and J. W. Tomm. How does external feedback cause AlGaAs-based diode lasers to degrade? *Applied Physics Letters* **102** (2013).

28. P. O. Leisher, *et al.* Root cause investigation of back-irradiance-failure of high power diode lasers. *2017 IEEE High Power Diode Lasers and Systems Conference (HPD)*, 53–44 (2017).

29. C. Li, K. P. Pipe, C. Cao, P. Thiagarajan, R. J. Deri, and P. O. Leisher. CCD-based thermoreflectance imaging of high-power diode lasers with back-irradiance. *Proc. SPIE* **10637** (2018).

30. C. Li, K. P. Pipe, C. Cao, P. Thiagarajan, R. J. Deri, and P. O. Leisher. Thermal imaging of high power diode lasers subject to back-irradiance. *Applied Physics Letters* **112** (2018).

31. P. W. Epperlein. Micro-temperature measurements on semiconductor laser mirrors by reflectance modulation: A newly developed technique for laser characterization. *Japanese Journal of Applied Physics* **32**, 5514-5522 (1993).

32. D. Guidotti and H. M. Driel. Spatially resolved defect mapping in semiconductors using laser-modulated thermoreflectance. *Applied Physics Letters* **47**, 1336-1338 (1985).

33. S. Dilhaire, S. Jorez, L.-D. Patiño-Lopez, W. Claeys, and E. Schaub. Laser diode light efficiency determination by thermoreflectance microscopy. *Microelectronics Journal* **32**, 899-901 (2001).

34. S. Dilhaire, S. Grauby, and W. Claeys. Thermoreflectance calibration procedure on a laser diode: application to catastrophic optical facet damage analysis. *IEEE Electron Device Letters* **26**, 461-463 (2005).

35. J. Christofferson and A. Shakouri. Thermoreflectance based thermal microscope. *Review of Scientific Instruments* **76** (2005).

36. P. Chan, K. Pipe, J. Plant, R. Swint, and P. Juodawlkis. Temperature mapping and thermal lensing in large-mode, high-power laser diodes. *Applied physics letters* **89** (2006).

37. M. Farzaneh, *et al.* CCD-based thermoreflectance microscopy: principles and applications. *Journal of Physics D: Applied Physics* **42** (2009).

38. W. B. Nelson, *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*, Wiley Interscience (2004).

39. HBM Prenscia, Reliasoft Weibul++ / ALTA Pro, ver. 11 (2017).

40. A. Luque and S. Hegedus, *Handbook of Photovoltaic Science and Engineering*, John Wiley & Sons (2003).

41. S. Rauch, C. Holly, and H. Zimer. Catastrophic optical damage in 950-nm broad-area laser diodes due to misaligned optical feedback and injection. *IEEE Journal of Quantum Electronics* **54** (2018).

Chapter 4
Thermal Mitigation at the Device Level – Thermionic Refrigeration at CNT-CNT Junctions

Applied Physics Letter **109**, 163901 (2016)

4.1 Introduction

Techniques to remove heat at the device level can be broadly separated into two categories: passive cooling methods that ultimately rely on natural convection for heat removal, and active cooling methods that utilize forced convection or thermoelectric/thermionic effects, etc. The second part of Chapter 3 investigates the usage of one-dimensional (1D) carbon-nanotubes (CNTs) as active cooling material via the solid-state thermionic (TI) effect. Carbon nanotubes are well known for their excellent electrical and thermal transport properties,[1-4] and have been considered for applications ranging from electrical interconnects[5-7] and field effect transistors[8-10] to heat spreaders[11]. For electronics applications that rely on ballistic electron transport to minimize electrical losses (heat generation), a critical length scale is the energy relaxation length, which defines the distance over which a non-equilibrium distribution of electrical carriers moving in an applied field or injected at a heterojunction exchanges sufficient energy by carrier and phonon scattering to return to a quasi-equilibrium distribution.

Among electronic device applications, the energy relaxation length is particularly critical for solid-state thermionic (TI) devices, in which ballistic carrier transport is utilized to increase the efficiency of power generation[12] or refrigeration.[13,14] In such devices, an interfacial energy filtering mechanism is used to block carriers with low energies ("cold carriers") and allow high-energy carries ("hot carriers") to pass through. For example, a conventional solid-state heterostructure TI cooler consists of three layers (cathode, barrier, and anode) in which the cathode/barrier junction presents a step potential that selectively transmits hot carriers[13]. Cold carriers accumulate at this junction and come into equilibrium by removing heat from the lattice, thus realizing solid-state cooling. Hot carriers transmitted into the barrier come into equilibrium over the energy relaxation

length, heating the lattice by phonon emission. For larger energy relaxation lengths, the barrier can be made thicker, increasing the thermal resistance through which this heat is transmitted back to the cold side and improving device efficiency.

CNTs have long been used to facilitate thermionic transmission of electrons into vacuum[15], and show promise for solid-state TI devices due to their long electron mean free path[16] (MFP) ($l_{CNT} \sim$ 1.6 um, compared to ~40 nm in doped InGaAs[17], a material commonly used for heterostructure TI coolers[13]), which suggests a long energy relaxation length. However, it is unclear to what extent the enhanced carrier ballistic transport intrinsic to CNTs outweighs their large thermal conductivity and enables high TI efficiency. To study solid-state thermionic transport of electrons within a CNT network, we use Monte Carlo (MC) simulation to examine energy relaxation of hot electrons in a CNT following carrier energy filtering at a CNT-CNT intertube contact.

4.2 CNT-based solid-state thermionic cooler

We consider a mesoscopic TI cooler composed of three (10, 0) single-walled carbon nanotubes (SWCNTs) connected head-to-tail via Van der Waals interactions (Fig. 4.1A). While interfacial band offsets are used for energy filtering in epitaxial heterostructure TI coolers[13,14], here we use energy-dependent electron transmission functions[18] to account for transport across intertube junctions in the SWCNT TI cooler.

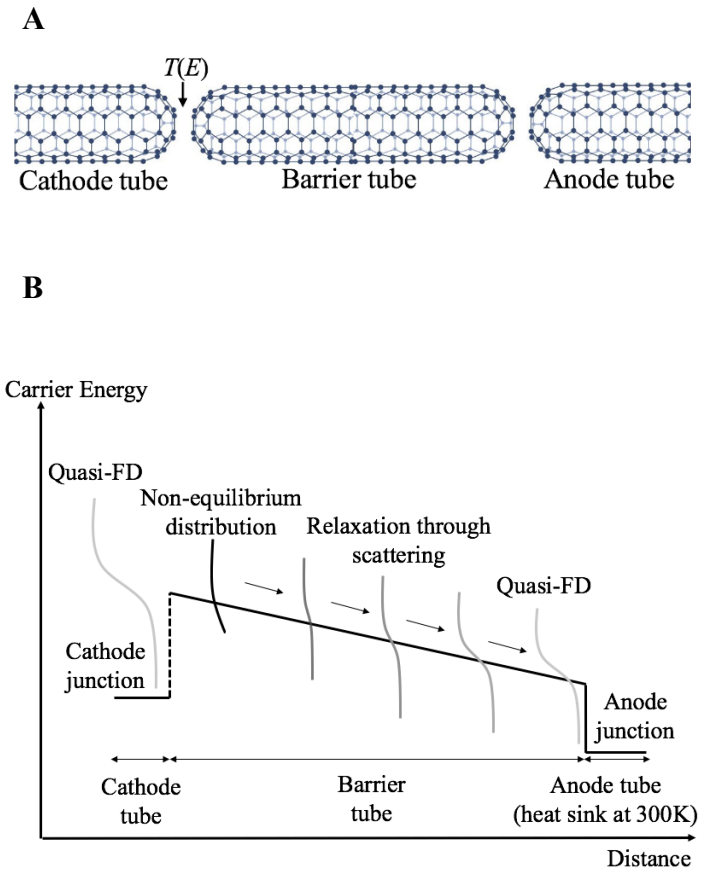

Figure 4.1 | CNT thermionic cooler. (A) Schematic of mesoscopic TI cooler structure consisting of three SWCNTs connected head-to-tail. **(B)** Evolution of carrier energy profile where the initial quasi-Fermi-Dirac (quasi-FD) distribution is filtered by the cathode-barrier junction, and the transmitted non-equilibrium distribution relaxes back to quasi-FD by scattering.

For computational convenience, we introduce $T_{prob}(E)$, which is the relative probability of an electron with energy E to cross the junction. This probability distribution is calculated by normalizing the VDW-bonded intertube junction transmission function $T_{jun}(E)$ (Fig. 4.2A) by the transmission function $T_{CNT}(E)$ of a covalently-bonded intertube contact (essentially a continuous SWCNT); the latter quantifies the number of conducting channels at a given energy[19,20] (Fig. 4.2B). $T_{jun}(E)$ for a contact between two (10, 0) SWCNTs is extracted from Ref. 18, and $T_{prob}(E)$ is normalized such that charge balance is maintained before and after transmission across the intertube junction. We note that $T_{prob}(E)$ gives the probability for an electron of a given energy to cross the junction relative to the probability of an electron with a different energy to cross the junction. Therefore, it can be used to calculate the shape of the carrier energy distribution that enters the barrier tube.

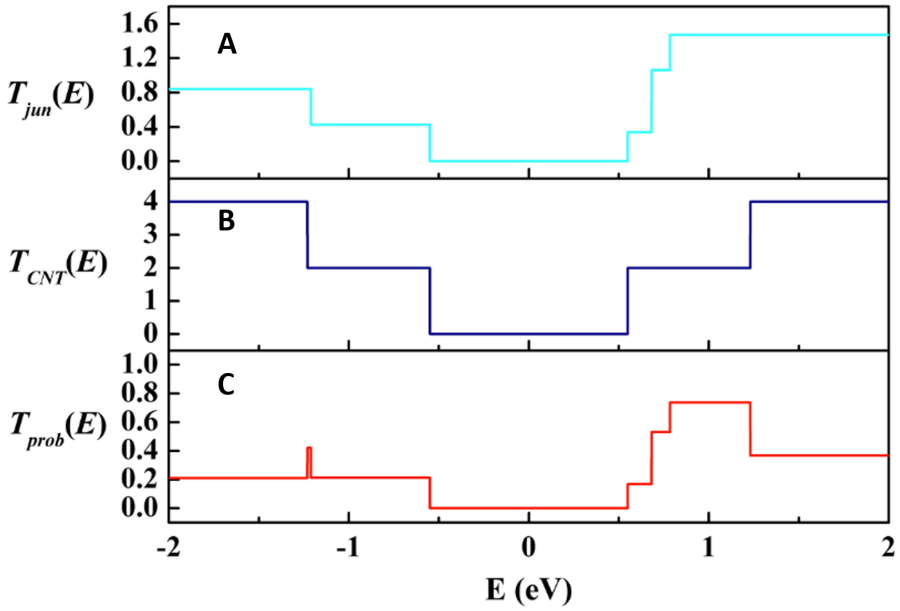

Figure 4.2 | Transmission function at intertube contacts. (A) Transmission for a VDW-bonded junction formed by two (10, 0) SWCNTs, **(B)** Transmission function for a perfect and pure (10, 0) SWCNT, **(C)** Normalized transmission probability of a junction between two (10, 0) SWCNTs.

As shown in Fig. 4.2C, the transmission probability $T_{prob}(E)$ shows several "steps" between 0.5 and 1.0 eV; placing the chemical potential near such steps could increase the asymmetry of the transmitted carrier energy distribution and hence the cooling capacity/energy conversion efficiency. Therefore, (10,0) *n*-type semiconductor SWCNTs with a chemical potential (0.70 eV) placed approximately one k_bT (T being the room temperature) below the step at 0.73 eV are chosen for MC simulation.

4.3 Calculation of the energy relaxation length for hot carriers

A Monte Carlo (MC) technique is implemented in the middle SWCNT ("barrier tube") to study the relaxation of transmitted hot electrons in this region. Scattering by acoustic and optical phonons (both emission and absorption) is considered; Pauli blocking is also taken into account by adding the Fermi occupation factor "1-f_0" (f_0 is the Fermi function) when calculating the scattering matrix, to ensure that the final state after scattering is available[21]. All calculations are performed at room temperature, and effects such as electron-electron scattering that are dominant at low temperatures[22] are neglected. Relevant SWCNT parameters and expressions for scattering rates are taken from prior work[21,23]. During simulation, the energies, momenta, and positions of electrons are calculated after each time step, and their energy-dependent scattering rates are

subsequently updated. This process is iterated until the carrier energy distribution returns to a quasi-Fermi-Dirac distribution in the barrier tube; for distances greater than this position, electrons deposit an amount of energy into the lattice equal to that which they gain from the external electric field. The transport of hot electrons is studied under different biases. We note that under low field conditions[19] (ε < 20 mV/μm), the electron mean free path (MFP) in SWCNTs has been found to be exceptionally long[15] (~1.6μm), which is beneficial for TI power generation/refrigeration applications.

After hot electrons are transmitted into the barrier SWCNT, they deposit heat through inelastic electron-phonon scattering. The net heat dissipated (transferred) from electrons to phonons as a function of distance traveled in the barrier tube is shown in Fig. 4.3. The distance at which heat dissipation becomes constant is the energy relaxation length (λ_E).

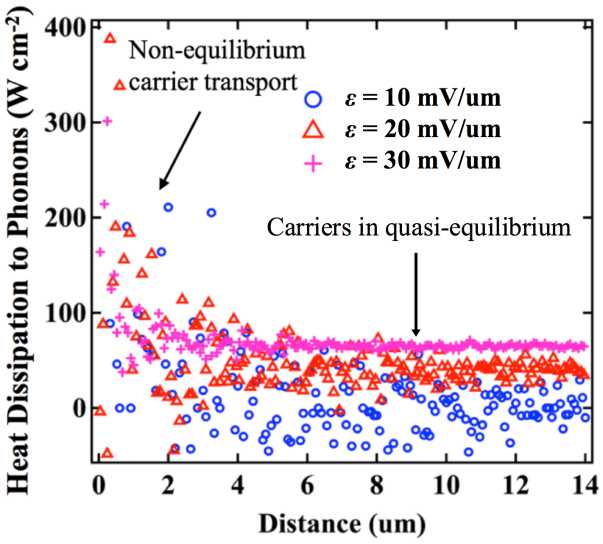

Figure 4.3 | Electron-phonon energy exchange in the barrier tube for three applied electric fields.

We find λ_E to be approximately 8 μm, which is nearly 5 times the MFP. This λ_E/MFP ratio is similar to that of typical III-V materials (for instance, hot electrons in doped InGaAs under room temperature have a MFP[16] of 44 nm while the energy relaxation length[24] is ~150 nm). The value for λ_E is also consistent with the typical relaxation time for photo-excited electrons in semiconducting SWCNTs of a similar diameter[25] multiplied by the Fermi velocity[19] (10ps × 10^6 m/s = 10 μm).

A quantity closely related to heat dissipation is the average electron energy E_{av} in the barrier tube, which undergoes a sharp increase due to junction filtering (i.e., hot electrons are more likely to be transmitted than cold electrons). The subsequent relaxation of E_{av} (Fig. 4) due to electron-phonon scattering has been studied under three applied electric fields: 10 mV/μm, 20 mV/μm and 30 mV/μm (high-field). Interestingly, we observe that under low fields, E_{av} decays more slowly as the field strength increases. In this low-field regime where ballistic/quasi-ballistic transport dominates, electrons travel greater distances (since drift velocity positively correlates with strength of external ε-field) before they attain enough energy to scatter with optical phonons. In contrast, once the bias exceeds a threshold value (~20 mV/μm for a semiconducting SWCNT[19]), transport enters the high-field regime in which transmitted electrons immediately gain sufficient energy to couple strongly with optical phonons, their scattering with which is inelastic. This leads to E_{av} in the high-field regime (30 mV/μm) decaying more rapidly than at 20 mV/μm.

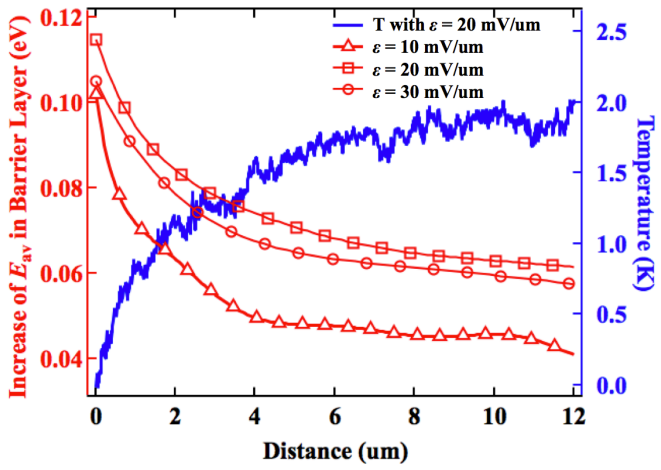

Figure 4.4 | Change in the mean carrier energy. The left y-axis shows the evolution of electron mean energy in the barrier layer normalized by electron energy prior to reaching the cathode junction as a function of distance traveled in the barrier tube. It can be seen that E_{av} decays more rapidly when transport regime transitions from low-field to high-field. The right y-axis represents the temperature profile (normalized by the cathode temperature) along the barrier tube with ε = 20 mV/μm.

Equally important to the electrical transport is the temperature distribution along the barrier tube, which can be obtained by tracking electron-phonon energy exchange. Since this energy exchange is essentially a heat source for phonons, the lattice temperature along the barrier tube (Fig.4.4) can be found by integrating this term over the entire barrier region (Eqn. 3.5) with suitable boundary conditions[24](heat sink at the anode with T_{anode} = 300K). Because low-field transport is most

relevant to thermionic applications, ε = 20 mV/μm is chosen for subsequent calculations. In this case, Fourier's law of heat diffusion applies, as hot phonon effects are negligible[24]:

$$\int \Delta Q_{phonon} dz = -\kappa \int \Delta T \, dz \tag{4.1}$$

Energy filtering at the SWCNT junction is clearly visible in Fig. 4 as a sharp increase of 0.115 eV in E_{av} relative to its value in the cathode tube at the cathode-barrier junction (which is set to zero). The effective Seebeck coefficient (S_{eff}) of a TI cooler associated with such a carrier average energy enhancement can be estimated as[24]:

$$\left|S_{eff}\right| = \frac{\Delta E_{av}}{eT} \sim 386 \ \mu\text{V/K} \tag{4.2}$$

where T represents the temperature in the barrier tube at the cathode-barrier junction. Our simulation therefore predicts that a highly asymmetric electron energy distribution due to junction filtering could significantly enhance the Seebeck coefficient beyond even the value of ~180 μV/K measured experimentally for oxygen-doped SWCNT mats[26].

In addition to large S_{eff} that promotes significant TI cooling power, long energy relaxation length is also critical for suppressing parasitic reverse heat current. As electrons traverse the barrier tube quasi-ballistically, they deposit heat via phonon emission and eventually establish a reverse temperature gradient from anode to cathode. This parasitic heat leakage is conducted back to the cathode, reducing the cooling efficiency. As λ_E increases, the optimal barrier tube length for maximized TI efficiency also grows (as shown in the calculation of cooling capacity Q_{TI} below), resulting in greater thermal resistance for this reverse heat conduction.

4.4 Cooling capacity of the TI cooler

To calculate the cooling capacity of the device, an energy balance in the barrier tube is invoked[13]:

$$Q_{TI} = \Delta E_{av} I - \left[\left(\frac{1}{2} - \frac{\lambda_E}{L} \right) - \frac{\lambda_E^{\,2}}{L^2} \left(e^{-L/\lambda_E} - 1 \right) \right] IV - \frac{\kappa}{L} \Delta T \tag{4.3}$$

where I, κ and L are the current density, thermal conductivity, and barrier tube length, respectively. The calculation of current density includes contributions from both electron thermionic emission and drift-diffusion[13]. The SWCNT thermal conductivity is length-dependent and has previously been both calculated[27] and measured[28,29]. The electronic contribution to thermal conductivity is neglected as it is secondary to the phononic contribution in the chosen semiconducting SWCNT[19].

The three terms on the right-hand side of Eq. 3.7 represent TI cooling at the cathode junction, Joule heating (Q_{Joule}) for non-equilibrium electrical transport, and reverse heat conduction (Q_{cond}) from anode to cathode, respectively. Since TI cooling is the only term that scales linearly with current density, by moving the last two terms to the left-hand side of Eqn. 3.3 and plotting the sum against I (Fig. 4.5), we can extract S_{eff} from the slope of the curve. S_{eff} of 366 μV/K in the ballistic thermionic limit and 37 μV/K in the diffusive thermoelectric limit are obtained (inset of Fig. 4.5). The ballistic S_{eff} agrees well with the estimate derived in Eq. 2, while the diffusive S_{eff} matches closely with the Seebeck coefficient predicted for an infinitely long SWCNT with a chemical potential of 0.7 eV (~42 μV/K).

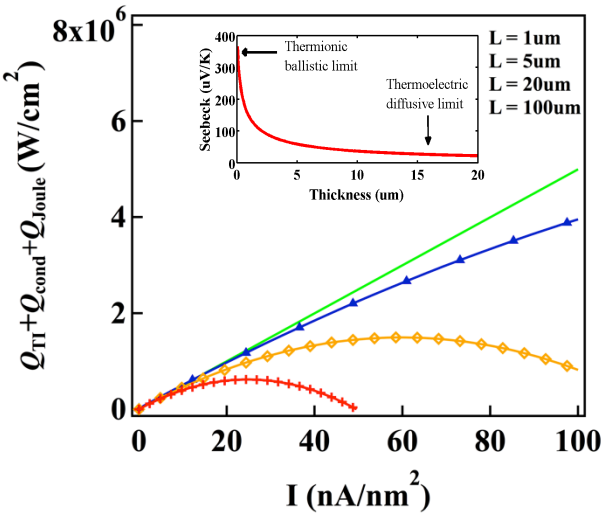

Figure 4.5 | The sum of cooling capacity, reverse heat load, and Joule heating as a function of current density. The slopes of the curves scale linearly with effective Seebeck coefficient. The inset shows the effective Seebeck coefficient as a function of barrier tube length.

The only remaining unknown parameter in Eq. 3.7 is the barrier tube length (L), a parameter that must be optimized to maximize cooling power. To this end, Q_{TI} is plotted versus L (Fig. 4.6). A constant electric field has been assumed for all barrier tube lengths and the length-dependent thermal conductivity of (10, 0) SWCNTs is adopted from ref. 27. Eq. 3.7 suggests that the total cooling capacity depends heavily on the interplay between Joule heating and reverse heat conduction, since both depend on L. As L increases, so does Joule heating, since more hot electrons travel a distance sufficient for energy relaxation. This can be seen in the expression for Joule heating in Eq. 3.7. On the other hand, as L decreases, the enhancement in reverse temperature gradient overwhelms the reduction in κ (as κ of the chosen SWCNT drops as L lessens), leading to an overall increase in reverse heat conduction. As a result, we expect to find a minimum for the sum of the two at an intermediate barrier tube length.

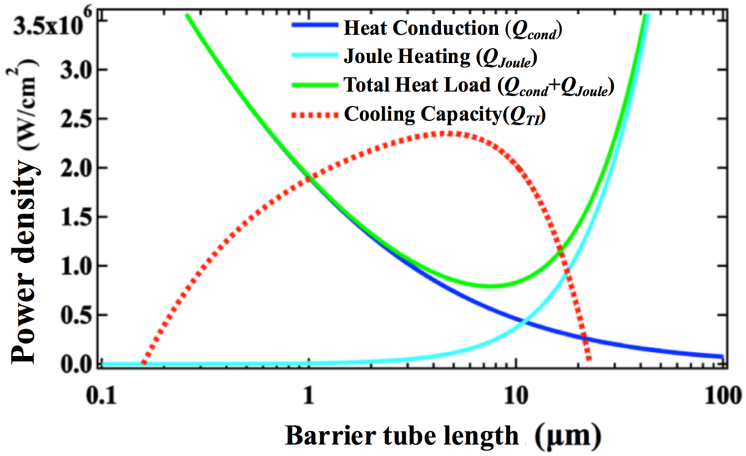

Figure 4.6 | Cooling capacity, reverse heat load, Joule heating, and total heat load as a function of barrier thickness. A minimum in total heat load and maximum in cooling capacity both occur at $L \sim 8\mu m$.

A minimum in total heat load is achieved at $L \sim \lambda_E$, which also corresponds to maximum cooling capacity. This highlights the important role that a long energy relaxation length plays in suppressing reverse heat conduction. We note that the SWCNT cross-sectional area is calculated using an annular shell with an outer diameter equal to the SWCNT diameter and shell width equal to the inter-graphene layer distance in graphite[19] (3.4 Å). The resultant Q_{TI} is higher than that of conventional solid-state heterostructure TI coolers[14,24], suggesting that SWCNT-based devices show promise for TI cooling applications.

4.5 Discussion and Summary

In summary, we have investigated thermionic energy filtering at SWCNT intertube junctions and find it to be effective at filtering low energy carriers, providing an elevated electron energy profile (ΔE_{av} = 0.115eV). The energy relaxation length for hot electrons is obtained via MC simulation and found to be quite long (~8 μm), making such SWCNTs promising for TI cooling and energy conversion applications. The optimal barrier thickness for a SWCNT TI cooler is found to be similar to the energy relaxation length, and our calculations suggest that its cooling capacity can be significant. In order to obtain a practical total cooling power, it could be promising to use vertically aligned SWCNTs assembled in parallel (given recent advancements in synthesizing high-quality vertical SWCNT bundles[30]) or SWCNT networks[31] to scale up the cross-sectional area.

References

1. T. T. Ebbesen, T. W., Lezec H. J., Hiura H., Bennett J. W., Ghaemi H. F. Electrical conductivity of individual carbo nanotubes. *Nature* **382**, 54 (1996).

2. S. A. Getty, E. Cobas, and M. S. Fuhrer. Extraordinary mobility in semiconducting carbon nanotubes. *Nano Lett.* **4**, 35-39 (2004).

3. S. Berber, Y. Kwon, and D. Tománek. Unusually high thermal conductivity of carbon nanotubes. *Phys Rev Lett.* **84**, 4613-4616 (2000).

4. E. Pop, D. Mann, Q. Wang, K. Goodson, and H. Dai. Thermal conductance of an individual single-wall carbon nanotube above room temperature. *Nano Lett.* **6**, 96-100 (2005).

5. G. F. Close, S. Yasuda, B. Paul, S. Fujita, and H. S. P. Wong. A 1 GHz integrated circuit with carbon nanotube interconnects and silicon transistors. *Nano Lett.* **8**, 706-709(2008).

6. P. Avouris, Z. Chen, and V. Perebeinos. Carbon-bsed electronics. *Nat. Nanotechnol*. **2**, 605-615 (2007).

7. T. Wang, K. Jeppson, L. Ye, and J. Liu. Carbon-nanotube through-silicon via interconnects for three-dimensional integration. *Small* **7** (2011).

8. J. A. Misewich, R. Martel, Ph. Avouris, J. C. Tsang, S. Heinze, J. Tersoff. Electrically induced optical emission from a carbon nanotube FET. *Science* **300**, 783-786 (2003).

9. A. Javey, J. Guo, Q. Wang, M. Lundstrom, and H. Dai. Ballistic carbon nanotube field-effect transistors. *Nature* **424**, 654-657 (2003).

10. R. Martel, T. Schmidt, H. R. Shea, T. Hertel, and Ph. Avouris. Single- and multi-wall carbon nanotube field-effect transistors. *Appl. Phys. Lett.* **73** (1998).

11. K. Zhang and M. M. F. Yuen, *7th International Conference on Electronics Packaging Technology* (2006).

12. D. Vashaee and A. Shakouri. Thermionic power generation at high temperatures using SiGe/Si superlattices. *J. Appl. Phys.* **101** (2007).

13. A. Shakouri, E. Y. Lee, D. L. Smith, V. Narayanamurti and J. E. Bowers. Thermoelectric effects in submicron heterostructure barriers. *Microscale Thermophys. Eng.* **2**, 37-47 (1998).

14. A. Shakouri and J. E. Bowers. Heterostructure integrated thermionic coolers. *Appl. Phys. Lett.* **71**, 1234-1236 (1997).

15. P. Liu, Y. Wei, K. Jiang, Q. Sun, X. Zhang, S. Fan, S. Zhang, C. Ning, and J. Deng. Thermionic emission and work function of multiwalled carbon nanotube yarns. *Phys. Rev. B* **73** (2006).

16. J. Y. Park, S. Rosenblatt, Y. Yaish, V. Sazonova, H. Üstünel, S. Braig, T. A. Arias, P. W. Brouwer, and P. L. McEuen. Electron-phonon scattering in metallic single-walled carbon nanotubes. *Nano Lett.* **4**, 517-520 (2004).

17. R. Teissier, J.-L. Pelouard, and F. Mollot. Measurement of electronic ballistic length in p-doped InGaAs. *Phys. Statu. Solidi B* **204** (1997).

18. K. Esfarjani, M. Zebarjadi, and Y. Kawazoe. Thermoelectric properties of a nanocontact made of two-capped single-wall carbon nanotubes calculated within the tight-binding approximation. *Phys. Rev. B* **73** (2006).

19. H. P. Wong and D. Akinwande, *Carbon Nanotube and Graphene Device Physics*. Cambridge University Press, New York (2011).

20. R. Nizam, S. M. A. Rizvi, and A. Azam. Calculating the electronic transmission properties of semiconducting carbon nanotube Schottky diodes with increase in diameter. *Semi. Phys. Quantum Elec. & Optoelec* **15**, 268-275 (2012).

21. A. Liao and E. Pop. Multiband mobility in semiconducting carbon nanotubes. *IEEE Electron Device Lett.* **30**, 1078-1080 (2009).

22. D. L. Maslov and M. Stone. Landauer conductance of Luttinger liquids with leads. *Phys. Rev. B* **52** (1995).

23. G. Pennington and N. Goldsman. Low-field semiclassical carrier transport in semiconducting carbon nanotubes. *Phys. Rev. B* **71** (2005).

24. M. Zebarjadi, A. Shakouri, and K. Esfarjani. Thermoelectric transport perpendicular to thin-film heterostructures calculated using the Monte Carlo technique. *Phys. Rev. B* **74** (2006).

25. O. A. Dyatlova, C. Koehler, P. Vogel, E. Malic, R. M. Jain, K. C. Tvrdy, M. S. Strano, A. Knorr, and U. Woggon. Relaxation dynamics of carbon nanotubes of enriched chiralities *Phys. Rev. B* **90**, 155402 (2014).

26. Y. Nakai, K. Honda, and K. Yanagi. Giant Seebeck coefficient in semiconducting single-wall carbon nanotube film. *Appl. Phys. Express* **7** (2014).

27. N. Mingo and D. A. Broido. Length Dependence of Carbon Nanotube Thermal Conductivity and the "Problem of Long Waves". *Nano Lett.* **5**, 1221-1225 (2005).

28. Z. L. Wang, D. W. Tang, X. B. Li, X. H. Zheng, W. G. Zhang, L. X. Zheng, Y. T. Zhu, A. Z. Jin, H. F. Yang, and C. Z. Gu. Length-dependent thermal conductivity of an individual single-wall carbon nanotube. *Appl. Phys. Lett.* **91** (2007).

29. Z. Wang, D. Tang, X. Zheng, W. Zhang, and Y. Zhu. Length-dependent thermal conductivity of single-wall carbon nanotubes: prediction and measurements. *Nanotechnology* **18** (2007).

30. S. J. Kang, C. Kocabas, T. Ozel, M. Shim, N. Pimparkar, M. A. Alam, S. V Rotkin, and J. A. Rogers. High-performance electronics using dense, perfectly aligned arrays of single-walled carbon nanotubes. *Nat. Nano.* **2**, 230-236 (2007).

31. H. Dai. Carbon Nanotubes: Synthesis, Integration, and Properties. *Acc. Chem. Res.* **35**, 1035-1044 (2002).

Chapter 5
Thermal Management at the Package Level – Electrostatically Extended Polymer with High Thermal Conductivity

Science Advances **3**, 1700342 (2017)

5.1 Introduction

The rising thermal challenge in high-power electronic and optoelectronic devices provides a strong incentive to develop novel E&O packages and heatsinks, from the fundamental material level all the way up to packaging architecture. In this part of the thesis we present a study that showcases a new route to enhancing the thermal conductivity of amorphous polymers, making it a promising candidate for novel thermal management materials.

As we have established in Chapter 1, effective thermal management in applications such as batteries, automobile cooling systems and high-power density electronic devices where heat accumulation can have deleterious effects is critically important to ensure system performance and reliability and enhance lifetime. Despite their poor thermal conductivity (κ), various advantages including light weight, low cost and easy processability make polymers the material of choice for several heat intensive applications like electronic chip encapsulation, cellphone casing, LED housing, etc. These existing applications along with emerging technologies such as flexible electronics, for which the requirements on flexibility and light weight simply cannot be met by most conventional thermal management materials (metals and ceramics), put greater technological incentives on developing thermally conductive polymers.

Blending with high- κ fillers such as metal or ceramic particles, carbon nanotubes (CNTs), or graphene flakes is the most commonly used method to enhance polymers' thermal conductivity[1]. However, the large volume fraction of fillers required to achieve appreciable enhancement in κ often leads to undesired optical or electrical properties, increased weight, high cost

(e.g., CNT: ~$1000/kg vs. PMMA: ~$2/kg), or loss of the easy processability generally associated with polymers.

In contrast to low κ in bulk samples, constituent individual polymer chains are believed to have very large κ. The thermal conductivity of a single polymer chain, in which the elastic disorder[2] between intra-chain covalent and inter-chain van der Waals bonds is absent, was calculated to be as large as few hundreds of $Wm^{-1}K^{-1}$.[4] Ultra-drawn crystalline nano-fibers with aligned polymer chains were measured to have κ over 100 $Wm^{-1}K^{-1}$ in the alignment direction[4]. The large thermal conductivities of single or few-chain fibers can be retained in amorphous polymers in the direction of chain orientation[5,6] along which heat propagation occurs predominantly through intra-chain transport. Singh et al. reported a significant increase in κ in amorphous polythiophene fabricated via a nanotemplate-assisted electrochemical method that allows polythiophene chains to be oriented in the vertical direction[7]. Thermal conductivity up to 2 $Wm^{-1}K^{-1}$ has been similarly reported for covalently grafted poly(3-methyl thiophene) brushes. Covalent grafting led to enhanced chain alignment as well as reduction in energetic and positional disorder in such surface-grafted films[8]. These high thermal conductivities reported in polymers with extended chain conformation stand in contrast to surface-grown polymer brushes[9] and polymer films under high pressure[10], in which the coiled conformation of polymer chains likely remained and enhancement in κ was found to be relatively moderate. However, these approaches either limit the orientation of chain extension to a certain direction or pose challenges in terms of scaling-up the nanoscale films for practical applications. Therefore, it's desirable to achieve high κ in both in- and out-of-plane directions in bulk amorphous polymers using common fabrication processes.

While the mechanisms of thermal transport in amorphous materials continue to be studied[10-13], it is generally believed that the thermal conductivity in bulk amorphous polymers (a class of disordered solids) is inhibited by: a) highly coiled and entangled intra-chain structure, b) loose chain packing with voids that dampen the speed at which vibrations propagate, and c) weak non-bonding inter-chain interactions (e.g., van der Waals, dipole-dipole)[14]. In this work, we demonstrate an unexplored molecular engineering route that attack these three bottlenecks simultaneously. By employing the coulombic repulsive forces between ionized pendant groups on the backbone of polyelectrolytes to "stretch" the main chain at the molecular level, we achieve significant enhancements of thermal conductivity in amorphous polymer with randomly oriented

yet superiorly packed extended polymer chains and strong ionic inter-chain interactions (Fig. 5.1A). Moreover, the favorable conformation change is accomplished in a thermodynamically driven process, making it amenable for scale-up.

5.2 Experimental design

The objective of this study was to recognize the structural bottlenecks to high thermal conductivities in amorphous polymers, and design and develop a polymeric material to address them. Two water soluble polymers with and without ionizable pendant groups were chosen to compare and contrast the effects of polymer ionization on their thermal transport properties. Material evaluation involved thermal conductivity measurement on spin-cast thin films and materials characterization in terms of degree of polymer ionization, viscosity, film density and elastic modulus. The materials properties were correlated with the measured thermal conductivity of the polymers. Morphological and theoretical analyses were done to rule out any extraneous contribution to measured high thermal conductivities. Additional salt composite films and thick films were fabricated to highlight the efficacy of the reported molecular design strategy to achieve high thermal conductivities in amorphous polymer films.

To test our idea, we used a weak polyelectrolyte, polyacrylic acid (PAA, atactic), made up of a C-C backbone with a carboxylic acid (-COOH) group at alternate carbon atoms that can be ionized to a carboxylate ($-COO^-$) by addition of a base, i.e., with increase of the polymer solution pH. The close proximity of the densely packed ionizable groups to the polymer main chain allows the effect of electrostatic repulsion between them to easily translate to the backbone, resulting in chain extension with increasing ionization. As a negative control, a water soluble polymer, poly(*N*-vinyl pyrrolidone) (PVP, atactic), without an ionizable pendant group, was employed. To fabricate thin films for thermal conductivity measurement, the polymer was dissolved in de-ionized (DI) water, and the pH of the solution was adjusted to the desired value by addition of 1M hydrochloric acid (HCl) or 1M sodium hydroxide (NaOH) solution. The final polymer concentrations were 0.5 wt.% and 1 wt.% for PAA and PVP, respectively. Polymer solutions were then spin-cast on a Si wafer with a ~100 nm oxide layer and annealed to obtain smooth films with thickness in the range of 10-35 nm. Cross-plane thermal conductivities of the polymer films were measured by a differential 3ω method, which is a standard technique for such measurements in films with thicknesses as small as few nanometers[15,16]. As shown in Fig. 5.1B, the thermal conductivity of PAA increased

from 0.34 ± 0.04 $Wm^{-1}K^{-1}$ at pH 1 when the PAA chains are completely unionized to 1.17 ± 0.19 $Wm^{-1}K^{-1}$ at pH 12 when the PAA chains are predominantly ionized (>90%, *vide infra*). The thermal conductivity of PVP, however, measured ~0.2 $Wm^{-1}K^{-1}$ across the entire pH range, consistent with its non-electrolyte nature.

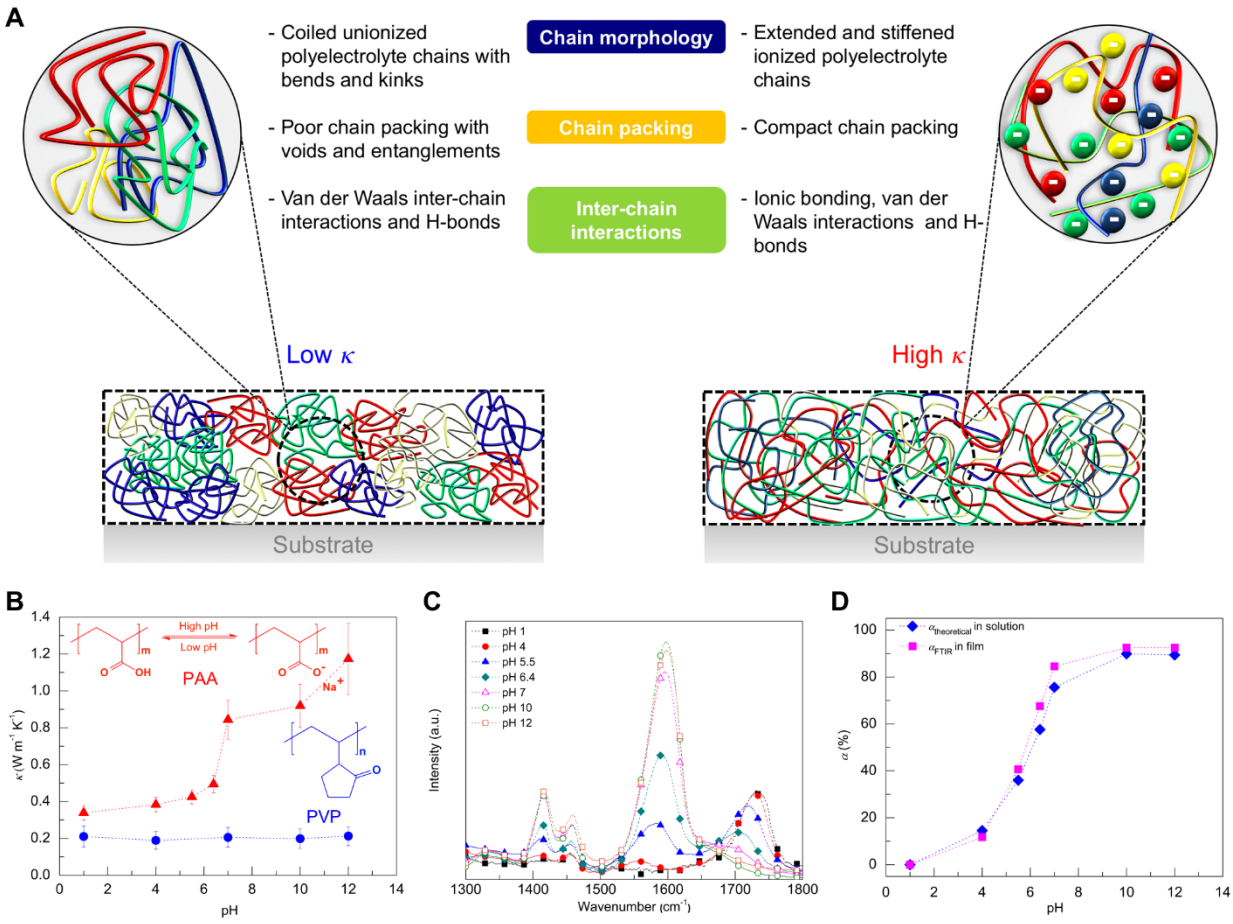

Figure 5.1 | High thermal conductivity in polyelectrolyte thin films via controlled ionization. (A) Illustrations of chain conformation and packing in spin-cast polymer films: coiled unionized polyelectrolyte (*left*) and extended ionized polyelectrolyte (*right*). The zoomed in images show chain confirmation at the molecular level. (**B)** Cross-plane thermal conductivity of a weak polyelectrolyte, PAA (MW 100kDa), and a non-ionizable water soluble polymer, PVP (MW 40kDa) thin films spin-cast from polymer solutions of different pH. Error bars were calculated based on uncertainties in film thickness, temperature coefficient of electrical resistance for the heater, and heater width. Chemical structures of the polymers and ionization reaction for PAA are also shown. **(C)** Fourier transform infrared (FTIR) spectra of PAA films spin-cast from solution of different pH. **(D)** Fraction of ionized carboxylic acid groups (α) as a function of solution pH: calculated from the FTIR spectra and by applying charge balance on PAA solutions.

5.3 Sample preparation, thermal conductivity measurement, and polymer characterization

To prepare samples for 3ω measurement, a part of the spin-coated or blade-coated film was removed using a steel blade, and the cleared area was further cleaned by a cotton swab dipped in water and ethanol to give a clean polymer-free reference region. Thin heater lines were deposited using a shadow mask (50 μm) and electron beam deposition (5 nm Ti / 200 nm Au) on both sample and reference region. To test the potential dependence of measured thermal conductivity on ambient humidity, a second series of spin-coated samples also included a 50 nm thick alumina capping layer that was sputter-coated on the polymer layer before the metal heater lines were deposited. Error analyses for 3ω data are described in later subsections in this Chapter (5.8). The film thicknesses of spin-cast films were measured by ellipsometry (Woollam M-2000DI Ellipsometer) and profilometry (Dektak XT Surface Profilometer). Only the latter was used for the blade-coated films because of their increased surface roughness.

Fourier transform infrared (FTIR) spectroscopy was used to confirm and quantify the extent of ionization of PAA. Figure 5.1C shows the FTIR spectra of PAA films spin-cast from solutions of different pH. The decrease in intensity of the carbonyl (–C=O) stretching band (1680-1730 cm^{-1}) of the carboxylic acid (-COOH) group with pH and concomitant increase in the intensity of the asymmetric carboxylate ($-COO^-$) stretching band (1556-1594 cm^{-1}) indicate ionization of the PAA chains[17]. The degree of ionization (α) of PAA as a function of solution pH (Fig. 5.1D), calculated from the areal ratio of peaks corresponding to ionized and unionized acidic groups fitted assuming Gaussian distributions and the same extinction coefficient for the two bands[18], matches with previously reported trends[17]. A theoretical charge balance calculation for the PAA solution yielded similar values for α confirming that PAA retains its ionization in the thin film. The FTIR spectra of PVP at different values of pH are nearly identical, consistent with its non-electrolyte nature. (Fig. 5.2A).

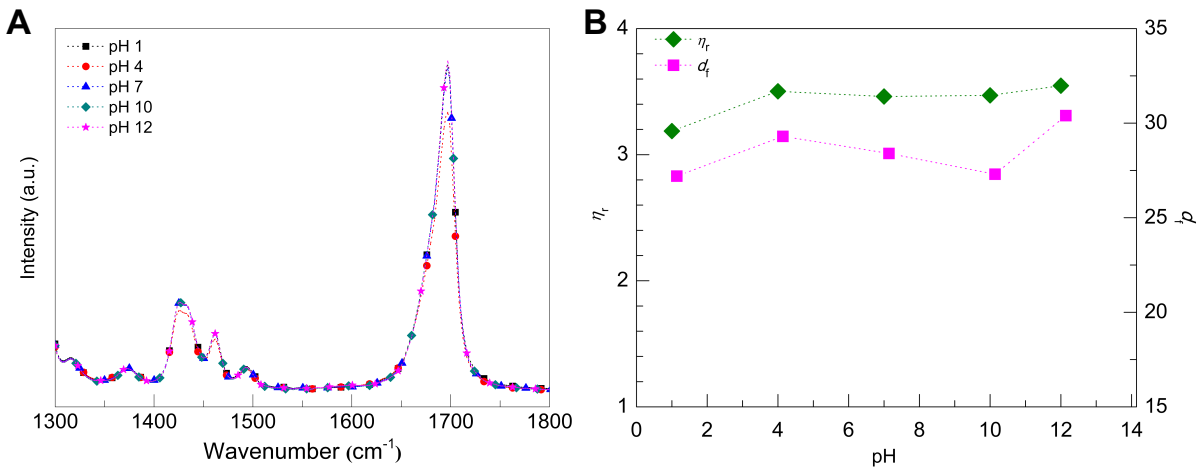

Figure 5.2 | Characterization of PVP films and solutions. (A) FTIR spectra of spin-cast PVP films as a function of pH. **(B)** Relative viscosities of 8wt.% aqueous solutions of PVP and film thicknesses for the spin-cast samples shown in Fig. 3.1B. Consistent with its non-electrolyte nature, FTIR spectra at different pH are nearly unchanged for PVP. Viscosities and film thicknesses remain unchanged across the entire pH range as well.

We further measured viscosities of PAA solutions at different pH, and porosities and elastic moduli of PAA thin films fabricated from solutions at different pH to quantify the three ionization-induced effects, viz., polymer chain extension[19], chain packing[19] and chain stiffening[20], respectively. As shown in Fig. 5.3A, the relative viscosity, η_r (= $\eta_{polymer}$ / η_{water}, η_{water} = 10^{-3} Pa.s), increases with solution pH, indicating that coulombic repulsion between ionized carboxylic acid groups stretches out the PAA chains, resulting in an extended morphology and hence increased solution viscosity[21]. Under the same spin-casting conditions, the trend in film thickness (d_f) matches well with that of solution viscosity, suggesting that the extended conformations of PAA chains in solution are likely preserved in the thin films. We note that it is likely that upon spin-casting the Na^+ ions condense close to the negatively-charged carboxylate pendant groups of PAA to maintain charge neutrality, reducing the degree of chain extension in the thin films. However, a previous AFM study on spin-cast samples of a brush polymer with grafted PAA side chains demonstrated systematic extension in chain morphology as pH increased[22], corroborating our assertion that at least some level of chain extension is preserved in the solid-state films. In contrast, the viscosities of PVP solutions as well as the film thicknesses for spin-cast PVP samples remained unchanged across the pH range as expected (Fig. 5.2B).

Since polymer density in confined films is known to differ from the bulk value and depends on the film thickness[23], we used Positronium annihilation lifetime spectroscopy (PALS)[24] to assess the change in density of PAA films with pH in order to clarify the potential density-related contributions to thermal conductivity. PALS data (Fig. 5.3B) shows a linear decrease in PAA film porosity with pH, which can be explained by better chain packing in the spin-cast thin films afforded by extended ionized chains. The measured ~33% drop in film porosity from pH 4 to pH 12 was consistent with the trend of ionization-dependent bulk density for partially ionized PAA[24.] Due to the inherent difficulty of performing nano-indentation[26,27] on nano-scale spin-cast films, micrometer-thick blade-coated PAA (MW 450kDa, atactic) films[28] were used for elastic modulus (E) measurement. As shown in Fig. 5.3C, elastic modulus increased with pH, i.e., with ionization of PAA chains. A similar chain stiffening effect caused by reduced chain segmental mobility due to strong ionic interactions between the negatively charged polymer chains and the surrounding positive cationic coordination sphere[29] is generally attributed for the large increase in glass transition temperature (T_g) of PAA with ionization[30].

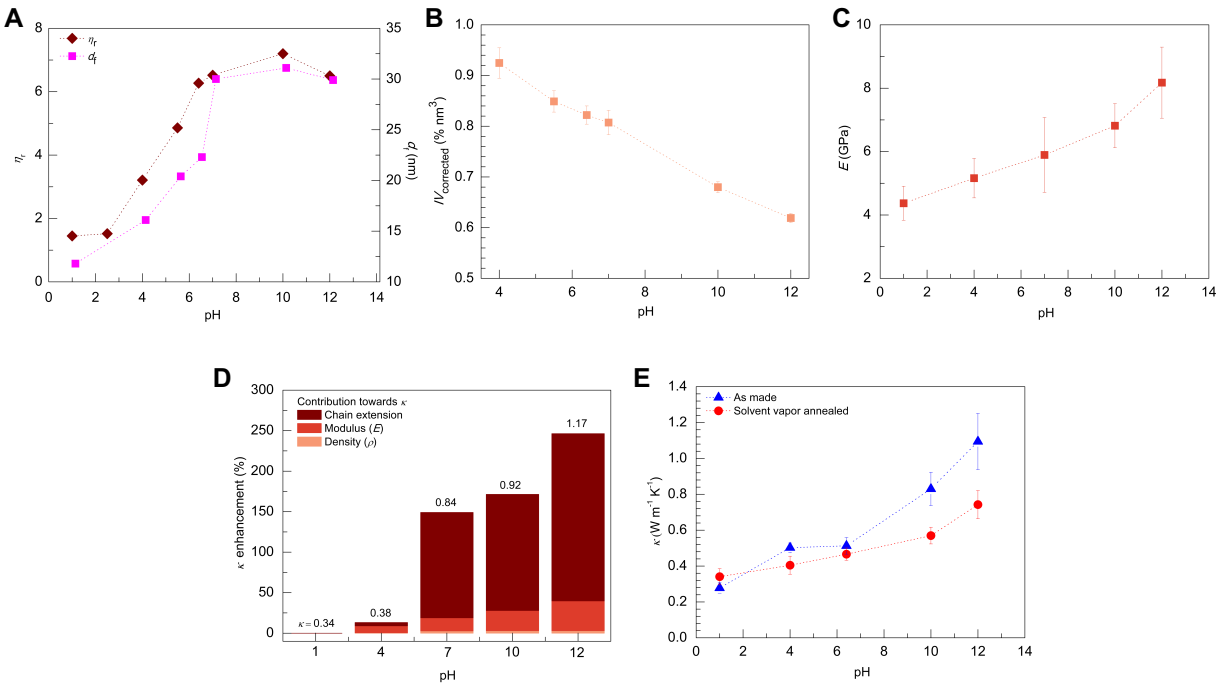

Figure 5.3 | Effects of PAA ionization and their contributions towards enhancement in κ. (A) Relative viscosity, η_r (= $\eta_{polymer}/\eta_{water}$, η_{water} = 10^{-3} Pa.s), of a 2 wt.% solution of PAA, and film thickness, d_f, of spin-cast samples (from 0.5wt.% solution) as a function of pH. **(B)** Positronium annihilation lifetime spectroscopy data for PAA films at different pH. The change in the product of positronium (Ps) intensity (*I*, %) and pore volume (*V*, in nm^3) represents the change in film porosity. Error bars were estimated based on the errors from fitting *I* and *V* plus an estimate of the error in the positron transmission correction. **(C)** Elastic modulus of blade-coated PAA (MW 450kDa) films measured by nano-indentation. The error bar shows standard deviation of measurements at four different points on the film. **(D)** Contributions from ionization effects shown in A-C towards enhancement in thermal conductivity of spin-cast PAA films. κ at different pH is noted above the bars. (E) Thermal conductivities of solvent vapor-annealed PAA films compared to those of as-made samples. PAA films were solvent-vapor annealed at 90°C for 30 minutes followed by annealing at 100°C for 15 minutes.

Furthermore, grazing-incidence x-ray diffraction measurements carried out on spin-cast PAA films did not show any sign of polymer crystallinity (Fig. 5.4) thereby ruling out any crystallinity-related contribution to the measured thermal conductivity. A broad diffused peak (known as the amorphous halo) has been seen for all samples, which is characteristic of amorphous polymers including PAA[31] and PAA salts[32].

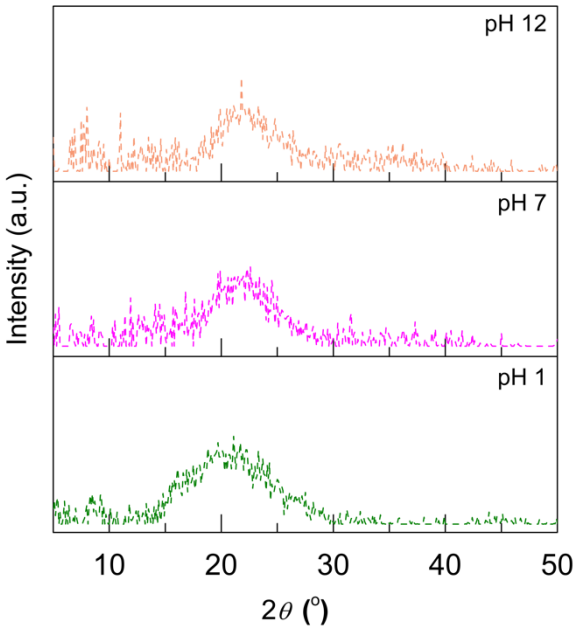

Figure 5.4 | Grazing-incidence x-ray diffraction spectra of PAA films at different pH. The broad diffused peak from ~15°-30°, called amorphous halo, is characteristic of amorphous polymers including PAA[31] and PAA salts[32].

5.4 Film morphology and contributions from NaOH inclusions

In order to rule out the possible contributions of NaOH crystals that could potentially act as high-κ fillers to measured thermal conductivity, we calculated the maximum possible volume fraction of NaOH crystals (V_{NaOH}) in the resulting polymer films. Such NaOH crystals formed from residual Na^+ and OH^- ions in the polymer solution could act as high-κ fillers and contribute to thermal conductivity enhancement. However, based on the known amount of PAA and NaOH added in the polymer solutions at various pH, the calculated V_{NaOH} was found to be negligible (V_{NaOH} ~1.65% for the highest pH 12 sample) except for an additional sample (not shown in the data of Fig. 5.1B) for which excess NaOH was added specifically to probe the potential contribution of NaOH crystals (shown in Figs. 5.5A and 5.5B). While crystals were not observed in the samples except

for the one with excess NaOH, a Maxwell model was nevertheless used to predict the thermal conductivities that would be expected if the NaOH crystals were homogeneously distributed within the film as nano-sized spherical fillers[33], for comparison with values measured for chain-extended PAA films. A volumetric percolation threshold equal to 25% is required for appreciable κ enhancement in such composites; this is not reached even for the highest pH PAA film (pH 12; V_{NaOH} = 1.65%). The fact that the measured PAA thermal conductivities are significantly greater than Maxwell-predicted values (Fig. 5.5C) indicates that κ enhancement due to ionization-induced effects dominates over the possible contributions of high-κ fillers over the range of pH selected.

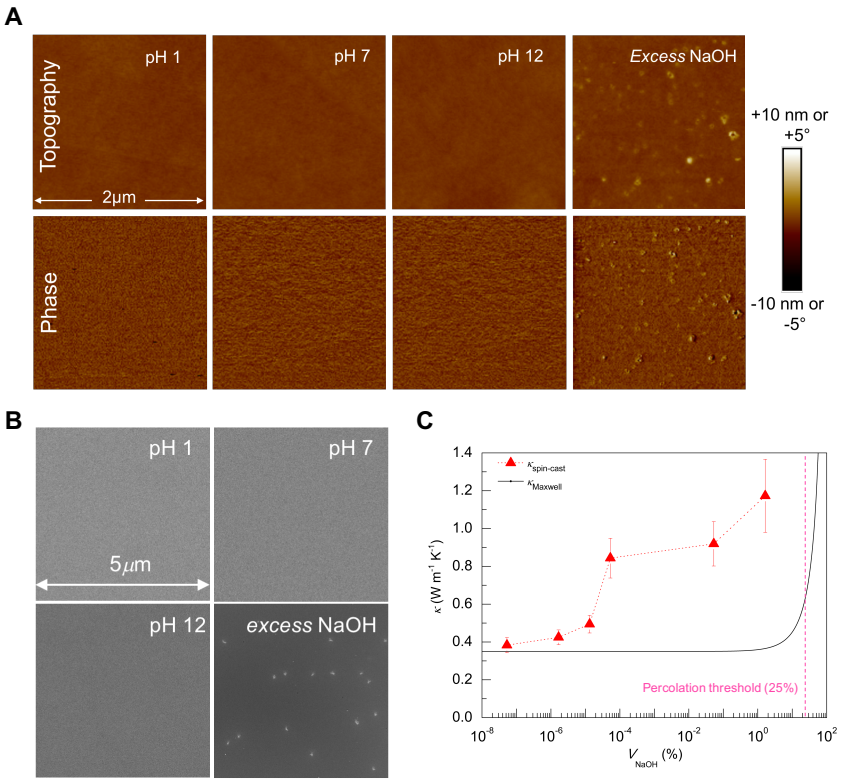

Figure 5.5 | Tapping-mode AFM and SEM analyses of PAA films. (A) Tapping-mode topography (top) and phase (bottom) images (2 μm x 2 μm) of PAA films spin-cast from solutions of different pH. AFM images have been shifted to zero mean values (i.e." flattened") for illustration purposes. Nano-sized NaOH crystals are only visible in sample with excess amount of NaOH added to the PAA solution. **(B)** SEM images of the same films analyzed by AFM. NaOH crystals can be seen only when excess NaOH is added, consistent with the AFM data. **(C)** Measured thermal conductivities, $\kappa_{spin\text{-}cast}$, for spin-cast films greatly exceeds the Maxwell-model predicted values, indicating enhancement is not primarily due to a high-κ filler effect.

Tapping-mode atomic force microscopy (AFM) and scanning electron microscopy (SEM) analyses of the PAA films further corroborate the theoretical calculation of V_{NaOH}. As can be seen in Fig. 5.5A, AFM topography images show a smooth featureless film surface morphology for values of pH up to 12. Small spherical NaOH crystals can be seen in the thin film spin-cast from polymer solution with excess NaOH added. SEM images were used to confirm surface morphology as well as investigate the potential presence of NaOH crystals buried within the film (Fig. 5.5B). We did not observe any sign of NaOH crystals except for the sample with excess NaOH added, which is consistent with the AFM topography images.

To investigate whether preferential crystallization of NaOH occurs in the polymer film beneath the gold heater lines (i.e., metal-polymer interface) due to heterogeneous nucleation[34], focused ion beam (FIB)-assisted SEM (FEI Nova Nanolab 200 FIB/SEM) was used to image the cross-section of a pH 10 sample beneath the gold heater line (Fig. 5.6). Focused ion beam (FIB) was used to etch out a 20 μm long and 10 μm deep indentation (Fig. 5.6, left panel) on a heater line, exposing a sharp cross-section centered around the boundary between sample and reference region comprising two Pt protective layers, 200nm thick gold heater line covered with a thin sputter-coated gold layer for SEM imaging, PAA film, SiO_2 layer and Si substrate (Fig. 5.6, center panel). From Fig. 5.6 (right panel), small grains (~10nm) of the Pt layer are clearly observed, indicating a SEM resolution that is sufficient to characterize features on similar length scale. However, we did not observe any sign of NaOH crystals buried beneath the heater line.

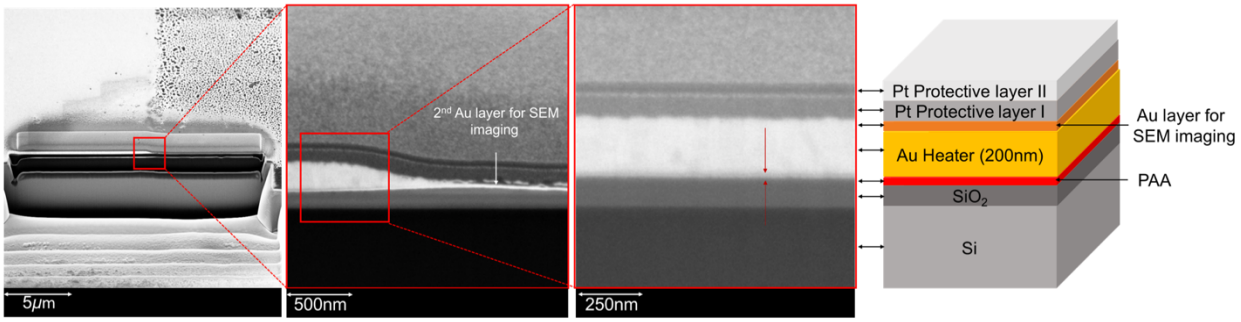

Figure 5.6 | Cross-section view of a pH 10 PAA film under the Au heater lines. Focused ion beams (FIB) were used to etch an indentation on a heater line at the edge of sample and reference region (left panel). Polymer film, Au heater line, deposited Au layer for SEM imaging and protective Pt layers can be seen in the cross-sectional view (center panel). Zoomed in image of the sample cross-section doesn't show any embedded NaOH crystals in the film (marked by red arrows) beneath the Au heater lines (right panel).

5.5 Factors contributing to enhanced κ

To deconvolute the contributions to measured κ from the three ionization induced effects, we employed the minimum thermal conductivity model (MTCM)[10,35] which describes thermal transport in amorphous and highly disordered materials. According to this model, κ scales with atomic density (ρ_{atom}) as $\rho_{atom}^{1/6}$, which has been approximated with mass density (ρ) here, and linearly with sound velocity which further depends on elastic modulus as $E^{1/2}$. Since the film density cannot be directly calculated from the PALS data, we interpolated densities at different degrees of ionization based on the bulk densities reported in ref. 24. Assuming that film density scales linearly with bulk density, a ~20% higher bulk density at pH 12 (α = 92.5%) compared to pH 1 (α = 0%) suggests a relatively small (~3%) density-related contribution to the enhanced κ. The modulus-related contribution to the measured κ was calculated to be ~37%. Based only on density- and modulus-related contributions to κ, the thermal conductivity at pH 12 is predicted to be ~0.47 $Wm^{-1}K^{-1}$, a 40% enhancement in κ over that of pH 1 (κ = 0.34 $Wm^{-1}K^{-1}$) that is substantially smaller than the ~250% enhancement measured. Fig. 5.3D shows the various contributions to measured κ for each pH calculated by taking κ for pH 1 as the baseline. Clearly, the MTCM, which is based on vibrational states that are neither fully localized nor propagating (diffusons), doesn't entirely capture the enhancement in κ measured in this system. We speculate that the extended and stiffened PAA chains may result in increased diffusion lengths for diffusons. Since long-range propagating modes have been previously shown to exist in disordered solids like

amorphous Si[13,36], it is also possible that a small population of "propagons" exists in the chain-extended PAA. However, further studies are necessary to understand the detailed heat transport mechanisms in these extended systems. We note that a prior work[37] has examined the thermal conductivity of ionically-crosslinked polymer salt with κ reaching 0.67 $Wm^{-1}K^{-1}$; we thus attribute the additional κ enhancements shown in this study to the added chain extension effect in NaOH-treated PAA, which gives rise to a greater persistence length and larger effective rigidity of the polymer chains[20]. It is likely that the predominant vibrational transfer of heat along the covalently bonded polymer backbone afforded by the extended and stiffened chain morphology as well as enhanced interchain conductance due to stronger ionic bonds result in the substantial increment in κ. The measured increases in κ are consistent with a recent computational study that predicts large enhancements in κ with increasing persistence length in amorphous polyethylene[38]. To further confirm the contributions of extended chain morphology of the ionized PAA chains to measured thermal conductivity, we performed solvent vapor annealing (SVA) on spin-cast PAA films. Absorption of solvent vapors during SVA increases chain mobility resulting in morphological equilibration of the polymer chains that had been kinetically frozen[39,40]. As shown in Fig. 5.3E, the differences in thermal conductivity between "as-made" and "solvent-annealed" samples are within experimental uncertainties for low pH samples, indicating that any disruption of ionic and H-bond interactions due to solvent annealing doesn't change κ significantly. However, for pH 10 and 12 samples much lower thermal conductivities were measured for solvent-annealed samples, which can be explained by the coiling up (i.e., relaxation) of PAA chains during the solvent-annealing process. At pH 12 specifically, thermal conductivity dropped by as much as ~32%. This signifies that kinetically frozen extended PAA chains are partially responsible for high thermal conductivities measured in the spin-cast films.

5.6 Comparison with composites

We compared the thermal conductivities of chain-extended PAA films with those of two types of composite films, PAA/NaCl and PVP/NaOH, composed of mutually unreactive polymer-salt mixtures. We assume that the salt added in these samples is proportionally retained in the thin film upon spin-casting from the polymer-salt solution and acts as a high-κ filler. As shown in Fig. 5.7A, salt fillers have miniscule effect on composite thermal conductivities till ~20% filler volume fraction. This signifies that extended chain morphology may be more effective at transferring heat

than composite strategies, where large thermal resistances may exist at filler-filler and filler-polymer interfaces.

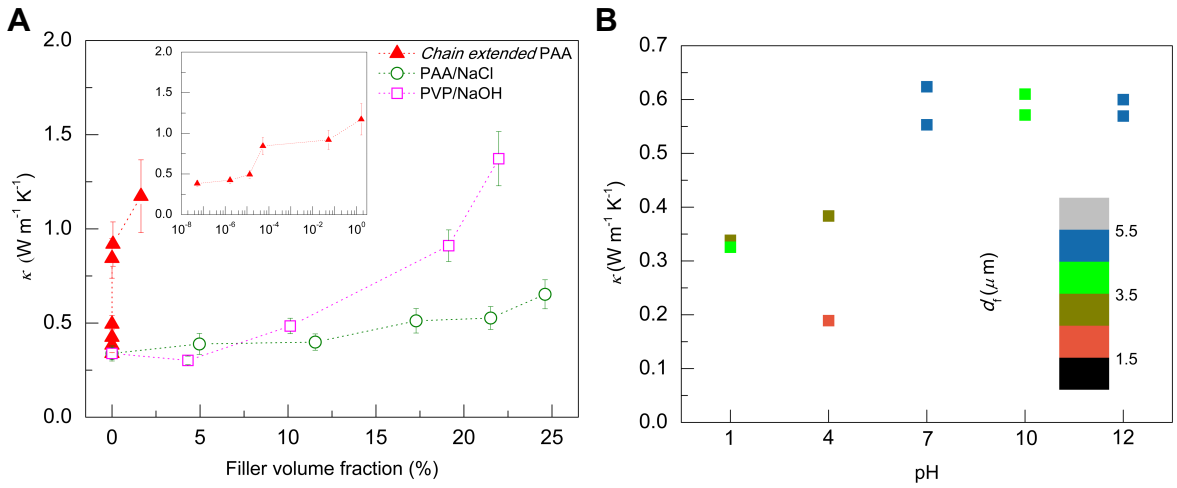

Figure 5.7 | Comparison of thermal conductivities of chain extended PAA and polymer-salt composites. **(A)** Thermal conductivity of thin films of water soluble polymers with added inorganic salts. Chain extended PAA refers to PAA films spin-cast from solutions at different pH. Salts added in PAA/NaCl and PVP/NaOH samples do not react with respective polymers and act as high-κ fillers. The inset shows data for chain-extended PAA with abscissa on log scale. **(B)** Thermal conductivity of thick PAA films blade-coated from solutions at different pH. The color map shows film thickness in micrometers. The error in κ was less than 4% for all samples and has not been shown.

5.7 Micrometer-thick amorphous films

As a demonstration of the applicability and potential for scale-up of chain-extended PAA, we fabricated ~1.5-4.5 μm thick PAA (MW ~450kDa, atactic) films by blade-coating[28], which is a method representative of the large-scale roll-to-roll processing. The average κ measured for chain-extended samples (pH 7-12) was ~0.59 $Wm^{-1}K^{-1}$, which is nearly 80% enhancement over the average κ (~0.33 $Wm^{-1}K^{-1}$) measured for the coiled-chain pH 1 samples (Fig. 5.7B). The lower κ in the blade-coated samples likely results from the thermodynamic nature of the blade-coating method. While spin-casting freezes the polymer chains in a thermodynamically high energy state upon rapid solvent evaporation, slow evaporation during blade-coating allows the chains to relax into a more thermodynamically favorable coiled-up morphology, which is consistent with the decline in κ observed for the spin-cast films subjected to solvent vapor annealing. The highest κ (0.62 ± 0.02 $Wm^{-1}K^{-1}$) measured among the thick films is more than 50% larger than the κ (~0.4 $Wm^{-1}K^{-1}$) achieved in un-stretched ultra-high MW semi-crystalline (crystallinity ~15%)

polyethylene (UHMWPE) films of comparable thickness[41]. The maximum value of V_{NaOH} in such films was calculated to be only 1.64% (corresponding to pH 12).

The measured thermal conductivities for these films at high pH (i.e., pH > 7) significantly exceed the Maxwell-predicted values (Fig. 5.8), ruling out any contribution to κ from NaOH crystals. Although directional shear force during blade-coating can potentially lead to some short-range ordering parallel to the substrate within the polymer films as has been previously reported for atactic PAA[42,43], any such ordering in the in-plane direction would likely cause the in-plane κ (κ_x) to be even greater than the measured cross-plane κ (κ_z). We also note that no such short-range ordering is deemed possible in the spin-cast films due to the kinetic nature of film formation.

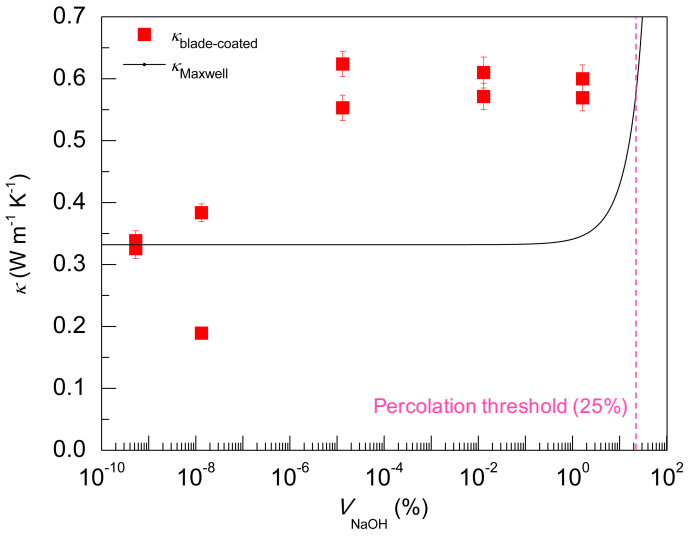

Figure 5.8 | Comparison of thermal conductivities of blade-coated films with Maxwell model. Measured thermal conductivities, $\kappa_{blade\text{-}coated}$, for blade-coated films greatly exceeds the Maxwell-model predicted values, indicating enhancement is not primarily due to a high-κ filler effect.

5.8 Error analysis for differential 3ω measurement

In the standard sample configuration (middle panel of Fig. 5.9A), a lithography-fabricated shadow mask with precisely defined geometry (± 1 μm) was attached on a silicon substrate (thickness = 500 μm) with 100 nm thermally-grown SiO_2 for electrical insulation and spin-cast polymer film (on one half of the substrate) for heater line deposition. A 5 nm adhesion layer (Ti) was first deposited via electron beam evaporation (SJ-20 Evaporator, Denton Vacuum) followed by 200 nm of heater line material (Au). Since PAA is hygroscopic and the absorbed humidity could alter charge density along the backbone as well as disrupt various hydrogen-bonds within the thin film, an alternative sample configuration (right panel in Fig. 5.9A) with a 50 nm sputter-coated alumina capping layer (to block humidity) was also used for 3ω measurement. As the alumina capping layer was deposited on sample and reference regions simultaneously, the temperature rises across it were approximately equal in both regions, making errors caused by the thin Al_2O_3 layer negligible. We note that, in order to minimize sample degradation due to humidity, all PAA samples were kept under vacuum prior to 3ω measurements, and consequently values of κ measured from both configurations are consistent within sample-to-sample variations (Fig. 5.9C).

A crucial assumption of the differential 3ω measurement is that the heater lines deposited in the sample and reference regions are identical. This is ensured by simultaneous deposition of closely-placed heater lines (distance ~10 mm) in both regions using shadow masks with precise and consistent dimensions. Four identical heater lines (indexed "1", "2", "3", "4") are deposited on sample and reference regions respectively to facilitate cross-checking (i.e., obtain κ values from different combinations of heater lines in both regions). Good agreements between κ measured from various heater line combinations indicate minuscule variations among heater lines deposited using the same mask (Fig. 5.9B).

For spin-cast films, an upper limit of heating frequency (f) equal to 442 Hz was chosen to satisfy the line source approximation adapted in the thermal model. This corresponds to a lower bound of thermal penetration depth (TPD) equal to 120 μm (TPD = $(\alpha_{sub}/2\omega)^{1/2}$, where α_{sub} = 0.8 cm^2s^{-1} is the thermal diffusivity of the silicon substrate and $\omega = 2\pi f$ (*16, 46*) that is much greater than heater line half-width ($b = w/2$ = 22.3 μm). Meanwhile, the lower limit of heating frequency was set at 175 Hz, yielding an upper bound of TPD equal to 191 μm that is well below the substrate thickness (500 μm), validating the 1-dimensionalilty of heat transfer as well as semi-infinite substrate

approximation. As for the blade-coated samples with large thicknesses ($d_f > 2$ μm), a lower range of heating frequencies were selected. A prior work[16] has examined the deviation from 1-D heat transfer due to lateral (in-plane) heat spreading, which can be sizable as d_f becomes comparable to heater line half width. Furthermore, it is possible that shear-force-induced short-range chain ordering in the blade-coating direction (in-plane)[28] gives rise to a higher in-plane thermal conductivity (κ_x) than its cross-plane counterpart (κ_z), amplifying lateral heat spreading[16]. However, using an analytical model developed in Ref. 16, we find the error in κ_z due to lateral heat spreading can be largely mitigated (<1% deviation from κ calculated using a 2-D thermal model) by replacing heater line width with a corrected value of ($w + 0.88d_f$). In addition, a range of heating frequencies was selected for blade-coated PAA films such that it was low enough to enlarge TPD/d_f and further suppress lateral heat transfer, but at the same time high enough to constrain the TPD (~460 μm for the lowest heating frequency used) within the substrate thickness (Fig. 5.9F). As a check for semi-infinite substrate approximation, it is known that parasitic effect due to insufficient TPD appears as a deviation from linearity in ΔT vs. $\log(\omega)$[16,44], which is not evident in our results (Fig. 5.9D and E). We note that in reality potential errors that usually stem from line source and 1-D heat transfer approximations can be largely mitigated by the differential nature of our experiment. Lastly, to check measurement accuracy, thermal conductivities of spin-cast poly(methyl methacrylate) (PMMA) and PAA (dissolved in DMF) were measured to be 0.21±0.02 $Wm^{-1}K^{-1}$ and 0.35±0.03 $Wm^{-1}K^{-1}$, respectively, agreeing well with reported values obtained by time-domain thermal reflectance for films of similar thicknesses and produced under identical conditions[34.]

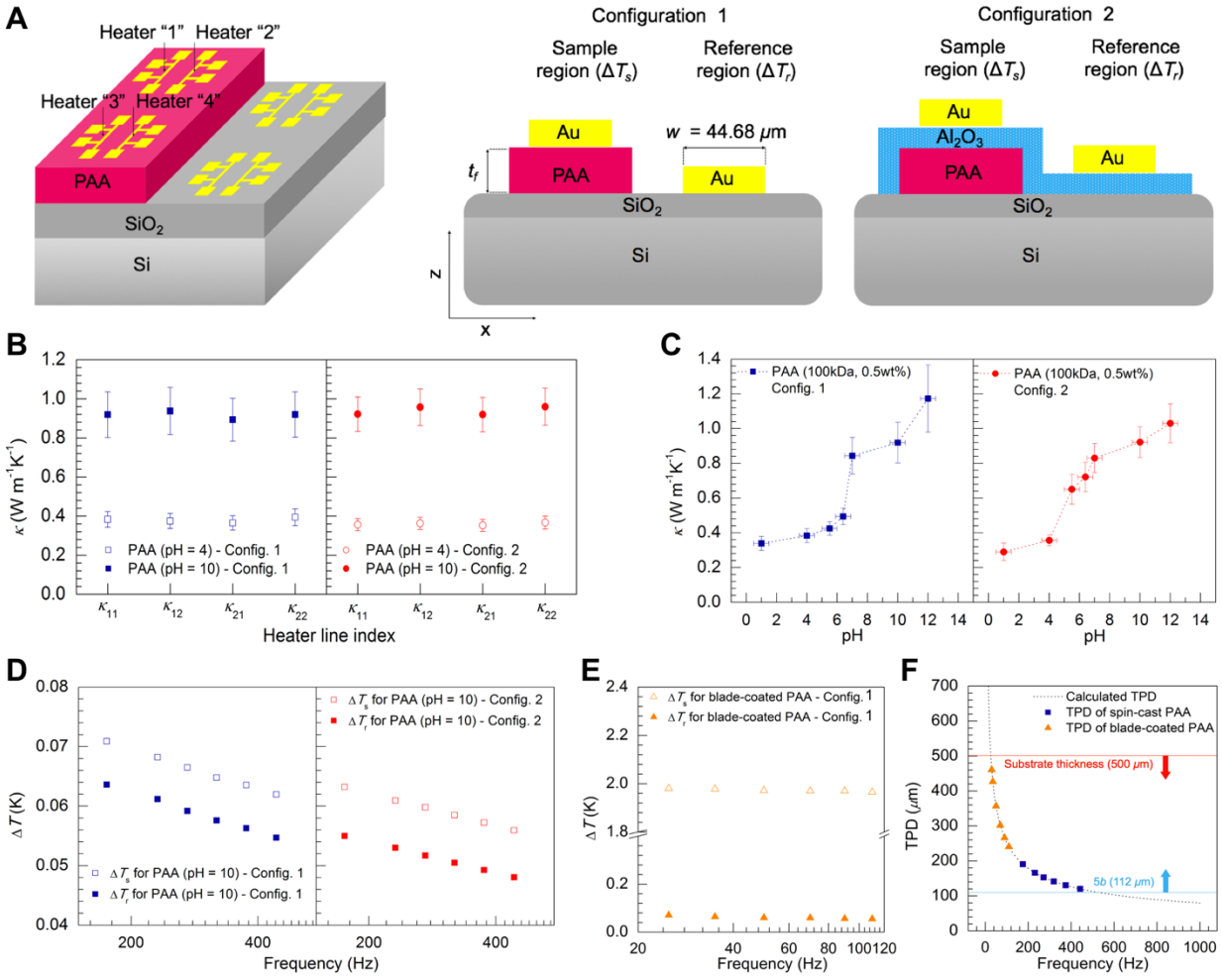

Figure 5.9 | Differential 3ω measurement of thermal conductivity (κ) (A) Cross-section (not drawn to scale) of sample geometry showing four identical heater lines (indexed "1", "2", "3", "4") deposited on both the sample and reference regions. Indices are the same for heater lines on both regions. Configuration 2 includes a 50 nm alumina capping layer to block humidity. **(B)** κ of pH=4 and 10 samples measured from different combinations of heater lines (for example, "κ_{11}" represents κ measured using heater line "1" in both sample and reference regions) using both sample configurations. **(C)** Thermal conductivities of PAA as a function of solution pH measured from both configurations, showing good agreement of measured κ. **(D) and (E)** Temperature rise on sample (ΔT_s) and reference (ΔT_r) regions for spun-cast **(D)** and blade-coated films **(E)**, the difference of which is the temperate rise across PAA films (ΔT_f). **(F)** TPD versus heating frequencies (f). Several heating frequencies were chosen such that their corresponding TPDs were greater than 5 times the heater line half-width ($5b$=112 μm) but lower than the substrate thickness (500 μm), validating the semi-infinite substrate and 1-D heat transfer approximations adopted in the thermal model.

As mentioned in Chapter 2, the accuracy of differential 3ω technique depends primarily on heater lines and substrate conditions being identical in the sample and reference regions. While good agreement between values of κ measured from different heater line combinations indicates measurement robustness, several parameters (heater line width, temperature coefficient of resistance, film thickness and thermal boundary conductance) are subject to experimental uncertainties and their consequential errors are analyzed below.

Heater line width, temperature coefficient of resistance and film thickness

Widths (w) of forty heater lines measured by an optical microscope yielded a standard deviation ($\sigma_{SD,\,w}$) of 0.933% (Fig. 5.10A). Since $\sigma_{SD,\,w}$ affects temperature rises in the sample and reference regions independently, it is magnified by the inverse of signal-to-noise ratio $(\Delta T_s^2 + \Delta T_r^2)^{1/2}/\Delta T_f$ (*46*) so that a small variation in w may lead to an enlarged uncertainty in measured κ depending on the magnitude of $(\Delta T_s^2 + \Delta T_r^2)^{1/2}/\Delta T_f$. Therefore, $\sigma_{SD,\,w} \times (\Delta T_s^2 + \Delta T_r^2)^{1/2}/\Delta T_f$ was used for error calculation. TCRs of forty heater lines were measured previously in a cryogenic setup with a homemade LabVIEW data acquisition program. Since TCR varies for heater lines deposited using different deposition tools, the same e-beam evaporator has been used for all the experiments. $\sigma_{SD,TCR}$ was found to be small (2.7%) and it is further minimized by the differential nature of our experiment.

Thickness (d_f) of spin-cast PAA films was measured by ellipsometry (Woollam M-2000DI Ellipsometer) and surface profilometry (Dektak XT Surface Profilometer) while d_f of blade-coated films were measured using only the latter due to their greater film roughness. Ellipsometry is ideal for measuring the thicknesses of smooth films, typically yielding a measurement error within 5 Å; assuming a lower bound of d_f equal to 10 nm, we set the thickness error to be 5 % (5 Å/10 nm) for error calculation. During ellipsometry, the laser beam spot was focused on the same region where the heater lines were to be deposited later to ensure d_f and κ were measured in an overlapping region, thereby minimizing errors in κ caused by spatial variation within the film. Thicknesses of twenty PAA samples were cross-checked using Dektak surface profilometer. Prior to profilometry, thin grooves were made by scratching the polymer film with a steel blade (hardness: steel < SiO_2) while keeping the SiO_2 layer intact. Since Dektak stylus radius (12 μm) is smaller than the width of the steel blade (25 μm), it is able to measure the depth of the groove as the stylus sweeps across. As shown in Fig. 5.10B, Dektak-measured thicknesses (averaged over three positions along the

groove) match well with the ellipsometry-measured values, which is consistent with smooth sample morphology. For blade-coated samples, the thickness was measured in a way that the stylus moved from a point close to the heater line in the sample region to the reference region. The error caused by intrinsic height variation due to this translational movement (which is around 1 cm) was estimated to be 20 nm, which is negligible compared to the film thickness (>2 μm).

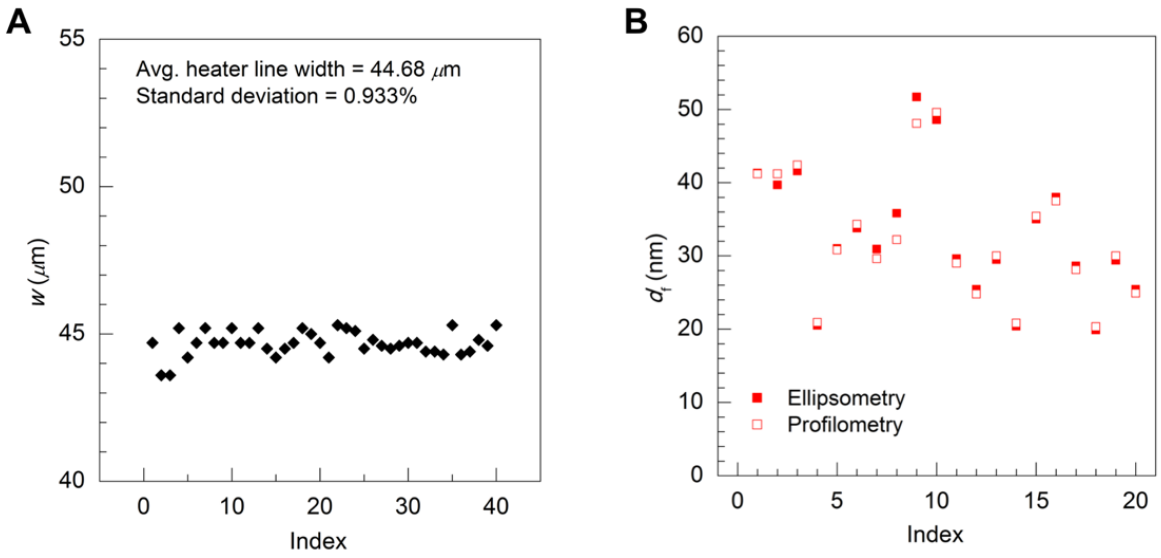

Figure 5.10 | Widths measured for forty heater lines and film thicknesses measured for twenty spin-cast PAA samples. (A) Small standard deviation in heater line width translates into a relatively small (though magnified by the inverse of signal-to-noise ratio) error in thermal conductivity measurement. **(B)** Good agreement was found between thicknesses measured using ellipsometry and surface profilometry for spin-cast films, underlying the smooth morphology of these films.

Thermal boundary conductance

It is well known that thermal boundary conductance (TBC) of polymer-inorganic interface could vary as the polymer film thickness becomes comparable to its radius of gyration, owing to polymer chain re-orientation near the substrate. Conceivably, ionization-induced chain extension could lead to better adhesion between polymer and substrate, thereby enhancing the measured effective thermal conductivities of PAA spin-cast from high pH solutions due to a larger PAA/SiO_2 TBC. However, we found the enhancement in TBC, if any, insufficient to account for the measured high thermal conductivities. Assuming an actual PAA (100 kDa, pH = 10) thermal conductivity of 0.35 $Wm^{-1}K^{-1}$ and a Ti/SiO_2 TBC of 100 $MWm^{-2}K^{-1}$ in the reference region, a diverging TBC on either side of PAA yielded only a measured κ_{eff} equal to 0.424 $Wm^{-1}K^{-1}$, significantly lower than the

values measured for high pH samples (κ = 0.922 $Wm^{-1}K^{-1}$ for pH 10 sample, see Fig. 5.11A). Similarly, as shown in Fig. 5.11B, we calculated the minimum PAA conductivities (κ_{min}) required to fit the measured values ($\kappa_{measured}$), assuming again an infinite TBC in the sample region and a Ti/SiO_2 TBC of 100 $MWm^{-2}K^{-1}$ in the reference region. In addition, since errors due to variations in TBC become negligible in thick (>2 μm) blade-coated samples due to large thermal resistance, the six-time increase in κ of blade-coated pH=7 film (0.624 ± 0.021 $Wm^{-1}K^{-1}$) compared to that of a pH=1 film (0.114 ± 0.004 $Wm^{-1}K^{-1}$) suggests a prominent effect from chain ionization and its effects rather than TBC enhancement.

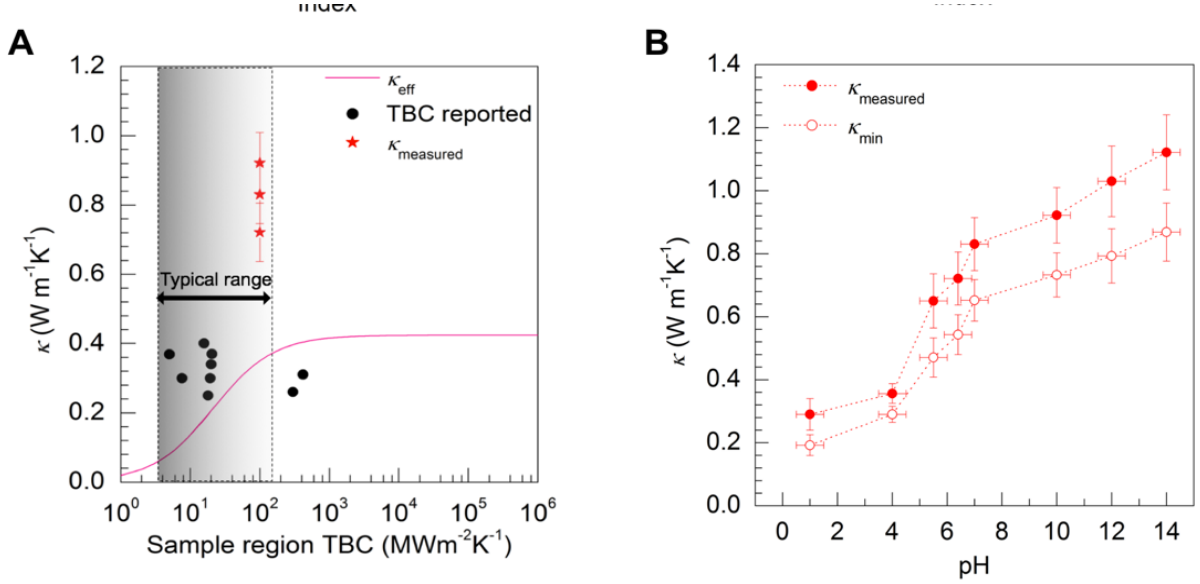

Figure 5.11 | TBC analysis and minimum κ. (A) The true thermal conductivity of the PAA film a a function of sample region TBC, assuming an actual PAA (100,000, pH = 10) thermal conductivity of 0.35 W m^{-1} K^{-1} and a Ti/SiO_2 TBC of 100 MW m^{-2} K^{-1} in the reference region. **(B)** $\kappa_{measured}$ (Configuration 2) and κ_{min} versus pH. κ_{min} is the minimum thermal conductivities PAA films need to have in order to fit the measured values, by assuming an infinite TBC in the sample region and a Ti/SiO_2 TBC equal to 100 MWm^{-}

5.9 Discussion and summary

In summary, we have employed electrostatic repulsive forces to stretch the polyelectrolyte backbone at the molecular level, resulting in extended conformations, better packed chains and enhanced modulus, all of which contribute to significantly enhanced thermal conductivities. For the spin-cast thin films, it is to be noted that centrifugal forces during spin-casting may cause polymer chains to be more expanded in the in-plane direction, possibly making in-plane thermal conductivity even greater than the measured cross-plane κ.[6] This unexplored route for molecular engineering of polymer thermal conductivity is also extended to making micron-thick blade-coated films, with thermal conductivity reaching over 0.6 $Wm^{-1}K^{-1}$.

References

1. H. Chen, V.V. Ginzberg, J. Yang, Y. Yang, W. Liu, Y. Huang, L. Du, B. Chen. Thermal conductivity of polymer-based composites: fundamentals and applications. *Prog Polym Sci* **59**, 41-85 (2016).

2. S. Gelin, H. Tanaka, A. Lemaitre. Anomalous phonon scattering and elastic correlations in amorphous solids. *Nature Mater.* **15**, 1177-1181 (2016).

3. A. Henry, G. Chen. High thermal conductivity of single polyethylene chains using molecular dynamics simulation. *Phys Rev Lett* **101**, 235502 (2008).

4. S. Shen, A. Henry, J. Tong, R.T. Zheng, G. Chen. Polyethylene nanofibers with high thermal conductivities. *Nature Nanotech* **5**, 251-255 (2010).

5. C.L. Choy. Thermal conductivity of polymers. *Polymer* **18**, 984-1004 (1977).

6. K. Kurabayashi, M. Asheghi, M. Touzelbaev, K.E. Goodson. Measurement of the thermal conductivity anisotropy in polyimide films. *J Microelectromech Sys.* **8**, 180-191, (1999).

7. V. Singh, T.L. Bougher, A. Weathers, Y. Cai, K. Bi, M.T. Pettes, S.A. McMenamin, W. Lv, D.P. Resler, T.R. Gattuso, D.H. Altman, K.H. Sandhage, L. Shi, A. Henry, B.A. Cola. High thermal conductivity of chain-oriented amorphous polythiophene. *Nature Nanotech* **9**, 384-390 (2014).

8. A. Roy, T.L. Bougher, R. Geng, Y. Ke, J. Locklin, B.A. Cola. Thermal conductivity of poly(3-methylthiophene) brushes. *ACS Appl. Mater Interfaces* **8**, 25578-25585 (2016).

9. M.D. Losego, L. Moh, K.A. Arpin, D.G. Cahill, P.V. Braun. Interfacial thermal conductance in spun-cast polymer films and polymer brushes. *Appl. Phys. Lett.* **97**, 011908 (2010).

10. W.P. Hsieh, M.D. Losego, P.V. Braun, S. Shenogin, P. Keblinski, D.G. Cahill. Testing the minimum thermal conductivity model for amorphous polymers using high pressure. *Phys Rev B* **83**, 174205 (2011).

11. P.B. Allen, J.L. Feldman, J. Fabian, F. Wooten. Diffusons, locons and propagons: character of atomic vibrations in amorphous Si. *Philos Mag B* **79**, 1715-1731 (1999).

12. S. Shenogin, A. Bodapati, P. Keblinski, A.J.H. McGaughey. Predicting the thermal conductivity of inorganic and polymeric glasses: The role of anharmonicity. *J Appl Phys* **105,** 034906 (2009).

13. K.T. Regner, D.P. Sellan, S. Zonghui, C.H. Amon, A.J.H. McGaughey, A.J. Malen. Broadband phonon mean free path contributions to thermal conductivity measured using frequency domain thermoreflectance. *Nat Commun* **4**, 1640 (2013).

14. A. Henry. Thermal Transport in Polymers, *Annual Review of Heat Transfer*, Vol 17, Chapter 13, 485-520 (2013).

15. S.M. Lee, D.G. Cahill. Heat transport in thin dielectric films. *J Appl Phys* **81**, 2590-2595 (1997).

16. T. Borca-Tasciuc, A.R. Kumar, G. Chen. Data reduction in 3 omega method for thin-film thermal conductivity determination. *Rev Sci Instrum* **72**, 2139-2147 (2001).

17. J. Choi, M.F. Rubner. Influence of degree of ionization on weak polyelectrolyte multilayer assembly. *Macromol* **38**, 116-124 (2005).

18. A.F. Xie, S. Granick. Local electrostatics within a polyelectrolyte multilayer with embedded weak polyelectrolyte. *Macromol* **35**, 1805-1813 (2002).

19. D.W. van Krevelen. Properties of Polymers: their correlation with chemical structure; their numerical estimation and prediction from additive group contributions. 4th, completely rev. ed. Amsterdam: Elsevier (2009).

20. S.W. Cranford, M.J. Buehler. Variation of weak polyelectrolyte persistence length through an electrostatic contour length. *Macromol* **45**, 8067-8082 (2012).

21. D. Stigter, K.A. Dill. Theory of radii and second virial coefficients 2. Weakly charged polyelectrolytes. *Macromol.* **28,** 5338-5346 (1995).

22. H. Lee, J.R. Boyce, A. Nese, S. S. Sheiko, K. Matyjaszewski. pH-induced conformational changes of loosely grafted molecular brushes containing poly(acrylic acid) side chains. *Polymer* **49**, 5490–5496 (2008).

23. A.B. Unni, G. Vignaud, J. P. Chapel, J. Giermanska, J. K. Bal, N. Delorme, T. Beuvier, S. Thomas, Y. Grohens, A. Gibaud. Probing the Density Variation of Confined Polymer Thin Films via Simple Model-Independent Nanoparticle Adsorption. *Macromol.* **50**, 1027−1036 (2017).

24. D.W. Gidley, H.-G. Peng, R.S. Vallery. Positron Annihilation as a Method to Characterize Porous Materials. *Annu. Rev. Mater. Res.* **36,** 49-79 (2006).

25. K. Hiraoka, H. Shin, T. Yokoyama. Density measurements of poly(acrylic acid) sodium salts. *Polym. Bull.* 8, 303-309 (1982).

26. W.C. Oliver, G.M. Pharr. An improved technique for determining hardness and elastic modulus. *J. Mater. Res* **7**, 1564-1583 (1992).

27. R. Akhtar, N. Schwarzer, M.J. Sherratt, R.E. Watson, H.K. Graham, A.W. Trafford, P.M. Mummery, B. Derby. Nanoindentation of histological specimens: Mapping the elastic properties of soft tissues. *J. Mater. Res* **24**, 638-646 (2009).

28. B. Kim, E.J. Jeong, J.W. Chung, S. Seo, B. Koo, J. Kim. A molecular design principle of lyotropic liquid-crystalline conjugated polymers with directed alignment capability for plastic electronics. *Nature Mater.* **12**, 659-664 (2013).

29. A. Eisenberg. Glass transitions in ionic polymers. *Macromol.* **4**, 125-128 (1971).

30. A. Eisenberg, H. Matsura, T. Yokoyama. Glass transition in ionic polymers: the acrylates. *J. Polym. Sci.: Part A-2* **9**, 2131-2135 (1971).

31. M. Todica, T. Stefan, S. Simon, I. Balasz, L. Daraban. UV-Vis and XRD investigation of graphite-doped poly(acrylic) acid membranes. *Turk J Phys* **38**, 261 – 267 *(*2014).

32. S. Komaba, N, Yabuuchi, T. Ozeki, Z.-J. Han, K. Shimomura, H. Yui, Y. Katayama, T. Miura. Comparative study of sodium polyacrylate and poly(vinylidene fluoride) as binders for high capacity si graphite composite negative electrodes in li-ion batteries. *J. Phys. Chem. C* **116**, 1380–1389 (2012).

33. V.R. Raghavan, H. Martin. Modelling of two-phase thermal conductivity. *Chem Eng Process* **34,** 439–446 (1995).

34. T. Brar, P. France, P.G. Smirniotis. Heterogeneous versus homogeneous nucleation and growth od zeolite A. *J. Phys. Chem. B* **105**, 5383-5390 (2001).

35. X. Xie, D. Li, T.-H. Tsai, J. Liu, P.V. Braun, D.G. Cahill. Thermal conductivity, heat capacity, and elastic constants of water soluble polymers and polymer blends. *Macromol.* **49**, 972-978 (2016).

36. X. Liu, J.L. Feldman, D.G. Cahill, R.S. Crandall, N. Bernstein, D.M. Photiadis, M.J. Mehl, D.A. Papaconstantopoulos. High thermal conductivity of a hydrogenated amorphous silicon film. *Phys. Rev. Lett.* **102**, 035901 (2009).

37. X. Xie, K. Yang, D. Li, T.-H. Tsai, J. Shin, P.V. Braun, D.G. Cahill. High and low thermal conductivity of amorphous macromolecules. *Phys Rev B* **95**, 035406 (2017).

38. T. Zhang, T. Luo. Role of chain morphology and stiffness in thermal conductivity of amorphous polymers. *J. Phys. Chem. B* **120**, 803-812 (2016).

39. J. Vogelsang, J. Brazard, T. Adachi, J.C. Bolinger, P.F. Barbara. Watching the annealing process one polymer chain at a time. *Angew. Chem.* **123**, 2305 –2309 (2011).

40. Y.S. Jung, C.A. Ross. Solvent-vapor-induced tunability of self-assembled block copolymer patterns. *Adv. Mater.* **21**, 2540-2545 (2009).

41. H. Ghasemi, T. Nagarajan, X. Huang, J. Loomis, X. Li, J. Tong, J. Wang, G. Chen. High thermal conductivity ultra-high molecular weight polyethylene (UHMWPE) films. *Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2014 IEEE Intersociety Conference on*, Orlando, FL, 235-239 (2014), doi: 10.1109/ITHERM.2014.6892287.

42. V.A. Kargin, S.Y. Mirlina, V.A. Kabanov, G.A. Mikheleva, A study of the structure of isotactic polyacrylic acid and its salts. Vysokomol. soedin. **3**, No. 1, 139-143 (1961).

43. M.L. Miller, K. O'Donnell, J. Skogman. Crystalline polyacrylic acid. *J. Colloid Sci.* **17**, 649-659 (1962).

44. D.G. Cahill. Thermal-Conductivity Measurement from 30K to 750K: the 3-Omega Method. *Rev Sci Instrum* **61**, 802-808, (1990).

45. T. Borca-Tasciuc, D. Song, J. L. Liu, G. Chen, K. L. Wang, X. Sun, M. S. Dresselhaus, T. Radetic, and R. Gronsky. Anisotropic Thermal Conductivity of a Si/Ge Superlattice. *Mat Res Soc Symp Proc* **545**, 473, (1998).

46. Y.K. Koh, S. L. Singer, W. Kim, J. M. O. Zide, H. Lu, D. G. Cahill, A. Majumdar, A. C. Gossard. Comparison of the 3 omega method and time-domain thermoreflectance for measurements of the cross-plane thermal conductivity of epitaxial semiconductors. *J Appl Phys* **105**, 054303, (2009).

Chapter 6
Thermal Management at the Package Level – Integrated Phase Change Heatsink for Thermal Buffering of an Accelerator Chip

6.1 Introduction

The ever-increasing transistor density and halted voltage scaling cause the power density of modern electronics to rise with each technology generation, posing a serious thermal management problem especially for hand-held smartphones and wearable electronics that comply to stringent device temperature limit[1-4]. Since the heat dissipation mechanisms in such devices are mainly constrained by passive convection, their clock frequencies are in general lower than that of larger electronics (for example, laptops and desktops) for which active cooling is feasible. Consequently, in such devices at each given time a fraction of transistors is off to maintain stable chip temperature, giving rise to the so-called "dark silicon" phenomenon[4-6].

A large fraction of existing and emerging mobile applications, such as speech recognition, visual recognition and web browsing, demands short bursts of intense computations separated by extended (at least tens of seconds) idle states waiting on user input[7,8]. Therefore, a strategy named "computational sprinting" in tune with the time trace of such applications is proposed[9-11], where a chip is run temporarily at a power density far exceeds the sustainable thermal budget (also known as the thermal design power or TDP) or by activating all reserved cores to meet the bursty computation demands, after which the chip is allowed to cooldown to its normal state. The sprinting strategy, with proper design, takes advantage of material thermal capacitance to buffer short time-scale but intense heat generation, thereby leaving enough headroom for thermal budget to avoid chip overheating at the end of each sprint to maintain device functionality[9].

Prior research by our group has examined the thermal performance of a silicon thermal test-chip (TTC) integrated with on-chip phase change materials (PCMs) as a proxy for mobile chips[12]. The TTC can switch between two power states (11 W and 1W, respectively) controlled by solid-state

relays, simulating the "on" and "off" states of computational sprinting. It was found that incorporating PCMs can either reduce peak and average chip temperatures at a given sprinting duration, or allow an extended sprint duration at a given thermal budget. However, the practical applicability of the measurement results is limited by the proxy chip that is unable to perform real computations.

In the last part of this chapter, we present an experimental study regarding the effectiveness of PCMs as thermal buffer materials in a realistic state-of-the-art neural network accelerator test-chip (ATC) designed for deep learning[13] (see Fig. 6.1A, *left*) that is able to produce a power contrast (i.e., ratio of power during sprint and in normal state) greater than 10. Fabricated in a 65nm process, the 0.9mm^2 ATC die (see Fig. 6.1A, *right*) runs at 4.5GHz[13], achieving a performance of 0.76TOPS and an area efficiency of 0.84TOPS/mm^2. We note that the designed ATC outperforms the latest deep learning accelerators on other chips[14-16] by at least 4 times with improvement on area efficiency by at least 8 times. By taking advantage of energy-efficient resonant clocking we use high-speed resonant stream buffers to provide weights to neurons without stalling. The resulting stream architecture with multi-GHz resonant clocking demonstrates a new design point that optimizes both area efficiency and power efficiency.

In order to implement sprinting, a clock power of 216m W with and without an additional core power of 2.136 W are supplied (giving rise to a total power = 2.35W) to simulate the ATC's sprinting state and normal state, respectively, giving rise to a power contrast close to 11 that satisfies the requirement of sprinting where device peak power should exceed sustainable TDP by at least an order of magnitude.

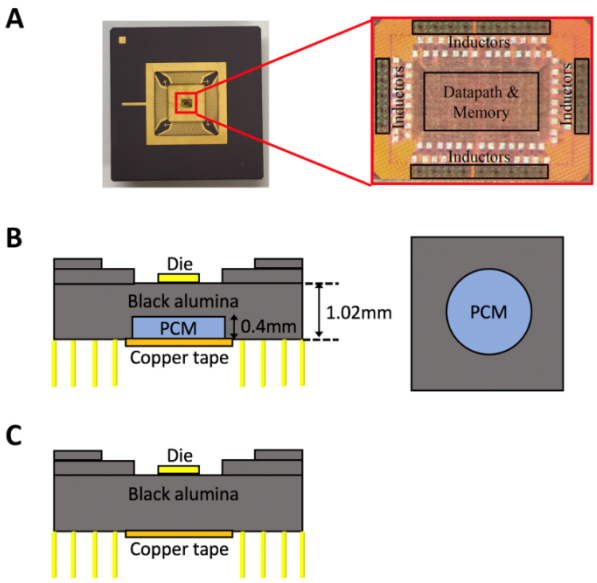

Figure 6.1 | Schematics of the die, ATC package and PCM-filled cavity. (A) Top side image of the ATC (*left*) and detailed floor plan of the die (*right*). **(B)** Cross-section view of the ATC package with integrated PCM (*left*) and bottom side view of the ATC (*right*). **(C)** Cross-section view of the ATC package without integrated PCM, used as a reference chip.

At the same time scale as human perception of mobile device responsiveness (sub-second to seconds), conventional PCMs such as paraffins and salt-hydrates become ineffective due to their low thermal conductivities that give rise to thermal constants exceeding the aforementioned time scale by orders of magnitude. To overcome this challenge, previous work has examined the use of high thermal conductivity PCM placed close to heat source to attain a quick thermal response[17,18].

To fully exploit the latent heat absorbed during phase change in the sub-second to seconds regime, we use a metallic alloy with low melting temperature, high thermal conductivity and relatively large volumetric latent heat as PCM (see Table 6.1).

TABLE 6.1
MATERIAL PROPERTIES OF THE PCM

PCM	Melting temp. (°C)	Latent heat (kJ/kg)	κ (W/mK)	Density (g/cm^3)
Roto136	56	28.9	10	9.01

6.2 Temperature measurements and PCM selection

Fig. 6.1B (*left*) depicts the cross-section view of the ATC used for temperature measurements. A metal alloy (Roto136) composed of 49% Bi, 21% In, 18% Pb, and 12% Sn (mass percentage) with a melting temperature of 58 °C is incorporated into a CNC-drilled cylindrical cavity (diameter = 9 mm and height = 4 mm) at the backside of the ATC. The relevant material properties of the alloy are summarized in Table 6.1. Prior to sealing the PCM with a copper tape, a *p*-type thermocouple is embedded in the cylindrical cavity to monitor backside temperature (T_2) *in-situ*. The die sits at the top of the ATC with its temperature (T_1) is measured by a precalibrated micro-thermocouple (diameter < 50 μm) connected to a TC1000 thermometer. T_1 was measured at 5 different spots over the die area and the results yield a spatial temperature variation less than 2 °C, verifying that a single point measurement can represent the overall die temperature without significant error. We note that the integrated PCM does not add to the volume of the ATC, which is important for compact design of mobile devices and wearable electronics. Another ATC (Fig. 6.1C) without the cavity (but with the same copper tape stick to the bottom side) was used as reference sample.

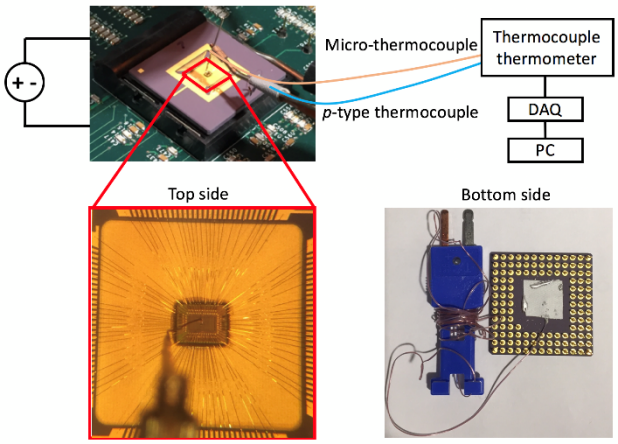

Figure 6.2 | Illustration of the ATC plugged into a breadboard, with its top and bottom side temperatures both measured and sent to a digital thermometer.

The ATC is embedded in a custom-made breadboard (acting as a heat sink) that is operated by two DC power supplies. The first power supply is always kept on to keep the clock running, while the second power supply is turned on to activate the die only during sprint. The TC1000 thermometer logs measured temperatures into the PC through a LabVIEW program with a time resolution of ~0.1s.

Prior work has defined a figure-of-merit[19] (*FOM*) for PCMs used as heat buffers. By considering the solutions of a one-dimensional Stefan problem for a semi-infinite PCM with a uniform initial temperature T_0 that is lower than the PCM's melting temperature T_m, explicit analytical solutions can be found for a constant-temperature heat source (constant boundary temperature u_L) for which

the heat flux at the boundary is expressed as $q = q_0 t^{1/2}$. In the case that convection is ignored as it is negligible compared to heat conduction in the melting region, u_L is derived as[19]:

$$u_L = \frac{q_0}{\kappa_L}\sqrt{\pi \cdot \alpha_L}\, erf(\lambda) \tag{6.1}$$

where κ_L, α_L and *erf* are liquid thermal conductivity, liquid thermal diffusivity and the error function, respectively. λ is a dimensionless parameter and is given by the following transcendental equation under the assumption that u_0 ($u_0 = T_m - T_0$) is small:

$$\lambda \cdot erf(\lambda) \cdot \frac{\sqrt{\pi}}{Ste_L} = e^{-\lambda^2} \tag{6.2}$$

where $Ste_L = C_p u_L / h$ is the liquid Stefan number and h is the latent heat. We note that in the actual temperature measurement the condition of negligible u_0 is not strictly met since the PCM is heated up from room temperature (~32.5 °C) to its melting point (we observe that Roto136 begins to melt at approximately 52 °C). Nevertheless, the physical insights we gain from the following analysis still holds as long as the sensible heat is small compared to latent heat. Since PCMs are suitable for situations where latent heat is not overwhelmed by heat flux, Ste_L is typically small, leading to a $\pi^{1/2}/(Ste_L)$ that is much greater than 1 (in Eq. 6) and a λ that is much smaller than 1. This results in $erf(\lambda) \approx Ste_L / \pi^{1/2} << 1$ and Eq. 6 becomes

$$u_L = \frac{{q_0}^2}{\kappa_L \cdot \rho \cdot h}\sqrt{\pi} \tag{6.3}$$

The product $\kappa_L \cdot \rho \cdot h$ contains only the PCM's material peripteries and is defined as its *FOM*; a larger *FOM* would lead to a lower boundary temperature. The *FOM* for Roto136 is 2603 $kW^2 \cdot s/(K \cdot m^4)$ based on published values of material properties (Table 5), comparing favorably with that of a paraffin wax (*FOM* =22 $kW^2 \cdot s/(K \cdot m^4)$ or inorganic salts such as $MgCl_2 \cdot 6H_2O$ (*FOM* =139 $kW^2 \cdot s/(K \cdot m^4)$). Though other lead-free metal alloys with similar *FOMs* are available, Roto136 is chosen due to its well-known material properties. In order to determine where the PCM should be placed relative to the heat source, it is also instructive to examine the other limiting

case where the material latent heat is overwhelmed by heat flux, representing the PCM being placed very close to the die. In this case, λ is large enough to make $erf(\lambda) \approx 1$ in Eq. 5, leading to

$$u_L = \frac{q_0\sqrt{\pi}}{\sqrt{\kappa_L \cdot \rho \cdot C_p}} \tag{6.4}$$

In this case, sensible heat rather than the latent heat should be exploited to suppress the device temperature. This implies it is not efficient to place PCM directly on top of a high-power-density heat source (such as the die of the ATC used in this work that produces a peak power density = 2.61×10^6 W/m^2) even for a PCM with large *FOM* ; they should be instead placed at the other side of the chip so that the heat generated from the heat source will spread over the entire chip before reaching the PCM. It is for this reason and the consideration to avoid electrically shorting the wires attached to the die that we integrate PCM to the backside of the ATC rather than directly on top of the die.

6.3 Experiment results for a single sprint and cooldown cycle

To study the thermal performance of the ATC during a single sprint, an ATC was operated while being mounted in air. A current/voltage equal to 0.072A/3.0V was supplied to the clock, giving rise to an idle state that consumed 216mW of power, while during sprinting an additional 2.136W (with a current/voltage of 1.78A/1.2V) of power was supplied to activate the die. The CPU of current wearable electronics typically consumes a few hundred miliwatts of power in its active state; for example, the Cortex A7 processor used in Android smartwatch consumes 185 mW at full usage[20] while the processor of Google glass draws 334 mW of power when the system is active[21]. Consequently, the 216mW power applied to the ATC at idle state was representative of the TDP of existing wearable technologies, whereas the power applied during sprinting was approximately an order of magnitude greater than the current wearable electronics maximum, approximating ten times as many active cores if parallel computation were to be assumed.

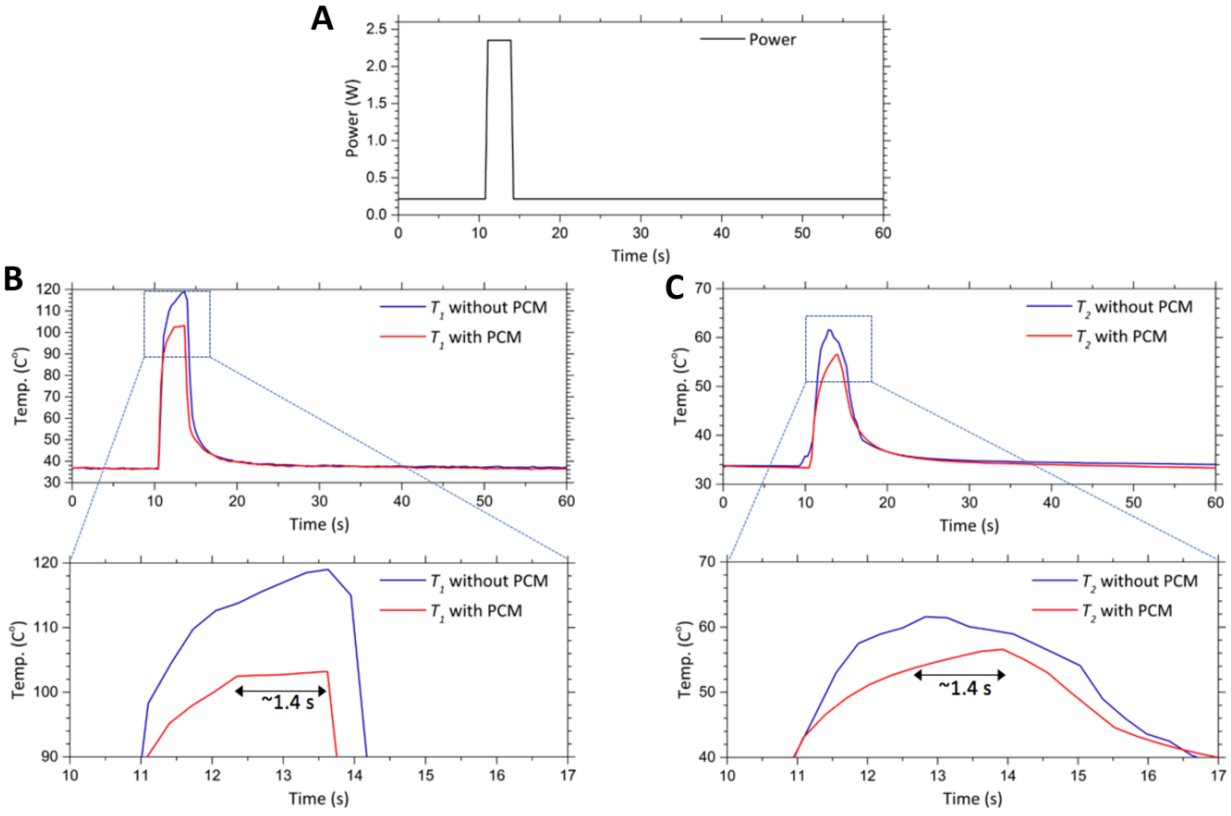

Figure 6.3 | Power trace and temperature transients of ATCs with and without integrated PCM for a single sprint-cooldown cycle. (A) Power trace. **(B)** Die temperature transients of the ATC (*top*) and close-up during the sprint (*bottom*). **(C)** Backside PCM temperature transients of the ATC (*top*) and close-up during the sprint (*bottom*).

Fig. 6.3 compares the temperature transients of the two ATCs, one with integrated PCM (Fig. 6.1B) and one without (Fig. 6.1C). The chips were operated in sprinting mode for 3.5 second as shown by their power traces (Fig. 6.3A), after which they were switched to idle state to cool down. As can be seen in Fig. 6.3B and Fig. 6.3C bottom figures that zoom-in on the temperature transients of the die and PCM temperatures, respectively, the die temperature for the ATC with PCM was held constant during the same period as PCM underwent melting, giving rise to a ~16^{o}C reduction in peak temperature compared to the chip without PCM. For the chip with integrated PCM, the backside temperature transient (Fig. 6.3C bottom figure) has a shoulder of 1.4 seconds starting at the onset of melting (~52 °C), demonstrating the metal alloy's phase change. The 1.4 second melting time converts to ~3.3 J of energy being absorbed for a heating power of 2.35W, which then translates to ~0.114 g of PCM that is melted. Therefore, only less than half of PCM filling up (0.23g as measured with a precision scale) the cavity underwent melting. This is due to the large PCM area relative to that of the chip, causing the heat to spread insufficiently prior to reaching the PCM such that the portion of PCM sitting directly beneath the die is heated up and melted faster than the peripheral PCMs. The layer of alumina sandwiched between the die and PCM provides an additional thermal resistance that delays the temperature rise of PCM, and therefore a relatively long sprint duration is required for the PCM to reach its melting temperate. The prolonged sprint can be mitigated if the interlayer material has a much higher thermal conductivity (for instance, crystalline Si). Further extending the sprinting duration may lead to more PCM melting and more heat being buffered, as described in the next section. We note that this integrated PCM configuration is far from optimized due to difficulty in precisely machining the alumina package without damaging the die, and there is still significant room for improvement. In addition to the peak temperature suppression, another important metric is the cooldown time required for the chip to restore the initial temperature after a sprint [9]. A comparison of the cooling transients of the two chips (Fig. 6.3B and Fig. 6.3C) reveals that they return to the initial temperature after a similar time of around 7 seconds due to the presence of the breadboard acting as a heatsink, indicating the re-solidification of the PCM does not delay the cooling process when the chip returns to idle state. This is expected as the integrated PCM changes little the overall thermal properties of the ATC due to its small volume compared to that of the alumina package (Fig. 6.1B), and the heat released during re-solidification can be easily dissipated via the breadboard heatsink.

6.4 Experiment results for multiple sprint and cooldown cycles

We next address the temperature transients of multiple sprint and cooldown cycles by comparing ATCs with and without PCM filling and study how the duration of the sprint could impact the performance improvement induced by the integrated PCM. A duty cycle (sprint duration per cycle) that is less than the inverse of the sprint power contrast (sprint power/sustained power) will provide an off percentage large enough for the chip to cool down after a sprint, while the peak temperature asymptotically reaches its maximum during continuous sprint and cooldown cycling. Since the supplied power in the idle and sprinting state are 0.216 W and 2.35W, respectively, the maximum duty cycle for this power contrast is 1/11 to avoid significant heat accumulation of the chip; however, this constraint can be relaxed if the chip is properly heatsinked. For this reason, two duty cycles D (0.3 and 0.4, with 3 and 4 seconds of sprint in a 10-second cycle, respectively) exceeding the inverse of power contrast were chosen without inducing significant idle state temperature build-up by the end of the 80-second experiment run (Fig. 6.4). The idle state die temperature gradually increases from 32.1 ºC to 36.2 ºC for the cycle with $D = 0.3$ and from 31.7 ºC to 34.6 ºC for the cycle with $D = 0.4$. Their difference is within spatial temperature variations, evidencing the effectives of the breadboard heatsink. Due to the presence of an alumina interlayer sandwiched between the die and PCM, the temperature rise of the PCM is delayed due to the added thermal resistance and we found a sprint duration greater than ~2.5s is necessary for the PCM to begin melting. Consequently, PCM will function as a heat buffer only when sprint duration exceeds 2.5s, and the benefits of using PCM increases as sprint time grows. To illustrate the effect of sprint duration on the performance of PCM, the temperature transients of the same ATC with integrated PCM were run at two sprint durations. For the cycle with 3-second sprints (Fig. 6.4A), the backside PCM is just reaching 52ºC by the end of each sprint and therefore the buffering effect is insignificant with a ~6 ºC peak temperature suppression for the chip with PCM compared to the chip without. On the other hand, for the cycle with 5-second sprints (Fig. 6.4B, most backside PCM underwent melting by the end of each sprint, giving rise to an average peak temperature decrease of 19ºC. Therefore, an interesting trade-off exists between the duration of the sprint and the effectiveness of PCM, i.e., a short sprint duration of less than 1s is more effective in enhancing the user perception of device responsiveness in wearable electronics and mobile phones, but any added thermal resistance between heat source (processor) and PCM would require an extended sprint duration to maximize the benefits of using PCMs. Though reducing the interlayer thickness by placing the PCM closer to the heat source can decrease thermal resistance, this approach would

likely cause the power density to grow, in which case the PCM latent heat can be more easily overwhelmed by power density as reasoned in Section II. B. Therefore, it is more practical to use high thermal conductivity material between processor and PCM as well as apply thermal grease/thermal interface materials to suppress the thermal resistance between various interfaces such that short sprint duration and effective utilization of PCM can be achieved simultaneously.

The simplified assumption that power consumed by the processor is proportional to the computational work indicates the ATC can get 11 times more work done in the sprinting state than in idle state, and the work published by Raghavan, *et al* suggests sprinting can get more computations done than sustained operation due to a race-to-idle effect[9]. Currently, compute intensive applications like vision and speech recognitions are routinely deferred to the cloud in wearable electronics and mobile devices. A sprinting device could potentially allow such tasks to be completed locally within the acceptable response-time limits of interactive applications, thereby giving rise to better control and flexibility of the device as its operations can be self-contained. Furthermore, using sprinting to execute typical wearable electronics and phone workloads, such as web browsing and application initialization, etc., may also greatly improve responsiveness. Therefore, sprinting can potentially significantly enhance the end-user perception of quality and performance for consumer wearable electronics and mobile devices.

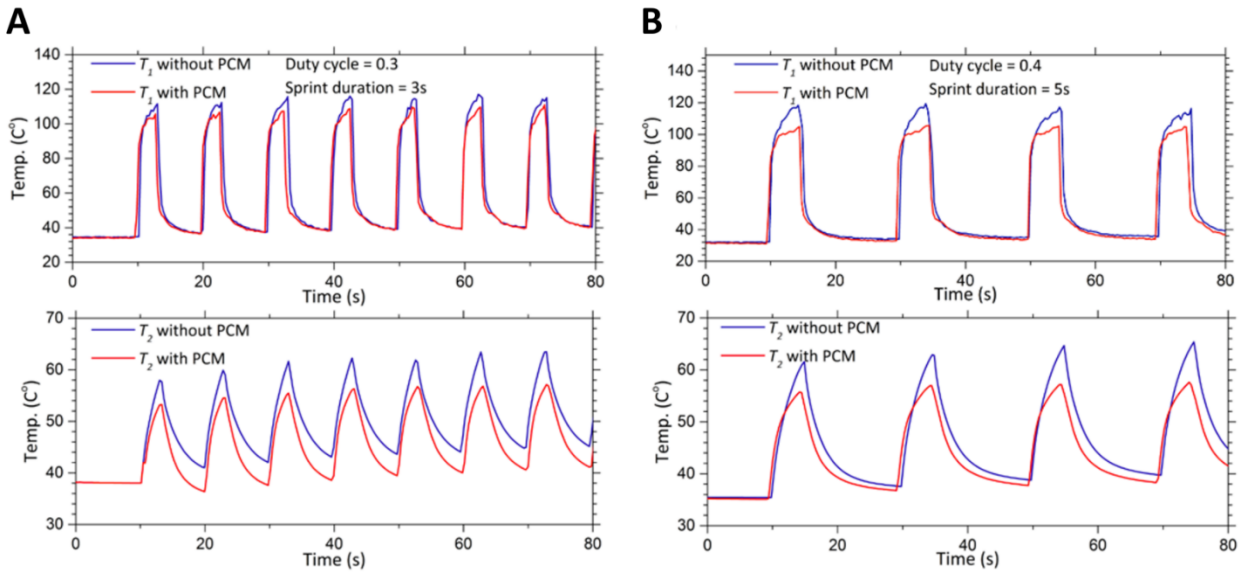

Figure 6.4 | Temperature transients of ATCs with and without integrated PCM for multiple sprint-cooldown cycles. (A) Die temperature (*top*) and backside PCM temperature transients (*bottom*) of the ATC for a duty cycle of 0.3 and sprint duration of 3s. **(B)** Die temperature (*top*) and backside PCM temperature transients (*bottom*) of the ATC for a duty cycle of 0.4 and a sprint duration of 5s.

6.5 Discussion and summary

In summary, the thermal implications of computational sprinting using a high-performance accelerator chip as test object were investigated. By integrating a fast-thermal-response metallic PCM into the chip package, a meaningful 17% drop in die peak temperature suggests a promising thermal buffering effect that can be explored and optimized further to manage heat dissipation in volume-sensitive devices. Multiple sprint and cooldown cycles were repeated to demonstrate the short-term repeatability of the PCM heatsink. The integrated PCM heatsink can potentially improve the reliability of the chip by keeping its peak temperature lower.

References

1. S. Borkar, A. A. Chien. The future of microprocessors. **54**, 67-77 (2011).

2. H. Esmaeilzadeh, E. Blem, R. S. Amant, K. Sankaralingam, D. Burger. Dark silicon and the end of multicore scaling. 38th *Annual International Symposium on Computer Architecture (ISCA)*. 365-376 (2011).

3. A. L. Moore, L. Shi. Emerging challenges and materials for thermal management of electronics. *Materials Today* **17**, 163-174 (2014).

4. P. Abdulla. 21st Century Computer Architecture. 1-16 (2013).

5. N. Hardavellas, M. Ferdman, B. Falsafi, A. J. I. M. Ailamaki. Toward dark silicon in servers. *IEEE Micro.* **31**, 6-15 (2011).

6. M. B. Taylor. Is dark silicon useful? Harnessing the four horsemen of the coming dark silicon apocalypse. *49th ACM/EDAC/IEEE Design Automation Conference (DAC).* 1131-1136 (2012).

7. L. Yan, L. Zhong, N. K. Jha. User-perceived latency driven voltage scaling for interactive applications. *Proceedings of the 42nd annual Design Automation Conference*. 624-627 (2005).

8. A. Shye, B. Scholbrock, G. Memik. Into the wild: Studying real user activity patterns to guide power optimizations for mobile architectures. *Proceedings of the 42nd Annual IEEE/ACM International Symposium on Microarchitecture*. 168-178 (2009).

9. A. Raghavan *et al.* Computational sprinting. *2012 IEEE 18th International Symposium on High Performance Computer Architecture (HPCA)*. 1-12 (2012).

10. A. Raghavan *et al.* Utilizing dark silicon to save energy with computational sprinting. *IEEE Micro* **33**, 20-28 (2013).

11. S. Fan, S. M. Zahedi, and B. C. Lee. The computational sprinting game. *Proc. of the Twenty-First International Conference on Architectural Support for Programming Languages and Operating Systems* **44**, 561-575 (2016).

12. L. Shao *et al.* On-chip phase change heat sinks designed for computational sprinting. *Proceedings of the 30th IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM).* 29-34 (2014).

13. S. Lu, Z. Zhang, M. Papaefthymiou. A 4.5 GHz 0.84TOPS/mm^2 Neural Network Engine with Stream Architecture and Resonant Clock Mesh. *IEEE Asian Solid-State Circuits Conference (A-SSCC).* 133-136 (2016).

14. J. K. Kim, P. Knag, T. Chen, Z. Zhang. A 640M pixels/s 3.65mW Sparse Event-Driven Neuromorphic Object Recognition Processor with On-Chip Learning. *IEEE VLSI Circuits*. C50-C51 (2015).

14. J. Sim *et al.* A 1.42TOPS/W Deep Convolutional Neural Network Recognition Processor for Intelligent IoE Systems. *2016 IEEE International Solid-State Circuits Conference (ISSCC)*. 264-265 (2016).

16. Y.-H. Chen, T. Krishna, J. S. Emer, V. J. I. J. o. S.-S. C. Sze, Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks. *2016 IEEE International Solid-State Circuits Conference* **52**, 127-138 (2017).

17. R. Kandasamy, X.-Q. Wang, and A. Mujumdar. Transient cooling of electronics using phase change material (PCM)-based heat sinks. Applied Thermal Engineering **28**, 1047-1057 (2008).

18. R. Kanapady. Experimental and computational evaluation of phase change materials for handheld computing. *15th IEEE ITHERM Conference*. 118-124 (2016)

19. L. Shao, *et al*. Figure-of-merit for phase-change materials used in thermal management. *International Journal of Heat and Mass Transfer* **101**, 764-771 (2016).

20. X. Liu, F. Qian. Poster: Measuring and optimizing Android smartwatch energy consumption. *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking. (ACM).* 421-423 (2016).

21. R. LiKamWa, Z. Wang, A. Carroll, F. X. Lin, and L. Zhong. Draining our glass: an energy and heat characterization of Google glass. *Proceedings of 5th Asia-Pacific Workshop on Systems. (ACM).* 1-6 (2014).

Chapter 7
Summary and Future Work

7.1 Summary

As device miniaturization and power scaling progress, the challenge posed by a rising heat flux and/or total power with increasingly stringent reliability constraints make thermal management a key enabling technology in future electronic and optoelectronic devices. The objective of the work presented in this thesis is to address several aspects of this challenge, starting from advanced microscale thermal characterization techniques, to heat mitigation strategies at the device and package levels.

Specifically, in applying a CCD-thermoreflectance technique to study GaAs-based high-power diode lasers with back-irradiance (BI), it is found that heating caused by substrate absorption of optical feedback can pose a significant reliability concern in devices fabricated on an absorbing substrate. Because the CCD-TR technique offers high spatial resolution, we were able to determine the temperature of the hotspot located at the active region of the device having a characteristic dimension of ~1.5 μm. Short-term reliability tests were implemented to quantify the correlation between the level of back-irradiance and a reduction in lifetime. By bounding the activation energy of BI-induced catastrophic optical damage (COD), we deduce that the effect of BI in accelerating device degradation is likely grounded in thermal. Studying diode lasers at different wavelengths and compare their susceptibility to BI-induced damage reveal a strong dependence on wavelength for this mode of feedback-induced heating. We then developed an active cooling strategy that can be scaled down to the size of hotspots in devices for high cooling efficiency, taking advantage of CNTs' exceptionally high electrical conductivity and an effective energy filtering effect afforded by intertube contacts. The CNT cooler exhibits a large effective Seebeck coefficient (386μV/K) and provides a promising miniaturized cooling solution that can effectively suppress hotspots.

The next two chapters of the thesis each describe an innovative strategy for heat mitigation at the package level. Firstly, a new approach has been proposed to improve the thermal conductivity of

amorphous polymers. Utilizing electrostatically repulsive forces between ionized pendant groups to linearize and expand the curved polymer backbones, we achieved a 3.5-time enhancement in thermal conductivity (0.34 $Wm^{-1}K^{-1}$ to 1.17 $Wm^{-1}K^{-1}$) of poly(acrylic acid). The strategy of isotropic chain extension can be potentially applied to increasing the thermal conductivity of polymeric matrix used in thermal interface materials, making them more effective in transferring heat across adjoining interfaces. Secondly, for devices subject to a bursty heat source in a small-volume package, a phase change material (PCM) heatsink is proposed for temporal heat storage. The thermal properties of the PCM is carefully chosen such that the melting time is comparable with the duration of intense computation, so that it can store heat via phase change. The PCM is back-filled in a real-world package of an accelerator chip to minimize package volume and demonstrates a meaningful reduction in the die peak temperature compared to a reference package without PCM.

7.2 Future work

This section discusses several directions in which the current work can be extended.

7.2.1 Improve the thermal and mechanical properties of graphene/polymer composite

An important aspect of polymer-based materials is their mechanical property, which is especially crucial in automobile and aircraft applications. Here in, we propose to engineer high performance polymer composites that integrate both heat-dissipating and structural functions, which would significantly increase the range of automotive applications for which polymers can be used.

Specifically, we propose to use experimental and computational methods to study the mechanisms that limit heat transfer in filled polymers, focusing in particular on interfaces between graphene (Gr) based fillers and the surrounding polymer matrix. Based on this knowledge, we will then design polymer composites with superior thermal and mechanical properties. Graphene has an extremely high intrinsic in-plane lattice thermal conductivity[1,2] (>2000 $Wm^{-1}K^{-1}$) and can form percolating thermal pathways at extremely small loading fractions. However, polymer/Gr composites typically have a thermal conductivity far below what the simple rule of mixtures would predict, as interactions between Gr and the surrounding matrix reduce its intrinsic thermal conductivity and thermal boundary resistance (TBR) additionally develops at the Gr/matrix interface. To systematically assess these interfacial effects, we will use Gr/SAM (self-assembled

monolayer) structures that offer molecular design tunability and synthetic reproducibility. Since the TBR of molecular interfaces has been shown to scale inversely with the strength of interfacial bonding, we will oxidize Gr in a controlled manner and functionalize it with chemical moieties capable of forming strong covalent bonds with SAM terminal groups. Experimental Gr/SAM results will be used to validate molecular dynamics simulations that are then used to design filled composites with superior thermal and mechanical properties. The thermal and mechanical properties of these composites will then be measured to assess the effectiveness of the molecular/bonding design strategies

7.2.2 Use CCD-TR to probe the early stages of catastrophic optical damage

While catastrophic optical damage (COD) is well-known as a failure mechanism in diode lasers, the extremely early stages of COD comprising a sequence of events that eventually leads to device failure are still an active area of research due to their complicated nature and fast dynamics[3]. While an occurrence of dark-line damage patterns at the outcoupling facet observed under microscope is usually regarded as the symptom of COD (Fig. 7.2.1), in one of our thermoreflectance experiment we saw what seems to be a precursor of those dark-line defects (DLDs, see Fig. 7.2.2).

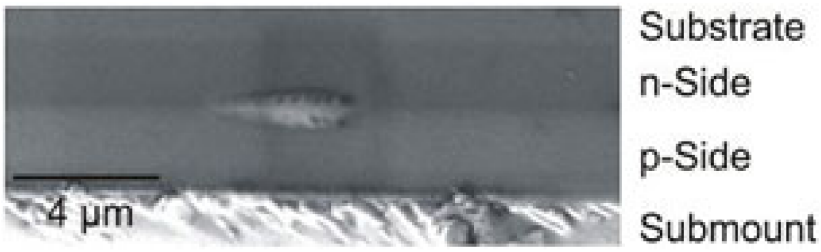

Figure 7.2.1 | Dark-line defects at the outcoupling facet observed under a microscope. Image courtesy of ref. [3].

Specifically, in one set of the TR experiment for the 800 nm laser with a bias current of 2A and an R_{eff} = 24.3% (with a BI power P_{BI} equal to 0.466 W), as the BI spot is swept from deep into the solder to the substrate, a small bright feature (possibly defects) appeared in the TR image when BI is located at X = -1 μm (the QW position is set to be the reference point with the direction going into the solder as negative). No such feature was found in the microscope image (first three bottom panels of Fig. 7.2.2). Then as more sets of TR experiments were taken with the BI spot moving towards the substrate in increment of 1 μm, the bright feature was observed to grow (Left to right in Fig. 18). Eventually, at the critical location (x = +2 μm) at which BI leads to maximal ΔT_{QW}, a line-defect symptomatic of COD appeared both in the TR and microscope image. Each TRBI experiment lasts for about 2 minutes with the bias current alternating between 0 and 2A at 4Hz with a 50% duty cycle, so the BI spot dwell time during each experiment is approximately 1 minute.

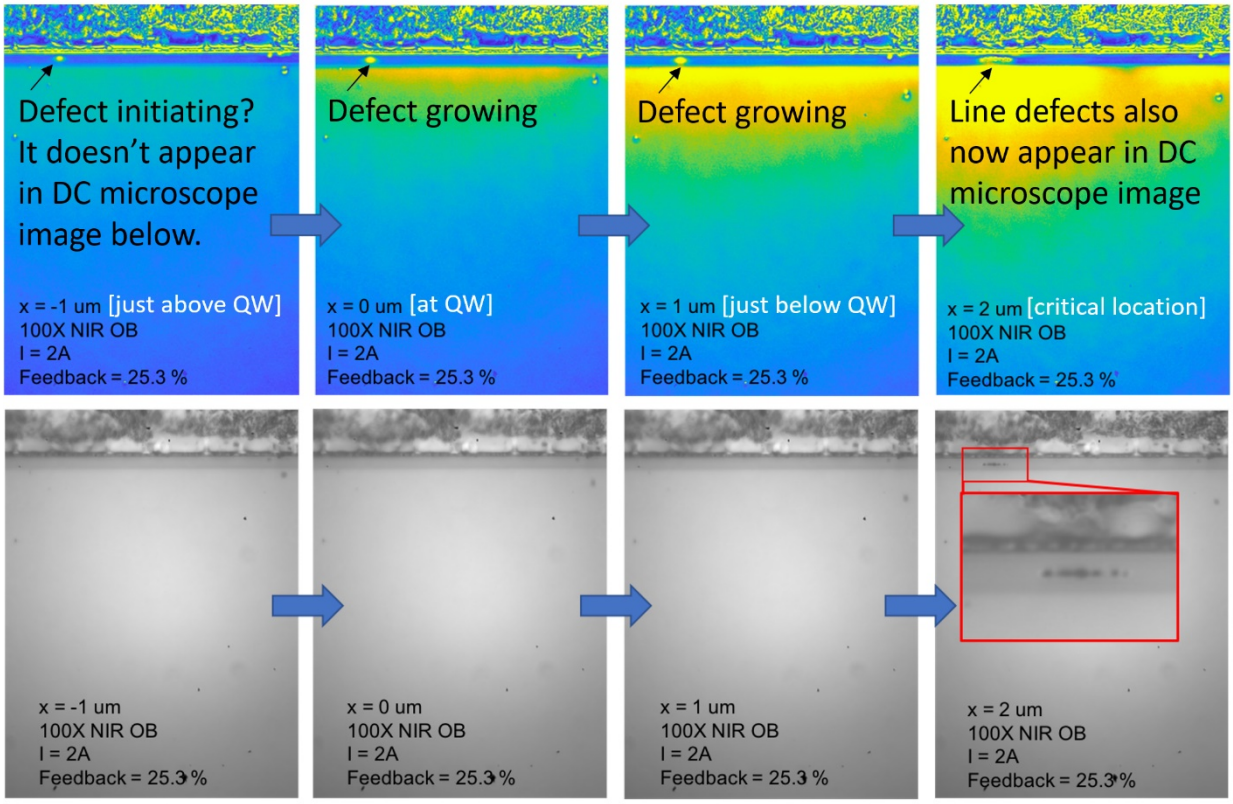

Figure 7.2.2 | Evolution of a potential precursor of the dark-line defects. The top panels depict the thermoreflectance images with a small defect that eventually develops into a dark line defect typically seen in lasers that suffers COD. The bottom panels depict the DC microscope images, only after the precursor turns into a full-fledge line defect does it appear in the microscope image.

The fact that the precursor appears brighter in the TR images suggests it is either a localized hotspot or a phase transformed defect having a different thermoreflectance coefficient or could be a combination of these two. In any case, the fact that it manifests itself before the dark-line defects appear in the DC image suggests it could be a precursor of the DLD. Therefore, we propose to replicate this precursor before it is fully developed, by performing a long-term thermoreflectance imaging test. We will run a fresh 800 nm device at high bias current (4.5 – 5A) while taking TR images continuously. Using 500 iterations for data averaging, we propose a moving-average scheme where new TR images (for instance, the TR images taken in the most recent 30 iterations) are constantly added to update the TR map with equivalent number of old TR images being excluded. This will be done continuously for weeks until a defect precursor is seen in the TR map, at which point the device will be shut off and be scrutinized under SEM and cathode luminescence to reveal more details regarding the material composition of the bright spot that appears in the TR image.

References

1. A. A. Balandin, *et al*. Superior thermal conductivity of single-layer graphene. *Nano. Lett*. **8**, 902-907 (2008).

2. E. Pop, V. Varshney, and A. K. Roy. Thermal properties of graphene: Fundamentals and applications. *MRS Bull.* **37**, 1273 (2012).

3. M. Hempel, *et al*. Microscopic origins of catastrophic optical damage in diode lasers. *IEEE Journal of Selected Topics in Quantum Electronics* **19**, 1500508 (2013).