THERMALLY AWARE DESIGN APPROACHES FOR HIGH POWER DENSITY ULTRA-WIDE BANDGAP POWER ELECTRONICS

A Dissertation
Presented to
The Academic Faculty

by

Samuel H. Kim

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
George W. Woodruff School of Mechanical Engineering

Georgia Institute of Technology
December 2022

COPYRIGHT © 2022 BY SAMUEL KIM

THERMALLY AWARE DESIGN APPROACHES FOR HIGH POWER DENSITY ULTRA-WIDE BANDGAP POWER ELECTRONICS

Approved by:

Dr. Samuel Graham, Advisor
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Sukwon Choi
School of Mechanical Engineering
The Pennsylvania State University

Dr. Suresh Sitaraman
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Andrew Green
Sensors Directorate
Air Force Research Laboratory

Dr. Satish Kumar
School of Mechanical Engineering
Georgia Institute of Technology

Date Approved: [August 16, 2022]

To my family

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to my advisor, Professor Samuel Graham, for the continuous support of my PhD study and research, for his patience, motivation, enthusiasm, and immense knowledge. I have gained an enormous amount of knowledge and learned several valuable lessons. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my study.

Besides my advisor, I would like to thank Professor Sukwon Choi (The Pennsylvania State University) for spending the time to discuss with me on a weekly basis about my research for last two years. The knowledge he has provided to me as well as the research questions he has pushed me to answer, have contributed significantly to results presented in this thesis. I am also sincerely grateful to AFRL for their financial support for which I would not have been able to complete my PhD without this. I would also like to thank Dr. Andrew Green, Dr. Kelson Chabak, Dr. Kyle Liddy, and Dr. Eric Heller for participating in fruitful conversations about my research.

I would also like to extend my sincere thanks to Professor Suresh Sitaraman and Professor Satish Kumar for serving on my thesis committee. I am grateful for their time taken to carefully review my work and provide me with advice. Several collaborators have made this thesis possible by either providing high quality devices or materials for analysis. I would like to acknowledge the following collaborators: Professor Asif Khan at University of South Carolina for providing me high Al-content AlGaN transistors, Dr. Sreekant Narumanchi, Dr. Paul Paret, and Dr. Gilbert Moreno at NREL for their deep knowledge in

thermal management of UWBG electronics, and Professor James Speck and Dr. Yuewei Zhang at UCSB for providing geometrical and electrical information of the device.

I would like to thank my fellow graduate students, past and present, for their support and scientific collaborations in Graham Lab: In no particular order, thanks are deserved to Dr. Georges Pavlidis, Dr. Luke Yates, Dr. Bikramjit Chatterjee, Nicholas Hines, Dr. Zhe Cheng, Dr. Jingjing Shi, Dr. Darshan Pahinkar, Dr. Chao Yuan, Dr. Riley Hanus, Dr. Brian Kelly, Dr. Jinho Hah, Dr. Kyungjin Kim, Dr. Gabe Cahn, Jason Hirschey, Chidnma Imediegwu, Robert Montgomery, Michael Sulkis, and Anusha Krishnan.

Outside of the Graham Group, many others contributed to my success. I would like to thank the students and friends that I worked with during the qualifying exams, Dr. Sungkun Hwang, Dr. Sangbeom Cho, Dr. Dudong Feng, Myung Joon Kim, and Daniel Lee. Also, I would like to thank colleagues from PSU: Dr. James Spencer Lundh, Daniel Shoemaker, and Yiwen Song. Without their help, I would not be able to finish my thesis.

Most of all, I am extremely appreciative to my parents who supported me throughout my academic journey. I would like to thank my parents for always believing in me, for their continuous love and their support in my decisions. Lastly, I would like to thank my wife, Dr. Sooji Ha, for accompanying me all the time and helping me go through all my lows and highs. I would not have completed my PhD degree without her heartwarming kindness and care.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

UWBG	Ultra-wide Bandgap
WBG	Wide Bandgap
β-Ga_2O_3	β-type Gallium Oxide
GaN	Gallium Nitride
SiC	Silicon Carbide
SEP	Solar Electric Propulsion
BFOM	Baliga Figure of Merit
JFOM	Johnson Figure of Merit
MOSFET	Metal-oxide Semiconductor Field Effect Transistors
CAVET	Current Aperture Vertical Electron Transistors
FET	Field Effect Transistors
HEMT	High Electron Mobility Transistor
TBC	Thermal Boundary Conductance
TBR	Thermal Boundary Resistance
IR	Infrared
TTI	Transient Thermoreflectance Imaging
HTC	Heat Transfer Coefficient
ΔT	Temperature Rise
V_{DS}	Drain-Source Voltage
V_{GS}	Gate-Source Voltage
L_{GD}	Gate-to-drain Spacing
CBL	Current Blocking Layer

SUMMARY

Ultra-wide bandgap (UWBG) semiconductors like β-type gallium oxide (β-Ga_2O_3) show promise for the development of next-generation high power density electronics devices such as RF and power electronics. The large bandgap (4.8 eV), high breakdown fields (8 MV/cm), and excellent thermal stability of β-Ga_2O_3 give promise to the production of low-loss power switching devices with large breakdown voltage, and potentially allows for high-temperature and deep space operation. However, a major drawback of β-Ga_2O_3 arises from its poor thermal conductivity, which results in devices with unacceptably high junction-to-package thermal resistance. While there is considerable promise for future devices made from UWBG materials, their adoption as a technology will hinge upon novel approaches to address heat dissipation at the die level which will enable high power density operation. The aims of this thesis are i) to develop novel thermal management strategies to reduce the junction-to-package thermal resistance for devices made from low thermal conductivity UWBG materials for both lateral and vertical devices, ii) to conduct an analysis of architectures for homoepitaxial β-Ga_2O_3 metal-oxide semiconductor field effect transistors (MOSFETs) to optimize the device thermal performance and verify experimentally, and iii) to optimize thermal management design for both steady-state and transient-state of UWBG transistors. Overall, the optimal thermally-aware design for vertical and lateral structures for steady-state and transient applications will be provided by investigating the device layout such as substrate orientation, configuration of electrodes (number of fingers, channel width, location of metallization pads), dielectric heat spreader, and thermal boundary conductance between metal and β-Ga_2O_3.

CHAPTER 1. INTRODUCTION

1.1 Background and Motivation

The demand of power electronics is increasing as the use of electric power in systems (e.g., industrial motors, smart grid, transportation) continues to grow. The future of these systems requires the use of electronics to process and control the flow of electric power to improve efficiency and functionality [1]. Electrical energy constitutes 40% of total primary energy consumption in the United States, and the percentage is expected to increase rapidly due to broad adoption of electric vehicles, production of renewables, and other factors. It is anticipated that 30% of all electrical energy passes through power electronics today and this number could reach 80% in the next decade [2], [3].

The semiconductor material that is most often used in power electronics is Silicon (Si) due to ease of processing and availability. However, the performance of Si power devices is now approaching the operational limits set by its intrinsic material properties. Wide-bandgap semiconductors (WBG), such as Gallium Nitride (GaN) and Silicon Carbide (SiC), are gaining more attention, since they can withstand higher voltages, temperatures, and electromagnetic radiation than Si power electronics before experiencing breakdown as shown in Table 1 and Figure 1 (a). Power electronics applications range from on-chip power converters to high voltage rectifiers for electric power transmission lines. Applications requiring high-voltage, high-current, radiation-tolerant power electronics include electronic thrust control actuators of rockets and missiles, and solar electric propulsion (SEP) of spacecraft. The SEP system proposed by NASA would enable human

exploration missions outside of earth's orbit, and requires functional power devices with high blocking voltages, high current capability, and high tolerance for radiation effects [4].

Table 1. Electronic properties of Ga_2O_3 and other semiconductor materials

Material Property	Si	4H-SiC	GaN	β-Ga_2O_3
Bandgap (eV)	1.1	3.25	3.4	4.6-4.9
Electron mobility (cm^2/V-s)	1480	1000	1500	200
Breakdown field (MV/cm)	0.3	2.5	2.8	8
Dielectric constant	11.7	9.8	10.4	10
Normalized BFOM	1	320	860	1100-3250
Normalized JFOM	1	8.2	22.9	37.5

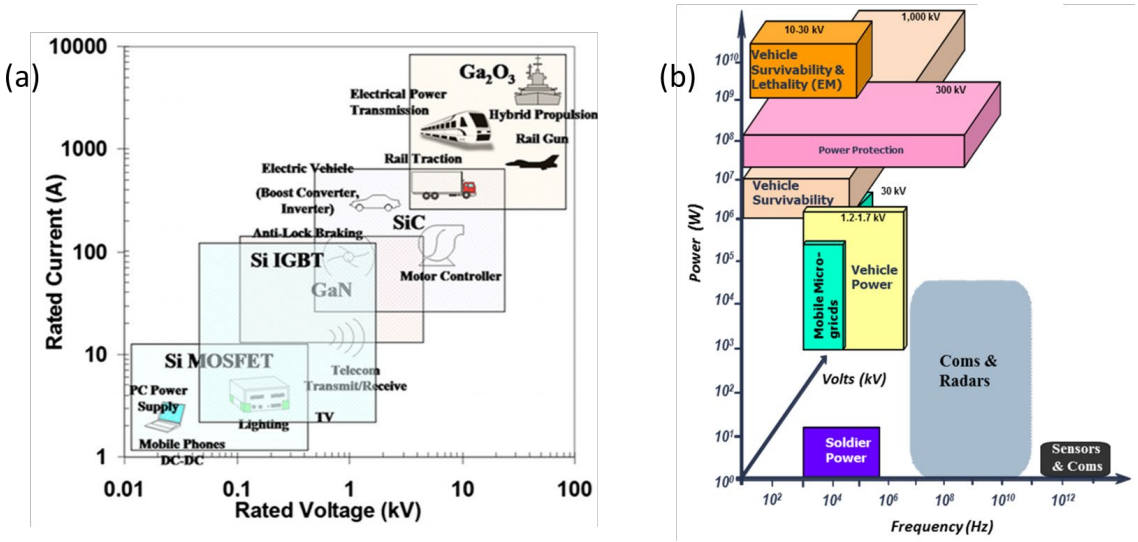

Figure 1. (a) Applications for Si, SiC, GaN, and Ga_2O_3 power electronics in terms of current and voltage requirements. Possible applications for Ga_2O_3 include fast chargers for electric vehicles, high voltage direct current for data centers, industrial motors, and alternative energy sources. (b) Map of applications for pulsed power devices, organized by switching frequency and power.

1.2 β-Ga_2O_3

Considering the range of WBG materials being studied for power electronics devices, beta-Gallium Oxide (β-Ga_2O_3) is a promising semiconductor material because of its unique combination of material properties – properties that make it excellent for power electronics applications. Major requirements for an efficient high-voltage switching transistors are (i) low on-resistance (R_{on}), (ii) low off-state leakage current, (iii) large breakdown voltage (V_{Br}), and (iv) high-temperature operation. As shown in Table 1, the large bandgap of β-Ga_2O_3 (4.6 - 4.9 eV) and high breakdown field (8 MV/cm) lead to an outstanding Baliga's figure of merit (BFOM), which is defined as V_{Br}^2/R_{on}, with superior thermal stability for low-frequency power switches [5]–[8]. The extraordinary BFOM of β-Ga_2O_3 due to high critical electric field strength, which is 4-10 times higher than GaN and SiC as shown in Table 1, enables the production of power electronics devices that possess low on-resistance (less conduction loss) and high breakdown voltage to block a large voltage during off state [9]. For higher frequency, from 100 kHz to 1 MHz, the device spends more time switching versus being on or off. The key measure of semiconductor for high-frequency power switches is called the Jonson figure of merit (JFOM), which can be defined as $v_{sat}E_c/2\pi$, where v_{sat} is the saturated carrier velocity and E_c is the critical electric field or breakdown field [10]. Losses during switching are the product of both the device's resistance and how much charge needs to accumulate on the transistor gate in order to make the switch [5]. JFOM of β-Ga_2O_3 is still 1.5-4.5 times higher than that of GaN and SiC as shown in Table 1. In addition to higher JFOM, there are benefits of faster switching in power electronics. The bulkiest part of that system are the transformer and other passive components, and smaller devices could be used if the frequency is increased. For example,

a 1200-V Si inverter switching at 20 kHz can deliver around 3 kilowatts, while by switching at 150 kHz, a SiC inverter delivering the same power can operate at higher temperature in a package that is one-third the size [10], [11]. Additionally, the greatest advantage of β-Ga_2O_3 may lies in the availability of affordable, high-quality, and large native substrate, which are the essential features for economical mass substrate production, alleviating concerns related to the high cost of SiC and GaN wafers [5], [12]. These properties of β-Ga_2O_3 hold promise for an improvement in the size, weight, and power as well as the cost of a broad range of power switching and RF components used in power supply, radar, electronic warfare, and communication systems as shown in Figure 1 (b) [5].

1.2.1 β-Ga_2O_3 Devices

A transistor is the most basic and important component for power semiconductor electronics. Various β-Ga_2O_3 transistors have been developed since 2012 as shown in Figure 2 to verify the viability of β-Ga_2O_3 as an electron device material. The swift progress of β-Ga_2O_3 metal-oxide-semiconductor field-effect transistors (MOSFETs) with breakdown voltages reaching greater than 2 kV has established β-Ga_2O_3 as a pertinent candidate for power switching technologies [13]–[16]. Normally-off field-effect transistors (FETs) are strongly demanded for switching device applications. However, preliminary reports [17]–[19] of the device characteristics of normally-off lateral FETs were far short of the requirements for practical applications that the development of devices is shifting from lateral to vertical geometry.

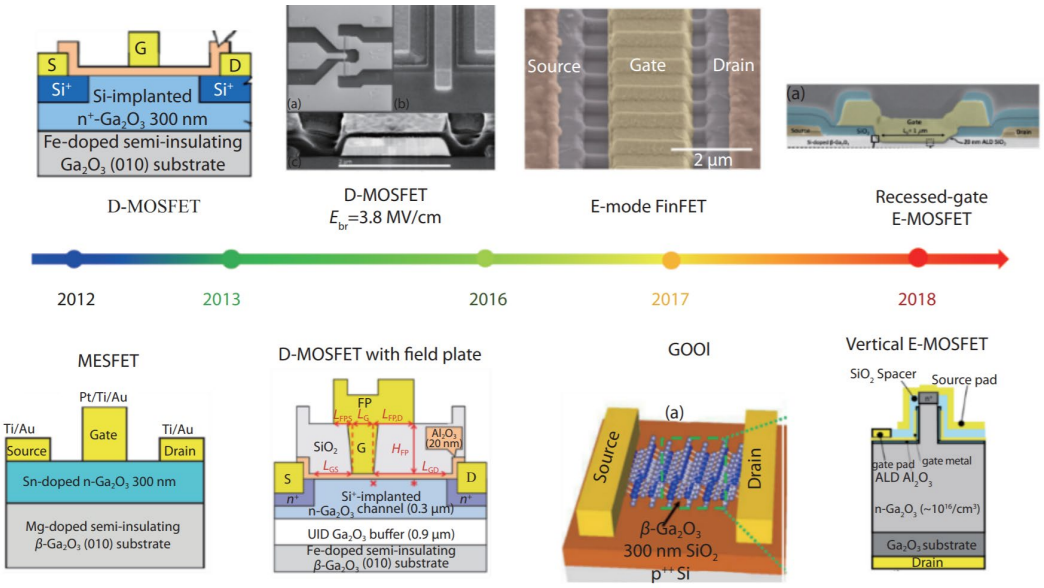

Figure 2. The development of β-Ga_2O_3 transistors in last 10 years [20].

The first β-Ga_2O_3 metal–semiconductor FET (MESFET) with a platinum Schottky gate was demonstrated in 2012 on Sn-doped β-Ga_2O_3 by molecular beam epitaxy on semi-insulating β-Ga_2O_3 substrates showing depletion-mode operation with 25 mA/mm maximum drain current and 250 V breakdown voltage [16]. A MOSFET with 40 mA/mm drain current and 370 V breakdown voltage showed an on-to-off current ratio of 10^{10} [21]. MOSFETs with extrinsic gate oxides such as Al_2O_3 are more suited to support the large gate swing [19]. About 750 V breakdown strength was achieved by changing from epi-based n-type doping to Si ion implantation of the channel area in combination with 11 μm thick buffer layer between channel and substrate and by using field plates in the transistor design [22]. The normally off performance was demonstrated for a 500 V device by recessing the Si-doped β-Ga_2O_3 channel at the gate position from initially 200 nm down to 70 nm [23]. The highest reported breakdown strength of 1850 V was achieved with a

MOSFET using SiO_2 as gate insulator [24]. The average field between gate and drain of 92 V/µm is already close to values obtained for state-of-the-art lateral GaN HFETs. However, the area-specific R_{ON} is still more than one order of magnitude worst [25]. Lateral β-Ga_2O_3 transistors need a large gate swing since a thin well-confined transistor channel such as the 2DEG in GaN HEMTs is not available.

The high breakdown field combined with the moderate mobility makes β-Ga_2O_3 a promising candidate for high-power switching with voltages >1000 V. Here, the vertical device concept is of advantage, but device development is just starting. A current aperture vertical electron transistor (CAVET) with a current density of 1050 A/cm^2 was realized by growing the 3×10^{16} cm^{-3}-doped drift layer on an Sn-doped β-Ga_2O_3 substrate [26]. Since there is no p-type doping available in β-Ga_2O_3, the current blocking layer (CBL) was defined by Mg-ion-implanted insulating β-Ga_2O_3. However, the presented device was not blocking because of gate–source leakage issues. A normally off FinFET structure with a current density of 320 A/cm^2 was demonstrated by adopting an HVPE-grown 10 µm thick Sn-doped drift layer with electron concentration [27]. The average drift region field is 117 V/µm and yet very similar to the average breakdown fields observed for lateral and vertical GaN transistors. Similar to vertical GaN devices, a technology for laterally structured p–n junctions are not yet available. Alternative concepts for gate finger field engineering and for an edge-termination structures must be developed to fully exploit the high material breakdown strength of β-Ga_2O_3 inside a real transistor.

1.2.2 Thermal Reliability

While there have been great steps in improving electrical device performance, thermal management is still a considerable bottleneck that is limiting device performance. Self-heating can be detrimental to device output performance and reliability due to physical damage in the device. Heat generation under high-power operation cannot be avoided that overheating is a critical challenge for the reliability of these state-of-the-art device technologies. A smaller device footprint combined with a greater power handling capability means substantially increased power densities for individual devices. The performance of many semiconductor devices is limited by heat dissipation capacity since the drift-layer resistance increases with rising operation temperature due to the decrease in electron mobility [8]. In addition, as shown in Figure 3 (a), increased localized power densities in smaller transistors (~90 kW/cm^2) can exceed the heat flux of sun surface (6.3 kW/cm^2), resulting in extremely high channel temperature (>200 °C) that degrade device performance and reliability [28], [29]. The lifetime of semiconductor devices is strongly related to its channel temperature that the lifetime of GaN devices could increase about ten times if the channel temperature decreases by 25 K as shown in Figure 3 (b) [30]. Also, extremely high channel temperature is the most critical stressors for power electronics for reliability concerns, while there are other stresses such as humidity, mechanical vibration, and radiation [31], [32].

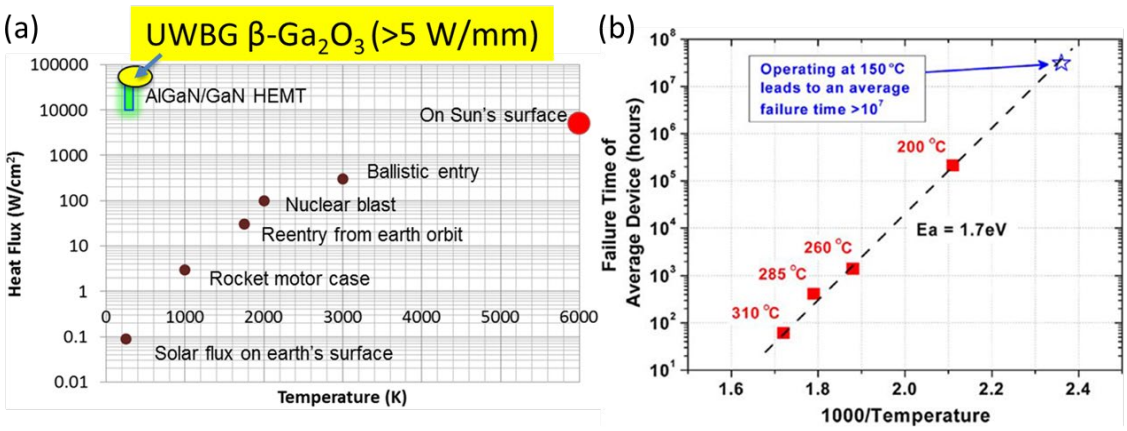

Figure 3. (a) Heat fluxes vs temperature for various occasions, (b) Lifetime of the GaN-on-Si HEMT device by Arrhenius relationship.

To understand the thermal limitations in the devices, for example, Near Junction Thermal Transport (NJTT) program through the Defense Advanced Research Projects Agency (DARPA) was able to achieve a 2.7× reduction in the thermal resistivity of a high electron mobility transistor (HEMT) device compared to a GaN-on-SiC device and demonstrated a 3× increase in the areal dissipation density of GaN-on-diamond comparatively with GaN-on-SiC [33]. It was found that the relatively large thermal boundary resistance (TBR) between GaN-diamond interface was 47.6 m^2K/GW, which was a major component in constraining the device performance so that several works were done to minimize TBR of GaN-diamond interface [34]–[36], and now it is reduced to the near of the limit of the calculated value of 3.1 W/m^2-K [37].

1.2.3 Limitation of β-Ga_2O_3

β-Ga_2O_3 possesses a poor thermal conductivity as compared to other WBG semiconductors (e.g., GaN and SiC), as shown in Figure 4. Therefore, β-Ga_2O_3 devices suffer from device self-heating under nominal operating conditions. Also, due to the highly

anisotropic monoclinic crystal structure of β-Ga_2O_3, the thermal conductivity of β-Ga_2O_3 fall in the range of 9 to 26 W/m-K at room temperature, as shown in Figure 4 and Figure 5 (b) , leads to excessively high operational channel temperatures that compromise the device performance and the reliability [38]. In addition, simulations indicate that at 25 °C, 1% and

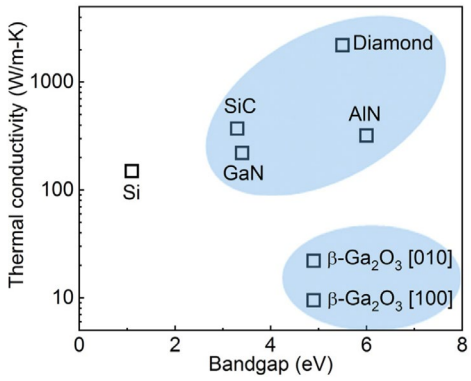

Figure 4. Thermal conductivity of wide/ultrawide bandgap semiconductors at room temperature [39]

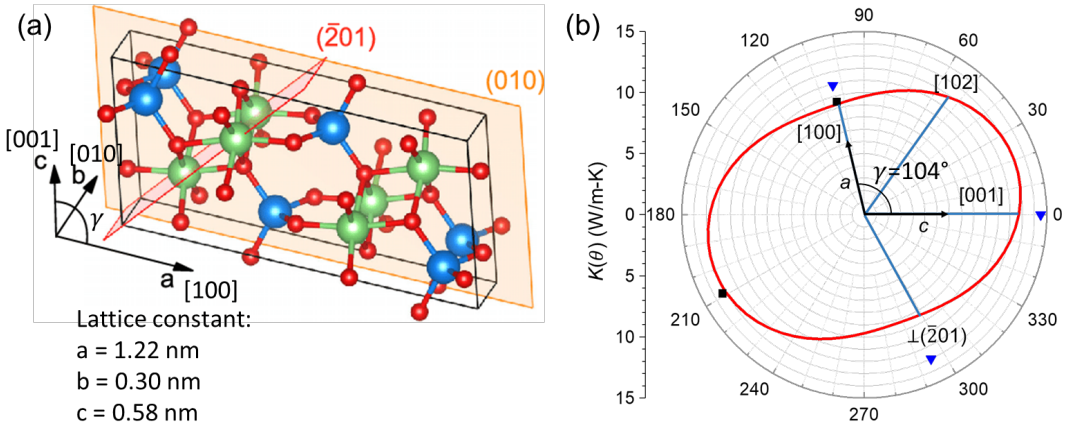

Figure 5. (a) Schematic representation of the unit cell of β-Ga_2O_3. The tetrahedrally coordinated Ga atoms are shown in blue whereas the octahedrally coordinated ones are shown in green. The oxygen atoms are marked in red. The (010) and the ($\bar{2}$01) planes are highlighted by the orange and red plane, respectively. (b) The directional dependence of the in-plane thermal conductivity of a (010) β-Ga_2O_3 substrate at room temperature. The red curve and square data points are from Jiang et al. [40], and the triangles are from Guo et al. [41].

2% oxygen vacancies decrease the thermal conductivity by 8.5% and 14.3% in [100] direction, 14.9% and 24.1% in [010] direction, 10.7% and 17.4% in [001] direction, respectively [7]. These values are one to two orders of magnitude lower than those of SiC and GaN semiconductors, while the size of the device is reduced for the equivalent capability resulting in higher power density. Thus, for smaller device with higher power density, much higher channel temperature will occur in low thermal conductive β-Ga_2O_3 devices compared to SiC and GaN devices, causing faster degradation and shorter device lifetime. Therefore, the low thermal conductivity is a serious potential weakness of β-Ga_2O_3 and will be one of the most important challenges among various R&D topics of β-Ga_2O_3 power device technologies.

1.3 Research Objectives

Based on the information outlined above, since the extremely high device temperature affects the performance and the reliability of the power electronics, heat removal is very important for the reliability concerns. Even though there are several studies for self-heating mitigation strategies for β-Ga_2O_3 transistors, we are far from having in hand optimal thermally aware design for these devices, and the thermal design for extremely low thermal conductivity material should be different from that of SiC and GaN devices. Therefore, this present work aims to continue the advancement in providing optimal thermal management solutions to β-Ga_2O_3 base power electronics. In order to better understand this and provide valuable insight to the scientific community, three overarching questions will be scrutinized with experimental methods and simulation techniques.

- **Is it possible to reduce the thermal resistance of device by adjusting design parameters related to the device layout? How does each parameter affect device self-heating?**

  This question is addressed through understanding the major factors for the thermal resistance in the devices: the thickness of the substrate, the highly anisotropic crystal structure of β-Ga_2O_3, and the thermal boundary conductance (TBC) between metal and oxide layer.

Figure 6 shows the heat generation region and the heat flow for both lateral and vertical devices, and possible design spaces for the optimal thermally aware design of β-Ga_2O_3 devices. Highly anisotropic thermal conductivity will affect thermal performance of

devices, and especially, lateral multi-finger device will most be influenced. Also, for multi-finger devices, gate-to-drain spacing and the width of the channel would have serious impact on electrical performance [42] and thermal performance. For both lateral and vertical structures, TBC between metal and oxide layer should affect the thermal performances as well. Few interfaces of metal/β-Ga_2O_3 have been reported to understand the role of the TBC in thermal transport [43]–[45].

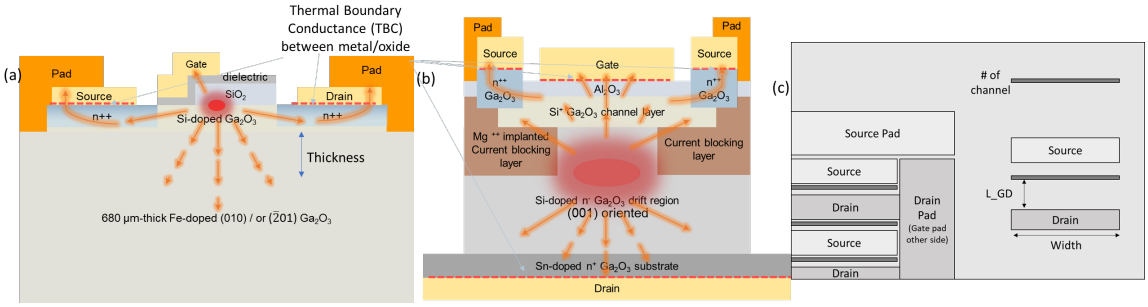

Figure 6. (a) Schematic of a typical lateral β-Ga_2O_3 structure (MOSFET) demonstrating the hotspot occurring in the device. Heat generated in the channel on the drain side of the gate must traverse low thermal conductive β-Ga_2O_3 substrate. (b) A vertical β-Ga_2O_3 structure (CAVET) that has a much more uniform thermal profile due to the device design. (c) Quarter symmetry model of lateral MOSFET showing possible design parameters: gate-to-drain spacing (L_{GD}), number of channels, width of the channel, size of the probe pad.

For high voltage and high power applications, vertical structures are preferred since chip area utilization is more efficient, device operation is insensitive to surface effect, and superior field termination is possible that reduce sharp regions of high electric field [46]–[48]. In addition, heat distribution is more uniform in vertical devices than in lateral devices since the latter tends to develop localized hot spots within a thin conducting channel confined near the surface [48]. Unlike lateral devices, however, thickness of the drift layer (>90% of the device thickness) defines the breakdown voltage of the device that cannot be thinned resulting in very high peak temperature due to high thermal resistive β-Ga_2O_3 layer.

Therefore, it is important to find the optimal external cooling solution. Here, the primary research question that will be addressed is:

- **Is it possible to add an external cooling structure like heat spreader on the current aperture vertical electron transistors (CAVET) to reduce thermal resistance?**

  This question is addressed through simulation developed by 2D Silvaco TCAD, which can model the Joule heat power profile of the device. For these computational studies, we explored the effectiveness of bottom-sided cooling and double-sided cooling methods with various design parameters such as thermal conductivity of die attach material, thermal boundary conductance (TBC) between metal and β-Ga_2O_3, and boundary conditions. Since there is no thermal management solution for the CAVET devices, this study will provide a set of useful plots that would allow a designer to quickly make decisions on the optimal packaging options for their specific operating parameters.

While steady-state power conversion applications have benefited from packaging and cooling improvements, these solutions have been shown to have the potential for detrimental effects in transient applications as well as overdesigned cooling systems [49]. Since the realistic operation will be transient profiles for RF and power devices applications, we must explore additional designs to be made or included that would not be allowable under steady-state operation. Unlike GaN and SiC, because of low thermal conductivity of β-Ga_2O_3, β-Ga_2O_3 has longer time constant than that of GaN and SiC that

optimal thermal design for transient operation shall be different from the design for steady-state application. Since the major applications for ultra-wide bandgap β-Ga_2O_3 power electronics are broad range of power switching and RF components, we should know the best cooling solution for the transient operation. For this study, the primary research question is:

- **Are optimized solutions made for DC operation still relevant for transient operational profiles expected for RF and power switching applications?**

  For this study, thermal management techniques for lateral devices will be applied to investigate the cooling effect for both steady-state and transient operation. The effectiveness of bottom-sided cooling, top-sided cooling, and double-sided cooling will be investigated for various timescale, and the effect of TBC between materials will be studied as well. Also, the effectiveness of transient cooling will be compared with high thermal conductivity material transistor, which has a shorter time constant.

The above questions and issues will be addressed throughout this dissertation, with an overall outline of each chapter and its contribution to the overall understanding of the document outlined below in Section 1.4.

1.4 Dissertation Outline

The present work aims to address the mentioned problems in the following chapters. Chapter 2 discusses the precedent works to improve thermal reliability of β-Ga_2O_3. Chapter

3 features temperature measurements methodology for ultra-wide bandgap devices. Theory of IR and Raman spectroscopy, implication of nano-particle assisted Raman and transient Raman thermometry are presented. Chapter 4 explains the importance of thermally-aware design of the layout of lateral MOSFETs. The effect of anisotropic behavior of β-Ga_2O_3, and the effect of metal spacing are explored. Then optimal thermal design is suggested for multi-finger device with studied design parameters. Chapter 5 demonstrates the optimal design process for the vertical device structure (CAVET). Several cooling schemes are investigated by bottom-sided, top-sided, and double-sided cooling, and optimal thermal design is suggested. Chapter 6 focuses on transient behavior of various possible steady-state cooling solutions for β-Ga_2O_3. It should be noted that optimized solutions for steady-state would not be applicable to transient state solution. Chapter 7 concludes the work and provides a summary and future recommendations.

CHAPTER 2. IMPROVING THERMAL RELIABILITY OF GALLIUM OXIDE

β-Ga_2O_3 is considered a potential candidate for next-generation power devices owing to its ultra-wide bandgap of 4.8 eV and breakdown electric field of 8 MV/cm^2. Moreover, high quality, large-size, and low-cost β-Ga_2O_3 substrates can be obtained by melt-growth techniques. The primary obstacle to β-Ga_2O_3 power devices is a low thermal conductivity (9–26 W/m-K), resulting in high thermal resistive devices, which causes the heat dissipation problem in high-power operations. In this chapter, several methods to cope with the thermal reliability of β-Ga_2O_3 are introduced.

2.1 Integration to the Composite Wafers

To understand the thermal limitations in the devices, the Near Junction Thermal Transport (NJTT) program through the Defense Advanced Research Projects Agency (DARPA) was able to achieve a 2.7× reduction in the thermal resistivity of a GaN high electron mobility transistor (HEMT) device compared to a GaN-on-SiC device and demonstrated a 3× increase in the areal dissipation density of GaN-on-diamond comparatively with GaN-on-SiC [33]. It was found that the relatively large thermal boundary resistance (TBR) between GaN-diamond interface was 47.6 m^2-K/GW, which was significantly hindering the device performance. Several works were done to minimize TBR of GaN-diamond interface[34]–[36], and it has been reduced to near the calculated limit of 3.1 m^2-K/GW [37].

Similar to GaN, integrating β-Ga_2O_3 thin films onto substrates with high thermal conductivity materials such as diamond and SiC is considered a potential method to mitigate thermal issues and, therefore, increase the performance and reliability of high-

power β-Ga_2O_3 electronic devices. Such an approach takes advantage of both the high thermal conductivity of the substrates and the reduction of thermal resistance (inversely proportional to the film thickness) resulting from the thin-film geometry. Thin β-Ga_2O_3 film can be mechanically exfoliated and transferred to polycrystalline diamond substrate integrated by Van der Waals forces [50], and deposit a 30 nm thick β-Ga_2O_3 layer on single crystal diamond grown via atomic layer deposition [51]. However, TBR of β-Ga_2O_3-diamond interface for exfoliated β-Ga_2O_3 is 58.8 m^2-K/GW , which is much higher than TBR of first generation of GaN-diamond interface, while that for ALD grown β-Ga_2O_3 is 7.3 m^2-K/GW . However, mechanical exfoliation results in the cleavage along the Ga_2O_3 (001) plane prevents the fabrication of a large β-Ga_2O_3 layer on the diamond, and a 30 nm thick β-Ga_2O_3 layer is unsuitable for the fabrication of electronic devices.

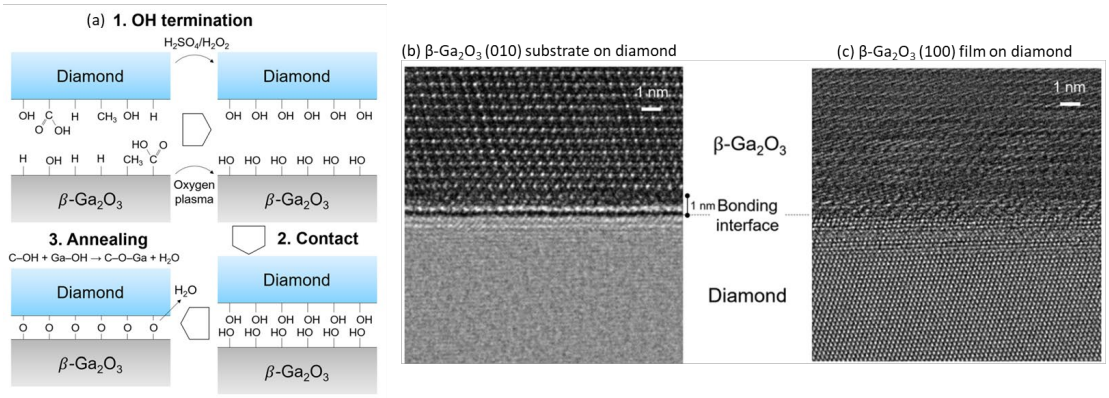

Figure 7. (a) β-Ga_2O_3 and diamond substrates were bonded by the hydrophilic bonding method. The OH-terminated surfaces formed direct bonding by a thermal dehydration reaction at 250 °C. (b-c) TEM images of the β-Ga_2O_3 /diamond bonding interface. They were bonded without nano-voids, cracks, or a significant loss in crystallinity [52].

To increase the quality of the bonded interface of β-Ga_2O_3 and diamond, the hydrophilic bonding of an oxygen-plasma-activated β-Ga_2O_3 substrate with an OH-terminated diamond substrate was reported as shown in Figure 7 [52]. The β-Ga_2O_3 surface was irradiated by oxygen plasma using our reactive ion etching at 250 °C. Meanwhile, after

ultrasonic cleaning in de-ionized water, the diamond substrates were OH terminated, then β-Ga_2O_3 and diamond surfaces were contacted with each other under atmospheric conditions without applying pressure. With this method, β-Ga_2O_3 (010) substrate was successfully bonded to diamond (111) substrate, and exfoliated β-Ga_2O_3 (100) film was also successfully bonded to diamond (111) substrate. Even though Newton's ring is observed in areas where the surfaces are not in contact with each other, this method shows ~70% of the diamond substrate directly bonded with the β-Ga_2O_3 substrate.

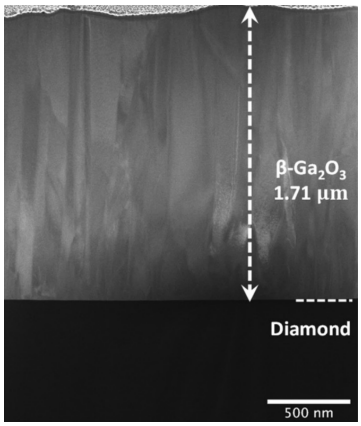

Figure 8. Low-magnification STEM image showing the cross-sectional morphology[53].

Moreover, the weak nature of the van der Waals bonding between the exfoliated β-Ga_2O_3 and the diamond substrate can limit the heat transfer rate across the heterointerface, which can be improved by depositing β-Ga_2O_3 directly on diamond. β-Ga_2O_3 can be grown using low pressure chemical vapor deposition on (100) oriented, single-crystalline diamond substrates [53]. The dominant growth direction of β-Ga_2O_3 films was along the <-201> direction. Figure 8 shows a low-magnification cross-sectional STEM image obtained from the sample, which shows a sharp contrast between the grown film and the substrate. The thickness of the films was measured to be 1.7 μm. No voids or exfoliation were observed indicating a very high-quality interface which is essential for high thermal boundary conductance.

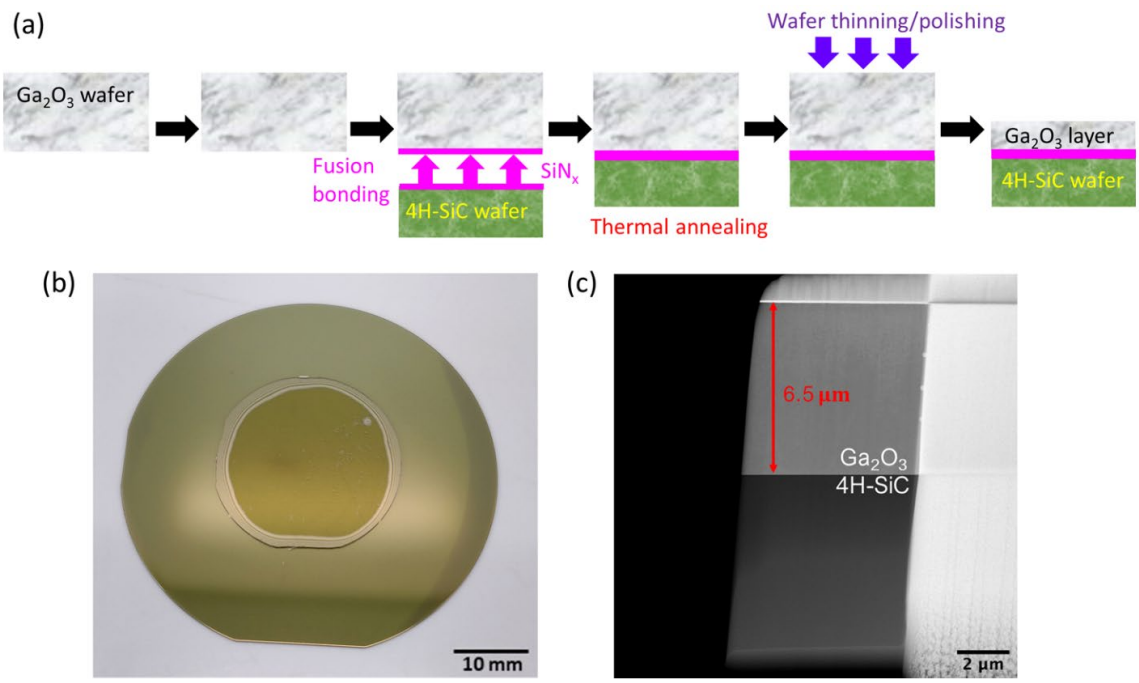

Figure 9. (a) Wafer-bonding and -thinning approach used to create the Ga_2O_3 composite substrate. (b) An image of Ga_2O_3 bonded onto 4H-SiC. The yield is nominally 100% except in the edge exclusion region. (c) Cross-sectional transmission electron microscopy image of the Ga_2O_3-on-SiC composite wafer [54].

Similar to β-Ga_2O_3/diamond composite wafer, a novel β-Ga_2O_3/4H-SiC composite wafer with high heat transfer performance and an epi-ready surface finish has been developed using a fusion-bonding method [54]. By taking advantage of low-temperature metalorganic vapor phase epitaxy, a Ga_2O_3 epitaxial layer was successfully bonded on the composite wafer while maintaining the structural integrity of the composite wafer without causing interface damage. Ga_2O_3 was thinned down using a series of lapping plates and a diamond abrasive, followed by a silica-based chemical-mechanical polishing process to remove subsurface damage and enable subsequent epitaxial growth for device processing. To minimize the overall thermal resistance of the composite substrate, and the final thickness of the Ga_2O_3 layer was determined to be ~6.5 µm. This novel power transistor topology resulted in a ~4.3× reduction in the junction-to-package device thermal resistance.

2.2 Dielectric on β-Ga_2O_3

For transistor applications, a gate dielectric should present low leakage currents, have low interface trap densities to achieve a controllable threshold voltage, and should also have a higher breakdown field. Many insulators such as Al_2O_3, Si_3N_4, SiO_2, and HfO_2 have been studied as a gate oxide material and passivation layers for gallium oxide devices [55]–[58]. Extreme permittivity materials such as $BaTiO_3$ were also studied and used as dielectric material in β-Ga_2O_3 transistors and heterojunction Schottky barrier diodes [59], [60]. Among all the dielectric materials, Al_2O_3 is studied and used most extensively for β-Ga_2O_3-based devices due to its compatibility with β-Ga_2O_3. Various lateral MOSFETs with Al_2O_3 as gate dielectric were demonstrated showing excellent electric field strength [61], [62]. Vertical device structures with outstanding figure of merit (FOM) using Al_2O_3 as gate dielectric were demonstrated as well [63], [64].

Table 2. Dielectric constant and thermal conductivity of dielectric materials

Material	Dielectric constant	Thermal conductivity @ Room Temperature (W/m-K)
Al_2O_3	8.5 – 9	25
SiO_2	3.5 – 4.5	1.4
Si_3N_4	6.2	27
HfO_2	25	1.1
AlN	8.9	321
Diamond	5.7	2000 (Single Crystalline)

However, in terms of thermal reliability, top-sided cooling is important as β-Ga_2O_3 is highly thermally resistive that one needs to consider not only the dielectric strength, but also the thermal conductivity of dielectric layer. As shown in Table 2, dielectric materials which are compatible with β-Ga_2O_3, has similar or even lower thermal conductivity than β-Ga_2O_3. For example, the allowable thickness of the SiO_2 gate dielectric in Ga_2O_3-based

devices is severely restricted due to the low thermal conductivity (13 W/m- K in the [100] crystallographic direction) of β-Ga_2O_3. Thick SiO_2 can enable the high breakdown voltage of the device, but it makes the poor thermal conductivity of β-Ga_2O_3 even worse.

Analogous to those gate barrier materials mentioned above, AlN can be utilized as a gate dielectric for the β-Ga_2O_3 based devices due to its large bandgap (6.2 eV) and high gate dielectric constant. It was presented by several reports that the nitridation is an effective way to decrease the interface state density and improve interface electrical characteristics [65], [66]. As compared with the SiO_2 gate barrier, AlN/β-Ga_2O_3 heterostructures might be a promising solution to solve the problem of β-Ga_2O_3 poor thermal conductivity. Moreover, the AlN/ β-Ga_2O_3 based devices can possess high breakdown voltage while maintaining the low total thermal resistivity of β-Ga_2O_3 to the substrate. In addition, the knowledge of energy band alignment in the AlN/ β-Ga_2O_3 system is of particular importance for the design and optimization of advanced β-Ga_2O_3 based devices [67].

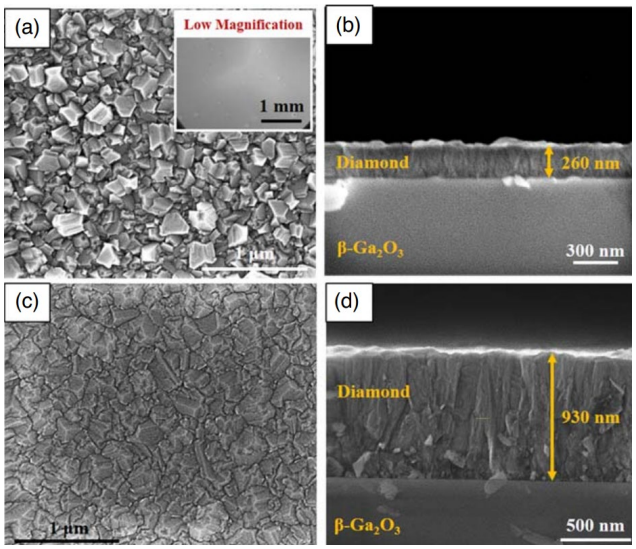

Figure 10. SEM micrographs of polycrystalline diamond films grown on SiO_2-coated β-Ga_2O_3. (a) and (b) Diamond were grown on 19 nm SiO_2/β-Ga_2O_3 with a thickness of ~260 nm. (c) and (d) Diamond were grown on a 100 nm SiO_2/β-Ga_2O_3 with an average grain size of ~400 nm and a thickness of 930 nm [68].

Similar to AlN, very high thermal conductivity material, diamond, can be also used not only as a gate dielectric, but also as a heat spreader. Using single crystalline diamond for β-Ga_2O_3 device cooling is an expensive approach due to the limited availability of large area single crystalline diamond wafers (>1 cm^2) and difficulty in single crystalline diamond growth on β-Ga_2O_3. Therefore, scalable growth of polycrystalline diamond on β-Ga_2O_3 wafers via chemical vapor deposition for thermal management purposes was reported [68]. With SiO_2 interlayer, this work was successful to grow the thickness of the diamond layers of ~260 nm when grown on a 19 nm SiO_2 interlayer and of 930 nm when grown on a 100 nm SiO_2 interlayer as shown in Figure 10. Unlike single crystalline diamond, thermal conductivity of polycrystalline diamond will be dependent on thickness and grain of diamond film. The directionally averaged thermal conductivity of a 267 ± 21 nm thick diamond film grown on an Sn-doped (-201)-oriented β-Ga_2O_3 substrate coated with 19 nm SiO_2 interlayer was 110 ± 33 W/m-K. This value is comparable to polycrystalline diamond thin films with similar thickness and grain size grown on GaN as shown in Figure 11 [69]. The effective TBR at the diamond/β-Ga_2O_3 interface, which includes the thermal resistance arising from the 19 nm SiO_2 interlayer, was determined to be ~30 m^2-K/GW.

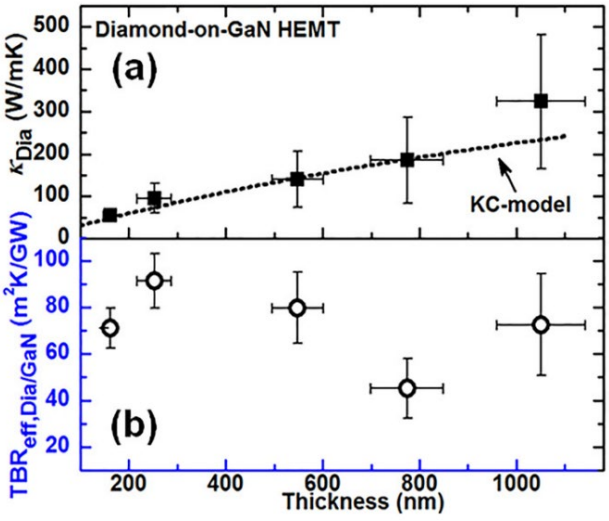

Figure 11. (a) Thermal conductivity of polycrystalline diamond (PCD) and (b) TBR between PCD and GaN as a function of the PCD film thickness.

2.3 Preceding Simulation Studies

In addition to the introduced experimental efforts, several simulation studies were reported to improve the thermal reliability of β-Ga_2O_3 devices. First of all, the effects of thermal boundary conductance (TBC) between β-Ga_2O_3 and substrate interface and thermal conductivity of substrate were briefly studied in the work that integrates exfoliated β-Ga_2O_3 on diamond [50]. An analytical solution for the temperature rise calculation in multilayer structures with discrete heat sources were used [70]. A 500-nm (100) Ga_2O_3 layer atop a substrate has anisotropic thermal conductivity with k_z = 12 W/m-K and k_r = 21 W/m-K. The modeled device structure and conditions were shown in Figure 12 (a): 10 fingers with 50 μm gate-to-gate spacing, 4 × 150 μm for each heat source, 2000 × 2000 μm total domain, and 10 W/mm total power density. The maximum device temperature with different TBC and different substrates (Si, SiC, and diamond) were calculated and the results are shown in Figure 12 (b). Compared with Si substrates, the maximum temperature

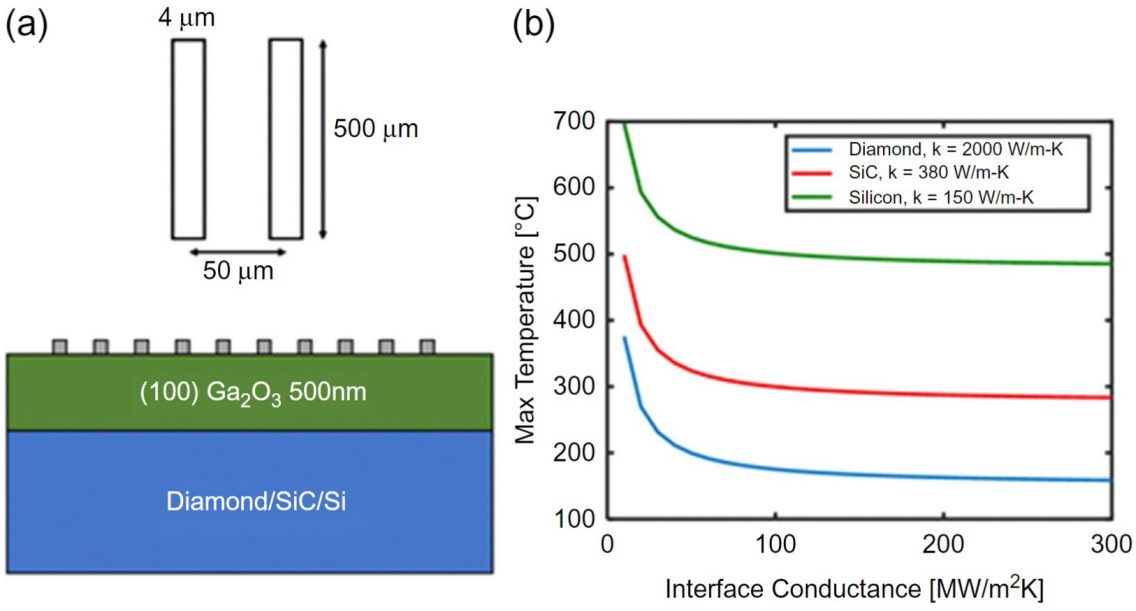

Figure 12. Effect of β-Ga_2O_3–substrate TBC on device thermal management. (A) Schematic diagram of device structure. (B) Effect of β-Ga2O3–substrate TBC on the max temperature of a device [39], [50]

of devices with high thermal conductivity substrates such as SiC and diamond are much lower. For a certain substrate, TBC plays an important role in limiting the maximum device temperature when TBC value is lower than 100 MW/m^2-K for the modeled device structures in this work. For single crystal diamond substrate with a relatively low TBC of the transferred β-Ga_2O_3/diamond interfaces, the cooling performance is excellent. For TBC higher than 200 MW/m^2-K, the maximum device temperatures are weakly dependent on TBC values for the modeled device structures in this work.

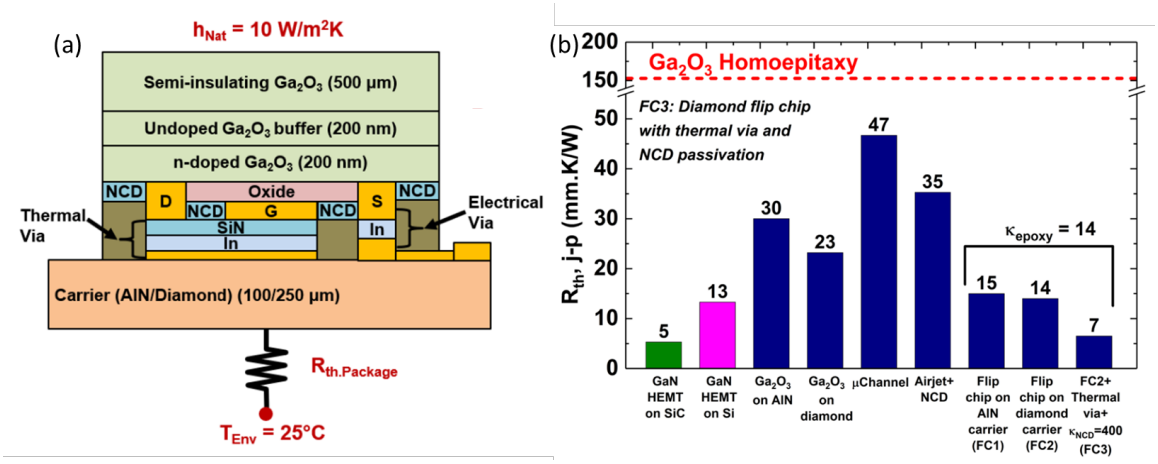

Figure 13. (a) Flip-chip hetero-integration of a Ga_2O_3 MOSFET combined with the use of NCD passivation and thermal bumps. (b) Comparison of the device thermal resistance associated with the thermal management schemes studied in this work as well as the benchmark GaN-on-Si HEMT technology [71].

The effectiveness of bottom-sided cooling methods (substrate engineering and microchannel cooling) and top-sided cooling methods (air-jet impingement cooling and flip-chip hetero-integration) has been demonstrated as shown in Figure 13 (b) [71]. The device model shows that a homoepitaxial device suffers from an unacceptable junction 10 W/mm, indicating the importance of employing device-level temperature rise of ~1500 °C under a power density of thermal managements to individual Ga_2O_3 transistors. The effectiveness of various active and passive cooling solutions was tested to achieve a goal of reducing the device operating temperature below 200 °C at a power density of 10 W/mm.

Results show that flip-chip heterointegration as shown in Figure 13 (a) is a viable option to enhance both the steady-state and transient thermal characteristics of Ga_2O_3 devices without sacrificing the intrinsic advantage of high-quality native substrates.

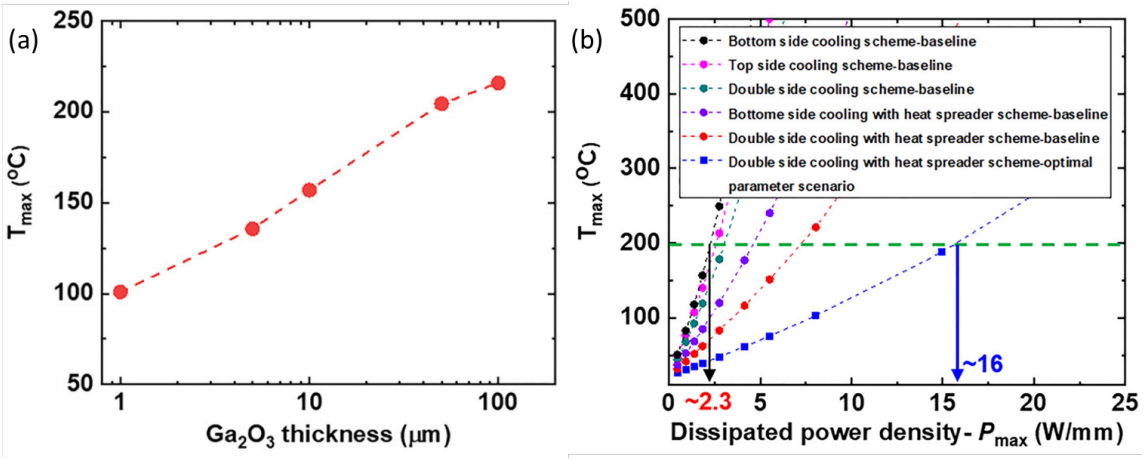

Figure 14. (a) The dependence of channel maximum temperature on Ga_2O_3 thickness, when the thermal boundary conductance and substrate thermal conductivity are kept with their baseline values, 20 MW/m^2-K and 400 W/m K, respectively. (b) The channel maximum temperature as a function of power density for the various cooling scheme [72].

Another study also comprehensively investigated the effects of the various cooling approaches on the device channel temperature along with guidance for material selection to enable the most effective thermal solutions [72]. Figure 14 (a) shows maximum temperature as a function of Ga_2O_3 thickness (from 1 μm to 100 μm), when the thermal boundary conductance and substrate thermal conductivity are kept with their baseline values, i.e., 20 MW/m^2-K and 400 W/m-K, respectively. The figure shows that thinning the Ga_2O_3 substrate from 100 μm to 1 μm results in a ~ 130 °C (~50%) reduction in maximum temperature. Among various cooling strategies, similar to previous work, double-sided cooling combined with a heat spreader used in the active region of the device can suppress the device thermal resistance to as low as 11 mm-°C/W, achieving a maximum dissipated power density as high as 16 W/mm for a junction temperature limit of 200 °C.

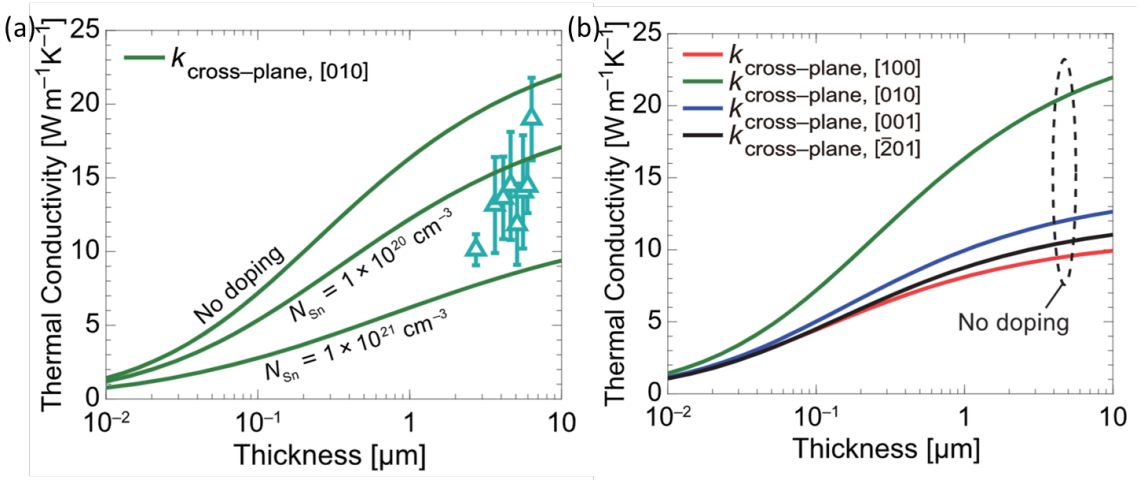

Figure 15. (a) Cross-plane thermal conductivity of (010)-oriented β-Ga_2O_3 thin films as a function of film thickness at room temperature. The predictions of the model with varying Sn doping concentration, N_{Sn}. (b) Cross-plane thermal conductivity of β-Ga_2O_3 thin films versus film thickness at room temperature [73].

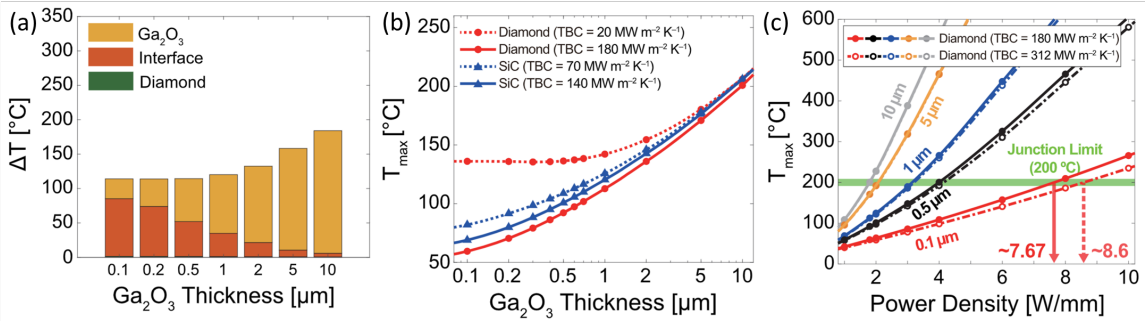

Figure 16. (a) Maximum device temperature rise, with the reference of a bottom boundary temperature of 22 °C, as a function of Ga_2O_3 layer thickness. (b) Maximum device temperature (T_{max}) as a function of Ga_2O_3 layer thickness for single-gate Ga_2O_3 devices on diamond and SiC at a power density of 1.8 W/mm. (c) T_{max} as a function of dissipated power density for single-gate Ga_2O_3 devices, where the thickness of the Ga_2O_3 layer ranges from 0.1 to 10 μm. With 200 °C assumed as a safe junction temperature limit [73].

Recent study addressed the fundamental conduction cooling limits for sub-1 μm β-Ga_2O_3 devices integrated with diamond via finite element simulations [73]. A semi-classical transport theory for phonons interacting with interfaces is employed to systematically calculate the thickness-dependent thermal conductivity of the β-Ga_2O_3 layers with different crystallographic orientations as shown in Figure 15. Figure 16 (a)

shows the thermal resistance of the Ga_2O_3 layer monotonically decreases with decreasing Ga_2O_3 layer thickness as the influence of TBC increases. This trend is strongly associated with the inherently low thermal conductivity of the Ga_2O_3 layer. However, if the TBC is too low as 20 MW/m^2-K, having a sub-1 μm β-Ga_2O_3 would not be affected as shown in Figure 16 (b). It was found that the maximum power density of sub-1 μm β-Ga_2O_3 devices on diamond, particularly that of the 0.1 μm device, can reach up to 7.7 W/mm with a junction temperature limit of 200 °C, considering an optimal device orientation as well as best-case experimental Ga_2O_3/diamond TBC. As the Ga_2O_3/diamond TBC approaches the limit predicted by the diffuse mismatch model, the fundamental limit to the maximum power density of these devices can reach up to 8.6 W/mm, which is comparable to those reported previously for costly augmented thermal management designs.

CHAPTER 3. EXPERIMENTAL METHODOLOGY OF TEMPERATURE MEASUREMENTS

3.1 Introduction

Accurate temperature measurement with high spatial resolution when applied to ultra-wide bandgap devices is a challenging research topic. Measurements down to the submicron regime are desired in order to capture both lateral and vertical thermal gradients (static and transient) in the transistor channels, where conventional methods are not able to make measurements. The development of a nanoscale thermometer is not only a matter of controlling the spatial resolution of the measurement technique, but also the development and improvement in the application of techniques and understanding their limitations. There are several applications of techniques used to measure the operational temperature of ultra-wide bandgap transistors: Thermoreflectance Imaging, Infrared (IR) thermometry, and Raman Thermometry.

Thermoreflectance Imaging measures the reflected visible wavelength illumination to provide the surface temperature distribution with submicron spatial resolution. Thermoreflectance is typically used in a pump probe setup known as Time Domain Thermoreflectance (TDTR) in which thermal properties such as thermal conductivity and thermal boundary resistance of various materials are extracted. Replacing the photodiode detector with a charge coupled device (CCD), thermoreflectance can be used as a thermal imaging instrument to monitor the change in thermoreflectance of every pixel in the CCD. To estimate the surface temperature rise via Transient Thermoreflectance Imaging (TTI), the correct thermoreflectance coefficient must be applied to the thermally induced optical reflectivity variation detected. Therefore, the accuracy of this methodology is based on

how well the thermoreflectance coefficient of the surface is estimated. In addition, the procedure requires additional complex equipment and long acquisition times.

Infrared thermography is based on the physical phenomenon of radiative emission from an object of finite temperature. Similar to TTI, to determine the emissivity of a sample using infrared thermography, a calibration procedure is required. The emissivity calibration involves heating the sample to a known temperature and measuring the total emitted radiation. The emissivity of the object can then be determined as the ratio of the measured radiance to the expected radiance of a blackbody at the known temperature. The major advantage of infrared thermography is the rapid measurement capabilities offering 2D temperature mapping of the entire field of view. Infrared thermography is proven to be much faster than other thermal characterization techniques. The entire measurement procedure can be completed in several minutes. However, the quality of the emissivity calibration and the lateral signal averaging due to the limited spatial resolution (~ 3 μm) limit this practice to be beneficial only in a qualitative manner, specifically for recent generations of power electronics with submicron length scales.

Raman spectroscopy provides non-contact and fast means to locally analyze micro-scale devices with sufficient high spatial resolutions of ~1 μm. Raman spectroscopy measures phonon frequency of semiconductor materials, which makes possible to measure both stress and temperature. Therefore, Raman spectroscopy can serve as an effective tool for examining operating devices since interference to the electrical performance of the device by photon irradiation can be minimized by utilizing a sub-band gap visible laser. However, in other words, the laser is transparent that will average the temperature of the ultra-wide bandgap semiconductor channel that will be limited to extract the exact peak channel temperature.

This chapter will introduce the principle of IR thermometry and Raman thermometry as they are utilized in further section in Chapter 4 and Chapter 6.

3.2 Infrared Thermography

Infrared (IR) thermography is based on the physical phenomenon of radiative emission from an object of finite temperature. A blackbody emits radiation as a function of wavelength (λ) and temperature (T) according to Planck's distribution:

$$\boldsymbol{E_b(\lambda, T) = \frac{2hc^2}{\lambda^5\left[exp\left(hc/\lambda k_B T\right) - 1\right]}} \quad (1)$$

where $E_b(\lambda, T)$ is the spectral radiance, h is Planck's constant, c is the speed of light, and k_B is the Boltzmann constant. Integrating Eq. 1 over the entire electromagnetic spectrum ($0 \leq \lambda \leq \infty$), the total radiance of a blackbody at a given temperature is determined and can be expressed through the Stefan-Boltzmann law:

$$\boldsymbol{E_b(T) = \sigma T^4} \quad (2)$$

where σ is the Stefan-Boltzmann constant. As the temperature of the object increases, the magnitude of radiance increases and the wavelength of maximum radiance decreases. So far, discussion has been limited to the blackbody, while the blackbody is a theoretical ideality and real objects are observed to emit less radiation than that of the blackbody. This observation leads to the introduction of the spectral emissivity coefficient (ε_λ) defined as a ratio of the actual spectral radiance of an object to that of the spectral radiance of a blackbody at the same temperature and wavelength:

$$\boldsymbol{\varepsilon_\lambda = E(\lambda, T)/E_b(\lambda, T)} \quad (3)$$

Integrating Eq. (3) over the entire electromagnetic spectrum and applying Eq. (2), the relationship between the temperature, total emissivity (ε), and total radiance (E) can be established:

$$\boldsymbol{E(T) = \varepsilon \sigma T^4} \quad (4)$$

If the emissivity of the object is known and the total emitted radiation is measured, then the temperature of the object can be determined. This relationship demonstrates the importance of accurate determination of the emissivity for the material of interest to obtain thermal measurements with the highest possible accuracy. It should be noted that infrared cameras are designed to detect radiation over a specific spectral range; for example, medium-wavelength infrared cameras detect radiation over a range of 3-5 μm.

To determine the emissivity of a sample using infrared thermography, a calibration procedure is required. At its most basic, the emissivity calibration involves heating the sample to a known temperature and measuring the total emitted radiation. The emissivity of the object can then be determined as the ratio of the measured radiance to the expected radiance of a blackbody at the known temperature. The main limitation of this one-temperature calibration procedure is its vulnerability to background radiation. To minimize the effects from the background radiation, a two-temperature emissivity calibration can be applied. However, background radiation can still be challenging for areas of low emissivity since the relative levels of radiative emission are similar. Moreover, the two-temperature calibration can result in misalignment of the acquired radiance images due to thermal expansion. This misalignment is disadvantageous as it can result in signal convolution around defining features of objects such as device edges [74]. These effects become

particularly significant and must be taken into consideration when characterizing devices with micron length scales. For example, GaN has a bandgap energy of 3.4 eV rendering it transparent to less energetic infrared radiation. This is detrimental to measurement accuracy because the measured radiation originates not only from the surface but also from underlying layers which may be significantly cooler. This can result in significant underestimation of the surface temperature.

The major advantage of infrared thermography is the rapid measurement capabilities offering 2-dimensional (2D) temperature mapping of the entire field of view. This can be extremely useful for the identification of hotspots which establishes significantly reduced regions of interest for further thermal probing. Infrared thermography is proven to be much faster than other thermal characterization techniques. The entire measurement procedure can be completed in several minutes. However, the quality of the emissivity calibration and the lateral signal averaging due to the limited spatial resolution (~3 μm) limit this technique to be useful only in a qualitative manner, especially for current generations of microelectronics with (sub)micron length scales. In this work, infrared thermal measurements were conducted on β-Ga_2O_3 MOSFETs using an InfraScope system (Quantum Focus Instruments). The system is equipped with a liquid nitrogen-cooled 512 × 512 pixel InSb focal plane array camera for MWIR detection. The full-field thermal map of the β-Ga_2O_3 MOSFETs provides an excellent qualitative sampling of the regions of temperature maxima as shown in Figure 23 in Chapter 4. After locating the hotspots on the device, complementary thermometric techniques that offer much greater spatial resolution and measurement accuracy can be utilized.

3.3 Raman Spectroscopy

3.3.1 Theory

Raman spectroscopy is a spectroscopic technique based on the inelastic scattering of monochromatic light with a material, usually from a laser source. In this method, photons are scattered by the sample elastically (Rayleigh scattering) or inelastically (Raman scattering) which results in a frequency shift in photons. The Raman scattering is based on molecular deformations in an electric field determined by the material's molecular polarizability. The photons can be considered as an oscillating electromagnetic wave with electrical vector. Upon interaction with the sample, it induces electric dipole moment, which interacts with the molecular vibrations.

Monochromatic laser light with frequency υ_0 excites molecules and transforms them into oscillating dipoles. Such oscillating dipoles emit light of three different frequencies (Figure 17) when:

1. A molecule with no Raman-active modes absorbs a photon with the frequency υ_0. The excited molecule returns back to the same basic vibrational state and emits light with the same frequency υ_0 as an excitation source. This type if interaction is called an elastic Rayleigh scattering.

2. A photon with frequency υ_0 is absorbed by Raman-active molecule which at the time of interaction is in the basic vibrational state. Part of the photon's energy is transferred to the Raman-active mode with frequency υ_m and the resulting frequency of scattered light is reduced to $\upsilon_0 - \upsilon_m$. This Raman frequency is called Stokes frequency, or Stokes scattering.

3. A photon with frequency υ_0 is absorbed by a Raman-active molecule, which, at the time of interaction, is already in the excited vibrational state. Excessive energy of

excited Raman active mode is released, molecule returns to the basic vibrational state and the resulting frequency of scattered light goes up to $\upsilon_0 + \upsilon_m$ This Raman frequency is called Anti Stokes frequency, or just Anti-Stokes scattering.

Since the energy loss or gained by the photons is directly related to the energy loss or gained by the phonon vibrational modes, this technique allows for the direct measurement of the vibrational energies of the zone centered phonons in the material.

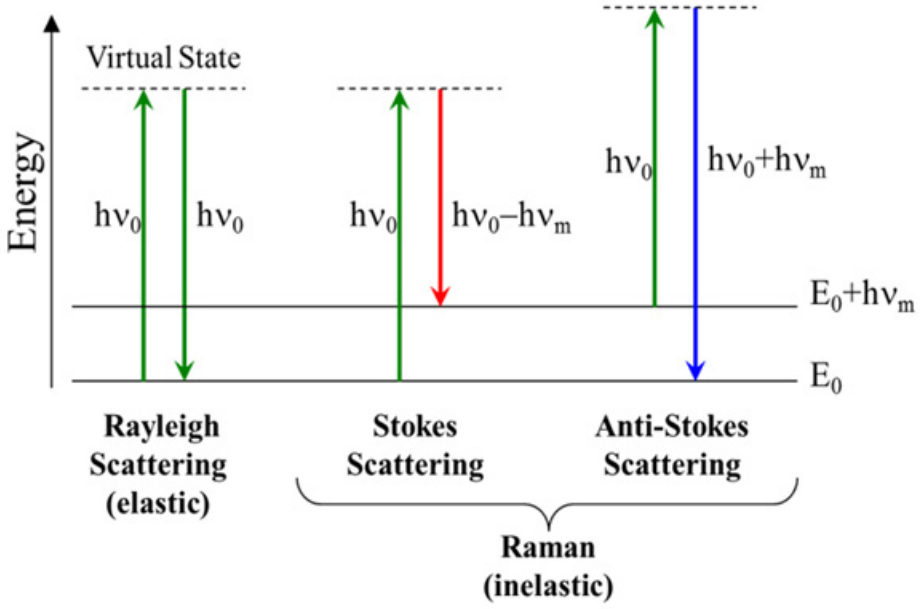

Figure 17. Energy level diagram for Rayleigh and Raman scattering processes

About 99.999% of all incident photons in spontaneous Raman undergo elastic Rayleigh scattering. Only about 0.001% of the incident light produces inelastic Raman signal with frequencies $\upsilon_0 \pm \upsilon_m$. Therefore, lasers are used as a light source that is capable of irradiation on a sample with very high photon density. Spontaneous Raman scattering is very weak and special measures must be taken to distinguish it from the predominant Rayleigh scattering. Instruments such as notch filters, tunable filters, laser stop apertures,

double and triple spectrometric systems are used to reduce Rayleigh scattering and obtain high-quality Raman spectra.

3.3.2 Raman Thermometry

Raman thermometry is a thermal characterization technique which makes use of Raman scattering phenomena to determine the local temperature in microelectronics systems. Any characteristics of the phonons, which vary with temperature, can be used to measure the thermal state of the system. For example, the change in Raman frequencies represents the change in temperature/stress states, and the change in line width of the peak, or full-width-half-maximum (FWHM), represents the change in temperature or quality of the crystal. Following Figure 18 shows the typical Raman spectrum of GaN, when it is heated or under tensile strain, the peaks show red-shifts, otherwise blue-shifts. For 180° back scattering geometry, GaN has E_2(High) and A_1(LO) Raman active modes.

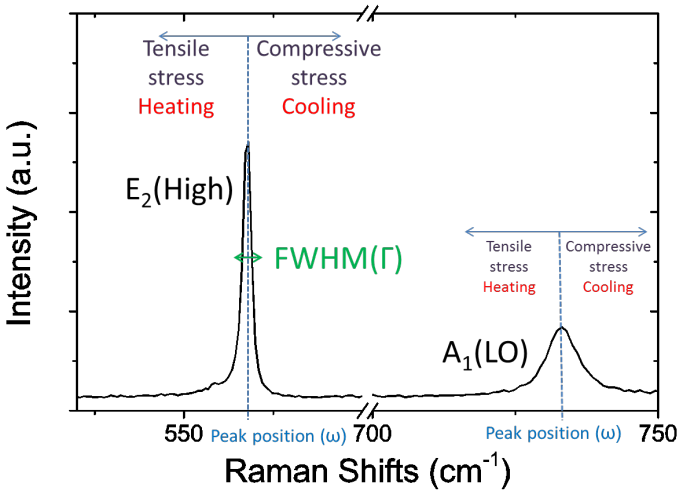

Figure 18. A schematic of a typical Raman spectrum of GaN showing with peak positions (ω) of two phonon modes (E_2(High) mode of strain-free GaN : ~568 cm^{-1} and A_1(LO) mode of it: ~ 732 cm^{-1}), and the line-width, or full width at half maximum (FWHM, Γ).

As the lattice is heated or cooled, the equilibrium positions of the atoms are displaced, resulting in a volumetric expansion or contraction of the lattice and a change in interatomic forces as a result of the anharmonicity of the bonds. These changes in the interatomic forces modify the phonon vibrational frequencies that results in the change of the Raman peak position. For example, the temperature dependent E_2(High) phonon frequency shift of Stokes Raman of AlGaN/GaN HEMTs is utilized [75]. However, temperature is not the only one that affects the peak position. As the volumetric changes, which contribute to peak shifts, result from the change of distances between the atoms, the peak position is sensitive to the lattice strain as well. During the operation of AlGaN/GaN HEMTs, thermo-elastic stress that comes from self-heating effect and inverse piezoelectric stress that is related with the magnitude of the vertical component of the electric field in GaN layer are developed. Therefore, the shifts in phonon frequency include both temperature and stress effects. Utilizing peak shift method to obtain the operating temperature of AlGaN/GaN HEMTs underestimates it that it is necessary to consider not only the thermal effects but the other factors that could affect the Raman spectra.

The linewidth of a Raman peak results from the lifetime of the phonon. The lifetime of the phonons can be determined from the Heisenberg uncertainty principle which states that the energy of the phonon can be measured only for a finite amount of time. The dominant contribution comes from the thermal expansion that the increase in interaction among optical phonons at high temperature causing increased phonon scattering and decreased phonon lifetime. Therefore rise in temperature accompanies phonon peak broadening since lifetime is decreased. Scattering of the phonon is dependent on a variety feature such as defects, material boundaries, and other phonons. Phonon-phonon scattering

is a dominant factor in broadening of line-width since the phonon population increases with elevation of lattice temperature, whereas other factors such as defects and grains are temperature independent. This increment in population reduces the phonon lifetime, and thus increasing the linewidth allowing temperature to be measured.

The final spectral feature used for thermal analysis is the ratio of the anti-Stokes intensity to the Stokes intensity for a given phonon mode. The temperature dependence of the anti-Stokes/Stokes intensity ratio is a result of the temperature dependence of the phonon population. Qualitatively, as temperature increases, the phonon population increases and there are more excited vibrational states. Therefore, it is more likely for a photon incident upon the material to interact with one of these phonons and absorb its energy. The sample then emits a photon with greater energy than the incident one, or anti-Stokes Raman scattering. Because of this, the anti-Stokes to Stokes intensity ratio is seen to increase with temperature.

However, applying this Raman thermometry to β-Ga_2O_3-based devices, there is a major limitation for thermal analysis. For GaN HEMTs, the device structure typically consists of a relatively thin (~1 μm) GaN channel/buffer layer where the active 2DEG channel is. Therefore, Raman thermometry probes the GaN layer averaging through the thickness of GaN, which is on the order of a micron, when using visible wavelengths below the bandgap of GaN. However, β-Ga_2O_3 devices are fabricated on thin films homoepitaxially grown on thick (~650 μm) substrates that through thickness averaging become more considerable. In addition, β-Ga_2O_3 has much lower thermal conductivity than GaN that there will be much greater temperature gradients through the thickness of β-Ga_2O_3. Averaging over these larger vertical temperature gradients will lead to greater

underestimation of the peak temperature rise in the channel, which is located near the device surface. In addition, if using Raman peak position-based methods for thermal analysis, the thermoelastic stresses induced by the large through-thickness temperature gradients need to be considered.

3.3.3 Nanoparticle-assisted Raman Thermometry

The use of Raman active nanoparticles has preliminarily shown to be a fine candidate to solve this challenge [76], [77]. If Raman spectroscopy can identify the various materials that are present in the probing volume of a given sample of interest, then it is reasonable to assume that a particle deposited onto the sample surface will introduce additional characteristic peaks in the Raman spectra. Spectral analysis can then be performed to measure the temperature rise of the particle on the sample surface. With the independent Raman sensors on top of the device, a strain free surface temperature can be measured.

There are several aspects of nanoparticle-assisted Raman thermometry that can be utilized for thermal analysis of emerging microsystems. While Raman spectroscopy cannot be used to directly probe metals, nanoparticle-assisted Raman thermometry can be used to indirectly probe the temperatures of metals [78]. This is because nanoparticles will be deposited on the surface of the device to be used as temperature transducers. As follows, this adaptation of Raman thermometry can also be used to measure the surface temperature of semiconductors whose bandgaps are greater than the Raman laser energy, where depth averaging would occur. This is crucial for UWBG semiconductor devices whose heat generation occurs within tens of nanometers of the device surface. Moreover, the nanoparticles are assumed to be in thermal equilibrium with the device surface, not alter the intrinsic temperature distribution, and experience negligible thermal stress.

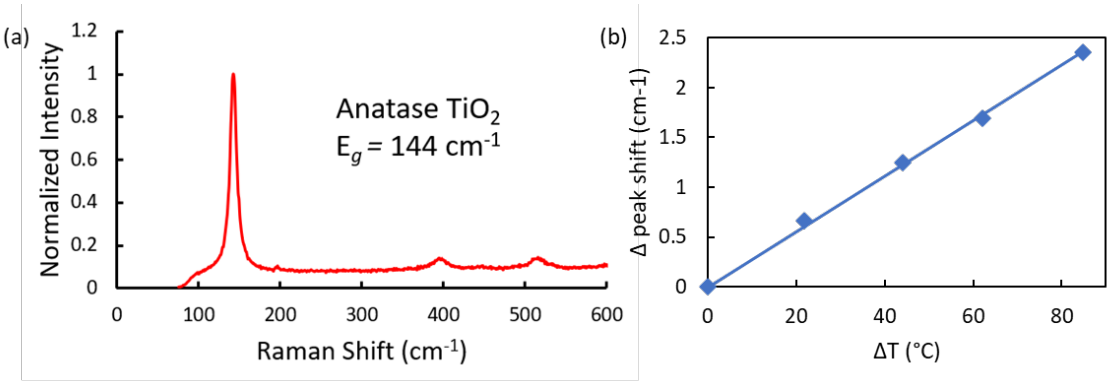

Figure 19. (a) Raman Spectrum of TiO_2 at Room Temperature with E_g = 144 cm^{-1}, (b) Raman temperature calibration of TiO_2 (99.98% purity).

The TiO_2 nanoparticles are applied by dropping a solution of isopropanol and the nanoparticles onto the device surface while it is heated on a temperature-controlled baseplate. The isopropanol is then allowed to evaporate by maintaining the temperature of the baseplate above the boiling point of isopropanol (~85 °C), leaving TiO_2 on the device surface. After applying TiO_2 to the device surface, the nanoparticles are allowed to reach thermal equilibrium with the device surface during operation and the Stokes Raman peak shift of the E_g phonon mode is used to evaluate the temperature response of the nanoparticles [79].

3.3.4 Transient Raman Thermometry

Several thermal characterization techniques are commonly used to quantify self-heating in electronics including infrared thermography, Raman thermometry, and thermoreflectance thermal imaging. While infrared thermography is the most common method employed, it has been shown to underestimate peak temperature rise in the transparent semiconductor channel. The temporal resolution of transient infrared thermography is limited to microsecond levels; therefore, it is not capable of capturing the thermal dynamics important in fast switching applications with high-power dissipation. Thermoreflectance thermal imaging is well-suited for steady-state and transient

microelectronics temperature assessment due to the abundance of metallization structures and two-dimensional mapping capabilities. However, due to low signal-to-noise ratios, it commonly employs an iterative lock-in measurement scheme, which forces synchronization of pulsed device operation and optical probing [80]. Improper use of this technique resulting from failure to fully understand the thermal dynamics of the system can result in reporting of quasi-steady state temperature rises in the device channel which can be significantly lower than the true steady-state value. In contrast, Raman thermometry is very effective as a point measurement technique to determine the temperature rise in the semiconductor channel under both steady-state and transient measurement conditions [81], [82].

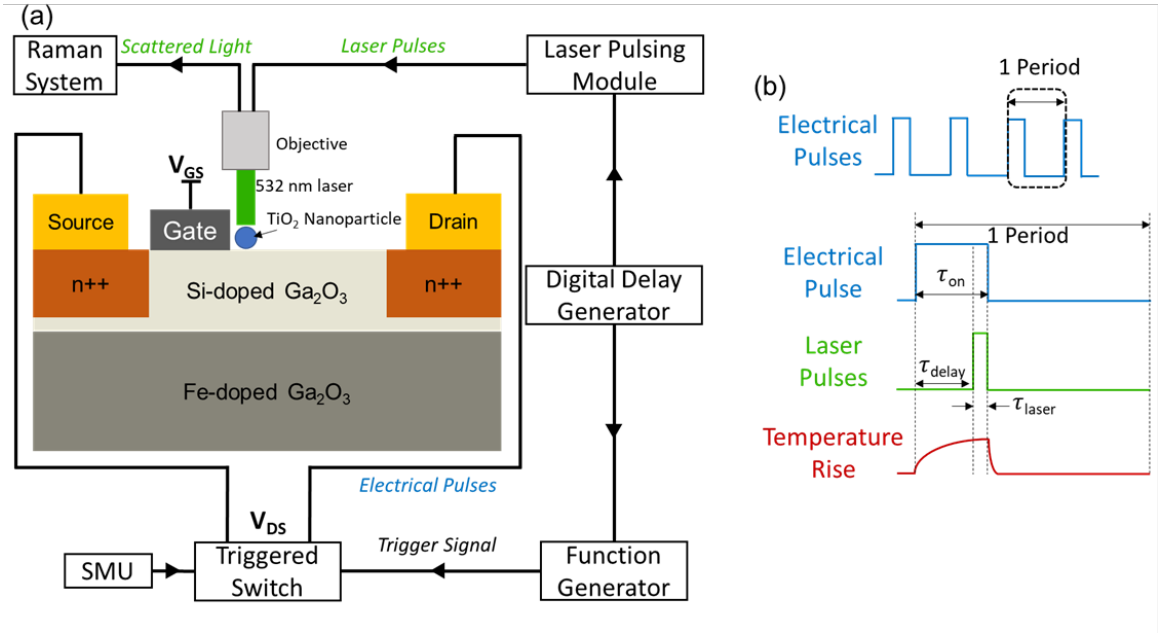

Figure 20. (a) Experimental setup used for transient Raman thermometry. (b) The synchronized pulsing scheme used to capture the transient thermal response of the β-Ga_2O_3 MOSFET.

The experimental setup used for transient Raman thermometry experiments is illustrated in Figure 20 (a). This setup adopts a lock-in modulation scheme, in which the electrical and laser pulse trains are synchronized while the Raman signal accumulates over

many periods. Using this experimental setup, a temporal resolution of 15 ns was achieved and used in this study. Figure 20 (b) shows the synchronized pulsing scheme that allows control of the electrical pulse width (τ_{on}) of the applied drain-source voltage (V_{DS}) and the laser pulse width (τ_{laser}) that produces a Raman signal, which is collected by the detector of the Raman system. The time delay (τ_{delay}) between the electrical and laser pulses is controlled by a digital delay generator which measures the full transient temperature rise of the device in response to a square electrical pulse. To initiate optical pulsing, the digital delay generator sends a signal to a fixed frequency driver which drives an acousto-optic modulator (AOM) for the Raman laser. To initiate electrical pulsing, the digital delay generator sends a signal to a function generator which modulates a switch in the biasing circuit of the device under test.

CHAPTER 4. THERMALLY-AWARE LAYOUT DESIGN OF LATERAL DEVICES: METAL-OXIDE SEMICONDUCTOR FIELD EFFECT TRANSISTOR (MOSFET)

Content in the chapter (figures and text) adapted from:

S. Kim, *et al.,* "Thermally-Aware Layout Design of β-Ga_2O_3 Lateral MOSFETs," in *IEEE Transactions on Electron Devices*, vol. 69, no. 3, pp. 1251-1257, March 2022 [38]

4.1 Overview & Approach

Recently, several experimental and computational studies have reported self-heating mitigation strategies for β-Ga_2O_3 transistors to reduce its' high thermal resistance. The thermal conductivity of β-Ga_2O_3 fall in the range of 9 to 26 W/m-K at room temperature, which is one or two order magnitude lower than that of SiC and GaN. The effectiveness of bottom-side cooling methods (substrate engineering and microchannel cooling) and top-side cooling methods (air-jet impingement cooling and flip-chip hetero-integration) have been demonstrated [71]. In addition, the transfer of thin β-Ga_2O_3 membranes onto a high thermal conductivity diamond substrate has been demonstrated [50], [83]. An improvement of both electrical and thermal performance by replacing the 200-μm-thick β-Ga_2O_3 substrate with a 50-μm-thick Cu substrate has been proposed [84]. However, these studies were based on single or two-finger lateral devices without any impact of highly anisotropic thermal conductivity of β-Ga_2O_3 nor optimization of the device layout.

In contrast, this study investigates how the device layout design of a homoepitaxial β-Ga_2O_3 MOSFET itself could be optimized to enhance the device thermal performance

without changing substrate. To be more specific, many device engineers are currently building their Ga_2O_3 devices based on device mask layouts they have been using for previous device development (e.g., GaN transistors), because guidelines for such thermally-aware device design is lacking in open literature. The orientation of β-Ga_2O_3 substrate, the thermal boundary conductance (TBC) between metal and oxide layers, the gate-to-gate spacing of multi-finger device, the width of the channel, and the thickness of lateral MOSFET β-Ga_2O_3 substrate should affect the device thermal performance and need to be optimized thermally. Device engineers have been overlooking these design parameters and thermal engineers have not yet provided device developers with this knowledge. Moreover, no one has demonstrated the impact of these design parameters via experiments in a way that decouples the effect of each variable, which is not easy nor trivial. Therefore, this study will provide the optimized thermal design of single-finger device, and that of multi-finger device.

To understand the effect of the in-plane anisotropy of the thermal conductivity of (010)-oriented β-Ga_2O_3 substrates, single-finger MOSFETs with various channel orientations (rotated by 0/30/60/90°) were fabricated and characterized. Electro-thermal device simulation was performed to quantify the orientation-dependence of the device self-heating behavior, isolated from effects from the metallization structure arrangement. After this, to exclusively study the cooling effect arising from the metallization layout, electrically-identical (identical heat source profile under a given bias condition) but thermally-different (different gate-to-drain contact spacing) devices were fabricated and tested via nanoparticle Raman thermometry and infrared (IR) thermography. Lastly, the effect of number of channels, the spacing of the channel, and the width of the channel will be investigated and optimized via computational study.

4.2 MOSFET Device

4.2.1 Principle of MOSFET

The MOSFET (Metal-oxide Semiconductor Field Effect Transistor) transistor is a semiconductor device that is widely used for switching purposes and for the amplification of electronic signals in electronic devices. A MOSFET is either a core or integrated circuit where it is designed and fabricated in a single chip because the device is available in very small sizes. FETs can be majority-charge-carrier devices, in which the current is carried predominantly by majority carriers, or minority-charge-carrier devices, in which the current is mainly due to a flow of minority carriers. The device consists of an active channel through which charge carriers, electrons or holes, flow from the source to the drain. Source and drain terminal conductors are connected to the semiconductor through ohmic contacts. The conductivity of the channel is a function of the potential applied across the gate and source terminals. The FET's three terminals are: Source (S), through which the carriers enter the channel. Drain (D), through which the carriers leave the channel. Conventionally, current entering the channel at D is designated by I_{DS}. Drain-to-source voltage is V_{DS}. And Gate (G), the terminal that modulates the channel conductivity. By applying voltage to Gate, one can control I_{DS} [85].

In an n-channel enhancement-mode device, a conductive channel does not exist naturally within the transistor, and a positive gate-to-source voltage is necessary to create one. The positive voltage attracts free-floating electrons within the body towards the gate, forming a conductive channel. But first, enough electrons must be attracted near the gate to counter the dopant ions added to the body of the FET; this forms a region with no mobile

carriers called a depletion region, and the voltage at which this occurs is referred to as the threshold voltage of the FET. Further gate-to-source voltage increase will attract even more electrons towards the gate which are able to active channel from source to drain.

4.2.2 Device Preparation

Figure 21 (a) shows the device cross-sectional schematic. A 65 nm thick Si-doped β-Ga_2O_3 channel layer was grown on a 680 µm thick Fe-doped (010)-oriented β-Ga_2O_3 substrate using metal organic vapor phase epitaxy (MOVPE). The device fabrication process started with depositing 200 nm of SiO_2 by plasma-enhanced chemical vapor deposition (PECVD) which acted as the gate dielectric as well as an implant cap.

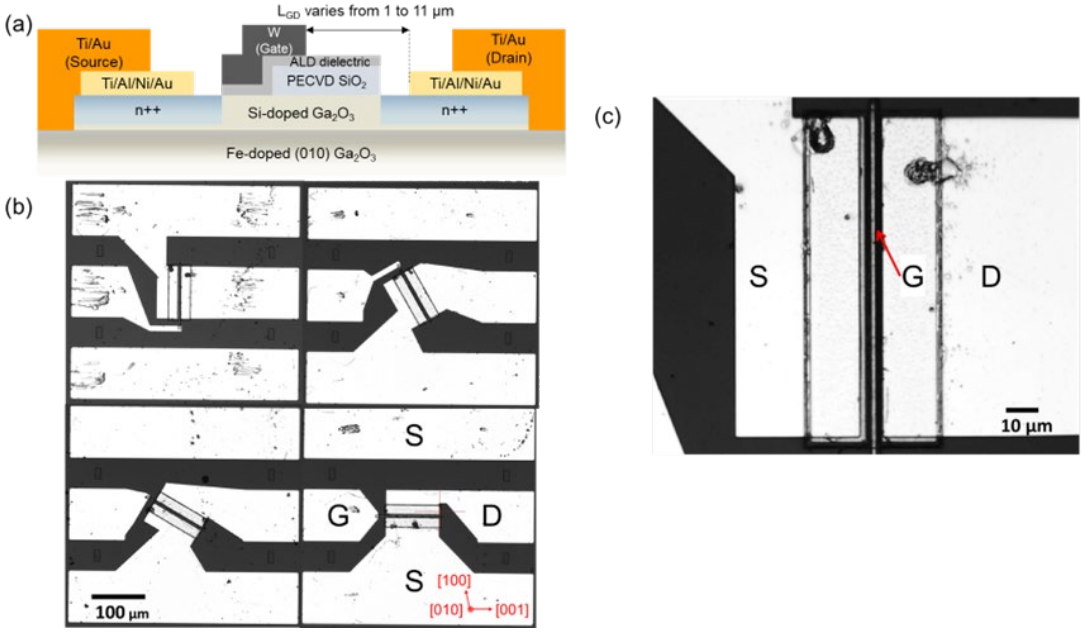

Figure 21. (a) Schematic cross-section of the β-Ga_2O_3 MOSFET. (b) CCD image of four different rotational MOSFETs with different in-plane orientation. The top-left device is denoted as a 0-degree device. (Similarly, top-right device: 30-degree device, bottom-left: 60-degree device, bottom-right: 90-degree device.) (c) Enlarged CCD image of 0-degree device showing the gate, source, and drain electrodes.

A tungsten (W) refractory metal layer was sputtered and patterned with a chromium (Cr) hard mask to define a 2.5 μm W/Cr gate electrode using a SF_6 reactive ion etch (RIE) chemistry. A refractory metal gate is crucial to the self-aligned process because an Au-based gate metal stack would not survive at the required implant activation temperature of greater than 900 °C. Si-implant regions were then patterned with the source-side of the W/Cr gate (L_G = 0.5 μm, W_G = 100 μm) exposed to eliminate the gate-source region (L_{GS} = 0 μm), while the gate–drain distance (e.g., L_{GD} = 2 μm) remained. A shallow Si-implant profile was designed with 10 and 35 keV energies with a total dose of 1×10^{15} ions cm^{-2} to achieve a target doping concentration of 1×10^{20} cm^{-3}. The Si-implant was activated at 900°C for 120 s using rapid thermal annealing (RTA) in a N_2 ambient. Ohmic contacts over the implanted regions were achieved with a Ti/Al/Ni/Au evaporated metal stack followed by a 470 °C RTA process for 1 min in a N_2 ambient, after removing the implant cap via RIE. Electrical isolation was achieved using inductively coupled plasma/reactive ion etching. The Ti/Au gate and interconnect metals were added for device characterization. From transmission line measurements of the implanted material, the average sheet resistance across the sample was 1.9 kΩ/sq. From Hall measurements using a Van der Pauw structure consisting of the non-implanted epitaxial material, the sheet resistance of the channel was 11.3 kΩ/sq. The average contact resistance across the sample was 1.2 Ω-mm. More fabrication details, including the gate metal, implant conditions, and implant activation can be found in [86]. For the orientation dependence study, the gate and the interconnect metal were rotated by 30/60/90 degrees. To be more specific, the channel width of the baseline device (heretofore to be called as a 0-degree device; Figure 21 (c)) was oriented along a direction close to [100]. Three additional devices were fabricated with

the gate metal rotated counterclockwise by 30 degrees, as shown in Figure 21 (b). To study the cooling effect by metal contact arrangement, "electrically-identical but thermally-different" devices were fabricated. Specifically, the length of the Si-implant region was kept at 2.5 μm for all devices (i.e., identical effective channel length for electron transport), while the distance between gate and drain metal contacts was varied from 1 μm to 11 μm.

4.2.3 Electrical Characteristics

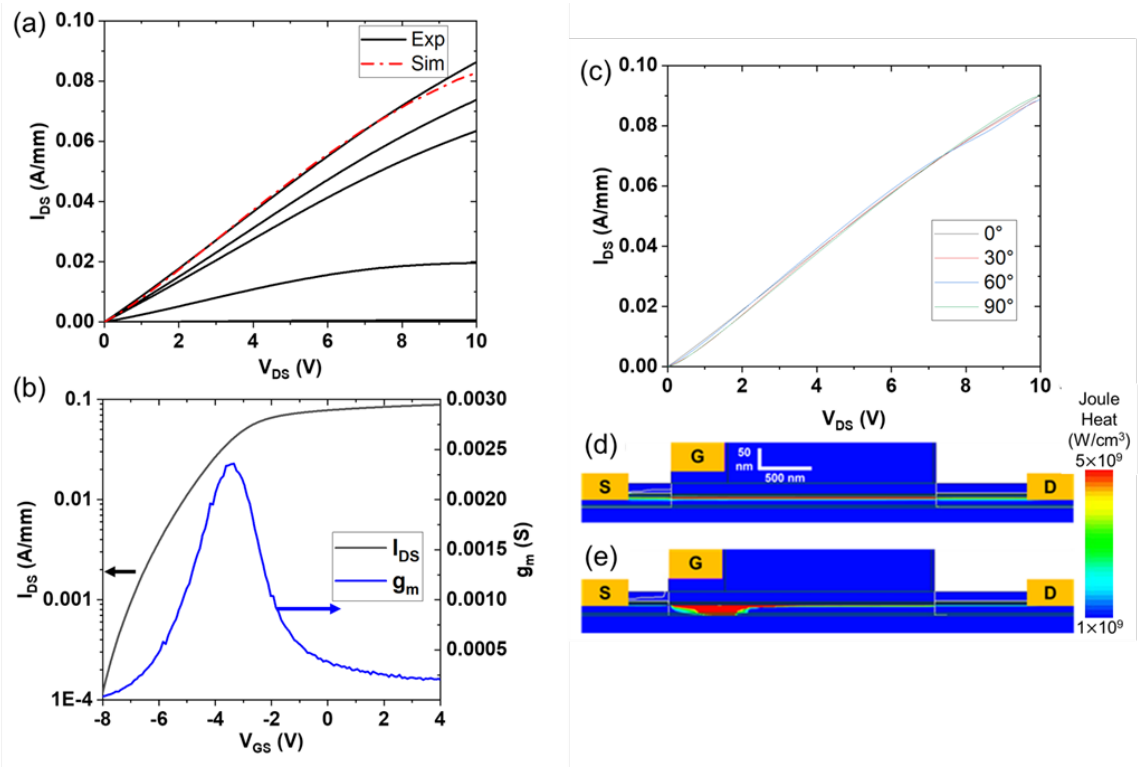

Figure 22. (a) I-V characteristics of the baseline (0-degree) β-Ga_2O_3 MOSFET for V_{GS} increasing from -8 to 4 V by 3 V steps. The simulated I-V curve for V_{GS} = 4 V is also shown. (b) The transfer characteristics of the baseline MOSFET at V_{DS} = 10 V. (c) I_{DS}-V_{DS} curves at V_{GS} = 4 V (fully-open channel conditions) of the four devices with different orientations (0°, 30°, 60°, and 90°). (d) Simulated Joule heating profile of the "fully open" channel condition (V_{GS} = 4 V) showing a relatively uniform heating profile throughout the entire channel. (e) Simulated Joule heating profile of the "partially open" channel condition (V_{GS} = -4 V) showing concentrated heating under the gate.

DC I-V curves were generated for all devices at room temperature. Figure 22 (a) shows the measured and simulated I-V characteristics of the baseline (0-degree) device. The channel temperature of the devices with different channel orientation were measured under fully-open channel conditions (gate-source voltage, V_{GS} = 4 V) as shown in Figure 22 (b) and (c). Figure 22 (c) shows the identical IV characteristics of the four devices with different orientation under fully open channel conditions. The orientation-dependent difference in the current and on-resistance among these devices is negligible. These bias conditions were used to minimize alteration of the heat generation profile arising from different voltage bias conditions required to operate the devices at an identical power level [87]. Figure 22 (d) and (e) show how bias conditions can affect the Joule heating profile for an identical P_{dis}. Under a fully-open channel condition (V_{GS} = 4 V, V_{DS} = 6.5 V), a relatively uniform heat generation profile forms between source to drain, whereas localized Joule heating occurs near the gate under partially-open (or pinched-off) channel conditions (V_{GS} = -4 V, V_{DS} = 14 V).

4.3 Effect of Anisotropic Thermal Behavior

Figure 23 (a) shows the temperature rise at the gate metallization of the four devices illustrated in Figure 21 (b), measured by nanoparticle-assisted Raman thermometry for two different power dissipation levels: 0.75 and 1 W/mm. Also shown are the modeling results of the gate surface temperature where the nanoparticles were located. IR thermography images are included as insets to qualitatively visualize the orientation-dependent self-heating effect. The 0-degree and 30-degree devices exhibited lower channel temperatures among the four devices. The IR images also confirm these results. It should be noted that all of the device models used to derive results in Figure 23 (a) employed the geometry of the metallization structures for the 0-degree device. For this reason, there are discrepancies

between the experimental and simulation results for the 30, 60, and 90-degree devices. In other words, the geometrical effect of the top metal structures on the device self-heating behavior was not isolated from the orientation-dependent effects.

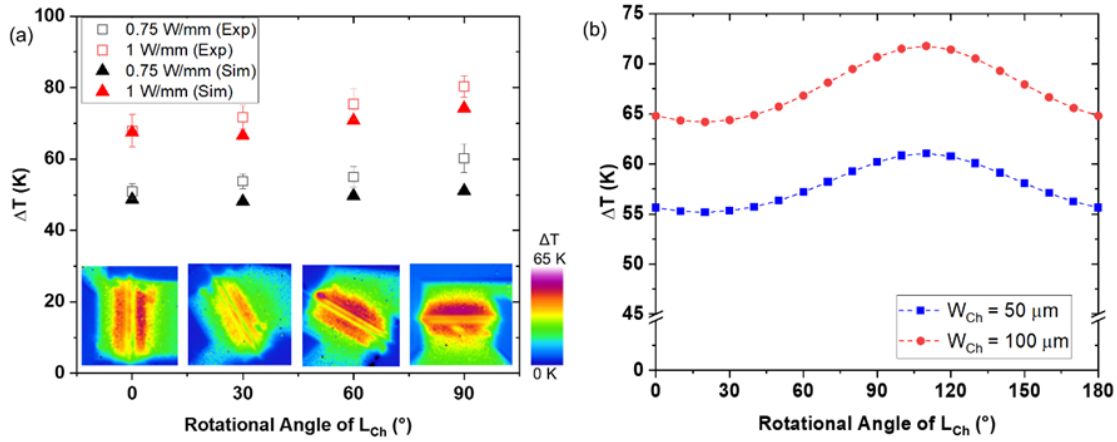

Figure 23. (a) The gate temperatures of MOSFETs with different orientation obtained by simulation and experiments. Results for two different power dissipation levels (P_{dis} = 0.75, 1 W/mm at V_{GS} = 4 V) are shown. Also, IR images of the four MOSFETs are displayed (P_{dis} = 0.75 W/mm). (b) Simulated MOSFET channel temperatures as a function of channel orientation for P_{dis} = 1 W/mm at V_{GS} = 4 V. This model does not include surface metallization structures to exclusively quantify the orientation-dependence of the device self-heating. Modeling was performed for two different channel widths (W_{ch}) of 50 µm and 100 µm.

Figure 23 (b) shows the simulated channel temperature rise as a function of different channel orientations. The simulation only accounts for the orientation-dependence because other variables such as metallization structures are excluded from the device model. The 30-degree device exhibits a lowest channel temperature rise among all the possible channel orientations. This is because, the thermal conductivity along the channel length is the highest (this direction is close to the direction perpendicular to $\perp(\bar{2}01)$ direction as shown in Figure 5 (b)) while the thermal conductivity along the channel width is the lowest (this direction is close to [100]). In other words, the channel width direction is less effective in terms of spreading the heat generated within the channel than the direction along the channel length. In contrast, a 110-degree device is subject to an opposite

condition, resulting in the highest temperature rise among all orientations. In this case, the thermal conductivity along the channel length, which is close to the direction, is the lowest. It was found that the channel orientation itself can result in a ~10% difference in the channel temperature rise for P_{dis} = 1 W/mm at V_{GS} = 4 V for MOSFETs fabricated on (010)-oriented β-Ga_2O_3 substrates. Also, devices with two different channel widths (50 and 100 μm) were studied via modeling to understand whether the self-heating of narrow channel devices or wide channel MOSFETs would be more influenced by the anisotropic thermal conductivity of β-Ga_2O_3. As plotted in Figure 23 (b), when the device width decreases from 100 μm to 50 μm (for a power dissipation level of 0.75 W/mm), the differences between minimum and maximum temperature decreases by ~10%, meaning a weaker anisotropic effect.

A previous report suggests a targeted power density of 10 W/mm for Ga_2O_3 MOSFETs, which is twice the operational power density of GaN power amplifiers [88]–[90]. However, this study [71] also states that the operating junction temperature should be kept below 200°C, which is based on studies on legacy GaN RF applications [88]–[90]. Using the calibrated electro-thermal device model and assuming a base temperature condition of 25°C, the 110-degree device (with the lowest in-plane thermal conductivity along the channel length direction) is able to operate up to a power density of ~2.1 W/mm at a channel temperature below 200°C. On the other hand, a 30-degree device (with the highest cross-plane thermal conductivity along the channel length direction) can operate up to ~2.4 W/mm. This equates to a ~14% increase in the power density by implementing the thermally-aware design. Augmenting the optimized device layout (30-degree device) with device- and package-level thermal management solutions will allow to achieve maximum power densities, while keeping the channel temperature below the safe operation limit.

4.4 Effect of Metal Dimension and Spacing

The discrepancy between the mean values of the experimental data and the modeling results in Figure 23 (a) indicates that the metal structure geometry may impact the device self-heating behavior. To quantify this effect, simulation was performed where the width (W_{hp}) of the interconnect between the drain electrode and the metal bond pad (where heat is extracted through the needle probes or wire bonds; in this work, we used needle probes to operate the devices) was varied from 10 μm to 100 μm, as shown in Figure 24 (a). While the area of the heat extraction region (bond pad) was kept invariant, the width of the heat pathway (interconnect), W_{hp}, was varied to mimic the different shape of the metallization structures for the devices with different channel orientations (Figure 21 (b)).

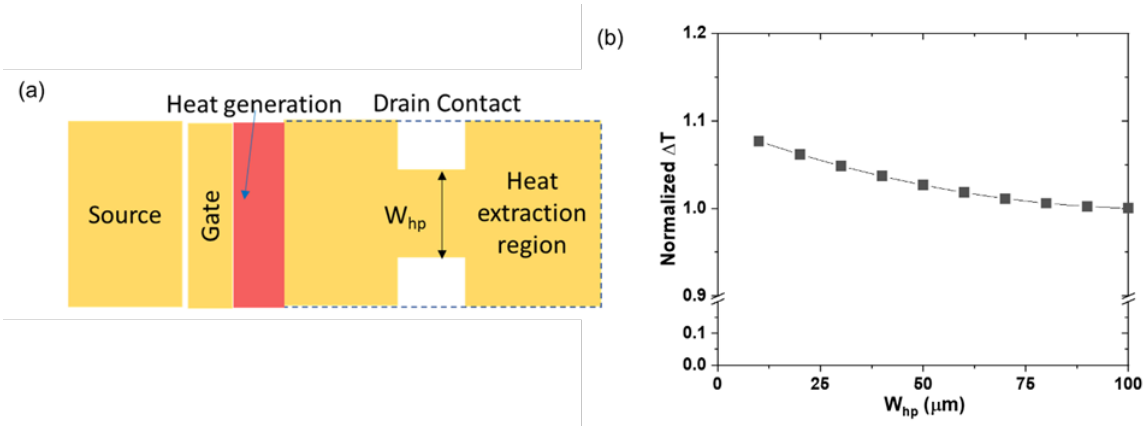

Figure 24. (a) Schematic of the drain metal contact: to consider the differences in the metal contact shape, the width of the heat path of the metal contact (i.e., interconnect), W_{hp} varies while the area of heat extraction region is fixed. (b) The temperature rise with reduced W_{hp} is normalized with respect to the 0-degree model (W_{hp} = 100 μm) results. It should be noted that W_{hp} is 10 μm for the 90-degree MOSFET.

As shown in Figure 24 (b), when W_{hp} decreases to 10% of the original width, the channel temperature rise increases by ~8%. Therefore, the geometry of the metallization structures near the device active region play an important role in dissipating heat away from the channel region. The relatively large difference between the thermal conductivities of the metal layers and the β-Ga_2O_3 is responsible for the observed geometrical effect of

the metallization structures on the device self-heating. In contrast, this effect is negligible for upright-configured devices based on SiC and GaN, because of the relatively high thermal conductivity of the semiconductor base materials.

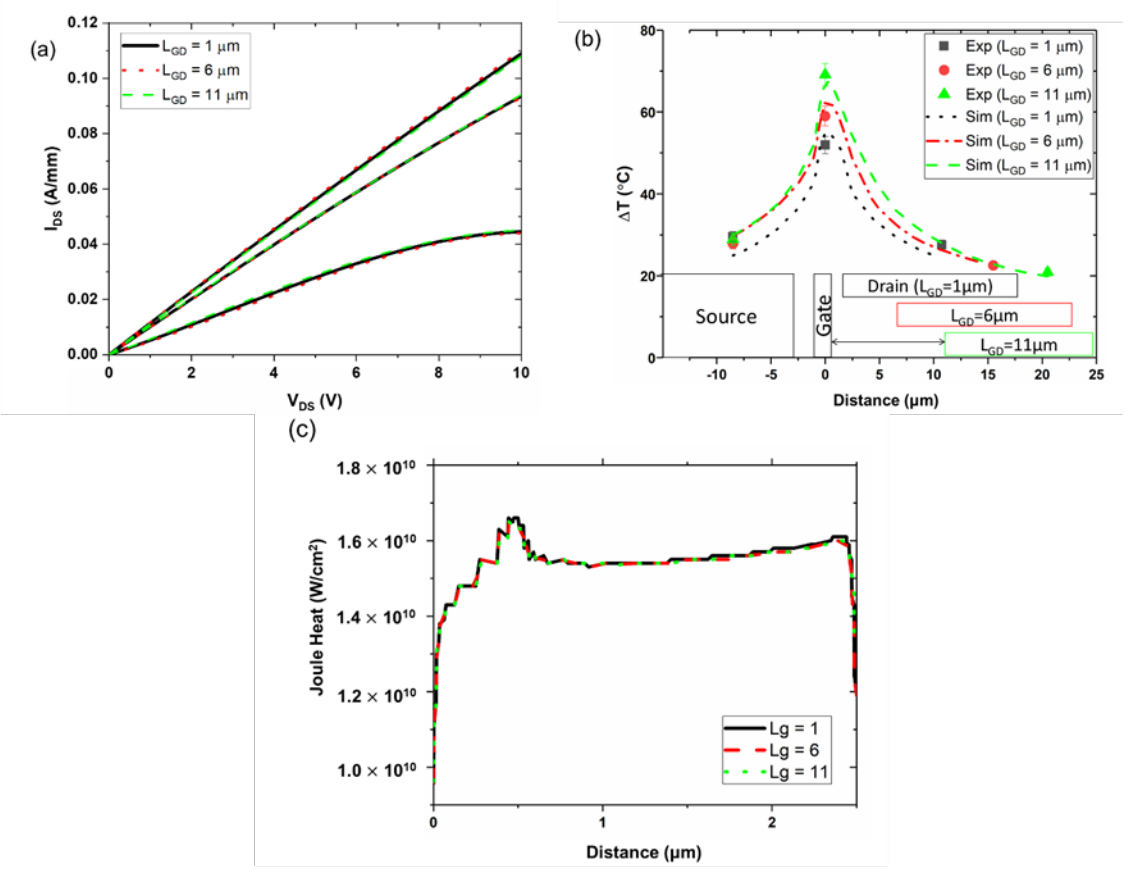

Figure 25. (a) Pulsed I-V curves for V_{GS} = -2, 1, 4 V, for the three "electrically-identical" devices with different L_{GD}. (b) Temperatures of the source – gate – drain electrodes along the centerline of the devices operating under P_{dis} = 0.8 W/mm at V_{GS} = 4 V. The length of the drain and the source electrodes are 16 μm and nanoparticles were measured near the center of the electrodes. (c) Heat flux profiles for the three gate-drain spacings at a power density of 0.75 W/mm. The two end points in the x-axis correspond to both ends of the channel region. In other words, the low resistance n^{++} regions outside the channel terminate at both ends.

Another important aspect related to the device layout that may impact the device thermal performance is the gate-to-drain distance, which is typically controlled to achieve a targeted device breakdown voltage. For this reason, "electrically-identical but thermally-

different" devices were fabricated and investigated. Figure 25 (a) shows the pulsed electrical output characteristics of the three devices with different gate-to-drain electrode distances (L_{GD}=1, 6, 11 μm) under three V_{GS} conditions, demonstrating the electrically-identical behavior, which is expected due to the low resistance of the n^{++} region. In other words, despite the metal electrode distances are different, the effective electron channel lengths are identical, which results in the identical electrical output characteristics. Accordingly, for identical bias conditions, the three devices will exhibit an identical heat generation profile. This is shown in Figure 25 (c) as line-plots of the integrated heat flux within the channel region. Therefore, the sole effect of the distance between the heat source (located near the drain side corner of the gate [87], [91]) and drain metal electrode on the device self-heating behavior can be evaluated. Utilizing nanoparticle- assisted Raman thermometry, the temperatures of the source, gate, and drain electrodes were measured under a fully-open channel condition (V_{GS} = 4 V, dissipated power = 0.8 W/mm), as plotted in Figure 25 (b). Since nanoparticle deposition (i.e., positioning individual particles) is not a fully controllable process, it was not possible to measure temperatures at the center of the device channels. Instead, temperatures at the drain side corner of the gate were measured using the nanoparticle-assisted Raman thermometry method. Although all three transistors were operating with identical heat generation profiles, the temperatures of the gate metal electrode show a large discrepancy. For the device with L_{GD} = 1 μm, the temperature rise of the gate electrode is the lowest since the drain electrode, which is acting as a heat sink, is closest to the heat source. Accordingly, a larger temperature rise at the drain electrode is observed, as compared to other devices with longer L_{GD}. When the drain metal electrode is further shifted by 10 μm away from the gate electrode (L_{GD} = 11 μm), a ~35% increase

in the gate temperature rise occurs. For the case of L_{GD} = 6 μm, the gate temperature rise increases by about 15%, as compared to the case of L_{GD} = 1 μm. The temperature rise of the drain electrode of the MOSFET with L_{GD} = 1 μm is ~10% higher than that for of the device with L_{GD} = 6 μm MOSFET and ~20% higher than that of the MOSFET with L_{GD} = 11 μm. There is a minor difference in the source electrode temperature among the three device structures. These results demonstrate the trade-off between increasing the device breakdown voltage and improving the device thermal performance by adjusting L_{GD} for homoepitaxial lateral transistors based on β-Ga_2O_3.

4.5 Optimal Design for Multi-finger Device

Based on previous studies, now this study expands to optimize multi-finger device structure as it will be the design for real application. As shown in Figure 26, various number of channels (from 2-finger to 20-finger), channel width (from 50 μm to 200 μm), gate-to-drain spacing (L_{GD}; from 1 μm to 20 μm), and the orientation of channel/substrate will be investigated and optimized. To reduce the computational loads, ¼ symmetry was applied so that thermal insulation boundary condition is applied to the symmetry surface.

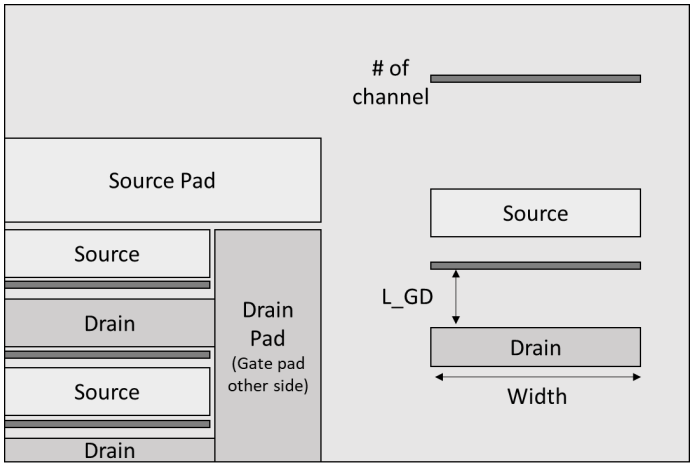

Figure 26. Quarter symmetry of top-view of the multi-finger device.

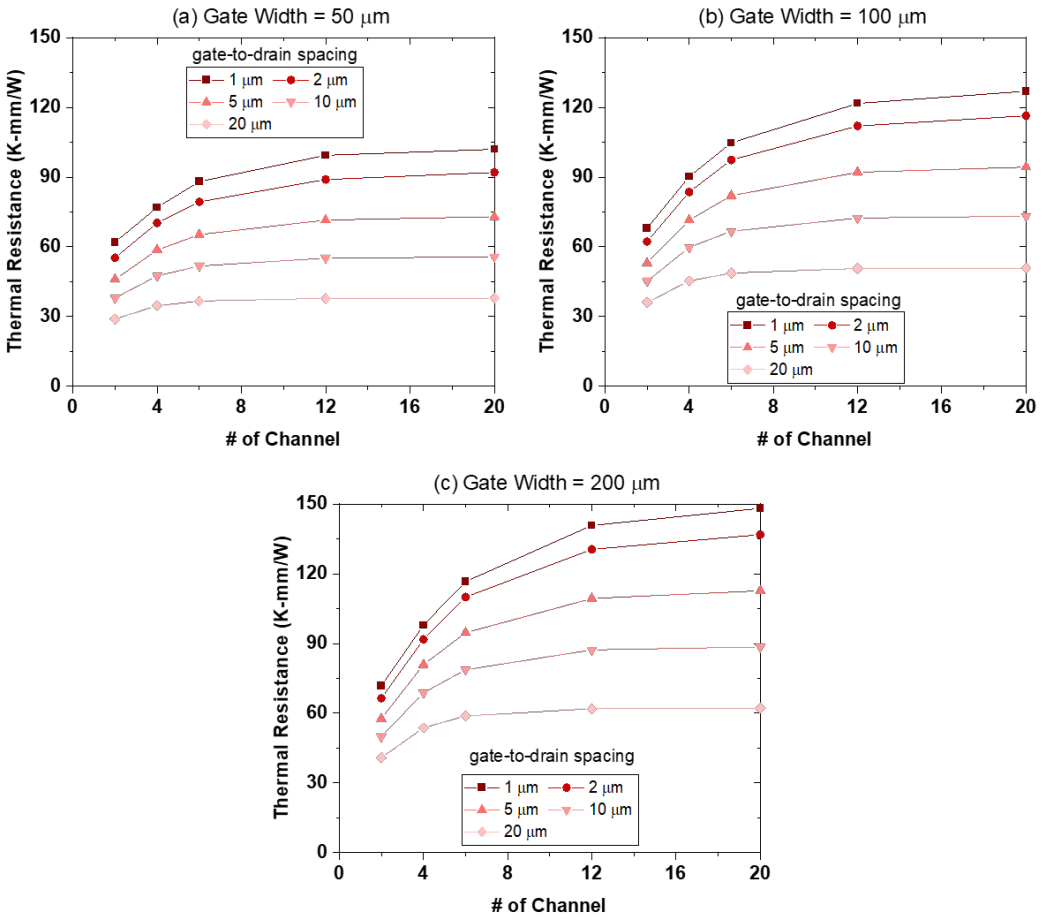

Figure 27. Total thermal resistance for various number of channels, gate-to-drain spacing (L_{GD}), and gate width.

Table 3. Selective thermal resistance (K-mm/W) from Figure 27.

L_{GD}	1 μm		20 μm	
Gate width	50 μm	200 μm	50 μm	200 μm
2-finger	62	72	29	41
20-finger	102	148	38	62

First of all, similar to Figure 23 (b), as gate width decreases, the temperature rise decreases for the same number of channels and same gate-to-drain spacing as shown in Figure 27. However, the decrease rates are different for number of channels and gate-to-drain spacing. For example, for two-finger device with 1 μm L_{GD}, the thermal resistance drops from 72 K-mm/W to 62 K-mm/W, 14% drop, from 200 μm gate width to 50 μm gate width. While for 20-finger device with 1 μm L_{GD}, the thermal resistance drops from 148 K-mm/W to 102 K-mm/W, 31% drop, from 200 μm gate width to 50 μm gate width. For 20 μm L_{GD} 2-finger devices, the thermal resistance drops from 41 K-mm/W to 29 K-mm/W, 29% drop, and for 20-finger device with same spacing, the thermal resistance drops from 62 K-mm/W to 38 K-mm/W, 39% drop. Therefore, the thermal resistance decreases more with longer spacing of L_{GD} for more channels, since as the channels are closer to each other, there would be more thermal cross-talk resulting in less decreasing rate of thermal resistance.

When L_{GD} = 20 μm, number of channels does not affect the channel temperature rise after 6-finger device indicating that there is no thermal cross-talk for all three gate widths structures. However, as the L_{GD} decreases to 1 μm, the channel temperature increases as the number of channels increases for all three different gate widths. Changing the number of channels from 2-finger to 20-finger device induces thermal resistance increase of 64.5% (from 62 to 102 K-mm/W), 86.8% (from 68 to 127 K-mm/W), 105.5% (from 72 to 148 K-mm/W) for 50 μm, 100 μm, 200 μm gate width, respectively. For the widest gate, the thermal resistance increases more than 2 × changing from 2-finger to 20-finger configuration that one should note the effect of thermal cross-talk when designing multi-finger device.

For fixed gate width, as number of channels increases the temperature, or thermal resistance of the device, clearly increases. Also, as gate-to-drain spacing, or L_{GD}, decreases, thermal cross-talk will increase so the thermal resistance increases. Comparing the effect

of L_{GD}, two-finger device shows less changes compared to 20-finger in terms of percentage. For 50 µm gate width, the thermal resistance drops from 62 K-mm/W to 29 K-mm/W, 53% drop for 2-finger device, and 102 K-mm/W to 38 K-mm/W, 63% drop for 20-finger device. After 6-finger device, the thermal resistance barely increases for 20 µm of gate-to-drain spacing, but the resistance of 1 µm gate-to-drain spacing keep increasing.

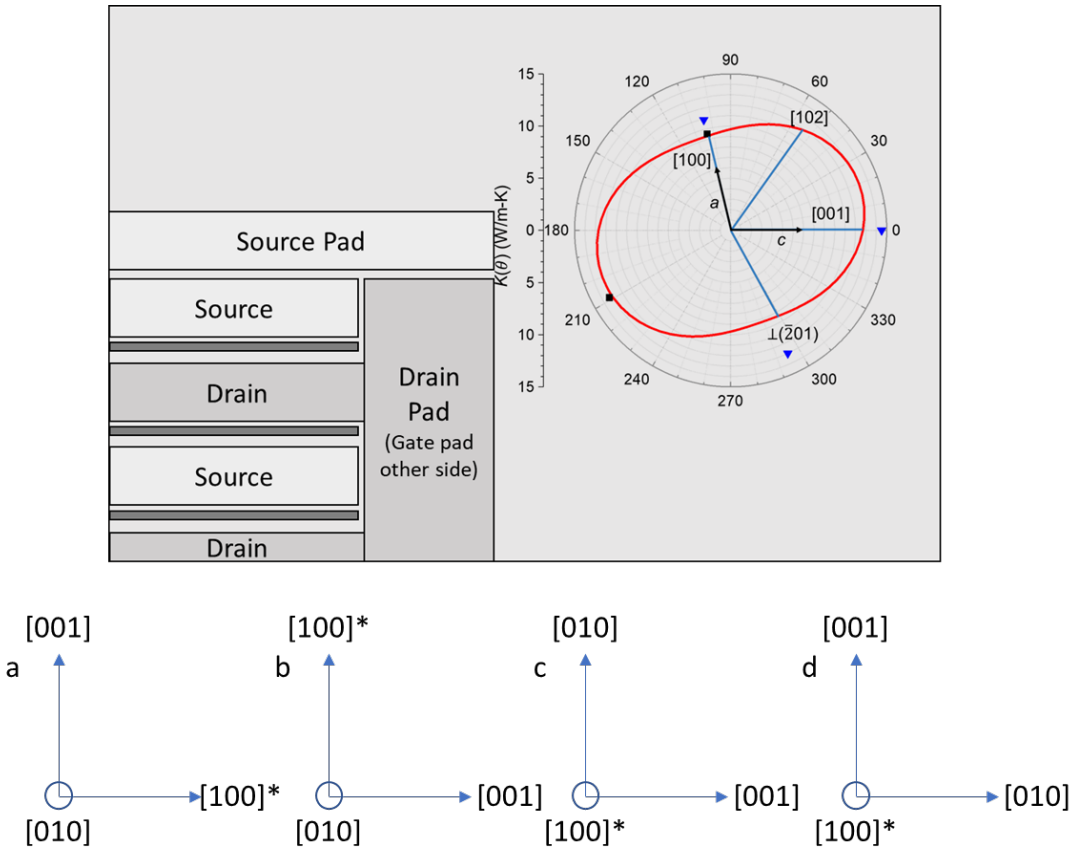

Figure 28. Schematic of multi-finger device structure with inset of in-plane thermal conductivity of (010)-oriented β-Ga_2O_3. Four different investigated orientation for multi-finger device. "a" orientation is based on 0-degree device from Chapter 4.3

So far, all the simulations were studied assuming (010)-orientation is the cross-plane, [100]* is oriented along with the channel width, and [001] is oriented along with the channel length. Please note that since [100] orientation is not perpendicular to [001] orientation as shown in Figure 5 (a), * is added to [100]. To investigate the effect of

anisotropic behavior of multi-finger device structure, four different possible orientation is studied as shown in Figure 28. Since [010] orientation has the highest thermal conductivity and [100]* orientation has the lowest thermal conductivity, both should affect the thermal cross-talk for in-plane and cross-plane for how they are placed.

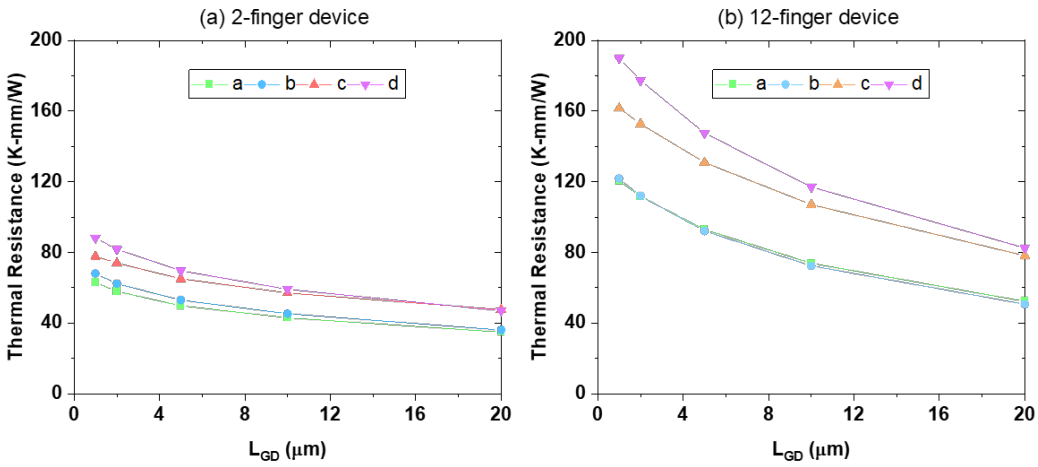

Figure 29. Thermal resistance of (a) 2-finger device and (b) 20-finger device for different orientation of β-Ga_2O_3 for various L_{GD}.

Figure 29 compares 2-finger device and 12-finger device for various L_{GD} with four different possible crystal orientation. "a" orientation from Figure 28 shows the lowest thermal resistance, while "d" orientation shows the highest thermal resistance. Orientation of "a" is the same orientation as 0-degree device from Chapter 4.3, and it still shows the lowest thermal resistance. As there are multiple metal including source, drain, gate, and the pads on top of the β-Ga_2O_3 channel layer, it is beneficial to align the higher thermal conductivity direction along the cross-plane direction for the purpose of facilitating efficient heat transfer from the active device region. If [010] orientation, the highest thermal conductivity direction, is aligned with the channel width, and the lowest thermal conductivity direction, [100]*, is aligned with the cross-plane direction, then the device

will suffer from the highest thermal resistance. If the orientation is not considered during design stage, the thermal resistance could increase up to 40% from "a" orientation to "d" orientation for 2-finger device with L_{GD} = 1 μm. When L_{GD} increased to 20 μm, the thermal resistance would decrease, but still the thermal resistance will increase ~35% if "d" configuration is utilized for the device instead of "a" orientation. When the number of channels increased to 12-finger, the differences increase more that the increase ratio becomes ~58% for both L_{GD} of 1 μm and 20 μm.

These results show corresponsive trend with Figure 27 that thermal resistance decreases with increased gate-to-drain spacing. Changing L_{GD} from 1 μm to 20 μm for 2-finger device, the thermal resistance will drop ~55% and ~53% for "a" and "d" configuration, respectively. For 12-finger device, the thermal resistance will drop ~43% for both "a" and "d" orientation. Similar to the results from Figure 27, as the number of channels of device increases, there is less chance to be affected by other design parameters that the drop rate is lower for the device with a greater number of channels.

In summary, from a thermal perspective, longer gate-to-drain spacing, small number of channels, narrow gate width with highest thermal conductivity aligned with cross-plane is desirable to have low thermal resistance. However, one should consider the changes in electrical characteristics as those geometrical parameters' changes. As shown in Figure 30 (a), as gate-to-drain spacing increases, the on-resistance between the channel increases resulting in higher breakdown voltage. Figure 30 (b) shows that a smaller drain-to-source spacing, L_{sd}, is better for characteristics of current and transconductance (g_m). Therefore, device engineers working in the promising field of the β-Ga_2O_3 device

technology should employ device layout co-design practices that account for both electrical and thermal effects.

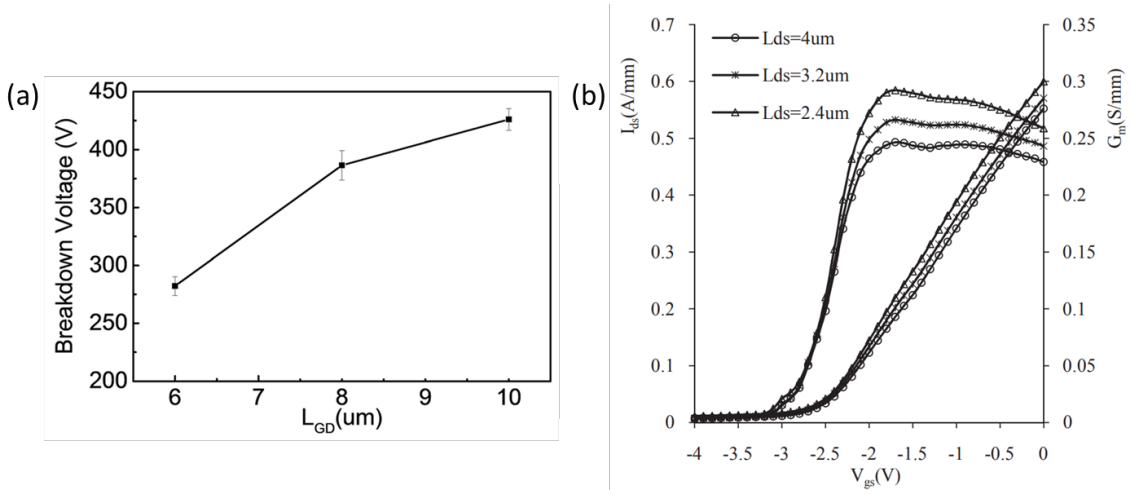

Figure 30. (a) Relationship between breakdown voltage and gate-drain spacing for the AlGaN/GaN HEMT device, L_{GD} [92]. (b) IV-characteristics and G_m characteristics for the AlGaN/GaN HEMT devices with different drain-source spacing, L_{ds} [93].

4.6 Conclusion

Previous studies have focused on the design of active and passive cooling solutions that add upon or alter the homoepitaxial configuration of β-Ga_2O_3 transistors. In contrast, this work has focused on how to optimize the device layout to mitigate self-heating, prior to implementing such engineering solutions. It was found that the channel orientation, distance between the gate and drain metal electrodes, and the geometry of the interconnects that link the metal electrodes and the bond pads can play a significant role in the dissipation of heat away from the device active region. These effects are pronounced in β-Ga_2O_3 devices as compared to GaN and SiC electronics, due to the relatively low and anisotropic thermal conductivity of the base material. It was found that aligning the gate/channel length along the orientation with the highest thermal conductivity is favorable for lateral devices built on (010)-oriented β-Ga_2O_3 substrates. While a longer gate-to-drain distance is

favorable in terms of increasing the device breakdown voltage, this is achieved at a price of sacrificing the device thermal performance. From a thermal standpoint, it is also recommended to use wide metal interconnects between the device metal electrodes and bond pads to enhance heat extraction by the bond wires. This work demonstrates that device engineers working in the emerging field of the β-Ga_2O_3 device technology should implement device layout co-design practices that account for both electrical and thermal effects.

CHAPTER 5. THERMALLY-AWARE DESIGN PROCESSES FOR VERTICAL DEVICES: CURRENT APERTURE VERTICAL ELECTRON TRANSISTORS (CAVET)

Content in the chapter (figures and text) adapted from:

S. Kim, *et al.,* "Thermal Management of β-Ga_2O_3 Current Aperture Vertical Electron Transistors," in *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 11, no. 8, pp. 1171-1176, Aug. 2021 [94]

5.1 Overview & Approach

Even though fail-safe operation and simplified designs are primarily engineered as lateral devices, many high-power infrastructures (*e.g.*, electrical power transmission, rail tracking, electric vehicle converter) demand voltages over 1 kV in combination with currents over 100 A as shown in Figure 1. Lateral devices are not ideal for those applications owing to a necessity for sizeable chip areas and potential reliability concerns arising from surface instabilities [48]. For applications demanding high voltage and high power levels, vertical transistors are required since they allow for superior field termination and current drive at the device-level while enabling fail-safe operation and simplified designs at the system level [18], [19], [47], [95], [96]. The average off-state breakdown voltage of lateral devices is ~ 600 V, while highest reported value is 2.32 kV [97], whereas the average breakdown voltage of vertical devices exceeds 1 kV [26], [27], [46], [98]. Most of the recently developed vertical devices are current aperture vertical electron transistors (CAVET), which were motivated by the commercially successful SiC double-implanted

MOSFET and modeled after the advanced GaN CAVET [99]–[103]. Several thermal management techniques for lateral structures were introduced in previous section, but most of them cannot be applied to the vertical structures. For example, a temperature reduction of 75% was demonstrated by reducing the thickness of β-Ga_2O_3 substrate [104]. In vertical architectures, on the other hand, the β-Ga_2O_3 substrate cannot be substituted and reducing the thickness of β-Ga_2O_3 will change the electrical performance that the breakdown voltage of vertical transistors scales with the drift layer thickness. Therefore, the options for thermal management in vertical devices are limited compared to thermal management of lateral devices.

For this study, we investigate to find the best way to reduce the junction-to-package thermal resistance and optimize for the vertical device, especially for CAVET structure by adding external solution since there is limitation of engineering the device layout. CAVET is a type of planar-gate vertical transistor capitalizes on deep-acceptor doping for junction formation with an adjoining n-type drift layer to establish a potential barrier for voltage blocking [26], [46], [48]. We investigated thermal management approaches for β-Ga_2O_3 CAVET to set out thermal guidelines when designing these devices and provide the optimal thermal design for CAVET structure. We developed CAVET by 2D Silvaco TCAD to obtain Joule heat profile, then imported the heating profile into 3D COMSOL thermal model making it one-way electro-thermo coupled model. From the baseline device, we are targeting to decrease the total thermal resistance less than 15 mm·°C/W, which is the resistance of current state-of-the-art GaN-on-Si HEMTs. In this work, we investigated bottom-sided cooling, top-sided cooling, and double-side cooling strategies for the CAVET structure, then provide the optimal cooling design for the CAVET.

5.2 Device Simulation

5.2.1 Principle of CAVET

The CAVET device design in β-Ga_2O_3 is likely to be inspired by the GaN and SiC CAVET design as shown in Figure 31. Due to the absence of p-type doping, the current blocking layer (CBL) in CAVET can be obtained by ion implantation. The role of the CBL in the CAVET is to block the flow of electrons from source to drain in both on state as well as off state and guide the electrons to the aperture of the device [48]. Since CBL can be achieved through ion implantation, the fabrication of the device can be relatively simplified. Also, the formation of the aperture without etching the material alleviates the etching issues impacting electrical properties such as new trap generation, leakage current, etc. The n-type layer on top of the CBL can be achieved either via ion implantation or by regrowth on top of the implanted CBL. The first CAVET was fabricated on bulk n-type β-Ga_2O_3 substrates with Mg as the CBL [105]. The Mg ion implantation in β-Ga_2O_3 has been

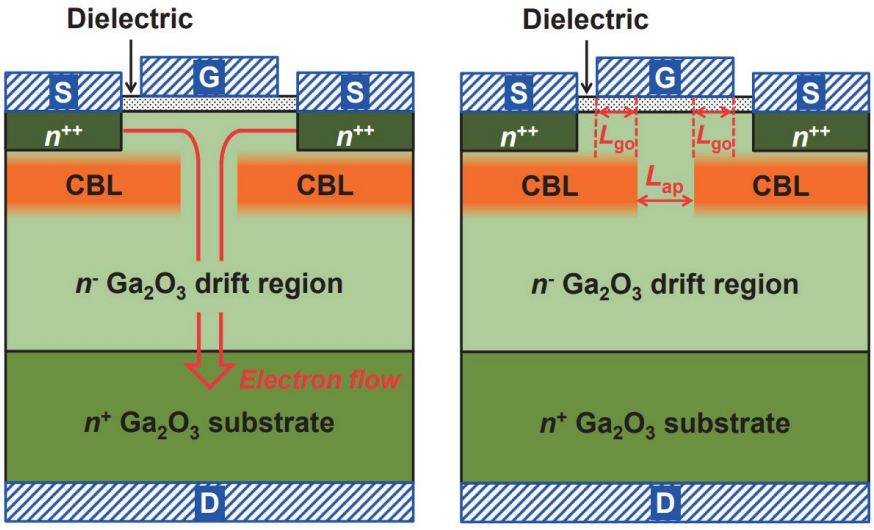

Figure 31. Structure and operation of a current aperture vertical Ga_2O_3 MOSFET (CAVET) [105]

shown to produce semi-insulating behavior as the fermi level is pinned relatively closer to the valence band (about 0.6–1 eV) [106]. The n-type β-Ga_2O_3 layer on top of the CBL and the n-type aperture region was achieved by Si ion implantation. This method avoided any regrowth and relied on ion implantations and thermal anneals to obtain the desired structure. The first reported CAVETs demonstrated current–voltage modulation but also exhibited significant leakage current in these devices. The high leakage current in these devices was attributed to the loss of Mg from the CBL during the ion implantation anneal process. This resulted in the background n-type carrier concentration to be higher than the Mg concentration leading to a high leakage current. Therefore, alternate species were sought to serve as the CBL in β-Ga_2O_3 CAVETs. Researchers demonstrated that nitrogen can be utilized instead of Mg as nitrogen can easily substitute on an oxygen site and also compensate for the n-type doping [107], [108]. The low energy requirement for nitrogen to occupy the oxygen site makes it thermodynamically favorable. The nitrogen was also significantly more resilient (in terms of diffusion) to high annealing temperatures compared to Mg, thereby, making it a better candidate for the CBL. Using nitrogen as the CBL, first, enhancement-mode β-Ga_2O_3 CAVET was demonstrated [48]. In this device, the n-type doping above the CBL and the aperture were carefully tailored to achieve a normally off device. A low doped region was maintained above the aperture and additional ion implantations were performed to obtain the heavily doped access and source regions.

5.2.2 *Electrical Simulation*

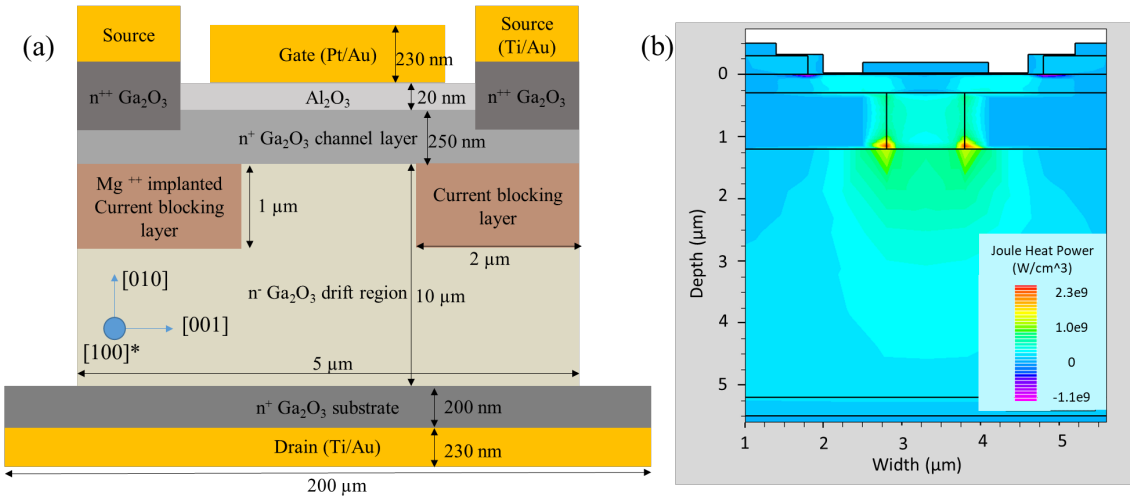

Figure 32. (a) (Not to scale) Schematic cross-section of β-Ga_2O_3 current aperture vertical electron transistor for 2-D device finite element simulation, (b) Joule heat power distribution of 2-D drift diffusion model using Silvaco ATLAS at bias condition of V_{GS} = 1 V and V_{DS} = 20 V.

The transistor structure we studied is shown in Figure 32 (a) with geometric details. The width of the device is 100 µm and the width of Ga_2O_3 substrate and drain are 200 µm. A 2-D drift diffusion model developed with Silvaco ATLAS was used to model the Joule heat power profile, then adapted to a 3-D thermal model using COMSOL Multiphysics. The electrical modeling scheme employed temperature dependent parameters such as electron mobility, electronic bandgap, and Schottky barrier height to accurately capture the negative differential resistance in the DC I-V characteristics. Also, device bias conditions and device geometry details should be known to calculate the heat generation distribution. Figure 32 (b) shows the calculated Joule heating profile, representing the heat dissipated due to the on-state resistance of the device, for the studied device under bias with a source-to-gate voltage of V_{GS} = 1 V and a source-to-drain voltage of V_{DS} = 20 V. Concentrated

heating exists at the edge of the current blocking layer (CBL) due to the electric field spreading through the structure. Using the map of the heat generation as a generalized guide to heating trends within the device, the channel region was chosen to act as a volumetric heat source within the device.

5.2.3 Thermal Simulation

The dissipated power density is on the order of 10 W/mm (or volumetric heat generation of $1x10^{16}$ W/m^3). For simplicity in the 3-D finite-element model we assumed a uniform heat generation, with the heating profile corresponding to that calculated by the Silvaco ATLAS simulation that was modeled as one slab concentrated between two current block layers. For various power dissipation densities we assumed that the channel is fully open and that there is no bias-dependence [87] which would affect the geometry of the Joule-heating region. Temperature dependent anisotropic thermal conductivity was adapted for β-Ga_2O_3, where $k_z = 23.4\times(300/T)^{1.27}$, $k_y = 10.7\times (300/T)^{1.21}$, and $k_x = 13.7\times(300/T)^{1.12}$ W/m-K in the [010], [100], and [001] crystallographic directions, respectively, as illustrated in Figure 32 (a) [109]. Since the effects of doping on the thermal conductivity of β-Ga_2O_3 is negligible [110], the channel, drift region, CBL, and substrate layers were lumped to have same thermal properties as bulk β-Ga_2O_3.

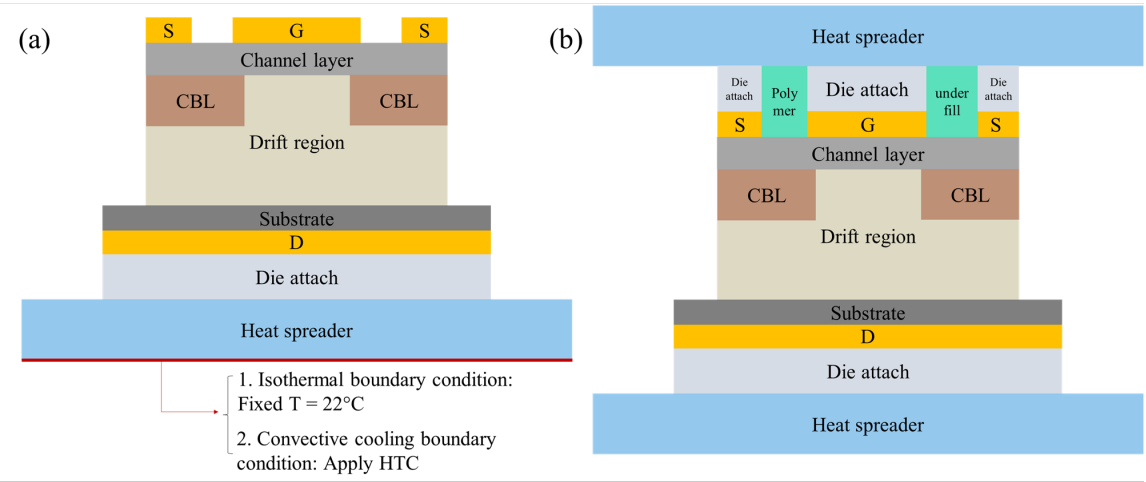

Figure 33. (a) Schematic of CAVET for 3D thermal simulation with two different boundary conditions for bottom-sided cooling, (b) Schematic for double-sided cooling, using the interconnection structure that the source, gate, and drain contact pads are connected onto a high-thermal-conductivity dielectric heat spreader using die attach material with a polymer-based underfill material for encapsulation.

Figure 33 (a) shows the bottom-sided cooling strategy that integrates the drain electrode with a high thermal conductivity die attach, which is in turn connected to a heat spreader that has a fixed temperature (isothermal boundary) or convective boundary condition applied to the bottom surface. The thickness of the die attach is 50 μm, that of the heat spreader is 2 mm, while the length and the width of the heat spreader are both 1 mm, respectively. For the isothermal boundary condition, the temperature of the heat spreader was fixed to be 22 °C, and a convective heat transfer coefficient (HTC) of 10 W/m^2-K was applied at all other surfaces exposed to ambient conditions to represent natural convection. For all modeling studies, ambient temperature was set to be 22 °C. To reduce computational complexity, a thin layer of Al_2O_3 (20 nm) and n^{++} Ga_2O_3 layer (200 nm) were neglected from the calculations and were instead modeled as β-Ga_2O_3 channel layer.

We first conducted the parametric studies for die attach and heat spreader thermal conductivities for the bottom-sided cooling with the isothermal boundary condition. For the die attach material, the thermal conductivity of sintered silver is about 200 W/m-K, which can also be modulated by the sintering temperature and particle size [111]. Solder could be used for die attach material, however it usually has a lower thermal conductivity than sintered silver that ranges from 10 to 60 W/m-K for a range of solders [112]. Thus, we set 10 W/m-K - 200 W/m-K as the range investigated in the parametric studies. As for the heat spreader, a thermal conductivity range from 100 W/m-K to 2,000 W/m-K was investigated, covering most of the heat dissipation substrate (such as Si, SiC, AlN, AlSiC, Cu, and diamond). For further study, we chose the baseline value for the thermal conductivity of the die attach and heat spreader: 130 W/m-K and 400 W/m-K, respectively, referred to [113].

Next, we changed the boundary condition from the isothermal boundary condition to a convective cooling boundary condition, to study convective cooling effects. A wide range of heat transfer coefficients (HTC) from 10 W/m^2-K to 100,000 W/m^2-K, were applied to the heat spreader uniformly, with a fluid temperature at 22 °C. The range of HTC covers the typical HTC values of moderate speed flow of air, active single-phase, and two-phase cooling methods.

Lastly, as shown in Figure 33 (b), along with the bottom-sided cooling, we added a top-sided cooling scheme where the source and the gate pads were connected to a high-thermal-conductivity dielectric heat spreader using die attach material with a polymer-based underfill material for encapsulation and insulation. The gate was insulated from the source by the polymer underfill, and we assumed metal connections were on the dielectric

heat spreader. With this setup, the heat can also be extracted from the top-side of the device into the higher thermal conductivity interconnects and further into the heat spreader and the heat sink. This top-sided cooling approach is attractive since the low-thermal-conductivity β-Ga_2O_3 layers have high thermal resistance hindering the heat dissipation through the bottom-side of the device. Here, we applied thermal conductivity of 200 and 1 W/m-k to the die attach material and polymer underfill, respectively.

5.3 Results and Discussion

5.3.1 Bottom-sided Cooling

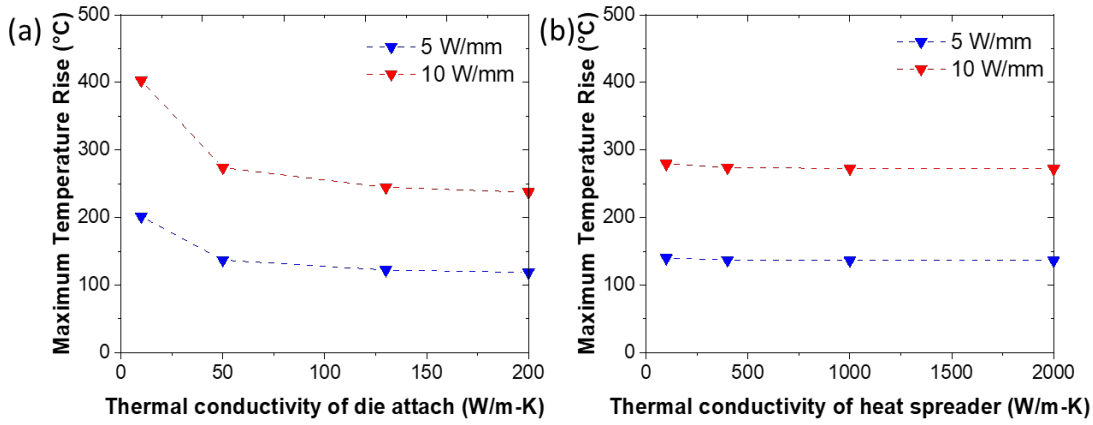

Figure 34. Bottom-sided cooling results: Modeling results of CAVET device with isothermal boundary condition, T_{bottom} = 22 °C, applying two different joule heating, 5 and 10 W/mm (or volumetric heat generation of 0.5×10^{16}, 1×10^{16} W/m^3) (a) Maximum channel temperature rise versus thermal conductivity of die attach (Thermal conductivity of heat spreader is fixed to 400 W/m-K), (b) Maximum channel temperature rise versus thermal conductivity of heat spreader (Thermal conductivity of die attach is fixed to 130 W/m-K).

To compare the effectiveness of changing the material for bottom-sided cooling for isothermal boundary condition, various thermal conductivity of die attach material and heat spreader were studied as shown in Figure 34 (a). When sintered silver was utilized for the

die attach material, k = 200 W/m-K the maximum temperature drops ~15% compared to using Au/Sn solder, k = 57 W/m-K. This demonstrates the importance of the die attach material in dissipating heat, particularly if the thermal conductivity of die attach is similar to temperature does not decrease as much once the thermal conductivity or lower than β-Ga_2O_3. Also, note that the maximum channel of die attach has increased to more than 130 W/m-K. This is because while the thermal resistance decreases in the die attach material with increased thermal conductivity, the high thermal resistance of low-thermal-conductivity β-Ga_2O_3 obstructs the flow of heat into the bottom-side.

Figure 34 (b) shows the maximum temperature as a function of the thermal conductivity of the heat spreader. The channel temperature showed little dependence on the thermal conductivity of the heat spreader, since 10-μm-thick β-Ga_2O_3 is acting as a thermal barrier with high thermal resistance. These results convey that rather than utilizing expensive high thermal conductivity material such as diamond (1000-2000 W/m-K), using common materials, such as Cu, AlN, or AlSiC (300-550 W/m-K), would be sufficient for bottom-sided cooling. For the concern of the thermal stresses, AlN and AlSiC have reasonable thermal expansion match to β-Ga_2O_3 compare to Cu, where the coefficients of thermal expansion (CTE) of AlN and AlSiC are 4.5×10^{-6} °C^{-1} and 6.5×10^{-6} °C^{-1}, respectively, and the range of the CTE of β-Ga_2O_3 is from 3.8 to 7.8×10^{-6} °C^{-1}, while Cu has high CTE with 16×10^{-6} °C^{-1} [113], [114].

5.3.2 Top-sided Cooling

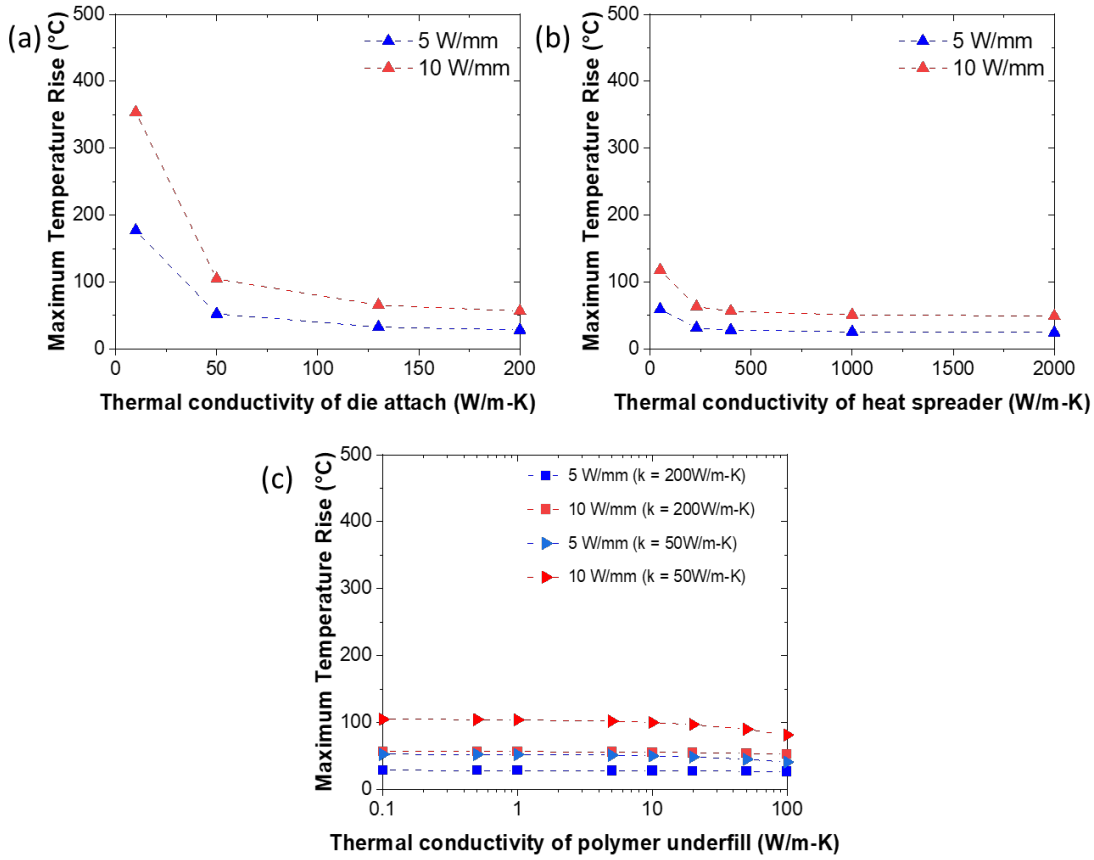

Figure 35. Top-sided cooling results: (a) Maximum channel temperature rise versus thermal conductivity of die attach material (Thermal conductivity of heat spreader is fixed to 400 W/m-K), (b) Maximum channel temperature rise versus thermal conductivity of heat spreader (Thermal conductivity of die attach is fixed to 130 W/m-K). (c) Maximum channel temperature rise versus thermal conductivity of polymer underfill for two different die attach material (k = 50 and 200 W/m-K),

To understand the effect of top-sided cooling scheme, a heat spreader is attached to the CAVET structure with polymer underfill to prevent any electrical shorts and die attach material that could provide an electrical pathway to gate and source. As heat source is close to the top-side of the device, maximum channel temperature rise decreases compared to bottom-sided cooling scheme. In addition, as it is closer to the heat source, the effects of thermal conductivity of die attach material and heat spreader of top-sided cooling are much more crucial than bottom-sided cooling. When sintered silver was utilized for the die attach

material, k = 200 W/m-K the maximum temperature drops ~46% compared to using Au/Sn solder, more than 3 × decreases compared to bottom-sided cooling (15% decrease). Moreover, changing the heat spreader from AlN or AlSiC to high thermal conductivity diamond, will decrease the maximum channel temperature to ~22%, which is significant compared to bottom-sided cooling.

Figure 35 (c) shows the maximum channel temperature rise as a function of underfill thermal conductivity (0.1-100 W/m-K). Polymer underfill materials such as epoxy resins and silicone gels typically have very poor thermal conductivities (~0.2 W/m-K). Thus, much effort has been made to increase their thermal conductivities, loading them with high thermal conductivity but electrically insulating particles, such as hexagonal boron nitride (hBN) or AlN [115]–[117]. Thermal conductivities up to 11 W/m-K have been achieved for hBN nanosheets/polybutylene terephthalate composites materials [118]. However, with this thermal conductivity value, 11 W/m-K, there is no change in maximum channel temperature. If we have poor die attach material with k = 50 W/m-K, then we could get a thermal pathway with a polymer underfill with thermal conductivity higher than 50 W/m-K. In this scenario, the temperature will drop ~12%. However, it is more plausible to have higher thermal conductive die attach material that with 200 W/m-K (utilizing sintered silver), the temperature rise would not be affected by the thermal property of polymer underfill.

5.3.3 Double-sided Cooling

Double-sided cooling scheme were explored by adding a heat spreader to the top-side of the device to allow for an additional pathway for heat dissipation. Applying

convective cooling boundary conditions to both the bottom- and top-side, the maximum channel temperature drastically dropped, as shown in Figure 36 (b).

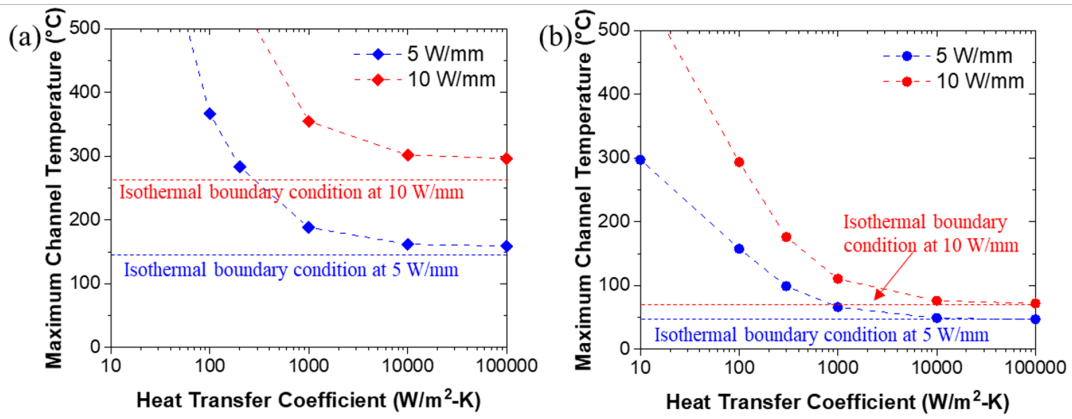

Figure 36. Modeling results of CAVET device with convective cooling boundary condition by applying heat transfer coefficient. (a) Bottom-sided cooling, (b) Double-sided cooling (dotted lines represent the data calculated with isothermal boundary condition with the baseline value for the thermal conductivity of the die attach and heat spreader: 130 W/m-K and 400 W/m-K, respectively).

Compared to bottom-sided cooling with a convective boundary condition, the temperature of double-sided cooling decreased about 50% for each HTC value as shown in Figure 36 (a). In addition, compared to bottom-sided cooling with isothermal boundary conditions, the maximum channel temperature of double-sided cooling with convective boundary condition was 45-60% lower than that of bottom-sided cooling, when we apply HTC between 1,000 and 10,000 W/m^2-K. When we pushed HTC to be 100,000 W/m^2-K, the maximum channel temperature dropped even more that the temperature approaches to the temperature of applying isothermal boundary condition to top- and bottom-side of the heat spreaders.

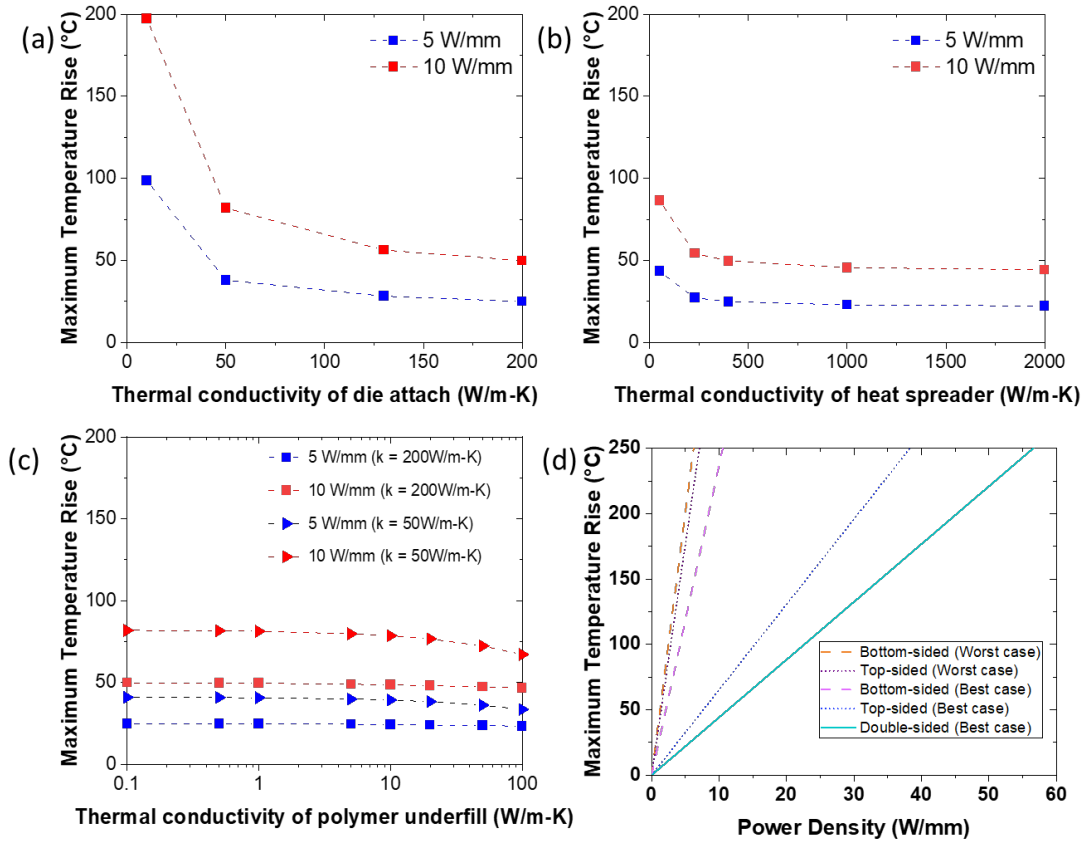

Figure 37. Double-sided cooling results: (a) Maximum channel temperature rise versus thermal conductivity of die attach material. (b) Maximum channel temperature rise versus thermal conductivity of heat spreader (Thermal conductivity of die attach is fixed to 130 W/m-K). (c) Maximum channel temperature rise versus thermal conductivity of polymer underfill for two different die attach material (k = 50 and 200 W/m-K), (d) Maximum channel temperature versus power density, showing that best cooling approach can increase power 5× higher than the baseline study for the same channel temperature.

The effect of die attach and underfill thermal conductivity were studied for the double-sided cooling structure. Die attach and underfill thermal conductivities were swept from 10 to 200 W/m-K, and from 0.1 to 100 W/m-K, respectively. Similar to the bottom-sided cooling, we found the thermal conductivity of die attach will significantly affect the heat flow on the top-side of the device. Figure 37 (a) plots the maximum channel temperature as a function of die attach thermal conductivity, when thermal conductivity of the underfill was fixed to 10 W/m-K. If we choose to use 10 W/m-K for the die-attach

material on the gate and source pads, then the temperature will only drop ~25% compared to bottom-sided cooling with HTC of 1,000 W/m^2-K. However, if we use a material with a thermal conductivity higher than 50 W/m-K, the maximum channel temperature will drop 65-75% when compared to bottom-sided cooling with HTC = 1,000 W/m^2-K as shown in Figure 37 (a).

Unlike bottom-sided cooling, but similar to top-sided cooling, Figure 37 (b) shows that thermal conductivity of heat spreader affects the maximum channel temperature rise of double-sided cooling. Changing the heat spreader from AlN or AlSiC (thermal conductivity of 200~320 W/m-K) to high thermal conductivity diamond (thermal conductivity of 2000 W/m-K), the maximum channel temperature will decrease down to ~20%, which is significant compared to bottom-sided cooling

Figure 37 (c) shows the maximum channel temperature rise as a function of underfill thermal conductivity (0.1-100 W/m-K), when the thermal conductivity of die attach was fixed to 200 W/m-K (utilizing sintered silver). Similar to top-sided cooling, Figure 37 (c) shows less than 5% difference in the maximum channel temperature for using polymer with thermal conductivity of 0.1 W/m-K and that of 100 W/m-K when using k = 200 W/m-K for die attach material. When the thermal conductivity of die attach is 50 W/m-K, then polymer with a thermal conductivity higher than 10 W/m-K must be used if we are to see a more than 10% decreases on the channel temperature.

Lastly, we studied the ideal case of using optimal thermal properties on a CAVET device with double-sided cooling, which is compared with multiple cases of bottom-sided and top-sided scheme. This ideal case was intended to determine what the expected limits of the double-sided cooling approach were. Thermal conductivity of 200 W/m-K and that of 10 W/m-K for the die attach material and polymer underfill, respectively, were utilized and the results are shown in Figure 37 (d). Comparing to the worst-case bottom-sided

cooling case, optimized bottom-sided cooling scheme shows 60% decrease in the device thermal resistance, while top-sided and double-sided cooling scheme show 84%, 89% decrease, respectively. For the same maximum channel temperature, not only the device power density increased up to a factor of 9 higher than the baseline structure, but the device thermal resistance decreased from 40.2 mm·°C/W to 4.42 mm·°C/W, 89% decrease in the channel temperature, beating the resistance of current state-of-the-art GaN-on-Si and GaN-on-SiC HEMTs, 15 mm·°C/W and 5 mm·°C/W, respectively. Because of the low thermal conductivity and highly anisotropic thermal conductivity of β-Ga_2O_3, heat could hardly be mitigated through the stack of the device. As such, it is strongly recommended to have double-sided cooling scheme to mitigate heat from the vertical device.

5.3.3.1 Anisotropic effect study

Similar to lateral device from previous chapter, the anisotropic effect on thermal resistance was studied that as the orientation of substrate changes, thermal resistance would change due to anisotropic thermal conductivity. As shown in Figure 38, various orientational options were investigated for the baseline material properties compared to the best-case scenario showed in Figure 39 (b). a-orientation in Figure 38 (b) is the preferrable orientation that was obtained from previous study that highest thermal conductivity in cross-plane (010) orientation, and for in-plane, having higher thermal conductivity toward y-orientation shows the lowest temperature rise as shown in Figure 39 (b). When the lowest thermal conductivity of (100) orientation is aligned with the substrate cross-plane, d-orientation, then the maximum temperature rise increases ~45% for baseline case, and ~60% for best-case. Therefore as we also learned from previous chapter, it is important to align highest thermal conductivity with the cross-plane.

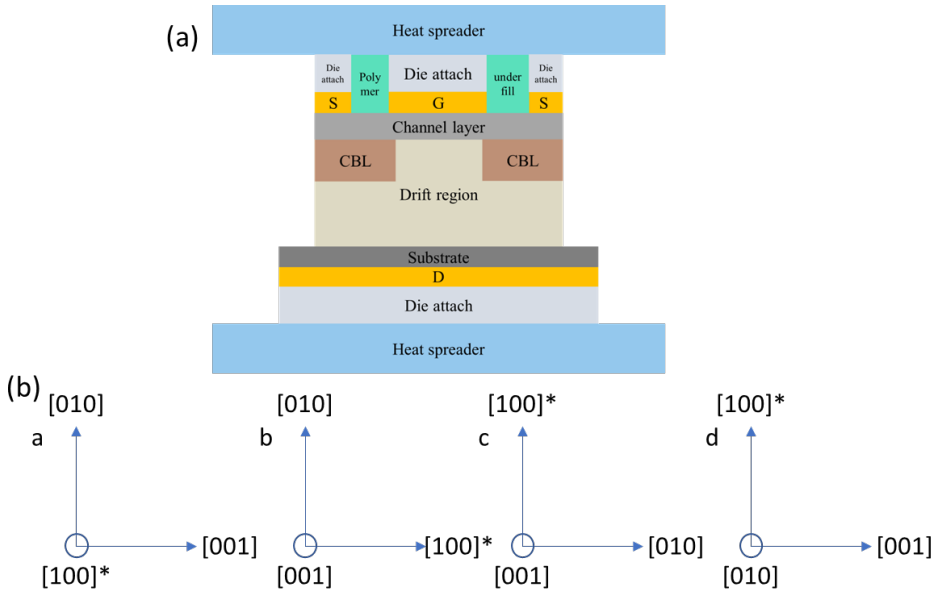

Figure 38. (a) Extracted from Figure 33 (b), (b) Studied orientation of β-Ga_2O_3 substrate.

(a)

	Die attach	Underfill	Heat spreader
Baseline	50	0.1	230
Best	200	1	400

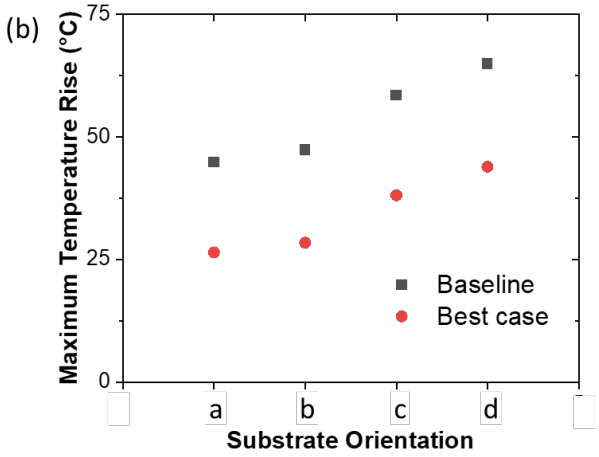

Figure 39. (a) Material properties used in this anisotropic effect study for baseline-case and best-case scenario, (b) Maximum temperature rise of two cases when 5 W/mm (or volumetric heat generation of 0.5×10^{16} W/m^3) is applied.

5.3.3.2 Error estimation from the assumption

So far, the study was based on a uniform heat generation that was modeled as one slab concentrated between two CBLs. And the 300 nm-thick channel and 200 nm-thick substrate layers were lumped to have same thermal properties as bulk β-Ga_2O_3. However, as shown in Figure 32 (b), highest Joule heat power occurs as the edge of the CBL. For the lateral FETs, gate controls the current flow from the source to drain, and there is a bottleneck of hot electrons at the gate edge on the drain side. Similarly, as current is blocked by the CBLs and the current flows from the source to the drain, the bottleneck occurs at the corner of the CBL. Thus, few cases were investigated to estimate the difference from the studied cases. In addition, thermal conductivity of thin film β-Ga_2O_3 is much lower than that of bulk β-Ga_2O_3, ~5 W/m-K for 200 nm-thick β-Ga_2O_3 film.

Figure 40 shows the maximum temperature rise for the assumed conditions (utilizing thermal conductivity of bulk β-Ga_2O_3 and applying heat flux between the CBL), when channel and substrate thermal conductivities are thickness dependent, when the maximum Joule heat flux is applied at the corner of the CBL, and considering the both. Also, Figure 40 shows the impact of the orientation that a- and d-orientation are studied as illustrated in Figure 38 (b), and the impact of the material properties considering two cases based on Figure 39 (a).

For a-orientation, which is preferrable orientation, shows ~5% increase in temperature for baseline device, when thickness dependent thermal conductivity is utilized. When better materials are utilized (green bar chart in Figure 40), even though the absolute difference remains the same, ~3 °C, as the maximum temperature drops, the error increases to ~9%. When the substrate is oriented as d-orientation, the difference increases more up to 13%. When the heat flux is applied at the edge of the current blocking layer (CBL), the temperature rise of a-orientation increases ~7%, while that of d-orientation increases

~14%. Therefore, when you utilize the right orientation, the error of the assumption of this study would be less than 10%, while the anisotropic behavior could impact the temperature rise up to 60%.

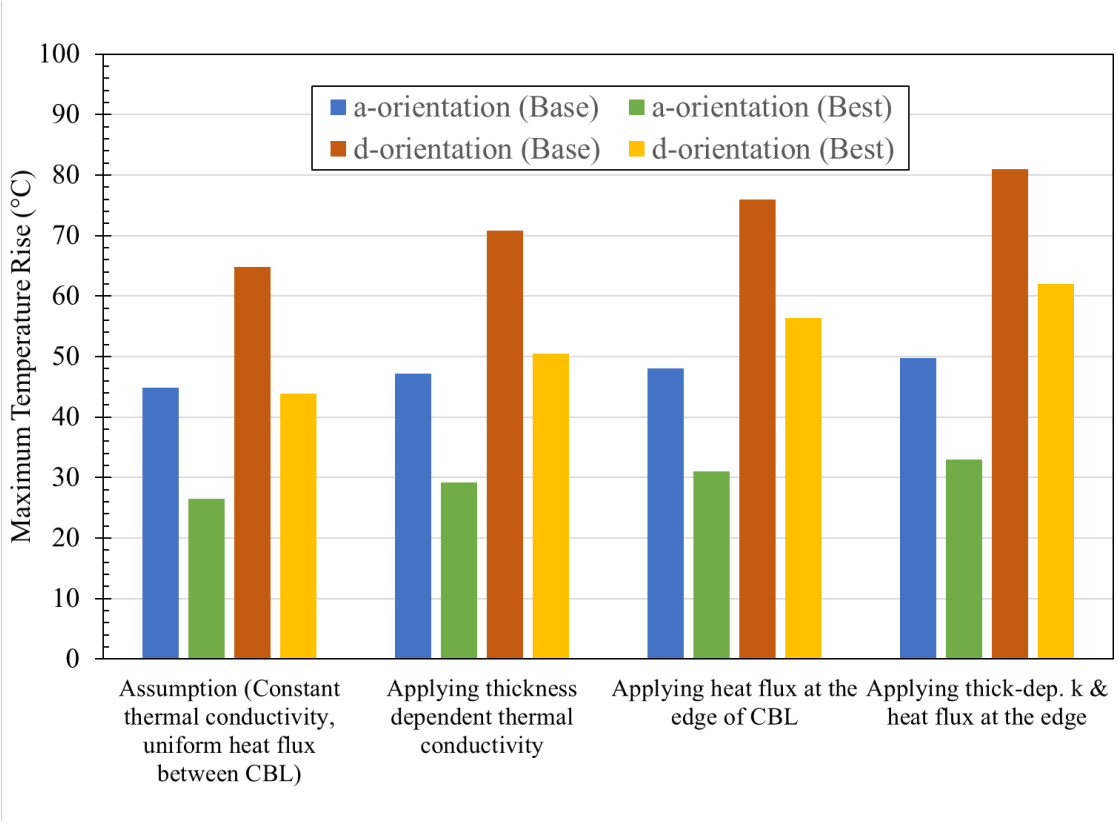

Figure 40. Comparing maximum temperature rise for different assumption conditions. a- and d- orientation are considered with two cases of Figure 39.

5.4 Conclusion

In summary, this work presents the comparative thermal modeling results for recently reported vertical β-Ga_2O_3 structures. It has been frequently pointed out that to fully achieve the high potential of β-Ga_2O_3 for power electronics applications, we have to thermally engineer the device architectures to counteract the effects of the low thermal conductivity of β-Ga_2O_3. The key issue is that the thermal resistance of the entire heat dissipation path and thermal conductivity of the β-Ga_2O_3 is only one factor that plays a role

in this limit. Thus, by using the appropriate materials and device architectures for CAVET devices, it is possible to improve the thermal performance of devices. Bottom-, top-, and double-sided cooling results show that the device temperature can be decreased by increasing the thermal conductivity of heat spreader, die attach material, and heat transfer coefficient. In addition, the model demonstrates the importance of top-sided cooling to alleviate the heat from the top-side of the vertical CAVET, which is closer to where joule heating occurs. Ultimately, architectures that utilize cooling from both sides of the device will perform best, enabling an 89% reduction in device temperature or 9× increased operational power at the same temperature limits.

CHAPTER 6. OPTIMIZATION OF DEVICE-LEVEL THERMAL SOLUTION CONSIDERING BOTH STEADY-STATE AND TRANSIENT REGIME

Content in the chapter (figures and text) adapted from:

S. Kim, *et al.,* "Device-level Transient Cooling of β-Ga_2O_3 MOSFETs," *2021 20th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (iTherm),* June. 2021

6.1 Overview & Approach

Package- and system-level thermal management solutions, which are designed based on steady-state operation, were shown to be ineffective for applications that operate under transient thermal loading, potentially leading to overdesigned cooling systems [49]. The device thermal time constant [71], [119], [120] (τ; the rise time for a device to reach ~63.21% of its steady-state temperature) [121] is inversely proportional to the thermal diffusivity. Since the thermal conductivity of β-Ga_2O_3 is an order of magnitude lower than those for GaN and SiC, the thermal diffusivity is also an order of magnitude lower. This renders β-Ga_2O_3 transistors to possess a significantly longer thermal time constant than those for GaN and SiC devices. Because of this relatively long thermal time constant, the heat diffusion length in β-Ga_2O_3 is limited for fast transient thermal loading. Therefore, this work highlights key considerations for the design of transient cooling solutions for high power β-Ga_2O_3 electronic devices using transient thermal modeling. Furthermore, this work will prove that traditional thermal solutions, such as those developed for relatively high thermal conductivity systems [122], are unsuitable to β-Ga_2O_3.

6.2 Device Model and Validation

Table 4. Thermo-physical properties of Ga_2O_3 and other semiconductor materials

Material Property	**β-Ga_2O_3** [41]	**Diamond** [123], [124]	**GaN** [125]	**4H-SiC** [68]
Density (g/m^3)	6.44	3.5	6.15	3.21
Specific heat (J/kg-K)	490	520	490	670
k_x (W/m-K)	13.7 × (300/T)$^{1.12}$ [001] direction	For 2 µm thick (First 500 nm: 85, Second 500 nm: 175, Third 1 µm: 309) Bulk: 2158	180 (4 µm thick)	490
k_y (W/m-K)	10.7 × (300/T)$^{1.21}$ [100] direction			
k_z (W/m-K)	23.4 × (300/T)$^{1.27}$ [010] direction	For 2 µm thick (First 500 nm: 142, Second 500 nm: 310, Third 1 µm: 510) Bulk: 2158		
Thermal Boundary Resistance at 300 K (m^2-K/GW)	7.3[83], [126], 30.2 [68], 47.1 [54]		4.35 [127]	
	-	7.3 [126]		

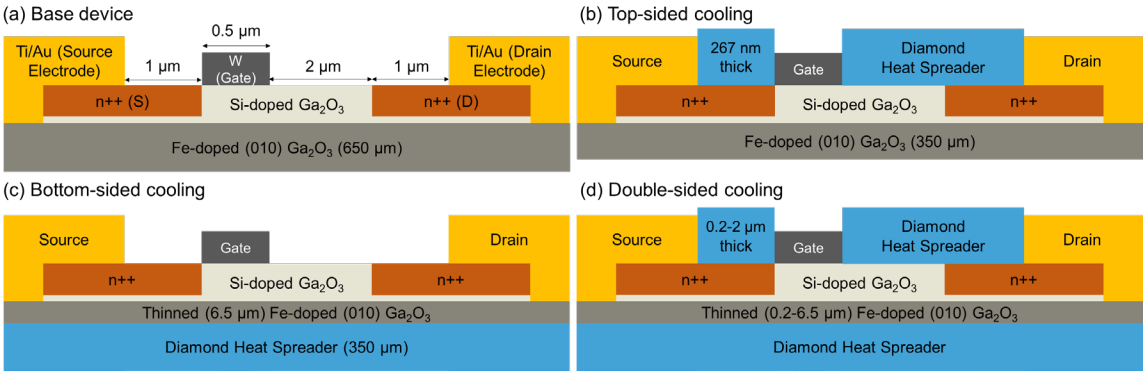

Figure 41. (a) Schematic of the base β-Ga_2O_3 MOSFET, (b) Top-sided cooling scheme: 267 nm thick diamond heat spreader was grown on β-Ga_2O_3, (c) Bottom-sided cooling scheme: Diamond heat spreader was bonded to thinned β-Ga_2O_3, (d) Double-sided cooling scheme: Various thickness options of top-side diamond and β-Ga_2O_3 were considered.

A transient device thermal model was created based on previous study, Chapter 4, Figure 21 (a). This model was then extended to investigate a hypothetical β-Ga_2O_3 device that employs a polycrystalline diamond passivation layer grown on top of the β-Ga_2O_3

channel (Figure 21 (b): top-sided cooling scheme), fabricated on a β-Ga_2O_3/diamond composite substrate (Figure 21 (c): bottom-sided cooling scheme), and lastly, adding both top and bottom heat spreader (Figure 21 (d): double-sided cooling scheme). For top-sided cooling, it is assumed that 267 nm thick polycrystalline diamond is grown on β-Ga_2O_3 (similar to the work [68]) with thermal conductivity of 110 W/m-K, with TBR of β-Ga_2O_3-diamond interface of 30.2 m^2-K/GW [68]. For bottom-sided cooling, the β-Ga_2O_3/diamond composite wafer is assumed to be constructed by bonding a 6.5 μm thick β-Ga_2O_3 layer thinned from the host substrate (similar to our previous work [54]) onto a polycrystalline diamond substrate with a thickness of 350 μm with TBR of 47.1 m^2-K/GW [54]. Throughout this study, as shown in Figure 21 (d), the thickness of top-side diamond heat spreader and the thickness of β-Ga_2O_3 layer will be optimized that may offer the thermal performance of β-Ga_2O_3 MOSFETs comparable to that for commercial GaN-on-Si and/or GaN-on- SiC technologies. For these simulated device structures, the device geometries (gate-to-source distance, gate length, gate-to-drain distance) were kept identical to the homoepitaxial MOSFET.

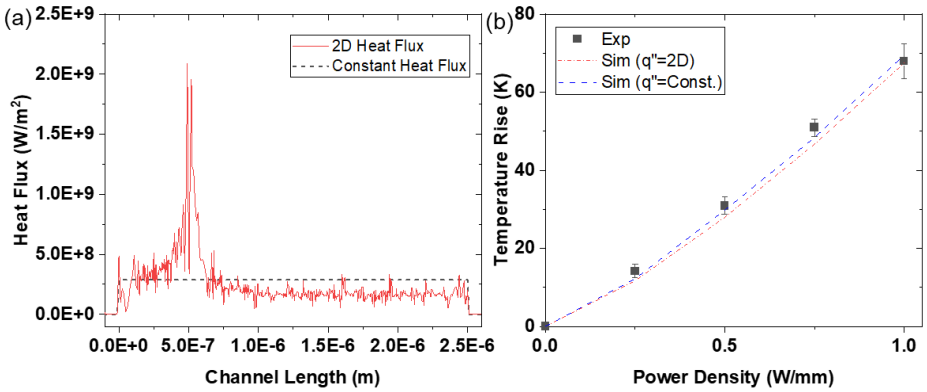

Figure 42. (a) Heat flux obtained from 2D electrical model (Synopsys Sentaurus TCAD software) for 1 W/mm, and corresponding constant heat flux, (b) Experimental results compared with 2D heat flux and constant heat flux.

In the previous work, 3D coupled electro-thermal modeling was performed by coupling a 2D electrical model (Synopsys Sentaurus TCAD software) [128], [129] with a 3D finite element thermal model with a detailed solid geometry that represents that of a real device (COMSOL Multiphysics) [38]. However, to reduce the computational loads for transient modeling study, the constant heat flux is applied to the fully-open channel [130] from source to drain, instead of applying the obtained 2D heat flux from the 2D electrical model, as shown in Figure 39 (a). Figure 39 (b) shows strong agreement between the experimental data and simulation results of both the 2D heat flux and constant heat flux. Throughout the remainder of this study, a constant heat flux corresponding to a power density of 1 W/mm was selected for use in the 3D thermal model due to the great agreement between the experiments, imported heat flux, and the constant heat flux.

6.3 Effect of Various Cooling Solutions

6.3.1 Transient thermal response of base device structures

Figure 40 (a) shows the normalized transient temperature rise with respect to their steady-state temperature rise of both the homoepitaxial β-Ga_2O_3 MOSFET and the GaN-on-Si HEMT under 1 W/mm and 1.6 W/mm power dissipation levels, respectively. Under steady-state, the β-Ga_2O_3 MOSFET exhibits a 2.7 × higher temperature rise than the GaN HEMT despite the β-Ga_2O_3 MOSFET is operating under a ~38% lower power density. The corresponding device-to-package thermal resistances of the β-Ga_2O_3 MOSFET and the GaN HEMT are 65 K-mm/W and 15 K-mm/W, respectively. The GaN device clearly exhibits a shorter thermal time constant than the homoepitaxial β-Ga_2O_3 MOSFET as

indicated with the green line in Figure 40 (a), which means its channel temperature reaches the steady-state value much faster than the β-Ga_2O_3 device.

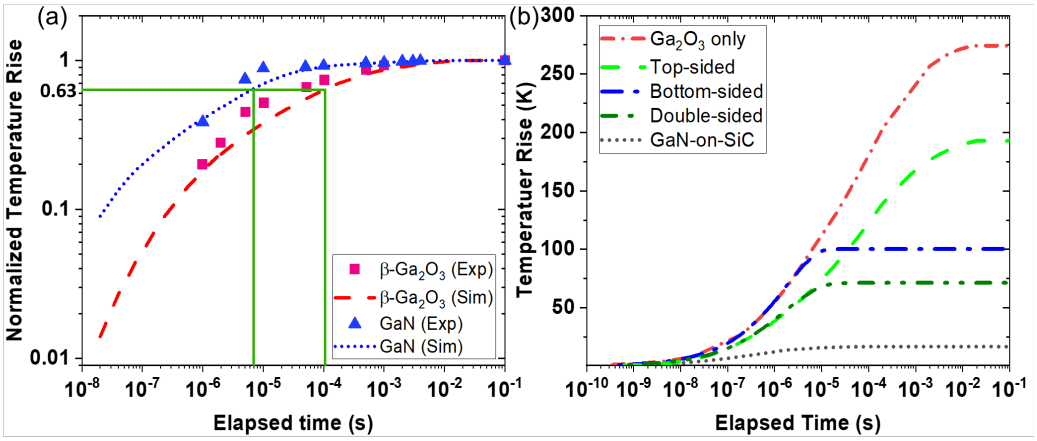

Figure 43. (a) Normalized temperature rises of the β-Ga_2O_3 MOSFET and GaN-on-Si HEMT with respect to their steady-state temperature rise. The steady-state temperature rise of the β-Ga_2O_3 MOSFET is 65°C for a power dissipation level of 1 W/mm [38]. The steady state temperature rise of the GaN-on-Si HEMT is 24°C for a power dissipation level of 1.6 W/mm [131]. (b) Transient temperature rise under a power density of 4 W/mm for a homoepitaxial β-Ga_2O_3 MOSFET (Figure 38 (a): Ga_2O_3-only), a diamond passivation layer on a β-Ga_2O_3 MOSFET (Figure 38 (b): Top-sided), a β-Ga_2O_3 MOSFET fabricated on the composite substrate (Figure 38 (c): Bottom-sided), a β-Ga_2O_3-on-diamond MOSFET further augmented by diamond passivation (Figure 38 (d): Double-sided), and a GaN-on-SiC HEMT.

To estimate the channel temperature rise in the four device architectures from Figure 38 under a realistic power dissipation level, device simulation was performed at a power dissipation level of 4 W/mm, and results are plotted in Figure 40 (b). For further study, current state-of-the-art of GaN-on-SiC HEMT is compared instead of a GaN-on-Si device. Today's GaN devices typically operate under 5–6 W/mm to ensure that the operating temperature does not exceed the safe allowable range for reliable operation [122]. Without any cooling solution applied, the steady-state channel temperature rise of the homoepitaxial β-Ga_2O_3 MOSFET (Figure 38 (a)) is 278°C (i.e., the channel temperature is

298°C while the base temperature is 20°C), which exceeds typical operational safety limits (e.g., 175°C for GaN and 125°C for Si devices) [71]. Because of the low thermal conductivity of β-Ga_2O_3, replacing the β-Ga_2O_3 substrate with diamond (Figure 38 (c)) reduces the steady-state temperature rise by ~64% (dropping from 278°C to 100°C). For high frequency power switching applications operating beyond the ~10^2 kHz range (elapsed time < ~10^{-5} s), however, employing a composite substrate (i.e., bottom-sided cooling) does not improve the transient thermal response (i.e., self-heating) of the device as shown in red and blue curves in Figure 40 (b). The channel temperature rises for both device structures are identical up to ~3×10^{-6} s, which corresponds to transient thermal loading under ~300 kHz. Therefore, solely relying on a bottom-sided cooling strategy (i.e., employing a composite substrate similar to the case of GaN-on-diamond devices)[122] is insufficient for the thermal management of pulse-powered β-Ga_2O_3 MOSFETs. The addition of a top-sided heat spreader (i.e., diamond passivation, Figure 38 (b)) reduces the steady-state temperature rise by ~28% (decreasing from 278°C to 198°C), and slightly reduces the transient temperature rise during short transient conditions (e.g., elapsed time < ~10^{-5} s). However, it is still not comparable to a GaN-on-SiC HEMT for bottom-sided and top-sided cooling scheme so that we configure a double-sided cooling scheme, as shown in Figure 38 (d). Utilizing double-sided cooling scheme would drop the steady-state temperature by ~75%, compared to the homoepitaxial β-Ga_2O_3 MOSFET, with the slight decrease in the short transient regime.

6.3.2 Top-sided Cooling

Recently, polycrystalline diamond was first epitaxially grown on β-Ga_2O_3 with thicknesses of 267 nm and 960 nm [68]. For this parametric study, the thickness of the

diamond was increased up to 2 µm and the corresponding thermal properties were adapted from Table 3. Figure 41 shows the impact of the thickness of the top-side diamond passivation layer. As the thickness was increased, not only the thermal conductivity of diamond increases, but also the ability to spread heat increases that both steady-state and transient channel temperature decrease due to the proximity to the heat source. Figure 41 also includes the effect of TBR. It was found that the TBR is negligible for the 267 nm thick diamond, but there would be ~8% and ~18% differences in the steady-state temperature rise between the lowest and the highest TBR for 1µm and 2 µm thick diamond (i.e., for Figure 41 (c), temperature rise with TBR of 47 m^2-K/GW is 71°C, while that of 7.3 m^2-K/GW is 58°C), respectively. The transient temperature rise also decreases with an increased thickness of the top-side diamond, but the time constant for all three thicknesses was found to be ~10^{-4} s, which corresponds to the time constant of the base device as shown in Figure 40 (a).

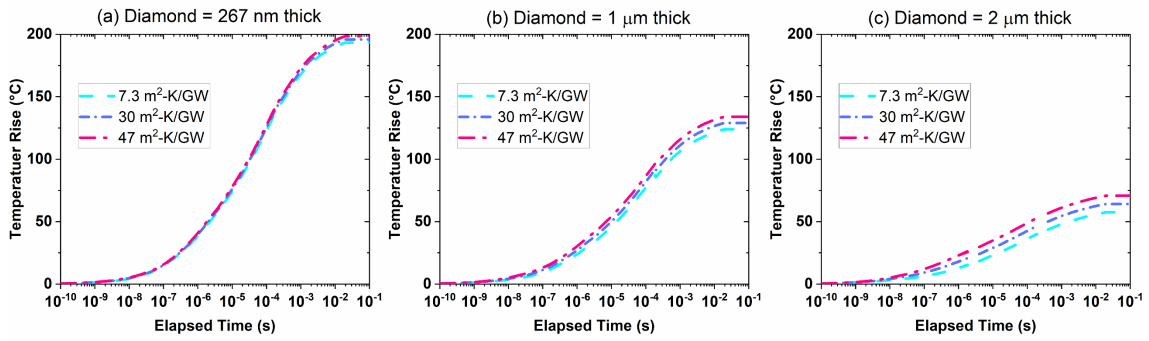

Figure 44. Top-sided cooling scheme with various diamond heat spreader thickness: (a) 267 nm thick, (b) 1 µm thick, (c) 2 µm thick

6.3.3 Bottom-sided Cooling

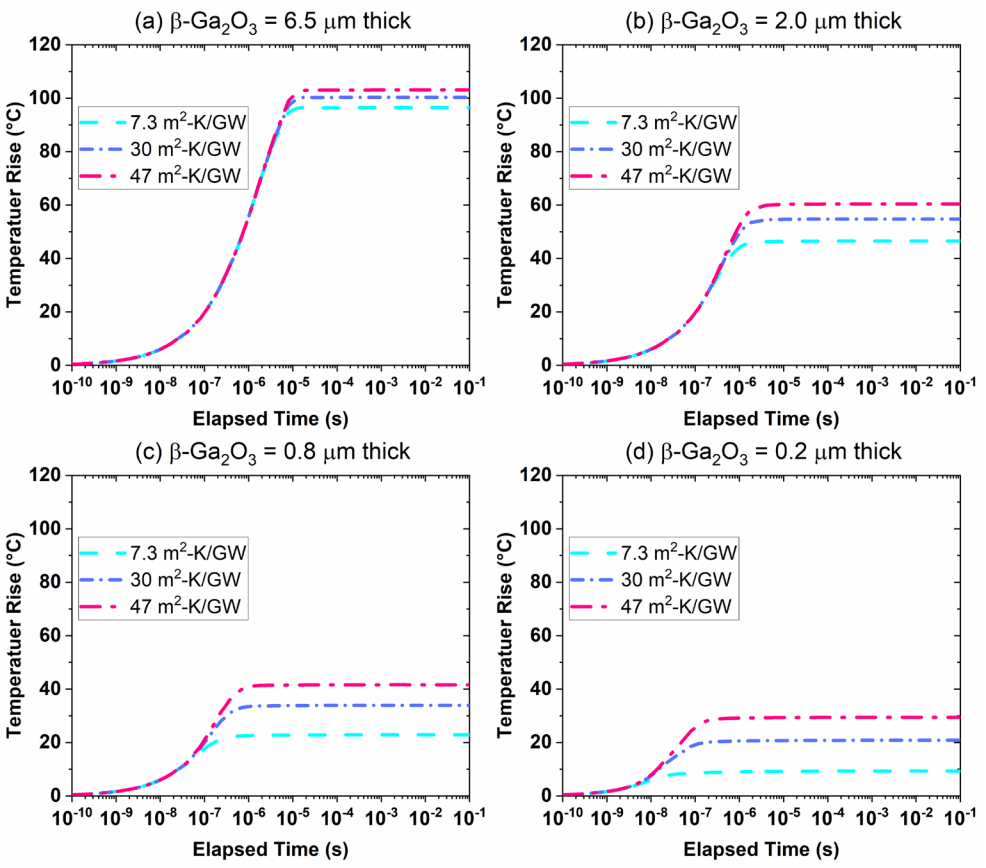

Figure 45. Bottom-sided cooling scheme with various β-Ga_2O_3 thickness: (a) 6.5 µm thick, (b) 2.0 µm thick, (c) 0.8 µm thick, (d) 0.2 µm thick.

In order to further enhance the cooling performance, the addition of a diamond composite wafer on the bottom surface of the β-Ga_2O_3 was investigated. In particular, the effect of the β-Ga_2O_3 thickness with various TBR values was studied in detail as shown in Figure 42. As the β-Ga_2O_3 was thinned from 6.5 µm to 0.2 µm thick, the steady-state temperature rise dropped by ~77% from 103°C to 23°C for a TBR of 47 m^2-K/GW. If the TBR of the bonded interface of the β-Ga_2O_3-diamond could be reduced to 7.3 m^2-K/GW, then by reducing the thickness of the β-Ga_2O_3 from 6.5 µm to 0.2 µm, the steady-state temperature rise will drop ~90% from 96°C to 9.3°C. Additionally, the thermal time constants as well as the transient temperatures decrease as the thickness of the β-Ga_2O_3

decreases. Moreover, as the TBR of the interface of β-Ga_2O_3-diamond decreases, the time constant decreases as well. For a TBR of 47 m^2-K/GW, the time constant of the 6.5 µm thick β-Ga_2O_3 was found to be ~1.5×10^{-6} s, while that of 0.2 µm thick β-Ga_2O_3 was found to be ~4×10^{-8} s, decreased by two orders of magnitude. For TBR of 7.3 m^2-K/GW, the time constant of 6.5 µm thick β-Ga_2O_3 is 1.25×10^{-6} s, while that of 0.2 µm thick β-Ga_2O_3 is ~9.5×10^{-9} s. Therefore, the time constant of bottom-sided cooling scheme decreases with decreased thickness of β-Ga_2O_3, while the time constant of top-sided cooling scheme is not affected by the thickness of the diamond layer.

6.3.4 Double-sided Cooling

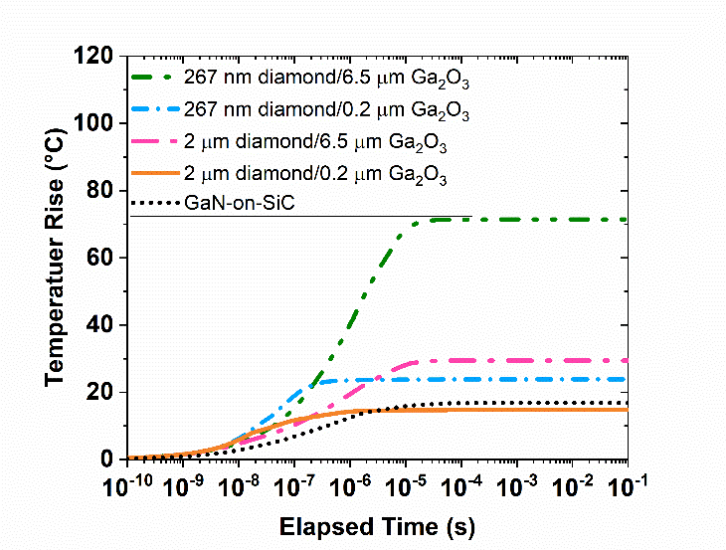

Figure 46. Transient channel temperature rises of double-sided cooling scheme compared to GaN-on-SiC HEMT.

Based on previous studies, the four different possible double-sided cooling solutions were compared, including a 267 nm or 2 µm thick top-side diamond heat spreader and 6.5 or 0.2 µm thick β-Ga_2O_3, as shown in Figure 43. For the top-side interface, a TBR of 30

m^2-K/GW was used, and for bottom-side interface, a TBR of 47 m^2-K/GW was used. Understandably, the combination that includes the thickest top-side diamond and the thinnest β-Ga_2O_3 shows the lowest temperature rise. This reduced temperature is comparable to GaN-on-SiC device, and by increasing the interface quality, it can be possible to reduce the temperature even further. The diamond passivation layer with a moderately high thermal conductivity effectively reduces the device temperature not only under steady-state conditions, but also under the high frequency operating regime, since it is located in proximity (i.e., less than several tens of nanometers) to the β-Ga_2O_3 devices active region where the Joule heating occurs. Therefore, device-level thermal management of β-Ga_2O_3 MOSFETs requires the combined use of a composite wafer and a top-side heat spreader in order to handle the thermal loading that occurs during both direct current (DC; steady-state) and pulsed (transient) operating conditions.

6.4 Effect of Base Temperature

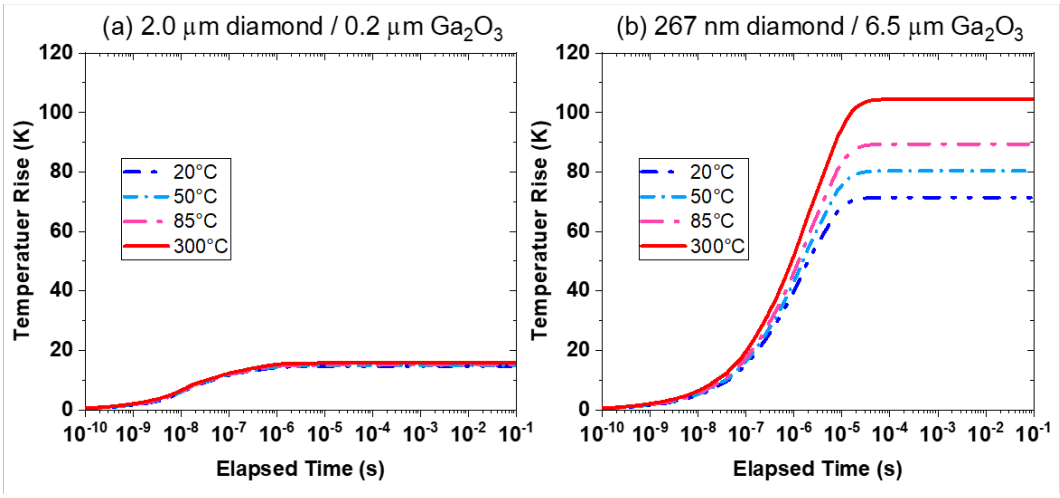

Figure 47. (a) Effect of base temperature on the channel temperature rise for double-sided cooling scheme of: (a) 2 µm thick diamond / 0.2 µm thick β-Ga_2O_3, and (b) 267 nm thick diamond / 6.5 µm thick β-Ga_2O_3.

The previous studies in this report were based on the base temperature of 20°C. However, in reality, many applications often result in different environment temperatures, thus changing the channel temperature. For example, a base temperature of 85°C is common for industrial applications, while a temperature of 300°C can be expected for space applications [5], [132]. Here, the best- and worst-case scenarios were compared for a device using the double-sided cooling from Figure 43. For the best-case scenario with the shorter time constant ($\sim 4 \times 10^{-8}$ s), no effects in the channel temperature rise are observed as the base temperature increases, as shown in Figure 44 (a). However, as shown in Figure 44 (b), with a lack of enough cooling power (or the system has the long time constant), the steady-state temperature can increase up to 50% (from 71°C to 104°C) if the base temperature increases from 20°C to 300°C, while the temperature rise in transient regime can increase up to 25% from 10^{-8} to 10^{-6} s (1-100 MHz regime).

6.5 Effect of Multi-pulses

Lastly, the effect of multi-pulses was investigated. Since the system has a relatively long time constant, the channel temperature may not be cooled down to the base temperature, resulting in higher peak temperature rise for the next pulse. For this study, a 20% duty cycle was applied for five different pulse periods, and the temperature rise for first six pulses of each pulse periods was investigated. As discussed in previous section, the time constant of the double-sided cooling scheme with 2 μm thick diamond / 0.2 μm thick β-Ga_2O_3 device is $\sim 4 \times 10^{-8}$ s, thus, the peak temperatures do not increase when the period is longer than 10^{-7} s. In contrast, even though the time constant of homoepitaxial β-Ga_2O_3 is $\sim 10^{-4}$ s, the peak temperature slightly increases for the period of 10^{-3} s, so that the peak temperature difference is ~5% between the temperature of the first peak and the sixth peak as shown in Figure 45 (d). This result indicates that not only the time constant matters for the cooling time, but the magnitude of the pulse's peak temperature is also important, since the temperature is not able to reach the base temperature.

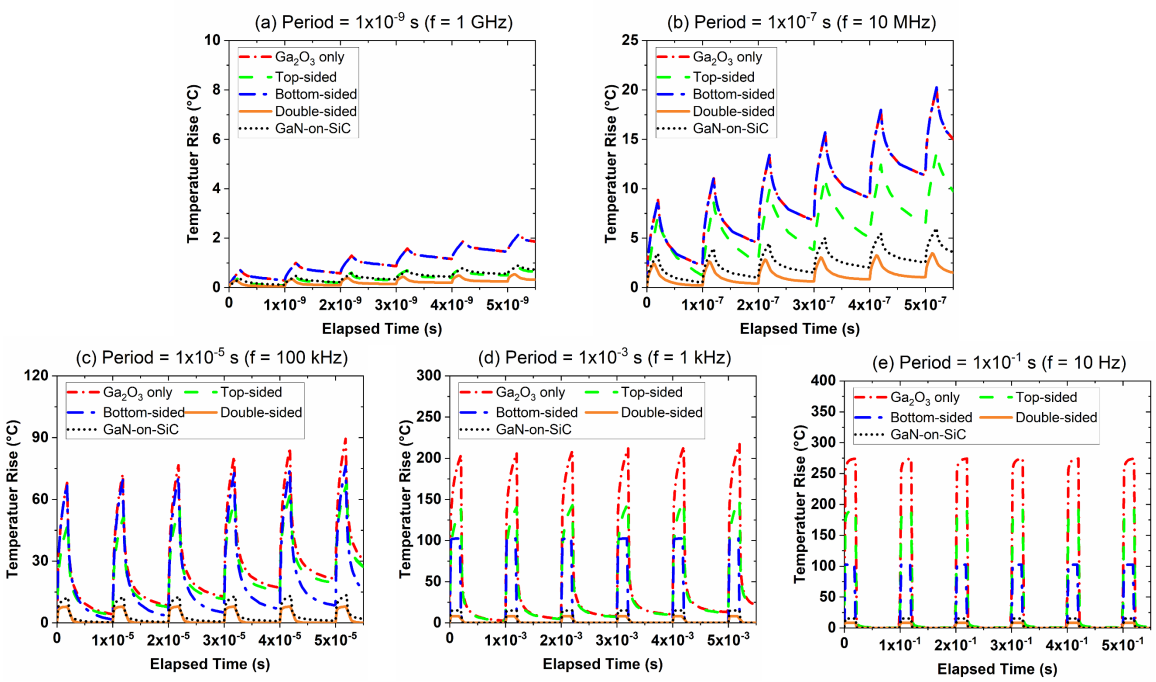

Figure 48. (a)-(e) Effect of multi-pulses for various time periods (or frequencies). 267 nm thick diamond was simulated for top-sided cooling with TBR of 30 m^2-K/GW, 6.5 µm thick β-Ga_2O_3 was used for bottom-sided cooling with TBR of 47 m^2-K/GW, and 2 µm thick diamond / 0.2 µm thick β-Ga_2O_3 were utilized for double-sided cooling.

6.6 Conclusion

In this study, we investigated the steady-state and transient self-heating behavior of a homoepitaxial β-Ga_2O_3 MOSFET and various cooling options of the MOSFET. The effectiveness of top-sided, bottom-sided, and double-sided cooling schemes using a polycrystalline diamond substrate and a diamond passivation layer were studied via transient thermal modeling with realistic parameters. Because of the low thermal diffusivity of β-Ga_2O_3, the use of a β-Ga_2O_3 composite substrate (bottom-sided cooling) must be augmented by a diamond passivation layer (top-sided cooling) to effectively cool the device active region under both steady-state and transient operating conditions. Replacing the substrate with polycrystalline diamond (under a 6.5 µm-thick β-Ga_2O_3 layer) could reduce the steady-state temperature rise by 64% compared to that for a homoepitaxial

β-Ga_2O_3 MOSFET. However, for high frequency power switching applications beyond the ~10^2 kHz range, bottom-side cooling (integration with a high thermal conductivity substrate) does not improve the transient thermal response of the device. Adding a diamond passivation over layer diamond not only suppresses the steady-state temperature rise, but also drastically reduces the transient temperature rise under high frequency operating conditions. Both steady-state and transient temperatures could be reduced more with better interface of β-Ga_2O_3-diamond and having short time constant will be less affected to the multi-pulses.

CHAPTER 7. SUMMARY AND CONCLUSIONS

7.1 Summary of Contributions

Considering the range of WBG materials being studied for power electronics devices, beta-Gallium Oxide (β-Ga_2O_3) is a promising semiconductor material because of its unique combination of material properties – properties that make it excellent for power electronics applications: (i) low on-resistance (R_{on}), (ii) low off-state leakage current, (iii) large breakdown voltage (V_{Br}), and (iv) high-temperature operation. While the development of high power electronics continues to expand swiftly, it is becoming more apparent that Si based technologies are approaching or have reached the theoretical limit of the materials capabilities. However, heat generation under high-power operation cannot be avoided that overheating is a critical challenge for the reliability of these state-of-the-art device technologies. A smaller device footprint combined with a greater power handling capability means substantially increased power densities for individual devices.

A major drawback of β-Ga_2O_3 arises from its poor thermal conductivity, which results in devices with unacceptably high junction-to-package thermal resistance. While there is considerable promise for future devices made from UWBG materials, their adoption as a technology will hinge upon novel approaches to address heat dissipation at the die level which will enable high power density operation. Thus, this thesis provides valuable insight to the device community on how GaN based electronics are thermally impacted at the most basic device level.

The purpose and contribution of this work was to deliver the reader with a sufficient motivation on the current and potential uses of β-Ga_2O_3 based electronics through a summary of the current state of the technology. This was followed by a more in-depth discussion of β-Ga_2O_3 devices by looking specifically at the β-Ga_2O_3 lateral MOSFETs

and a vertical CAVET structure. The significance of β-Ga_2O_3 as used in a power electronics is emphasized in this work. A major emphasis of this work stems from the fact that typical thermal management techniques for WBG electronics such as integrating high thermal conductive substrate at the bottom of β-Ga_2O_3 substrate would not be sufficient for β-Ga_2O_3 devices due to very low thermal conductivity or highly resistive thermal layer.

Previous studies have focused on the design of active and passive cooling solutions that add upon or alter the homoepitaxial configuration of β-Ga_2O_3 transistors. In contrast, this work has focused on how to optimize the device layout to mitigate self-heating, prior to implementing such engineering solutions. For lateral β-Ga_2O_3 MOSFETs, the layout configuration was investigated including the anisotropic behavior of β-Ga_2O_3 substrate, spacing and geometry of metal contacts, and number of channels for multi-finger devices. The key finding is that due to the highly anisotropic thermal conductivity, it is important that the highest thermal conductivity orientation should be aligned with the cross-plane of β-Ga_2O_3 substrate. In addition, for the in-plane orientation, highest thermal conductivity should be aligned with the channel length so that the heat can be dissipated easily towards the metal contacts. These results were confirmed with single-finger MOSFET by experiment and simulation, and with multi-finger MOSFET by simulation.

One way to reduce the thermal resistance is to reduce the thickness of β-Ga_2O_3, thus the thermal resistivity of the device will be decrease. In vertical architectures, on the other hand, the β-Ga_2O_3 substrate cannot be substituted and reducing the thickness of β-Ga_2O_3 since it will change the electrical performance that the breakdown voltage of vertical transistors scales with the drift layer thickness. Therefore, the options for thermal management in vertical devices are limited compared to thermal management of lateral devices. In this work, several cooling scheme options were investigated to find the best way to reduce the junction-to-package thermal resistance and optimize for the vertical

device, especially for CAVET structure by adding external solution since there is limitation of engineering the device layout. Bottom-, top-, and double-sided cooling results show that the device temperature can be decreased by increasing the thermal conductivity of heat spreader, die attach material, and heat transfer coefficient. In addition, the model demonstrates the importance of top-sided cooling to alleviate the heat from the top-side of the vertical CAVET, which is closer to where joule heating occurs. Ultimately, architectures that utilize cooling from both sides of the device will perform best, enabling an 89% reduction in device temperature or 9× increased operational power at the same temperature limits.

Lastly, package- and system-level thermal management solutions, which are designed based on steady-state operation, were shown to be ineffective for applications that operate under transient thermal loading, potentially leading to overdesigned cooling systems. Especially due to low thermal conductivity of β-Ga_2O_3, low thermal diffusivity will cause much longer time constant compared to WBG materials. Therefore, thermal management transient-state operation should be differed from WBG power electronics. This work showed that for high frequency power switching applications beyond the ~10^2 kHz range, bottom-side cooling (integration with a high thermal conductivity substrate) does not improve the transient thermal response of the device. Adding a diamond passivation over layer diamond not only suppresses the steady-state temperature rise, but also drastically reduces the transient temperature rise under high frequency operating conditions. Both steady-state and transient temperatures could be reduced more with better interface of β-Ga_2O_3-diamond and having short time constant will be less affected to the multi-pulses.

7.2 Future Works

The β-Ga_2O_3 materials system presents promising aspects in terms of creating next-generation power electronic devices (i.e., low-cost substrate manufacturability and outstanding electronic properties). The material's low thermal conductivity resulting in overheating has become a major bottleneck to maximize the performance of β-Ga_2O_3 device technologies. A paradigm shift in the device design process, i.e., electro-thermal co-design, is essential to conquer the thermal obstructions. To implement such co-design techniques, the development of novel thermal characterization and multi-physics, multi-scale device modeling schemes are necessary. These innovations in convergent research will allow the full exploitation of the favorable benefits of the ultra-wide bandgap material.

With the success of integration of β-Ga_2O_3 on to high thermal conductivity diamond by either growth of diamond on β-Ga_2O_3, or low-temperature bonding of β-Ga_2O_3 to diamond, it is necessary to fabricate the β-Ga_2O_3 device with that configuration. Similar to DARPA NJTT program (i.e., comparison of GaN-on-SiC HEMT and GaN-on-diamond HEMT), as electro-thermal co-design of β-Ga_2O_3 approaches are available by this work, the experimental study will provide us a new perspective in terms of thermal management of UWBG power electronics.

Additional methods to cool high-power β-Ga_2O_3 devices will likely involve active cooling strategies. While air cooling is desired for low cost and high reliability strategies, active liquid cooling methods are expected to significantly enhance the operational power densities achieved by these devices and may find their position in high performance applications.

Lastly, regardless of all the cooling solutions, the importance of design and experimental study of β-Ga_2O_3 devices in transient-state cannot be emphasized enough as they are intended to be utilized for power switching applications. All the cooling solutions

should be verified its cooling ability during the high frequency application regime so that the solutions are not over-/under-designed for their demand. The applications of β-Ga_2O_3 electronics are not limited to power switching. From the perspective of device functionality, other prospective applications for β-Ga_2O_3 include high-temperature signal processing, harsh-environment electronics, and wireless communication devices/circuits. With respect to high-temperature and/or harsh-environment operation, it is important that β-Ga_2O_3 devices would be survivable under such conditions that proper thermal management is crucial.

REFERENCES

[1] K. O. Armstrong, S. Das, and J. Cresko, "Wide bandgap semiconductor opportunities in power electronics," in *WiPDA 2016 - 4th IEEE Workshop on Wide Bandgap Power Devices and Applications*, 2016, doi: 10.1109/WiPDA.2016.7799949.

[2] U. S. IEA, "US Energy Information Administration, International Energy Outlook 2017 Overview," *Int. Energy Outlook*, vol. IEO2017, no. 2017, p. 143, 2017, doi: www.eia.gov/forecasts/ieo/pdf/0484(2016).pdf.

[3] C. F. Heuberger and N. Mac Dowell, "Real-World Challenges with a Rapid Transition to 100% Renewable Power Systems," *Joule*, vol. 2, no. 3, pp. 367–370, 2018, doi: https://doi.org/10.1016/j.joule.2018.02.002.

[4] D. J. Hoffman *et al.*, "Concept design of high power solar electric propulsion vehicles for human exploration," *62nd Int. Astronaut. Congr. 2011, IAC 2011*, vol. 9, no. December, pp. 7463–7473, 2011.

[5] J. Y. Tsao *et al.*, "Ultrawide-Bandgap Semiconductors: Research Opportunities and Challenges," *Adv. Electron. Mater.*, vol. 4, no. 1, pp. 1–49, 2018, doi: 10.1002/aelm.201600501.

[6] M. Higashiwaki, K. Sasaki, A. Kuramata, T. Masui, and S. Yamakoshi, "Development of gallium oxide power devices," *Phys. Status Solidi Appl. Mater. Sci.*, vol. 211, no. 1, pp. 21–26, 2014, doi: 10.1002/pssa.201330197.

[7] S. J. Pearton, F. Ren, M. Tadjer, and J. Kim, "Perspective: Ga_2O_3 for ultra-high power rectifiers and MOSFETS," *J. Appl. Phys.*, vol. 124, no. 22, 2018, doi: 10.1063/1.5062841.

[8] M. Higashiwaki, H. Murakami, and Y. Kumagai, "State-of-the-art technologies of gallium oxide power devices Current status of Ga_2O_3 power devices," *J. Phys. D Appl. Phys. Masataka Higashiwaki al J. Phys. D Appl. Phys*, vol. 50, 2017.

[9] B. J. Baliga, "Power Semiconductor Device Figure of Merit for High-Frequency Applications," vol. 10, no. 10, pp. 455–457, 1989.

[10] B. Bayraktaroglu, "Assessment of Ga_2O_3Technology," *APL Mater.*, pp. 1–95, 2017.

[11] B. Kaplar and S. N. Labs, "Ultra-Wide-Bandgap Semiconductors : Outline - Applications Motivating the Development of UWBG Semiconductors - Impact of UWBG Semiconductor Properties on Devices Considerations," 2019.

[12] S. B. Reese, T. Remo, J. Green, and A. Zakutayev, "How Much Will Gallium Oxide Power Electronics Cost?," *Joule*, pp. 1–5, 2019, doi: 10.1016/j.joule.2019.01.011.

[13] K. Sasaki, Q. T. Thieu, D. Wakimoto, Y. Koishikawa, A. Kuramata, and S. Yamakoshi, "Depletion-mode vertical Ga_2O_3 trench MOSFETs fabricated using Ga_2O_3 homoepitaxial films grown by halide vapor phase epitaxy," *Appl. Phys. Express*, vol. 10, no. 12, p. 124201, 2017, doi: 10.7567/APEX.10.124201.

[14] J. Yang, F. Ren, M. Tadjer, S. J. Pearton, and A. Kuramata, "2300V Reverse Breakdown Voltage Ga_2O_3 Schottky Rectifiers," *ECS J. Solid State Sci. Technol.*, vol. 7, no. 5, pp. Q92–Q96, 2018, doi: 10.1149/2.0241805jss.

[15] F. G. O. Mosfets, K. Zeng, A. Vaidya, U. Singisetti, and S. Member, "1.85 kV Breakdown Voltage in Lateral," vol. 39, no. 9, pp. 2018–2021, 2018.

[16] M. Higashiwaki, K. Sasaki, A. Kuramata, T. Masui, and S. Yamakoshi, "Gallium oxide (Ga_2O_3) metal-semiconductor field-effect transistors on single-crystal β-Ga_2O_3 (010) substrates," *Appl. Phys. Lett.*, vol. 100, no. 1, pp. 1–4, 2012, doi: 10.1063/1.3674287.

[17] M. J. Tadjer *et al.*, " Editors' Choice Communication—A (001) β-Ga_2O_3 MOSFET with +2.9 V Threshold Voltage and HfO_2 Gate Dielectric ," *ECS J. Solid State Sci. Technol.*, vol. 5, no. 9, pp. P468–P470, 2016, doi: 10.1149/2.0061609jss.

[18] M. H. Wong, Y. Nakata, A. Kuramata, S. Yamakoshi, and M. Higashiwaki, "Enhancement-mode Ga_2O_3 MOSFETs with Si-ion-implanted source and drain," *Appl. Phys. Express*, vol. 10, no. 4, p. 041101, 2017, doi: 10.7567/APEX.10.041101.

[19] K. D. Chabak *et al.*, "Enhancement-mode Ga_2O_3 wrap-gate fin field-effect transistors on native (100) β -Ga_2O_3 substrate with high breakdown voltage," *Appl. Phys. Lett.*, vol. 109, no. 21, p. 213501, 2016, doi: 10.1063/1.4967931.

[20] H. Dong *et al.*, "Progress of power field effect transistor based on ultra-wide

bandgap Ga_2O_3 semiconductor material," *J. Semicond.*, vol. 40, no. 1, p. 011802, Jan. 2019, doi: 10.1088/1674-4926/40/1/011802.

[21] S. J. Pearton *et al.*, " A review of Ga_2O_3 materials, processing, and devices ," *Appl. Phys. Rev.*, vol. 5, no. 1, p. 011301, 2018, doi: 10.1063/1.5006941.

[22] M. H. Wong, K. Sasaki, A. Kuramata, S. Yamakoshi, and M. Higashiwaki, "Field-Plated Ga_2O_3MOSFETs With a Breakdown Voltage of Over 750 V," *IEEE Electron Device Lett.*, vol. 37, no. 2, pp. 212–215, 2016, doi: 10.1109/LED.2015.2512279.

[23] K. D. Chabak *et al.*, "Recessed-Gate Enhancement-Mode β-Ga_2O_3MOSFETs," *IEEE Electron Device Lett.*, vol. 39, no. 1, pp. 67–70, 2018, doi: 10.1109/LED.2017.2779867.

[24] K. Zeng, A. Vaidya, and U. Singisetti, "1.85 kV Breakdown Voltage in Lateral Field-Plated Ga_2O_3 MOSFETs," *IEEE Electron Device Lett.*, vol. 39, no. 9, pp. 1385–1388, 2018, doi: 10.1109/LED.2018.2859049.

[25] O. Hilt *et al.*, "Lateral and vertical power transistors in GaN and Ga_2O_3," *IET Power Electron.*, vol. 12, no. 15, pp. 3919–3927, 2019, doi: https://doi.org/10.1049/iet-pel.2019.0059.

[26] M. H. Wong, K. Goto, H. Murakami, and Y. Kumagai, "Current Aperture Vertical β -Ga_2O_3 MOSFETs Fabricated by N- and Si-Ion Implantation Doping," vol. 40, no. 3, pp. 431–434, 2019, doi: 10.1109/LED.2018.2884542.

[27] Z. Hu *et al.*, "Enhancement-Mode Ga_2O_3 Vertical Transistors With Breakdown Voltage > 1 kV," vol. 39, no. 6, pp. 2018–2021, 2018.

[28] D. J. Cheney *et al.*, "Reliability studies of AlGaN/GaN high electron mobility transistors," *Semicond. Sci. Technol.*, vol. 28, no. 7, 2013, doi: 10.1088/0268-1242/28/7/074019.

[29] S. Mukherjee, Y. Puzyrev, J. Chen, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, "Hot-Carrier Degradation in GaN HEMTs Due to Substitutional Iron and Its Complexes," *IEEE Trans. Electron Devices*, vol. 63, no. 4, pp. 1486–1494, 2016, doi: 10.1109/TED.2016.2532806.

[30] C. Suckling and D. Nguyen, "Thermal Analysis of GaN Devices," *ARMMS Conf.*

RF Microw. Soc., 2012.

[31] J. Falck, C. Felgemacher, A. Rojko, M. Liserre, and P. Zacharias, "Reliability of Power Electronic Systems," *IEEE Ind. Electron. Mag.*, vol. 12, no. 2, pp. 24–35, 2018, doi: 10.1109/MIE.2018.2825481.

[32] S. Yang, A. Bryant, P. Mawby, D. Xiang, L. Ran, and P. Tavner, "An industry-based survey of reliability in power electronic converters," *IEEE Trans. Ind. Appl.*, vol. 47, no. 3, pp. 1441–1451, 2011, doi: 10.1109/TIA.2011.2124436.

[33] D. Altman *et al.*, "Analysis and characterization of thermal transport in GaN HEMTs on Diamond substrates," *Thermomechanical Phenom. Electron. Syst. - Proceedings Intersoc. Conf.*, pp. 1199–1205, 2014, doi: 10.1109/ITHERM.2014.6892416.

[34] D. Liu *et al.*, "Impact of diamond seeding on the microstructural properties and thermal stability of GaN-on-diamond wafers for high-power electronic devices," *Scr. Mater.*, vol. 128, pp. 57–60, 2017, doi: 10.1016/j.scriptamat.2016.10.006.

[35] L. Yates *et al.*, "Low Thermal Boundary Resistance Interfaces for GaN-on-Diamond Devices," *ACS Appl. Mater. Interfaces*, vol. 10, no. 28, pp. 24302–24309, 2018, doi: 10.1021/acsami.8b07014.

[36] Z. Cheng, F. Mu, L. Yates, T. Suga, and S. Graham, "Interfacial Thermal Conductance across Room-Temperature-Bonded GaN/Diamond Interfaces for GaN-on-Diamond Devices," *ACS Appl. Mater. Interfaces*, vol. 12, no. 7, pp. 8376–8384, 2020, doi: 10.1021/acsami.9b16959.

[37] M. Malakoutian *et al.*, "Record-Low Thermal Boundary Resistance between Diamond and GaN-on-SiC for Enabling Radiofrequency Device Cooling," *ACS Appl. Mater. & Interfaces*, vol. 13, no. 50, pp. 60553–60560, Dec. 2021, doi: 10.1021/acsami.1c13833.

[38] S. H. Kim *et al.*, "Thermally-Aware Layout Design of β-Ga_2O_3 Lateral MOSFETs," *IEEE Trans. Electron Devices*, vol. 69, no. 3, pp. 1251–1257, 2022, doi: 10.1109/TED.2022.3143779.

[39] Z. Cheng, J. Shi, C. Yuan, S. Kim, and S. Graham, "Chapter Four - Thermal science and engineering of β- Ga_2O_3 materials and devices," in *Ultrawide Bandgap Semiconductors*, vol. 107, Y. Zhao and Z. Mi, Eds. Elsevier, 2021, pp. 77–99.

[40] P. Jiang, X. Qian, X. Li, and R. Yang, "Three-dimensional anisotropic thermal conductivity tensor of single crystalline β-Ga_2O_3," *Appl. Phys. Lett.*, vol. 113, no. 23, pp. 8–13, 2018, doi: 10.1063/1.5054573.

[41] Z. Guo *et al.*, "Anisotropic thermal conductivity in single crystal β-gallium oxide," *Appl. Phys. Lett.*, vol. 106, no. 11, pp. 1–6, 2015, doi: 10.1063/1.4916078.

[42] A. Bhattacharyya, S. Roy, P. Ranga, D. Shoemaker, and Y. Song, "Temperature MOVPE-Regrown Ohmic Contacts," pp. 1–5.

[43] H. T. Aller *et al.*, "Chemical Reactions Impede Thermal Transport across Metal/β-Ga_2O_3 Interfaces," *Nano Lett.*, vol. 19, no. 12, pp. 8533–8538, 2019, doi: 10.1021/acs.nanolett.9b03017.

[44] L. A. M. Lyle *et al.*, " Electrical and chemical analysis of Ti/Au contacts to β-Ga_2O_3 ," *APL Mater.*, vol. 9, no. 6, p. 061104, 2021, doi: 10.1063/5.0051340.

[45] J. Shi *et al.*, "Thermal Transport across Metal/ β-Ga_2O_3 Interfaces," 2021, doi: 10.1021/acsami.1c05191.

[46] J. Yang, F. Ren, S. J. Pearton, and A. Kuramata, "Vertical Geometry, 2-A Forward Current Ga_2O_3 Schottky Rectifiers on Bulk Ga_2O_3 Substrates," *IEEE Trans. Electron Devices*, vol. 65, no. 7, pp. 2790–2796, 2018, doi: 10.1109/TED.2018.2838439.

[47] M. H. Wong *et al.*, "First Demonstration of Vertical Ga_2O_3 MOSFET : Planar Structure with a Current Aperture," vol. 95, no. 2004, pp. 4–5, 2016.

[48] M. H. Wong, "Vertical β-Ga_2O_3 Power Transistors: A Review," *IEEE Electron Device Lett.*, vol. 67, no. 10, 2020, doi: 10.1109/led.2019.2926202.

[49] N. R. Jankowski and F. P. McCluskey, "Modeling transient thermal response of pulsed power electronic packages," *PPC2009 - 17th IEEE Int. Pulsed Power Conf.*, pp. 820–825, 2009, doi: 10.1109/PPC.2009.5386368.

[50] L. Yates, J. Shi, M. J. Tadjer, K. D. Hobart, and S. Graham, "Thermal conductance across-Ga_2O_3-diamond van der Waals heterogeneous interfaces," vol. 031118, no. March, 2019, doi: 10.1063/1.5089559.

[51] Z. Cheng *et al.*, "Integration of polycrystalline Ga_2O_3 on diamond for thermal management," *Appl. Phys. Lett.*, vol. 116, no. 6, 2020, doi: 10.1063/1.5125637.

[52] T. Matsumae *et al.*, "Low-temperature direct bonding of β-Ga_2O_3 and diamond substrates under atmospheric conditions," *Appl. Phys. Lett.*, vol. 116, no. 14, pp. 2–6, 2020, doi: 10.1063/5.0002068.

[53] M. R. Karim *et al.*, "Two-step growth of β-Ga_2O_3 films on (100) diamond via low pressure chemical vapor deposition," *J. Vac. Sci. \& Technol. A*, vol. 39, no. 2, p. 23411, 2021, doi: 10.1116/6.0000854.

[54] Y. Song *et al.*, "Ga_2O_3-on-SiC Composite Wafer for Thermal Management of Ultrawide Bandgap Electronics," *ACS Appl. Mater. Interfaces*, vol. 13, no. 34, pp. 40817–40829, 2021, doi: 10.1021/acsami.1c09736.

[55] T.-H. Hung *et al.*, "Energy band line-up of atomic layer deposited Al_2O_3 on β-Ga_2O_3," *Appl. Phys. Lett.*, vol. 104, no. 16, p. 162106, 2014, doi: 10.1063/1.4873546.

[56] A. J. Green *et al.*, "β-Ga_2O_3 MOSFETs for Radio Frequency Operation," *IEEE Electron Device Lett.*, vol. 38, no. 6, pp. 790–793, 2017, doi: 10.1109/LED.2017.2694805.

[57] N. Moser *et al.*, "Ge-Doped β-Ga_2O_3 MOSFETs," *IEEE Electron Device Lett.*, vol. 38, no. 6, pp. 775–778, 2017, doi: 10.1109/LED.2017.2697359.

[58] P. H. Carey, J. Yang, F. Ren, R. Sharma, M. Law, and S. J. Pearton, "Comparison of Dual-Stack Dielectric Field Plates on β-Ga_2O_3 Schottky Rectifiers," *{ECS} J. Solid State Sci. Technol.*, vol. 8, no. 7, pp. Q3221--Q3225, 2019, doi: 10.1149/2.0391907jss.

[59] Z. Xia *et al.*, "Metal/$BaTiO_3$/β-Ga_2O_3 dielectric heterojunction diode with 5.7 MV/cm breakdown field," *Appl. Phys. Lett.*, vol. 115, no. 25, p. 252104, 2019, doi: 10.1063/1.5130669.

[60] S. Roy, A. Bhattacharyya, P. Ranga, H. Splawn, J. Leach, and S. Krishnamoorthy, "High-k Oxide Field-Plated Vertical (001) β-Ga_2O_3 Schottky Barrier Diode With Baliga's Figure of Merit Over 1 GW/cm^2," *IEEE Electron Device Lett.*, vol. 42, no. 8, pp. 1140–1143, 2021, doi: 10.1109/LED.2021.3089945.

[61] K. D. Chabak *et al.*, "Enhancement-mode Ga_2O_3 wrap-gate fin field-effect transistors on native (100) β-Ga2O3 substrate with high breakdown voltage," *Appl. Phys. Lett.*, vol. 109, no. 21, p. 213501, 2016, doi: 10.1063/1.4967931.

[62] A. J. Green *et al.*, "3.8-MV/cm Breakdown Strength of MOVPE-Grown Sn-Doped β-Ga_2O_3 MOSFETs," *IEEE Electron Device Lett.*, vol. 37, no. 7, pp. 902–905, 2016, doi: 10.1109/LED.2016.2568139.

[63] W. Li, K. Nomoto, Z. Hu, D. Jena, and H. G. Xing, "Field-Plated Ga_2O_3 Trench Schottky Barrier Diodes With a $V_B^2/R_{on,sp}$ of up to 0.95 GW/cm^2," *IEEE Electron Device Lett.*, vol. 41, no. 1, pp. 107–110, 2020, doi: 10.1109/LED.2019.2953559.

[64] W. Li, K. Nomoto, Z. Hu, T. Nakamura, D. Jena, and H. G. Xing, "Single and multi-fin normally-off Ga_2O_3 vertical transistors with a breakdown voltage over 2.6 kV," in *2019 IEEE International Electron Devices Meeting (IEDM)*, 2019, pp. 12.4.1-12.4.4, doi: 10.1109/IEDM19573.2019.8993526.

[65] N. Tanaka, Y. Sumida, H. Kawai, and T. Suzuki, "Delay Time Analysis of AlGaN/GaN Heterojunction Field-Effect Transistors with AlN or SiN Surface Passivation," *Jpn. J. Appl. Phys.*, vol. 48, no. 4, p. 04C099, Apr. 2009, doi: 10.1143/jjap.48.04c099.

[66] N. Tsurumi *et al.*, "AlN Passivation Over AlGaN/GaN HFETs for Surface Heat Spreading," *IEEE Trans. Electron Devices*, vol. 57, no. 5, pp. 980–985, 2010, doi: 10.1109/TED.2010.2044675.

[67] J.-X. Chen *et al.*, "Band alignment of AlN/β-Ga_2O_3heterojunction interface measured by x-ray photoelectron spectroscopy," *Appl. Phys. Lett.*, vol. 112, no. 26, p. 261602, 2018, doi: 10.1063/1.5035372.

[68] M. Malakoutian *et al.*, "Polycrystalline diamond growth on β- Ga_2O_3 for thermal management," *Appl. Phys. Express*, vol. 14, no. 5, 2021, doi: 10.35848/1882-0786/abf4f1.

[69] Y. Zhou *et al.*, "Thermal characterization of polycrystalline diamond thin film heat spreaders grown on GaN HEMTs," *Appl. Phys. Lett.*, vol. 111, no. 4, 2017, doi: 10.1063/1.4995407.

[70] K. R. Bagnall, Y. S. Muzychka, and E. N. Wang, "Analytical solution for temperature rise in complex multilayer structures with discrete heat sources," *IEEE*

Trans. Components, Packag. Manuf. Technol., vol. 4, no. 5, pp. 817–830, 2014, doi: 10.1109/TCPMT.2014.2299766.

[71] B. Chatterjee *et al.*, "Device-Level Thermal Management of Gallium Oxide Field-Effect Transistors," vol. 9, no. 12, pp. 2352–2365, 2019.

[72] C. Yuan *et al.*, "Modeling and analysis for thermal management in gallium oxide field-effect transistors," *J. Appl. Phys.*, vol. 127, no. 15, p. 154502, Apr. 2020, doi: 10.1063/1.5141332.

[73] T. Kim, S. Il Park, C. Song, H. Lee, and J. Cho, "Fundamental conduction cooling limits for sub-1 μm Ga_2O_3 devices integrated with diamond," *Int. J. Heat Mass Transf.*, vol. 191, p. 122864, 2022, doi: https://doi.org/10.1016/j.ijheatmasstransfer.2022.122864.

[74] P. W. Webb, "Thermal imaging of electronic devices with low surface emissivity," *IEE Proc. G (Circuits, Devices Syst.*, vol. 138, no. 3, pp. 390-400(10), Jun. 1991, [Online]. Available: https://digital-library.theiet.org/content/journals/10.1049/ip-g-2.1991.0065.

[75] S. H. Kim, "ADDRESSING THERMAL AND ENVIRONMENTAL RELIABILITY IN GAN BASED by," 2014.

[76] N. Lundt *et al.*, "High spatial resolution Raman thermometry analysis of TiO_2 microparticles," *Rev. Sci. Instrum.*, vol. 84, no. 10, 2013, doi: 10.1063/1.4824355.

[77] R. B. Simon, J. W. Pomeroy, and M. Kuball, "Diamond micro-Raman thermometers for accurate gate temperature measurements," *Appl. Phys. Lett.*, vol. 104, no. 21, p. 213503, 2014, doi: 10.1063/1.4879849.

[78] J. Dallas *et al.*, "Thermal characterization of gallium nitride p-i-n diodes," *Appl. Phys. Lett.*, vol. 112, no. 7, pp. 1–5, 2018, doi: 10.1063/1.5006796.

[79] M. J. Šćepanović, M. Grujić-Brojčin, Z. D. Dohčević-Mitrović, and Z. V. Popović, "Characterization of anatase TiO_2 nanopowder by variable-temperature raman spectroscopy," *Sci. Sinter.*, vol. 41, no. 1, pp. 67–73, 2009, doi: 10.2298/SOS0901067S.

[80] D. Kendig, A. Tay, and A. Shakouri, "Thermal analysis of advanced microelectronic

devices using thermoreflectance thermography," in *2016 22nd International Workshop on Thermal Investigations of ICs and Systems (THERMINIC)*, 2016, pp. 115–120, doi: 10.1109/THERMINIC.2016.7749037.

[81] M. Kuball *et al.*, "Time-Resolved Temperature Measurement of AlGaN/GaN Electronic Devices Using Micro-Raman Spectroscopy," *IEEE Electron Device Lett.*, vol. 28, no. 2, pp. 86–89, 2007, doi: 10.1109/LED.2006.889215.

[82] G. J. Riedel *et al.*, "Nanosecond Timescale Thermal Dynamics of AlGaN/GaN Electronic Devices," *IEEE Electron Device Lett.*, vol. 29, no. 5, pp. 416–418, 2008, doi: 10.1109/LED.2008.919779.

[83] Z. Cheng *et al.*, "Integration of polycrystalline Ga_2O_3 on diamond for thermal management," *Appl. Phys. Lett.*, vol. 116, no. 6, p. 062105, Feb. 2020, doi: 10.1063/1.5125637.

[84] M. J. Tadjer, "Cheap Ultra-Wide Bandgap Power Electronics? Gallium Oxide May Hold the Answer," *Electrochem. Soc. Interface*, vol. 27, no. 4, pp. 49–52, 2018, doi: 10.1149/2.f05184if.

[85] J. Millman, *Electronic Devices and Circuits*. McGraw Hill, 1967.

[86] K. J. Liddy *et al.*, "Thin channel β-Ga_2O_3MOSFETs with self-aligned refractory metal gates," *Appl. Phys. Express*, vol. 12, no. 12, pp. 0–4, 2019, doi: 10.7567/1882-0786/ab4d1c.

[87] S. Choi, E. R. Heller, D. Dorsey, R. Vetury, and S. Graham, "The impact of bias conditions on self-heating in AlGaN/GaN HEMTs," *IEEE Trans. Electron Devices*, vol. 60, no. 1, pp. 159–162, 2013, doi: 10.1109/TED.2012.2224115.

[88] A. Bar-Cohen, J. D. Albrecht, and J. J. Maurer, "Near-junction thermal management for wide bandgap devices," *Tech. Dig. - IEEE Compd. Semicond. Integr. Circuit Symp. CSIC*, pp. 10–14, 2011, doi: 10.1109/CSICS.2011.6062454.

[89] X. Chen, F. N. Donmezer, S. Kumar, and S. Graham, "A numerical study on comparing the active and passive cooling of AlGaN/GaN HEMTs," *IEEE Trans. Electron Devices*, vol. 61, no. 12, pp. 4056–4061, 2014, doi: 10.1109/TED.2014.2360504.

[90] S. Khanna, P. McCluskey, A. Bar-Cohen, B. Yang, and M. Ohadi, "Thin thermally efficient ICECool defense semiconductor power amplifiers," *J. Microelectron. Electron. Packag.*, vol. 14, no. 3, pp. 77–93, 2017, doi: 10.4071/imaps.456518.

[91] B. Chatterjee *et al.*, "Nanoscale electro-thermal interactions in AlGaN/GaN high electron mobility transistors," *J. Appl. Phys.*, vol. 127, no. 4, 2020, doi: 10.1063/1.5123726.

[92] D. Li *et al.*, "AlGaN/GaN MISHEMTs with AlN gate dielectric grown by thermal ALD technique," *Nanoscale Res. Lett.*, vol. 10, p. 109, 2015, doi: 10.1186/s11671-015-0802-x.

[93] G. Dechun, Q. Kankan, C. Junfeng, L. Xiaobin, and Y. Chao, "A simulation about the influence of the gate-source-drain distance on the AlGaN/GaN HEMT performance at Ka-band," in *2012 IEEE MTT-S International Microwave Workshop Series on Millimeter Wave Wireless Technology and Applications*, 2012, pp. 1–4, doi: 10.1109/IMWS2.2012.6338225.

[94] S. Kim *et al.*, "Thermal Management of β- Ga_2O_3 Current Aperture Vertical Electron Transistors," *IEEE Trans. Components, Packag. Manuf. Technol.*, vol. 11, no. 8, pp. 1171–1176, 2021, doi: 10.1109/TCPMT.2021.3089321.

[95] H. Zhou, K. Maize, G. Qiu, A. Shakouri, and P. D. Ye, "β- Ga_2O_3insulator field-effect transistors with drain currents exceeding 1.5 A/mm and their self-heating effect," *Appl. Phys. Lett.*, vol. 111, no. 9, 2017, doi: 10.1063/1.5000735.

[96] Z. Hu *et al.*, "1.6 kV Vertical Ga_2O_3 FinFETs With Source- Connected Field Plates and Normally-off Operation," *2019 31st Int. Symp. Power Semicond. Devices ICs*, vol. 3, no. 001, pp. 483–486, 2019, doi: 10.1109/ISPSD.2019.8757633.

[97] J. K. Mun, K. Cho, W. Chang, H.-W. Jung, and J. Do, " Editors' Choice—2.32 kV Breakdown Voltage Lateral β-Ga_2O_3 MOSFETs with Source-Connected Field Plate ," *ECS J. Solid State Sci. Technol.*, vol. 8, no. 7, pp. Q3079–Q3082, 2019, doi: 10.1149/2.0151907jss.

[98] A. Phys, "Breakdown mechanism in 1 kA/cm^2 and 960 V E-mode β-Ga_2O_3 vertical transistors," vol. 122103, no. September, pp. 3–8, 2018, doi: 10.1063/1.5038105.

[99] T. Kimoto, "Material science and device physics in SiC technology for high-voltage power devices," *Jpn. J. Appl. Phys.*, vol. 54, no. 4, p. 040103, Apr. 2015, doi:

10.7567/JJAP.54.040103.

[100] J. N. Shenoy, S. Member, J. A. Cooper, M. R. Melloch, and S. Member, "High-Voltage Double-Implanted Power MOSFET ' s in 6H-SiC," vol. 18, no. 3, pp. 93–95, 1997.

[101] M. Kanechika *et al.*, "A vertical insulated gate AlGaN/GaN heterojunction field-effect transistor," *Japanese J. Appl. Physics, Part 2 Lett.*, vol. 46, no. 20–24, pp. 4–7, 2007, doi: 10.1143/JJAP.46.L503.

[102] H. Nie *et al.*, "1.5 kV Vertical GaN Transistors on Bulk-GaN Substrates," vol. 35, no. 9, pp. 939–941, 2014.

[103] S. Chowdhury, B. L. Swenson, M. H. Wong, and U. K. Mishra, "Current status and scope of gallium nitride-based vertical transistors for high-power electronics application," *Semicond. Sci. Technol.*, vol. 28, no. 7, 2013, doi: 10.1088/0268-1242/28/7/074014.

[104] J. W. Pomeroy *et al.*, "Raman Thermography of Peak Channel Temperature in β-Ga_2O_3 MOSFETs," vol. 40, no. 2, pp. 2018–2021, 2019.

[105] M. H. Wong *et al.*, "First demonstration of vertical Ga_2O_3 MOSFET: Planar structure with a current aperture," in *2017 75th Annual Device Research Conference (DRC)*, 2017, pp. 1–2, doi: 10.1109/DRC.2017.7999413.

[106] Z. Galazka *et al.*, "On the bulk β-Ga_2O_3 single crystals grown by the Czochralski method," *J. Cryst. Growth*, vol. 404, pp. 184–191, 2014, doi: https://doi.org/10.1016/j.jcrysgro.2014.07.021.

[107] T. Gake, Y. Kumagai, and F. Oba, "First-principles study of self-trapped holes and acceptor impurities in Ga_2O_3 polymorphs," *Phys. Rev. Mater.*, vol. 3, no. 4, p. 44603, Apr. 2019, doi: 10.1103/PhysRevMaterials.3.044603.

[108] J. L. Lyons, "A survey of acceptor dopants for Ga_2O_3," *Semicond. Sci. Technol.*, vol. 33, no. 5, p. 05LT02, Apr. 2018, doi: 10.1088/1361-6641/aaba98.

[109] M. Higashiwaki *et al.*, "Anisotropic thermal conductivity in single crystal β-gallium oxide," *Appl. Phys. Lett.*, vol. 106, no. 11, p. 111909, 2015, doi: 10.1063/1.4916078.

[110] J. A. Phys, “Anisotropic thermal conductivity of β- Ga_2O_3 at elevated temperatures : Effect of Sn and Fe dopants,” vol. 235104, no. April 2017, 2018, doi: 10.1063/1.4986478.

[111] M. Li, Y. Xiao, Z. Zhang, and J. Yu, “Bimodal Sintered Silver Nanoparticle Paste with Ultrahigh Thermal Conductivity and Shear Strength for High Temperature Thermal Interface Material Applications,” 2015, doi: 10.1021/acsami.5b01341.

[112] J. Kuang, M. Sheen, C. H. Chang, C. Chen, and G. Wang, “Effect of Temperature Cycling on Joint Strength of PbSn and AuSn Solders in Laser Packages,” vol. 24, no. 4, pp. 563–568, 2001.

[113] D. G. Pahinkar *et al.*, “Transient Liquid Phase Bonding of AlN to AlSiC for Durable Power Electronic Packages,” vol. 1800039, pp. 1–9, 2018, doi: 10.1002/adem.201800039.

[114] M. E. Liao *et al.*, “Coefficients of thermal expansion of single crystalline β-Ga_2O_3 and in-plane thermal strain calculations of various materials combinations with β-Ga_2O_3,” *APL Mater.*, vol. 7, no. 2, 2019, doi: 10.1063/1.5054327.

[115] C. Yuan, B. Duan, L. Li, B. Xie, M. Huang, and X. Luo, “Thermal Conductivity of Polymer-Based Composites with Magnetic Aligned Hexagonal Boron Nitride Platelets,” pp. 1–7, 2015, doi: 10.1021/acsami.5b03007.

[116] C. Yu *et al.*, “Composites : Part A Enhanced through-plane thermal conductivity of boron nitride / epoxy composites,” *Compos. Part A*, vol. 98, pp. 25–31, 2017, doi: 10.1016/j.compositesa.2017.03.012.

[117] T. Morishita and H. Okamoto, “Facile Exfoliation and Noncovalent Superacid Functionalization of Boron Nitride Nanosheets and Their Use for Highly Thermally Conductive and Electrically Insulating Polymer Nanocomposites,” pp. 2–11, 2016, doi: 10.1021/acsami.6b08404.

[118] Y. Lin, T. V. Williams, and J. W. Connell, “Soluble, exfoliated hexagonal boron nitride nanosheets,” *J. Phys. Chem. Lett.*, vol. 1, no. 1, pp. 277–283, 2010, doi: 10.1021/jz9002108.

[119] K. R. Bagnall and E. N. Wang, “Theory of Thermal Time Constants in GaN High-Electron-Mobility Transistors,” doi: 10.1109/TCPMT.2017.2773065.

[120] K. R. Bagnall, O. I. Saadat, S. Joglekar, T. Palacios, and E. N. Wang, "Experimental Characterization of the Thermal Time Constants of GaN HEMTs Via Micro-Raman Thermometry," *IEEE Trans. Electron Devices*, vol. 64, no. 5, pp. 2121–2128, 2017, doi: 10.1109/TED.2017.2679978.

[121] J. Dallas *et al.*, "Thermal characterization of gallium nitride p-i-n diodes," *Appl. Phys. Lett.*, vol. 112, no. 7, 2018, doi: 10.1063/1.5006796.

[122] A. Bar-Cohen, J. J. Maurer, and D. H. Altman, "Embedded cooling for wide bandgap power amplifiers: A review," *J. Electron. Packag. Trans. ASME*, vol. 141, no. 4, pp. 1–14, 2019, doi: 10.1115/1.4043404.

[123] J. Anaya *et al.*, "Simultaneous determination of the lattice thermal conductivity and grain/grain thermal resistance in polycrystalline diamond," *Acta Mater.*, vol. 139, pp. 215–225, 2017, doi: 10.1016/j.actamat.2017.08.007.

[124] C. Song, J. Kim, and J. Cho, "The effect of GaN epilayer thickness on the near-junction thermal resistance of GaN-on-diamond devices," *Int. J. Heat Mass Transf.*, vol. 158, p. 119992, 2020, doi: 10.1016/j.ijheatmasstransfer.2020.119992.

[125] A. A. Allerman *et al.*, "Size dictated thermal conductivity of GaN," *J. Appl. Phys.*, vol. 120, no. 9, p. 095104, 2016, doi: 10.1063/1.4962010.

[126] D. Shoemaker *et al.*, "Diamond-Incorporated Flip-Chip Integration for Thermal Management of GaN and Ultra-Wide Bandgap RF Power Amplifiers," *IEEE Trans. Components, Packag. Manuf. Technol.*, vol. 11, no. 8, pp. 1177–1186, 2021, doi: 10.1109/TCPMT.2021.3091555.

[127] E. Ziade, J. Yang, G. Brummer, D. Nothern, T. Moustakas, and A. J. Schmidt, "Thermal transport through GaN-SiC interfaces from 300 to 600 K," *Appl. Phys. Lett.*, 2015, doi: 10.1063/1.4930104.

[128] B. Chatterjee *et al.*, "Electro-thermal co-design of β-$(Al_xGa_{1-x})_2O_3/Ga_2O_3$ modulation doped field effect transistors," *Appl. Phys. Lett.*, vol. 117, no. 15, 2020, doi: 10.1063/5.0021275.

[129] B. Chatterjee, J. S. Lundh, J. Dallas, H. Kim, and S. Choi, "Electro-thermal reliability study of GaN high electron mobility transistors," *Proc. 16th Intersoc. Conf. Therm. Thermomechanical Phenom. Electron. Syst. ITherm 2017*, pp. 1247–1252, 2017, doi: 10.1109/ITHERM.2017.7992627.

[130] E. Heller, S. Choi, D. Dorsey, R. Vetury, and S. Graham, “Electrical and structural dependence of operating temperature of AlGaN/GaN HEMTs,” *Microelectron. Reliab.*, vol. 53, no. 6, pp. 872–877, 2013, doi: 10.1016/j.microrel.2013.03.004.

[131] H. P. de Bock *et al.*, “A System to Package Perspective on Transient Thermal Management of Electronics,” *J. Electron. Packag.*, vol. 142, no. 4, pp. 1–11, 2020, doi: 10.1115/1.4047474.

[132] X. Guo, Q. Xun, Z. Li, and S. Du, “Silicon Carbide Converters and MEMS Devices for High-temperature Power Electronics: A Critical Review,” *Micromachines*, vol. 10, no. 6, 2019, doi: 10.3390/mi10060406.