Thermal Conductivity of Power Semiconductors—When Does It Matter?

L. Boteler[1,a], A. Lelis[1], M. Berman[1], and M. Fish[1]
[1] U.S. Army Research Laboratory, Adelphi, MD 20783, USA, [a] lauren.m.boteler.civ@mail.mil

***Abstract*— A study comparing the interaction between wide bandgap (WBG) materials, chip thickness, and heat sink selection on the total temperature rise in a standard power electronics package was performed. The thermal calculations, using the ARL ParaPower tool, showed that chips with thermal conductivities of less than 150 W/mK (including β-Ga_2O_3) had a dominant effect on total temperature rise, whereas thermal conductivities greater than 400 W/mK (including SiC and diamond) had little impact. In designs with devices having thermal conductivities between 150 and 400 W/mK (including GaN) the temperature rise resulting from the chip is a significant fraction of the total rise for heatsinks with heat transfer coefficients of 50,000 W/m^{2}K or higher. For chips with thermal conductivities less than Si, the thickness of the chip can be a significant factor affecting the overall temperature rise.**

***Keywords*—Thermal management, power electronics packaging, ARL ParaPower, gallium oxide, diamond, silicon carbide, gallium, ultra wide bandgap**

I. INTRODUCTION

Wide bandgap (WBG) and ultra wide bandgap (UWBG) semiconductors are excellent candidates for expanding the power, voltage, frequency and temperature range of power devices compared to standard silicon devices [1]. In fact, WBG silicon carbide (SiC) power devices have long been successfully commercialized, and in more recent years WBG GaN power devices are finding their way into the power device market as well. Most recently, UWBG materials such as β-Ga_2O_3, high Al-content AlGaN, and diamond are gaining increased interest as possible candidates for even higher power density applications. [1]

For many power electronics applications, thermal effects are now limiting the performance of the device; and the methods of thermal management, including the packaging and heat sink used, are driving the size and weight of the system. [2] Given the wide range of thermal conductivity values for various WBG and UWBG materials, it is useful to consider under which circumstances these differences are critical in designing appropriate thermal-management solutions. Analysis tools able to perform quick, reduced order modeling are critical to design-space optimization. The ARL ParaPower tool was previously developed for this purpose [3], and is used in this work to quickly and efficiently perform the thermal analysis.

In order to determine the thermal effects of UWBG devices, this paper uses a standard power-electronics module configuration and simply replaces the existing devices in the module with various device materials and thicknesses. Determining the appropriate heat sink for an application is critical to reducing overdesign and minimizing operation in a temperature sensitive regime wherein small changes in the heat transfer coefficient, h, can lead to large increases in the temperature rise between the package and the outside ambient. This paper will investigate the impact on device temperature rise based on the semiconductor material choice, the package, and the heat sink. As WBG material options continue to increase, it is critical to understand how each option affects the overall thermal performance to ensure the proper heat sink selection, and package design.

II. BACKGROUND

A. Standard Power Modules

A schematic of a standard power module is shown in Fig. 1. The primary functional component of the module is the power device, labeled as a SiC chip in the figure. The devices are soldered to a direct bond copper (DBC) substrate using gold tin (AuSn) or other electrically conductive die-attach material. The DBC substrate is a three layer board: copper (0.35 mm), ceramic (0.635 mm), copper (0.35 mm), and is patterned on the side connected to the devices. The DBC substrate functions as the electrical circuity to connect the devices to each other and the external connections. The ceramic is either aluminum nitride (AlN) or alumina (Al_2O_3), depending on cost and performance criteria, with the AlN having both higher performance and higher cost. Thick aluminum wirebonds (not pictured in the figure) are used to electrically connect the topside of the devices to other devices and to the DBC substrate. The underside of the DBC is attached to a heat spreader through a thermal interface material (TIM). The devices and wirebonds are encapsulated in a dielectric gel primarily for voltage isolation. The structure is completed with a surrounding housing which includes electrical contacts. The bottom side of the heat spreader, which also forms the module bottom, is then attached to a separate heat sink. The separate heat sink is defined by a heat transfer coefficient (h). A heat transfer coefficient of around 5-10 W/m^{2}K represents natural air convection, a heat transfer coefficient of around 100-1000 W/m^{2}K represents forced air cooling, a heat transfer coefficient of around

1,000 – 100,000 W/m^{2}K represents single-phase liquid cooling. Past 100,000 W/m^{2}K requires advanced cooling technologies (two-phase, jet impingement, etc.). [4]

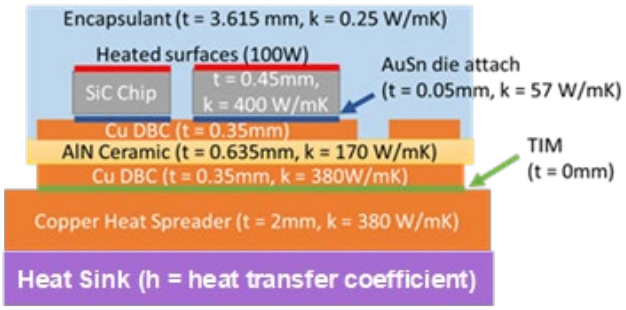

Fig. 1 Standard power module configuration with baseline thicknesses and thermal conductivity.

B. Power Module Case Study Referenced in this Work

In order to demonstrate the thermal impact of moving towards UWBG devices, a case study has been performed using a 1.2kV, 400A all SiC power module previously built and published by the Army Research Laboratory. [5, 6] This module was built using standard fabrication methods such that the conclusions made on this power module could be applied to other power modules fabricated in a similar fashion. A drawing of the module is shown in Fig. 2 and consists of four almost-symmetric quadrants, each of which comprises an individual DBC board with three diodes and four MOSFETs (metal-oxide semiconductor field-effect transistors) for a total of 16 MOSFETs and 12 diodes. The final fabricated module (without the lid) is shown in Fig. 3.

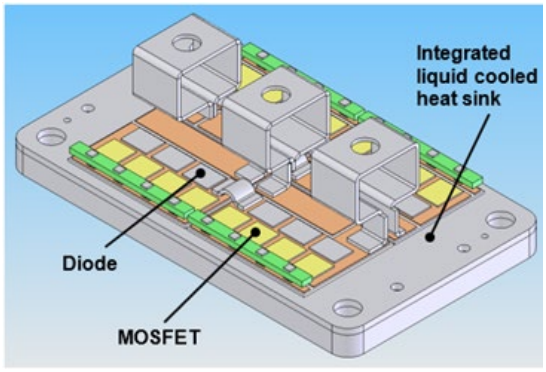

Fig. 2 Schematic of the module used in the case study.

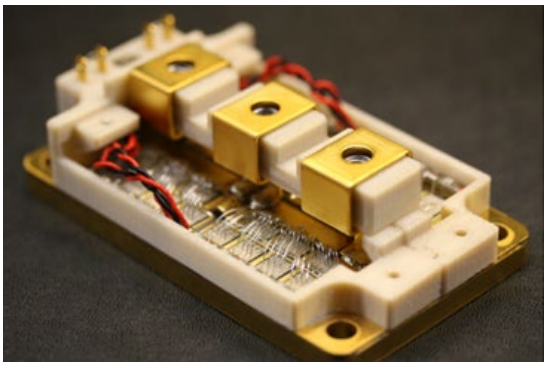

Fig. 3 Image of the fabricated module used for case study.

C. ARL ParaPower Modeling Tool

As electronics become smaller and more power dense, there is an increasing need for modeling tools which can quickly and accurately model the design space. ARL ParaPower is an open-source design tool which can quickly analyze parametric spaces including material types, material properties, layout designs, geometry, heat-sink selections, and heat-sink placement. Varying multiple parameters can quickly become computationally expensive for standard finite element analysis (FEA) and computational fluid dynamics (CFD) tools. ParaPower eliminates computer-aided design (CAD) in favor of numerical parameters that can be easily and quickly varied over a wide range. The tool is based on a resistor network which solves quickly in MATLAB, enabling fast, iterative thermal analyses and designs. A thermal resistor network is a foundation of thermal analysis and a 1D thermal resistance network is used by many thermal designers to determine steady-state temperatures. [7] ParaPower expands the 1D network into a 3D space with easily definable features which are further divided into elements. [8] Each element has thermal resistors in all six directions from the center of the element, as is shown in Figure 4.

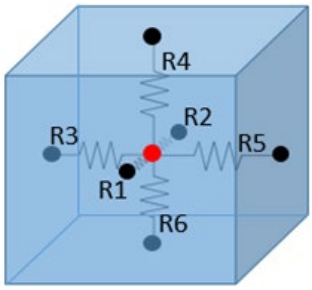

Fig. 4 Single node resistor network.

Although the study in this paper is a static analysis, at a general level ParaPower is based on an implicit Euler scheme solver to determine the temperature at each timestep. The implicit, or backward Euler method was selected due to its inherent stability. [7] The equation for a single node is given by:

$$\sum_{i=1}^{6}\left(\frac{1}{R_i}\left(T_i^{p+1}-T_0^{p+1}\right)\right)-CT_0^{p+1}=-Q_0-CT_0^{p} \qquad (1)$$

where

$$C=\frac{\rho_0^p c_0^p V_0}{\Delta t} \qquad (2)$$

For static analyses, used for all the analysis in this paper, the capacitive (transient) terms involving C are dropped.

Each element is then connected to each adjacent element forming a 3D thermal resistance network, as is shown in Fig. 5 for two dimensions. Each element creates an equation as shown above, and the series of equations is solved through a matrix inversion in MATLAB to calculate the temperatures at the center of each element. Fig. 5 is a representation of two features (yellow and blue) which have been divided into three elements each. The perimeter of the network model is comprised of ambient heat transfer coefficients (h_∞) and ambient temperatures (T_∞) which can vary over the perimeter of the structure. The model is adaptable to any number of features, heat loads, and heat transfer conditions.

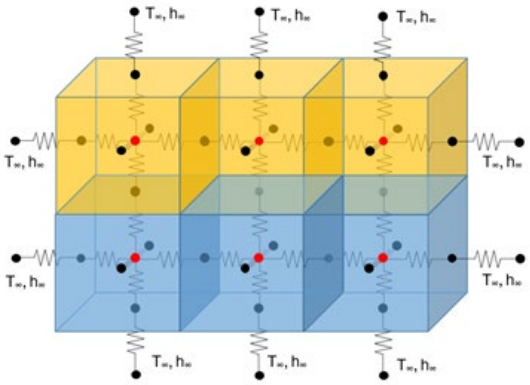

Fig. 5. Thermal Network Formulation.

In order to quickly run the parametric modeling analysis, the open source ARL ParaPower modeling tool was used. The geometry of the power module allowed analysis of a quarter symmetry model of the power module as shown in Fig. 6.

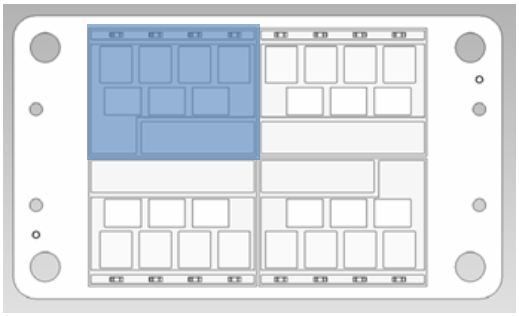

Fig. 6 View of the top of the baseline power module indicating the section for further analysis.

The geometry used in ARL ParaPower is shown in Fig. 7 without the encapsulation and natural convection boundary conditions. 100 W of heat is applied to each of the four MOSFETs, and no heat is generated by the diodes (100 W/chip times four chips = 400 W total). The heat dissipated is applied evenly to the top surface of the die. This is a reasonable, first-order approximation of the problem. A more thorough examination would consider local heating only on some portions of the chip surface, as well as heat generation within the bulk of the device.

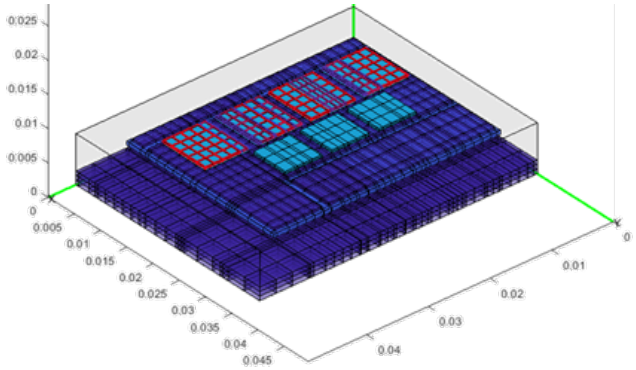

Fig. 7 ARL ParaPower model.

The ARL ParaPower results and the experimental results are shown in Figure 8 indicating good agreement. The results are at a heat transfer coefficient of 50,000 W/m^{2}K which corresponds to a flow rate of 7.6 L/min in the attached heat sink (not modeled) which has been previously calculated through CFD analysis [8].

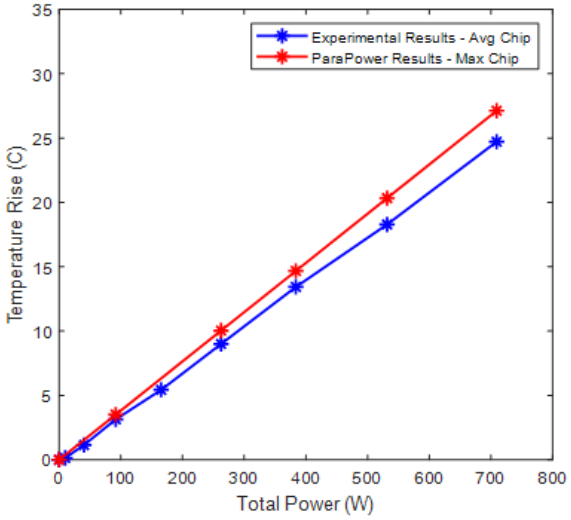

Fig. 8 Plot showing good agreement between the experimental results and the ARL ParaPower results.

Having experimentally validated the ARL ParaPower results in this configuration, subsequent analysis will only be performed using the ARL ParaPower tool.

III. RESULTS AND DISCUSSION

The first analysis performed simply replaced the SiC chip with other material types. The results are shown in Fig. 9 of the maximum temperature rise calculated for 0.45-mm thick chips, each composed of a different power semiconductor material, assuming the standard package described above, and a backside heat transfer coefficient of h = 50,000 W/m^{2}K applied evenly to the bottom of the power module. In all cases the plots are for a chip temperature rise so the analysis is independent of a heat sink temperature or a maximum chip temperature. To determine this for any application, one would simply add the temperature rise to the fluid temperature. For example, if the fluid was at 85 °C and the chip temperature rise was 45 °C, then the chip temperature is 130 °C. The analysis assumed 100 W of heat was dissipated by each of the four power switches and no heat generated by the three diodes.

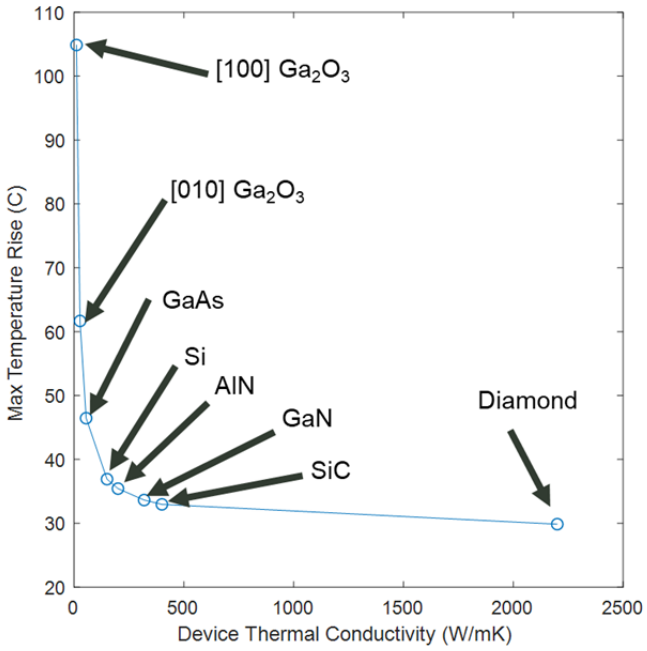

Fig. 9 Maximum chip temperature versus the thermal conductivity of various power semiconductors (h = 50,000 W/m^{2}K, 100W per MOSFET, and chip thickness of 0.45 mm).

Since this is a steady state analysis, the only material property that affects the solution is the thermal conductivity. The thermal conductivities used in this analysis are shown in Table 1. Although the thermal conductivity is a temperature dependent parameter, and the values vary greatly in the literature, for ease of comparison we assume a constant, temperature-independent thermal conductivity.

It is important to note that these analyses are performed assuming an equivalent comparison based on the rise in chip temperature. The maximum reliable operating temperature of each of these devices varies and could cause one material to perform better than another material. For example, silicon is usually rated to 125 °C and silicon carbide can operate above 175 °C, thus giving the SiC a larger margin. But to reduce complexity, all devices are compared against simply their temperature rise and not the absolute temperatures.

Table 1. UWBG materials and the thermal conductivity used in the analysis.

Material	Thermal Conductivity (W/mK)
[100] Gallium Oxide (Ga_2O_3)	11 [9]
[010] Gallium Oxide (Ga_2O_3)	27 [9]
Gallium Arsenide (GaAs)	55
Silicon (Si)	150
Gallium Nitride (GaN)	200
Aluminum Nitride (AlN)	319
Silicon Carbide (SiC)	400
Diamond	2200

It is evident from Figure 9 that a low thermal conductivity results in a significant rise in chip temperature, whereas a high thermal conductivity results in a very small rise. For example, the difference in chip temperature rise between [100] Ga_2O_3 and [010] Ga_2O_3—with thermal conductivities of 11 and 27 W/mK, respectively—is 43.2 °C. On the other hand, moving from SiC (k=400 W/mK) to diamond (k=2200 W/mK) only improves the die temperature by 3.1 °C. Thus, for chip thermal conductivites less than 50 W/mK, the temperature is very sensitive to thermal conductivity and small improvements can make a large impact, whereas beyond 400 W/mK, further improvements in thermal conductivity have little impact on the die temperature. Therefore, unless the electrical benefit of moving to those chip material types makes sense, there is little reason to do so unless a new, improved packaging approach causes the chip temperature rise to dominate.

The previous analysis was performed using a heat transfer coefficient of 50,000 W/mK, which assumes a relatively good single-phase liquid-cooled heat sink. The heat sink also has a significant impact on the overall rise in die temperature. Analysis was performed—for each of the thermal conductivities considered in Fig. 9—for a wide array of potential heat sinks, with various heat transfer coefficients ranging from 5,000 W/mK up to 1,000,000 W/mK. These results are shown in Fig. 10, which considers the maximum temperature rise versus the heat sink's heat transfer coefficient, as a function of the thermal conductivity of the chip. The temperature rises steeply with small changes in h below 50,000 W/m^{2}K, whereas values of h above 200,000 W/m^{2}K have very little effect. Minimal additional improvement occurs past 1,000,000 W/m^{2}K, and the temperature goes too high for anything less than 5,000 W/m^{2}K. Therefore, no heat sinks beyond this range are worth considering.

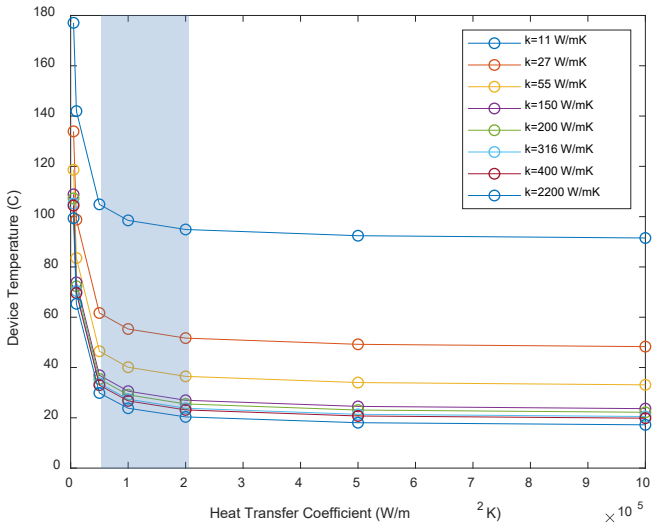

Fig. 10 Device temperature rise versus heat sink heat transfer coefficient for chip thickness of 0.45 mm. The optimum range is highlighted.

Thus, Fig. 10 is an important plot when trying to determine the appropriate heat sink for a given application and package. Not only does a heat sink with a heat transfer coefficient beyond about 200,000 W/m^{2}K not significantly improve the device temperature (for this example module), it also requires thermal management systems with increasingly larger pumping power and other complications which should be avoided if possible. The figure also shows that device temperature is very sensitive to a heat transfer coefficient less than around 50,000 W/m^{2}K. This means that the device temperature increases significantly with only small reductions in the heat transfer coefficient. This is also not an ideal regime because small changes in heat transfer coefficients can cause large fluctuations in chip temperature. Therefore, for this power module, for any chip selection, the heat transfer coefficient should be between 50,000 and 200,000 W/m^{2}K—this is shown as a light blue box on the plot in Fig. 10.

The results shown in both Fig. 9 and Fig. 10 were calculated by assuming a die thickness of 0.45 mm. But since there is a trend towards thinner die, the effect of die thickness was investigated as well. Fig. 11 shows the temperature rise as a function of chip thickness (which was varied between 0 and 500 µm in 100 µm increments), assuming a backside heat transfer coefficent of 50,000 W/m^{2}K, which is consistent with the low end of the ideal heat sink operating regime identified in Fig. 10. A thickness of 0 µm is not physically realizable, but it shows the limit of an infinitely thin chip. The figure indicates that for lower thermal conductivity chip materials (< 150 W/mK), reducing the thickness can have a significant improvement on the chip temperature. For chips with thermal conductivities > 150 W/mK, the thermal improvement by thinning the chips is not as critical but small improvements in chip temperature are observed. For values of thermal conductivity > 400 W/mK (including SiC and diamond) the impact is minimal.

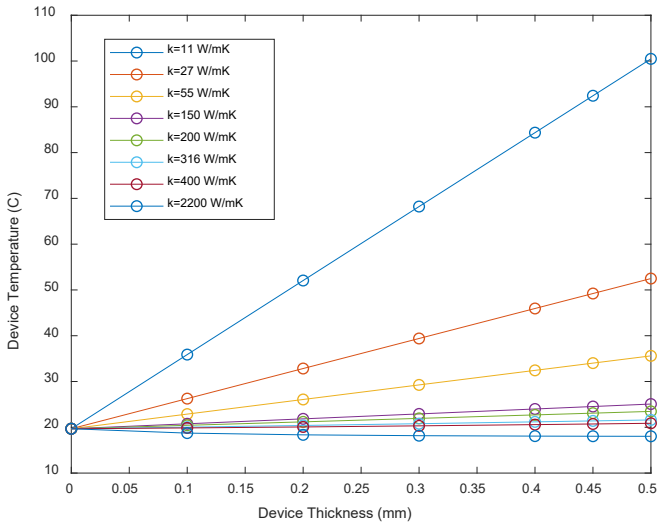

Fig. 11 Variation of maximum chip temperature with chip thickness for various power semiconductor materials, assuming a heat transfer coefficient of 50,000 W/m^{2}K.

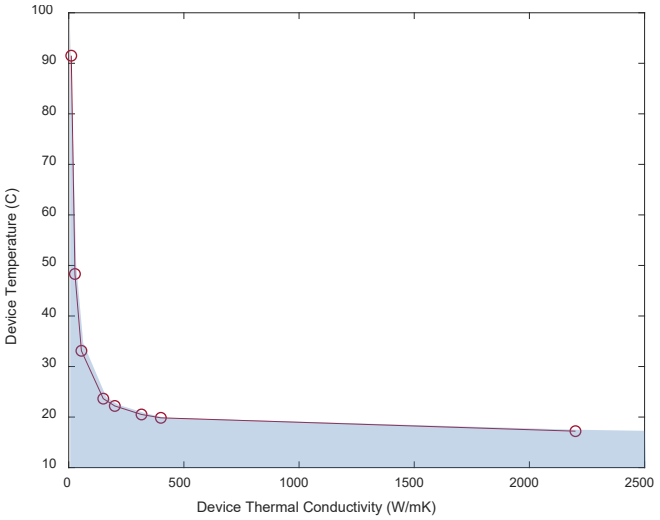

Fig. 12 Device temperature based on device thermal conductivity assuming h=1,000,000 W/m^{2}K, used to calculate package thermal resistivity

When looking at datasheets, the thermal resistance is a common metric which refers to the junction to case thermal resistance ($R_{\Theta JC}$), and is measured between the top of the device and the base of the module. $R_{\Theta JC}$ can be reasonably approximated by assuming an infinitely good heat sink (h = 1,000,000 W/m^{2}K) on the backside of the module, which eliminates the effect of the temperature rise of the heat sink and just considers the temperature rise of the packaging itself. The results of this are shown in Fig. 12. This assumes a device thickness of 0.45 mm, and 100 W power dissipation per each of four MOSFETs. Dividing the device temperature by 100 W gives the junction to

case thermal resistance ($R_{\Theta JC}$), with units of K/W. The $R_{\Theta JC}$ varies from 0.915 K/W for Al_2O_3 to 0.172 K/W for diamond. It is important to emphasize that this calculation assumes an extremely good heat sink. Therefore, to know what the actual device temperature will be, it is always necessary to add the effects of the actual heat sink to be employed.

Fig. 13 shows more generally the effects of the heat sink on the device temperature, showing how the device temperature rises for heat transfer coefficients between 5,000 W/m^{2}K to 1,000,000 W/m^{2}K. The figure also shows the effect of device thermal conductivity on chip temperature indicating that for all heat transfer coefficeints, all device thermal conductivities greater than 400 W/mK do not significantly affect the temperature rise of the device, whereas thermal conductivites less than 150 W/mK, and especially less than 100 W/mK can have profound effects on the die temperature since this is a most sensitive regime.

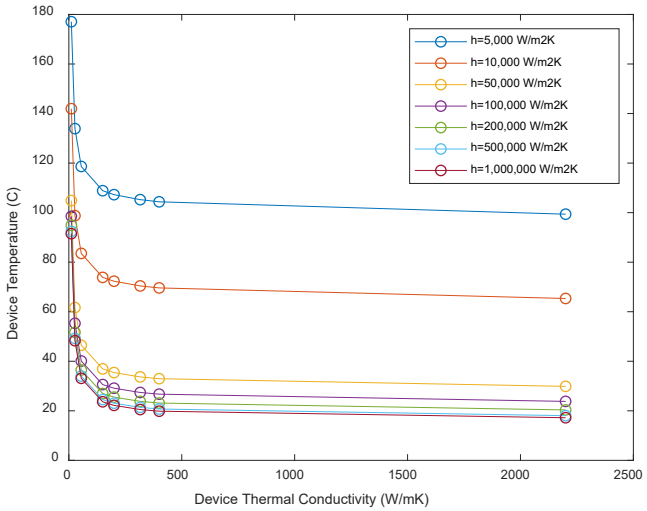

Fig. 13 Device temperature rise versus device thermal conductivity for chip thickness of 0.45 mm.

The device temperature strongly depends on both the chip temperature and heat sink. This notion is reinforced and summarized by the comparison shown in Fig. 14. The plot shows the temperature rise for the combination of chip and package along the positive y-axis, and the additional temperature rise from the heat sink chosen along the negative y-axis. The better the heat sink, the greater the impact of the chip thermal conductivity and thickness. Clearly, the very poor thermal conductivity of Ga_2O_3 results in a significant temperature rise, whereas for SiC and diamond the temperature rises are exceedingly small.

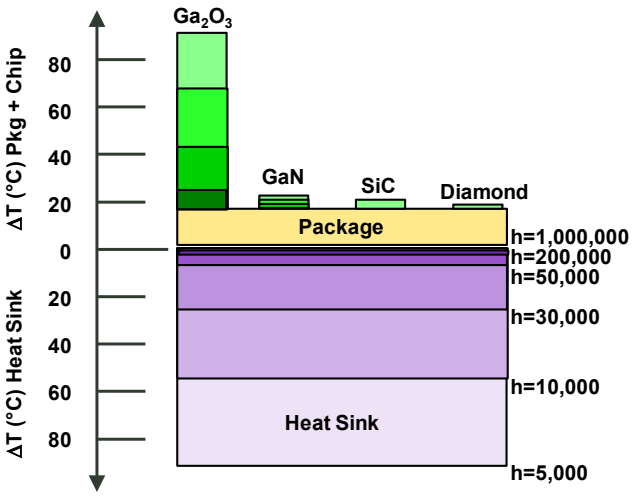

Fig. 14 Variation of chip temperatures with chip thickness for various power semiconductor materials, with a separate temperature delta calculated for various heat sink HTC values.

Fig. 14 demonstrates why constant temperature boundary conditions are a poor assumption. By setting the bottom temperature of the module to a fixed temperature, that is assuming an infinitely good heat sink and not accounting for the additional temperature rise due to the heat sink. This is also why using the $R_{\Theta JC}$ value alone to calculate temperature is also a poor approximation. There will be additional temperature rise due to the heat sink which could be significant depending on the heat sink performance and is non-negligible.

IV. CONCLUSION

In this paper the thermal impact of different WBG and UWBG materials were investigated for various chip thicknesses and heat sinks using a standard power electronics module configuration. The thermal conductivities considered range from 11 W/mK for [100] β-Ga_2O_3 to 2,200 W/mK for diamond. The results show a knee in the response curve lying between 150 W/mK for Si and 400 W/mK for SiC. For devices with thermal conductivities less than 150 W/mK, the effect on temperature is significant and efforts should be taken to reduce the thickness if possible. For devices with thermal conductivities greater than 400 W/mK, the effect on temperature is relatively negligible compared to the effect of the packaging and heat sink, and these materials should only be considered if the electrical and/or thermal properties justify their use. All of these analyses were performed using standard power module metrics with variable heat sink performance and 100 W power dissipation per chip. Due to the low thermal conductivity of the Ga_2O_3 and the high thermal conductivity of the diamond, it is possible that moving to UWBG devices will necessitate the move towards advanced power module configurations and or heat sinks.

V. REFERENCES

[1] J. Y. Tsao, et al., "Ultrawide-Bandgap Semiconductors: Research Opportunities and Challenges," Adv. Electron. Mater. 4, 1600501 (2018).

[2] Boteler, L. M., & Sharar, D. (2016, March). Thermal and Packaging Challenges of Power Electronics Modules. In Gomac Tech 2016 Proceedings (pp. 14-17)

[3] Boteler, L., Smith, A. "3D thermal resistance network method for the design of highly integrated packages", Proceedings of the ASME Summer Heat Transfer Conference, July 2013.

[4] Mudawar, Issam. "Assessment of high-heat-flux thermal management schemes." *IEEE Transactions on Components and Packaging Technologies* 24.2 (2001): 122-141.

[5] Boteler, L., Urciuoli, D., Ovrebo, G., Ibitayo, D., & Green, R. (2010, February). Thermal performance of a dual 1.2 kV, 400 a silicon-carbide MOSFET power module. In 2010 26th Annual IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM) (pp. 170-175). IEEE.

[6] Urciuoli, D., Green, R., Lelis, A., & Ibitayo, D. (2010, September). Performance of a dual, 1200 V, 400 A, silicon-carbide power MOSFET module. In 2010 IEEE Energy Conversion Congress and Exposition (pp. 3303-3310). IEEE.

[7] Incropera, Frank P., et al. Fundamentals of heat and mass transfer. Wiley, 2007.

[8] Anderson, Dale, John C. Tannehill, and Richard H. Pletcher. Computational fluid mechanics and heat transfer. CRC Press, 2016.

[9] Zhi Guo, Amit Verma, Xufei Wu, Fangyuan Sun, Austin Hickman, Takekazu Masui, Akito Kuramata, Masataka Higashiwaki, Debdeep Jena, and Tengfei Luo, " Anisotropic thermal conductivity in single crystal Beta-gallium oxide," Appl. Phys. Lett., 106, 111909 (2015).