FULL SCALE SIMULATION OF AN INTEGRATED MONOLITHIC HEAT SINK FOR THERMAL MANAGEMENT OF A HIGH POWER DENSITY GaN-SiC CHIP

Tanya Liu ∗
NanoHeat Laboratory
Department of Mechanical Engineering
Stanford University
Stanford, California 94305 - USA
Email: tanyaliu@stanford.edu

Farzad Houshmand
Catherine Gorle †
Sebastian Scholl ‡
Hyoungsoon Lee
Yoonjin Won
Mehdi Asheghi
Kenneth Goodson
NanoHeat Laboratory
Department of Mechanical Engineering
Stanford University
Stanford, California 94305 - USA

Hooman Kazemi
Kenneth Vanhille
Nuvotronics LLC
Radford, VA 24141 - USA

ABSTRACT
Advances in manufacturing techniques are inspiring the design of novel integrated microscale thermal cooling devices seeking to push the limits of current thermal management solutions in high heat flux applications. These advanced cooling technologies can be used to improve the performance of high power density electronics such as GaN-based RF power amplifiers. However, their optimal design requires careful analysis of the combined effects of conduction and convection.

Many numerical simulations and optimization studies have been performed for single cell models of microchannel heat sinks, but these simulations do not provide insight into the flow and heat transfer through the entire device. This study therefore presents the results of conjugate heat transfer CFD simulations for a complex copper monolithic heat sink integrated with a 100 micron thick, 5 mm by 1 mm high power density GaN-SiC chip. The computational model (13 million cells) represents both the chip and the heat sink, which consists of multiple inlets and outlets for fluid entry and exit, delivery and collection manifold systems, and an array of fins that form rectangular microchannels. Total chip powers of up to 150 W at the GaN gates were considered, and a quarter of the device was modeled for total inlet mass flow rates of 1.44 g/s and 1.8 g/s (0.36 g/s and 0.45 g/s for the quarter device), corresponding to laminar flow at Reynolds numbers between 19.5 and 119.3. It was observed that the mass flow rates through individual microchannels in the device vary by up to 45%, depending on the inlet/outlet locations and pressure drop in the manifolds. The results demonstrate that full device simulations provide valuable insight into the multiple parameters that affect cooling performance.

∗Address all correspondence to this author.
†Current affiliation: Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, NY, 10027
‡Visiting researcher at Stanford University. Permanent address: Von Karman Institute for Fluid Dynamics, Chaussee de Waterloo 72, 1640 Rhode-Saint-Genese - Belgium
INTRODUCTION

Since the initial proposal of microchannel heat sinks for cooling applications by Tuckerman and Pease [1], many analytical and numerical studies have been performed to characterize the heat transfer performance of these heat sinks [2–8]. Due to their compact design and effective heat dissipation capability, microchannel heat sinks have emerged as the cooling devices of choice for high heat flux electronics. Advances in manufacturing techniques have created the possibility of increasingly complex and novel microchannel heat sink geometries, but a detailed understanding of the combined effects of conduction and convection in the heat sink is required for optimal design.

Many efforts have been made to both analyze and optimize microchannel heat sink performance based on a unit cell approach, where only a single channel is considered due to symmetry and periodicity conditions [9–17]. These studies optimize the heat transfer performance of a single channel over a variety of geometric parameters, operating conditions, and various constraints such as allowable pressure drop and manufacturing feasibility. Certain approaches have utilized evolutionary algorithms for multi-objective optimization [10–12], while other studies have found approximate analytical models sufficient for optimization [14]. Fedorov and Viskanta [2] presented a 3D numerical model of a single microchannel with conjugate heat transfer and found the average convective heat transfer coefficient to be largest near the channel inlet regions where the flow is thermally developing. This supports the concept of the manifold microchannel (MMC) heat sink, which compounds upon this inlet effect by using a series of parallel manifolds to distribute fluid to the microchannels through multiple inlets [16,18]. Lee et al. [4] compared experimental results for the heat transfer in a single microchannel to a range of numerical models and found that a simplified “thin wall” model (a boundary condition of axially uniform heat flux and circumferentially uniform temperature) could accurately predict the experimental results at a reduced computational cost compared to full 3D conjugate analysis. Liu and Garimella [14] proposed that a 1D thermal resistance model could be sufficient for the optimization of a single microchannel.

While valuable information regarding the conditions necessary for optimal heat transfer performance in a single channel can be extracted from these studies, these studies fail to take into account the fluid flow behavior and heat transfer throughout the entire device. This is particularly important in the case of MMC heat sinks, where the manifold design plays a large role in the fluid flow distribution.

Chein and Chen [19] conducted a series of simulations on microchannel heat sinks with five different inlet and outlet arrangements and found that varying the inlet and outlet locations affected both the fluid flow and heat transfer in the sink. Different inlet and outlet configurations caused varying degrees of flow maldistribution and subsequent temperature nonuniformity throughout the device. Tiselj et al. [20] performed simulations for a complete cooling device including a silicon chip and inlet/outlet collectors. Their full device simulation results differed significantly from other single channel studies, with the cited reason being axial heat conduction effects near the inlet and outlet regions that were not captured in single channel simulations [7,21]. Escher et al. [22] performed a numerical optimization of an MMC heat sink and found that the choice of optimum design parameters was strongly dependent on the fluid manifold design.

Based on all the complexities and intertwined parameters that affect the performance of a heat sink, performing a full device simulation is a critical step before fabrication of such complicated structures. Though computationally expensive, full device simulations provide valuable insight into the effects of various parameters often not captured by single cell simulations. This paper presents results from numerical simulations with 3D...
conjugate heat transfer for a complex copper monolithic heat sink integrated with a high power density GaN-SiC chip. The heat sink is designed for fabrication with the ChipChill process developed by Nuvotronics LLC. The computational model of the heat sink is complete with multiple inlets and outlets for fluid entry and exit, delivery and collection manifold systems, and an array of rectangular microchannels. The fluid flow behavior and temperature distribution throughout the entire device are analyzed, and the effects of various system parameters on the overall cooling performance of the device are discussed.

**POLYSTRATA PROCESS**

Nuvotronics has developed and commercialized the PolyStrata® process, which is an additive multi-layer microfabrication process that has been optimized for microwave and millimeter-wave circuits [23]. The ChipChill fabrication process uses many of the same base materials and fabrication techniques, but is now being optimized to address thermal management solutions for high power density integrated circuits. This process allows nearly-arbitrary combinations of a permanent polymer, copper, and air to be defined layer by layer in a wafer-scale, batch process that can be carried out on 150-mm-diameter and 200-mm-diameter substrates with fabrication tolerances on the order of a few microns. The features of each stratum across the wafer are defined using photolithography, and the photoresist is used as a mold for plating of the copper features once the pattern has been defined and developed. The copper is planarized using a chemical-mechanical polishing (CMP). At this juncture, photo-patternable permanent dielectric supports or sheets may be embedded in the device, or the photolithography process begins anew, and the steps repeat themselves. This process continues until the entire height of the structure has been achieved. The photoresist is then dissolved to leave air-filled copper structures with dielectric supports for the center conductor or other applications. The resulting structures can have individual strata thickness values from 5 to over 100 µm with a variety of possible surface coatings. This process provides the ability to precisely define integrated heat sink geometries in three dimensions to optimize the heat spreading and heat exchanging properties of the overall device. For the design under consideration, the entirety of the heat sink including inlets, outlets, channels, and delivery and collection manifolds are fabricated following this process.

**NUMERICAL MODEL**

The geometry of the copper heat sink consists of 48 rectangular channels with \( W_{ch} = 50 \, \mu m \), \( H_{ch} = 450 \, \mu m \), and \( L_{ch} = 1320 \, \mu m \), two inlets 300 \, \mu m in diameter, and eight outlets 300 \, \mu m in diameter. The device is symmetric in two directions, so only a quarter of the full device geometry is represented with symmetry boundary conditions applied as shown in Fig. 1. The inlets are located on the top surface of the device along the longitudinal axis, and the outlets are spaced around the periphery of the device. Fluid flows into the device through the inlets and enters a delivery manifold that guides the fluid to the center of each microchannel. The fluid enters perpendicular to the microchannel,
makes a 90° turn, then leaves the device through an exit manifold connected to the outlets. The exit manifold delivers the fluid to the outlet channels where it makes another 90° turn and leaves the heat sink through the top surface of the device. The bottom of the heat sink is attached to a 920 µm x 5260 µm GaN-SiC chip with a 10 µm thick layer of thermal interface material (TIM). The chip model consists of a 100 µm thick layer of SiC and 80 GaN gates (40 in the quarter device model) 2 µm x 320 µm in area.

Though there has been some previous discrepancy as to the nature of fluid flow and heat transfer in microchannels, studies have found that the conventional macroscale Navier-Stokes and energy equations are sufficient to describe the heat transfer behavior of microchannel heat sinks \[2, 21, 24, 25\]. With the assumptions of steady-state, laminar, incompressible flow, the three-dimensional governing equations of continuity, momentum, and energy can therefore be written as follows:

\[
\frac{\partial}{\partial x_i} (\rho u_i) = 0 \tag{1}
\]

\[
\frac{\partial}{\partial x_i} (\rho u_i u_j) = -\frac{\partial P}{\partial x_j} + \frac{\partial}{\partial x_i} \left( \mu \frac{\partial u_i}{\partial x_i} \right) \tag{2}
\]

\[
\frac{\partial}{\partial x_i} \left( k_f \frac{\partial T_f}{\partial x_i} \right) = \frac{\partial}{\partial x_i} (\rho u_i C_p T_f) \tag{3}
\]

\[
\frac{\partial}{\partial x_i} \left( k_s \frac{\partial T_s}{\partial x_i} \right) = 0 \tag{4}
\]

The heat transfer behavior in a microchannel heat sink is characterized as a conjugate heat transfer problem, as there is combined conduction and convection in the microchannels. As such, the following conditions for continuity of temperature and heat flux at the solid-fluid interfaces hold:

\[
T_{s,int} = T_{f,int} \tag{5}
\]

\[
-k_s \left( \frac{\partial T_s}{\partial n} \right)_{int} = -k_f \left( \frac{\partial T_f}{\partial n} \right)_{int} \tag{6}
\]

<table>
<thead>
<tr>
<th>Material</th>
<th>Thermal Conductivity [W/m·K]</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaN</td>
<td>$-0.1623T + 214.17$</td>
</tr>
<tr>
<td>SiC</td>
<td>$0.0038T^2 - 4.1734T + 1259$</td>
</tr>
<tr>
<td>Epotek</td>
<td>30</td>
</tr>
<tr>
<td>Copper</td>
<td>387.6</td>
</tr>
<tr>
<td>Water</td>
<td>0.6</td>
</tr>
</tbody>
</table>

TABLE 2. MESH INDEPENDENCE STUDY RESULTS.

<table>
<thead>
<tr>
<th>Number of cells</th>
<th>$\Delta P$ [kPa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.9 million</td>
<td>41.3</td>
</tr>
<tr>
<td>13.2 million</td>
<td>40.7</td>
</tr>
<tr>
<td>19.7 million</td>
<td>40.5</td>
</tr>
</tbody>
</table>

A range of uniform heat fluxes from $q = 1.56 \times 10^9$ W/m$^2$ to $2.93 \times 10^9$ W/m$^2$ is applied to the bottom of the GaN surfaces to represent a total input thermal power range of 80 W to 150 W from the chip. The working fluid is water that enters the inlets at 343 K. The thermophysical properties of the water and copper are held constant, but the thermal conductivities of the GaN and SiC are allowed to vary as functions of temperature. Table 1 lists the thermal conductivity polynomials used for the GaN and SiC, as well as the constant thermal conductivity values used for all other materials in the simulations. Two different quarter device inlet mass flow rates of 0.36 g/s and 0.45 g/s are specified at the inlets (thereby representing 1.44 g/s and 1.8 g/s to the entire device), and the outlets are set as atmospheric pressure outlets. Solid-fluid interfaces are treated as coupled walls to satisfy the conjugate heat transfer conditions given by Eqn.’s (5) and (6), and the no-slip condition is applied in the fluid domain to solid-fluid interfaces. All external walls are treated as adiabatic boundaries.

The overall system of equations and applied boundary conditions are solved using the commercially available ANSYS Fluent® finite-volume CFD solver. The pressure-velocity coupling is treated using the SIMPLE algorithm, and a second order numerical scheme is used for all discretization schemes. The device model is discretized using a hexahedral mesh, and a grid-independency study was performed to ensure grid-independent results. Three different mesh sizes of 4.9 million cells, 13.2 million cells, and 19.7 million cells were considered, and Table 2 shows the pressure drop results across the entire device for each
mesh. Moving from the coarse mesh to the intermediate mesh showed a 1.5% change in $\Delta P$, while moving from the intermediate mesh to the fine mesh showed only a 0.5% change in $\Delta P$. For the sake of computational savings, the intermediate mesh was deemed sufficient and chosen for use in all subsequent simulations. Convergence was determined in all cases by monitoring the device $\Delta P$ and temperature in the GaN gates until variation between iterations was within 1%. An energy balance check was performed to ensure that the increase in internal energy of the fluid matched the applied power at the GaN gates, and the net energy imbalance was less than 0.25%. In addition, a mass balance check on the inlet and outlets showed a net mass imbalance of less than 0.01%.

**RESULTS AND DISCUSSION**

**Fluid flow**

There are 24 channels in the quarter device geometry, and the channels are numbered such that channel 1 is located furthest from the inlet, and channel 24 is located in the center of the full device (adjacent to the symmetry plane in the quarter device domain). The channel Reynolds numbers based on the hydraulic diameter $D_h$ range from 19.5-119.3 for the two different inlet mass flows, confirming that the flow is indeed laminar.

Figure 2 shows the mass flow distribution per channel for both the 0.36 g/s quarter device inlet mass flow and the 0.45 g/s inlet mass flow. The fluid flow distribution through the device is clearly non-uniform, with channel mass flow variations of up to 45% and 36% for the 0.36 g/s and 0.45 g/s mass flow rates, respectively. This channel flow non-uniformity is very important in terms of heat sink design, as channels with lower mass flow rates may potentially have weaker cooling performance than channels with higher mass flow rates, resulting in temperature non-uniformity in the solid.

Relatively higher mass flow rates are observed in channels located near the device inlet and outlets. Figure 3 shows a more in depth view of the fluid flow behavior across individual channels for an inlet mass flow of 0.36 g/s. As shown in Fig. 3(a), channel 14 is located directly beneath the inlet and receives a relatively high amount of mass flow. There appears to be a recirculation zone in the first layer of the inlet manifold, though the majority of the fluid enters the microchannel and experiences a velocity increase from the sudden contraction at the microchannel inlet. A stagnation point is visible as the low velocity region in the outlet manifold, at which the fluid exiting channel 14 turns and flows along the outlet manifold in the direction of outlets 1 and 2. Fig. 3(b) shows the velocity vectors for channel 7, a chan-
nel with a relatively lower amount of mass flow and not located directly beneath the inlet. The fluid exiting channel 7 leaves at a low velocity and joins the stream of faster moving fluid composed of all the fluid exiting from channels located upstream.

Figure 4 shows a section view of the velocity vectors for the 0.36 g/s mass flow rate as the fluid flows through the microchannels and enters the outlet channels. The stagnation point occurs in the outlet manifold in between channels 14 and 15, and the fluid flow from the two channels splits into opposite directions. These two streams accelerate in the outlet manifold as they merge with flow from the other channels and eventually reach the outlet channels. The sudden contraction and 90° bend at the outlets causes a sharp increase in fluid velocity, and regions of high fluid velocities up to 9.1 m/s are observed at these outlet entry points. In comparison, the highest average channel velocity observed for the same inlet mass flow rate is only 1.1 m/s, though velocities up to 5 m/s are reached at the channel inlets. The high velocity regions near the outlet entrances can largely be explained by the outlet spacing. Outlet 1 is relatively ineffective in removing fluid from the outlet manifold, and the resulting mass flow distribution among the three outlets with the symmetry condition taken into account is 23%, 28%, and 49%, respectively. Outlet 3 receives a disproportionately high amount of mass flow, exacerbating the high velocity effects caused by the contractions and bends at the outlet channels. This issue of uneven mass flow distribution to the outlets could be solved in future designs by spacing the outlets more uniformly relative to the microchannels, as outlet 3 alone currently receives fluid from 10 microchannels (20 microchannels for the full device), while the remaining 14 channels are split amongst outlets 1 and 2. Having larger entrances to the outlet channels could also help to alleviate the high velocity regions.

**Pressure drop**

The pressure drop across each channel for both inlet mass flows is shown in Fig. 5. The pressure drop distribution closely follows the mass flow distribution, with the exception of channels 14 and 15, where the pressure drops are lower than surrounding channels with lower mass flow rates. The total pressure drop across the device is 40.7 kPa for the 0.36 g/s mass flow and 59.3 kPa for the 0.45 g/s mass flow. Analysis of the pressure contours in the fluid domain reveals that a major part of the pressure drop in the device actually occurs in the outlet manifold, and not across the microchannels. The maximum pressure drop across the channels for the 0.36 g/s mass flow is 10.2 kPa, while the pressure drop across the outlet manifold is as large as 30 kPa. A similar trend is observed for the 0.45 g/s mass flow. This is highly significant, as the cooling performance of the device is limited by the total pressure drop. Raising the inlet mass flow rate increases the overall heat transfer performance of the device, but this improvement in heat transfer efficiency also requires an increase in pumping power to accommodate higher pressure drops across the device. Optimal heat sink performance therefore desires the highest heat transfer efficiency possible with the lowest pumping power requirements. Above a certain point, the pumping power requirement for a device becomes so high that the benefit from increased heat transfer performance becomes insignificant.

The existence of this trade off once again reiterates the importance of overall device design. Optimization based on single cell pressure drops may be insufficient when, as is the case in this heat sink structure, the overall pressure drop across the device is dominated by the inlet and outlet manifold designs. For optimal device cooling performance, modifications should be made so that higher inlet mass flow rates can be supported while maintaining relatively low pressure drops across the entire device. In the current design, most of the pressure drop occurs in the outlet manifold, so increasing the number of outlets or widening the width of the outlet manifold channel are both potential solutions that could lead to significant reduction in overall pressure drop.

**Temperature distribution in the GaN gates**

During operation of the chip, the electrical/RF performance of the chip is limited by the temperature at the GaN gates. Risk of component failure arises when the gate temperatures rise above a certain threshold value, and a uniform temperature distribution is desired among the gates for the chip to maintain maximum efficiency. A range of different heat fluxes from $\dot{q} = 1.56 \times 10^9$ W/m² to $2.93 \times 10^9$ W/m² are applied to the gates to represent total chip thermal powers ranging from 80 W to 150 W. Based on the fluid flow field results, the non-uniform nature of the mass flow distribution in the microchannels might imply a correspondingly non-uniform gate temperature distribution, however, the observed variation in gate temperature distribution for each heat flux is relatively small. A large temperature drop is observed in
the gates located near the end of the chip (approximately gates 1-8), but this temperature drop is largely due to edge effects and not the fluid flow non-uniformity in the channels above.

Further analysis reveals that the temperature uniformity in the gates is a result of heat spreading in the SiC layer. Figure 6 shows the maximum temperature at each gate for an inlet mass flow of 0.36 g/s and applied heat flux of $\dot{q} = 1.56 \times 10^9$ W/m$^2$ for two different SiC layer thicknesses of 100 $\mu$m and 25 $\mu$m. With the exception of the edge gates, the variation in gate temperature in the remaining gates with the 100 $\mu$m thick layer of SiC is only 0.6%. When the SiC layer is reduced to 25 $\mu$m, greater temperature variation is observed in the gates, as the thinner SiC layer limits the amount of heat spreading possible and increases the effect of the fluid flow non-uniformity on the gate temperatures. Even then, however, the overall variation in the maximum gate temperatures excluding the edge gates is still less than 1.5%. Analysis of the temperature contours in the SiC for both cases further confirms these results. Figure 7 shows a temperature contour comparison between the 100 $\mu$m thick case and 25 $\mu$m thick case for (a), the GaN side surfaces, and (b), the TIM side surfaces. In accordance with the gate temperature distribution results, the larger role played by lateral conduction in the presence of a thicker SiC layer leads to a greater temperature uniformity in the 100 $\mu$m SiC than in the 25 $\mu$m SiC.

With the exception of the edge gates, the rest of the gates manage to achieve a relatively uniform temperature distribution even in the presence of non-uniform mass flows. The temperature variation in the edge gates is significant, however, as there is a nearly 25 K drop from the central gates to the outermost gate for both SiC thicknesses. A future design might consider removal of the outermost microchannel (channel 1) above the edge gates entirely to prevent excess cooling and decrease the temperature drop.

It should be noted that the numerical model presented here does not take into account the thermal boundary resistance (TBR) between material layers, most notably the TBR between the GaN gates and SiC substrate. This TBR based on acoustic or diffuse mismatch models is typically on the range of $1 \times 10^{-9}$ m$^2$·K/W, though this is known to be an underestimate, and experimental values have been found to be closer to $4 \times 10^{-9}$ m$^2$·K/W [26, 27]. For the heat flux under consideration of $1.56 \times 10^9$ W/m$^2$, this TBR would lead to an additional temperature drop across the GaN-SiC interface of 6-7 K.

Possibility of two-phase flow  For higher heat fluxes, the possibility of two-phase flow in the device arises. The maximum wall temperatures observed in the microchannels for both inlet mass flows approach values well above the saturation temperature of water as the applied heat flux is increased, so the possibility of subcooled boiling was investigated. As noted in the previous section, these wall temperatures do not include the effect of the GaN-SiC TBR, however, so the wall temperatures in reality may be lower. Using correlations from Sato and Mat-
The maximum power before boiling for each mass flow was calculated based on the maximum wall temperatures from simulations with different inlet heat fluxes. For the 0.36 g/s mass flow rate and 0.45 g/s mass flow rate, applied heat fluxes of 1.66x10^9 W/m^2 and 1.85x10^9 W/m^2 (85 W and 95 W total thermal power), respectively, were estimated as the maximum powers before the possibility of boiling occurs. Figure 8 shows where these transition points occur for each flow rate on a plot of the maximum temperature at the GaN gates versus total input thermal power. All thermal powers below these two transition points can be categorized as single phase, affirming that the results presented in the previous section for a heat flux of 1.56x10^9 W/m^2 with a single phase model are valid.

CONCLUSIONS

A full scale CFD simulation with conjugate heat transfer was performed for a complex microstructure. The complete computational domain encompassed the fluid inlet and outlet delivery manifolds, microchannels, and integrated GaN-SiC chip. The inlet and outlet manifold designs were found to have a large effect on the fluid flow distribution throughout the device, with relatively higher mass flows observed in channels located near the inlets and outlets. Channel mass flow variations of up to 45% and 36% were observed for two different inlet mass flow rates of 0.36 g/s and 0.45 g/s. For both inlet mass flow rates, it was observed that a large portion of the pressure drop in the device occurs in the outlet manifold. Due to significant heat spreading in the SiC layer, the temperature distribution at the GaN gates was found to be relatively uniform in spite of the channel mass flow variations.

These results emphasize the importance of performing full geometry simulations as part of the path to achieving an optimal microchannel heatsink design. Though computationally expensive, full scale simulations reveal valuable information about which design parameters may be the most important in obtaining the best device cooling performance. Effects due to factors such as the inlet and outlet manifold designs and heat spreading in the solid are visible in full simulation results, but cannot always be captured by a single cell model. Full scale simulations therefore play a crucial role in understanding the fluid flow and heat transfer throughout a device, particularly so for geometries of increasing complexity.

ACKNOWLEDGMENT

This project was supported in part by the U.S. Defense Advanced Research Projects Agency Microsystems Technology Office ICECool Applications Program under award number NuvoSFD-1401-1066. Disclaimer: The views, opinions, and/or findings contained in this article/presentation are those of the author/presenter and should not be interpreted as representing the official views or policies, either expressed or implied of the Defense Advanced Research Projects Agency or Department of Defense. This project was also supported in part by the Stanford Graduate Engineering fellowship program.

REFERENCES


