Design of Integrated Thin Film Coolers

Chris LaBounty*, Ali Shakouri**, Gerry Robinson*, Patrick Abraham*, and John E. Bowers*
*Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106
**Jack Baskin School of Engineering, University of California, Santa Cruz, CA 95046
labounty@opto.ucsb.edu

Abstract

Thin film coolers can be monolithically integrated with optoelectronic and high-speed electronic components. Important parameters in the design of such coolers are investigated theoretically and experimentally. A three-dimensional finite element simulator (ANSYS) is used to model self consistently thermal and electrical properties of a complete device structure. Heat conduction, Joule heating, thermoelectric and thermionic cooling are included as well as non ideal effects such as contact resistance, finite thermal resistance of the substrate and the heat sink, and heat generation in the wire bonds. Various substrate materials (InP, Si, Cu, Diamond) and thicknesses are studied as well as the effect of current spreading in different sample geometries. Comparisons to experimental results with InGaAsP-based thermionic coolers are made. Simulations demonstrate that single stage thin film coolers can provide up to 20-30 degrees centigrade cooling with cooling power densities of several 1000's W/cm^2.

Introduction

Thermoelectric (TE) coolers are used to control heat generation in numerous microelectronic and optoelectronic devices. TE coolers have become essential in modern optical telecommunications to control the characteristics of laser sources, switching/routing elements, and detectors used in wavelength division multiplexed systems. Cooling requirements in microprocessors and other integrated circuits have also risen dramatically in recent years due to the increase in speed and reduction in feature size. Generally, as these devices have become smaller, faster, and more dense, the power density has greatly increased. Since the cooling power density for TE coolers is dependent on the length of the n- and p-type semiconductor thermoelements, conventional TE coolers are limited due to their incompatibility with integrated circuit fabrication processes. Furthermore, this bulk fabrication technology makes integration with microelectronic and optoelectronic devices difficult, resulting in a high cost of packaging. The solution to these problems is to move from bulk thermoelements to integrated thin film coolers.

The greatest advantage of integrated thin film coolers is the dramatic gain in cooling power density as it is inversely proportional to the length of the thermoelements. With thin films on the order of microns, cooling power densities greater than 1000 W/cm^2 should be possible. While the efficiency of these devices will be reduced due to the heat conduction between the heating and cooling regions, several technologies, such as thermionic emission in heterostructures [1-2] and decreased thermal conductivity in superlattices [3-4], are being explored to regain this loss in efficiency and even improve beyond what is possible with bulk thermoelements. Integrated thin film coolers can also be made in large quantities using well known integrated circuit fabrication methods. Integration with microelectronic and optoelectronic devices is also possible for active localized cooling on demand.

Many non-idealities become apparent and must be considered when moving from bulk to thin film coolers. Where as contact resistance, thermal resistance of the substrate and heat sink, and heat generation in the current carrying connections are secondary effects in bulk TE coolers, they all become critical in thin film coolers. In the following, these non-ideal effects, as well as heat conduction, Joule heating, thermoelectric and thermionic cooling, are investigated. A three-dimensional self consistent thermal/electrical software simulator [5] is used to model these effects (Fig.1), and the results are compared to experimental measurements. Thermionic cooling is considered throughout to make fair comparisons to experimental data, however the results are applicable to TE thin film coolers as well.

Non-Ideal Effects

The most obvious non-ideal effect for thin film coolers is the substrate and package thermal resistance. Qualitatively speaking, if the distance between the cooling and heating regions is several orders of magnitude smaller than the distance between the heating and heat sink regions, most of the heat will flow back to the cold side of the device if the thermal conductivity's of the thin film and substrate are comparable. Figure 2 shows simulation results of substrate thermal resistance versus thickness for various substrate materials. The simulation is performed assuming a rectangular etched mesa (5000 μm^2), thin film cooler with a semi-infinite substrate in the planar direction. The sides and top of the substrate are assumed to be adiabatic while the

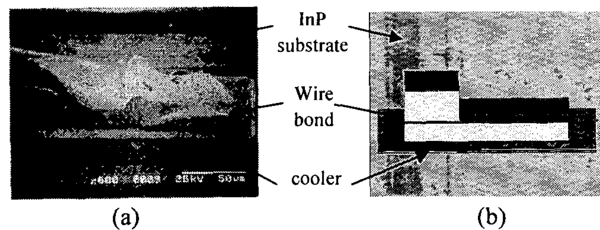

Fig. 1 (a) SEM of a heterostructure integrated thermionic (HIT) thin film cooler (InGaAs/ InGaAsP/ InGaAs 0.3μm/1μm/ 0.5μm) and (b) the simulated structure. The uppermost wire bond volume is scaled to reduce the element count in the mesh.

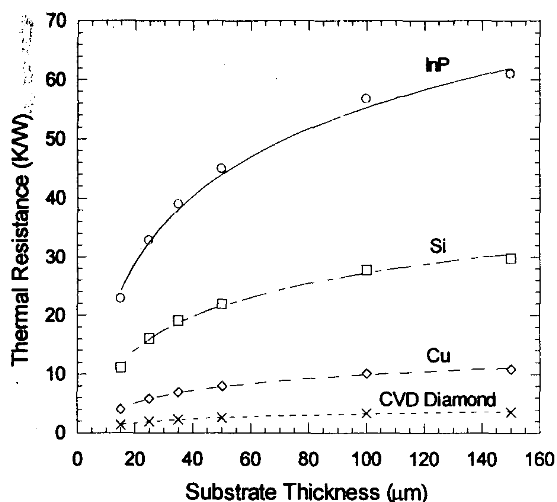

Fig. 2 Thermal Resistance versus substrate thickness for InP(β=71 W/m·K), Si(β=145 W/m·K), Cu(β=398 W/m·K), and CVD Diamond(β=1200 W/m·K). The points correspond to 3D simulation results, and the solid curves are the theoretical log(x) curve fits.

bottom is isothermal. The relatively good fit to a ln(x) function is indicative of the thermal spreading in two and three-dimensional heat flow [6,7]. Below 15 μm, the heat flow becomes purely one dimensional and the thermal resistance scales linearly with substrate thickness. Clearly it is beneficial to use substrates with high thermal conductivity and a minimum thickness. However, thin film coolers that are grown on typical substrate materials (InP, Si, etc.) have a limited amount by which the substrates can be lapped before the material begins to warp. This warping occurs with InP substrates when they are lapped below 100 μm, which for the device size simulated above corresponds to a thermal resistance of 57 K/W. Using a 1 μm film of InGaAsP for the cooler (β=3.3 W/m·K) gives a thermal resistance of 61.6 K/W showing that roughly half the heat flows back to the cold side of the device.

Several methods for transferring the thin films to surrogate substrates are possible. Epitaxial lift-off (ELO) and grafting of epitaxial films has been discussed in the literature [8-11] for various semiconductor material systems. This approach seems to work well for devices that do not require current conduction in the substrate, however neither the electrical or thermal contact resistance of the interface has been studied to the knowledge of the authors. Both of these parameters are critical in thin film coolers. Another possibility is to use flip chip bonding techniques to accomplish the same substrate transfer. While this alternative seems promising, problems still arise in processing the coolers when there is a thermal expansion mismatch between the thin films and the new substrate. For example during the alloying of ohmic metal contacts to InGaAs, the sample is subjected to temperatures greater than 400 °C. If the thin InGaAs film was mounted to a Cu substrate, extreme stress would be present on the film possibly damaging the material. To remedy this situation either low annealing temperature contacts are needed or substrates with high thermal conductivity and a thermal expansion coefficient that is close to the thin film material.

A more general concern than just the substrate thermal resistance is the issue of packaging. There is a thermal resistance associated with the solder layer (used to attach the die to the package) and the package itself. Besides providing good thermal properties, the package also needs to provide a way of carrying current to the device. Even using the largest diameter gold wire available, current experimental measurements of cooling in single stage devices is hampered by the large Joule heating due to the wire bonds. An optimum wire length can be determined by considering the Joule heating, thermal conduction away from device, wire diameter, and the temperature difference expected across the wire. Ultimately when the thin film coolers are integrated with real devices or are packaged in a conventional TE configuration (n- and p-type legs electrically in series, thermally in parallel), the issue of wire bonding will be less of a concern. Fig. 3 shows the increase in cooler performance for four generations of packages. As expected, when the packaging improves the cooling coefficient remains relatively constant while the heating coefficient is reduced resulting in greater cooling.

Another important non-ideal effect for thin film coolers is the electrical contact resistance. This additional Joule heating effect occurs very close to the cooling region of the device and must be minimized to attain any appreciable cooling. In order to determine the quality of our ohmic contacts, contact resistivity studies were performed with four point probe

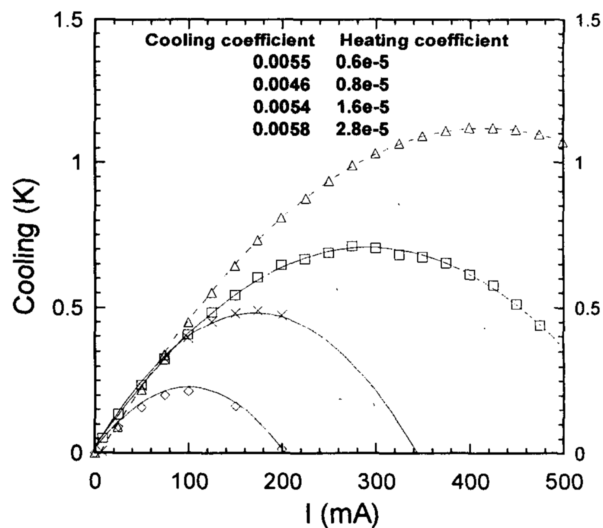

Fig. 3 Measured cooling for a 1μm heterostructure integrated thermionic (HIT) cooler (InGaAs/InGaAsP Superlattice) with various packages. The cooling (linear) and heating (quadratic) coefficients correspond to a second order polynomial fit. All temperatures are relative to the value at zero current with a heat sink temperature of 20 °C.

measurements on transmission line model (TLM) patterns [12]. The lowest measured value of specific contact resistance was roughly 5.5×10^{-7} Ω/cm^2 for alloyed Ni/AuGe/Ni/Au contacts to 0.5 µm thick n+ InGaAs (2×10^{19} cm^{-3}). Similar analyses were also performed on other thin film cooler material systems. It was determined in the InGaAs alloyed contacts that rapid heating and cooling produced the lowest contact resistance in agreement with Ogawa [13]. The alloying depth should also be considered when designing thin film coolers. This depth is equal to the amount of semiconductor that mixes with the contact metal and is roughly 0.1-0.2 µm in most III-V systems [13,14]. In our designed thin film coolers, the top n+ or p+ contact regions are as short as possible for low resistance but safely greater than the expected alloying depth. The effects of contact resistance on cooler performance are discussed further in the next section.

Complete Device Simulation

Complete three-dimensional device structures with all non-ideal effects included are simulated to fit experimental data and determine which areas of the thin film cooler designs need to be improved. These simulations model self consistently the thermal and electrical operation, thus requiring knowledge of numerous material properties. To make this process easier and the model as accurate as it can be, all possible device parameters are measured experimentally in order to leave the fewest number of fitting parameters. Once the simulation is in agreement with experimentally measured temperature profiles, particular non-ideal effects can be removed one at a time and the dominating ones determined.

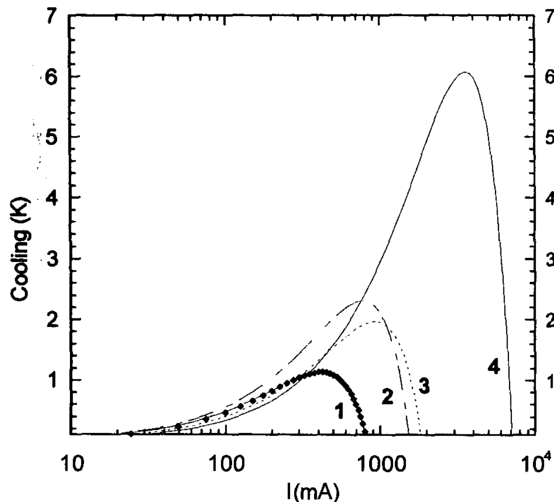

Fig. 4 Experimental and simulation results for the 1µm HIT cooler shown in fig. 3. The points and solid line (1) correspond to the simulated fit of the experimental data with all non-ideal effects. The dashed (2) and dotted (3) curves are the repeated simulation results when the contact resistance and substrate thermal resistance are taken away respectively. The solid line (4) corresponds to both non-ideal effects removed.

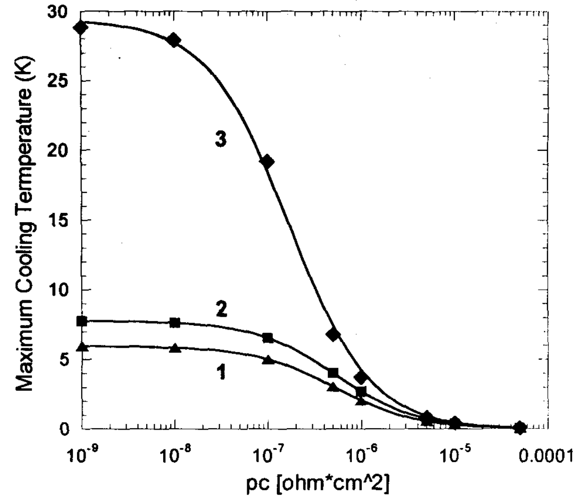

Fig. 5 Simulated maximum cooling temperature versus contact resistance for an optimized, 3 µm thick, thin film cooler for (1) all non-ideal effects included, (2) reduced substrate thermal resistance, and (3) with a shorter wire bond. The points are simulation results while the solid curves are fitted using a derived analytical expression.

Measured cooling for a 1µm HIT cooler (best device from fig. 3) is shown with the simulated curve in fig. 4 (curve 1). With an accurate model of the device in hand, the contact resistance was set to zero and the simulation was repeated (curve 2). The maximum cooling temperature increased from 1.14 °C to 2.3 °C. This was repeated with the contact resistance reset to its original value and the InP substrate replaced with a 10 µm copper substrate resulting in a similar improvement in performance (curve 3). The simulation was then again performed with the top wire bond removed (not shown), resulting only in a very small increase in cooling. Finally, the simulation was repeated once more with both the contact resistance removed and the Copper substrate resulting in a maximum cooling temperature of 6.07 °C or a cooling power density of 1821 W/cm^2 (curve 4). Therefore in this particular device structure, both the contact resistance and the substrate thermal resistance will need to be reduced to see substantial improvement.

The preceding illustrates the fact that decreasing only one non-ideal effect will increase device performance only until some other non-ideal effect becomes dominating. To see this relation better, another more optimized thin film cooler structure is simulated in fig. 5. The data points are simulation results while the solid curves are fitted using the following derived equation for maximum cooling temperature relative to the heat sink:

$$\Delta T_{max} := \frac{\left[\left(\phi_{TI} + S_M \cdot T_c\right)\cdot\left(Rth_d + Rth_{sub}\right) - S\cdot Rth_{sub}\cdot T_h\right]^2}{4\cdot\left(\frac{R_w}{2} + Rc\right)\cdot\left(Rth_d + Rth_{sub}\right)}$$

where ϕ_{TI} is the thermionic cooling coefficient (in the limit of Boltzmann statistics, it is (ϕ_B+2$k_B T_c$/e) where ϕ_B is the barrier height, k_B is Boltzmann's constant, and e is the charge of an electron), S_M is the metal-semiconductor Seebeck coefficient, T_c and T_h are the cold junction and hot junction temperatures, Rth_d and Rth_{sub} are the device and substrate thermal resistance, and R_w and R_c are the wire and contact resistance respectively. The first curve includes all of the same non-ideal effects previously discussed. In the second curve the InP substrate was again replaced with a 10μm thick copper substrate, and maximum cooling was seen to increase by a few degrees for low values of contact resistance. The wire bond was then removed and a dramatic increase in cooling was observed. In each case the three curves approached the same limit when high values of contact resistance became dominating. Experimental testing of the optimized thin film structure will be necessary to confirm the simulation predictions and to determine whether the wire bonding is indeed the dominating non-ideal effect. Better packaging, integration with devices, and conventional TE configurations should approximate the removal of the wire bond for increased cooling performance.

Conclusions

Important parameters and non-ideal effects in thin film coolers have been discussed through experimental and simulation results. A three-dimensional finite element simulation has been developed and used to determine the dominating non-ideal mechanisms for thin film coolers and the impact of changing device characteristics. Contact resistance, finite thermal resistance of substrate and heat sink, and heat generation in wire bonds have all been identified as limitations in thin film cooler performance. Experimental results in thin film thermionic emission coolers have demonstrated cooling by 1.1-1.2 °C with cooling power densities of several 100's W/cm^2, and simulations have predicted cooling of 20-30 degrees with cooling power densities of several 1000's W/cm^2 for more optimized structures and packaging.

Acknowledgments

This work was supported by DARPA-HERETIC program and the Office of Naval Research.

References

1. Shakouri A., LaBounty C., Abraham P., Piprek J., Bowers J.E., "Enhanced Thermionic Emission Cooling in High Barrier Superlattice Heterostructures," Mat. Res. Soc. Symp. Proc. Vol. 545, 1999, pp. 449-458.
2. Shakouri A., LaBounty C., Piprek J., Abraham P., Bowers J.E., "Thermionic emission cooling in single barrier heterostructures," Appl. Phys. Lett., Vol. 74, No.1, (1999), pp. 88-89.
3. Chen G., "Thermal conductivity and ballistic-phonon transport in the cross-plane direction of superlattices," Phys. Rev. B (Condensed Matter), Vol.57, (1998), pp. 14958-73.
4. Hyldgaard P., Mahan G.D., "Phonon superlattice transport," Phys. Rev. B, Vol.56, No.17, (1997), pp. 10754-57.
5. ANSYS Release 5.5.3 (1999) by Swanson Analysis Systems, Inc., Houston, PA.
6. Coldren, L., Corzine, S., Diode Lasers and Photonic Integrated Circuits, John Wiley & Sons, Inc. (New York, 1995), pp. 55-57.
7. Masana, F, "A Closed Form Solution of Junction to Substrate Thermal Resistance in Semiconductor Chips," *IEEE Trans-CPMT-A*, Vol. 19, No. 4 (1996), pp. 539-545.
8. Fan J.C., Lee C.P., Tsai C.M., Wang S.Y. Tsang J.S., "Optical and structural properties of epitaxially lifted-off GaAs films," Journal of Applied Physics, Vol.83, No.1, (1998), pp. 466-468.
9. Morrison C.B., Strijek R.L., Bechtel J.H., "New processing approach for grafting optoelectronic devices and applications to multichip modules," SPIE-Int. Soc. Opt. Eng, San Jose, CA, Feb. 1997, pp.120-127.
10. Yi-Yan A., Chan W.K., Ravi T.S., Gmitter T.J., Bhat R., Yoo K.H., "Grafted InGaAsP Light Emitting Diodes on Glass Channel Waveguides," Electronic Letters, Vol. 28, No.4, (1992), pp. 341-342.
11. Yablonovitch E., Hwang D.M., Gmitter T.J., Florez L.T., Harbison J.P., "Van der Waals bonding of GaAs epitaxial liftoff films onto arbitrary substrates," Appl. Phys. Lett., Vol. 56, No. 24, (1990), pp.2419-2421.
12. Reeves G.K., Harrison H.B., "Obtaining the Specific Contact Resistance from Transmission Line Model Measurements," Elec. Dev. Lett., Vol. EDL-3, No.5, (1982), pp. 111-113.
13. Ogawa M., "Alloying behavior of Ni/Au-Ge films on GaAs," J. Appl. Phys. Vol.51, No.1, (1990), pp. 406-412.
14. Reeves G.K., Leech P.W., Harrison H.B., "Understanding the Sheet Resistance Parameter of Alloyed Ohmic Contacts Using a Transmission Line Model," Solid-St. Electron., Vol.38, No.4, (1995), pp. 745-750.