Optimization of Thermoelectric Coolers for Hotspot Cooling in Three-Dimensional Stacked Chips

Matthew Redmond
G. W. Woodruff School
of Mechanical Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332
e-mail: mredmond3@gatech.edu

Satish Kumar
G. W. Woodruff School
of Mechanical Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332
e-mail: satish.kumar@me.gatech.edu

Three-dimensional (3D) chip stacking architecture is expected to reduce form factor, improve performance, and decrease power consumption in future microelectronics. High power density and nonuniform power distribution in stacked dies are expected to bring significant thermal challenges for 3D packages due to localized hot spots. Embedded thermoelectric coolers (TECs) have potential to provide reliable and localized cooling at these hot spots. In this work, peak package temperature or active cooling per power consumption of TECs are optimized, considering applied current and thickness of TECs as parameters, for a 3D electronic package with two stacked dies. Each die has two hot spots and one TEC is paired with each hot spot. Three different optimization methods are considered in order to ensure a robust solution. The optimization suggests that both the peak temperature in package and energy efficiency of the cooling system can be significantly improved through the optimization of TECs. TECs are also compared against a configuration where they are replaced by copper blocks or thermal vias. A total of 4.7 °C of additional localized cooling is observed using TECs which is beyond what is achievable with copper vias in place of the TECs. The study also suggests that it is better to use TECs to cool only the hottest portions of the package to avoid introducing additional thermal resistance and Joule heating in the package. [DOI: 10.1115/1.4028254]

Keywords: thermoelectric coolers, optimization, 3D-ICs, thermal vias, embedded

1 Introduction

As microelectronic device size continues to be reduced, a physical limit to the component size will eventually be reached. Current research efforts focus on new ways to improve the performance of microelectronic devices without reducing component size. One such avenue, 3D stacked chips, involves stacking several active dies interconnected through silicon vias. This novel architecture may reduce microelectronic form factor, improve performance by reducing interconnect delay, and provide a more energy-efficient chip design [1].

However, the electrical and packaging benefits of 3D stacked chips may be offset by thermal concerns. It is expected that widespread adoption of 3D technology will bring severe thermal management challenges at the package level beyond the challenges already encountered by two dimensional chips [2]. The 2011 International Technology Roadmap for Semiconductors projects that average power dissipation on chip will soon reach up to 200 W/cm^2 and local hotspot heat fluxes up to 500 W/cm^2 [3,4]. These hotspots can degrade chip performance and reliability and can bring additional restrictions in chip design [5]. Effective thermal management of chips requires heat removal of the overall power dissipation and also the efficient control of hot spots due to the nonuniform power dissipation [6].

Typically, heat is removed from a microelectronic package using a heat spreader paired with an air cooled heat sink. Many other cooling technologies have been investigated; each one has advantages and disadvantages relative to conventional air cooling [6,7]. Large scale adaptation of cooling technologies is dependent on its reliability, energy efficiency, cost effectiveness, and form factor in comparison to traditional air cooling. TECs have the potential to meet these requirements. The suitability of a material for TE applications is quantitatively described by its nondimensional figure of merit zT, shown below:

$$zT = \frac{\alpha^2 \sigma T}{k} \tag{1}$$

In this equation, zT is the nondimensional figure of merit of a material, α is the Seebeck coefficient, T is temperature, and k is the thermal conductivity. A higher figure-of-merit of a TE material leads to more efficient cooling by TECs and more efficient power generation by thermoelectric generators [8].

In modern microelectronics, hotspot temperature often governs chip design and reliability. Using microscale TECs to control the hotspot temperature in combination with air cooling at the package level can be an appealing approach as it may turn into an efficient cooling solution in terms of both power consumption and cost [9,10]. Furthermore, microscale TECs have an advantage of high heat flux pumping capacity as this is inversely proportional to the thickness of the TE material used in the TECs [9]. A variety of microscale TEC configurations for on-chip hotspot cooling have been demonstrated by the previous studies: they have been recessed into the heat spreader [10], placed in contact with the chip [11], placed on the back-side of the chip [12], and integrated into the package with microchannels [13].

Regardless of the TEC configuration used, it is imperative to optimize the TE device [9]. The importance of optimization has been demonstrated by previous researchers by comparing performance of off-the-shelf components with optimized components [14,15]. In a typical TEC, the independent variables for an optimization include the TE material, TE element thickness, number of TE couples, packing density of TE couples, and operating current. The previous studies have observed that the number of TE couples

Contributed by the Electronic and Photonic Packaging Division of ASME for publication in the Journal of Electronic Packaging. Manuscript received February 6, 2014; final manuscript received August 7, 2014; published online October 6, 2014. Assoc. Editor: Madhusudan Iyengar.

does not significantly affect the performance of TECs, except to adjust the operating current, voltage, and packing density of the couples [16,17]. The packing density of the couples should be maximized to the extent allowed by current, voltage, and manufacturing constraints.

Normally, optimization studies seek to either maximize coefficient of performance (COP), or minimize the cold side temperature; however, it is possible to simultaneously maximize the COP and minimize the cold side temperature when both the heat load and heat sink thermal resistance are known and constant [18]. Although less common, optimizations have been also performed to minimize the operational costs of a TEC [19]. The effects of thermal and electrical contact resistances, which become important for microscale TECs, have also been accounted for in recent optimization studies [16,20,21]. However, it is important to note that all the optimization studies referenced here assume the sides of the TEC are thermally insulated and that the heat flux through the TEC is uniform. While the resulting optimizations are still very useful, these assumptions are not valid when embedded TECs are used for hotspot thermal management and 3D effects become relevant. A new optimization approach is needed for embedded TECs, which considers 3D effects.

Furthermore, little attention has been given to the application of TECs for thermal management of stacked chip architecture. One study envisioned the used of TE heat spreading using lateral Peltier devices in stacked chips [22]. In another study, a TEC is combined with a silicon interposer to cool hotspots on chip [23]. This study concluded that at low power dissipations the TEC can improve the cooling in a stacked chip, but at high power dissipation levels it would be better to use a copper spreader without a TEC to cool the package. A third study examined the effects of thermal contact resistance of TECs embedded in a stacked chip package in both steady state and transient operation [24]. To our knowledge, the optimization of TECs for stacked chip architecture has not been performed to date.

In this study, the TE material thicknesses and current magnitudes were considered as variables for the performance optimization of TECs embedded in a 3D stacked chip package. Four independent variables were considered: top TEC current (I_t), bottom TEC current (I_b), top TE material thickness (t_t), and bottom TE material thickness (t_b). These variables were chosen because TEC current and TE material thickness represent the primary design parameters for a TEC, aside from material selection, and are frequently optimized for TECs [16,18–21]. A variety of optimization methods are compared and analyzed with regards to their solution speed and accuracy. Two different cases of optimization were considered, each with a unique objective. In the first case, the objective was to minimize the maximum temperature occurring at any point in the computational domain of the 3D package. In the second case, a combined temperature and power optimization was performed using the active cooling divided by the energy consumption of the TECs as the objective function.

2 Computational Methodology

In order to optimize embedded TECs, a computational model of four TECs integrated into a stacked chip electronic package was developed using the commercial finite element software COMSOL; the geometric details and material properties are based on our previous study in Ref. [24]. It should be noted that the configuration of a 3D stacked chip and methodology of chip bonding may vary depending on the application. The investigation of configuration and bonding of stacked dies is still an area of active research. The commercial applications of 3D technology are currently limited to imaging sensors and dynamic random access memory (DRAM). Even so, it is expected that there will be sufficient space to mount microscale TECs on the different dies of a 3D package. There are numerous die bonding techniques available for 3D stacking, all of which leave a gap between the two dies ranging from 2 to 70 μm [25,26]. This is in the range of thicknesses for state-of-the-art microscale TE devices, which have total device thicknesses of less than 100 μm with TE material thickness ranging from 0.5 to 25 μm [9,12,27–32]. If TECs were fabricated directly on-chip using MEMS processes, the total thickness of the TEC can approach the TE material thickness.

A schematic of the electronic package, TECs, and hotspot locations are shown in Fig. 1. Four TECs, each 100 μm thick and composed of 7×7 p–n couples, are paired with two hotspots on the bottom and two hotspots on the top chip. It should be noted that the top TECs cool both hotspots, but the bottom TECs cool the bottom hotspot but heat the top hotspot. The bonding method of the chips is left unspecified because bonding in 3D stacked chips has been proposed using many different methods [25,26]. For generality, the bond is represented by a thin effective resistance layer (ERL) and an infill layer. The infill is presumed to be a compliant polymer, similar to modern underfill or thermal interface material (TIM), which enhances both mechanical stability and thermal performance of the electronic package. The TE properties ($k = 1.2$ W/m-K, $\alpha = 300\,\mu$V/K, and $\sigma = 1.08 \times 10^{-5}$ Ω-m) are taken from the experimental measurement in Ref. [12]. In each TEC module, a Bi_2Te_3 based superlattice TE material is sandwiched between two copper layers. The dimensions of the TEC module are held constant at $3 \times 3 \times 0.1$ mm^3 even though the thickness of the TE material is variable. Electrical and thermal resistances at the interface of the TE-copper layers (10^{-11} Ωm^2, 1×10^{-6} m^2 K/W), inside the TEC module, and at the interface of the TEC-spreader (8×10^{-6} m^2 K/W) are also included in the model, based on the values reported in Ref. [12]. Selected dimensions and material properties for the model are listed in Table 1.

In the model, the substrate is considered as an adiabatic interface, and the heat sink is represented by a 2050 W/m^2-K convective boundary condition with ambient temperature of 300 K for computational simplicity. The four hotspots are represented by four high heat flux sources of magnitude 1000 W/cm^2 and area 500 μm × 500 μm located at the bottom of their respective chips. A uniform heat flux of 14.5 W/cm^2 is considered for the rest of the chip area. The total power dissipation is 29.0 W per chip, or 58.1 W for the electronic package. The effective heat flux of the top and bottom chips are described by Eq. (2).

$$q'' = \begin{cases} 1000 \dfrac{w}{\text{cm}^2} & \text{at the hotspots} \\ 14.5 \dfrac{w}{\text{cm}^2} & \text{elsewhere} \end{cases} \tag{2}$$

Peltier cooling is a surface effect which is incorporated by adding heat ($\sim \alpha I T_h$) at the hot side and subtracting heat ($\sim \alpha I T_c$) from the cold side of the TE material. Here, T_h and T_c are the hot and cold junction temperatures and I is the applied current. Joule heating inside the TEC module and at the interfaces (∼electrical contact resistance at TE-copper interface) is modeled by adding source terms of magnitude I^2R at the corresponding volumes and interfaces.

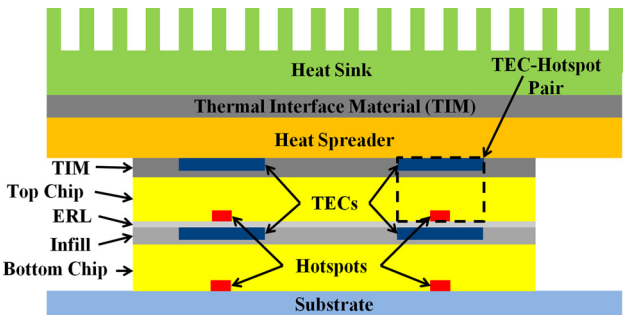

Fig. 1 Schematic of the electronic package. The computational domain includes the heat spreader, chips, TIM, ERL, infill, hotspots, and TECs [24].

Table 1 Dimensions and material properties for the stacked chip TEC model [23]

Components	Thermal conductivity (W/m-K)	Dimensions (mm × mm × mm)
ERL	5	12 × 11 × 0.025
Chip	140	12 × 11 × 0.5
Heat spreader	400	30 × 30 × 1
Infill and TIM	1.75	12 × 11 × 0.125

The Peltier cooling and Joule heating effects are implemented in the model with a volumetric heating/cooling approach. Although the Peltier effect and the Joule contact heating effect are interfacial phenomena, they are represented by thin volume elements on either side of the TEC in our model. The total power dissipation due to Joule heating in the bulk TE material is defined by Eq. (3). The Peltier effect is implemented using a negative heat source at the cold side of the TEC and a positive heat source at the hot side of a TEC according to Eqs. (4) and (5). The Joule contact heating for a single contact is given by Eq. (6)

$$P_{\text{bulk}} = \frac{I^2(2N)^2 t}{A\Phi\sigma} \tag{3}$$

$$P_{\text{c}} = -2N\alpha I T_{\text{c}} \tag{4}$$

$$P_{\text{h}} = +2N\alpha I T_{\text{h}} \tag{5}$$

$$P_{\text{contact}} = \frac{I^2(2N)^2 R''_{\text{elec}}}{A\Phi} \tag{6}$$

In these equations, I is applied electrical current, N is the number of TE elements, t is the TE material thickness, A is the total TEC area, Φ is the packing factor, σ is the electrical conductivity of the TE material, α is the Seebeck coefficient, T_{c} is the cold side temperature, T_{h} is the hot side temperature, and R''_{elec} is the electrical contact resistance.

This method of modeling a TEC has been validated against the experimental results for a 2D electronic package using the commercial software FLUENT [33]. An identical 2D electronic package model was developed in COMSOL and validated against the FLUENT model in Ref. [33] and the hotspot cooling experiment and model described in Ref. [12]. The advantage of using COMSOL in this application is that it is capable of parametric variation of the TE material thickness. The hot spot temperature estimated from different models and the experimental results from Ref. [12] for different applied current to the TEC is shown in Fig. 2 to validate the COMSOL model.

The present computational model for the 3D stacked chip package contains 69,695 quadratic finite elements. Grid independence tests are performed using 176,208 elements which yield very small changes (<1%) in temperature distribution and verify that 69,695 elements are sufficient for the numerical simulations. A model of this size is small enough to be solved on a typical desktop PC in only a few minutes, which greatly simplified the optimization process.

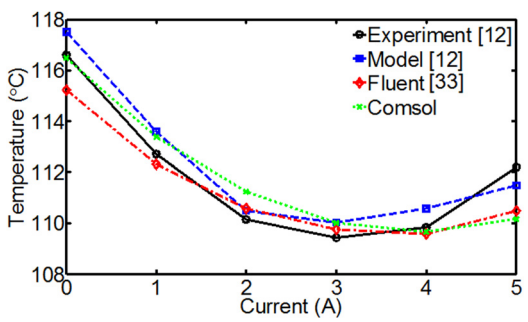

Fig. 2 Comparison of the hot spot temperature for various applied current magnitudes to the TEC. Results from the model developed in COMSOL compared against the experimental data and modeling results reported in Refs. [12] and [33].

A large variety of optimization methods exist, but two of the most common methods include the gradient descent method and Newton's method [34]. The gradient descent method is a first order method and uses a function's gradient to find local extrema, while Newton's method is a second-order method, which uses the second derivative of a function to find its local extrema. Both these methods are very effective for optimization problems where the gradient is continuous and differentiable. However, for non-continuous or nondifferentiable gradients, the optimization method can oscillate around the extrema or become unstable. An alternative class of "direct search" optimization methods also exists. These methods can be especially useful for problems where the objective function or its gradient is nondifferentiable or discontinuous. One of the simplest methods to implement is the Luus–Jaakola method [35]. In this method, the solution space is sampled and the size of the space is incrementally decreased. As the sample space size is decreased, the center of the sample space tracks the minimum or maximum point which has been sampled so far.

In this work, both the gradient descent method and the Luus–Jaakola method will be used. The gradient descent method is computationally efficient, but is limited in accuracy if the objective function has a discontinuous derivative or if there are multiple local minima. The Luus–Jaakola method is more computationally expensive, but also more robust, because discontinuous gradients do not affect the solution and the entire solution space is sampled. This increases the likelihood that the global, rather than local, extrema are found.

3 Numerical Optimization of TE Material Thickness and Current Magnitude

An optimization of the TE material thickness and current magnitude was performed using COMSOL. Four independent variables were considered for the optimization: top TEC current (I_{t}), bottom TEC current (I_{b}), top TE material thickness (t_{t}), and bottom TE material thickness (t_{b}). Two different objective functions for the optimization were considered. In the first case, the objective was to minimize the maximum temperature occurring at any point in the computational domain. This maximum temperature always occurred on either the top or bottom hotspot. In the second case, a combined temperature and power optimization was performed using the active cooling divided by the energy consumption of the TECs as the objective function. Active cooling was defined as the decrease in the maximum temperature beyond that achievable with 100 μm copper blocks acting as thermal vias. Before the optimization was performed, a simulation was performed where all four TECs were replaced by 100 μm thick copper thermal vias. The maximum chip temperature for this configuration was 110.84 °C. This scenario represents purely passive thermal control and is a benchmark to which all TEC run cases are compared in this paper.

Constraints were imposed on the optimization problem in order to ensure that results were physically realistic and to reduce the size of the parametric space. The current through the TECs was limited to between 0 and 3 A because extremely large currents are difficult to supply to TECs. The thickness of the top and bottom TE material was constrained by a lower limit of 0 μm and an upper limit of 50 μm. Even though the total TEC thickness is 100 μm, it is expected that about 25 μm copper on each side of the TEC would be necessary for structural integrity in the manufacturing approach taken in Ref. [12]. A special case of this four variable optimization occurs when the thickness of the top TECs is equal to zero. In this case, the top TECs are effectively replaced by 100 μm copper thermal vias and the bottom TECs can be

optimized for two variables alone: bottom TEC current (I_b), and bottom TE material thickness (t_b).

The resolution considered for the parametric space of optimum operating current and TE material thickness is 0.05 A and 1 μm, respectively. With these resolutions, the top and bottom TEC optimization case (four independent variables) has 9×10^6 possible combinations, but the special case when the thickness of the top TECs equals zero only has 3000 combinations. A parametric sweep can be performed to find the true optimum at the desired resolution for this special case (bottom TECs only). This special case serves as a useful tool for determining an appropriate optimization method for the case containing both top and bottom TECs. This is then compared to the results of the gradient descent and Luus–Jaakola methods. A summary of the optimizations performed in this paper are shown in Table 2. A total of two objective functions (maximum temperature and active cooling/power), two geometries (bottom TECs only and top and bottom TECs), and three optimization methods (parametric sweep, gradient descent, and Luus–Jaakola) were used. Of all the possible combinations, only one case corresponding to the parametric sweep optimization for the top and bottom TECs geometry was not considered in the present study. This case has 9×10^6 points in its parametric space, which will make the optimization process very expensive.

3.1 Bottom TEC Only Optimization. For the special case where the thickness of the top TE material thickness is equal to zero, the gradient descent method, Luus–Jaakola direct search method, and a parametric sweep of the solution space were all used to optimize the variables according to both objective functions. In the gradient descent method, the initial guess was 1.75 A and 8 μm for current and TE material thickness, respectively, because these values were used extensively in Ref. [24]. From this starting point, the partial derivative of the objective function with respect to each independent variable was calculated and the independent variables were subsequently changed by taking steps in the direction of the negative gradient of the objective function until a minima was reached. For the Luus–Jaakola method, the parametric space or solution space was sampled with a total of 25 points (five in each dimension). The solution space dimensions were reduced in size by 33% in each direction and centered on the extrema found in the previous step. This process was repeated until the specified resolution was achieved. The parametric sweep used a grid with a resolution of 0.05 A and 1 μm to sweep the solution space in search of an extrema.

The objective of the first optimization case is to minimize the maximum chip temperature. Results from each of the three optimization methods were obtained and are shown with a maximum temperature contour plot in Fig. 3. Out of all three methods, the lowest maximum temperature was found to be 106.14 °C with a bottom TEC current of 1.18 A and a bottom TEC thickness of 38.8 μm by the Luus–Jaakola method. This corresponds to 4.7 °C of cooling beyond what is achievable with copper thermal vias; this is significantly better than the cooling (2.05 °C) reported in Ref. [24] without any optimization.

Even though the independent variables estimated for the lowest maximum temperature by all three methods were quite different, the maximum temperature in all three solutions is similar. This is

Table 2 Summary of the optimizations performed

	Bottom TECs only		Top and bottom TECs	
	Maximum temperature	Active cooling/power	Maximum temperature	Active cooling/power
Parametric sweep	Yes	Yes	No	No
Gradient descent	Yes	Yes	Yes	Yes
Luus–Jaakola	Yes	Yes	Yes	Yes

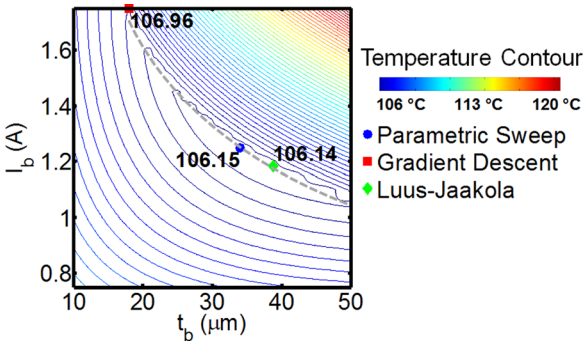

Fig. 3 Contour plot of the objective function (maximum temperature) for the temperature only optimization. Maximum temperature (°C) for the optimum point found by the gradient descent method, Luus–Jaakola method, and parametric sweep have been indicated. The *y*-axis (I_b) is the applied current to the bottom TEC, and the *x*-axis (t_b) is the thickness of the bottom TEC. The dashed line represents the thin region where the lowest maximum temperature occurs, and the top and bottom hot spot temperatures are equal.

because the lowest maximum temperature occurs along a long and thin region where the top and bottom hot spot temperatures are the same, represented by a dashed curve in Fig. 3. This agrees with the expected behavior since the objective was to minimize the maximum temperature of the package, which occurs when the top and bottom hotspots are equal in temperature. Because the bottom TEC cools the bottom hotspot and heats the top hotspot, having top and bottom hotspots with different temperatures is evidence of a nonoptimal system. In an optimal system, the bottom TEC is tuned so that it cools the bottom hotspot and heats the top hotspot just enough so that their temperatures are the same. A parametric sweep of the solution space revealed that the bottom and top hotspot temperatures can be represented by two smooth surfaces which intersect each other. The objective function is the maximum of these two surfaces, and is nondifferentiable at the surfaces' intersection. This intersection is shown in Fig. 4. This gives insight as to why the gradient descent method gave the least accurate optimization, since the gradient was discontinuous in the region of the lowest maximum temperature.

The objective of the second optimization case is to maximize the active cooling divided by power consumption. Active cooling is defined as the decrease in the maximum chip temperature which can be achieved with TECs compared to thermal vias, where the TEC is replaced by a copper block. The power consumption is defined as the total power consumed by all TECs in the package. In this case, the results obtained by the gradient descent method, Luus–Jaakola method, and parametric sweep were all within the study's resolution of 0.05 A and 1 μm of each another. The optimum points, along with a contour plot of the objective function, are shown in Fig. 5. In this region, the temperature of the top hotspot is less than the temperature of the bottom hotspot. The objective function is smooth because the optimum point occurs away from the region where the top and bottom hotspot temperatures are equal. Since the gradient is continuous, both the gradient descent and Luus–Jaakola methods achieve similar accuracy. A total of 2.32 °C of cooling is achieved for each Watt of power consumed in the package. The total cooling is 1.13 °C higher than what is achievable with copper thermal vias.

3.2 Top and Bottom TEC Optimization. When considering both the top and bottom TECs simultaneously, a parametric sweep is not feasible because the solution space is very large. For the temperature-only optimization with top and bottom TECs, the gradient descent method optimization was attempted; however, it was found to yield inaccurate optimization results when compared to the Luus–Jaakola method, as would be expected. This is due to

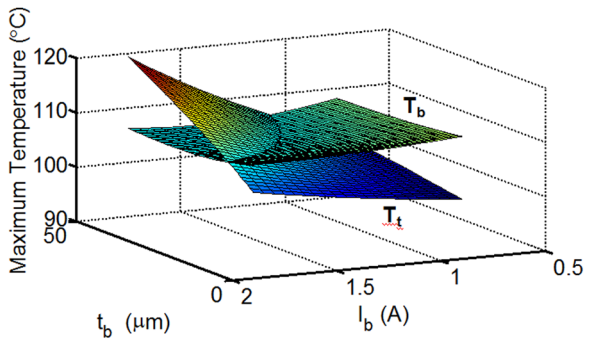

Fig. 4 Intersection of the top and bottom hotspot temperatures in the region of the optimum solution. I_b is the applied current to the bottom TEC, t_b is the thickness of the bottom TEC, T_t is the peak temperature on the top chip, and T_b is the peak temperature on the bottom chip.

the discontinuous derivative which occurs in the region where the top and bottom hotspots are equal, as discussed in Sec. 3.1. As a result, the Luus–Jaakola method was relied upon for the temperature only optimization considering both the top and the bottom TECs.

In order to achieve high confidence in the optimization results, the Luus–Jaakola optimization method was performed using two different criteria. In the first (criterion A), a total of 625 points were sampled in the solution space (five in each direction), and the solution space dimensions were decreased by 33% in each iteration. In the second (criterion B), a total of 256 points were sampled in the solution space (four in each direction), and the solution space dimensions were decreased by 25% in each iteration.

Both sets of criterion produced very similar optimization results, even though the path to the solution for each criterion was quite different. The path to solution along with the result for each criterion is shown in Fig. 6. There is high degree of confidence in the robustness of this solution since both optimizations yielded similar results. The lowest maximum chip temperature for the optimum point is 108.08 °C, which corresponds to 2.76 °C of cooling beyond what is achievable with copper thermal vias. The total power consumed by the TECs is 6.27 W.

When the active cooling per power consumption is optimized for the top and bottom TECs simultaneously, both the gradient descent and Luus–Jaakola methods can be utilized since the optimum point should occur away from the region where the top and bottom hot spot temperatures are equal. When this optimization is performed, both the gradient descent and the Luus–Jaakola methods indicate that the optimum TE material thickness of the top TEC approaches 0 μm. This solution represents the special case which was discussed in detail in Sec. 3.1. This result suggests that

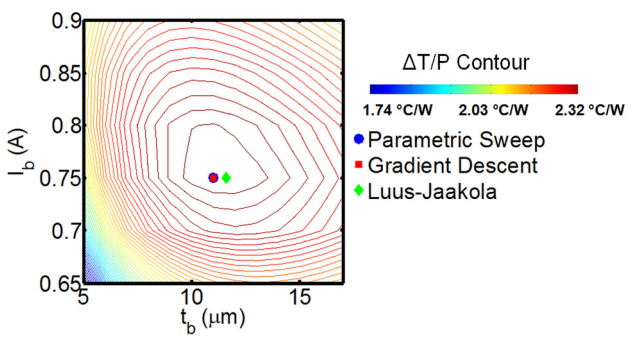

Fig. 5 Contour plot of the objective function (active cooling divided by power consumption). Optimum points found by the gradient descent method, Luus–Jaakola method, and parametric sweep have been indicated.

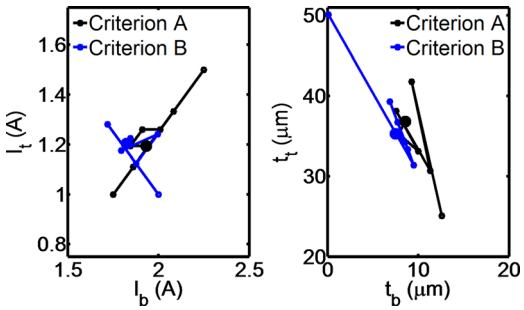

Fig. 6 The temperature only optimization path for the Luus–Jaakola method using two different criteria is shown on two plots. Two plots are necessary to show the solution path because a total of four independent variables are optimized simultaneously. Two of the four independent variables make-up the axes on each plot. These variables are top TEC current (I_t, left), bottom TEC current (I_b, left), top TE material thickness (t_t, right), and bottom TE material thickness (t_b, right). The similar results obtained using different criteria and solution paths gives confidence in the robustness of this solution. The small circles represent intermediate steps and the large circles represent the optimum.

it is preferable to use TECs on only the bottom chip when energy efficiency is considered.

3.3 Comparison of the Optimization Results. The results of the optimizations described in Secs. 3.1 and 3.2 are compared and summarized in Fig. 7. Using a copper block in place of the top TEC reduced the maximum chip temperature and increased the cooling efficiency. This result suggests that using TECs on each layer of a stacked chip can actually be detrimental for the package. This holds true when considering both maximum cooling and cooling per power consumption (∼energy-efficient cooling) as an objective function. TECs should be used primarily to cool only the hottest portion of a package, which would be the bottom chip in a stacked chip package. It is expected that as the number of chips in the package increases past two, this trend will not change. One interesting result of this optimization is that for the temperature only optimization with TECs on the top and bottom (Fig. 7(*a*)), the top TEC has a relatively small TE thickness and large

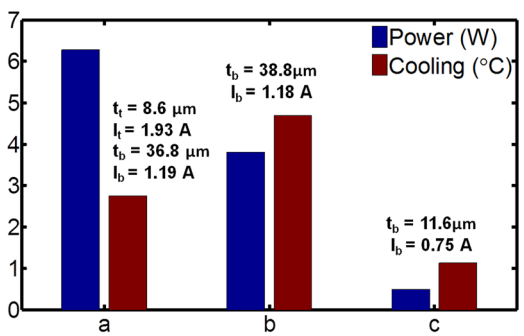

Fig. 7 Hot spot cooling using TECs in comparison to using 100 μm thick copper blocks in place of all four TECs and the total power consumption in all TECs. The optimums are labeled with the corresponding independent variables. (*a*) Temperature-only optimization with TECs on the top and bottom. (*b*) Temperature-only optimization with copper blocks on the top chip and TECs on the bottom chip. (*c*) Combined temperature and power consumption optimization with copper blocks on the top chip and TECs on the bottom chip. The combined temperature and power consumption optimization with top and bottom TECs (not shown) suggests that the thickness of the top TEC should be zero, which is the same as the result for (*c*).

applied current while the bottom TEC has a relatively large TE thickness and small applied current. This makes physical sense since smaller TE thickness and larger applied currents are known to be optimal for higher heat flux TECs (top TEC in present case), while larger TE thickness and smaller applied currents known to be optimal for lower heat flux TECs (bottom TEC in present case).

4 Conclusion

This study optimized the TE material thickness and current applied to TECs for hotspot cooling in stacked chip packaging. The optimization suggests that both the package temperature and energy efficiency of the cooling system can be improved by replacing the top TEC with a copper thermal via. It is essential to use TECs to primarily cool only the hottest portions of the package to avoid introducing additional thermal resistance and Joule heating in the package. For stacked chips, configurations where TECs cool only the hottest chip in the package can sometimes be superior to other configurations even though other components in the package can increase in temperature. This study also provided insights into various optimization methods which may be used to optimize TECs embedded in a 3D stacked chip package. Although the gradient descent method yielded satisfactory results with minimal computational expense, the direct search Luus–Jaakola method yielded improved accuracy with a computational cost that can be managed by a single desktop computer. Future work could include performing optimizations of additional variables which were beyond the scope of this work including TEC footprint area and optimizations related to transient pulse cooling using TECs.

Acknowledgment

The authors would like to thank the Turkish Fulbright Commission and National Science Foundation Grant No. ECCS-1028569 for supporting part of this work.

References

[1] Black, B., Annavaram, M., Brekelbaum, N., DeVale, J., Jiang, L., Loh, G. H., McCauley, D., Morrow, P., Nelson, D. W., Pantuso, N., Reed, P., Rupley, J., Shankar, S., Shen, J., and Webb, C., 2006, "Die Stacking (3D) Microarchitecture," 39th IEEE/ACM International Symposium on Microarchitecture (MICRO-39), Orlando, FL, Dec. 9–13, pp. 469–479.

[2] Jain, A., Jones, R. E., Chatterjee, R., and Pozder, S., 2010, "Analytical and Numerical Modeling of the Thermal Performance of Three-Dimensional Integrated Circuits," IEEE Trans. Compon. Packag. Technol., **33**(1), pp. 56–63.

[3] ITRS, 2011,"The International Technology Roadmap for Semiconductors," *System Drivers*, Semiconductor Industry Association (SIA), Washington, DC, pp. 1–29.

[4] ITRS, 2009, "The International Technology Roadmap for Semiconductors," *Assembly and Packaging*, Semiconductor Industry Association (SIA), Washington, DC, pp. 1–62.

[5] Hamann, H. F., Weger, A., Lacey, J. A., Hu, Z., Bose, P., Cohen, E., and Wakil, J., 2007, "Hotspot-Limited Microprocessors: Direct Temperature and Power Distribution Measurements," IEEE J. Solid-State Circuits, **42**(1), pp. 56–65.

[6] Mahajan, R., Chiu, C.-P., and Chrysler, G., 2006, "Cooling a Microprocessor Chip," Proc. IEEE, **94**(8), pp. 1476–1486.

[7] Mudawar, I., 2001, "Assessment of High-Heat-Flux Thermal Management Schemes," IEEE Trans. Compon. Packag. Technol., **24**(2), pp. 122–141.

[8] Snyder, G. J., and Toberer, E. S., 2008, "Complex Thermoelectric Materials," Nature Mater., **7**(2), pp. 105–114.

[9] Sharp, J., Bierschenk, J., and Lyon, J. H. B., 2006, "Overview of Solid-State Thermoelectric Refrigerators and Possible Applications to On-Chip Thermal Management," Proc. IEEE, **94**(8), pp. 1602–1612.

[10] Snyder, G. J., Soto, M., Alley, R., Koester, D., and Connor, B., 2006, "Hot Spot Cooling Using Embedded Thermoelectric Coolers," 22nd IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM), Dallas, TX, Mar. 14–16, pp. 135–143.

[11] Yang, B., Wang, P., and Bar-Cohen, A., 2007, "Mini-Contact Enhanced Thermoelectric Cooling of Hot Spots in High Power Devices," IEEE Trans. Compon. Packag. Technol., **30**(3), pp. 432–438.

[12] Chowdhury, I., Prasher, R., Lofgreen, K., Chrysler, G., Narasimhan, S., Mahajan, R., Koester, D., Alley, R., and Venkatasubramanian, R., 2009, "On-Chip Cooling by Superlattice-Based Thin-Film Thermoelectrics," Nat. Nanotechnol., **4**(4), pp. 235–238.

[13] Sahu, V., Joshi, Y. K., and Fedorov, A. G., 2010, "Experimental Investigation of Hotspot Removal Using Superlattice Cooler," 12th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), Las Vegas, NV, June 2–5.

[14] Bierschenk, J., and Johnson, D., 2004, "Extending the Limits of Air Cooling With Thermoelectrically Enhanced Heat Sinks," 9th Intersociety Conference on Thermal and Thermochemical Phenomena in Electronic Systems (ITHERM'04), Las Vegas, June 1–4, pp. 679–684.

[15] Chu, R. C., and Simons, R. E., 1999, "Application of Thermoelectrics to Cooling Electronics: Review and Prospects," 18th International Conference on Thermoelectrics, Baltimore, MD, Aug. 29–Sept. 2, pp. 270–279.

[16] Hodes, M., 2012, "Optimal Design of Thermoelectric Refrigerators Embedded in a Thermal Resistance Network," IEEE Trans. Compon. Packag. Technol., **2**(3), pp. 483–495.

[17] Taylor, R. A., and Solbrekken, G. L., 2008, "Comprehensive System-Level Optimization of Thermoelectric Devices for Electronic Cooling Applications," IEEE Trans. Compon. Packag. Technol., **31**(1), pp. 23–31.

[18] Taylor, R. A., and Solbrekken, G. L., 2006, "Optimization of Thermoelectric Cooling for Microelectronics," 10th Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, (ITHERM'06), San Diego, CA, May 30-June 2, pp. 483–489.

[19] Xuan, X. C., 2002, "Optimum Design of a Thermoelectric Device," Semicond. Sci. Technol., **17**(2), pp. 114–119.

[20] Pettes, A. M., Hodes, M. S., and Goodson, K. E., 2009, "Optimized Thermoelectric Refrigeration in the Presence of Thermal Boundary Resistance," IEEE Trans. Adv. Packag., **32**(2), pp. 423–430.

[21] Hodes, M., "Optimal Pellet Geometries for Thermoelectric Refrigeration," IEEE Trans. Compon. Packag. Technol., **30**(1), pp. 50–58.

[22] Solbrekken, G. L., 2005, "Peltier Enhanced Heat Spreading for Localized Hot Spot Thermal Management," ASME Paper No. IPACK2005-73471.

[23] Li, S.-L., Chung-Yen, H., Chun-Kai, L., Ming-Ji, D., Chieh, C. H., and Ra-min, T., 2011, "Hot Spot Cooling in 3DIC Package Utilizing Embedded Thermoelectric Cooler Combined With Silicon Interposer," 6th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT), Taipei, Oct. 19–21, pp. 470–473.

[24] Redmond, M., Manickaraj, K., Sullivan, O., Mukhopadhyay, S., and Kumar, S., 2013, "Hotspot Cooling in Stacked Chips Using Thermoelectric Coolers," IEEE Trans. Compon. Packag. Manuf. Technol., **3**(5), pp. 759–767.

[25] Knickerbocker, J. U., Andry, P. S., Dang, B., Horton, R. R., Interrante, M. J., Patel, C. S., Polastre, R. J., Sakuma, K., Sirdeshmukh, R., Sprogis, E. J., Sri-Jayantha, S. M., Stephens, A. M., Topol, A. W., Tsang, C. K., Webb, B. C., and Wright, S. L., 2008, "Three-Dimensional Silicon Integration," IBM J. Res. Dev., **52**(6), pp. 553–569.

[26] Dang, B., Wright, S. L., Andry, P. S., Sprogis, E. J., Tsang, C. K., Interrante, M. J., Webb, B. C., Polastre, R. J., Horton, R. R., Patel, C. S., Sharma, A., Zheng, J., Sakuma, K., and Knickerbocker, J. U., 2008, "3D Chip Stacking With C4 Technology," IBM J. Res. Dev., **52**(6), pp. 599–609.

[27] Bottner, H., Nurus, J., Gavrikov, A., Kuhner, G., Jagle, M., Kunzel, C., Eberhard, D., Plescher, G., Schubert, A., and Schlereth, K.-H., 2004, "New Thermoelectric Components Using Microsystem Technologies," J. Microelectromech. Syst., **13**(3), pp. 414–420.

[28] Huang, I.-Y., Lin, J.-C., She, K.-D., Li, M.-C., Chen, J.-H., and Kuo, J.-S., 2008, "Development of Low-Cost Micro-Thermoelectric Coolers Utilizing MEMS Technology," Sens. Actuators A, **148**(1), pp. 176–185.

[29] Fan, X., Zeng, G., Croke, E., LaBounty, C., Ahn, C. C., Vashaee, D., Shakouri, A., and Bowers, J. E., 2001, "High Cooling Power Density SiGe/Si Microcoolers," Electron. Lett., **37**(2), pp. 126–127.

[30] Fan, X., Zeng, G., LaBounty, C., Bowers, J. E., Croke, E., Ahn, C. C., Majumdar, S. H. A. A., and Shakouri, A., 2001, "SiGeC/Si Superlattice Microcoolers," Appl. Phys. Lett., **78**(11), pp. 1580–1582.

[31] Snyder, G. J., Lim, J. R., Huang, C.-K., and Fleurial, J.-P., 2003, "Thermoelectric Microdevice Fabricated by a MEMS-Like Electrochemical Process," Nature Mater., **2**(8), pp. 528–531.

[32] Zhang, Y., Zeng, G., and Shakouri, A., 2006, "Silicon Microrefrigerator," IEEE Trans. Compon. Packag. Technol., **29**(3), pp. 570–576.

[33] Gupta, M. P., Sayer, M.-H., Mukhopadhyay, S., and Kumar, S., 2011, "Ultrathin Thermoelectric Devices for On-Chip Peltier Cooling," IEEE Trans. Compon. Packag. Technol., **1**(9), pp. 1395–1405.

[34] Nesterov, Y., 2003, *Introductory Lectures on Convex Optimization: A Basic Course (Applied Optimization)*, Kluwer Academic Publishers, Norwell, MA.

[35] Luus, R., and Jaakola, T. H. I., 1973, "Optimization by Direct Search and Systematic Reduction of the Size of Search Region," Am. Inst. Chem. Eng. J., **19**(4), pp. 760–766.