2012

Topological Design of Channels for Squeeze Flow Optimization of Thermal Interface Materials

Nikhil Bajaj  
*Purdue University*

Ganesh Subbarayan  
*Purdue University*

S V. Garimella  
*Purdue University, sureshg@purdue.edu*

Follow this and additional works at: [http://docs.lib.purdue.edu/coolingpubs](http://docs.lib.purdue.edu/coolingpubs)

http://dx.doi.org/http://dx.doi.org/10.1016/j.ijheatmasstransfer.2012.03.020

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Topological Design of Channels for Squeeze Flow Optimization of Thermal Interface Materials

Nikhil Bajaj, Ganesh Subbarayan, Suresh V. Garimella

Abstract

Performance of thermal interface materials (TIMs) used between a microelectronic device and its associated heat spreader is largely dependent on the bulk thermal conductivity of the TIM, but the bond-line thickness (BLT) of the applied material as well as the interfacial contact resistances are also significant contributors to overall performance. Hierarchically Nested Channels (HNCs), created by modifying the surface topology of the chip or the heatsink with hierarchical arrangements of microchannels in order to improve flow, have been proposed to reduce both the required squeezing force and the final BLT at the interfaces. In the present work, a topological optimization framework that enables the design of channel arrangements is developed. The framework is based on a resistance network approximation to Newtonian squeeze flow. The approximation, validated against finite element (FE) solutions, allows efficient, design-oriented solutions for squeeze flow in complex geometries. A comprehensive design sensitivity analysis exploiting the resistance network approximation is also developed and implemented. The resistance approximation and the sensitivity analysis is used to build an automated optimal channel design framework. A Pareto optimal problem formulation for the design of channels is posed and the optimal solution is demonstrated using the framework.

Keywords:

1. Introduction

The development of higher performance microelectronic devices that can reliably operate at high speeds with complex functionality is a challenge that is tied to not just the electrical design of such devices, but also the design of the associated thermal management system.
In many electronic devices, the inability to remove heat adequately is a bottleneck to improving performance.

The need for improved thermal management has steadily increased since the introduction of the integrated circuit, and the trend is projected to continue, especially in high-performance computation and power electronics applications. In lower-power embedded and consumer electronics applications, the push for smaller, lighter, and thinner form factors also drives a need for improved thermal management. As there may not be room for large heat sinks and traditional thermal management approaches in this case, creative solutions are necessary to reduce form factors further.

A common thermal solution is illustrated in Figure 1. The choice of heat sink, whether it is air-cooled or cooled via more sophisticated and costly means such as liquid- or phase change-cooling, only addresses a portion of this overall resistance extending from the interface with the device to the ambient. In some packaging arrangements, the thermal interface resistance can account for up to 50% of the total thermal budget [2]. Reducing the interface resistance is thus a critical requirement in such systems.

In Figure 1, TIM1 and TIM2 are intended to improve thermal transport by filling in gaps between structures that exist because of the manufacturing processes used to create the heat spreader, heat sinks, and other devices. These processes typically result in both large scale defects such as surface bowing as well as small-scale defects (asperities, depicted in Figure 1) that limit the actual heat transfer contact area to some small fraction of the nominal value [3]. Previous work has been done to model and experimentally validate predictions of behavior in bare joint contact situations [4] and contact between coated surfaces [5]. Polymeric TIMs, or thermal greases, are a particularly common form of TIMs, as they are relatively easy to apply on a wide variety of surfaces, are removable, and have relatively good thermal performance given their cost. Many greases utilized as TIMs, especially the higher-performance varieties, are particle-filled greases (often of particle size 1 micron or less) in a grease or oil binder. Alumina and other ceramic materials, in addition to metallic
particles such as silver, are common choices for the particles, while silicone and other similar materials are common choices for the grease matrix. The particles are typically present in the fluid at loading volume fractions of 30-60% [6], enabling a considerable improvement of the bulk thermal conductivity of the TIM—often an order of magnitude improvement or more. However, loading the TIM with high volume fractions of particles only helps up to a certain point, after which the changes to the fluid’s rheological behavior become so significant as to cause increases in bond-line thickness that overshadow the gains that increase in the bulk thermal conductivity might provide. This behavior is seen in other particle-filled classes of TIMs as well [7].

The broad contributors to the overall thermal resistance of the interface are commonly understood through the following one-dimensional heat transfer model [8] (see also Figure 2):

\[
R_{\text{interface}} = R_{c1} + \frac{h}{k_{\text{TIM}} A_{\text{interface}}} + R_{c2} 
\]

where, \( R_{c1} \) and \( R_{c2} \) are components of the interfacial contact resistance. The remaining component is clearly dependent on the Bond-Line Thickness (BLT) \( h \), the bulk conductivity of the TIM \( k_{\text{TIM}} \), and the area of the interface \( A_{\text{interface}} \). Reducing \( h \) proportionally decreases this component of the resistance, as does increasing \( k_{\text{TIM}} \) and \( A_{\text{interface}} \).

[Figure 2 about here.]

The mechanism by which \( k_{\text{TIM}} \) is typically increased for a given material set is in conflict with \( h \) because, as the volume fraction of the filler particles increases, the rheological behavior of the fluid changes. While a grease matrix such as silicone oil can have behavior on its own that is well approximated by a Newtonian fluid or a Power Law fluid model, adding particles beyond a critical volume fraction (dependent on the specific material set) makes it more accurately modeled by non-Newtonian constitutive laws that include a yield stress, such as Bingham fluids, or Herschel-Bulkley fluids [9]. A key effect of the presence of a yield stress is that a system undergoing squeeze flow between two parallel plates (as during the assembly of a thermal interface) will have a finite final BLT, as opposed to the zero BLT.
that a model without a yield stress would predict. In an effort to avoid a larger BLT, Brunschwiler et al. [10] proposed the use of “Hierarchically Nested Channels” (HNC) in which channels are cut into the heat spreader surface to decrease transient squeezing forces and final bond-line thickness. This approach has been explored further in [11], and different channel arrangements have been designed and experimentally tested in order to evaluate effectiveness with regards to (among others) both final BLT and improved squeezing speed. Other researchers have also attempted to optimize design parameters such as channel sizing and spacing [12].

Inspired by these complex channel arrangements, in this work, we develop a design procedure for systematically (and formally) optimizing the channel configuration to achieve the optimal trade-off between minimal BLT and acceptable squeezing force. Towards this end, a computationally efficient resistance network model for approximating squeeze flow behavior for Newtonian fluids is developed. The model is integrated with a numerical topology optimization framework in order to generate optimal heat spreader/heat sink surface topologies. The optimization problem is formulated for Newtonian fluids. Analytical design sensitivity procedures are developed to improve solution efficiency, and candidate designs for improved surface topologies are explored.

2. Background and Relevant Literature

The idea of surface modification to improve thermal performance can be addressed as an optimization problem that involves modifying $A_{interface}$ and arranging it in different configurations to modify $h$ and as a whole, reducing $R_{interface}$. Understanding the effect of a surface modification of $A_{interface}$ on the quantity $h$ requires the ability to estimate the changes in the pressure and flow patterns of the material during assembly. For determining the constitutive modeling parameters of such viscous fluids, the use of squeeze-flow rheometry (or plastometry) is often suggested. Early work by Dienes and Klemm in 1946 [13] discussed the characterization of Newtonian fluids. The work of Covey and Stanmore [14] extended the work of Dienes and Klemm further to viscous fluids with a yield stress. The appropriate modeling of squeeze flow has been a topic of discussion in the literature, and as a result,
many different models for the same process exist. A survey of literature on squeeze flow and its applications to rheometry was compiled by Engmann et al. [15].

For a Newtonian fluid in squeeze flow between parallel circular plates, Dienes and Klemm [13] described the squeezing force $F$ in terms of the squeezing rate $\dot{h}$ of the gap $h$ between the two plates as:

$$F = \frac{-3\mu \pi r^4 \dot{h}}{2h^3}$$  \hspace{1cm} (2)

where, $\mu$ is the Newtonian viscosity, and $r$ is the radius of the plates. This model is based on a number of assumptions including, quasi-static flow or very low flow rate, symmetry about the $z$-axis, negligible flow in the $z$-direction, incompressible fluid, no-slip boundary conditions at the walls, and zero back-pressure (the pressure on the cylindrical face of the control volume). The equation may also be manipulated to describe a time-evolving squeeze flow progression of the gap $h$:

$$h(t) = \frac{1}{\sqrt{\frac{4F}{3\pi \mu r^4} + \frac{1}{h_0^2}}}$$  \hspace{1cm} (3)

where, $h_0$ is the initial height from which the time-evolution proceeds. It is important to note that one can easily modify the derivation of Dienes and Klemm [13] for the current system with a non-zero back pressure, $p_0$, resulting in the following relation:

$$F = \frac{-3\mu \pi r^4 \dot{h}}{2h^3} + \pi r^2 p_0$$  \hspace{1cm} (4)

For a given geometry, the viscosity $\mu$ is the key determining factor of the amount of time it takes for squeezing together the two plates (aside from the squeezing force, which in electronics applications is typically limited by the structural limitations). In Figure 3, the effect of changes in viscosity on transient response is illustrated.

[Figure 3 about here.]

Brunschwiler et al. [10] originally presented the general concept of reducing flow resistance of TIMs into and out of narrow bond lines. This was accomplished by using micrometer-scale orthogonal or diagonal channels that separate posts. Outside the arrays of posts, wider channels were responsible for longer-range (chip-scale) evacuation of the
thermal paste. The arrangement of smaller and larger channels is hierarchical in nature, hence the name Hierarchically Nested Channels. The configuration of channels specified is a compromise between the conflicting interests of the need for densely packed posts, large flow channels, and a large thermal contact area: low-density packing of posts and large flow channels will improve flow performance, but negatively impact the thermal contact area. Brunschwiler et al. reviewed squeeze flow (between two parallel plates) and Poiseuille flow (pipe flow) and used these to derive analytical expressions for estimating bond-line thickness for a given HNC design, after combining the two based on an interpretation of the design geometry. They noted agreement to within $\pm 10\%$ of experimental values for Newtonian fluids, but the method was not as successful at predicting Bingham or Herschel-Bulkley fluid performance.

Brunschwiler et al. [11, 16] expanded on the HNC concept and provided new designs that focused on the flow performance as well as the idea of controlling “particle stacking.” This is the phenomenon where particles in the TIMs tend to “stack up” along flow stagnation and bifurcation lines in the flow pattern, and limit BLT due to the stacking. The new designs tended to result in triangular sections of plate separated by the hierarchical channels, as opposed to the original design, which utilized square plate sections separated by channels. In a subsequent study [17], they extended these studies with numerical (CFD/FEA) modeling, and discussed appropriate usage and design of HNCs.

Davidson and Sammakia [12] also carried out research related to HNCs. They analyzed the squeezing flow of a power-law fluid between channeled plates, and suggested design guidelines for selecting better HNC configurations (based on tabulated data) with this more complex fluid type. They recommended an exploration of other fluid constitutive laws with HNCs.

3. A Resistance Network Model for Analyzing Complex Channel Geometries

The fundamental design strategy used in the present study, is to model a complex channel configuration as a grid of posts of varying heights. In the limit of refinement of the grid spacing, the varying heights of the posts describe continuous channels. Such a procedure
allows channel topologies to naturally emerge by varying individual post heights. A pictorial illustration of this concept is shown in Figure 4.

[Figure 4 about here.]

Resistance/conductance network analogues are commonly used for approximating solutions of partial differential equations in the literature. The discretization in finite element approaches may be thought of as resulting in a conductance network between the nodes. Conductance networks have been constructed to study near-percolation thermal transport behavior of TIMs [18, 6]. Flow resistances are commonly used when modeling flow through pipes in fluid systems, as well. Resistance network analogs have also been used in solving other PDE systems, such as in electromagnetic problems [19].

3.1. Methodology

In order to derive the expression for a local resistance in this problem, Equation (4) serves as the starting point. Dividing both sides of Equation (4) by the area of the plates, \( \pi a^2 \), yields:

\[
\frac{F}{\pi r^2} = \bar{\rho} = -\frac{3\mu \dot{h} r^2}{2h^3} + p_0
\]

where \( \bar{\rho} \) is an average pressure across the plate due to the squeezing force. Now, defining \( \Delta p = \bar{\rho} - p_0 \), one can rearrange the above expression as:

\[
\Delta p = -\frac{3\mu \dot{h} r^2}{2h^3}
\]

Finally, the effective flow resistance is obtained by dividing the above pressure difference by the flow rate \( q = \pi r^2 \dot{h} \). Thus, the average resistance for the region under consideration is:

\[
\frac{\Delta p}{q} = \hat{R}_{fl} = -\frac{3\mu}{2\pi h^3}
\]

To construct the resistance network, we treat the above derived expression for resistance as corresponding to a local “unit cell.” Within the cell, the flow is allowed to traverse along \( m \) paths from the center of the cell. When several of these \( m \)-resistor cells are combined
together, a plate with a complex geometry can be approximated by varying the height within each cell.

The model uses a global bond line thickness $h$ together with local cell gap variations represented by $x$, which is a vector of length equal to the number of cells. Thus, the gap used to compute the resistance for the $i$th cell is the sum of the global gap (datum or BLT) and the individual cell gap given by $h_i = h + x_i$. The resistance value of an individual resistor connecting the cell’s center node $i$ (or cell $i$) with a surrounding node $j$, once the parallel flow along the arbitrarily chosen $m$ discrete paths is taken into account, is:

$$R_{ij} = -m \frac{3\mu}{2\pi h_i^3}$$

In the particular implementation used in the present study, we specifically choose eight flow paths from the center, corresponding to the four diagonal directions and to the positive as well as negative directions of the two orthogonal coordinate directions. Thus, the overall resistance of each cell is further broken into eight resistances corresponding to the aforementioned directions. A $3 \times 3$ cell configuration is illustrated in Figure 5.

Now, as the governing principle, mass conservation at each node is approximately enforced by utilizing the resistance calculated in Equation (8). Thus, if at the $i$th node in Figure 5, the flow is summed over each of the $n$ number of connections at that node, then due to mass conservation, the flow at the node must sum to the average value due to squeezing of the region:

$$\sum_{j=1}^{n} \Delta p_{ij}/R_{ij} = \pi \dot{h} r^2$$

where, $r$ is the equivalent radius of the region influenced by the node (explained below). Also, the number of resistors connecting at a node of the cell, $n$, varies depending on the node being considered. For a cell’s corner nodes, $n = 4$, for a cell’s side nodes, $n = 2$, and for the center nodes, $n = 8$. In matrix-vector notation, the above equation can be written
as:

\[
\begin{bmatrix}
\frac{1}{R_{i1}}, & \frac{1}{R_{i2}}, & \cdots, & \frac{1}{R_{in}}
\end{bmatrix}
\begin{pmatrix}
\Delta p_{i1} \\
\Delta p_{i2} \\
\vdots \\
\Delta p_{in}
\end{pmatrix}
= \pi \dot{h} r^2
\]  

(10)

When this equation is constructed for each node, a set of simultaneous linear equations is formed, which can be expressed as the following matrix equation with a row and column for each node in the system:

\[
Cp = q
\]  

(11)

where \( C \) is a flow conductance matrix, \( p \) is a vector of local pressure difference \( \Delta p_{ij} \), and \( q \) is a vector of local flow inputs. \( q \) is expected to be the same for each node since the squeezing rate (averaged at each node) is constant over the entire plate. Solving the above equation for \( p \) yields the local pressures. Since Equation (7) is derived using an axisymmetric approximation, and since the geometry of each cell is square, in calculating the resistances, a reasonable approximation to the radius \( r \) is the average of the midside and the diagonal, i.e., \( r = \frac{1+\sqrt{2}}{2}a \), where \( a \) is the side of the cell.

In order to find the required squeezing force, the pressure is integrated over the domain. Since the network approximation is a discrete one, this force can be written as:

\[
\Psi = a^T p
\]  

(12)

where \( \Psi \) is the squeezing force, and \( a \) is the vector of areas of each cell (squares with side length equal to \( 2a \), and area \( 4a^2 \)).

3.2. Implementation

The resistance network solution is implemented as a collection of functions and script files written within the MATLAB\(^1\) environment. The structure of the program follows methodologies in part inspired from the implementation of finite element (FE) codes. First, the

\(^1\)A registered trade mark of The MathWorks, Inc.
geometry of the domain is identified, and discretized into the arrangement of cells. This collection of cells is then overlaid with the 9-node “stencils” of each cell as discussed in the derivation of the resistance network. The overlaying process establishes the position of each node and manner in which the nodes are connected by resistors. Then, a list of resistors is constructed, with their corresponding nodes. This is then used to assemble the flow conductance matrix. Boundary conditions are applied, and then the matrix system is solved.

3.2.1. Grid Generation

Two automated grid generation functions are implemented in MATLAB; they generate either a discretized circular or rectangular plate geometry. The connectivity list encodes the connectivity of the nodes and keeps track, using a template, of the direction of nodes relative to one another. The connectivity list is then traversed to create a resistor list (one resistor per connection). Each resistor is also associated with a cell, from which it derives its local \( h + x_i \) value, which is composed of the local cell deviation from a datum, together with the distance of the datum from the squeezing plate. This sum is used as the BLT in computing the flow resistance.

3.2.2. Matrix Assembly

If \( k \) and \( w \) are the indices of two nodes defining a resistor, then the corresponding conductance matrix becomes:

\[
\begin{bmatrix}
1/R_{kk} & -1/R_{kw} \\
-1/R_{wk} & 1/R_{ww}
\end{bmatrix}
\begin{bmatrix}
p_k \\
p_w
\end{bmatrix} = 
\begin{bmatrix}
q_k \\
q_w
\end{bmatrix}
\]

(13)

The simplest approximation is to assume the above resistances to be of equal value; this is done in the present implementation. This matrix is then treated as an individual element matrix as done in finite-element codes. The individual entries in the matrix are assigned to positions in a corresponding global matrix of the entire system. The global conductance matrix \( C \) is square, with dimensions \( N_{\text{nodes}} \times N_{\text{nodes}} \), where \( N_{\text{nodes}} \) is the number of nodes in the system.
3.2.3. Boundary Conditions

Two alternative boundary conditions may be imposed. The first is a no-flow boundary condition, and the second is a known-pressure boundary condition. The no-flow boundary condition is easily implemented using the connectivity list. If a node is not connected to an adjoining one, then, as expected, the lack of connection results in a lack of flow between the neighbors. This type of boundary condition is required to enforce symmetry.

An extended attribute in the connectivity list (in an augmented matrix) is used to hold information regarding the nodes that have known pressure boundary conditions. The known pressure boundary condition may, in general, be enforced either exactly by eliminating the corresponding row and column from the global matrix or approximately through a penalty method as is common in finite element codes [20]. Due to the nature of sparse matrix data structures, eliminating rows and columns of the global matrix is very computationally inefficient. Thus, the known pressure boundary conditions are enforced in the implementation through a penalty method. This method simply forces the desired \( p_i \) value to be close to zero by multiplying the diagonal entry in \( C \) associated with \( p_i \) by a quantity that is several orders of magnitude larger than other terms of the corresponding row. The corresponding value in \( q, q_i \), is then set to that same multiplier times the desired value for \( p_i \).

Once boundary conditions are applied, the matrix \( C \) is symmetric and positive definite (SPD) as are the matrices that occur in finite element and finite difference methods to describe physical systems.

3.2.4. Matrix Solution

After the construction of \( C \) and \( q \) matrices and the application of boundary conditions, solution for \( p \) may be obtained. Because the matrix is SPD, and because the sensitivity analysis described later in the optimization part of the work exploits it, the Cholesky factorization is utilized to solve the matrix system \( Cp = q \) for \( p \). The sparse matrix version of Cholesky decomposition together with their appropriate forward and backward substitution routines implemented in MATLAB are utilized for the solution.
3.3. Validation

The resistance network model was validated through a comparison against results from analytical and finite element simulations of the same geometries and input parameters as those used in the resistance network model. In Figure 6, error relative to the analytical solution versus level of discretization for a circular plate with identical post heights is shown. The model stabilizes to a converged value relatively quickly as the number of cells is increased. The deviation from the analytical solution evaluated with the same inputs (i.e., the error) converges to a value of approximately 2%. This deviation may be due to the assumptions in the original problem formulation (Dienes and Klemm) or to the axisymmetric resistance.

[Figure 6 about here.]

The validation of the resistance network model was also carried out through comparison with COMSOL Multiphysics\textsuperscript{2} finite element results for various configurations. The geometry in these cases varied from flat plates to significantly more complex ones with many posts. In Figure 7, the results corresponding to the squeezing of 2 cm by 2 cm plates, with posts spaced 100 µm apart, squeezed at 1 µm/s, containing a Newtonian fluid with a viscosity of 400 Pa·s, are shown. The contour plots correspond to one-quarter of the plate since symmetry was exploited in generating the solutions.

The calculated force is the integral of pressure over the inlet portion of the fluid boundary in the case of finite element solution, and is the value obtained through Equation (12), $\Psi$, for the resistance network model. The finite element model was of a full-plate (no quarter symmetry) in which the gap between the two plates was the control volume, with an input flow velocity forced to be normal to the surface of the flat plate. The input flow velocity was the same as the squeezing rate used in the resistance network model.

[Figure 7 about here.]

[Figure 8 about here.]

\textsuperscript{2}A registered trademark of COMSOL AB.
Simulations corresponding to three more complex geometries are illustrated in Figure 8, where subfigures (a), (d), and (g) are the plate configurations, with the channeled area indicated in blue and the raised area in red. Subfigures (b), (e), and (h) are the corresponding resistance network solutions, while (c), (f), and (i) are the finite element analysis results for the same geometries. The resistance network method appears capable of capturing the general trends of behavior for even these complex geometrical configurations. Table 1 contains a direct comparison of the computed pressure values for these examples. The resistance network consistently under-predicts the force but to within an acceptable error for the purposes of design (at most 15% and considerably less for simpler geometries). Since the overall pressure contours are captured well, and since the numerical error is a consistent one, the resistance network framework appears to be a reasonable one for purposes of design and for comparisons of relative performance between designs.

4. Design Optimization

In the present section, the design problem is formulated, the objective for the optimization is defined, a computationally efficient sensitivity procedure to determine the derivatives of the objective at one order faster computational effort is developed, and the optimization is performed. Results and discussion follow the implementation.

4.1. Thermal Resistance Objective Function

Given that the developed resistance network-based reduced order model solves for the pressure and subsequently the squeezing force (Ψ, Equation (12)) based on an input squeezing rate, it is clear that the minimum force solution is a trivial one - to not squeeze at all! Then, the gaps would be at their upper design limit, as increasing the gaps monotonically decreases the squeezing force. As the goal of the project is to minimize the thermal resistance of the interface, the squeezing force objective must be traded-off against the need for minimal thermal resistance. Therefore, a thermal resistance objective is also introduced.
The thermal resistance objective function makes the assumption that the fluid-filled gaps corresponding to the posts can be treated as a large number of parallel, one-dimensional thermal resistances. In order for this to be valid, spreading in other, lateral dimensions must be negligible, and thus the model relies on the aspect ratio of an individual cell being \( g_i << a \) where \( a \) is the radius of one of the post cells, and \( g_i \) is the gap at the \( i \)th cell. The resistance network model for the pressure solution also relies on this same assumption.

If this requirement is satisfied, the individual (\( i \)th) cell thermal resistance can be written as

\[
R_{thermal,i} = \frac{g_i}{k_i A_i} = \frac{h + x_i}{k_i A_i}
\]

where, \( h \) and \( x_i \) are as described earlier. In the utilized implementation, \( k_i \) and \( A_i \) do not vary between cells, and thus are written as \( k \) and \( A \), respectively. For a parallel network, the thermal resistance relationship is

\[
\frac{1}{R_{thermal,total}} = \sum_{i=1}^{N} \frac{1}{R_{thermal,i}}
\]

And thus the total thermal resistance is,

\[
\Gamma = R_{thermal,total} = \left[ \sum_{i=1}^{N} \frac{1}{R_{thermal,i}} \right]^{-1}
\]

4.2. Multi-Objective Optimization

The thermal resistance objective function \( \Gamma \) in Equation (16) represents a measure of the effective thermal resistance of the interface through a parallel resistance approximation. Equation (12), in which \( \Psi \) is defined, is the instantaneous squeezing force (i.e., at a chosen \( h \)), and is a measure of the amount of effort required to move the plate at a given velocity \( \dot{h} \). The two objectives are in conflict, as minimizing \( \Gamma \) would require maximizing \( \Psi \). Thus, a composite objective function is used here that combines the two, with a selectable weighting, namely the Pareto weight, that determines the relative importance of the two objective functions. Also, here, the global BLT, \( h \), is treated as an independent parameter of fixed value, while the individual cell gaps \( x_i \) are the optimization variables. Thus, the Pareto-
optimal or multi-objective problem is formulated here as follows:

$$\min_{x} Y(x) = w \Psi(h, x) + (1 - w) \Gamma(x)$$ (17)

such that $l_b \leq x \leq u_b$ (18)

Here, $w$ is the Pareto weight, $x$ is the vector of input local plate heights and $Y$ is the overall objective function to be minimized. As stated earlier, the $x$ values directly correspond to the individual cell heights in the resistance network formulation for the squeezing force. $l_b$ and $u_b$ are the lower and upper bounds on the design variables, respectively. The lower bound of the individual cell gap variables, $x_i$, was kept at 0 units while the bond-line thickness as well as the cell gap was varied during the optimization process to make the gap at any cell $h + x_i$.

4.3. Sensitivity Analysis

Gradient-based optimization methods require derivatives to determine the search direction at each iteration. This in turn requires the computation of sensitivity, or partial derivatives with respect to design variables. These can be computed via a number of methods. The simplest, and most computationally intensive, is the method of finite differences, where the derivative is approximated as:

$$\frac{\partial f}{\partial x} \approx \frac{f(x + \Delta x) - f(x)}{\Delta x}$$ (19)

It is important to note that $\frac{\partial f}{\partial x}$ is a vector quantity with length equal to the number of design variables, and where the individual entries are the partial derivatives of the function $f$ with respect to the corresponding variable. The accuracy of finite difference derivatives in general depends on the relative magnitudes of the truncation and roundoff errors. Therefore, the selection of $\Delta x$ that leads to the least error in the computed derivative can be non-trivial. The method is computationally inefficient because for each perturbation of the design variable, an additional function evaluation is necessary. This means that given $n$ design variables, $n$ additional analyses are required at each iteration of optimization.

Below, we develop an alternative derivative calculation procedure that avoids the $n$ additional analyses required for finite difference derivatives. Considering the objective function,
Equation (17), and taking partial derivatives with respect to an individual \(i\)th component of \(\mathbf{x}\) yields:

\[
\frac{\partial Y}{\partial x_i} = w \frac{\partial \Psi}{\partial x_i} + (1 - w) \frac{\partial \Gamma}{\partial x_i}
\]

(20)
since \(w\) represents a pre-specified design preference and is independent of \(x_i\). Thus, the two terms of the objective function sensitivity can be computed separately. Note that this quantity must be computed for each set of design variable values.

4.3.1. Sensitivity Analysis for the Thermal Resistance Objective

In the case of the thermal resistance objective function, it is possible to derive an explicit (analytical) form of the sensitivity. Beginning with Equation (16),

\[
\Gamma = \left[ \sum_{i=1}^{N} \frac{kA}{h + x_i} \right]^{-1}
\]

(21)

Taking the derivative of both sides with respect to \(x_i\) and utilizing the chain rule, we get

\[
\frac{\partial \Gamma}{\partial x_i} = -(kA) \left[ -(h + x_i)^{-2} \right] \left[ \sum_{i=1}^{N} \frac{kA}{h + x_i} \right]^{-2}
\]

(22)

Thus,

\[
\frac{\partial \Gamma}{\partial x_i} = kA(h + x_i)^{-2} \Gamma^2
\]

(23)

4.3.2. Squeezing Force Sensitivity Analysis for Newtonian Fluids

An analytical sensitivity analysis procedure is developed [21] for calculating the partial derivatives for the squeezing force objective function for Newtonian fluids. The procedure begins with the squeezing force objective function, Equation (12), and a recognition that the function depends on the design variables \(x_i\) corresponding to the gap in the \(i\)th cell in addition to the current datum (BLT), \(h\). Taking the derivative with respect to the \(i\)th design variable \(x_i\), we obtain

\[
\frac{\partial \Psi}{\partial x_i} = \mathbf{a}^T \frac{\partial \mathbf{p}}{\partial x_i}
\]

(24)

where, \(\mathbf{a}\) is a vector containing the areas of each cell. Now, considering the mass conservation condition, Equation (11), and taking the derivative with respect to the \(i\)th design variable \(x_i\)
gives:
\[
\frac{\partial C}{\partial x_i} p + C \frac{\partial p}{\partial x_i} = \frac{\partial q}{\partial x_i}
\]  \hspace{1cm} (25)

Rearranging the above equation, we get:
\[
C \frac{\partial p}{\partial x_i} = \frac{\partial q}{\partial x_i} - \frac{\partial C}{\partial x_i} p
\]  \hspace{1cm} (26)

In the above expression, the two partial derivatives on the right hand side are explicitly and analytically known, and the current solution to \( p \) is known once Equation (11) is solved. Further, when the system governing equation is solved, the Cholesky factors of \( C = LL^T \) are obtained as an intermediate step. These factors are stored and reused during the sensitivity analysis when solving the system of Equations (26). The largest computational expense associated with the analysis is the Cholesky factorization, which is of order \( n^3/6 \). The formation of \( C \) is in general less expensive. Thus, the quantities \( \frac{\partial q}{\partial x_i} \) and \( \frac{\partial C}{\partial x_i} \) may be explicitly obtained using Eqs. (8) and (10), and the system represented by Equation (26) solved in a manner identical to the original solution of Equation (11). The forward and backward substitutions required to solve Equation (26) require order \( n^2/2 \) computations, and thus, the sensitivity analysis for \( n \) variables requires the same order of effort as one factorization of the original system of equations, Equation (11). This represents an order \( n \) speed-up in derivative computation relative to using finite difference derivatives. In Figure 9, an example computation is demonstrated, with the analytical sensitivity and the finite-difference sensitivity both plotted. The analytical and finite-difference derivatives are identical to within limits of machine precision.

[Figure 9 about here.]

4.4. Implementation

The optimization objectives as well as the sensitivity analysis described above were implemented in the MATLAB programming environment. Sensitivity analysis for both the thermal and flow resistance objective functions was implemented as sub-functions, evaluated using the analytical expressions derived in prior sections. The thermal resistance sensitivity
was implemented as a direct computation using a vector-optimized format for speed, and the squeezing force sensitivity was implemented using MATLAB's sparse matrix data structures and functions just as was done for the evaluation of the objective function.

A trust-region method (specifically, MATLAB's trust-region-reflective method) was used for the optimization, and in order to increase the effective resolution of the design variables for square and rectangular geometries without impacting the computational complexity, quarter symmetry was utilized in the solved examples. Trust-region methods [22] are a class of unconstrained and constrained non-linear programming algorithms that function by constructing local approximations to the objective function that are more accurate than either steepest descent (locally linear) search direction or the Newton (locally quadratic) search direction for a complex non-quadratic problem. A constrained version implemented as an option for the fmincon routine, built-in to MATLAB, was utilized in this work. The stop conditions for the optimization runs ranged from taking too many iterations or function evaluations to the detection of the optimum by using relative tolerances in the gradient, function values, or step size. The collections of scripts and functions that made up the optimization program for this study also include the retention of the history of each iteration of the optimization, along with corresponding configurations and pressure plots.

4.5. Optimization Results and Discussion

In non-convex optimization problems, the final optimized design depends on the initial guess. In this study, the starting conditions were simply the initial plate geometry. Two initial configurations were used. They are shown to produce different final configurations. The first initial configuration corresponds to the subtractive method, which cuts away from one of the two plates that are very close together, with each component of the gap vector $\mathbf{x}$ starting at the lower bound, $l_b$. The second initial configuration corresponds to the additive method, where each component of the gap vector $\mathbf{x}$ is initialized at the upper bound, $u_b$. This is akin to the algorithm adding material to one of two plates that are relatively far apart. The model is run with the individual heights of the cells being $x_i$ (the design variable) added to the datum $h$ (bond-line thickness).
4.5.1. Local Minima

Table 2 details the input parameters for the representative optimization runs that are illustrated in Figures 10, 11, 12, and 13. Figures 10 and 11 correspond to an optimization run using the subtractive method, and Figures 12 and 13 correspond to optimization using the additive method. The optimum designs are similar in performance between the two methods for this case in terms of final objective function values, while the actual final plate configurations are different. The design space contains multiple local minima, since the gradient-based trust-region methods converged to different designs. This is relatively common in topology optimization problems.

4.5.2. Pareto Weight Sweep

Additional runs of the optimization framework were performed in order to determine the effects of the Pareto weight on the solutions. These were performed with both the additive and subtractive methods, by finding optimized solutions for a range of Pareto weights. Figure 14 is a scatter plot, showing the final value of the squeezing force objective function versus the corresponding value of the thermal resistance objective function at varying Pareto weights. It can be seen that in terms of the objective functions, in this case, the performance of the design methodologies is very close, as the curves (Pareto fronts that indicate a family of solutions that depend only on changing the Pareto weight) nearly overlap. Figure 15 shows selected designs that correspond to points along the Pareto fronts for the subtractive method, and Figure 16 shows selected designs along the Pareto front for the additive method.
Qualitatively, it appears that the designs resulting from the two methods is similar at the extreme values of the Pareto weight (close to 1 or close to 0). Final configurations seem to deviate significantly over the Pareto weight range of 0.1 to 0.7. This behavior can be explained by the fact that at the extremes of the Pareto weight values, the individual objective functions dominate and thus tend toward their trivial solutions. As for the thermal objective function $\Gamma$, the trivial solution is to minimize the gap while the trivial solution for the squeezing force objective function $\Psi$ is to drive all design variables towards the largest allowed gap. It is important to keep in mind that every Pareto-optimal solution corresponding to a particular value of the Pareto weight is valid as an optimal solution since the chosen weight indicates the designer’s relative preference between the two objectives.

[Figure 14 about here.]

[Figure 15 about here.]

[Figure 16 about here.]
4.5.3. *Effects of Increasing Design Resolution*

The number of design variables $x_i$ for a given plate geometry can be adjusted by changing the number of cells that are used in the model. The discrete representation developed in the resistance network model allows a tradeoff between computational expense and the ability to generate a relatively smoothly varying channels. With an appropriate choice of the number of computational cells, one can match the smallest feature sizes reproducible by the desired manufacturing process, and thus optimize the design for that process.

Increasing the resolution of the design should be performed with some caution, as the validity of the model for the squeeze flow resistance as in Equation (8) depends on the aspect ratio of the individual squeeze flow cells, as discussed earlier. Thus, when performing optimizations with refined cells, it may be necessary to decrease the initial value of the datum for the BLT ($h$) correspondingly to mitigate the impact on the aspect ratio.

Relative to the ones shown in Figure 14, decreasing BLT would increase squeezing force and therefore move the vertical asymptote of the curve to the right. It would also, by virtue of decreasing the thermal resistance due to a smaller gap, move the horizontal asymptote downwards. The shape of the curve would also change, though it would maintain the same general trend, decreasing from left to right and curving between the asymptotes. In Figure 17, optimized designs are shown generated using parameters of Table 2 for an individual Pareto weight, using an additive method. It can be observed that even with a coarse discretization, the basic diagonal ridge is still present. As finer discretization is utilized, the detail of the design evolves into shapes with finer features.

[Figure 17 about here.]
4.6. Comparison of Optimized Designs to Ad-Hoc Straight-Channel Configurations

A final study was conducted in order to compare the results obtained using the optimization framework against trials of ad-hoc straight-channel designs. A number of simple channel arrangements were considered and their required squeezing force plotted against the thermal resistance objective function, evaluated in the same way as performed during the optimization sweep. Input parameters were the same as in Table 2 with the exception of having no associated Pareto weight (as this was not an optimization procedure, but merely an evaluation of a fixed design). The ad-hoc geometries used (here and elsewhere also) are those shown in Figure 18.

Ad-hoc designs were constructed by selecting arrangements of channels (such as a simple diagonal cross pattern) and then producing design variants with increasing channel area of the same overall pattern. As the thermal resistance objective function was computed from parallel one-dimensional thermal resistances corresponding to cell gaps, designs with identical overall channeled area have identical values of the thermal resistance objective function. This characteristic assisted in the comparison of the flow performance of designs with similar channeled areas.

The above described ad-hoc designs are, interestingly, similar to those that emerge in the course of optimization as shown in Figure 19 for various Pareto weights. The two sets of designs are also indicated in the plot of relative objective function values in Figure 20.

The first conclusion that can be drawn is that amongst the ad-hoc, channeled designs, simple diagonal patterns appear to perform more poorly compared to those with horizontal/vertical channels. This is reflected in the designs that are produced by the optimization algorithms, as the optimized designs tend to move material toward the diagonals in some manner. The designs considered in the literature tend to utilize a combination of diagonal and horizontal/vertical channels and justify their use of the diagonal channels by arguing that it reduces particle stacking in the material. However, the homogenized Newtonian flow fluid model used here does not consider the effect of filler particles, and therefore the possible effects of particle stacking are not modeled. The indication given here, though, is that for
TIMs or fluids with conditions close to Newtonian flow, the diagonal channels are not the optimal place to remove material.

A second conclusion from the present work is that the optimized designs are superior when compared to ad-hoc designs with improved values of both thermal and flow objectives.

5. Concluding Remarks

A reduced-order model based on a resistance network was developed for simulating squeeze flow between parallel plates that are separated by distances common in thermal interfaces. Individual resistances were computed based on smaller squeeze flow sub-problems, which were then linked together with appropriate conservation equations to simulate flow around posts. This reduced-order model allowed for efficient evaluation of the performance of the modified heat transfer surfaces used with particle-filled greases, and is extendable to non-Newtonian fluid behavioral models. The model lends itself to efficient analytical sensitivity analysis, which can be used in design optimization through efficient, gradient-based search algorithms.

The resistance network model together with the sensitivity analysis procedure was used to generate optimal designs in a Pareto optimal framework that traded-off thermal resistance against squeezing force. Families of optimal designs were generated, depending on the user preference. Finally, the generated optimal designs were compared to ad-hoc designs similar to those in the literature. The optimal designs were superior in their performance relative to those with regular horizontal/vertical or diagonal channel arrangements.

Further study is required with regards to the modeling and design of such thermal interfaces when more complex fluid types such as the Bingham fluids are involved. For the design of channels for such fluids, the flow resistance network model as well the sensitivity
analysis procedure must be updated. Such an effort together with experimental validation of generated optimal designs is currently ongoing.

6. Acknowledgements

Financial support for this work was provided by the Cooling Technologies Research Center, a National Science Foundation Industry/University Cooperative Research Center at Purdue University.

References


List of Figures

1 Schematic diagram of a common configuration of a thermal management solution for microelectronics. .......................................................... 29
2 Schematic diagram of the components of the one-dimensional thermal resistance model. .......................................................... 30
3 Time integration of Equation (2). Squeezing force: 40 N. Plate radius: 1 cm. .......................................................... 31
4 Schematic diagram of channel topology representation through a collection of posts. .......................................................... 32
5 Resistance Network schematic diagram. The thickened lines indicate resistors, and the thickened dots correspond to nodes where pressure is evaluated. Individual cells are shown as squares separated by thin lines. .......................................................... 33
6 Convergence of the resistance network solution to the analytical solution with increasingly finer discretization. .......................................................... 34
7 Pressure distribution over one quarter of 2 cm by 2 cm flat plates containing a Newtonian fluid with $\mu = 400$ Pa·s squeezed at a rate of 1 $\mu$m/s: (a) resistance network solution and (b) finite element solution. Estimated resistance network force was 26.222 N, and that obtained using the finite element analysis was: 26.975 N. Color range is 0 to 150 kPa. .......................................................... 35
8 Validation of pressure plots from the resistance network solutions against finite element solutions for complex geometries. .......................................................... 36
9 Sensitivity of the squeezing force objective function, by cell, for a system using quarter symmetry and containing 30 by 30 cells for design. .......................................................... 37
10 Convergence trend for an optimization run using the subtractive method with the input parameters in Table 2. .......................................................... 38
11 Selected plate configurations for an optimization using the subtractive method with the input parameters in Table 2. Red indicates cells where the gap is low, and blue indicates cells where the gap is high. .......................................................... 39
12 Convergence trend for an optimization run using the additive method with the input parameters in Table 2. .......................................................... 40
13 Selected plate configurations for an optimization using the additive method with the input parameters in Table 2. Red indicates cells where the gap is low, and blue indicates cells where the gap is high. .......................................................... 41
14 Pareto weight sweeps for both additive and subtractive methods of optimization. Size of the scatter points indicates the relative value of the overall objective function, $\Psi$. The input parameters, with the exception of the Pareto weight, are those in Table 2. .......................................................... 42
15 Selected plate configurations for final designs at various Pareto weights, generated using the subtractive method. The input parameters, with the exception of the Pareto weight, are those in Table 2. Red indicates cells where the gap is lowest, and blue indicates cells where the gap is highest. .......................................................... 43
16 Selected plate configurations for final designs at various Pareto weights, generated using the additive method. The input parameters, with the exception of the Pareto weight, are those in Table 2. Red indicates cells where the gap is lowest, and blue indicates cells where the gap is highest. ........................................ 44

17 Optimized designs generated using parameters of Table 2 with $w = 0.5$, and increasing the number of design variables (cells), using the additive method. .................................................. 45

18 Ad-hoc designs used to test against optimized-geometries, with results in Figure 20. .................................................................................................................................................. 46

19 Optimized designs generated using parameters of Table 2 with various Pareto weights (indicated on the side of the figures) on 20 by 20 cells with quarter-symmetry (and then mirrored to produce the full geometry) used in Figure 20. ........................................ 47

20 Squeezing force performance of a number of substrates plotted versus their thermal resistance objective function performance. ................................................................. 48
Figure 1: Schematic diagram of a common configuration of a thermal management solution for microelectronics.
Figure 2: Schematic diagram of the components of the one-dimensional thermal resistance model.
Figure 3: Time integration of Equation (2). Squeezing force: 40 N. Plate radius: 1 cm.
Figure 4: Schematic diagram of channel topology representation through a collection of posts.
Figure 5: Resistance Network schematic diagram. The thickened lines indicate resistors, and the thickened dots correspond to nodes where pressure is evaluated. Individual cells are shown as squares separated by thin lines.
Figure 6: Convergence of the resistance network solution to the analytical solution with increasingly finer discretization.
Figure 7: Pressure distribution over one quarter of 2 cm by 2 cm flat plates containing a Newtonian fluid with \( \mu = 400 \text{ Pa} \cdot \text{s} \) squeezed at a rate of 1 \( \mu \text{m/s} \): (a) resistance network solution and (b) finite element solution. Estimated resistance network force was 26.222 N, and that obtained using the finite element analysis was: 26.975 N. Color range is 0 to 150 kPa.
Figure 8: Validation of pressure plots from the resistance network solutions against finite element solutions for complex geometries.
Figure 9: Sensitivity of the squeezing force objective function, by cell, for a system using quarter symmetry and containing 30 by 30 cells for design.
Figure 10: Convergence trend for an optimization run using the subtractive method with the input parameters in Table 2.
Figure 11: Selected plate configurations for an optimization using the subtractive method with the input parameters in Table 2. Red indicates cells where the gap is low, and blue indicates cells where the gap is high.
Figure 12: Convergence trend for an optimization run using the additive method with the input parameters in Table 2.
Figure 13: Selected plate configurations for an optimization using the additive method with the input parameters in Table 2. Red indicates cells where the gap is low, and blue indicates cells where the gap is high.
Figure 14: Pareto weight sweeps for both additive and subtractive methods of optimization. Size of the scatter points indicates the relative value of the overall objective function, Ψ. The input parameters, with the exception of the Pareto weight, are those in Table 2.
Figure 15: Selected plate configurations for final designs at various Pareto weights, generated using the subtractive method. The input parameters, with the exception of the Pareto weight, are those in Table 2. Red indicates cells where the gap is lowest, and blue indicates cells where the gap is highest.
Figure 16: Selected plate configurations for final designs at various Pareto weights, generated using the additive method. The input parameters, with the exception of the Pareto weight, are those in Table 2. Red indicates cells where the gap is lowest, and blue indicates cells where the gap is highest.
Figure 17: Optimized designs generated using parameters of Table 2 with $w = 0.5$, and increasing the number of design variables (cells), using the additive method.
Figure 18: Ad-hoc designs used to test against optimized-geometries, with results in Figure 20.
Figure 19: Optimized designs generated using parameters of Table 2 with various Pareto weights (indicated on the side of the figures) on 20 by 20 cells with quarter-symmetry (and then mirrored to produce the full geometry) used in Figure 20.
Figure 20: Squeezing force performance of a number of substrates plotted versus their thermal resistance objective function performance.
List of Tables

1 Comparison of squeezing force values from resistance network and finite element solutions, corresponding to Figure 8. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 50
2 Parameters for the optimization examples shown in Figures 10, 11, 12, and 13. 51
Table 1: Comparison of squeezing force values from resistance network and finite element solutions, corresponding to Figure 8.

<table>
<thead>
<tr>
<th>Geometry Case</th>
<th>Resistance Network</th>
<th>Finite Element Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.955 N</td>
<td>2.060 N</td>
</tr>
<tr>
<td>2</td>
<td>0.9545 N</td>
<td>1.021 N</td>
</tr>
<tr>
<td>3</td>
<td>0.4004 N</td>
<td>0.4617 N</td>
</tr>
</tbody>
</table>
Table 2: Parameters for the optimization examples shown in Figures 10, 11, 12, and 13.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Squeezing Velocity $h$, $\mu$m/s</td>
<td>1</td>
</tr>
<tr>
<td>Viscosity $\mu$, Pa·s</td>
<td>400</td>
</tr>
<tr>
<td>Total Plate Width, cm</td>
<td>2</td>
</tr>
<tr>
<td>Number of Cells Along Width</td>
<td>20</td>
</tr>
<tr>
<td>Number of Cells Along Length</td>
<td>20</td>
</tr>
<tr>
<td>Total Plate Length, cm</td>
<td>2</td>
</tr>
<tr>
<td>BLT ($h$), $\mu$m</td>
<td>100</td>
</tr>
<tr>
<td>Upper Bound on $x_i$ ($u_b$), $\mu$m</td>
<td>300</td>
</tr>
<tr>
<td>Lower Bound on $x_i$ ($l_b$), $\mu$m</td>
<td>100</td>
</tr>
<tr>
<td>Pareto Weight $w$</td>
<td>0.5</td>
</tr>
</tbody>
</table>